Institut für Lebensmittel- und Ressourcenökonomik der
Rheinischen Friedrich-Wilhelms-Universität zu Bonn

EU wide analysis of the Common Agricultural Policy using spatially disaggregated data

I n a u g u r a l - D i s s e r t a t i o n

zur

Erlangung des Grades

Doktor der Agrarwissenschaften

(Dr. agr.)

der

Hohen Landwirtschaftlichen Fakultät

der

Rheinischen Friedrich-Wilhelms-Universität

zu Bonn

vorgelegt am 21.12.2012

von

Markus Kempen

aus

Geilenkirchen

Referent:	Prof. Dr. Thomas Heckelei
Korreferent:	Prof. Dr. Ernst Berg
Tag der mündlichen Prüfung:	11. Oktober 2013
Erscheinungsjahr:	2013

Diese Dissertation ist auf dem Hochschulschriftenserver der ULB Bonn elektronisch publiziert (http://hss.ulb.uni-bonn.de/diss_online).

Danksagung

„Good things come to those who wait.”

(Englisches Sprichwort)

In diesem Sinne danke ich Prof. Dr. Thomas Heckelei für seine Geduld und Ermutigung, die vorliegende Dissertation zu Ende zu bringen. Sein Verständnis für alle landwirtschaftlichen Belange und Zeitzwänge, aber auch seine offene Perspektive gegenüber anderen beruflichen Zielsetzungen trägt sehr zu der angenehmen Arbeitsatmosphäre hier am Institut bei. Ich danke Prof. Dr. Ernst Berg für die Übernahme des Koreferats und der damit verbundenen Hilfestellung zur Beendigung meiner Promotion. Auch bedanke ich mich bei Dr. Heinz-Peter Witzke und Dipl. Inf. Andrea Zintl für Ihre Unterstützung, meine verschiedenen Arbeitstätigkeiten zusammen zu bringen. Weiterhin möchte ich mich bei dem CAPRI-Team vom Jahr 2004 bedanken, damals bestehend aus Dr. Wolfgang Britz, Dr. Marcel Adenäuer, Dr. Torbjörn Jansson, Dr. Ignacio Perez und Dr. Christine Wieck, die mich in die „faszinierende Welt“ des CAPRI Universums eingeführt haben.

Meine Arbeit war geprägt durch die Kooperation mit vielen Partnern in den verschiedenen europäischen Ländern. Hier möchte ich mich besonders bedanken bei Dr. Berien Elbersen von Alterra, Dr. Adrian Leip vom JRC in Ispra, Dr. Tim Kränzlein vom der Forschungsanstalt für Agrarwirtschaft und Landtechnik und Foppe Bouma vom LEI. Die vielen Kontakte, Emails, Anrufe, Besuche und gemeinsame Stunden über CAPRI-Code haben mir viel Freude bereitet und meine Promotionszeit bereichert.

Ich habe viele Doktoranden kommen und gehen sehen. Ich danke allen ehemaligen und jetzigen DoktorandInnen und Mitarbeitern am Institut für gute Ideen, Interesse und Ablenkung, was zum Gelingen dieser Arbeit beigetragen hat.

Großer Dank gilt meinen Eltern, die mich in meinem Studium und auf meinem wissenschaftlichen Weg immer unterstützt haben. Last but not least danke ich meiner Familie, die alle meine Unterfangen mit Unterstützung und Ideen begleiten.

Zusammenfassung

Die jüngsten Reformen der Gemeinsamen Agrarpolitik zielten auf eine verstärkte Förderung der Wettbewerbsfähigkeit des Agrarsektors, des ländlichen Raumes und der umweltverträglichen Landwirtschaft ab. Diese Reformen trugen damit auch der besonderen Rolle der Landwirtschaft beim Schutz von Natur und Landschaft Rechnung. Deutliche Fortschritte bei der Evaluierung von Politikreformen können erreicht werden, wenn die bestehenden ökonomischen und bio-physikalischen Modelle verknüpft würden. Ein wichtiges methodisches Problem liegt in diesem Zusammenhang in der Überbrückung von verschiedenen „Modellskalen“: Während die meisten bio-physikalischen Modelle auf der Ebene des Feldschlages arbeiten, modellieren EU-weite agrarökonomische Modelle in der Regel vergleichsweise große administrative Regionen.

Der Forschungsbeitrag dieser Dissertation zielt auf eine Verbesserung der integrierten Bewertung der europäischen Agrarpolitikreformen ab. Hierfür werden Methoden entwickelt, die räumlich explizite landwirtschaftliche Informationen zu Bodennutzung und Anbausystemen liefern. Dabei wird zunächst in Kapitel 2 ein Verfahren zur Abschätzung der landwirtschaftlichen Bodennutzung entwickelt. Dies geschieht durch die Verbindung hochaufgelöster Informationen zur pflanzlichen Bodennutzung mit aggregierten Daten aus administrativen Regionen. Ein statistischer Ansatz, der eine Kombination aus einem binären choice Modell mit einem Bayesian highest posterior density estimator darstellt, erlaubt die Disaggregation von regionalen Landnutzungsanteilen auf 100,000, so genannte homogene räumliche mapping units. Die angewandte Bayes'sche Methode erlaubt eine vollständige und transparente Darstellung der prior information - Mittelwert und Varianz der Landnutzungsanteile aus den binären choice Modellen - bei der Suche nach Konsistenz zwischen den verschiedenen Skalen.

In Kapitel 3 wird ein Ansatz zur räumlichen Verteilung von landwirtschaftlichen Betrieben entwickelt, da EU-weite Betriebsinformationen nur auf einer hoch aggregierten Ebene erhältlich sind. Der entwickelte Allokationsalgorithmus ordnet jedem Testbetrieb eine räumliche Dimension zu, die es erlaubt, die Betriebe sowohl natürlichen als auch niedrigeren administrativen Skalen zu zuordnen. Dieser Allokationsalgorithmus ist als Optimierungsmodell mit Nebenbedingungen definiert, die bei der Suche nach einer optimalen Konsistenz zwischen betrieblichen Attributen und räumlichen Eigenschaften helfen. Die Zielfunktion wird von einem Bayesian highest posterior density estimator Ansatz abgeleitet.

Kapitel 4 stellt eine Methode zur Integration von räumlich expliziten Betriebsinformationen in das landwirtschaftliche Sektormodell CAPRI vor. Dieser Ansatz wurde im Rahmen einer Studie zu den wirtschaftlichen und ökologischen Auswirkungen der Abschaffung der EU-Milchquote entwickelt. Dabei wurden ökonometrische Schätzungen aus Testbetriebsdaten genutzt, um die regionalen Milchquotenrenten im CAPRI-Modell zu aktualisieren. Die Ergebnisse zeigen, aggregiert für die EU für das Jahr 2020, dass die Produktion sich um circa 5% erhöhen wird während der Preisrückgang für Rohmilch bei etwa 10% liegt. Weiterhin wurden Regionen identifiziert, in denen die wirtschaftlichen und ökologischen Veränderungen

wesentlich die Änderungen auf Ebene der Mitgliedstaaten überschreiten. Regionale Nitratauswaschungsprobleme können sich in Folge der Quotenabschaffung verschärfen, wohingegen es nur schwache Hinweise auf eine Zunahme des Brachlandes in marginalen Gebieten gibt.

Summary

Recent reforms of the Common Agricultural Policy shifted the emphasis towards competitiveness of the agricultural sector, rural development and environmentally sound farming approaches, acknowledging the considerable role agriculture plays in protecting nature and landscape. Significant progress in the evaluation of policy reform scenarios can be made if it is be possible to link existing economic and environmental models. An important methodological problem in this context is “bridging” the scales: whereas most bio-physical models work on field scale, comprehensive EU wide economic models generally work with large administrative regions.

The research presented in this thesis aims at improving integrated assessment of European policy options by developing methodologies that deliver spatially explicit agricultural data regarding crop shares and farming systems. In doing so, first, in Chapter 2 a procedure for estimating agricultural land use choices is developed bringing together high resolution information on crops and land cover as well as aggregate information from administrative regions. Combining a binary choice model with a Bayesian highest posterior density estimator, a statistical approach to break down land use choices from European administrative regions to about 100.000, so called Homogeneous Spatial Mapping Units is developed. The applied Bayesian method fully and transparently accounts for the prior information – mean and variance of land use shares obtained from binary choice models – when searching for consistency between the different scales.

Next, an approach for the spatial allocation of farm information is developed. European wide farm information is so far only available at a rather aggregated administrative level. The suggested allocation approach adds a spatial dimension to all sample farms making it possible to aggregate farm types both to natural and to lower scale administrative regions. The allocation approach is implemented as a constrained optimization model searching for an optimal match between farm attributes and spatial characteristics subject to consistency constraints. The objective functions are derived from a Bayesian highest posterior density framework.

Chapter 4 presents an approach to integrate spatially explicit farm information in an agricultural sector model in the context of a study on the abolition of the EU milk quota. It presents an economic and environmental impact analysis using the CAPRI model, which has been updated with econometric estimates of milk quota rents from sample farms. Aggregated at EU level for the year 2020, production may increase by 5% while the price drop for raw milk is about 10%. Regions are identified where economic or environmental changes substantially exceed those at the Member State level. While regional nitrate leaching problems could be exacerbated, there is only weak evidence of an increased risk of land abandonment in marginal areas.

Table of Contents

Chapter 1: Introduction

1 Motivation and research objective

The Common Agricultural Policy (CAP) has its roots in the 1960s when agricultural productivity was low and food supplies could not be guaranteed. The emphasis of the early CAP was on encouraging better agricultural productivity so that consumers had a stable supply of affordable food and on ensuring that the European Union (EU) had a viable agricultural sector. From the 1980s onwards, the EU had to contend with almost permanent surpluses of the major farm commodities, leading to drastic changes of the CAP support mechanism in the beginning 1990s. At that time, agricultural economists started to develop models assessing impacts of policy changes on the agricultural sector. The focus was on supply and demand balances as well as agricultural income and budgetary effects. Examples of this type of models include early versions of the CAPRI model[1], the AGLINK model[2], and others.

Further reform steps in the years 1999 (Agenda 2000), 2003 (Mid Term Review) and 2008 (Health Check) shift emphasis more and more towards competitiveness, rural development and environmentally sound farming. A study by Buckwell (1997) proves the considerable role agriculture plays in protecting nature and landscape. Various models simulating environmental effects of agriculture were developed (e.g. DNDC[3], see Li, 1992). In the beginning years of 2000, it became evident that significant progress in judging further policy reform scenarios could be made, if it would be possible to link economic and environmental models. Several EU funded research projects addressed this issue as for example CAPRI-Dynaspat[4], SEAMLESS[5], or SENSOR[6].

Four key challenges and requirements to make research tools more useful for integrated assessment in the EU were defined in interactions between scientists and the European Commission (EC): (1) Overcoming the gap between micro–macro level analysis, (2) the bias in integrated assessments towards either economic or environmental issues, (3) the poor re-use of models, and (4) hindrances in technical linkage of models (Ittersum et al, 2008).

---

[1] The CAPRI modelling system is to a large extent developed by the Institute for Food and Resource Economics of the University of Bonn. Further details can be found at: www.capri-model.org.

[2] AGLINK is maintained by the Organisation for Economic Cooperation and Development (OECD) in Paris. Further details can be found at: http://www.oecd.org/site/oecd-faoagriculturaloutlook/oecd-faoagriculturaloutlook-tools.htm.

[3] DNDC (DeNitrification DeComposition).

[4] See http://www.ilr.uni-bonn.de/agpo/rsrch/dynaspat/dynaspat_e.htm.

[5] See http://www.seamlessassociation.org/.

[6] See http://www.sensor-ip.org/.

Linking models working at different levels typically requires scaling of data and transfer of data between the components. Whether data manipulation refers to changes in extent, coverage, or resolution (Volk and Ewert, 2011, Ewert et al., 2011; Bierkens et al., 2000), different methods must be chosen. Accordingly, we distinguish methods related to:

- Change in extent as extrapolation and singling out,
- Change in coverage as interpolation and sampling,
- Change in resolution as aggregation and disaggregation,

Economic models typically operate at aggregate administrative regions, sometimes even states, where harmonized statistical data on agriculture is available. Contrary, bio-physical and, or process based models often simulate at field scale. Consequently, linking those tools in integrated assessment requires changes in resolution. Due to the non-linear character of most process based models, the results depend largely on the resolution of input data (Mulligan, 2006; Montzka et al., 2008). Therefore disaggregation techniques have to be applied to transfer results from economic to environmental models.

The research presented in this thesis aims at improving integrated assessment of European policy options by developing methodologies delivering spatially explicit agricultural data. In doing so, first, a procedure for estimating agricultural land use choices is developed bringing together high resolution information on crop land cover and aggregate information from administrative regions. The next improvement for integrated assessment models presents an approach to spatially allocate farm information to specific agri-environmental zones. The method adds a spatial dimension to all sample farms allowing to aggregate farm types both to natural regional types and to lower administrative scales. Both procedures became relevant features for data-processing and allocation in the CAPRI modelling system supporting the regional and farm type-related model analysis. Finally, using this disaggregated information now available in the CAPRI model, an integrated assessment of policy options for milk quota abolishment in the EU is performed with special focus on the economic and environmental regional impacts of such a reform. Each of the three thesis projects has been published as a separate paper in a journal or on a conference. They are presented as separate chapters in the thesis in the order, as they were performed over time.

In order to provide the overall methodological context of the research done in this thesis, the following section of this introduction describes procedures generating spatially explicit data on agricultural land use or farming systems. Some of these procedures are embedded in integrated assessment tools, whereas others aim at making spatial data on agriculture available for further analysis without the link to specific modelling system. The final section provides a conclusion on the state-of-the-art and limitations of downscaling methods in the context of European data availability and discusses potential for further research in this area.

2 State-of-the-art methodologies to derive spatially explicit agricultural data

Generally, agricultural data can be collected at any desired resolution. When the area under investigation is small, for example a few thousand hectares in a river basin or natural protection area, field observations and farmer interviews are appropriate means to collect detailed data. Observation of land use in medium size regions (e.g. NUTS 2[7] regions) can be supported by high resolution remote sensing technologies, differentiating crops at field level. The limits of this technique are set by the availability of comparable, high resolution satellite images and their time consuming interpretation. Nonetheless, medium resolution images allow differentiation of land cover, for example arable land or pasture at country or even continent level. Section 2.1 describes exemplarily remote sensing based land cover observations.

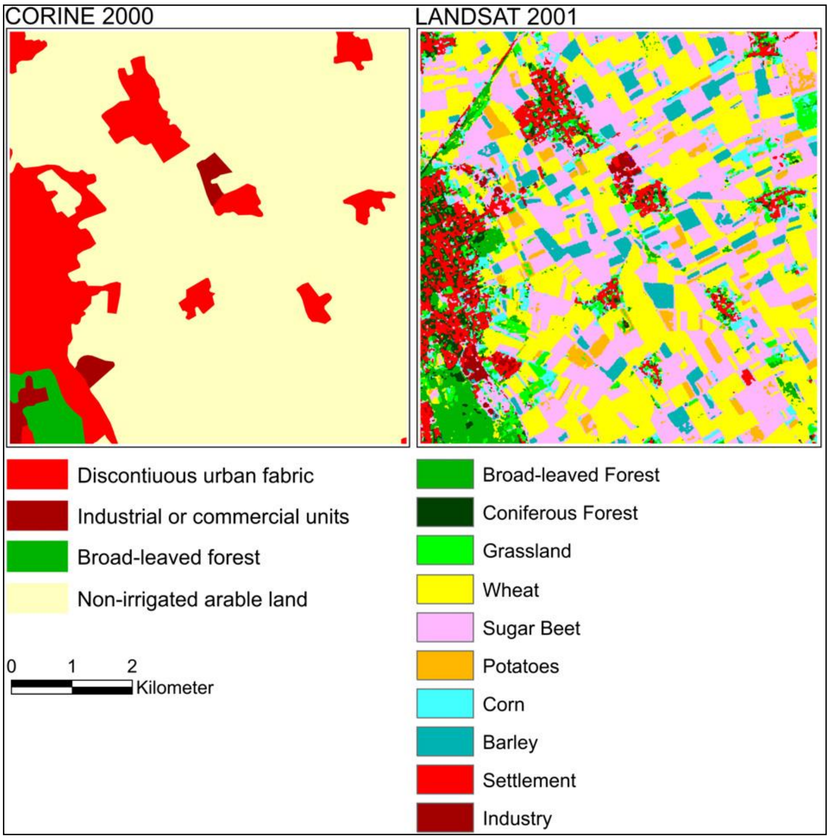
Alternatively, techniques aiming at detailed spatially explicit information relate agricultural land use or farming systems to natural or socio-economic characteristics, where spatial information is available. The easiest way is to link spatial information and agricultural data by defining expert rules. Section 2.2 describes several studies of this kind, which are typically focusing on developing countries. More complicated approaches regress sample data on spatial characteristics and extrapolate to the entire area. When additionally aggregate data at administrative level is available, disaggregation procedures can be applied (see Section 2.3). Disaggregation is typically a two step procedure, combining prior information derived from sample data with a reconciliation step ensuring consistency with aggregate statistics.

2.1 Remote sensing based land cover observations

The CORINE land cover map (European Topic Centre on Terrestrial Environment, 2000) describes land cover (and partly land use) according to a nomenclature of 44 classes, based on the visual interpretation of satellite images and ancillary data (aerial photographs, topographic maps etc.). The CORINE classification system distinguishes 11 agricultural classes (non-irrigated arable land, permanently irrigated land, rice fields, vine yards, fruit and berry plantations, olive groves, annual crops associated with permanent crops, complex cultivation, pasture, marginal areas and forestry). Some of the classes as “Rice fields”, “Olive groves”, “Vineyard”, “Pasture” or “Arable land” clearly indicate a special agricultural use. A minimum of 25 hectare (ha) of homogeneous land cover is defined to build one CORINE mapping unit. That definition of the minimum mapping unit leads to two effects. Firstly, “pure” classes such as “Arable land” may in reality comprise small parcels of other land cover classes as well if these are smaller than 25 ha. Secondly, so-called heterogeneous agricultural areas as e.g. “Land principally occupied by agriculture with significant areas of natural vegetation (marginal area)” comprise no pre-dominant land use that is larger than 25 ha and give only limited information about the type of agricultural use. The 25 ha limit results from the mapping conventions and the interpretative limits set by the spatial and spectral resolution of the satellite images.

[7] The nomenclature d'unités territoriales statistiques (NUTS) refers to administrative units in the EU context where the layers of NUTS 1, NUTS 2, and NUTS 3 are usually distinguished with NUTS 1 referring to the highest administrative level below state level.

Figure 1.1 Land cover derived from CORINE (left) and LANDSAT (right)



Source: Montzka et al., 2008.

Montzka et al. (2008) discussed the advantages of using multispectral remotely sensed data for modelling nitrate concentrations in a river catchment. In this context, it was shown that the identification of main crops and annual crop rotation by the satellite "LANDSAT" (see Figure 1.1), provides the key for a spatial and thematic enhancement of the model results. The spatial resolution of the nitrogen surplus data, taken from the agricultural sector model RAUMIS[8], is enhanced from district level to field/pixel level. In parallel, the empirical water balance model GROWA[9] is enhanced to differentiate between agricultural crops in the real evapotranspiration calculation. Results show an average nitrate concentration in the leachate of 42 mg NO_3/l in the relatively wet year of 2002 and almost 62 mg NO_3/l in the dry year of 2003. There is a 20 mg NO_3/l weather-induced difference which can be modeled in a more detailed way using self-processed remotely sensed data. The model results were compared to nitrate concentrations observed in the top parts of multi-level wells. In this way the related coefficient of determination (R^2) has been improved from a value of 0.50 using coarse land use data to 0.59 by

[8] RAUMIS is an agricultural modelling system maintained by the von Thünen Institute in Braunschweig. See http://www.vti.bund.de/de/startseite/ institute/lr/forschungsbereiche/politik folgenabschaetzung/vti-modellverbund/raumis.html.

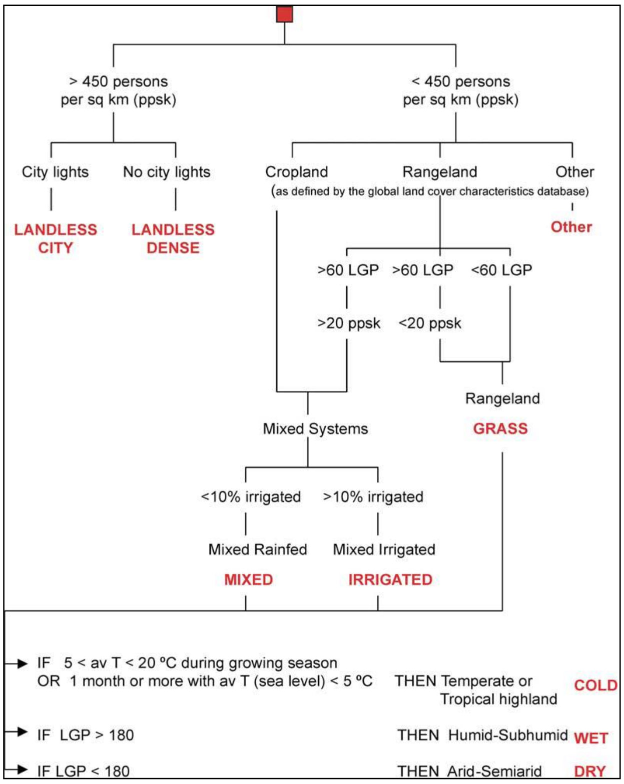
[9] GROWA is a water balance modelling system maintained by the FZ Jülich. See http://www2.fz-juelich.de/icg/icg-4/index.php?index=759.

using self-processed remotely sensed data, thus demonstrating the potential of the enhanced modelling system.

2.2 Spatial allocation procedures based on expert data

Kruska et al. (2003) presented a methodology for mapping livestock production systems in the developing world. The mapping is based on spatially explicit data on agri-climatology (length of growing period), land cover, and human population density. Based on rules developed by Sere and Steinfeld (1996), the approach allows differentiating 12 livestock systems (see Figure 1.2).

Figure 1.2: Decision tree for mapping the livestock systems classification



Source: Kruska (2003). Note: LGP = length of growing period. PPSK = persons per square kilometer. Codes in bold capitals identify the different systems.

The resulting spatially explicit global systems database can be a key component in a wide range of analyses. It can assist in assessing potential changes and adaptations at different levels. Depending on the situation, these adaptations can be designed either to ameliorate expected negative changes (for example an increase in rainfall variability or a decrease in rainfall amount) or to analyse beneficial impacts (such as greater market access or an increase in the length of growing period). Analyses using the systems classification can be of considerable value, not least as a first step in a two-tiered approach that involves identification of hotspots of rapid change, with the second step then involving zooming in to these areas for more detail. At a global level, and even with relatively coarse data sets, hot spots where system changes are likely to be substantial over the next three to five decades can be identified.

Cecchi et al. (2010) refined the livestock mapping for Djibouti, Eritrea, Kenya, Somalia, Uganda, and parts of Ethiopia and Sudan by analyzing datasets collected in the framework of a livelihood analysis where socio-economic rural household data is linked to geographic livestock information. The quantitative definition of the production systems is adopted, based on the ratio of livestock to crop derived income. The geographic distribution of the livestock production systems was modelled using multivariate analysis of remotely sensed and other geospatial datasets from that region. The results were used to fill gaps in the observed distribution of livestock production systems. Finally, selected environmental factors underpinning the systems (agri-climatology, human and livestock populations and land cover) were added, allowing the spatial mapping of livestock production systems and to examine the relationships between these systems and the environment.

2.3 Disaggregation procedures using sample data

Van der Steeg et al. (2010) describes a method to identify the spatial distribution of types of farming systems without the need to extensively map all farming systems across a large region. Moreover, it explains differences between farming systems based on spatial variation in environmental and socio-economic conditions. In the study area, farming systems were characterized and classified based on the criteria (1) area under cultivation of food and cash crops, (2) milk production, and (3) the usage of fertilizers. Logit models were fitted to explain differences in farming system using location factors and household characteristics of about 3300 surveyed farms. A model based on an integrated set of household and location factors best described the diversity of farming systems across the region. However, the location factors alone also described a larger part of the diversity. The spatial variation in location factors, household, and socio-economic characteristics were used to determine the likelihood of occurrence of the different farming systems across the study area. By assigning the farming system to a location that best fits the local conditions based on the logit model, a regional level farming systems map for the Kenyan Highlands was created. The methodology provides a tool of analyzing spatial variation in farming systems, complementary to the analysis of farming systems at the household level and provides insight in the spatial determinants of these systems. The map representing the spatial distribution of farming systems shows a pattern that is too ‘smooth’ when compared to the variation found in the field. This ‘smoothing’ is caused by the absence of household level information on the variation in household characteristics covering the entire region. In case more spatially detailed information of the household characteristics was available, it is likely that a more realistic image of the distribution of farming systems was obtained. So while using the map it is important to realize that at each location a mix of farming systems occurs, while the map only displays the most likely dominant farming system.

Howitt and Reynaud (2003) developed a dynamic, data-consistent method for estimating agricultural land use choices at a disaggregate level (district level) using more aggregate data (regional level). In this context, the term “data-consistent” means that the newly calculated disaggregated data is consistent to the data given at more aggregate levels. The disaggregation procedure proposed by Howitt and Reynaud requires two steps. The first step consists in specifying and estimating a dynamic Markov model of land use at the regional level. In the second step, outcomes of the aggregate model are disaggregated using Generalized Maximum Entropy (GME). The GME disaggregation procedure was applied to a sample of California data. The GME approach gives an optimal solution using the Kullback-Leibler Cross-Entropy criterion in cases where traditional inversion methods do not result in identifying a set of parameters. The resulting disaggregated, district level data are consistent with priors, given by the Markov metrics, and with given data, which consisted of aggregated land use shares at regional level. The GME approach is flexible enough to take into account out-of-sample information. Any specific out of-sample information may be added to the disaggregation program via additional constraints. Specific out-of-sample information on transition probabilities may be added to the model via specification of priors. The disaggregation procedure enabled the recovery of land use at the district-level with an out-of-sample prediction error of 16%.

Verburgh et al. (2003) developed the CLUE-S model[10] specifically for spatially explicit simulation of land use changes based on an empirical analysis of location suitability combined with the dynamic simulation of competition and interactions between the spatial and temporal dynamics of land use systems. The model is sub-divided into two distinct modules, namely a non-spatial demand module and a spatially explicit allocation procedure. The non-spatial module calculates the area change for all land use types at the aggregate level. Within the second part of the model, these demands are translated into land use changes at different locations within the study region using a grid-based system. The probability of land use changes is estimated through logistic regression using the actual land use pattern on the bio-physical and socio-economic location characteristics.

The CLUE-S model supports the spatial allocation of land use change. For the land use demand module different model specifications are possible ranging from simple trend extrapolations to complex economic models. The choice for a specific model is very much dependent on the nature of the most important land use conversions taking place within the study area and the scenarios that need to be considered. The results from the demand module need to specify, on a yearly basis, the area covered by the different land use types, which is a direct input for the allocation module. Several studies (Verburgh et al., 2006; Britz et al., 2011) are based on the CLUE-S model. Remaining challenges are the further downscaling of the simulated land cover changes to the fundamental determinants of the landscapes, including the field size and structure, management intensity, and landscape elements. Such assessment of landscape change trajectories could be linked to the current downscaling procedure and complement the toolbox to discuss the future of Europe’s landscape and spatial planning policies.

3 Conclusions

Although high resolution satellite images can deliver spatially explicit agricultural land use, expanding the procedure to a European scale fails due to technical limits of the availability of comparable satellite images across Europe and training plots in every region. The studies described in the previous section suggest that estimation based on de-

[10] The “Conversion of Land Use and its Effects at Small regional extent” model.

tailed sample data in combination with reconciliation procedures can produce spatially explicit agricultural data at an acceptable precision. As maps on natural conditions, socio-economic data, aggregate regional land use, and farm structure data as well as sample farms and point observations of land use are available in Europe, a statistical procedure achieving consistency of spatially explicit and regional data seems to be a promising methodology. Thus, the development of methodologies delivering spatially explicit agricultural data was the objective of the work undertaken to complete this thesis over the course of the last years. An initial and up to that point unique approach in that direction was presented by Kempen et al. (2005) in the context of the CAPRI-DynaSpat project, where a land use map, differentiating about 40 crops for the whole territory of the EU, was built. Another European research project aiming at spatially explicit land use shares, GENEDEC[11], took over the CAPRI-DynaSpat concept with some technical modifications. In parallel, over the course of the project, the CAPRI-DynaSpat land use map was refined and extended.

Building on Kempen et al. (2005), Chapter 2 describes a procedure for estimating agricultural land use choices in about 100.000 homogeneous spatial units all over the EU territory, using 100.000 sampling points and aggregate data from administrative regions. The disaggregation procedure requires two steps. In the first step, the share of a specific crop is regressed on natural conditions (soil, relief, climate) using the information from sampling points. The estimated coefficients are then used to predict land use choices in each homogeneous spatial unit. Consistency with the administrative statistics is achieved by maximizing the posterior density of estimated a priori information. The downscaling procedure became a feature of the CAPRI modelling system (Britz and Witzke, 2008). This feature enables an automated linkage to the biogeochemistry model DNDC and further modules calculating landscape and biodiversity indicators (for example Shannon index or High Natural Value Farmland index). Both model linkages are regularly applied to analyze policy options with respect to greenhouse gas emissions as well as landscape and biodiversity (Britz and Leip, 2009; Leip et al., 2008; Paracchini and Britz, 2010).

