DEPARTMENT OF FOOD AND RESOURCE ECONOMICS
UNIVERSITY OF COPENHAGEN

IFRO Working Paper

Forecasting the potential of Danish biogas production – spatial representation of Markov chains

Mikkel Bojesen
Hans Skov-Petersen
Morten Gylling

2013 / 16

IFRO Working Paper 2013 / 16

Forecasting the potential of Danish biogas production– spatial representation of Markov chains

Authors: Mikkel Bojesen, Hans Skov-Petersen, Morten Gylling

www.ifro.ku.dk/english/publications/foi_series/working_papers/

Department of Food and Resource Economics (IFRO)
University of Copenhagen
Rolighedsvej 25
DK 1958 Frederiksberg DENMARK
www.ifro.ku.dk

Forecasting the potential of Danish biogas production– spatial representation of Markov chains

Bojesen, Mikkel[1]; Skov-Petersen, Hans[2]; Gylling, Morten[1]

[1]Department of Food and Resource Economics, University of Copenhagen, Rolighedsvej 25, 1958 Frederikberg C, Denmark. email: mbo@ifro.ku.dk

[2] Department of Geosciences and Natural Resource Management, University of Copenhagen, Rolighedsvej 23, 1958 Frederiksberg C, Denmark.

Abstract

This paper forecasts the spatial distribution of Danish husbandry production from 2009 until 2025. The study builds on a time series data set (1999 – 2009) for the number of livestock units (including piglets, finishers, sows, dairy cattle and young stock) measured by 1km^2 grid cells for the whole of Denmark. A Markov Chain Model (MCM) was applied to estimate transition probabilities for the future livestock intensity composition, divided into state classes. Neighbouring effects between grid cells were not included. The modelled transition probabilities fit the data very well in all state classes, except for those farms which are in the largest state class. Regional differences in development trends were documented. The strategic objective of the model is to provide data for the spatial assessment of the potential of biogas production which can form the basis for a location analysis for future biogas plants.

Key words: structural changes, Markov Chain Models, biogas, agriculture

JEL: Q47, Q16, Q21

1. Introduction

Markov chain models have been widely recognised as a means of modelling and forecasting structural changes within agricultural production (see e.g. Allen 1994, St-Pierre and Jones 2001, Skinner 1981, Matis et al 1985, Gillespie and Fulton 2001, Karantininis 2002, Jongeneel 2002, Jongeneel et al. 2005, Stokes 2006, Tonini and Jongeneel 2009, Huettel, Jongeneel 2008 and Zimmermann et al. 2009). In large parts of the Markov chain literature within agricultural production, the intention has been to investigate drivers of structural changes within specific agricultural sectors, as well as to predict structural changes within these sectors (see e.g. Gillespie and Fulton 2001, Karantininis 2002, Jongeneel 2002, Jongeneel et al. 2005, Stokes 2006, Tonini and Jongeneel 2009, Huettel and Jongeneel 2008, Zimmermann et al. 2009).

The aim of this article is to develop a model based on a Markov Chain Model (MCM) approach, as described by, e.g. Lee et al. (1970) in order to forecast the spatial distribution of structural changes in Danish husbandry production. The suggested MCM follows the model specifications as suggested throughout the Markov chain literature with respect to state class numbers, number of years of observation and level of data. Our contribution to the existing literature is that we consider space as an important dimension when applying the standard Markov Chain Model. The inclusion of space is made possible by applying extremely detailed data (both in terms of spatial/temporal resolution and with respect to the types of livestock involved) based on a census of every individual animal production unit in the country. Further standard GIS operations are applied in order to build a regional MCM involving disaggregation into five regions and five livestock types. This enables us to predict the future geographical distribution of agricultural husbandry production at the local level. Within the regions, the unit of observation is a 1x1 km square cell. We do not consider neighbour effects between cells and assume spatial autocorrelation to be