Chapter 3 presents an approach to spatially allocate farm information to a specific environmental context. Data from the European wide Farm Structure Survey (FSS) is only available at a rather aggregated administrative level. Single farm records are available from the Farm Data Accountancy Network (FADN) sample, but the published location of the farm is vague for confidentially reasons. The suggested allocation approach adds a spatial dimension to all sample farms making it possible to aggregate farm types both to natural and to lower scale administrative regions. This spatial flexibility allows providing input data to economic or bio-physical models at a desired resolution. The allocation approach is implemented as a constrained optimization model searching for an optimal match between farm attributes and spatial characteristics subject to consistency constraints. The objective functions are derived from a Bayesian highest posterior density framework (Heckelei et al., 2008). The allocation procedure recovers the spatial farm

[11] http://www.grignon.inra.fr/economie-publique/genedec/eng/enpub.htm.

type distributions quite well, thereby providing information of significant value for further analysis in a multidisciplinary context. Results can serve as input for generic template models on farming systems (e. g. FSSIM[12], see Louhichi et al., 2007), which is a key component of the SEAMLESS integrated framework. Moreover, every analysis based on FADN sample farms can be scaled to any desired spatial resolution.

FADN data was used by Sckokai (see Kempen et al., 2011) and Wieck and Heckelei (2007) to estimate marginal costs of milk production in Europe. In case of production quotas, marginal cost are essential in the calibration phase of economic models based on positive mathematical programming (Adenäuer, 2006), but the spatial resolution of the FADN regions do not match those of the agricultural sector model CAPRI. Chapter 4 presents a feasible approach to integrate FADN with CAPRI information in the context of a study on the abolition of the EU milk quota. The chapter contains an impact analysis of milk production quota expiry using the CAPRI model, which was updated with econometric estimates of milk quota rents at the level of about 230 European regions. The milk quota rents were disaggregated from about 100 FADN clusters using spatially explicit farm information.

The procedures described in the following chapters show that it is feasible to add a spatial dimension to land use and farming systems to improve the results of modelling. Aiming at European wide applications of the procedures, only EU wide harmonized data sources should be used. Nonetheless, problems in mapping definitions are obvious. This can be shown for example for the case of grassland types which vary over a wide range and definitions of land use like “Temporary grazing”, “Permanent pasture”, “Natural grassland”, “Shrub land”, and “Agro forestry” can differ among sources leading to sometime contradictious data. Hence, in particular for grassland systems, this approach could be very useful in the future as the design of the reconciliation procedures used in chapter 2 and 3 allow the prioritizing of sources by the design of the constraint optimization model and manual adjustment of the weight of the data sources in the objective function and.

A general drawback of statistical procedures is that the estimated crop pattern cannot easily be validated against the observed pattern. When the most detailed harmonized European data sources are used in the procedures, there is no comprehensive out-of-sample data left for validation or parameterization. Only for some regions detailed data for validation was accessible. The validation exercises performed in this thesis in Chapter 2 and 3 generally show satisfying results and give clear hints how to specify parameters, but also point out weaknesses which could not be solved immediately.

However, since there is a need for spatially explicit agricultural data by several institutions, as for example the EU Joint Research Centre (JRC) in Ispra (Italy), these methodologies will be further developed and some of the drawbacks will be overcome in the future.

[12] Farm System Simulator (FSSIM), http://www.seamlessassosiation.org.

4 References

Adenäuer, M., 2006. Modelling the European Sugar Sector - Incentives to Supply Sugar Beets and Analysis of Reform Options; Institute for Food and Resource Economics, University of Bonn (downloadable at http://hss.ulb.uni-bonn.de:90/2006/0759/0759.pdf).

Bierkens, M.F.P., Finke, P.A., de Willigen, P., 2000. Upscaling and Downscaling Methods for Environmental Research Kluwer Academic Publishers.

Britz, W., Verburg, P. H. und Leip, A., 2011. Modelling of land cover and agricultural change in Europe: Combining the CLUE and CAPRI-Spat approaches, Agriculture, Ecosystems & Environment, 142(1-2), p. 40-50.

Britz, W., Mittenzwei, K., 2008. Spatial down-scaling as a tool to improve multifunctionality indicators in economic models, in: Paul N. Findley (ed.). Environmental Modelling: New Research, Nova Publishers.

Britz, W., Leip, A., 2009. Development of marginal emission factors for N losses from agricultural soils with the DNDC-CAPRI meta-model. Agriculture, Ecosystems & Environment. 133, 3-4. p. 267-279.

Britz, W., Witzke P., 2008. CAPRI model documentation 2008: Version 2. Institute for Food and Resource Economics, Bonn.

Buckwell, A., 1997. Towards a Common Agricultural and Rural Policy for Europe. European Commission Directorate General for Economic and Financial Affairs, Reports and Studies No 5, 1997, Brussels.

Cecchi, G., Wint, W., Shaw, A., Marletta, A., Mattioli, R. and Robinson T., 2010. Geographic distribution and environmental characterization of livestock production systems in Eastern Africa. Agriculture, Ecosystems and Environment, 135,1-2, 98-110.

European Commission, 2003. The Lucas survey. European statisticians monitor territory. Theme 5: Agriculture and fisheries, Series Office for Official Publications of the European Communities, Luxembourg.

European Topic Centre on Terrestrial Environment, 2000. Corine land cover database (Version 12/2000).

Eurostat, 2000. Manual of Concepts on Land Cover and Land Use Information Systems. Theme 5: Agriculture and Fisheries: Methods and Nomenclatures. Series Office for official Publications of the European Communities, Luxembourg, 2000, pp. 110-ff.

Ewert, F., Van Ittersum, M., Heckelei, T., Therond, O., Bezlepkina, I. und Andersen, E., 2011. Scale changes and model linking methods for integrated assessment of agri-environmental systems, Agriculture, Ecosystems and Environment, 142,1-2, p. 6-17.

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

Heckelei, T., R. C. Mittelhammer and W. Britz, 2005. A Bayesian Alternative to Generalized Cross Entropy Solutions to Underdetermined Models. Contributed paper presented at the 89th EAAE Symposium „Modelling agricultural policies: state of the art and new challenges“, February 3-5, Parma, Italy.

Heckelei, T., Mittelhammer R., Jansson T., 2008. A Bayesian Alternative to Generalized Cross Entropy Solutions for Underdetermined Econometric Models. Agricultural and Resource Economics, Discussion Paper, Institute for Food and Resource Economics, Bonn.

Howitt, R., 1995). Positive Mathematical Programming. American Journal of Agricultural Economics, 77(2), pp. 329-342.

Howitt, R., Reynaud A., 2003. Spatial Disaggregation of Agricultural Production Data by Maximum Entropy. European Review of Agricultural Economics 30(3):359-387.

Kempen, M., T. Heckelei, W. Britz, R. Leip and R. Koeble, 2005. A Statistical Approach for Spatial Disaggregation of Crop Production in the EU, in „Modelling agricultural policies: state of the art and new challenges“, MUP, Parma, Italy.

Kempen, M., Witzke, P., Pérez Domínguez, I., Jansson T., Sckokai, P. , 2011. Economic and environmental impacts of milk quota reform in Europe. Journal of Policy Modeling, Volume 33, 1, 29–52.

Kruska, R.L., Reid, R.S., Thornton, P.K., Henninger, N., Kristjanson, P.M., 2003. Mapping livestock-oriented agricultural production systems for the developing world. Agricultural systems, 77, 39-63.

Leip A., Marchi G., Koeble R., Kempen M., Britz W., Li C., 2008. Linking an economic model for European agriculture with a mechanistic model to estimate nitrogen and carbon losses from arable soils in Europe. Biogeosciences, 5, 73-94.

Li, C., Frolking, S., Frolking, T. A., 1992. Model of nitrous oxide evolution from soil driven by rainfall events: 1. Model structure and sensitivity, Journal of Geophysical Research, 97, 9759–9776.

Louhichi, K., Blanco Fonseca, M., Flichman, G., Janssen, S.J.C., Hengsdijk, H., 2007. A generic template for FSSIM for all farming systems. PD3.3.11,SEAMLESS integrated project, EU 6th Framework program, contract no. 010036-2 (downloadable from http://www.SEAMLESS-IP.org).

Montzka, C., Cantya, M., Kunkela, R., Menzb, G., Vereeckena, H.,Wendlanda, F., 2008. Modelling the water balance of a mesoscale catchment basin using remotely sensed land cover data. Journal of Hydrology , 353, 322–334.

Mulligan, D.T., 2006. Regional modelling of nitrous oxide emissions from fertilised agricultural soils within Europe. Series PhD thesis, submitted to the University of Wales, Bangor.

Paracchini, M.L., Britz, W., 2010. Quantifying effects of changed farm practices on biodiversity in policy impact assessment – an application of CAPRI-Spat. Paper presented at OECD Workshop on Agri-Environmental Indicators, Leysin, Switzerland.

Van der Steeg, J.A., Verburg, P.H., Baltenweck, I., Staal, S.J., 2010. Characterisation of the spatial distribution of farming systems in the Keynyan Highlands. Applied Geography 30, 2, 239-253.

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., 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.

Verburg, P.H., Soepboer, W., Limpiada, R., Espaldon, M.V.O., Sharifa, M.A. & Veldkamp, A., 2002. Modelling the spatial dynamics of regional land use: The CLUE-S model. Environmental Management, 30, 391-405.

Verburg, P.H., Schulp, C.J.E., Witte, N., Veldkamp, A., 2006. Downscaling of land use change scenarios to assess the dynamics of European landscapes. Agriculture, Ecosystems and Environment 114(1): 39-56.

Volk, M., Ewert, F., 2011. Scaling methods in integrated assessment of agricultural systems - State-of-the-art and future directions. Agriculture, Ecosystems & Environment, 142, 1–2, p. 1–5.

Wieck, C., & Heckelei, T., 2007. Determinants, differentiation, and development of short-term marginal costs in dairy production: an empirical analysis for selected regions of the EU. Agricultural Economics, 36(2), 203-220.

Chapter 2: Computation of a European agricultural land use map – statistical approach and validation[13]

Abstract

Combining a binary choice model with a Bayesian highest posterior density estimator, this paper develops a statistical approach to break down land use choices from European administrative regions to about 100.000, so called *Homogeneous Spatial Mapping Units*. The applied Bayesian method fully and transparently accounts for the prior information – mean and variance of land use shares obtained from binary choice models – when searching for consistency between the different scales. The paper validates the results of the disaggregation procedure with out-of-sample data.

1 Introduction

Not at least due to the so-called multi-functional model of European agriculture, there is growing interest in modelling environmental effects of the agricultural sector in the EU. In many cases, results beyond rather crude passive indicators can only be obtained linking bio-physical models to economic models for policy impact analysis. An important methodological problem in this context is "bridging" the scales: whereas most bio-physical models work on field scale, comprehensive EU wide economic models generally work on large administrative regions.

Within these administrative boundaries the natural conditions of soil, relief and climate usually differ in such a manner, that the assumption of identical cropping pattern, yields or input use cannot be maintained. Simulations with bio-physical models thus require breaking down results from the economic models into a smaller regional scale. In this paper, we aim to overcome this shortcoming by providing an approach that integrates spatially explicit data and statistics of administrative units in a consistent way. We develop a procedure that computes a land use map at a spatial resolution of 1x1km for 30 cropping activities and one aggregated non-agricultural land use class covering the entire EU.

The issue of spatial allocation of land use has been addressed from different perspectives during the last years which could be assigned to three major categories: (1) Modelling approaches, (2) remote sensing, and (3) soil suitability.

---

[13] This Chapter is based on the following two conference papers: 1) "A Statistical Approach for Spatial Disaggregation of Crop Production in the EU" presented at the 89th EAAE Seminar, Parma, 3-5 February 2005, together with Thomas Heckelei (University of Bonn), Wolfgang Britz (University of Bonn), Adrian Leip (JRC Ispra), Renate Koeble (JRC Ispra), and Giulio Marchi (JRC Ispra); and 2) "Validation of Spatially Explicit Land Use Choices Based on Probabilistic Theory" presented at the International Conference on Regional and Urban Modelling, Brussels, 1-2 June 2006, together with Thomas Heckelei (University of Bonn) and Wolfgang Britz (University of Bonn).

Most of the approaches developed in recent years were applied at a scale comparable to a 1x1km grid. The accuracy of the employed land use classes varies depending on the desired results and the complexity of the allocation procedure.

Howitt and Reynaud (2003) for example, propose a method where models based on analysis of available small scale data or expert rules predict values for small spatial entities. Following data on larger regional units can be used in a procedure ensuring consistency across different scales. Expected values from the first step serve as prior information in a constrained optimization problem ensuring consistency of regional data und corresponding spatial units.

A rough distinction of about 10 different land cover classes related to agricultural use is provided by the CORINE land cover map (European Topic Centre on Terrestrial Environment, 2000), which is based on visual interpretation of satellite images. Sophisticated remote sensing techniques technically allow differentiation of single crops at a scale desired in our study. Although these techniques have been successfully applied in small scale applications like river basins (e.g. Montzka et al., 2008), there is no application at European level.

Soil suitability is the degree to which soil physical and/or chemical characteristics are in agreement with characteristics that are required for a certain land use. Soil maps give a number of permanent characteristics of the soil which in combination with other characteristics can be used to determine the suitability of locations for certain crops.

The procedure developed in this paper to establish a European land use map combines characteristics of various above mentioned approaches. The CORINE land cover map serves as a starting point which has to be refined to meet the aimed land use classes. This subdivision is based on a model regressing point observations of cropping decisions all over the European territory on soil, relief, and climate parameters describing the land suitability. Spatial statistical techniques are used to allow for spatial heterogeneity of the coefficients using a locally weighted logit model. Since statistical data of agricultural production in Europe is available at NUTS[14] regions, a consistency step will finalize the disaggregation procedure. The applied Bayesian method fully and transparently accounts for the available information – prior distributions derived from the binary choice model and aggregate information on regional crop area.

The paper is structured as follows: Section 2 describes the data base used while section 3 explains and justifies the statistical procedure in detail. Section 4 validates the results with out-of-sample data and the final section concludes.

[14] The nomenclature d'unités territoriales statistiques (NUTS) refers to administrative units in the EU context where the layers of NUTS 1, NUTS 2, and NUTS 3 are usually distinguished with NUTS 1 referring to the highest administrative level below state.

2 Database and definition of Homogeneous Spatial Mapping Units (HSMUs)

The description of the database is divided in two main parts: (1) Sources and definitions of the natural location factors, (2) the construction of HSMUs.

2.1 Maps of natural location factors

The competitiveness of an agricultural crop at a certain location is determined by natural factors, technology, and market conditions. While market conditions and the available technology are assumed to be rather invariant within an administrative region, differences in natural conditions will lead to heterogeneity regarding the optimal crop mix between different locations inside the NUTS 2 region. Therefore, this study concentrates on natural location factors.

Table 2.1: Relevant maps of natural conditions

Data sources	Indicators
European Soil Database V2.0 [15]	Set of Soil Code (World Reference Base)
	Drainage / water management
	Presence of stones
Digital Elevation model [16]	Slope
	Elevation
JRC-MARS meteodata [17]	Annual rainfall
	Cumulative temperature sum

Source: Own compilation.

According to crop science literature, yield potentials of agricultural crops are mostly affected by soil quality, relief and climate conditions. Small scale information on location factors from different sources (see Table 2.1) was prepared with the help of geographical information system (GIS).

The European territory is divided into spatially referenced Soil Mapping Units (SMU). Each SMU consist of up to ten Soil Topological Units (STU). These STU are not spatially referenced and only the percentage of each STU within the corresponding SMU is

[15] European Commission, 2004. European Soil Database (version V2.0), CD-ROM EUR 19945 EN, March 2004, European Commission, DG JRC, Institute for Environment and Sustainability.

[16] CCM DEM 250, 2004. EuroLandscape/Agri-Environment Catchment Characterisation and Modellig Activity, Land Management Unit, Institute for Environment and Sustain-ability, EC-Joint Research Centre. 250 Meter DEM, compiled on the basis of data acquired from data providers and national mapping agencies over Europe for internal use.

[17] Orlandi, S., Van der Goot, E., 2003: Technical description of interpolation and processing of meteorological data in CGMS, Available under http://agrifish.jrc.it/marsstat/Crop Yield Forecastingcgms.htm, European Commission, DG JRC, Agrifish Unit

known. The parameters of interest are given at STU level, mostly as qualitative information like “well drained” or “poorly drained”. Afterwards the percentage of “well” or “poorly” drained areas may be calculated for the SMU or to simplify, the attribute of the dominant STU is assigned to the entire SMU. In the following analysis the percentage values are used.

According to the World Reference Base (WRB), each STU is given a soil code like “Albic Luvisol”. These soil codes already indicate the suitability for farming. In the relevant SMU, 95 WRB soil codes are present. A first clustering was made based on Driessen et al. (2001) who rearranged the 30 WRB soil groups into 10 so-called “sets”, based on the dominant soil forming factors that determined the soil conditions.

Based on expert knowledge, van Diepen and Hazeu (2005) aggregated the soil codes occurring in Europe into 9 of those sets. Next, as some sets were heterogeneous in its characteristics relevant for land use they subdivided those sets further into more homogeneous sets, while other sets were combined into one as the soils were rather similar from the land use perspective. Finally, from the resulting sets, some soil units have been placed in other sets because of some common prominent feature. The distinction of new sets was based on similarity of STU in terms of soil characteristics which are judged relevant for land use, notably rooting depth, organic matter, texture, drainage class, available water holding capacity, and presence of stones and slope.

2.2 CORINE Land Cover Map

The general distinction of different land cover classes is based on the CORINE land cover (CLC) map (European Topic Centre on Terrestrial Environment, 2000) describing land cover (and partly land use) according to a nomenclature of 44 classes, based on the visual interpretation of satellite images and ancillary data (aerial photographs, topographic maps, etc.).

The CORINE classification system distinguishes 11 agricultural classes: Non-irrigated arable land, permanently irrigated land, rice fields, vine yards, fruit and berry plantations, olive groves, annual crops associated with permanent crops, complex cultivation, pasture, marginal areas, and forestry. Some of the classes as “Rice fields”, “Olive groves”, “Vineyard”, “Pasture”, or “Arable land” clearly indicate a special agricultural use. A minimum of 25ha of homogeneous land cover is defined to build one CORINE mapping unit. That definition of the minimum mapping unit leads to two effects. Firstly, “pure” classes such as “Arable land” may in reality comprise small parcels of other land cover classes as well if these are smaller than 25ha. Secondly, so-called heterogeneous agricultural areas as for example “Land principally occupied by agriculture with significant areas of natural vegetation (marginal area)” comprise no pre-dominant land use larger than 25ha and give only limited information about the type of agricultural use. The 25ha limit results from the mapping conventions and the interpretative limits set by the spatial and spectral resolution of the satellite images.

In this study we assume that only the agricultural classes are suitable for farming. The reader is reminded that agricultural classes may comprise small parcels of non-agricultural uses and that agricultural use may be found in non-agricultural classes.

2.3 Motivation and construction of Homogeneous Spatial Mapping Units

The aim of building HMSUs, as broadly discussed by Leip et al. (2008), is the definition of areas inside an administrative region where approximate homogeneity according location factors may be assumed. The HMSUs serve then as simulation units for the biophysical models and are constructed by overlaying different maps (land cover, soil map, climatic factors etc.). In order to allow for a manageable number of HSMUs, the most important characteristics must be selected, and continuous parameters must be grouped in classes. The CORINE land cover map was used here in combination with two further main factors relating to soil (Soil Mapping Units) and relief (slope in 5 classes). Each HSMU has identical values for these three items, other parameters (such as annual rainfall) may vary inside the HSMUs. For those characteristics weighted averages are calculated for each HSMU using GIS techniques.

The HMSUs approach was deemed superior to a grid layout, mainly because factors determining optimal cropping patterns may be identical across very large regions (say Northern Finland) so that grid units would be "wasted", whereas in other regions especially such which high relief changes, the grid units may already comprise differences in natural conditions. Relevant units can be defined so that they do not cross administrative borders, and grid data may be redefined based on the HSMUs.

2.4 Land Use/Cover Area Frame Statistical Survey (LUCAS)

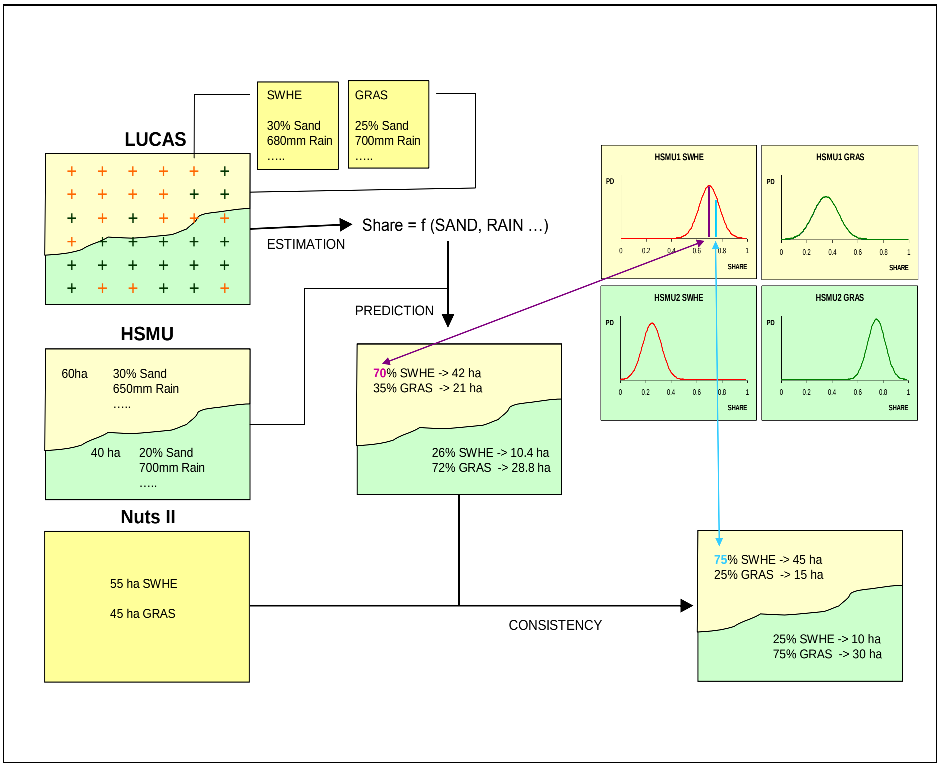
In opposite to mapping approaches, area frame surveys based on a common statistical sampling method gather land cover and land use data (EUROSTAT, 2000) at specific sample points only, and extrapolate from these to the entire area under investigation. The LUCAS survey (European Commission, 2003) covers the territory of all EU Member States and all kinds of land uses, and is based on a two-stage sampling design: at the first level, so-called Primary Sampling Units (PSUs) are defined as cells of a regular grid with a size of 18 × 18 km, while the Secondary Sampling Units (SSUs) are 10 points regularly distributed (in a rectangular of 1500 × 600 m side length) around the center of each PSU resulting in approximately 10.000 PSUs for the whole EU.

Due to possible measurement errors regarding the geo-references in the CORINE maps (Gallego, 2002), about 30% of the LUCAS points closer than 100 m to the border of a CORINE class were not considered in here. The 38 agricultural classes found in LUCAS (36 crop land, 2 permanent grassland classes) were re-grouped according to the crops represented in CAPRI. All other classes (artificial areas, woodland, water, etc.) are aggregated in a residual class termed "Other".

3 Disaggregation procedure

Before describing the crucial steps in detail, the general approach of the disaggregation procedure is illustrated in Figure 2.1. Suppose there is a NUTS 2 region divided in only two HSMUs each comprising two crops – grassland (GRAS) and soft wheat (SWHE). Combining the LUCAS survey with digital maps provides us with several observations of crops grown at a defined point with a set of natural conditions. Using an adequate estimation model, we can regress the probabilities of finding a crop at a certain location on the natural conditions. As this probability can be interpreted as the share of the crop in a homogeneous region, applying these estimated coefficients to the average natural conditions in a certain HSMU, yields normally distributed predictions of crop shares for this HSMU under corresponding assumptions on the stochastic processes governing crop choice. This a priori information on cropping shares is generally not consistent with the "known" cropping area in the NUTS 2 region. The "best" set of consistent shares given the prior information is identified by a Bayesian highest posterior density (HPD) approach (Heckelei et al., 2005). The concept of the HPD estimator allows the direct inclusion of the uncertainty of the prior mean. The variance can be derived from asymptotic properties or bootstrapping procedures.

Figure 2.1: Scheme of disaggregation procedure



Source: Own compilation.

3.1 Locally weighted binomial logit estimation

Generally, shares for each crop, $\hat{Y}_C$, are regressed on the following explanatory variables describing natural conditions:

- Set of soil code [15 categories]
- Drainage [Yes/No]
- Presence of stones [Yes/No]
- Slope [%]
- Elevation [m]
- Rainfall [mm/year]
- Sum of temperature in vegetation period [°C]

The regressions were estimated independently for each crop *c* in each CORINE class *clc*. The arguments for using specific coefficients for each CORINE class are as follows. Assume grass land parcels are found in the LUCAS survey in the "non-irrigated land" CORINE class. We would assume that slope has a positive effect on the probability to find grass. In the "pasture" class of CORINE, we would eventually find the opposite effect: with increasing slope, grass land could be replaced by forest. For convenience the indices *c* and *clc* are omitted in the following.

The LUCAS survey reports one point in time observations and hence does not deliver cropping shares (or rotations), but requires a binary choice model. Both logit and probit models (see e.g. Green, 2000) were originally tested, with the logit approach giving slightly better results in terms of precision of the estimates. The likelihood function of finding crop *c* at a specific LUCAS point *i* for the binomial logit model is defined as:

$$\log L = \sum_{i=1}^{n} [y_i \log \Lambda(\boldsymbol{\beta}' \mathbf{x_i}) + (1 - y_i)(1 - \Lambda(\boldsymbol{\beta}' \mathbf{x_i}))]$$

with $\Lambda(\boldsymbol{\beta}' \mathbf{x_i}) = \dfrac{e^{\boldsymbol{\beta}' \mathbf{x_i}}}{1 + e^{\boldsymbol{\beta}' \mathbf{x_i}}}$

where y_i is a dummy vector indicating whether a certain crop was observed at a location *i* (y_i=1), $\mathbf{x_i}$ is the design matrix containing data on natural conditions and $\Lambda(\boldsymbol{\beta}' \mathbf{x_i})$ is the probability that a specific crop is grown at location *i*.

Applying the estimated $\widehat{\boldsymbol{\beta}}$ to the average natural conditions in a HSMU ($\mathbf{x}_h$) give us a prior estimate for the share of a specific crop in a certain HSMU:

$$\hat{\mathbf{Y}} = \Lambda(\boldsymbol{\beta}' \mathbf{x}_h) = \frac{e^{\boldsymbol{\beta}' \mathbf{x}_h}}{1 + e^{\boldsymbol{\beta}' \mathbf{x}_h}}$$

3.2 Binomial versus multinomial regression

The approach discussed above examines the crops independently from each other and thus neglects the information that crops compete for the available land. This has two possible effects. Firstly, the error terms for the different crops are probably correlated, and secondly, the individual estimated shares do not add up to unity. The multinomial probit model would be ideal as it allows for an unrestricted variance covariance structure of the error terms and satisfies the additivity condition, but is computationally infeasible for 30 crops and 10.000 points. The assumption of an identity matrix for the variance covariance matrix underlying the multinomial logit model was deemed as too inflexible (Nelson et al., 2004), albeit it is easier to solve. The way out might be a nested logit model, a possible expansion in further analysis.

However, both problems were not deemed crucial for the application at hand. Given the large number of observations, the possible gain of taking correlations between the error terms across crops into account is most probably small. Furthermore, the violation of the adding up condition for the shares is explicitly accommodated in the second step of the disaggregation procedure, where the estimated shares serve as prior information only.

3.3 Single crops versus groups of crops

The land use shares are estimated separately for 30 crops and for 10 groups of crops with similar natural requirements (see also Table 2.2). The grouping of the crops was done with respect to natural requirements of specific crops. This can help to allocate crops where sparse information from the LUCAS sample is available, because the overall procedure will tend to allocate those crops in HSMUs where crops requiring similar natural conditions are assigned to.

3.4 Local versus global regressions

The assumption of European wide invariant relationships between the share of each crop and a limited number of location factors describing natural conditions may be problematic if other omitted explanatory factors are not randomly distributed in space, but "clustered". Suppose, for example, two HSMUs with identical natural conditions, the first one close to a sugar refinery, and the second one far away from the next sugar plant. The share of sugar beets in the first unit will probably be much higher, an effect not linked to the natural conditions. Clearly, omitted variables as the effect of sugar refineries could lead to seriously biased parameter estimates. Adding more explanatory variables would certainly help, but it is simply impossible to collect information on all potentially relevant factors at European level (market points, transport infrastructure, environmental legislation, etc.). Instead, spatial econometric techniques are applied to overcome the problem of omitted variables that are correlated over space.

The basic idea behind Locally Weighted Regression, which was proposed by Cleveland and Devlin (1988), is to produce site specific coefficient estimates using Weighted Least Squares to give nearby observation more influence than those far away. Furthermore, the estimation for any specific site is limited to a number of observations within a certain bandwidth around the site. Locally Weighted Regression are mostly found combined with Least Squares estimators, but applications to Maximum Likelihood Estimation as needed in the case of discrete dependent variables are described as well (Anselin et al., 2004).

The weight given to any observation *i* in constructing the estimate for site *j* is given by ω_{ij}. The tri-cube is a commonly used weighting function:

$$\varpi_{ij} = \left[1 - \left(\frac{\delta_{ij}}{d_j}\right)^3\right]^3 I(\delta_{ij} < d_j)$$

,

where δ_{ij} is the distance between site *i* and observation *j*, d_j is the bandwidth and $I(.)$ is an indicator function equal to one when the condition is true. The effect of any one location in space on near points thus falls depending on the distance and becomes zero once the distance exceeds the bandwidth. There are other common weighting schemes like the Gaussian function or several Kernel weighting, but it seemed to be that opting for a proper bandwidth is more significant than choosing a certain spatial weighting function functions (see: Anselin et al., 2004; or Fotheringham et al., 2002).

When there is no prior justification for applying a particular bandwidth, an appropriate bandwidth can be found by the minimizing either the cross-validation score (CV), the Akaike Information Criterion (AIC) or the Schwartz Criterion (SC). The AIC and the SC are offered by most software packages. The CV is calculated as:

$$CV = \sum_{i=1}^{n} \left(y_i - \hat{y}_{i \neq i} \right)^2 ,$$

where n is the number of data points and the prediction for the ith data point $\hat{y}_{i \neq i}$ is obtained with the weight for that observation set to zero. Each of the criteria can be minimized by a golden section search (see Press et al. 1989). In our study all criteria led to similar results. We opted to minimize the Schwartz Criterion, because according to Boots et al. (2002) it seems to have better large sample properties.

In typical applications, sites and observations would be identical. In our context, that would require estimates per crop and CORINE class for each LUCAS point, which is computational impossible. Instead, the NUTS 2 regions were chosen as sites. When estimating a particular NUTS 2 region, all LUCAS point inside that NUTS 2 region received uniform unity weight, and points in neighboring NUTS 2 regions weights received equal to or smaller than unity according to the tri-cube weighting function. Weighting each likelihood contribution with ω_{ij} gives (Fotheringham et al., 2002):

$$\log L = \sum_{i=1}^{n} \omega_{ij} [y_i \log \Lambda(\boldsymbol{\beta}_{\mathbf{j}}^{'} \mathbf{x_i}) + (1 - y_i)(1 - \Lambda(\boldsymbol{\beta}_{\mathbf{j}}^{'} \mathbf{x_i}))]$$

3.5 Attaining variance of land use shares

Given the non-linear character of the estimations, the variance-covariance matrices offered by the statistical packages are not analytically calculated but instead numerically approximated which proved to be not suitable in the overall approach. Quite small predicted mean values in combination with incredibly high variances led to shaky final results, since high variances result in low penalties in the reconciliation step. Consequently, the estimation of the prior variance attracts our attention. Statistical formulas can be used to derive the variance of a predicted mean.

The prior variance $\hat{\mathbf{Y}}$ is based on the asymptotic covariance matrixes for the coefficients. A robust covariance matrix can be calculated analytically (see White, 1982; Green, 2003, p. 673):) as:

$$V_\beta = Cov\left|\hat{\boldsymbol{\beta}}\right| = \hat{\mathbf{H}}^{-1} \hat{\mathbf{B}} \hat{\mathbf{H}}^{-1} ,$$

where for the weighted logit model the elements of Hessian ***H*** and the Brendt, Hall, Hall and Hausman matrix ***B*** are given by (Green, 2003, p. 672):

$$\mathbf{H} = \frac{\partial^2 LogL}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^{'}} = -\sum_i \omega_i \Lambda_i (1 - \Lambda_i) \mathbf{x_i} \mathbf{x_i^{'}} ,$$

$$\mathbf{B} = \sum_i \omega_i (y_i - \Lambda_i)^2 \mathbf{x_i x_i'}$$
.

As insignificant parameter estimates might influence the efficient calculation of a robust covariance matrix although they do not influence the forecasted value, insignificant variables were removed from the estimations. The variance of $\hat{\mathbf{Y}}$ builds upon the calculated covariance matrix $\mathbf{V_\beta}$ (Green, 2003, p. 674):

$$\mathbf{V_Y} = Var[\hat{\mathbf{Y}}] = \Lambda_i (1 - \Lambda_i) \mathbf{x}' \mathbf{V_\beta} \mathbf{x}$$

Using specific characteristics of a HSMU x_h yields variances of the predicted land use share in each HSMU.

3.6 Consistent disaggregation

The second step of the disaggregation procedure identifies crop shares in each HSMU using the prior information on the estimated crop shares from the first estimation step under two data constraints: Firstly, adding up the areas per crop in each HSMUs must recover the cropping areas *CA* for that crop at NUTS 2 level. Secondly, the posterior shares in each HSMU must add to unity, including all non-agricultural land use from the LUCAS survey aggregated to the category "OTHER". In opposite to the first step, this requires simultaneous accounting for all crops *c* in all relevant HSMUs *h*. The notation is therefore extended, for example from *Y* to $Y_{c,h}$.

The crop areas in each HSMU $C_{c,h}$ are defined by multiplying the posterior shares $Y_{c,h}^{con}$ with the entire area A_h thus,

$$Y_{c,h}^{con} A_h = C_{c,h}$$

Crop areas of the HSMU are then aggregated to the crop area corresponding Nuts region $C_{c,N}$:

$$\sum_{h \in N} C_{c,h} = C_{c,N}$$

Crop shares of single crops must be in line with the corresponding groups of crops $Y_{g,h}^{con}$:

$$\sum_{c \in g} Y_{c,h}^{con} = Y_{g,h}^{con}$$

and the adding up of crop shares to unity must be imposed:

$$\sum_C Y_{c,h}^{con} = 1$$

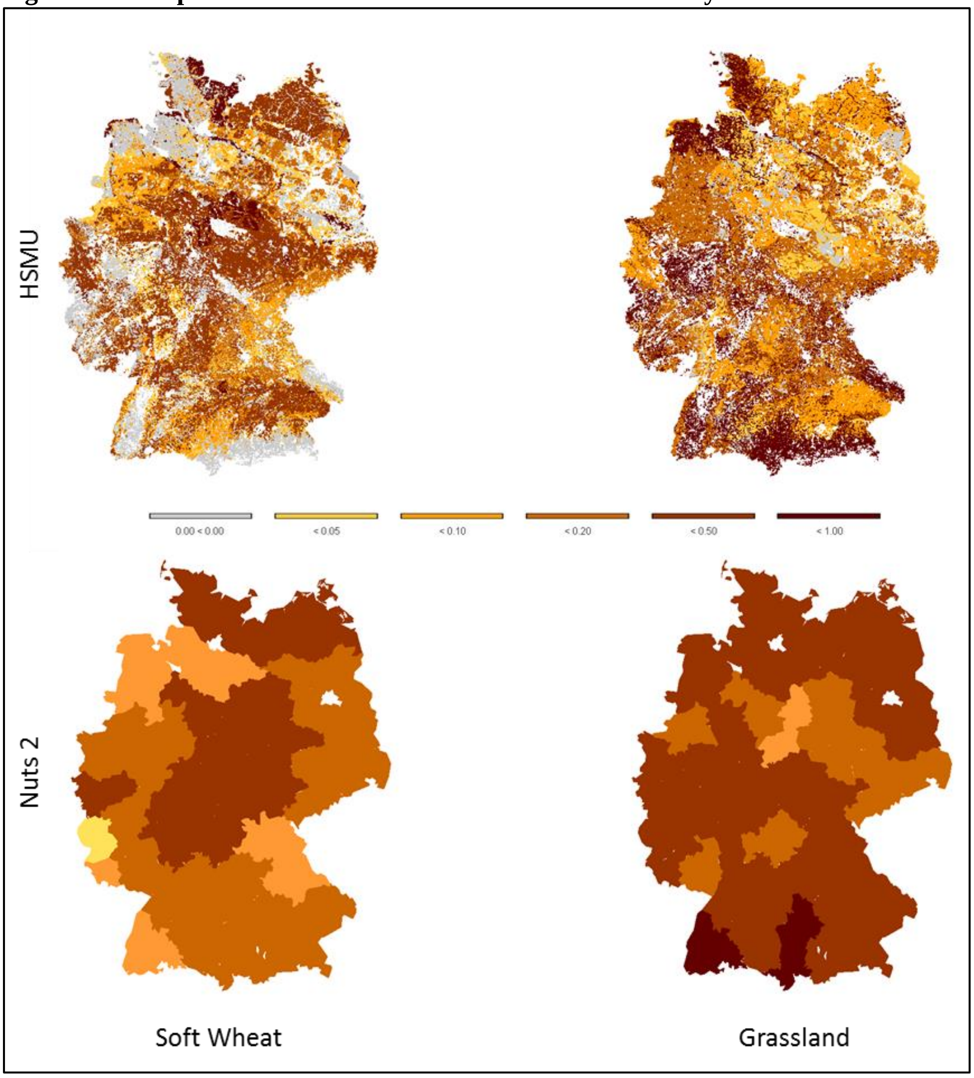
As the predicted unrestricted shares will typically violate the constraints, a penalty function is necessary to define the optimal deviations from the predictions. Generalized Maximum Entropy (GME) techniques (Golan, Judge and Miller, 1996) have often been used for this type of data balancing exercises in recent times. Here, however, a *Bayesian highest posterior density (HPD) estimator* is applied allowing for a direct and transparent formulation of prior information and considerably reducing the computational complexity compared to the GME approach (Heckelei et al.,2005 and 2008). The prior information for singe crops c and groups of crops g is expressed as normal densities of predicted shares, with mean vector $\hat{\mathbf{Y}}_{cg,h}$ and variance derived by the methods described before. After taking logs, the prior density function for the consistent shares $Yc^{con}_{g,h}$ is:

$$-\sum_{cg}\sum_{h}\left[\log\left(\sqrt{2\pi}\mathbf{V}_{\mathbf{Y}_{cg,h}}\right)+\frac{\left(\mathbf{Y}^{con}_{cg,h}-\hat{\mathbf{Y}}_{cg,h}\right)^{2}}{2\mathbf{V}_{\mathbf{Y}_{cg,h}}{}^{2}}\right]$$

4 Results and validation

The basic outcomes of the disaggregation procedure are maps of crop shares for 30 activities in 100.000 HSMUs at a resolution of 1x1 km. Figure 2.2 presents exemplary results for soft wheat and grassland in Germany.

Figure 2.2: Crop shares at HSMU and Nuts 2 level in Germany

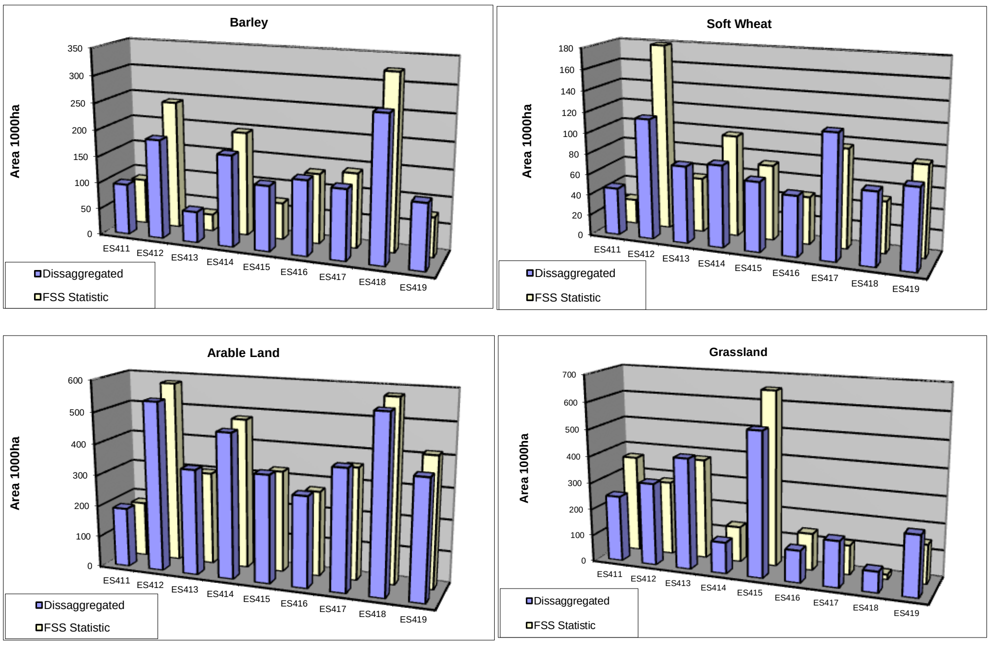


Source: Own compilation.

As expected the spatial heterogeneity of crop shares significantly increases at the resolution of 1x1 km. It can be expected that the results of spatially explicit bio-physical models are improved, if the accuracy of the disaggregation procedure is acceptable. Therefore various approaches of validation are discussed in the following sections.

For some European regions, land use statistics at a lower administrative level, called NUTS 3, are available from the farm structure survey (FSS; EUROSTAT, 2002). This information is used as out-of-sample observation to validate the results of the disaggregation algorithm, which predicts cropping areas for the HSMUs consistent to NUTS 2[18]. Adding up over the corresponding HSMU yields crop areas at NUTS 3 level that can be compared to the observed data. Figure 2.3 exemplary contrasts actual and predicted cropping areas for selected crops in the nine NUTS 3 regions in Castilla-Leon, Spain. Although the disaggregation reflects the principal pattern quite well there are sometimes large differences.

Figure 2.3: Comparison of estimated and observed shares in NUTS 3 region for different crops in NUTS 2 region Castilla Leon ES410



Source: Own compilation.

Error assessment analyses of the agricultural land use maps have been performed both at the regional scale, using district to regional-scale statistics from an agricultural census of the year 2000 covering the EU15 Member States, and at the local scale, using commune-level statistics of the Lombardia region in Italy and in the Netherlands.

[18] Usually this disaggregation procedure is applied to the complete and consistent NUTS 2 database of the CAPRI modelling system, but any other statistic can be used as well. In order to allow a consistent analysis based on FSS NUTS 3 data the corresponding NUTS 2 information was used here.

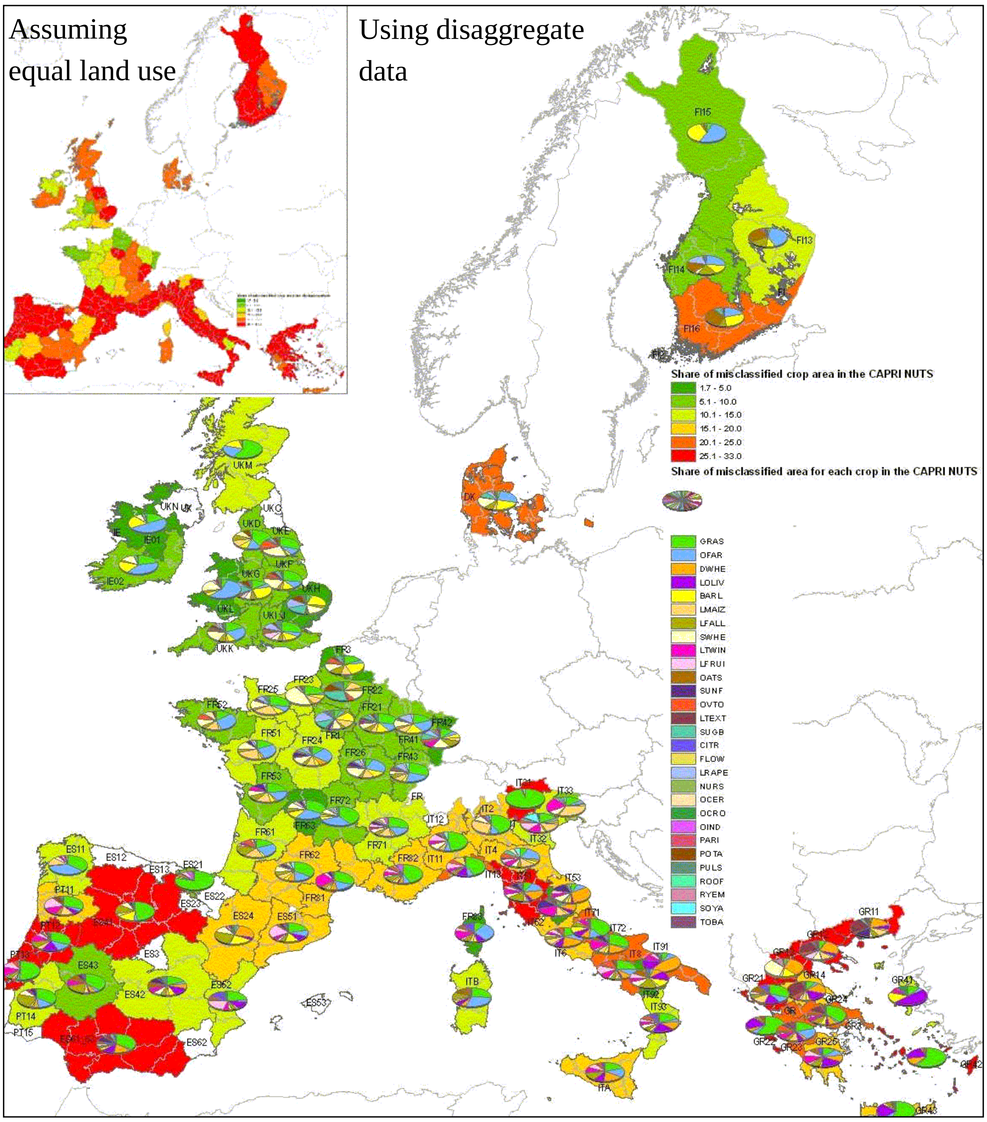
More specifically, the disaggregation results were compared with the data from the FSS agricultural census for the year 2000 (FSS2000, European Commission, 2003b). For some European regions, land use statistics from the FSS2000 are available at a lower administrative level, i.e. NUTS 3. Within the area where both data sets were available the NUTS 2 regions are subdivided into a minimum of 2 and a maximum of 10 NUTS 3 regions. For the comparison, distribution results at the HSMU level were aggregated to the NUTS 3 level and compared with the FSS2000 statistics as out-of sample data. For each individual crop, the difference between the crop area given by FSS2000 and the area of the disaggregation result was calculated. All positive area differences were summed up for all crops and expressed as percentage of the total NUTS 2 agricultural area. In this way we obtained the share of misclassified agricultural area in a NUTS 2 region for all regions where FSS2000 data at NUTS 3 level were available (see Figure 2.4 and Table 2.2).

Table 2.2: Percentage of misclassified area for different crops in Castilla Leon ES410

Crop	Missclassified Area in Nuts III (% of UAA in Nuts II)			
	Single Crop	Groups		
Soft Wheat	3.39	6.44	Arable Land 3.96	UAA 8.43
Durum Wheat	0.32			
Barley	5.02			
Rye	0.87			
Oats	1.46			
Maize	1.24			
Other Cereal	0.06			
Fallow Land	2.95	2.95		
Rice	0.00	1.09		
Sunflower	1.19			
Soya	0.00			
Texture Crops	0.59			
Pulses	0.34			
Other Crops	0.00			
Potatoes	0.24	0.83		
Sugar Beet	0.60			
Root Crops	0.01			
Rape	0.02			
Tobacco	0.01	0.18		
Other Industrial	0.04			
Tomatoes	0.00			
Other Vegetable	0.17			
Flowers	0.00			
Other Fodder	1.88	Fodder Production 10.15		
Grassland	9.27			
Nursery	0.01	Permanent Crops 0.4		
Fruits	0.10			
Citrus	0.00			
Olive	0.11			
Vine	0.40			
Nuts II	**30.29**	**22.39**	**14.51**	**8.43**

Source: Own compilation.

Figure 2.4: Percentage of misclassified areas in validated NUTS 2[19].



Source: Own compilation.

In regions with a high percentage of misclassified area often grassland accounts for a significant part of the errors. This is somewhat astonishing since grassland has its "own" CORINE land cover class and indicates that misclassification might not only be a consequence of a poor disaggregation procedure but also a result of problematic data sources[20].

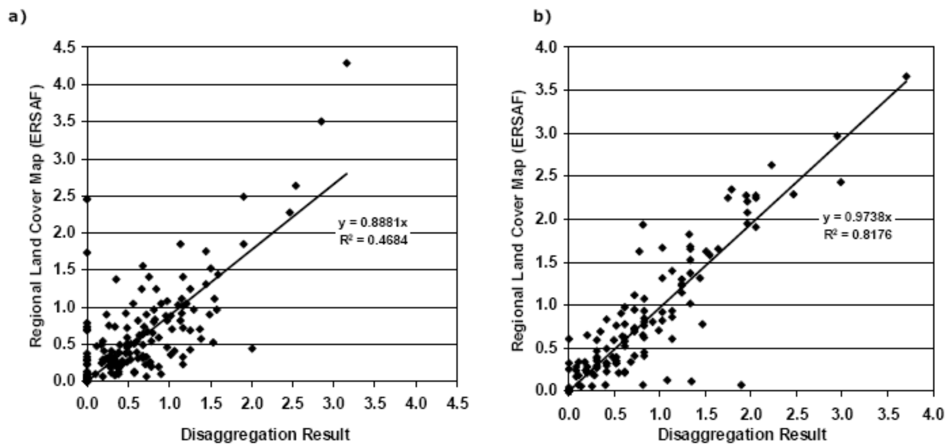
[19] Note: The pies show the contribution of different crop groups to the total error in the region (Cereals: soft wheat, durum wheat, barley, rye, oats, maize, other cereal; Fallow: fallow land; Rice and Oil Seeds: rice, sunflower, soya, texture crops, pulses, other crops; Root Crops: potatoes, sugar beet, root crops, rape, nurseries; Permanent/Industrial Crops: tobacco, other industrial, vegetables, flowers, citrus trees, fruit trees, olive trees, vineyards; Grassland: grassland, fodder production). Note that the size of the pie is related to the area of the NUTS 2 region for visualization purposes only.

[20] The CORINE land cover map reports indeed about 2 mio hectares "Pasture" and "Natural Grassland" in Spain while, in the FSS statistic, about 9 mio hectares of grassland are declared.

Nonetheless the disaggregation is a significant improvement compared to the assumption of identical cropping pattern within each NUTS 2 region (see upper left in Figure 2.4)

In addition, the pie charts in Figure 2.4 depict the contribution of each crop to the total error. The misclassified agricultural area within NUTS 2 regions ranges between 2 and 35%. We obtained an area-weighted mean error of _12.2% for Europe. With the developed disaggregation procedure very good results (2–15% misclassified area) have been obtained for United Kingdom, Ireland, France and southern Spain. The errors are slightly higher in northern/central Spain and Portugal. For southeastern Italy, Greece and some regions in Sweden and Finland errors of about 25–35% occur. The higher errors in Sweden and Finland can be explained by the very small agricultural area which has to be located in quite large HSMUs. High errors can be also a consequence of inaccuracies and inconsistencies in the input data for the dis-aggregation (CORINE land use/cover, LUCAS survey, agricultural statistics, etc.).

Figure 2.5: Comparison of communal data (ERSAF, 2005) and dis-aggregation results in the Italian Pavia province (Mortara, IT208) for the 190 single communes.

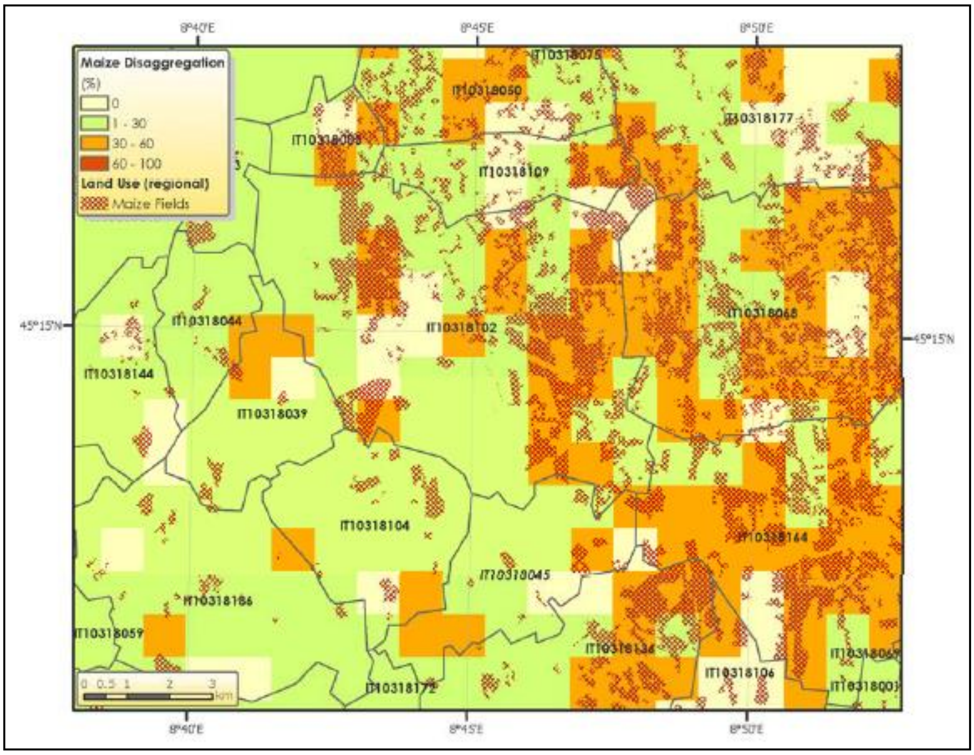


Note: Maize (a) and rice (b) distribution as percentage of the total maize (rice) area within the province.
Source: Own compilation.

Error assessments of the agricultural land use maps have also been performed at the local scale, using year 2003 commune-level statistics of the Lombardia region in Italy (ERSAF, 2005) and Netherlands. We present the former results here.

For the Lombardia region, we compared the rice and maize distribution in 190 communes with the results of the disaggregation. For illustration, Figure 2.5 shows the disaggregation result (1 km×1 km grid resolution) and the maize fields based on data for a set of communes (ERSAF, 2005). The maize pattern (light brown areas) indicating a maize share of 30% from the disaggregation result corresponds with the main maize field distribution based on ERSAF. But looking at the scatter plot (Figure 2.6) comparing ERSAF and disaggregation data for maize in all 190 communes, it can be seen that generally the disaggregation blurs the distribution that is more distinct in reality. To interpret this comparison, however, one has to keep in mind that in this region the areas of the single communes are close to the mean HSMU area and sometimes even larger. Our approach does not allow distributing crop area below the HSMU level and therefore some discrepancies are unavoidable. Thus, we reach herewith the maximum level of detail that can be considered. Furthermore maize is a crop that has no single corresponding CORINE land cover class in which it occurs but is distributed over a range of classes. The contrary holds for rice as a separate rice field class is given in CORINE thus the disaggregation for rice (Figure 2.5) corresponds closely to the communal data. We learned from this comparison that a large portion of the error was introduced when resampling the original CORINE land cover map at the resolution of 100m into the 1 km×1 km pixels. This was necessary because of computing resources, as CORINE was used for the delineation of the HSMU. By leaving out the CORINE layer in constructing HSMUs but using instead percentages of each CORINE class as an attribute of new mapping units, we expect to improve future versions. As a positive side effect, this would reduce the numbers of mapping units making the reconciliation step simpler and faster.

Figure 2.6: Disaggregation result for maize and maize fields given in the ERSAF (2005) agricultural land use map.



Note: The grey borders outline individual communes.
Source: Own compilation.

5 Conclusions

Our study has shown that meaningful disaggregation results for crop shares can be obtained from the combination of available data sources with the HSMU approach and the appropriate statistical methodology. The procedure was developed as a template model that works for each NUTS 2 region in the entire EU15 and can be expanded to any region where the same type of data is available. The approach was further developed to make also yield spatially explicit and, finally, to link input use such as fertilizers to the results (Britz and Witzke, 2008). These combined results were then fed into bio-physical models to replace previously used crude passive indicators at NUTS 2 level by results from the bio-physical application. (Leip et al., 2008).

The construction of the a priori information for the reconciliation step requires an estimation model allowing for spatial heterogeneity. A locally weighted maximum likelihood estimator is proposed here, and requires considerable computational resources as repeated estimations of about 15.000 equations are necessary to define the optimal bandwidth for the approach. The investigation of different model specifications (e.g. nested multinomial logit or semi parametric approaches) might improve the results in future.

The framework of Highest Posterior Density estimation proves to be a powerful tool regarding the reduction of computational complexity compared to widely used cross entropy approaches. It comes up with well-behaved results. The a priori information of the estimation step can directly be used since their distribution is known. The HPD framework depends on accurate estimates of mean and variance. Since the constraints in the optimization link every crop in every HSMU with each other, a single outlier in the estimation can harm final results considerably. Visual checking of the prior and posterior crop shares reveal that the HPD framework is rather sensitive against outliers in estimating the variance. Alternative approaches have been tested and compared. Kempen et al. (2006) reveals slightly better properties when using bootstrap procedures. Conceptually a multivariate distribution should be superior over a univariate distribution as used here, but all attempts to specify a proper covariance matrix failed so far.

The CORINE land cover map as well as the LUCAS survey were repeated in the following years. Since the land use map proved to be useful in analyzing environmental impacts of agriculture an update using most recent data was launched in 2012 by the Joint Research Center in Ispa (Italy). Shortcomings of the current approach will be addressed and the methodology will be further developed. New results can be expected in 2013.

6 References

Anselin, L., Florax, R.J.G.M., Rey, S.J. (2004): Advances in Spatial Econometrics, Springer Verlag, Berlin.

Boots, B., Okabe, A., Thomas, R. (2002): Modelling Geographical Systems, Kluwer Academic Publishers, Dordrecht.

Cleveland, W.S., Devlin, S.J. (1988): Locally Weighted Regression: An Approach to Regression Analysis, Journal of the American Statistical Association, Vol. 83, No. 403 (Sep., 1988), pp. 596- 610.

European Commission (2003): The Lucas survey. European statisticians monitor territory. Theme 5: Agriculture and fisheries, Series Office for Official Publications of the European Communities, Luxembourg.

European Topic Centre on Terrestrial Environment (2000): Corine land cover database (Version 12/2000).

Eurostat (2000): Manual of Concepts on Land Cover and Land Use Information Systems. Theme 5: Agriculture and Fisheries: Methods and Nomenclatures. Series Office for official Publications of the European Communities, Luxembourg, 2000, pp. 110-ff.

Fahrmeier, L., Tuts G. (1994): Multivariate Statistical Modelling Based on Generalized Linear Models, Springer Verlag, New York.

Fotheringham, A. S., Brusdon, C., Charlton, M. (2002): Geographically Weighed Regression, John Wiley & Sons, Chichester.

Gallego, J. (2002): Fine scale profile of CORINE Land Cover classes with LUCAS data, in European Commission (ed.): Building Agro Environmental Indicators. Focussing on the European area frame survey LUCAS, Vol. EUR Report 20521 EN.

Golan, A., G. Judge, Miller, D. (1996): Maximum Entropy Econometrics, Chichester UK: Wiley.

Greene, W.H. (2003): Econometric Analysis, Macmillian Publishing Company, New York.

Heckelei, T., Mittelhammer, R.C., Britz, W., (2005): A Bayesian Alternative to Generalized Cross Entropy Solutions to Underdetermined Models. Contributed paper presented at the 89th EAAE Symposium „Modelling agricultural policies: state of the art and new challenges“, February 3-5, Parma, Italy.

Heckelei, T., Mittelhammer, R.C., Jansson, T. (2008): A Bayesian Alternative to Generalized Cross Entropy Solutions for Underdetermined Econometric Models, Discussion Paper. (Downloadable at http://www.ilr.uni-bonn.de/agpo/publ/dispap/download/dispap08_02.pdf)

Howitt, R., Reynaud A. (2003): Spatial Disaggregation of Agricultural Production Data by Maximum Entropy. European Review of Agricultural Economics 30(3):359-387.

Kempen, M., Heckelei, T., Britz W. (2006): Validation of Spatially Explicit Land Use Choices Based On Probabilistic Theory. Paper presented at the International Conference on Regional and Urban Modeling Brussels, June 1-2, 2006.

Kempen, M., Heckelei, T., Britz, W., Leip, A., Koeble, R. (2005): A Statistical Approach for Spatial Disaggregation of Crop Production in the EU, in „Modelling agricultural policies: state of the art and new challenges“, MUP, Parma, Italy.

Montzka, C., Cantya, M., Kunkela, R., Menzb, G., Vereeckena, H.,Wendlanda, F. (2008): Modelling the water balance of a mesoscale catchment basin using remotely sensed land cover data. Journal of Hydrology ,Volume 353, Pages 322–334.

Mulligan, D.T. (2006): Regional modelling of nitrous oxide emissions from fertilised agricultural soils within Europe. Series PhD thesis, submitted to the University of Wales, Bangor.

Nelson, G., De Pinto, A., Harris, V., Stone, S. (2004): Land Use and Road Improvements: A Spatial Perspective, International Regional Science Review.

Orlandi, S., Van der Goot, E.(2003): Technical description of interpolation and processing of meteorological data in CGMS, Available under http://agrifish.jrc.it/marsstat/Crop Yield Forecastingcgms.htm, European Commission, DG JRC, Agrifish Unit.

Press, W.H., Flannery, B.P., Teukolsky, S.A. and Vetterling, W.T. (1989): Numerical Recipes - The Art of Scientific Computing. Cambridge University Press, Cambridge, UK.

Roekaerts, M. (2002): The Biogeographical Regions Map of Europe, The data and document can be found on http://dataservice.eea.eu.int/dataservice/.

van Diepen, K., Hazeu, G., (2005) personal communication November 2005 (see Appendix).

White, H. (1982): Maximum Likelihood Estimation of Misspecified Models, Econometrica, 53: 1-16.

7 Appendix: Note on clustering of STUs for HSMU analysis.

By Kees van Diepen and Gerard Hazeu 18-November-2005

Received data sets and data needs

Markus Kempen provided Alterra with a list of 95 WRB soil unit codes of the STUs that occurred in the intersection of the SGDBE and the applied selection of 38 Corine land cover classes that cover the EU15. The question was how these 95 soil codes could be clustered into up to 20 groups of soils with similar land use potential.

The clustering method

The clustering was achieved in two steps:

A first clustering was made based on Driessen et al (2001) who rearranged the 30 WRB soil groups into 10 so-called Sets, based on the dominant soil forming factors that determined the soil conditions. The list of 95 soil codes could be aggregated into 9 of those Sets.

Next, as some Sets were heterogeneous in its characteristics relevant for land use we have subdivided those Sets further into more homogeneous Sets, while other Sets were combined into one as the soils were rather similar from the land use perspective. Finally, from the resulting Sets some soil units have been placed in other Sets because of some common prominent feature. .The distinction of new Sets was based on similarity of STUs in terms of soil characteristics which are judged relevant for land use, notably rooting depth, organic matter, texture, drainage class, water holding capacity, presence of stones and slope. The grouping was based on judgment. In order to maintain the logic of the distinction of the soil units on the soil map, we preferred that the Sets were defined by the highest hierarchical level in the WRB, the Soil Reference Group and then by the second level of the Soil Units.

This process has resulted in the following 15 soil Sets and one non-soil Set listed in the following Table under the header SET2. In this SET2 subdivision:

Set 0 holds all non-soils (Towns, water, glaciers, rocks, marshes)

Set 1 holds all organic soils (Histosols)

Set 2 holds all soils which have in common a high content of sand (Arenosol, Podzol, Arenic Umbrisol, Plaggic Anthrosol)

Set 3 holds all Regosols, characterized as soils from uplands developed in unconsolidated materials, in itself a very heterogeneous cluster

Set 4 holds all shallow soils typically occurring in sloping rocky landscapes (Leptosols)

Set 5 holds most Cambisols, which have as common feature that they are relatively young soils, Cambisols are not related to any specific landscape position.

Set 6 holds the soils of the forest-steppe transition zone, with dark topsoils Chernozems and Phaeozems (climate feature is the equilibrium in the annual moisture balance)

Set 7 holds the soils of the drier steppe, with dark topsoils, Kastanozems. (the climate feature is a water deficit in the annual moisture balance)

Set 8 holds the Albeluvisols: medium textured soils of the humid temperate region where leaching is the dominant process. Albeluvisols have soil properties in between Luvisols and Planosols

Set 9 holds the Luviosols: Medium textured soils of the humid temperate region where leaching of clay is a prominent process. Luvisols have sandy topsoil and clay enriched subsoil.

Set 10 holds the Planosols, medium textured soils of the humid temperate region where leaching is the dominant process. Planosols have white sandy topsoil and dense clayey subsoil.

Set 11 holds heavy clay soils that swell when wet and shrink when dry. This Set 11 contains the Vertisols and the vertic subgroups of Cambisols and Luvisols.

Set 12 holds all soils associated with flooding or wetness, usually located in lowlying flat terrain Fluvisols and Gleysols.

Set 13. Holds all Andosols (soils developed in volcanic deposits)

Set 14 Hold all soils with characteristics of subtropical weathering, leading to deep red soils of relatively low fertility (Acrisols and Alisols)

Set 15 holds soils of arid and semi-arid regions characterized by a shortage of water and accumulation of salts, lime or gypsum in the soil or at the surface (Solonchaks, Solontchak, Gypsisol and Calcisol).

The definition of the 15 sets in SET2 is meant as a preliminary initial set as a start for further testing and analysis by comparing the occurrence of soil units with land use within a NUTS2

On the other hand, some Sets which include a relatively large number of STUs in EU15 may still be split, e.g the Dystric soil units may be separated from the other units in the same set, and be combined with Dystric units from other Sets. Dystric refers to chemical soil conditions, and because the criteria on which the homogeneity check was carried out were mainly physical, a difference between Dystric and the other soil units was s not observed, but it could be relevant to explain differences in land use pattern. Also the relevance of gleyic units for the distinction of land use could be evaluated only in an analysis where land use class is paired to soil unit.

Literature

Food and Agriculture Organization of the United Nations , 1998. World Reference Base for Soil Resources World Soil Resources Reports 84. ISSN 1014-8531. FAO, ISRIC and ISSS.

Driessen, Paul, Joseph Deckers, Otto Spaargaren and Freddy Nachtergaele.2001 Lecture Notes on the major Soils of the World. World Soil Resources Reports 94. ISSN 0532-0488. FAO.

Table: Clustering of the STU´s

WRB Code	Soil Unit Qualifier	WRB Soil Ref Group	SET Driessen et al	SET DynaSpat
N1	Town	Town	0	0
N2	disturbed	disturbed	0	0
N3	Water	Water	0	0
N4	Marsh	Marsh	0	0
N5	Glacier	Glacier	0	0
N6	Rock	Rock	0	0
HS	Histosol	Histosol	1	1
HSdy	Dystric	Histosol	1	1
HSeu	Eutric	Histosol	1	1
HSge	Gelic	Histosol	1	1
ARab	Albic	Arenosol	3	2
ARha	Haplic	Arenosol	3	2
ATpa	Plaggic	Anthrosol	2	2
PZgl	Gleyic	Podzol	9	2
PZha	Haplic	Podzol	9	2
PZle	Leptic	Podzol	9	2
PZpi	Placic	Podzol	9	2
PZum	Umbric	Podzol	9	2
UMar	Arenic	Umbrisol	9	2
RGca	Calcaric	Regosol	4	3
RGdy	Dystric	Regosol	4	3
RGeu	Eutric	Regosol	4	3
RGha	Haplic	Regosol	4	3
LPca	Calcaric	Leptosol	4	4
LPdy	Dystric	Leptosol	4	4
LPeu	Eutric	Leptosol	4	4
LPha	Haplic	Leptosol	4	4
LPrz	Rendzic	Leptosol	4	4
CMca	Calcaric	Cambisol	5	5
CMcr	Chromic	Cambisol	5	5
CMdy	Dystric	Cambisol	5	5
CMeu	Eutric	Cambisol	5	5

WRB Code	Soil Unit Qualifier	WRB Soil Ref Group	SET Driessen et al	SET DynaSpat
CMgl	Gleyic	Cambisol	5	5
CMha	Haplic	Cambisol	5	5
CMmo	Mollic	Cambisol	5	5
CMvr	Vertic	Cambisol	5	11
CHcc	Calcic	Chernozem	8	6
CHgl	Gleyic	Chernozem	8	6
CHha	Haplic	Chernozem	8	6
CHlv	Luvic	Chernozem	8	6
PHca	Calcaric	Phaeozem	8	6
PHgl	Gleyic	Phaeozem	8	6
PHha	Haplic	Phaeozem	8	6
PHlv	Luvic	Phaeozem	8	6
KScc	Calcic	Kastanozem	8	7
KSha	Haplic	Kastanozem	8	7
KSlv	Luvic	Kastanozem	8	7
ABeun	Endoeutric	Albeluvisol	9	8
ABgl	Gleyic	Albeluvisol	9	8
ABha	Haplic	Albeluvisol	9	8
LVab	Albic	Luvisol	9	8
LVar	Arenic	Luvisol	9	9
LVcc	Calcic	Luvisol	9	9
LVcr	Chromic	Luvisol	9	9
LVdy	Dystric	Luvisol	9	9
LVfr	Ferric	Luvisol	9	9
LVgl	Gleyic	Luvisol	9	9
LVha	Haplic	Luvisol	9	9
LVvr	Vertic	Luvisol	9	11
PLdy	Dystric	Planosol	9	10
PLeu	Eutric	Planosol	9	10
PLha	Haplic	Planosol	9	10
VRcr	Chromic	Vertisol	3	11
VRha	Haplic	Vertisol	3	11

WRB Code	Soil Unit Qualifier	WRB Soil Ref Group	SET Driessen et al	SET DynaSpat
VRpe	Pellic	Vertisol	3	11
FLca	Calcaric	Fluvisol	4	12
FLdy	Dystric	Fluvisol	4	12
FLeu	Eutric	Fluvisol	4	12
FLgl	Gleyic	Fluvisol	4	12
FLha	Haplic	Fluvisol	4	12
FLmo	Mollic	Fluvisol	4	12
FLti	Thionic	Fluvisol	4	12
GLca	Calcaric	Gleysol	4	12
GLdy	Dystric	Gleysol	4	12
GLeu	Eutric	Gleysol	4	12
GLha	Haplic	Gleysol	4	12
GLhu	Humic	Gleysol	4	12
GLmo	Mollic	Gleysol	4	12
GLti	Thionic	Gleysol	4	12
ANdy	Dystric	Andosol	3	13
ANhu	Humic	Andosol	3	13
ANmo	Mollic	Andosol	3	13
ANvi	Vitric	Andosol	3	13
ACfr	Ferric	Acrisol	6	14
ACgl	Gleyic	Acrisol	6	14
ACha	Haplic	Acrisol	6	14
AChu	Humic	Acrisol	6	14
ACpl	Plinthic	Acrisol	6	14
ALpl	Plinthic	Alisol	6	14
CLad	Aridic	Calcisol	7	15
GYad	Aridic	Gypsisol	7	15
SCgl	Gleyic	Solonchak	7	15
SCha	Haplic	Solonchak	7	15
SCty	Takyric	Solonchak	7	15
SNmo	Mollic	Solonetz	7	15

Chapter 3: Spatial allocation of farming systems and farming indicators in Europe[21]

Abstract

In this article an approach to spatially allocate farm information to a specific environmental context is presented. At this moment the European wide farm information is only available at a rather aggregated administrative level. The suggested allocation approach adds a spatial dimension to all sample farms making it possible to aggregate farm types both to natural and to lower scale administrative regions. This spatial flexibility allows providing input data to economic or bio-physical models at their desired resolution. The allocation approach is implemented as a constrained optimization model searching for an optimal match between farm attributes and spatial characteristics subject to consistency constraints. The objective functions are derived from a Bayesian highest posterior density framework. The allocation procedure recovers the spatial farm type distributions quite well, thereby providing information of significant value for further analysis in a multidisciplinary context.

1 Introduction

Environmental benefits and decentralized policy implementation become more important in the reorientation of European Common Agricultural Policy (CAP) making integrated assessment of agricultural policy measures increasingly relevant (Van Ittersum et al., 2008). Integrated assessment models combine economic and environmental models which could both benefit from spatial explicit land use and management data. Environmental effects are often modelled by process based bio-physical models whose results depend on the spatial resolution of input data (Mulligan, 2006). Policy measures are increasingly targeted at specific areas like Nitrate Vulnerable Zones (NVZ), Less Favoured Areas (LFA) or NATURA2000 regions.

While maps on land cover and partially land use can be based on remote sensing, other data on agriculture are often available for administrative regions only, matching neither the boundary of specific targeted areas nor the spatial resolution required in environmental modelling. Large scale studies often apply downscaling techniques to get information at the relevant scale, since comprehensive field studies are too costly.

In recent years efforts have been made to estimate a European wide land use map. Howitt and Reynaud (2003) proposed a methodology to predict spatial explicit land use choices at a field level in a two step procedure. First a Markov model is estimated and applied to predict land use choices at field level. Results enter a cross entropy based rec-

[21] This Chapter is based on an article published together with Berien S. Elbersen (Alterra,Wageningen), Igor Staritsky (Alterra,Wageningen), Erling Andersen (University of Copenhagen), and Thomas Heckelei (University of Bonn) in the journal Agriculture, Ecosystem & Environment 1-2 (2011): 51-62.

onciliation procedure ensuring consistency with more aggregate data. This basic approach was adapted to the data bases available in Europe by Kempen et al. (2005 and 2007). Land use shares are derived regressing observed agricultural use on soil climate and topographic information. Spatial estimation techniques are employed to account for non measured characteristics like socio economic conditions. Leip et al. (2008) added information on fertilizer use, manure application and yield to the land use map. However, this downscaled data base defines average agricultural production activities. The heterogeneity within one production activity under different farming systems is not captured. Currently harmonized EU wide farm information is only available at the level of about 150 administrative regions while an allocation to a specific environmental endowment is desirable.

Various attempts have been made to disaggregate farming system information to a desired spatial resolution. Kruska et al. (2003) describe a methodology for mapping livestock-oriented agricultural production systems for the developing world. Since statistical data on livestock production are often completely missing in this case, each location is assigned a farming system based on expert rules. Farming systems are allocated using spatially explicit climate, soil and socio-economic criteria. Van der Steeg et al. (2010) present a methodology to derive a spatially explicit distribution of farming systems in the Kenyan Highlands. Their approach starts with the definition of farming systems based on a sample of about 3000 farms. Since the exact location of each holding is known, a regression model predicting the probability to observe a farming system based on relevant environmental and socio economic drivers is estimated. The estimated model parameters are then used to predict shares of farming systems for the whole study area.

The data availability in the EU differs significantly from developing countries where all previous studies were performed. The Farm Structure Survey (FSS) collects information on the whole population of farms each 2nd or 3rd year and publishes results for administrative regions, called Nomenclature of Territorial Units for Statistics (NUTS) regions. The Farm Accountancy Data Network (FADN) contains 75,000 individual farm records representing all commercial farms in 150 regions. The exact location of the sample farms is not made public due to confidentiality reasons.

This paper contributes to the literature on spatially locating farming systems by developing, applying, and validating a methodology to add a spatial dimension to the FADN sample farms. This spatial dimension is a reference to small scale spatial units, so called Farm Mapping Units (FMU), where relatively homogenous conditions for farming can be expected. For presentation and further use of results in economic or bio-physical analysis, farms and FMUs can be aggregated to any spatial unit or farm typology. The definition of a farm typology is independent from the downscaling procedure making the use of allocation results more flexible and avoids uncertainty that might be related to a classification. Since the allocation approach is based on harmonized data sets the requirements of large scale integrated assessment- approaches are met (Janssen et al., 2009). Our results can serve as input for generic template models on farming systems (Louhichi et al., 2007)

Our farm allocation approach follows a two step procedure. First we derive a prior probability of allocating a specific farm in a certain FMU, then a reconciliation step ensures consistency. Yet we cannot base prior expectation on an empirical model since the exact location of farms is not published. However, farm records include some information limiting the number of FMU where the farm might be allocated. Further we make various assumptions on regional land use areas, land use shares and yields. For example, farms realizing relatively high yields are more likely located in areas where potential yields are high. This seems a plausible assumption, but it could neither be verified from literature nor own empirical data. Furthermore, there is no clear methodology how the difference between realized and potential yield can be translated into an a priori probability. In order to define appropriate prior information we compare various sets of assumptions by validating the results against out-of-sample data. We found one set of prior information almost dominating all other specifications.

Socio economic characteristics are not considered in our prior expectations since we assume these conditions to be relatively homogeneous within our model regions. FADN regions are about half the size of the Kenyan Highlands investigated by van der Steeg et al. (2010) and European infrastructure is likely better. Furthermore, socio economic aspects were captured implicitly in the land use map by Kempen et al. (2007) which is used to formulate parts of the prior information.

The article is structured as follows: First we describe the data base on farms and spatial attributes. Following the allocation procedure is presented in detail. Then a validation of model results is performed to identify suitable settings for the derivation of EU wide results presented following. We finish with a conclusion and an outlook on promising future research in this field beyond the scope of this article.

2 Data

2.1 Farm data

The Farm Accountancy Data Network is a European system of sample surveys conducted every year to collect structural and accountancy data on farms, with the aim of monitoring the income and business activities of agricultural holdings and evaluating the impact of the measures taken under the Common Agricultural Policy. The FADN is the only source of micro-economic data harmonized across the EU, i.e. the same bookkeeping principles apply in each member country. FADN data are collected in about 100 administrative regions which are equal to countries, NUTS 1 or NUTS 2 regions. Exact natural conditions and location of the holdings cannot be derived from the data set mainly for confidentiality reasons. However some elements of the FADN data represent spatial characteristics relevant for our analysis:

For each sample farm, FADN records report whether it is located in a specific altitude zone and in a Less Favoured Area (LFA). Furthermore, many farms are assigned sub region codes which can be used to identify lower levels of administrative units (typically NUTS 2 or NUTS 3). Additionally, the land use patterns and crop yields recorded give hints for the spatial location of the farm.

Farms are selected for the database according to a sampling plan aiming at representativity of the sample for the population in a FADN region with respect to a classification by type of farming, economic size and region. To allow for corrections of deviations from a perfect stratified sampling, an individual weight is provided for each farm in the sample calculated as the ratio between the total number of holdings in the population over the sampled number of holdings in the same classification. The total number of holdings is obtained from the Farm Structure Survey (FSS) collecting information on the whole farm population every two to three years.

With respect to a consistent allocation of farms in space, however, some problems arise from this procedure. Representativity for sub regions, altitude zones and less favoured areas might be aimed for in the selection of farms by local agencies but is not guaranteed by the sampling plan and cannot be achieved with the available individual weights. Furthermore, the FADN survey covers only farms above a minimum size (threshold) which might lead to underrepresentation of agricultural activity in some areas.

2.2 Spatial information

The most important spatial data are provided by the CAPRI-DynaSpat project (see www.capri-model.org). Within this project homogeneous spatial mapping units (HSMU) were defined using a Geographical Information System (Kempen et al., 2007; Leip et al., 2008). For the spatial allocation of the FADN farm information, the land use information and other attributes assigned to the HSMUs in the CAPRI-DynaSpat project are taken as the main input basis. The aim of building HSMUs was to define areas inside an administrative region with approximate homogeneity with respect to land cover, soil and slope. The HSMUs were constructed by overlaying the CORINE land cover map (European Topic Centre on Terrestrial Environment, 2000) with spatial soil (Soil Mapping Units) and slope (5 classes) information. Land use shares and expected yields were assigned to each HSMU by a statistical procedure combining grid observations on land use with available aggregate information at regional level. Information on less favoured areas and altitude zones can be added by overlaying HSMU boundaries with specific maps.

2.3 Land use

Kempen et al., (2005 and 2007) and Leip et al., 2008 describe a statistical approach for spatial allocation of crop production in the EU. The resulting detailed land use map, available for EU25, is a core input for the spatial allocation of the FADN farms. The map provides land use shares on about 30 crops for approximately 150,000 HSMUs. The procedure employed to arrive at this map combines a locally weighted logit model estimating probabilities of observing a certain crop in a HSMU using European wide grid point information on land use and spatial soil- climate - and relief information. In a second step, a Bayesian highest posterior density estimator consolidates these estimates with regional information on crop production. Socio economic factors have been implicitly captured by a spatial estimation technique (Anselin et al., 2004). Important for the approach in this paper is that the uncertainty related to the predicted land use shares can be calculated from the standard errors of the estimators and may serve to adequately formulate relevant prior information..

2.4 Yield

Within the MARS project, yield potentials for specific crops were calculated linking expert information to soil and climate data (Boogaard et al. ,2002, Genovese et al., 2004 and 2007). Potential and rain fed yields for 7 relevant crops are available for each HSMU. A reconciliation procedure described by Britz and Witzke (2008) achieves consistency to regional production statistics. Assuming that potential yields can only be realized when irrigation is applied, average yields and shares of irrigation are estimated simultaneously for each crop at HSMU level.

2.5 Less favoured area and altitude zone

Less favoured areas and altitude zones were not considered in the delineation of HSMUs. However, the percentage of each HSMU belonging to a certain combination of less favoured area and altitude zone can be calculated overlaying HSMU boundaries with maps on LFA boundaries and altitude. As HSMUs are quite small spatial units, many of them belong exclusively to a specific combination of less favoured area and altitude zone. In the other cases one combination is usually dominant. Assigning the dominant attribute to the whole HSMU is considered here a justifiable simplification.

Table 3.1: Description and references of datasets

Name/Description		References and Links
farm data		
FADN	Farm Data Accountancy Network (2005)	European Comission, CD recieved 2009 URL:http://ec.europa.eu/agriculture/rica/index_en.cfm
FSS	Farm Structure Survey (2005) - Structure of agricultural holdings by Nuts region, main indicators	European Comission, download September 2009 URL:http://nui.epp.eurostat.ec.europa.eu/nui/show.do?dataset=ef_r_nuts&lang=en
GIAB	Geographical Information System for Agricultural Businesses (The Dutch IACS database)	Naeff, H.S.D., 2006. Geactualiseerde GIAB Nederland 2005. I. rapport (Ed.), Alterra Wageningen, recieved December 2009
spatial data		
Altitude Zones	Own compilation of altitude zone 0 - 300m, 300 - 600m and >600m bades on Digital Elevation Model	European Comission, JRC-IES Digital Elevation Model (CCM DEM, 250 meters), recieved 2004
CORINE	CORINE Land Cover	CORINE Land Cover 2000 national databases: European Topic Centre on Terrestrial Environment, Torre C5-S, 4a planta, Edifici C - Facultat de Ciencies, Universitat Autònoma de Barcelona, 08193 Bellaterra (Barcelona), Catalunya (Spain). http://terrestrial.eionet.eu.int/CLC2000
LFA	Less Favoured Area	European Comission, JRC, LFA boundaries map, recieved 2006
NUTS	Nomenclature of Territorial Units for Statistics	European Comission: regulation (EC) No 1059/2003 of the European Parliamt and of the council of 26 May 2003 on the establishment of a common classification of territorial units for statistics (NUTS)
SMU	Soil Mapping Unit	European Commission: European Soil Database (version V2.0), CD-ROM EUR 19945 EN, Directorate General Joint Research Centre, Institute for Environment and Sustainability, recieved 2004
MARS	potentail and rain fed yield	Gevonese et al. (2002), recieved 2004
Land use map	Application of the CAPRI modelling system, base year 2004 in September 2009.	Britz et al., 2008 URL:http://www.capri-model.org/spatDown.htm URL:http://afoludata.jrc.ec.europa.eu/DS_Free/AF_Agri.cfm

Source: Own compilation.

3 Methodology

3.1 Overview

FADN data are available for about 150 regions and various years. We develop in the following a template model that can be applied to each region and year independently.

Farms shall be mapped to continuous regions with homogeneous conditions for farming. These regions, so called Farm Mapping Units (FMU), are first defined. Then we develop a procedure to achieve the allocation of farms to these FMUs in order to achieve the highest possible consistency between characteristics of FMUs and allocated farms is achieved. Our allocation approach is a two step procedure. First we measure the statistical fit of certain characteristics between all available farms in a FADN region and the corresponding FMUs. Then a reconciliation step ensures consistency by maximizing the similarity over all farms and FMUs. For this purpose, a Bayesian highest posterior density concept (see Heckelei et al., 2008) is applied allowing to measure "similarity" with respect to several criteria simultaneously satisfying regional consistency constraints. The model specification aims at allocating a specific farm exclusively to one FMU. The motivation for this is to identify those farms that might represent the FMU in the best way. To find the best model set up with respect to certain issues we test out various settings and validate the results against out-of-sample data.

For confidentiality reasons, results cannot be shown for single farms. For presentation and further use of the results, the spatially allocated FADN farms are aggregated to farm types (see Andersen et al., 2006 and 2007) and then to larger agri-environmental zones (Hazeu et al., 2006).

3.2 Definition of Farm Mapping Units (FMU)

The spatial information is available as attributes of HSMU. Hence we want to build our definition of FMUs upon them. The HSMU were originally delineated by NUTS boundaries, CORINE land cover, soil mapping units and slope classes. However, HSMUs seem to be inappropriate units for mapping farms for reasons of content and computational performance.

The delineation according to CORINE land cover might be too detailed in order to describe an environment where farms can be located. For example, CORINE distinguishes grassland and arable land at a high resolution. A location with a diverse mixture of arable fields and pastures would be scattered in many HSMUs whereas one might expect (similar) mixed farms in appropriately defined FMUs can create continuous regions. While land cover information might be misleading, dominant Less Favoured Area status and altitude zone are key characteristics of farms, but not yet used for delineation.

As the complexity of the allocation procedure increases with the number of mapping units, we had to limit the number of FMUs in order to ensure feasibility in reasonable time. Slope classes can be neglected without loosing much information as we found them to be highly correlated to altitude zones. Other attributes should not be omitted from delineation. Soil mapping is highly relevant for yield and land use. NUTS 2/3 boundaries enable links to regional statistics during the reconciliation step.

Hence we define FMUs as a collection of HSMU which are uniform regarding administrative region, soil mapping unit, dominant less favoured area status and dominant altitude zone. All relevant attributes of the HSMUs are then aggregated to approximately 15,000 FMUs.

3.3 The constrained optimization model

Our basic idea is to allocate a farm if possible exclusively to one FMU, implying only a few farms should be located in a specific FMU. However consistency constraints in the model will sometimes hinder farms to be allocated completely to one FMU. In this case only a certain percentage of the farm's area is located in one FMU and the rest in another. The final result of our allocation procedure is a matrix $p_{f,fmu}$ indicating the percentage of a farm *f* located in a FMU. As a single farm in the FADN sample represents many similar farms, this percentage can also be understood as the share of these farms being allocated to a specific FMU. An obvious constraint in the allocation procedure is that the percentages for each farm over all FMUs must add up to 1:

$$\sum_{fmu} p_{f,fmu} = 1$$

Another obvious constraint refers to the utilizable agricultural area (UAA). The UAA of a FMU should be filled exactly with the UAA represented by the farms assigned to it.

$$UAA_{fmu} = \sum_{\mathrm{f}} p_{f,fmu} weight_f UAA_f$$

where UAA_{f} is the utilizable area operated by a FADN farm, $weight_f$ the representativity weight taken from FADN records, and UAA_{fmu} the agricultural area in a FMU.

3.4 LFA and altitude zone

From the FADN statistics it can be exactly derived which farms are located in a certain altitude zone and in a LFA. This information is taken as fixed and given, i.e. if the FADN farm and the FMU do not belong to the same qualification regarding LFA and altitude zone, $\mathrm{p}_{\mathrm{f,fmu}}$ is set to zero.

Since FADN data do not fully represent the agricultural area in a region, consistency with the area derived from other sources cannot be expected. An adjustment factor is calculated ensuring that the sum over all areas of farms allocated to a certain FMU adds up to its agricultural area.

$$UAA_{fmu} = adjustfactor_{fmu} \sum_{f} p_{f,fmu} weight_f UAA_f$$

The adjustment factor can be interpreted as a reconciliation of farm specific FADN weights. The same scaling factor is applied to all farms characterized by a specific combination of less favoured area status and altitude zone. (see Table 3.2).

Table 3.2: Utilizable agricultural area in Bavaria (Germany) - comparison of FADN and FMU

LFA status	altitude zone	UAA (1000ha)		*adjustfactor*
		FADN	FMU	
less favoured area	0 -300 m	260	418	1.61
	300 -600 m	1176	1533	1.30
	>600 m	323	444	1.37
non less favoured area	0 -300 m	198	284	1.43
	300 -600 m	350	487	1.39
	>600 m	16	22	1.41

Source: Own compilation.

3.5 Yield

In the case of yields, the findings on a single farm should be similar to those in a FMU. It is assumed that yields observed on a farm differ from the average yields because of some random deviation of management from the average technology. It is assumed that for each crop *c*, the observed yield of a FADN farm is an outcome from a normal distribution around the mean $\mu_{c,fmu}$ of the FMU with a variance $\sigma_{c,fmu}$. Whereas the mean yield can be taken from the corresponding HSMU, the variance is unknown. We derived a variance from FMU and FADN farm yield distributions, assuming that the variance is equal over all FMUs.

When mean and variance of the yield per FMU are available, probability density functions (*pdf*) can be applied to measure the chance of observing a certain farm in a certain FMU considering those crops where data is available for the farm and the FMU.

$$pdf_{YIELDf,fmu} = \prod_{c} N(\mu_{YIELD\ c,fmu}, \sigma_{YIELD\ c,fmu})$$

The *pdf* values are unfortunately non-intuitive and numerically difficult to handle since values are most frequently rather small. The absolute *pdf* value differs also systematically with the number of crops grown on farms. The more crops are cultivated on a farm the lower the absolute *pdf* values are in general. Although this should not matter theoretically we observed numerical problems. Assuming that for each farm the *pdf* value is proportional to the probability $p_{YIELD\ f,fmu}$ of observing a farm in a FMU, we get a more convenient number by simply scaling values so that they add up to 1 for each farm. Farms cultivating no relevant crops are assigned equal probabilities for each FMU.

The optimal allocation based on the yield observations can henceforth be found by maximizing

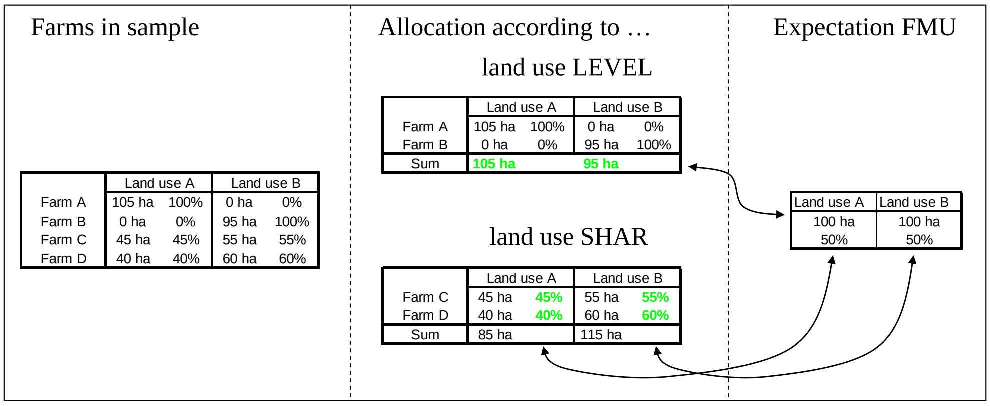
$$obje_{YIELD} = \sum_{f}\sum_{fmu} p_{f,fmu}\, p_{YIELD\ f,fmu} ,$$

where $obje_{YIELD}$ is the value of the objective function with respect to yield information. $p_{YIELD\ f,fmu}$ does not depend on model variables. If there were no constraints, the optimization model will set $p_{f,fmu}$ to 1 in the FMU where the highest value for $p_{YIELD\ f,fmu}$ is calculated.

3.6 Land use

Whereas in the case of yield the observation on a single farm should be similar to those in a FMU, land use information can be interpreted in different ways. On the one hand, it could be assumed that farms in a FMU look alike and therefore the predicted land use shares in FMU should be similar to that of the allocated farms. On the other hand, a region could also be managed by different specialized farms. In this case, the aggregated land use levels of all farms allocated to a FMU should be close to the predictions. The different concepts are visualized in Figure 3.1.

Figure 3.1: Concept of allocating farms according to land use level and land use share.



Source: Own compilation.

The a priori information on crop areas in the FMU is given in the form of probability density functions coming from Kempen et al. (2007). We assume a normal distribution characterized by mean $\mu_{LEVELc,fmu}$ and variance $\sigma_{LEVELc,fmu}$ aggregated from the HSMU land use data, $N(\mu_{LEVELc,fmu}, \sigma_{LEVELc,fmu})$.

After taking logs and summing over all crops and FMUs, the objective function based on the highest posterior density concept is consequently

$$obje_{LEVEL} = -\sum_{c}\sum_{fmu} \log N(\mu_{LEVELc,fmu}, \sigma_{LEVELc,fmu}),$$

where $LEVEL_{c,fmu}$ are land use levels aggregated over all farms allocated to the specific FMU, i.e.

$$LEVEL_{c,fmu} = \sum_{\mathrm{f}} p_{f,fmu}\, weight_f\, LEVEL_{\mathrm{c,f}},$$

with the land use levels $LEVEL_{c,f}$ of each FADN farm.

Similarity of crop shares is measured analogously to yield as described above.

$$pdf_{SHAR\ f,fmu} = \prod_{c} N(\mu_{SHAR\ c,fmu}, \sigma_{SHAR\ c,fmu}),$$

where $\mu_{SHAR\ c,fmu}$ is simply the cropping area of each crop divided by the total area of a FMU. The variance is set according to a coefficient of variation of 10%. Covariance is ignored. After scaling we get $p_{SHAR\ f,fmu}$.

The objective function is accordingly:

$$obje_{SHAR} = \sum_{f}\sum_{fmu} p_{f,fmu}\, p_{SHAR\ f,fmu}.$$

The complete optimization problem can finally be written as:

$$\max\ weight_{YIELD}\, obje_{YIELD} + weight_{SHAR}\, obje_{SHAR} + weight_{LEVEL}\, obje_{LEVEL}$$

$$\begin{aligned}
s.t.\quad &(1)\ obje_{YIELD} = \sum_{f}\sum_{fmu} p_{f,fmu}\, p_{YIELD\ f,fmu} \\
&(2)\ obje_{SHAR} = \sum_{f}\sum_{fmu} p_{f,fmu}\, p_{SHAR\ f,fmu} \\
&(3)\ obje_{LEVEL} = -\sum_{c}\sum_{fmu} \log N(\mu_{LEVELc,fmu}, \sigma_{LEVELc,fmu}) \\
&(4)\ LEVEL_{c,fmu} = \sum_{\mathrm{f}} p_{f,fmu}\, weight_{f}\, LEVEL_{\mathrm{c,f}} \\
&(5)\ UAA_{fmu} = adjustfactor \sum_{\mathrm{f}} p_{f,fmu}\, weight_{f}\, UAA_{f} \\
&(6)\ \sum_{fmu} p_{f,fmu} = 1
\end{aligned}$$

where $weight_{YIELD}, weight_{SHAR}, weight_{LEVEL}$ must be set a-priori. In a validation process, various settings will be tested and compared to find out which setting might produce the best overall results. Setting a weight to 0 means that the corresponding information is not used.

3.7 Validation

We follow different methods to validate the results of the allocation procedure and to determine preferable weights in the objective function.

The FADN data contain a sub region code that allows checking whether the allocation results are in line with the records. For several FADN regions, the sub region codes allow to identify the NUTS 2 region where the farm is actually located. For those FADN regions consisting of more than one NUTS 2 region, we calculate the percentage of farms with matching allocation result and sub region information.

The Farm structure survey (FSS) gives information on the area covered by different farm types according to the EU Digit 1 level (see Table 3.3) at detailed regional level. The total area covered by farm types differs systematically between FSS and FADN since FSS covers all farms and FADN only those above a certain size. Calculating shares of farm types makes numbers comparable and assumes implicitly that there is no systematic difference in farm type distribution depending on the farm size. The differences found per farm type are aggregated to the share of misclassified UAA for each administrative region.

The Dutch Geographical Information System for Agricultural Businesses (GIAB) provides numbers of holdings at a level of about 3200 postal codes belonging to 462 communes in 12 Provinces. Farms are classified according to the EU Digit 2 code. To obtain a manageable number of farm types the farm types were aggregated to some extent (see Table 3.3). Because of the large number of Postal Code regions and communes and some inconsistencies between GIAB and FADN, we calculate correlations of various indicators based on data and model results.

Table 3.3: Definition of Farm types based on EU classification

EU Classification			classification used
1-Digit Code	2-Digit Code	Label	(short name)
1	13	Specialist cereals, oilseed and protein crops	arable
	14	General field cropping	
6	60	Mixed cropping	
2	20	Specialist horticulture	horticulture
3	31	Specialist vineyards	permant crops
	32	Specialist fruit and citrus fruit	
	33	Specialist olives	
	34	Various permanent crops combined	
4	41	Specialist dairying	dairy
	42	Specialist cattle-rearing and fattening	beef
	43	Cattle-dairying, rearing and fattening combined	
	44	Sheep, goats and other grazing livestock	sheep and goat
5	50	Specialist granivores	granivores
7	71	Mixed livestock, mainly grazing livestock	mixed livestock
	72	Mixed livestock, mainly granivores	
8	81	Field crops-grazing livestock combined	mixed farms
	82	Various crops and livestock combined	

Source: Own compilation based on FSS and FADN.

3.8 Processing of allocation results for further use

The result of the allocation is that spatial information is added to each individual farm contained in the FADN data base. This locational dimension comprises a reference to a FMU in which the farm is most likely to be located. The individual FADN farm can then be aggregated to any cluster of farms per any cluster of FMUs. This aggregated information may be presented provided that FADN disclosure rules, prescribing a minimal representation of at least 15 FADN sample farms, are not violated. However, information on the share of the agricultural land managed by the different farm types can always be presented as this is not linked to the FADN variables as such, but is merely a calculated probability.

In order to present allocation results at European scale individual farm data from FADN have been aggregated to farm types (see Table 3.3, Andersen et al., 2006) and allocation results have been processed to agri-environmental zones with approximately homogenous conditions for farming (see Hazeu et al., 2006).

4 Results and Discussion

As described in section 3.2, assumptions have to be made on the weighting factors in the objective function. Hence we start with comparing different model specifications in order to find a suitable combination of settings. Following, we will evaluate the overall quality of allocation results with detailed data for the Netherlands. Finally, results on farming systems and farm structure will be demonstrated at European level.

4.1 Model comparison

In the following paragraphs model specifications are compared differing in applied weighting factors. Table 3.4 gives an overview on the different model specifications and their performance according to two validation criteria. The validation was done for about 60 NUTS 2 regions where sub region codes from FADN and out-of-sample data from FSS on farm types are available. These NUTS 2 regions belong to 15 more aggregate FADN regions covering Austria, Ireland, the Slovak Republic and Sweden as well as parts of Germany, Greece and UK. Percentage values refer to the UAA of the region to which farms were allocated correctly. The numbers presented are the arithmetic means of the NUTS 2 regions.

Table 3.4: Overview on model specifications and validation results

model name	weighting scheme			correct allocation (% of UAA)	
	yield	*shar*	*level*	Nuts2 farm type	Nuts2 sub codes
YIELD	1			75%	32%
SHAR		1		80%	35%
LEVEL			1	68%	29%
YIELD_SHAR	0.5	0.5		79%	35%
YIELD_LEVL	0.5		0.5	73%	34%
LEVEL_SHAR		0.5	0.5	79%	35%
YIELD_LEVEL_SHAR	0.33	0.33	0.33	78%	36%

Source: Own compilation.

As the absolute percentage of correctly allocated farms depends on the complexity of the FADN regions, we focus on the ranking of models. Although the absolute value of correctly allocated farms differs significantly between the two validation criteria, the ranking is quite similar.

Since differences in average values are small, we also calculated additional attributes to compare the models in more detail. We considered it interesting to know whether a model often performs best or worst compared to the other specifications. Since there might be "good" models that are not the best, we also report whether a model performs better than the average of the models tested. Results for the different validation criteria are shown in Table 3.5 and 3.6. The results confirm the tendencies already visible in Table 3.4 but differences are more pronounced. The model "LEVEL" is selected as the worst model in about 50, respectively 60 percent of the regions. As one might expect, models combing various objective functions seems to be better than those using only one source of information. The model using all information performs best most frequently, is better than average in almost 80% of the tested regions but also does poorly in a few regions.

Table 3.5: Validation against NUTS 2 sub code information

model name	Percentage of regions where model performs			relative difference to best model (avrerage over regions)
	best	worst	better than average	
YIELD	6%	19%	39%	-23%
SHAR	15%	11%	61%	-15%
LEVEL	16%	50%	31%	-29%
YIELD_SHAR	3%	2%	60%	-14%
YIELD_LEVL	21%	13%	45%	-19%
LEVEL_SHAR	11%	3%	60%	-14%
YIELD_LEVEL_SHAR	29%	5%	77%	-12%

Source: Own compilation.

As no model clearly dominates the others, we present additionally the average the relative difference compared to the best model in that region. Again, the model combining all information shows the smallest loss of accuracy on average across both types of validation information. However, the computational less demanding models “YIELD_SHAR” and “SHAR” also produce good results in this category and overall.

Table 3.6: Validation against NUTS 2 farm type information

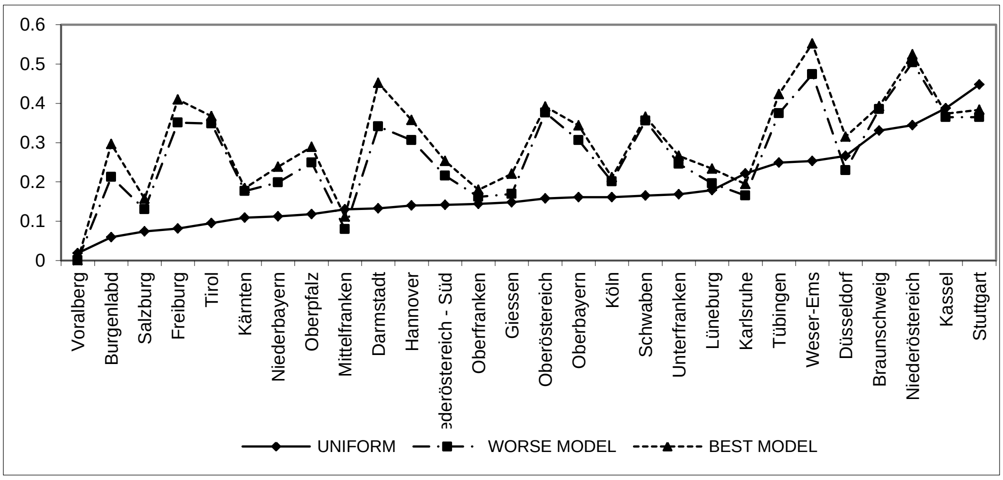
model name	Percentage of regions where model performs			relative difference to best model (avrerage over regions)
	best	worst	better than average	
YIELD	13%	13%	63%	-62%
SHAR	25%	0%	75%	-19%
LEVEL	13%	63%	13%	-112%
YIELD_SHAR	13%	13%	75%	-23%
YIELD_LEVL	0%	13%	13%	-72%
LEVEL_SHAR	13%	0%	88%	-21%
YIELD_LEVEL_SHAR	25%	0%	75%	-21%

Source: Own compilation.

In those FADN regions where out-of-sample validation is possible, we use the model which performs best according to the validation. In all other regions we apply the “YIELD_LEVEL_SHAR” model as it produces favourable results on average over all validated regions.

Using the sub code information the percentages of correctly allocated farms are around 35% – 40% on average (see Table 3.4). This is fairly low at first glance. It should be considered, however, what the alternative to this proposed allocation procedure would be. When it is not possible to collect or not allowed to use information on the location of a specific farm, the default assumption would be that farms are distributed equally in space. The performance of this uniform allocation depends on the number and relative size of sub regions. Figure 3.2 shows the results compared to the best and worst weighted model for selected German and Austrian regions. The share of correctly allocated farms with the weighted model strongly outperforms the uniform distribution.

Figure 3.2: Model results compared to uniform allocation for selected German and Austrian regions.



Source: Own compilation.

Even though the models outperform a uniform distribution, the allocation to sub regions is clearly not as successful as the allocation of farm types. This indicates that the allocation procedure mixes up the NUTS location of similar farms fairly often. Expressed differently, even if the farms are not located in the correct NUTS region, they are assigned to a fairly suitable environment.

4.2 Detailed validation in Netherlands

The Netherlands is a single FADN region. Since farms are uniform regarding LFA and altitude, this information cannot be effectively used in the allocation. Data used for validation of results in the Netherlands come from the GIAB data base providing the number of holdings differentiated by farm types for 3200 postal codes which can be aggregated to about 462 communes. However, it was not possible to differentiate the economic size of the farms, making comparisons of FADN and GIAB data difficult. Since FADN does not consider small, non commercial farms, the number of holdings reported in FADN is generally lower compared to GIAB (see Table 3.7). While for some farm types such as dairy and mixed livestock, numbers are quite similar, they differ significantly for others. Only about one third of the sheep and goat farms in the population are above the FADN cut off criteria. Consequently, a comparison of shares of farms types between the allocation results and the GIAB needs to be corrected by a farm type specific scaling factor adjusting for the number of farms considered in both data bases. It should also be noted that the shares of each farm type with respect to number of holdings differs significantly from the area share. For example, arable and dairy farms together make up about 40% of the holdings, but manage almost 75% of the UAA.

Table 3.7: Comparison of FADN and GIAB data in the Netherlands

	Item	Unit	arable	dairy	beef	sheep and goat	mixed farming	mixed livestock	permanent	horti-culture	granivores
			Farm Types								
GIAB	Holdings	Total number	14787	20390	3307	19304	5091	2012	4181	9511	5878
		Share	18%	24%	4%	23%	6%	2%	5%	11%	7%
FADN	Holdings	Total number	8284	19510	1725	6768	2395	2056	3890	8359	3539
		Share	15%	35%	3%	12%	4%	4%	7%	15%	6%
	UAA	1000 ha	389	733	19	158	75	42	38	67	24
		Share	25%	47%	1%	10%	5%	3%	2%	4%	2%
	UAA (farm)	ha	46,9	37,6	10,8	23,4	31,3	20,5	9,8	8,0	6,9
Scaling factor: Holdings(GIAB)/ Holdings(FADN)			1,79	1,05	1,92	2,85	2,13	0,98	1,07	1,14	1,66

Source: Own compilation.

Correlations of GIAB data and allocation results are presented in Table 3.8. The correlation differs significantly between farm types and model specifications. Correlations at commune level are generally higher but do not differ systematically from those at the very detailed postal code level. Comparing farm types, the correlation for dairy and arable systems is generally very high. The farm types beef, sheep and goat, mixed farming, mixed livestock and granivores have lower correlations. The allocation results for permanent and horticultural systems perform very poorly in this comparison. The models YIELD_LEVEL and YIELD_LEVL_SHAR perform best across all farm types. The latter nevertheless performs very heterogeneous between farm types. While dairy and arable systems seem to benefit from including crop share information, other farm types perform worse. We speculate that either the farm type as such or the general lower agricultural area managed by these farm types could explain this observation.

Table 3.8: Correlation at commune and postal code level for different model specifications

model name	correlation at commune level per farm type								
	arable	dairy	beef	sheep and goat	mixed farming	mixed livestock	permanent	horti-culture	granivores
YIELD	0,39	0,66	0,28	0,32	0,30	0,14	-0,05	-0,05	0,09
SHAR	0,67	0,80	0,25	0,20	0,22	0,11	0,03	-0,02	-0,01
LEVEL	0,86	0,80	0,21	0,30	0,36	0,08	0,10	-0,03	0,08
YEILD_SHAR	0,63	0,78	0,11	0,09	0,03	0,03	-0,08	-0,04	-0,13
YIELD_LEVEL	0,72	0,80	0,07	0,19	0,41	0,31	0,10	0,02	0,23
LEVEL_SHAR	0,59	0,78	0,07	0,03	0,08	-0,05	-0,10	-0,04	-0,13
YIELD_LEVEL_SHAR	0,76	0,77	0,11	0,37	0,21	0,12	-0,08	-0,04	0,10
	correlation at postal code level per farm type								
	arable	dairy	beef	sheep and goat	mixed farming	mixed livestock	permanent	horti-culture	granivores
YIELD	0,34	0,50	0,04	0,35	0,26	0,11	-0,04	-0,04	0,06
SHAR	0,49	0,72	0,02	0,05	0,05	0,12	0,09	-0,04	-0,06
LEVEL	0,78	0,68	0,03	0,49	0,35	0,14	0,08	-0,01	0,10
YEILD_SHAR	0,48	0,74	0,10	0,00	0,07	0,19	0,08	-0,03	-0,06
YIELD_LEVEL	0,66	0,71	0,05	0,17	0,23	0,28	0,07	0,00	0,19
LEVEL_SHAR	0,47	0,74	0,08	0,01	0,04	0,14	-0,05	-0,03	-0,05
YIELD_LEVEL_SHAR	0,76	0,77	0,11	0,37	0,21	0,12	-0,08	-0,04	0,10

Source: Own compilation.

Comparing Table 3.7 and 3.8 we find that accuracy of the allocation of farm types seems to increase with the land managed by an average farm. This is plausible since our allocation procedure makes use of land based characteristics. This is encouraging as it implies that the share of the area that is assigned to a farm type in the allocation procedure is likely to be more in line with what happens on the ground than the number of farms assigned to a certain location.

We also checked to what extent the dominant farm type is assigned correctly to a location. Overall, about 65% of the communes are assigned the correct dominant farm types. In order to better understand the conditions of good and bad performance for this classification, we clustered communes by different characteristics (Table 3.9). For communes with a very mixed farm type presence (e.g. the dominant farm type share below 40%), correct predictions are expected to be more difficult. The validation with the Dutch GIAB data indeed confirms this. The break point seems to be around a 40% dominant farm type share. When the share of the dominant farm type increases, the percentage of correct classifications also increases up to 75%. Due to a strong negative correlation between the number of holdings and the share of the dominant farm types, the performance of the allocation decreases when the share of the dominant farm goes above 80%. The number of holdings per commune range from 1 to 1300. Hence we clustered the communes with respect to number of holdings and found that the performance of our allocation is getting better with increasing agricultural importance of the commune. When we cluster communes with respect to dominant farm types, we observe the same pattern as found in the correlation. Either arable or dairy systems are dominant in 282 communes and 221 are predicted correctly. Sheep and goat systems are the dominant type in 24 communes, but only 8 are classified properly. The low representativity of the

FADN sample with respect to these types of holdings is likely to be blamed for this poor performance.

Table 3.9: Correctly classified dominant farm types for different characteristics of communes

Communes, where …	GIAB data	correct prediction YIELD_LEVEL	
dominant farm type <30%	37	11	30%
dominant farm type 30-40%	93	39	42%
dominant farm type 40-50%	91	54	59%
dominant farm type 50-60%	100	76	76%
dominant farm type 60-70%	61	43	70%
dominant farm type 70-80%	55	38	69%
dominant farm type 80-90%	16	10	63%
dominant farm type >90%	9	5	56%
<50 holdings	100	54	54%
50-150 holdings	161	91	57%
150-300 holdings	118	72	61%
>300 holdings	82	61	74%
arable farms dominant and >40%	61	36	59%
dairy farms dominant and >40%	221	185	84%
beef farms dominant and >40%	1	0	0%
sheep and goat farms dominant and >40%	24	8	33%
mixed farmsdominant and >40%	6	0	0%

Source: Own compilation.

4.3 Allocation Results

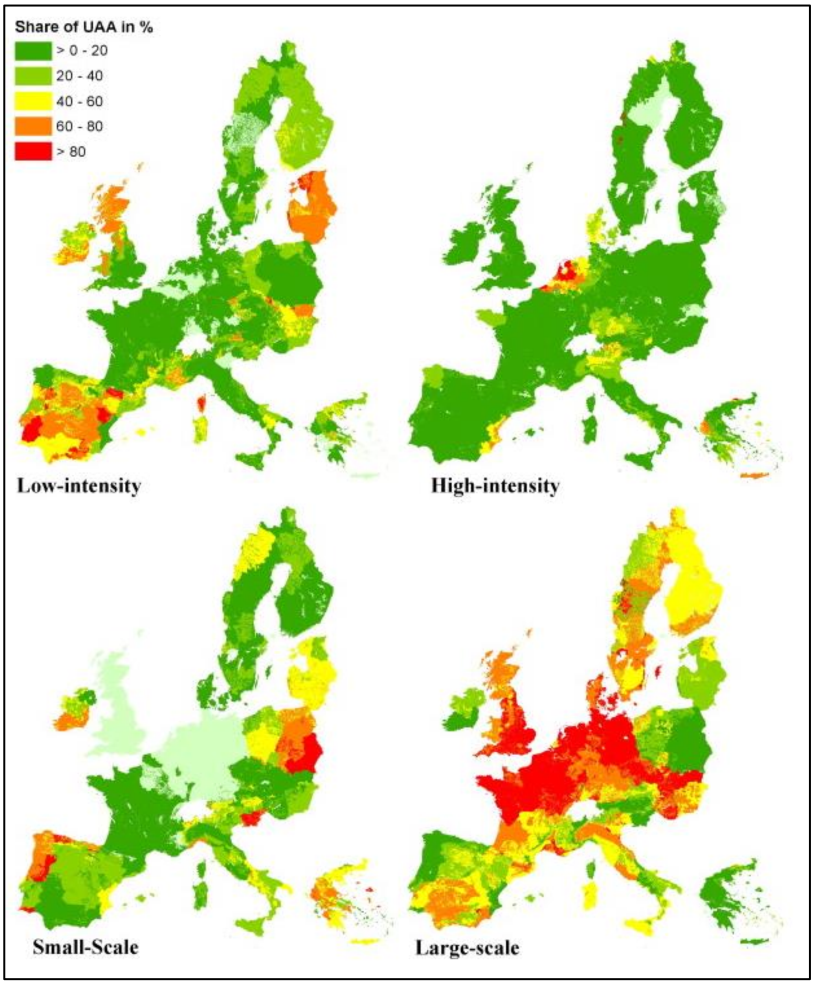
In the following subsections selected results of the allocation of the farm types are presented. The results are presented at EU-level for the allocation of farm types according to intensity, size and specialisation/land use. These three dimensions are the ones that are used to define the farm types in the SEAMLESS farm typology (see Andersen et al., 2006 and 2007). For the purpose of presenting the results in this paper, we have chosen the so-called agri-environmental zones defined in the SEAMLESS project. The agri-environmental zones are defined by relatively homogenous conditions for farming in terms of climate and soil characteristics and are furthermore within only one administrative region (see Hazeu et al., 2006).

4.4 Allocation of farm types according to farm size

The standard gross margin (SGM) can be used to determine the economic size of farms. In the agricultural statistics the SGM is calculated by the national statistical bureaus based on regional standard values for each crop and livestock item. This again is summarized per farm and expressed in terms of European Size Units (ESU), where 1 ESU corresponds to 1,200 Euro. In Figure 3.3 the results of the allocation of farm types to agri-environmental zones are shown in relation to the size dimension of the allocated farm types (two maps in the lower part of the Figure). The results are shown for the small (<

16 ESU) and large (> 40 ESU) scale farm types, whereas the results for the medium sized farms are not shown.

Figure 3.3: The distribution of low-intensity, high intensity, small-scale and large scale farm types on agri-environmental zones in the European Union



Note: Bulgaria, Cyprus, Malta and Romania are not included. The distribution of the medium scale and medium intensity farm types is not included in the illustration. The lightest green indicates regions where the farm type in question is not present.
Source: Own compilation based on data from EU/FADN, G3.

As can be seen in the Figure large scale farms dominate in the North-western part of the European Union, except in Ireland where small scale farming is dominated. The Southern and Mediterranean part of the Union shows is greater diversity in farm types according to size. In most Member States in this part of the Union both regions dominated by small scale and large scale farms can be found. The exception is Greece where only small scale farms dominate. The results for the new Member States in Eastern Europe also show a diverse picture, where both small and large scale farm types can be found as dominating. The results show some differentiation according to the agri-environmental zones within the administrative regions. This is for example the case for Denmark, where the large scale farms have a higher occurrence in the Eastern part of the country, where the soils are relatively better than in the Western part. Another example is Reggio Emilia, where small scale farms dominate in the mountains to the south and large scale farms dominate in the Po Valley to the North.

4.5 Allocation of farm types according to farm intensity

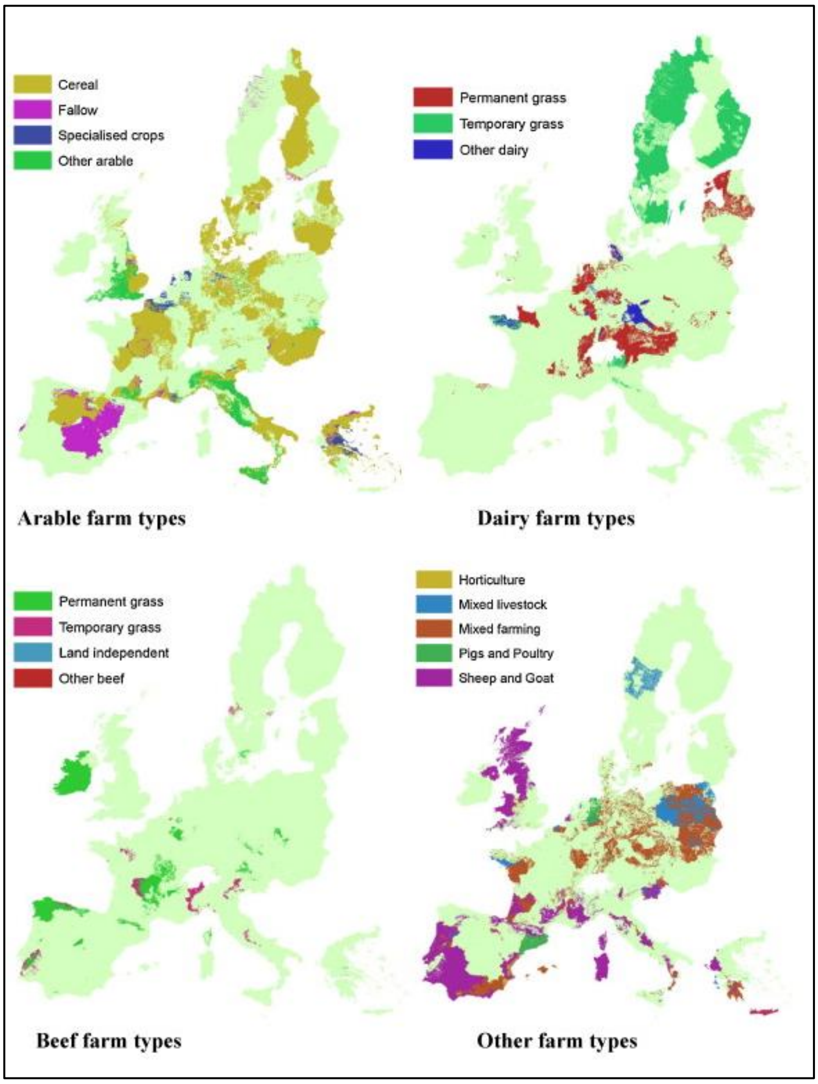
The results of the allocation of farm types according to the intensity of farming are shown in Figure 3.3 (two maps at the upper part of the Figure). Intensity is defined as total output (€) per ha. The results are shown for the low intensity (< 500€) and high intensity (> 3000€) farm types, whereas the results for the medium intensity farms are not shown.

As can be seen in the Figure high intensity farming is dominating in very few regions. Some clusters can be found in the Netherlands and the bordering regions of Germany and Belgium and in the Eastern part of Spain. Low intensity farming dominates in a more scattered pattern across the European Union. Three larger clusters are found on the Iberian Peninsula, in Scotland, Northern England, Wales, Northern Ireland and Ireland and in the Baltic States. An example of the differentiation in the allocation results within a region can be seen in Scotland, where the low-intensity farming to a higher degree dominates in the Highlands.

4.6 Allocation of farm types according to farm specialisation

The results of the allocation of farm types according to the specialisation of farming are shown in Figure 3.4. The maps shows in which regions a farm types is the most dominating in terms of agricultural area managed. For three of the main specialisations (Arable, Dairy and Beef) the dominating land use type is presented, whereas the remaining main types of specialisations are shown in the same map without information on dominating land use type. The largest part of the agricultural area of the European Union is dominated by arable farms, in most cases based on cereal production. Arable systems characterized by a high degree of fallow land dominates in parts of Spain, farm types with a high degree of specialised crops (potato, cotton, sugar beet etc.) dominate in parts Belgium, the Netherlands, Germany and Greece and mixed arable systems dominates especially in parts of England and Italy. Dairy farm types dominate in Central and Northern parts of the European Union. Farm types based on temporary grassland dominates in Sweden and Finland and in smaller parts of Brittany and Northern Italy, whereas farm types based on permanent grassland dominates in parts of Estonia, Latvia, Poland, Austria, Germany, the Netherlands and France. Beef farm types dominate fewer areas of the European Union and are mainly based in permanent grassland. Important areas are Ireland, Central France and North-Western Spain. Finally, of the remaining farm types sheep and goat and mixed farming are the most important. Sheep and goat dominates in several areas scattered across the European Union with clusters in the Northern and western parts of United Kingdom and in Portugal and the Southern part of Spain. For the mixed farm types Poland forms an important cluster with a mix of mixed and mixed livestock farms and also larger areas in France, Spain and Greece are dominated by mixed farming.

Figure 3.4: Distribution of farm types



Note: Distribution of arable farm types in agri-environmental zones dominated by arable farm types, Dairy farm types in agri-environmental zones dominated by dairy farm types, beef farm types in agri-environmental zones dominated by beef farm types and other farm types in agri-environmental zones dominated by other farm types.
Source: Own compilation based on data from EU/FADN, G3.

A special pattern occurs in Germany where mixed farming is scattered in smaller areas across the country. The remaining farm types only dominate in a few areas: Pigs and Poultry in Catalonia and parts of Lower Saxony and horticulture in areas along the Mediterranean coast of Spain. Examples of a differentiated distribution of farm types within regions can be seen in Estonia with Arable/cereal farm types dominating in the Eastern part of the country and dairy farming dominating in the Western part. Another example is Brittany with dairy farm types dominating to the North and mixed livestock farm types dominating to the South.

5 Conclusion

A methodology allowing a spatial allocation of farms in the FADN data base was developed and successfully applied in the European Union. Various models using all or sub-sets of information on yields, land use shares and levels were specified and validated. The model using all available information with equal weight is the best overall, but does not dominate. The suitability of prior information seems to depend on the characteristics of the farm. For example, the prior information on land use shares improves the allocation results for arable and dairy systems, which have strong land dependence and land use share, but negatively influences the correct allocation of other farm types, with lower land dependency and or area share. The prior information used here seems insufficient to allocate some farm types. Hence further information might contribute to improve the allocation results, for example herd sizes at administrative level below FADN regions for land independent systems like granivores.

The FADN database has considerable deficiencies which should be kept in mind when working with the allocated farm data. The FADN sample does not sufficiently represent small and part- time farms. This likely implies that farms in the more marginal farming areas are not well represented. The weights offered by FADN do not guarantee representativity at the level of LFA and altitude zones. An explicit (re-)estimation of representativity weights for individual farms might be a useful extension in further developing the allocation procedure.

After clustering single farms to farm types we could validate our results against out-of-sample data. The validation revealed bad matches mainly with respect to land independent systems. Nonetheless, the percentage of UAA assigned to the correct farming system is quite high, because the procedure performs well for farm types representing significant shares of land in important agricultural regions. Most of the validation criteria revealed that the accuracy of the allocation model is in a range of 60-70%. This seems acceptable given that our model results should mainly serve as input for models working with European coverage. The allocation procedure recovers the spatial farm type distributions quite well, thereby providing information of significant value for further analysis in a multidisciplinary context. The allocation of specific farms to the spatial units performs less well, but is still clearly better than a uniform distribution of farms in space as often implied by aggregate analysis.

6 Acknowledgements

The work was partially funded by the Directorate-General for Research of the EU-Commission in the context of the SEAMLESS Integrated Project, EU 6th Framework Programme, Contract No. 010036-2.

7 References

Andersen, E., Verhoog, A.D., Elbersen, B.S., Godeschalk, F.E., Koole, B., 2006. A multidimensional farming system typology, SEAMLESS Report No.12, SEAMLESS integrated project, EU 6th Framework Programme, contract no. 010036-2, www.SEAMLESS-IP.org, 30 pp, ISBN no. 90-8585-041-X.

Andersen, E., Elbersen, B., Godeschalk, F., Verhoog, D., 2007. Farm management indicators and farm typologies as a basis for assessments in a changing policy environment. Journal of Environmental Management, 82 (3), 353-362.

Anselin, L., Florax, R. J. G. M., Rey, S. J., 2004. Advances in Spatial Econometrics, Springer Verlag, Berlin.

Boogaard, H.L., Eerens, H., Supit, I., Diepen, C.A. van, Piccard, I., Kempeneers, P., 2002. Description of the MARS Crop Yield Forecasting System (MCYFS). METAMP-report 1/3, Alterra and VITO, JRC-contract 19226-2002-02-F1FED ISP NL.

Britz, W., Witzke P., 2008. CAPRI model documentation 2008: Version 2. Institute for Food and Resource Economics, Bonn.

Genovese, G. (ed.), 2004. Methodology of the MARS Crop Yield Forecasting System. Eur Rep 21291 EN/1-4.

Genovese, G., Baruth, B., Royer, A., Burger, A., 2007. Crop and yield monitoring activities – MARS STAT action of the European Commission, Geoinformatics, 10, 20–22.

Hazeu, G.W., Elbersen, B.S., van Diepen, C.A., Baruth, B., Metzger, M.J., 2006. Regional typologies of ecological and bio-physical context, SEAMLESS Report No.14, SEAMLESS integrated project, EU 6th Framework Programme, contract no. 010036-2, www.SEAMLESS-IP.org, 55 pp, ISBN no. 90-8585-042-8.

Heckelei, T., Mittelhammer R., Jansson T., 2008. A Bayesian Alternative to Generalized Cross Entropy Solutions for Underdetermined Econometric Models. Agricultural and Resource Economics, Discussion Paper, Institute for Food and Resource Economics, Bonn.

Howitt, R., Reynaud A., 2003. Spatial Disaggregation of Agricultural Production Data by Maximum Entropy. European Review of Agricultural Economics 30(3):359-387.

Janssen, S., Andersen, E., Athanasiadis I.N., Van Ittersum M.K., 2009. A database for integrated assessment of European agricultural systems. Environmental Science & Policy, 12(5), 573-587.

Kempen, M., Heckelei T., Britz W., Leip A., Koeble R., 2005. A Statistical Approach for Spatial Disaggregation of Crop Production in the EU, in „Modelling agricultural policies: state of the art and new challenges“, MUP, Parma.

Kempen M., Heckelei T., Britz W., Leip A., Koeble R., Marchi G., 2007. Computation of a European Agricultural Land Use Map – Statistical Approach and Validation, Technical Paper, Institute for Food and Resource Economics, Bonn.

Kruska, R.L., Reid, R.S., Thornton, P.K., Henninger, N., Kristjanson, P.M., 2003. Mapping livestock-oriented agricultural production systems for the developing world. Agricultural systems, 77, 39-63.

Leip A., Marchi G., Koeble R., Kempen M., Britz W., Li C., 2008. Linking an economic model for European agriculture with a mechanistic model to estimate nitrogen and carbon losses from arable soils in Europe. Biogeosciences, 5, 73-94.

Louhichi, K., Blanco Fonseca, M., Flichman, G., Janssen, S.J.C., Hengsdijk, H., 2007. 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.

Mulligan, D. T., 2006. Regional modelling of nitrous oxide emissions from fertilised agricultural soils within Europe. PhD thesis, University of Wales, Bangor.

Van der Steeg, J.A., Verburg, P.H., Baltenweck, I., Staal, S.J., 2010. Characterisation of the spatial distribution of farming systems in the Keynyan Highlands. Applied Geography 30(2): 239-253.

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., 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.

Chapter 4: Economic and environmental impacts of milk quota reform in Europe[22]

Abstract

The year 2008 'Health Check' decisions on the Common Agricultural Policy (CAP) included the expiry of the milk quota system. This paper presents an impact analysis using the CAPRI model, which has been updated with econometric estimates of milk quota rents at the level of ca. 230 European regions. Production may increase by 5% while the price drop is about 10%. Regions are identified where economic or environmental changes substantially exceed those at the Member State level. Regional nitrate leaching problems could be exacerbated. But there is only weak evidence of an increased risk of land abandonment.

1 Introduction

The dairy sector makes a substantial contribution to agricultural income in many Member States (MS) of the European Union (EU27). The EU dairy market is regulated by the Common Market Organisation (CMO) for milk and milk products, of which the milk quota regime is one of the most noticeable elements. In the early 1980's it became evident that the price support system ensuring profitable producer prices also triggered production increases far beyond self sufficiency, resulting in high spending for subsidised exports and storage. Facing the choice between a reduction of administrative prices or a regulated limitation of production EU policy makers opted for the milk quota system in 1984. Since its introduction, the milk quota has become a scarce production factor, on the one hand limiting milk production and, as a consequence, allowing profitable milk producer prices and maintaining dairy activities in less competitive regions. Recent policy developments, including reductions of intervention prices and specific quota increases of various amounts to MS, together with most recent market developments, have rendered the quotas non-binding in some MS and regions of the EU. The 'Health Check' decisions of the Common Agricultural Policy (CAP) of 2008 included the expiry of the milk quota system after 2014 and an increase of quotas by 1% annually from 2009 to 2013 to allow for a "soft landing" of the milk sector with expiring quotas.

An assessment of economic effects of an abolition of the milk quota regime is of great interest, since milk is one of the main agricultural commodities produced in the EU. Milk production takes place in all MS and represents for the EU27 a share of ca.13.7% of total agricultural production in 2006 or more than 42.5 billion € at the farm gate (European Commission, 2008a). The shares of milk in total revenues range from 6.7% to 33.5%

[22] This Chapter is based on an article published together with Peter Witzke (EuroCare/University of Bonn), Ignacio Pérez Domínguez (JRC Sevilla), Torbjörn Jansson (LEI, Wageningen), and Paolo Sckokai (Catholic University of Piacenza) in the Journal of Policy Modelling 33 (2011): 29-52.

across MS and tend to be higher in northern Europe and lower in the Mediterranean basin. Within the EU27, the size and relative importance of the dairy sector varies considerably between MS and across regions within MS, reflecting differences in climatic conditions, natural resource endowments, demand characteristics and traditions (e.g. cheese industries). The observed heterogeneity calls for a regionalised analysis.

Recent studies concerning milk market reform policies are often focused on market effects at European or national level. Most approaches were either based on existing computable general equilibrium models (Lips & Rieder, 2002; Lips & Rieder, 2005; Van Tongeren, 2002; Isermeyer, Brockmeier, Gömann, Hargens, Klepper, Kreins, Offermann, Osterburg, Pelikan, Salamon & Thiele, 2006) or partial equilibrium models (Kleinhanss, Manegold, Bertelsmeier, Deeken, Giffhorn, Jägersberg, Offermann, Osterburg & Salamon, 2002; Chantreuil, Donnelan, van Leeuwen, Salamon, Tabeau & Bartova, 2008; Bartova, Fellmann & M'barek, 2008). An INRA-Wageningen Consortium (2002) developed a detailed agro-econometric model (EDIM) focusing on the dairy sector. Regional programming models (Helming & Peerlings, 2005; Kleinhanss, Manegold, Bertelsmeier, Deeken, Giffhorn, Jägersberg, Offermann, Osterburg & Salamon, 2002; Isermeyer, et. al., 2006; Colman, 2002) focused on specific countries or regions.

In this article milk reform policies were simulated with the CAPRI model which links 270 regional supply models covering the entire EU27 with a spatial market model for agricultural commodities (see Britz & Witzke, 2008). The regional supply models derive from the positive mathematical programming tradition maximizing farmer's profit. The model was extended in order to better represent the specificities of a milk quota abolition scenario. Most importantly, the significance of correct quota rents was recognized, and therefore considerable resources were committed to econometric estimations of marginal cost functions and milk quota rents, which were used to calibrate the supply response of dairy producers. Market responses are simulated within a partial equilibrium framework for agricultural commodities. Parameters of the market model were updated, mainly based on results from INRA-Wageningen Consortium (2002) and the update Bouramra-Mechemache, Jongeneel & Requillart (2008).

The paper continues in section 2 with more details on the CAPRI modelling framework and the extensions underlying this study. After defining the main scenario assumptions in section 3, some selected economic and environmental impacts of milk quota abolition in the EU27 are presented in section 4. Section 5 is dedicated to a sensitivity analysis, and section 6 concludes.

2 Modelling Framework

2.1 The CAPRI model

The CAPRI model is an agricultural sector model covering the whole of EU27, Norway and Western Balkans at regional level (270 regions) and global agricultural markets at country or country block level. CAPRI consists of supply and market models which interact iteratively via exchange of prices and quantities. The supply model makes use of non linear mathematical programming to maximise regional agricultural income with explicit consideration of the CAP instruments of support in an open economy, subject to technical constraints for feeding, young animal trade, fertilization, set-aside, land availability and production quotas (see Britz & Witzke, 2008). Major outputs of the supply module are crop areas and animal numbers at regional level, with their associated revenues, costs and incomes.

Roughage is assumed to be non tradable among regions since transport costs are considered prohibitive. Components of feed concentrates, mainly cereals and cakes, are traded among regions and their prices are endogenous in the market model, which indirectly links the individual regional models and also introduces competition for feed among different animal sectors such as beef, pork and poultry. The equations in the supply model account for a balanced, cost minimizing feed composition for each animal category. Availability of land is bounded by constraints. Grassland and arable land are treated as different resources since EU legislations tend to prevent conversion from grassland to arable land[23]. About 30 annual or permanent crop activities compete for the available arable land whereas for permanent grassland there is only a choice of different intensity levels. In addition to technical constraints, supply behaviour is steered by a quadratic cost function in the tradition of positive mathematical programming. The specification of the quadratic terms was extended as compared to Howitt's (1995) original specification. In particular, the behavioural parameters of all crop sectors, including rotational interaction terms, were econometrically estimated by Jansson & Heckelei (2008) whereas the animal husbandry sectors except for dairy were calibrated based on exogenous supply elasticities. The existing milk quota is represented as a (binding) constraint in the regional supply models. The quota rents and elasticities of supply in the dairy sector are introduced as external information in the calibration process. Estimates are derived from a newly developed econometric estimation framework of regional marginal cost curves (Section 2.2). Milk yields of dairy cows are endogenously depending on milk prices by competition of low and high yielding technologies. Structural change effects are not explicitly considered but they are implicit in the marginal cost function estimates.

The market module consists of a system of equations representing supply and demand within a spatial world trade model for agricultural commodities. Major outputs of the market module include bilateral trade flows, market balances and nominal producer and consumer prices for the products for world country aggregates (i.e. the EU15 as one aggregate).

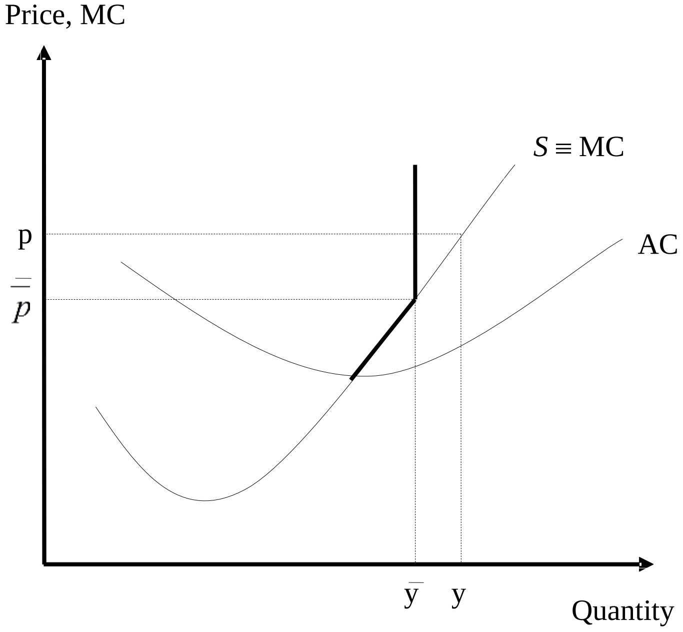
The CAPRI version used for this study is standard comparative-static, i.e. adjustment costs are not considered and policy simulations apply to a situation where dairy farmers were given time to adjust to the new policy framework.

[23] This refers to Council Regulation (EC) No 1782/2003. Member States shall ensure that land which was under permanent pasture at the date provided for the area aid applications for 2003 is maintained under permanent pasture. However some exemptions are allowed in the legislation but not modelled in CAPRI since national implementation could not considered so far.

2.2 Marginal cost and quota rents

In Figure 4.1 milk quota and milk quota rent are represented at the producer level. The average cost (AC) curve is assumed to be U-shaped and the MC curve crosses the AC curve at the minimum. In the absence of a quota, the supply curve *S* coincides with the increasing part of the marginal cost (MC) curve which is above the intersection with the average cost curve. At a given price (marginal revenue) *p*, supply is determined by the condition MC = *p*.

Figure 4.1: Representation of milk quota rents at the producer level.



Given a fixed price p, a marginal cost function MC and a binding quota $\bar{y}$ that is smaller than the unconstrained equilibrium y, the producer earns a quota rent corresponding to the difference between the price and the marginal cost at quota level of production, $p - \bar{p}$, and the supply schedule (heavy line) is kinked and infinitely inelastic at the binding quota.

Source: http://ipts.jrc.ec.europa.eu/publications/pub.cfm?id=2719, p.3

The introduction of a quota creates a departure from standard competitive market pricing, where profit-maximizing agents equate marginal revenue to marginal cost. If a quota is binding[24], production will be smaller than the unrestricted market equilibrium. The

[24] In terms of implementation, milk quotas are imposed through the payment of a fine (the super-levy). When the super-levy is applied at producer level, it means that for excess production the producer receives the market price less the fine. Usually the fine is that large that net return for a kilogram of surplus milk will by far not cover cost. However, a super levy only has to be paid when the national quota is overshot, i.e. a farmer producing beyond his quota might expect that he has to pay no super levy has because other framers do not fill their quota. In fact a farmer considers an expected super levy, which can be significantly lower than the declared fine, in his production decision.

new level of production will be fixed at $\bar{y}$ which represents the binding quota, which is less than the unconstrained equilibrium *y*. The supply curve will be kinked and becomes perfectly inelastic at the quota level (see thick line in Figure 4.1) so that it is no longer possible to directly observe production responses to price changes if quotas are binding. At $\bar{y}$ marginal revenue is greater than marginal cost and marginal cost coincides with the so-called output shadow price. The milk shadow price is the producer price that would induce a profit-maximizing producer to produce the current quota level in the absence of production restrictions. The difference between the market price and the shadow price defines the so-called unit quota rent, which corresponds to $p-\bar{p}$.

In a quota abolition scenario milk production will be determined again according to the usual 'marginal cost equals price' optimality condition. At the EU level it can be expected that milk production will increase while milk prices will decrease. However the regional production effects might be heterogeneous. In regions where the quota rent was low in the reference situation, the reduced milk price might settle below marginal cost in the calibration point, and in those regions production will decrease. In regions where the final milk price remains above marginal cost, production will increase.[25] This leads to a redistribution of production between NUTS 2 regions when national quota trade restrictions are removed. As neither quota rents nor marginal cost can be observed, regional marginal cost functions have to be estimated econometrically.

2.3 Econometric estimation

To estimate marginal costs for milk producers in the EU we have resorted to a cost minimization approach, as in a number of similar studies (Moschini, 1988; Guyomard, Herrard, & Mahè, 1995, Guyomard, Delache, Irz, & Mahè, 1996; Colman, Burton, Rigby & Franks, 1998; Wieck & Heckelei, 2007): if farmers adjust production inputs *x* given quasi-fixed inputs *z*, marketing quotas y^0, output level *y* and input prices *w*, then the following minimization problem [2-1] describes their behaviour:

$$C(y^0, y^1, w, z) \equiv \min_{x} \left\{ wx \;\middle|\; F(y^0, y^1, x, z) = 0 \right\} \qquad [2\text{-}1]$$

where $F(y,x,z)=0$ is a standard production function. Milk marginal cost is then the first derivative of equation [2-1] with respect to the restricted output y^0 (milk) and the unit quota rent r^0 is given by the difference between the milk farm gate price, p^0, and the marginal cost.

[25] In addition to the standard milk quota and milk quota rent description, there are at least four additional cases where farmers do not respond according to the magnitude of the quota rent, but rather according to the difference between milk market price and the average cost at quota level. For more details see Tonini and Pérez Dominguez (2008).

For the empirical specification of the cost function in [2-1], a flexible functional form (FFF) was selected because of its theoretical properties. Following Moschini (1988) the hybrid-translog cost function was chosen. One of the main advantages of the hybrid-translog cost function is that standard U-shaped marginal and average cost curves can be obtained.

The dataset used in the empirical investigation is an unbalanced panel data set of milk producing farms (both specialized and non-specialised), surveyed across the EU countries from 1990 to 2005. The source of the data is the European Farm Accountancy Data Network (FADN)[26]. Each farm in the FADN sample has a weight corresponding to the number of agricultural holdings it represents: thus, each record (sample unit) has a different representativeness within the reference population, and this has to be taken into account in the estimation phase. This database contains considerable information on farm structure and economic activities, as well as information on input costs by category, although variable input prices are not provided by the FADN database; thus, input price indexes have been taken from the official Eurostat statistics[27].

Since the objective of this analysis is to obtain marginal cost estimates to be used for long term scenario analyses, we have estimated the above cost function for a long term period, assuming that farmers may adjust all their factors of production, except unpaid (family) labour, given that, at least in the EU, farming is intimately linked to the family farm. A removal of the quota system requires long-run decisions on all structural characteristics of the farm, such as machinery, buildings, cow stock and also land renting. Family labour has been considered a fixed input even in the long run since the decision on family labour is mainly that of remaining in the sector, thus not strictly related to the quota removal scenario, but also to the opportunity costs of family labour that may depend on the alternatives in the economy (wages and unemployment rates).

Given the structure of the estimated model, non-linear in variables and parameters, the estimation of a fixed-effect model is not straightforward; therefore, we have decided to address the issue of farm heterogeneity by estimating the model for (small) samples of quite homogeneous farms, leaving heterogeneity among and not within samples. Therefore, to account for structural and regional aspects characterizing milk production, the

[26] The FADN database provides data only for the period in which each country has been member of the EU. Thus, for the New Member States, data are available only for the years 2004-05. However, given the large number of observations, the estimations performed reasonably well also on these prevalently cross-section databases.

[27] The EU individual marketing quotas refer to a standard fat content and, in order to compare milk production with the corresponding quota, the former must be corrected for its specific fat content. Unfortunately, the FADN database does not report the fat content of milk produced in each farm. Thus, milk production data could not be corrected. This may imply some bias in marginal cost estimations for those countries in which the deviation of the average milk fat content from the standard one is substantial. In addition, for some countries in some specific years, the FADN database does not provide any data on individual marketing quotas. In these cases, we have assumed that milk production of each farm coincides with the individual quota.

long-run costs have been estimated on well defined farm sub-samples in each EU member country. The FADN records allow distinguishing regional location, altitude above sea and size class of milk operations. The hypothesis underlying this choice is that farms distinguished by these factors display different cost structures.

The plausibility of estimation results were checked with the help of information from quota markets. Since the regional peculiarities of these markets may play a decisive role, the information was used to derive conservative threshold values for detecting and correcting outlier estimations.

2.4 Linking the marginal cost estimation to CAPRI

The PMP concept underlying the CAPRI supply models allows accounting for external information when calibrating the non-linear cost functions for each production activity (Heckelei, 2002). For the case of dairy cows we can choose parameters so that (1) quota rent, i.e. the shadow value of the quota constraint in the CAPRI model, is equal to the quota rent calculated in the MC estimation and (2) the slope in the calibration point reflects the elasticity derived from the MC (or supply) curve. Regional impacts of the milk quota abolition therefore depend upon different regional cost functions, reflecting the regional heterogeneity.

In our database we have regions with small positive quota rents even though the regional quota is not completely filled. This may be explained firstly with the uncertainty at the farm level. A farmer maximizing his expected income has to consider the uncertainty from natural conditions (diseases of cows, fodder quality, weather) and the likelihood of the super levy applied (in case of a national quota overshot). A single farmer will plan to produce less than his quota if the expected revenue from higher production would fall short of his marginal cost (see also Adenäuer, 2006). Nonetheless his marginal cost would be less than the milk price and his decisions would be determined by his quota endowment, although the quota would not be filled in most years. The second explanation for positive regional quota rents in spite of incomplete quota fill are positive transaction costs of quota trade that prevent an immediate flow of ‘unused’ quota endowments to farmers with higher willingness to pay. While both arguments are not explicitly modelled in CAPRI, they motivate our implementation of observed production as a (technical) quota with a small rent even in those cases where regional quotas are not fully used.

Two more problems need to be overcome before the results from the econometric estimation can be used for model calibration purposes since there is an imperfect match both in the regional and in the temporal dimension. Firstly, the estimation of marginal production cost is done at the regional resolution of FADN regions which are often more aggregated than the NUTS 2 regions underlying the CAPRI supply model. However, for many FADN regions the estimates are differentiated by altitude and size. The shares of altitude and size classes for each NUTS 2 region were known from the SEAMLESS project (Elbersen, Kempen, van Diepen, Andersen, Hazeu & Verhoog, 2006). These shares were used to calculate average MC specific for each NUTS 2 region even though the underlying econometric estimation may have been on the national level only with a distinction of size and altitude classes.

Another issue is the projection of quota rents from the ex post situation where the econometric estimation was carried out into the future where reform scenarios are simulated. The quota rent depends on the development of prices and quota endowments[28] but also on structural and technological changes in the dairy sector. To capture these drivers we built on the recent analysis at the MS level by Réquillart, Bouamra-Mechemache, Jongeneel & Penel (2008). More precisely we derived shift factors for the percentage quota rents in each MS from the European Dairy Industry Model (EDIM) model results and applied these shifts to all corresponding regions in a given MS. Thus the econometrically based pattern of quota rents within each MS is maintained in the baseline but relative competitiveness among countries may change. Results in 0 4.1 show, for example, that the new MS (NMS) are assumed to become more competitive in the forthcoming years. The increase of milk prices over time is driven by a strong growth of demand for protein rich dairy products on a global scale (Réquillart, Bouamra-Mechemache, Jongeneel & Penel, 2008).

Table 4.1: Comparison of Quota rents (Qrent) in Base year (2004) and Baseline (2020)

	Base year (2004)			Baseline (2020)			Baseline/Base year CAPRI		
	Milk Price	Qrent	Qrent	Milk Price	Qrent	Qrent	Milk Price	Qrent	Qrent
Quota rents	[€/t]	[%]	[€/t]	[€/t]	[%]	[€/ton]	[% diff]	[% diff]	[€/t diff]
Austria	252.0	30.1	75.9	281.7	29.9	84.2	11.8	-0.2	8.4
Belgium-Lux.	256.0	28.1	71.9	285.4	27.6	78.7	11.5	-0.5	6.8
Denmark	308.4	12.5	38.7	332.8	9.3	31.1	7.9	-3.2	-7.6
Finland	341.7	2.2	7.4	379.1	3.5	13.2	11.0	1.3	5.8
France	274.3	17.1	46.8	300.0	12.6	37.7	9.4	-4.5	-9.1
Germany	281.4	16.5	46.3	313.0	17.8	55.7	11.2	1.3	9.4
Greece	324.9	35.1	114.0	357.8	11.7	41.9	10.2	-23.4	-72.1
Ireland	260.3	24.6	64.0	284.2	20.9	59.4	9.2	-3.7	-4.6
Italy	339.6	20.0	67.8	369.5	14.0	51.5	8.8	-6.0	-16.2
Netherlands	318.5	32.9	104.9	353.8	27.8	98.2	11.1	-5.2	-6.7
Portugal	297.6	17.5	52.1	334.9	10.4	34.7	12.5	-7.1	-17.3
Spain	275.6	28.0	77.0	305.9	22.3	68.1	11.0	-5.7	-8.9
Sweden	312.0	4.2	13.2	341.4	3.0	10.3	9.4	-1.2	-2.9
United Kingdom	254.0	3.9	10.0	277.6	3.2	8.8	9.3	-0.8	-1.2
EU15	286.2	18.1	51.8	315.2	15.4	48.6	10.1	-2.7	-3.3
Cyprus	387.1	1.0	3.9	460.7	5.9	27.0	19.0	4.9	23.1
Czech Republic	242.7	1.2	2.9	282.4	9.6	27.1	16.4	8.4	24.2
Estonia	201.1	1.3	2.6	245.0	6.2	15.1	21.8	4.9	12.6
Hungary	254.3	1.4	3.5	269.8	12.8	34.4	6.1	11.4	31.0
Latvia	157.4	2.1	3.3	196.3	6.9	13.6	24.7	4.8	10.3
Lithuania	151.9	5.0	7.7	182.7	9.9	18.1	20.2	4.9	10.5
Malta	334.5	1.0	3.4	365.2	5.8	21.3	9.2	4.8	17.9
Poland	175.0	2.7	4.7	212.8	14.6	31.1	21.6	11.9	26.4
Slovac Republic	242.6	1.2	2.8	276.9	6.0	16.7	14.2	4.9	13.9
Slovenia	235.3	3.3	7.7	258.8	8.1	21.1	10.0	4.8	13.3
10 New MS	195.3	2.3	4.5	231.1	11.9	27.4	18.3	9.6	23.0
Bulgaria	193.7	0.0	0.0	233.7	8.9	20.8	20.6		
Romania	187.4	0.0	0.0	173.2	13.8	23.8	-7.6		
Bulgaria/Romania	188.7	0.0	0.0	186.1	12.5	23.2	-1.4		
EU27	269.1	15.9	42.9	298.3	15.0	44.6	10.9	-1.0	1.7

Source: Own calculation.

28 Quotas are expanded over time according to the Health Check decisions in all MS. Some specific countries, e.g. Greece, receive additional quota once in time according to EC decisions

3 Scenario Design

Impacts of a milk quota reform are simulated for the year 2020 based on a comparative static model design assuming that farmers fully adapted their behaviour to their desired optimal production programme after the expiry of the quota system in 2015. This requires two scenarios, the continuation of the "status quo" – often called baseline – and a countervailing scenario simulating the milk quota reform policy.

3.1 Baseline scenario

The baseline policy contains significant changes of the ex-post policy: dairy quotas are expanded and market protection is lowered for many products. The sugar CMO is reformed and brought in line with the dairy CMO. Direct payments are to a large extent replaced by decoupled payments, but Member States are given the opportunity of maintaining a certain level of coupled support for selected sectors. There is no multilateral trade agreement but the expansion of the single market to Bulgaria and Romania, together with the full implementation of the "Everything But Arms" (EBA)[29] initiative and the replacement of the ACP sugar protocol by the European Partnership Agreement (EPA)[30], significantly reduces overall market protection. The baseline may be interpreted as a projection covering the most probable future development of the European agricultural sector under the *status quo* policy and including all future changes already decided when preparing this study.

Creating a baseline scenario requires projection of future developments in European agriculture, e.g. production levels, prices, input and output coefficients. Some expert data on future trends are based on projections by the European Commission (2006, 2007a, 2007b and 2008b) and international agencies like FAPRI (2007), in particular for non-EU regions. However for many dairy related variables like milk prices we used Réquillart, Bouamra-Mechemache, Jongeneel & Penel (2008). This expert information is merged with trend projections of regional time series in the process of the CAPRI baseline construction (Britz & Witzke, 2008, section 3). This information and trend projections of regional time series for yields and production are jointly used in the baseline construction (Britz & Witzke, 2008)[31].

[29] EBA provides duty-free and quota-free access for products from the 49 Least Developed Countries.

[30] EPA negotiations offer additional market access opportunities for the African, Caribbean and Pacific countries

[31] Expert information and trend projections are fed into a Bayesian estimator, which selects the most likely combination of forecast values subject to a larger set (e.g. closed area and market balances, feed requirements, production quotas, etc.). This procedure ensures the compatible use of a large amount of projected variables in a stylised form.

Table 4.2: Exogenous drivers considered for the baseline construction

Exogenous drivers	Value
Inflation	1.9 % per annum
Growth of GDP per capita	2.0 % nominal per annum for the EU10, 5 % for India, 1.5 % for USA, 4 % for Russia, 1.5 % for Least Developed countries and ACPs, and 1 % for the rest.
Demographic changes	EUROSTAT projections for Europe and UN projections for the rest of countries in the world
Technical progress	0.5% input savings per annum (affecting exogenous yield trends), with the exemption of N, P, K needs for crops where technical progress is trend forecasted
Domestic Policy	National decisions on coupling options and premium models, with their expected implementation date for the EU25 MSs (25 different premium schemes, compilation by Massot Martí, 2005)
Common Market Organisations	Supply and demand shifted according to the expert forecasts (Commission of the European Communities, 2005)
Trade policy	Final implementation of the 1994 Uruguay round plus some further elements as NAFTA.
World markets	Supply and demand forecasts (FAO, 2003).

Source: Own compilation.

3.2 Milk quota reform scenario

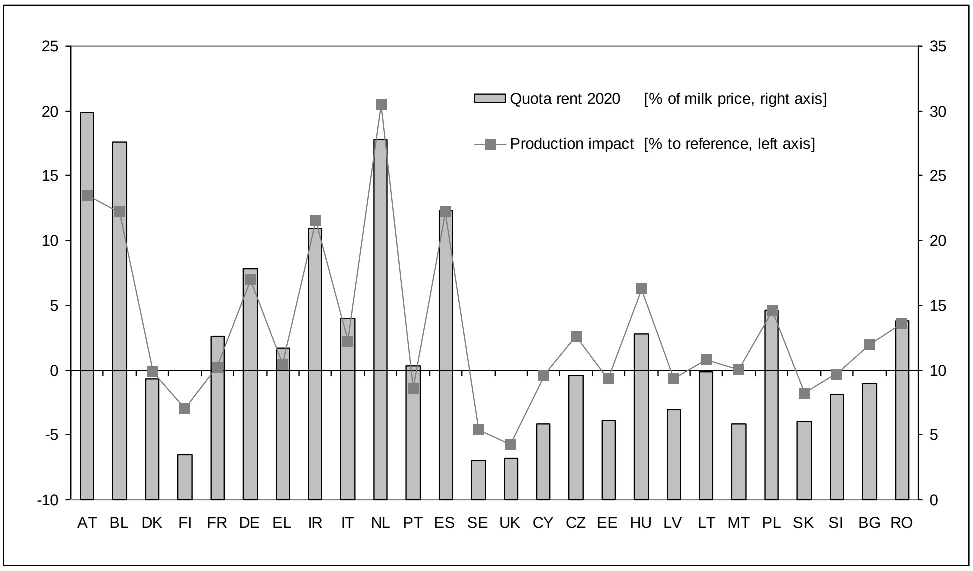
The common approach to identify the implications of a certain policy scenario such as the abolition of milk quotas is to repeat the baseline scenario with all parameters and exogenous inputs maintained except those under investigation, in this case the milk quotas. In the reform scenario milk production will be determined according to the usual ‘marginal cost equals price’ optimality condition, with marginal cost determined from the calibrated cost functions, endogenous quantities and shadow prices of fixed factors.

4 Impacts of Milk Quota Reform

4.1 Milk supply and dairy herds

Recent studies have confirmed the importance of quota rents in the quantitative results obtained in milk quota abolition scenarios (see Requillart, Bouamra-Mechemache, Jongeneel & Penel, 2008; Witzke & Tonini, 2009). This study has therefore devoted considerable efforts to estimate quota rents and supply elasticities at regional level, to merge the results with existing information from quota markets and to perform plausibility checks on the obtained results. As expected, the final specification of quota rents appears to be a key determinant for the milk production results (see correlation patterns in 0 4.2).

Figure 4.2: Quota abolition impacts on production of cow milk in EU27 and baseline quota rents, year 2020



Source: Own calculation.

In Figure 4.2, milk production and quota rents per MS do not perfectly match because there are other determinants for production in the model, such as demand elasticities by the dairy industry, regional supply elasticities and regional fodder production constraints. Nonetheless, the key message is that regional production impacts are crucially depending on the quota rent specification. Production impacts go back to changes in dairy herds and yields, which are the starting point for the detailed analysis of economic impacts in the European agricultural sector, as presented in Table 4.3.

Table 4.3 shows a 4.4% average increase in milk production, mainly due to a change in dairy herds, while milk yields are almost stable. This increase in dairy herds usually translates into a modest increase of overall cattle density because other cattle activities (i.e. animals for fattening) will not be significantly affected and suckler cows even decline (see also explanations in section 4.2). The Netherlands appear as a special case, since the cattle density strongly increases here (by 12.5%) together with the increasing dairy herd. However, environmental regulations on manure disposal which are not reflected in CAPRI could dampen the expansion of dairy production in this MS compared to the presented results. Milk yields tend to increase in most MS after the reform, because the milk quota rent ($p - \overline{p}$ in Figure 4.1) jumps to the market level (p). However regional shifts of production within an MS also impact on the change in the average yield.

Table 4.3: Changes in dairy herds, cattle density, yields, and cow milk production, year 2020

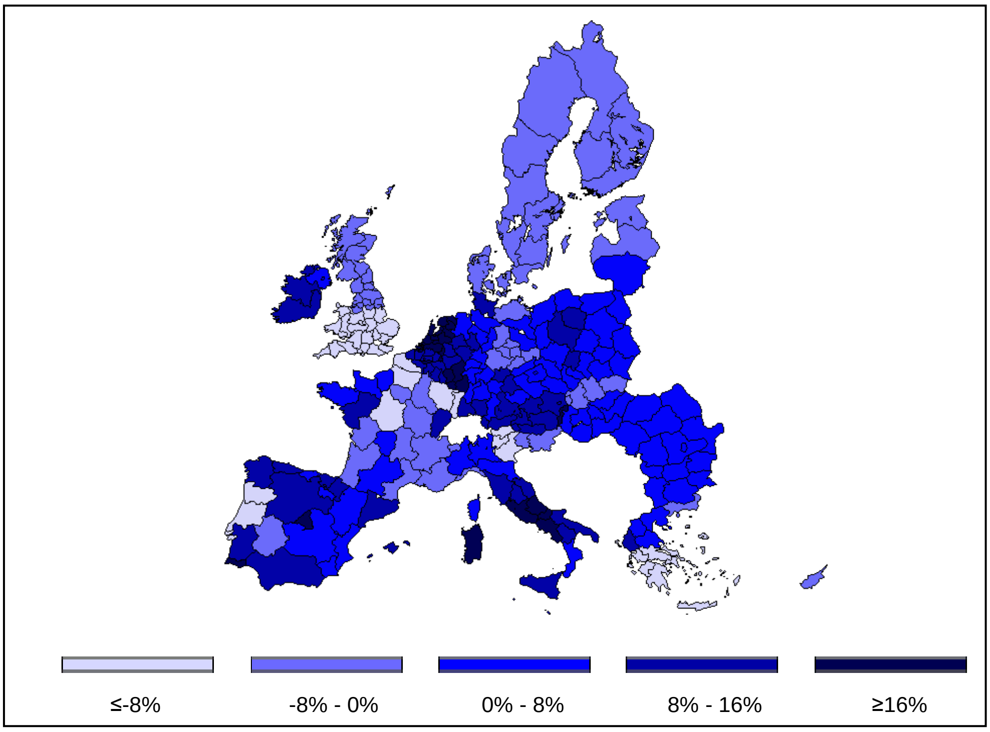
	Baseline				Milk Quota Reform			
	Dairy herd [1000 hd]	Cattle density [LU/ha]	Milk yield [kg/hd]	Production [1000 t]	Dairy herd [% diff]	Cattle density [% diff]	Milk yield [% diff]	Production [% diff]
Austria	445	0,41	7170	3193	13,8	3,8	-0,3	13,5
Belgium-Lux.	524	1,19	6603	3460	11,9	2,5	0,3	12,2
Denmark	519	0,38	9092	4715	-0,3	0,4	0,2	-0,1
Finland	283	0,27	8906	2518	-3,2	-1,3	0,2	-3,0
France	3473	0,45	7244	25157	-0,3	-0,9	0,5	0,2
Germany	3887	0,47	7538	29297	6,9	3,5	0,1	7,0
Greece	128	0,09	6076	776	0,0	-1,3	0,4	0,4
Ireland	1066	1,12	5036	5369	11,1	1,4	0,4	11,6
Italy	1857	0,39	6110	11343	1,9	0,2	0,3	2,2
Netherlands	1366	1,23	8185	11179	20,0	12,5	0,5	20,5
Portugal	286	0,31	7180	2056	-0,3	-1,0	-1,1	-1,4
Spain	931	0,25	7048	6563	11,1	0,2	1,0	12,2
Sweden	360	0,31	9198	3314	-4,8	-2,2	0,2	-4,6
United Kingdom	1883	0,43	8001	15063	-5,8	-2,4	0,1	-5,7
EU15	*17007*	*0,41*	*7291*	*124003*	*4,6*	*0,7*	*0,1*	*4,7*
Cyprus	24	0,30	6304	150	-0,5	-0,4	0,1	-0,4
Czech Republic	326	0,19	8320	2713	2,7	1,0	-0,1	2,6
Estonia	98	0,20	6840	670	-0,8	0,0	0,2	-0,7
Hungary	244	0,07	7720	1882	6,1	4,0	0,1	6,2
Latvia	171	0,17	4843	827	-0,8	-0,1	0,1	-0,7
Lithuania	366	0,22	5206	1903	0,7	1,1	0,1	0,8
Malta	7	1,10	6696	44	-0,2	-0,8	0,3	0,1
Poland	2030	0,21	5577	11322	4,5	3,4	0,1	4,7
Slovac Republic	144	0,13	7194	1037	-2,0	-1,4	0,2	-1,8
Slovenia	111	0,68	6103	676	-0,4	-0,8	0,1	-0,3
10 New MS	*3519*	*0,18*	*6031*	*21222*	*3,2*	*2,3*	*0,1*	*3,3*
Bulgaria	342	0,14	3686	1260	1,4	0,9	0,5	2,0
Romania	1289	0,18	3623	4671	3,0	2,6	0,6	3,6
Bulgaria/Romania	*1631*	*0,17*	*3636*	*5931*	*2,7*	*2,3*	*0,6*	*3,3*
EU27	*22157*	*0,35*	*6822*	*151156*	*4,2*	*1,0*	*0,2*	*4,4*

Source: Own calculation.

Increasing production exerts downward pressure on producer prices, which are declining on average by 10%. As raw milk is badly tradable, price formation is assumed to occur on the national level such that percentage changes in producer prices may be heterogeneous. High production increases tend to trigger strong price drops but markets for dairies also intervene. Basically, the profitability of the dairy processing industry and, hence, the equilibrium prices for particular deliveries also depend on changing prices of dairy products and on their weights in the national industry. Therefore, we can expect a decline in raw milk prices also in those MS where production is likely to decrease. In fact lower dairy prices (i.e. at the processing level) indirectly depress raw milk prices in the whole EU27 and explain why production is declining at all in MS with zero or small positive rents in the baseline.

Figure 4.3 visualizes regional effects on milk production, i.e. the percentage changes in the quota abolition scenario compared to the baseline. In bigger countries like Germany, France and UK, there are quite significant differences within regions. For instance, in Germany a significant reduction of milk production is expected in the Eastern part, while most of the remaining regions expand their production. On average German milk production would moderately increase. In the UK we observe an overall reduction of milk supply, whereas this decline is more considerable in the southern part than in the North. Finally it may be seen that the increase in the Netherlands is quite exceptional at MS level but that some other regions are responding in a similar way

Figure 4.3: Percentage change of milk production in European regions, year 2020 [% difference to the baseline]



Source: Own calculation.

4.2 Beef and cattle herds

Dairy markets are related to meat markets over several channels. In the cattle sector an expansion of the dairy herd will directly release some meat from old cows and render young animals cheaper, but it also means increasing competition for fodder. Beef meat activities – fattening of bulls, heifers and calves as well as suckler cows – compete with dairy cows for regional feed resources. This would result in opposite effects in beef and dairy sectors. When demand from an increasing dairy cow herd bids up the prices of non tradable fodder produced in the regions, beef production loses profitability. This effect is moderated when tradable feedstuff can be adopted or fodder areas can be adjusted. The previous considerations reveal that the interdependencies among cattle activities can lead to parallel as well as antagonistic changes in dairy and beef meat activities.

Table 4.4: Impacts on selected cattle types and beef production, year 2020

	Baseline				Milk Quota Reform			
	Dairy cows	Suckler cows	Bulls and heifers fattening	Beef	Dairy cows	Suckler cows	Bulls and heifers fattening	Beef
	[1000 heads]			[Mio t.]	[% diff to Baseline]			
Belgium and Lux.	524	638	128	246	11,9%	-3,5%	-1,5%	2,3%
Denmark	519	17	226	107	-0,3%	-20,3%	2,8%	1,3%
Germany	3887	459	1347	1017	6,9%	-9,2%	0,2%	2,8%
Austria	445	261	328	170	13,8%	-5,0%	-0,7%	2,1%
Netherlands	1366	93	6	272	20,0%	-7,3%	-1,3%	8,9%
France	3473	4024	2039	1466	-0,3%	-3,4%	-0,3%	-0,3%
Portugal	286	505	168	94	-0,3%	-2,5%	0,0%	0,1%
Spain	931	2725	2508	618	11,1%	-4,4%	1,0%	0,8%
Greece	128	155	128	39	0,0%	-4,5%	0,2%	0,4%
Italy	1857	454	2134	865	1,9%	-6,9%	-0,2%	-0,1%
Ireland	1066	1343	1475	514	11,1%	-4,8%	0,5%	0,7%
Finland	283	20	144	64	-3,2%	-6,9%	1,4%	-0,3%
Sweden	360	136	241	120	-4,8%	-4,1%	0,3%	-1,2%
United Kingdom	1883	1563	1862	726	-5,8%	-5,1%	0,7%	-1,0%
European Union 15	*17007*	*12392*	*12732*	*6316*	*4,6%*	*-4,4%*	*0,3%*	*0,9%*
Czech Republic	326	97	73	76	2,7%	-4,3%	-2,3%	-0,1%
Estonia	98	0	25	15	-0,8%	-6,3%	0,3%	0,1%
Hungary	244	20	31	37	6,1%	-5,3%	-2,0%	3,1%
Lithuania	366	0	60	37	0,7%	-15,2%	-0,3%	0,9%
Latvia	171	0	64	16	-0,8%	-20,8%	0,2%	0,6%
Poland	2030	2	638	319	4,5%	-31,9%	-1,2%	1,1%
Slovenia	111	92	73	50	-0,4%	-1,6%	-1,6%	-1,4%
Slovak Republic	144	14	34	36	-2,0%	-3,0%	-3,1%	-2,4%
Cyprus	24	0	16	4	-0,5%	0,0%	-0,4%	-0,7%
Malta	6,5	0,00	2,49	1	-0,3%	0,0%	-2,8%	-1,8%
European Union 10	*3519*	*225*	*1016*	*590*	*3,2%*	*-3,5%*	*-1,2%*	*0,6%*
Bulgaria	342	39	201	49	1,4%	-3,5%	0,0%	0,5%
Romania	1289	55	835	205	3,0%	-2,1%	2,3%	2,4%
Bulgaria and Romania	1631	94	1036	255	2,7%	-2,6%	1,9%	2,0%
European Union 27	22157	12712	14783	7161	4,2%	-4,4%	0,3%	0,9%

Source: Own calculation. Note that calves are omitted from the table.

The net impact of increased availability of calves and reinforced competition for fodder seems to have a very small expansionary effect on beef supply, except for the Netherlands where it is sizeable. The resulting decline of beef prices (about -2.8% in EU15) stimulates demand. In those NMS with a declining dairy herd, beef supply reacts negatively, whereas in others, most importantly Poland, there is an increase in supply giving a total increase for EU10 of 0.6%. Prices in the NMS are declining similar to EU15. It is evident that the overall change in beef production is closely correlated to the change in milk supply although the magnitude of effects is smaller

Percentage changes in suckler cow herds are often high because of low absolute numbers. Suckler cow herds tend to decrease in all countries suggesting that lower prices for calves all over Europe reduce their profitability. Significant changes in suckler cow herds can result in a change in net trade of calves without affecting the national beef production. In EU15, dairy cow herds increase by 780.000 heads, whereas suckler cows decrease by 550.000. Hence about 70% of the additional supply of calves stemming from dairy cows is compensated due to declining suckler cow herds. In the NMS suckler cows herds are almost negligible and so there is no such compensatory effect. Fattening activities are negatively affected due to decreasing beef prices but benefit from cheaper calves at the same time. Profitability and herd sizes are almost unchanged. Comparing herd sizes of dairy and fattening activities, results might suggest that there is a weak negative correlation in the NMS. Competition for fodder areas may explain this observation as argued above.

We conclude that the effects of the quota abolition on the beef sector are quite small. The effects are moderated by opposite changes in the suckler cow herd, but the overall beef supply tends to increase due to the increasing number of slaughtered dairy cows and cheaper calves.

4.3 Other animal herds

The quota abolition indirectly affects other animal sectors. Increased competition for compound feed (mostly from cereals) increases the production costs for pigs and poultry activities. Moreover, the expansion in dairy cow herds exerts further pressure on sheep and goats activities via increased demand for non tradable roughage. In addition, the demand system of CAPRI allows lower beef and milk prices on the market to affect the prices of meat from other animals: negatively from the direct substitution effect and positively from the positive income effect. The observed effects are generally small and reported in Table 4.5. Changes in the number of sheep and goats are moderate but higher than changes for pigs and poultry.

Table 4.5: Aggregated effects on pigs, poultry, sheep and goat, year 2020

	Baseline			Milk Quota Reform		
	pigs	poultry	sheep and goat	pigs	poultry	sheep and goat
		[Mio head]			[% diff to Baseline]	
European Union 15	236,1	354,7	111,0	-0,21%	-0,06%	-1,04%
European Union 10	38,8	100,7	2,7	-0,25%	0,00%	-0,81%
Bulgaria and Romania	4,1	14,8	12,9	-0,61%	-0,61%	-0,42%
European Union 27	279,0	470,2	126,7	-0,22%	-0,07%	-0,97%

Source: Own calculation.

4.4 Land use changes

The land allocation, i.e. the production level of crops on arable land, could be influenced by the quota abolition since fodder production activities on arable land compete with other crops for the fixed resource land. The CAPRI model distinguishes permanent grassland and arable land, but both land qualities are fixed during simulation. Changes in cattle herds could hence only change fodder production on arable land, mainly silage maize and temporary grazing, whereas for permanent grassland only the intensity can be adjusted. Among the other uses, cereals are most frequently occupying the largest part of arable land such that indirect impacts from scarce area are best visible here. On the other hand cereals are the most important tradable feedstuff which may be expected to increase in demand if production of milk and beef is dominating the decline in pork and poultry production.

However, model results show only moderate changes in fodder production on arable land. Since the area of this activity aggregate is small compared to total fodder production and cereals, in most MS no considerable changes in land use can be found for those aggregates.

Table 4.6: Aggregated land use changes, year 2020

	Baseline			Milk Quota Reform		
	Cereals	Fodder (total)	Fodder (on arable land)	Cereals	Fodder (total)	Fodder (on arable land)
		[Mio ha]			[% diff to Baseline]	
European Union 15	35,7	62,6	13,2	-0,01%	0,03%	0,14%
European Union 10	13,9	9,8	2,4	-0,03%	0,15%	0,63%
Bulgaria and Romania	7,0	7,3	0,9	0,05%	0,03%	0,26%
European Union 27	56,6	79,8	16,5	-0,01%	0,05%	0,22%

Source: Own calculation.

Changes in land allocation become visible when we focus our analysis on specific production activities. Looking at changes of the activities fodder maize and temporary grazing we observe more sizable effects. However these effects are often antagonistic, i.e. when fodder maize increases, temporary grazing goes down and vice-versa. The grass yield on permanent pastures is almost unchanged at European level, but varies among regions between -1% and +8%. Most significant increases can be found in the Netherlands, where milk production goes up strongly.

4.5 Agricultural income

Table 4.7: Income effects of quota abolition in agriculture, year 2020 [million €]

	Baseline				Milk Quota Reform (diff to Baseline)			
	agricultural income	from cow milk	from meat	non fodder feed cost	agricultural income	from cow milk	from meat	non fodder feed cost
Austria	3752	899	1884	737	-78	-5	-11	45
Belgium-Lux.	4463	987	3636	1714	-127,0	-37,7	-16,8	41,4
Denmark	4492	1569	3974	2312	-152,7	-130,7	-15,2	5,9
Finland	1543	955	828	586	-69,0	-69,5	-9,5	-9,8
France	37921	7548	16868	8509	-1070,0	-803,3	-255,1	-76,9
Germany	24004	9170	13788	6981	-870,2	-530,0	-42,9	227,4
Greece	11175	278	1440	1089	-130,0	-25,3	-17,1	-14,8
Ireland	3483	1526	2600	1076	-157,3	-22,9	-40,3	54,8
Italy	38678	4191	10594	6072	-538,9	-317,5	-136,5	10,7
Netherlands	12565	3955	4974	3012	-107,2	207,0	67,8	203,1
Portugal	3843	688	1726	1397	-71,0	-57,7	-19,9	-15,4
Spain	42087	2008	12706	7310	-386,3	-52,9	-105,7	50,1
Sweden	2114	1132	1103	380	-109,7	-105,7	-21,2	-8,1
United Kingdom	13585	4182	9032	4545	-358,3	-422,3	-160,0	-140,6
EU15	203705	39087	85152	45720	-4225,4	-2373,6	-783,7	372,4
					-2,1%	-6,1%	-0,9%	0,8%
Cyprus	459	69	267	194	-10,5	-3,3	-3,2	-1,7
Czech Republic	2201	766	1067	741	-56,4	-37,1	-7,0	10,3
Estonia	318	164	163	127	-10,6	-10,0	-1,1	-2,9
Hungary	4044	508	2065	1388	-40,1	-14,5	-9,0	19,4
Latvia	344	162	128	96	-12,3	-9,6	-0,8	-1,1
Lithuania	945	348	405	231	-36,2	-22,7	-1,4	0,1
Malta	56	16	41	48	-1,3	-0,4	-1,0	-0,3
Poland	10765	2409	6246	3254	-220,8	-122,7	-23,7	66,3
Slovac Republic	887	287	627	367	-24,9	-21,0	-7,7	-1,1
Slovenia	564	175	326	172	-16,3	-12,6	-6,8	-1,6
10 New MS	20584	4903	11334	6618	-429,4	-253,9	-61,7	87,5
					-2,1%	-5,2%	-0,5%	1,3%
Bulgaria	1949	295	609	313	-16,6	-1,6	-4,5	8,8
Romania	7163	809	1558	1294	1,8	11,1	-4,0	60,6
Bulgaria/Romania	9112	1103	2167	1607	-14,8	9,5	-8,5	69,4
					-0,2%	0,9%	-0,4%	4,3%
EU27	233400	45094	98653	53945	-4669,6	-2618,0	-853,8	529,3
					-2,0%	-5,8%	-0,9%	1,0%

Source: Own calculation.

The regional income effects follow from price and quantity impacts on the input and output side. The bottom line in terms of agricultural income is crucially determined from the impacts on revenues from raw milk and meat products, and the related impacts on non-fodder feed items.

The overall loss of agricultural income is expected to be nearly 4.7 billion € or 2.0%. It may be attributed to a large extent to the above mentioned components. In some MS, such as the Netherlands, there is a higher use of non-feed inputs, related to a high intensity of production, which are also increasing if production is expanding as projected (see Table 4.3). Hence the three components presented in Table 4.7 only explain a smaller part of the overall income effect for the Netherlands[32] than in many other MS.

4.6 Environmental effects

The calculation of a selected number of environmental indicators is included in the CAPRI modelling system (see Pérez Domínguez, 2006 for a description of nitrogen flows and calculation of gaseous emissions). While gaseous losses can be seen as a "global" problem, pollution of soil and ground water are site specific negative external effects of agricultural production. Nitrogen losses from soils in the form of nitrate leaching in particular are worth analyzing in more regional detail since average figures at MS or EU27 level often ignore environmental pressure at specific hot spots. The regional analysis of a milk quota reform scenario gives insights into whether a further concentration of animal husbandry at already highly affected regions can be expected.

Table 4.8: Environmental indicators in the EU27, year 2020

Environmental Indicators		Milk Quota reform vs. Baseline [% diff]
atmospheric	Gaseous loss mineral fertilizer (NH3 & N20 & NOX)	0,76%
	Ammonia loss mineral fertilizer (NH3)	0,76%
	Gaseous loss manure (NH3 & N2O & N2 & NOX)	0,66%
	Ammonia loss manure (NH3)	0,70%
	Methane(CH4)	1,41%
soil	N Surplus at soil level	1,05%
	Nitrate Leaching	1,33%
	Denitrification	0,98%

Source: Own calculation.

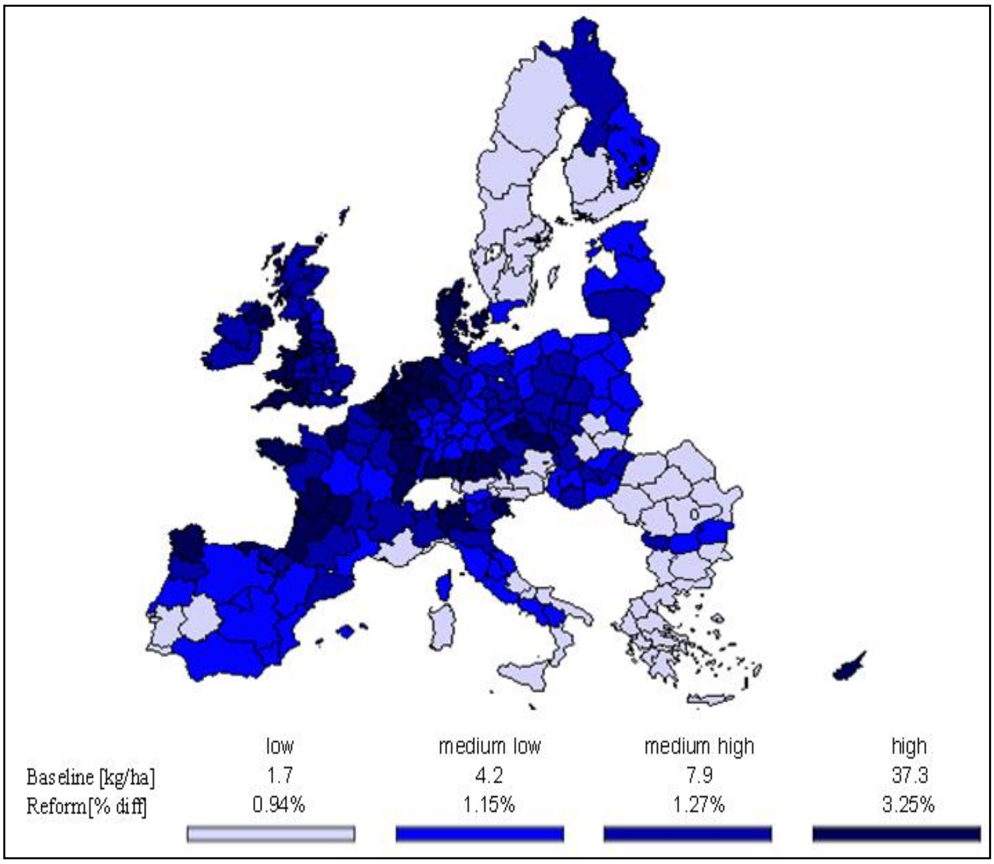
Environmental effects in CAPRI are strongly linked to the nitrogen (N) cycle of agricultural activities. Since dairy herds are simulated to increase after the milk quota reform, N losses are estimated to follow the same path. However the actual increase is quite moderate (0.66% - 1.41%). This is because the effect on the nitrogen cycle of 930.000 (4.6%) additional dairy cows is attenuated by a parallel decline of 560.000 suckler cows, even though less emissions per cow are associated with the latter. For gaseous losses, it can be seen that the increase in environmental pressure is only small, except for methane. N

[32] The share of 'industrial' inputs like fertilisers, pharmaceuticals, energy, capital etc. in total input is 66% in the Netherlands, whereas it is only 45% in EU15. The cost for those inputs increases by 154 million Euros after the reform for the Netherlands, and thus nearly fully explains the loss in income in spite of increased revenues.

losses from soil require a more detailed spatial analysis since unfavourable concentration of animal production might occur in a fully liberalized dairy sector. The fact that the aggregate increase of nitrate leaching (1.33%) exceeds the nitrogen surplus at soil (1.05%) indicates that production will increase in particular in those regions where the risk of nitrate leaching is relatively high. Whereas these changes are plausible it has to be admitted that they are quite small and the usual modelling uncertainties also apply to the environmental indicators.

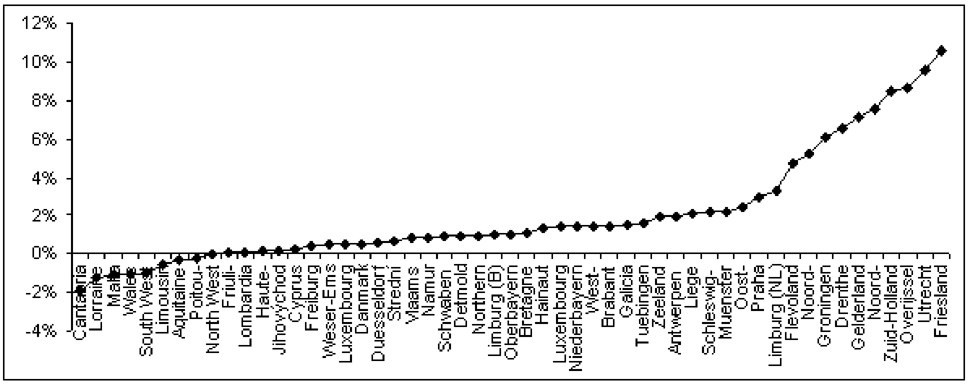
The division of the EU27 in 4 clusters ordered by nitrate leaching (see Figure 4.5) reveals that higher effects can be expected in regions with high environmental pressure in the baseline. Looking in more detail at the "high" group, the analysis shows that only in a limited number of regions, mainly 11 Dutch regions, drastic increases in nitrogen losses from soils are expected. These exceptional effects in the Netherlands can be explained since Dutch dairy farmers are highly competitive relative to other regions and, hence, tend to increase their herds significantly. Furthermore, the Dutch beef cattle herd is small relative to the dairy cow herd, so that reductions in beef cattle herds are insufficient to compensate for the effect of dairy herd increases.

Figure 4.5: Nitrate leaching in European regions, year 2020 (four equal groups, % difference to the baseline)



Source: Own calculation.

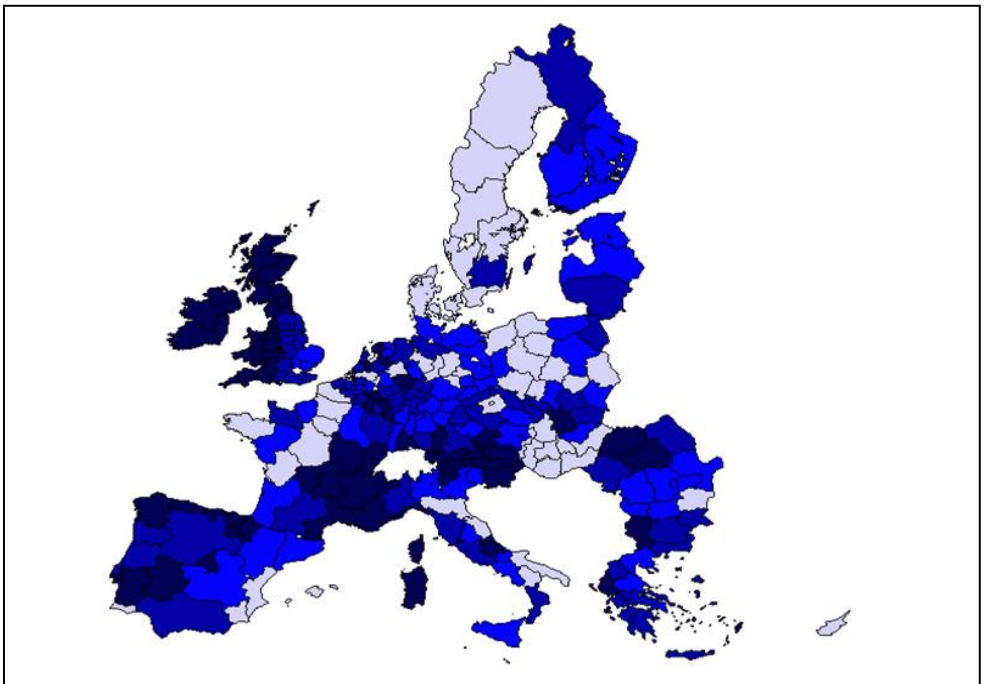
Figure 4.6: Changes in nitrate leaching of highly affected regions (% difference to the baseline)



Source: own calculation.

However, the environmental regulations in the Netherlands are quite strict and effective but their representation in CAPRI is limited (see previous comments in section 4.1). Thus it can be assumed that negative external effects in the Netherlands would be in fact lower than projected. Other regions where environmental pressure might become serious are located in Belgium and North-West Germany as well as specific regions in France (Bretagne) and Spain (Galicia), see Figure 4.6.

Figure 4.7: Regions classified by share of permanent pastures (four equal groups) combined with cattle density characteristics



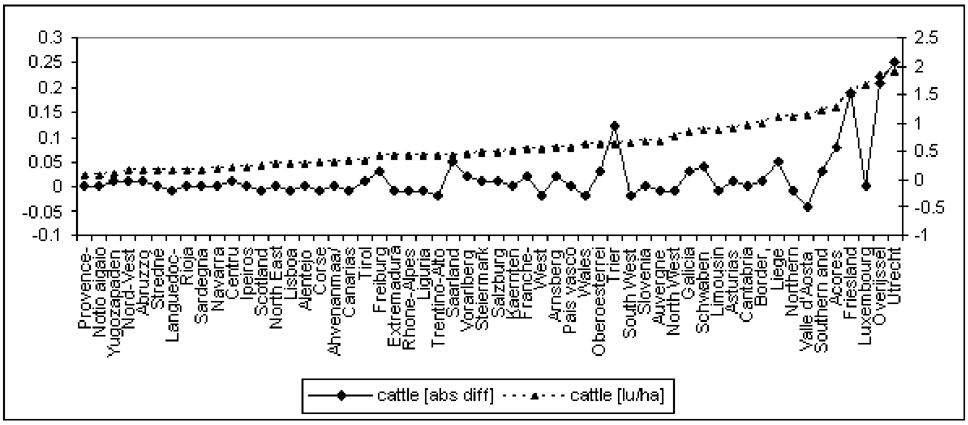
	low	medium low	medium high	high
Grassland as % of UAA	12.0	25.1	39.0	63.9
livestock density (cattle)				
Baseline [LU/ha]	0.24	0.35	0.48	0.59
Reform [abs diff]	0.01	0.01	0.02	0.01

Source: Own calculation.

Another important environmental issue is land abandonment in marginal areas, currently used for animal grazing. After grouping European regions by their share of permanent pasture on total agricultural land, the livestock density for each cluster is calculated. The livestock density per hectare is clearly positively correlated with the importance of grassland in the region.

However, impacts of the quota abolition on the cattle density are very homogeneous. Naturally, we find some less competitive pasture regions, but there is no overall movement of dairy and cattle herds from pasture regions towards arable regions, where maize is typically a major feed component. Hence concerns about particularly strong impacts on cattle herds in regions dominated by grassland are not confirmed by the CAPRI results. In order to present more details on the dominant pasture regions we sorted those regions by increasing livestock density. We can see in Figure 4.7 that regions with already low cattle densities in the reference situation do not change much after the reform. In these regions, mainly located in Southern France and the Spanish Peninsula, extensive beef cattle is accounting for most of the livestock and hence effects of the quota abolition are quite limited.

Figure 4.8: Change in cattle density in regions with high share of pastures (ordered by increasing overall cattle density, % difference to the baseline)



Source: Own calculation.

5 Sensitivity Analysis

Although the results of the milk quota reform presented in the previous sections are in line with results of other studies the simulations are based on model parameters that cannot be estimated accurately. The milk supply elasticity and the quota rents used to calibrate the CAPRI supply part are significantly influencing the simulation results. Hence we calculated additional quota abolition scenarios where those exogenous model parameters were varied.

5.1 Elasticity of milk supply

The supply elasticity derived from econometric estimation enters the CAPRI model as an exogenous parameter used to determine the slope of the marginal cost function underlying the regional supply functions. The supply models can be successively calibrated to different supply elasticities before the quota abolition scenario S4 is simulated. Here 4 scenarios are calculated: (1) ELAS_150, with the milk supply elasticity increased by 50%, (2) ELAS_125, with the milk supply elasticity increased by 25%, (3) ELAS_75, with the milk supply elasticity reduced by 25%, and (3) ELAS_50, with the milk supply elasticity reduced by 50%

Compared to the standard quota abolition scenario the effect of elasticity variation on overall milk supply is quite small. Looking at prices it becomes clear that lower elasticities of milk production would lead to a lower price decline (i.e. higher prices) for milk. The effect on the overall agricultural income is negligible.

Table 4.9: Summary of simulation results with respect to different supply elasticities

	Baseline	Milk Quota Reform	Milk Quote Reform with different supply elasticities			
			ELAS_150	ELAS_125	ELAS_75	ELAS_50
Milk Production	[1000t]	[% diff]	[% diff]	[% diff]	[% diff]	[% diff]
EU15	124003	4,68	4,94	4,93	4,88	4,65
10 New MS	21222	3,30	3,50	3,43	3,37	3,31
Bulgaria/Romania	5931	3,28	3,61	3,49	3,00	2,32
EU27	151156	4,43	4,66	4,66	4,61	4,37
Producer Price Milk	[€/t]					
EU15	315	-10,27	-10,93	-10,48	-8,96	-7,52
10 New MS	231	-8,21	-8,57	-8,19	-6,98	-5,86
Bulgaria/Romania	186	-2,34	-2,39	-2,20	-1,74	-1,40
EU27	298	-9,80	-10,41	-9,97	-8,51	-7,13
Agricultural Income	[Mio €]					
EU15	203705	-2,07	-2,22	-2,11	-1,74	-1,41
10 New MS	20582	-2,08	-2,14	-2,04	-1,71	-1,43
Bulgaria/Romania	9112	-0,16	-0,40	-0,69	-0,65	-0,04
EU27	233399	-2,00	-2,15	-2,05	-1,70	-1,35

Source: Own calculation.

5.2 Quota rents

Similarly to supply elasticities, quota rents enter the calibration of the CAPRI supply models as exogenous parameters. In order to assess the effects of different quota rents[33] the following alternative scenarios were calculated: (1) quota rent increased by 5ct/kg milk, (2) quota rent increased by 2ct/kg milk, (3) quota rent decreased by 2ct/kg milk, and (4) quota rent decreased by 5ct/kg milk.

The quota rents were shifted in absolute terms since a percentage variation would not affect all those regions where the quota rent is assumed to be 0. Given that quota rents in the baseline differ among countries in a range of 0 – 10 ct/kg milk an error of +/- 5 ct/kg has to be considered possible. In this sensitivity analysis the change in milk supply can range from almost 0% to +14% in the European Union. At the same time milk prices

[33] Note: for this sensitivity analysis quota rents are considered positive or negative variations in variable costs of the corresponding regional supply model (e.g. an increase of quota rent by +5ct/kg milk means a corresponding reduction of 5ct/kg milk in production costs).

vary from about -20% to 0%. The rather inelastic demand for milk products leads to declining agricultural income when quota rents are assumed to be higher, i.e. the increase in milk production cannot compensate the drastic drop in prices. The prices decrease although the budgets for export subsidies of dairy products rise significantly.

Table 4.10: Summary of simulation results with respect to different quota rents

	Baseline	Milk Quota Reform	Milk Quote Reform with different marginal cost			
			-5ct/kg milk	-2ct/kg milk	+2ct/kg milk	+5ct/kg milk
Milk Production	[1000t]	[% diff]	[% diff]	[% diff]	[% diff]	[% diff]
EU15	124003	4,68	10,94	7,49	3,09	-0,10
10 New MS	21222	3,30	11,05	6,64	0,89	-3,47
Bulgaria/Romania	5931	3,28	10,65	6,40	0,47	-4,08
EU27	151156	4,43	10,94	7,33	2,68	-0,73
Producer Price Milk	[€/t]					
EU15	315	-10,27	-20,51	-14,20	-5,66	0,85
10 New MS	231	-8,21	-21,42	-13,23	-2,07	6,34
Bulgaria/Romania	186	-2,34	-6,20	-3,52	-0,13	2,28
EU27	298	-9,80	-20,26	-13,80	-5,05	1,63
Agricultural Income	[Mio €]					
EU15	203705	-2,07	-4,16	-2,83	-1,10	0,14
10 New MS	20582	-2,08	-5,21	-3,16	-0,42	1,18
Bulgaria/Romania	9112	-0,16	-1,62	-0,47	-0,37	0,59
EU27	233399	-2,00	-4,15	-2,77	-1,01	0,25
Export subsidies dairy products	[Mio €]	[Mio €]	[Mio €]	[Mio €]	[Mio €]	[Mio €]
EU27	6	55	212	102	23	4

Source: Own calculation.

6 Conclusions

This paper reports on the use of econometric estimates of marginal costs of milk producers in the CAPRI model. This new information increases the validity of the analysis, as it provides careful estimates of regional quota rents and price-supply elasticities for raw milk based on historical FADN records. Moreover, expert data and projections from other recent studies on dairy commodities have been included in the analysis. This provides the basis for a comprehensive quantitative assessment of possible implications of the milk quota abolition that was a part of the 2008 CAP Health Check decisions.

The impacts follow from a comparison of the results of a milk quota reform scenario (year 2020) and a baseline situation (2020 with quotas in place). Simulation results indicate that the abolition of the milk quota regime is likely to increase milk production on average by 4.4% in EU27, and to decline raw milk prices by -10%. Agricultural income would decline on average by -1.6% since increasing production cannot compensate lower milk prices. These results are in line with results of other recent studies. The study providing important input to the CAPRI baseline by Réquillart, Bouamra-Mechemache, Jongeneel & Penel 2008 obtained an increase in production of 5.2% with prices declining by 11%. Witzke and Tonini (2009) reported a production increase of only 3% with a price drop of 7%. Chantreuil, Donnellan, van Leeuwen, Salamon, Tabeau, & Bartova (2008) reported a production increase of 4.8% with prices declining by 7%.

Even though the results are in line with those of other studies it would have been desirable to perform a related ex-post validation exercise to compare model outcomes with real data. Such a validation has not been undertaken with CAPRI and is difficult to perform for several reasons: Firstly, there is no suitable historical evidence of a similar policy change in the EU as the quota regime is in place for many years with moderate variations only over time. Checking prior ex-ante simulations with CAPRI on the 2003 CAP 'Mid Term Review' such as that by Wieck, Britz, Pérez Dominguez and Jansson (2004) runs into the difficulty that unanticipated events such as the rapid growth of the bioenergy sector and the recent food price "crisis" have dominated past developments such that little may be inferred from such an analysis.

Our model follows basically the profit maximization hypothesis and simulates the effects after an adoption period of 5-10 years. Other determinants of farmers' decisions like liquidity, risk attitudes and expectations on future developments and prices are not explicitly addressed. Some of these play a major role in short term decisions. Hence our analysis is not suitable, for example, to assess the short term impacts in specific regions of the price fluctuations in recent years from 2006 to 2010.

Other assumptions are embedded in the parameters of the model and their uncertainty is inherited to all results. A sensitivity analysis has been performed on two potentially critical ones. It turned out that the assumed supply elasticities are far more influential on regional impacts than the variation of the quota rents. It has to be stressed that different quota rents would have had significant effects on the results of milk production as well as on milk prices and agricultural income.

An explicit focus of this paper is on regional effects in the EU27. The impacts on dairy herds are quite heterogeneous among regions. Highly competitive regions tend to expand their milk production up to 30% and thus may be able to increase their revenues. Less competitive regions will lose revenues both from price and quantity sides (up to -20%). Incomes within EU MS are most heterogeneously affected in Germany, Portugal and Spain. In Germany, income gains up to 5% are observed in benefiting regions, with income losses amounting to 7% in most negatively affected regions. The sensitivity analysis revealed that the higher the assumed elasticity of milk supply, the wider the variety of regional effects. While high supply elasticities tend to increase the gap between winning and losing regions, lower supply elasticities favour more uniform changes among regions.

Moreover, the model results shed light on questions where regional extreme values are more important than average effects. The analysis allows the identification of regions where specific problems might need special attention, due to increased nitrate leaching or due to a risk of land abandonment. Some regions were identified where corresponding measures might be needed to counteract an increase of nitrate leaching. However, animal density and agricultural income are fairly stable in marginal areas at the spatial resolution of this analysis, suggesting that the quota abolition does not involve a marked increase in the risk of land abandonment.

7 Acknowledgements

The paper is based on a study carried out for the European Commission's Joint Research Centre, Institute for Prospective and Technological Studies (JRC-IPTS). We thank Daniele Moro and John Helming for their contributions to the development of the model and database and Axel Tonini for useful comments. We thank ESTAT, DG-AGRI market unit (C4) and DG-AGRI quantitative analysis (L2) for providing additional data and evaluation of results. The views expressed are purely those of the authors and may not in any circumstances regarded as stating an official position of the European Commission.

8 References

Adenäuer, M. (2006): Modelling the European Sugar Sector - Incentives to Supply Sugar Beets and Analysis of Reform Options; Institute for Food and Resource Economics, University of Bonn, Retrieved from http://hss.ulb.uni-bonn.de:90/2006/0759/0759.pdf.

Bartova, L., Fellmann T., & M'barek R. (Eds.). (2008). Modelling and Analysis of the European Milk and Dairy Market. AGMEMOD Consortium. European Commission, Joint Research Centre, Institute for Prospective Technological Studies, Seville.

Bouamra-Mechemache, Z., Jongeneel, R., & Réquillart, V. (2008). Impact of a gradual increase in milk quotas on the EU dairy sector. European Review of Agricultural Economics, 35(4), 461–491.

Britz, W., & Witzke H. P. (2008). CAPRI model documentation. Institute for Food and Resource Economics, University of Bonn. Retrieved from http://www.capri-model.org/docs/capri_documentation.pdf.

Chantreuil, F., Donnellan, T., van Leeuwen, M., Salamon, P., Tabeau, A., & Bartova, L. (2008). EU dairy quota reform – AGMEMOD scenario analysis, Paper prepared for presentation at the XII[th] Congress of the European Association of Agricultural Economists, People, Food and Environment: Global Trends and European Strategies, Ghent: University of Ghent, 2008.

Colman, D., Burton, M., Rigby, D., & Franks, J. (1998). Economic evaluation of the UK milk quota system. Final report to the MAFF. Centre for Agricultural, Food and Resource Economics, School of Economics Studies, University of Manchester.

Colman, D. (Ed.) (2002). Phasing Out Milk Quotas in the EU. Centre for Agricultural, Food and Resource Economics, School of Economic Studies, University of Manchester.

Elbersen, B.S., Kempen, M., van Diepen, C.A., Andersen, E., Hazeu, G.W., & Verhoog, D. (2006). Protocols for spatial allocation of farm types. SEAMLESS integrated project, EU 6th Framework Programme, contract no. 010036-2, SEAMLESS report no.19.

European Commission (2006). Milk and Milk Products in the European Union. Brussels.

European Commission (2007a). Preparing for the Health Check of the CAP reform. Communication from the Commission to the European Parliament and the Council. Retrieved from http://ec.europa.eu/agriculture/healthcheck/index_en.htm.

European Commission (2007b). Market Outlook for the Dairy Sector. Retrieved from http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2007:0800:FIN:EN:PDF, Brussels.

European Commission (2008a). The agricultural situation in the European Union. Retrieved from http://ec.europa.eu/agriculture/publi/agrep/index_en.htm, Brussels.

European Commission (2008b). Prospects for Agricultural Markets and Income in the European Union 2007-2014 Market Outlook for the Dairy Sector. Retrieved from http://ec.europa.eu/agriculture/publi/caprep/prospects2007b/fullrep.pdf. Brussels.

FAPRI (2007), U.S. and World Agricultural Outlook, FAPRI Staff Report 07-FSR 1, Food and Agricultural Policy Research Institute, Ames, Iowa, USA. Online with data tables at http://www.fapri.iastate.edu/outlook2007/.

Guyomard, H., Delache, X., Irz, X., & Mahè, L.P. (1996). A microeconometric analysis of milk quota transfer: application to French producers. Journal of Agricultural Economics, 47(2), 206-223.

Guyomard, H., Herrard, N., Mahè, L. P. (1995). Modelling a siphon on quota transfers: an application to milk in france. Oxford Agrarian Studies, 23(1), 15-26.

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

Helming, J., & Peerlings, J. (2005). The impact of milk quota abolishment on Dutch agriculture and economy – applying an agricultural sector model integrated into a mixed input-output model. EAAE proceedings, Vol 10, 10th Congress of the European Association of Agricultural Economists, Zaragoza, August 28-31, 2002.

Howitt, R. (1995). Positive Mathematical Programming. American Journal of Agricultural Economics, 77(2), pp. 329-342.

Jansson, T. & Heckelei, T. (2008). Regional crop supply behaviour in the EU. Presented on the XII EAAE Congress, on August 26-29, Ghent.

INRA-Wageningen Consortium (2002). Study on the impact of future options for the milk quota system and the common market organisation for milk and milk products. V. Réquillart (coordinator), European Commission, DG Agri.

Isermeyer, F., Brockmeier, M., Gömann, H., Hargens, R., Klepper, R., Kreins, P., Offermann, F., Osterburg, B., Pelikan, J., Salamon, P., & Thiele, H. (2006). Analyse politischer Handlungsoptionen für den Milchmarkt. Landbauforschung Völkenrode Sonderheft 300, Braunschweig.

Kleinhanss, W., Manegold, D., Bertelsmeier, M.,Deeken, E., Giffhorn, E., Jägersberg, P., Offermann, F.,Osterburg, B., & Salamon, P. (2002). Phasing out MilkQuotas – Possible Impacts on German Agriculture. Bundesforschungsanstalt für Landwirtschaft (FAL), Arbeitsbericht 01/2002, Braunschweig.

Lips, M., & Rieder, P. (2002). Endogenous adjusted production quotas – the abolishment of the raw milk quota in the European Union. EAAE proc. Vol 10, 10thCongress of the European Association of Agricultural Economists, Zaragoza, Spain, August 28-31, 2002.

Lips, M., & Rieder, P. (2005). Abolition of Raw Milk Quota in the European Union: A CGE Analysis at the Member Country Level. *Journal of Agricultural Economics* 56(1), 1–17.

Moschini, G. (1988). The cost structure of Ontario dairy farms: a microeconometric analysis. Canadian Journal of Agricultural Economics 36(2), 187-206.

Pérez Domínguez, I. (2006). Greenhouse Gases: Inventories, Abatement Costs and Markets for Emission Permits in European Agriculture - A Modelling Approach. Peter Lang, European University Studies, Frankfurt.

Réquillart, V., Bouamra-Mechemache, Z., Jongeneel, R. & Penel, C. (2008). Economic analysis of the effects of the expiry of the EU milk quota system. Report prepared for the European Commission, IDEI, March 2008, Toulouse.

Tonini A & Pérez Domínguez I (2008). Review of main methodological approaches quantifying the economic effects of the European milk quota scheme, internal report prepared for the European Commission (project AGRI-2007-0444), IPTS, January 2008, Seville.

Van Tongeren, F. (2002). Forward-looking analysis of reforms of the EU dairy policy. EAAE proceedings Vol 10, Xth Congress of the European Association of Agricultural Economists, Zaragoza, Spain, August 28-31, 2002.

Wieck, C., & Heckelei, T. (2007). Determinants, differentiation, and development of short-term marginal costs in dairy production: an empirical analysis for selected regions of the EU. Agricultural Economics, 36(2), 203-220.

Wieck, C., Britz, W., Pérez Domínguez, I. and Jansson, T. (2004). Auswirkungen der MTR-Beschlüsse auf die Landnutzung in den europäischen Regionen und Betrieben - Eine quantitative Analyse mit CAPRI Modellsystem. Schriften der Gesellschaft für Wirtschafts- und Sozialwissenschaften des Landbaues e.V. 39:87-96.

Witzke, H.P, & Tonini, A (2009). Milk quota expiry impacts and sensitivity analyses using the CAPSIM model, Special issue on "Milk Production and Dairy Markets", Agrarwirtschaft, 58(5/6), 228-237.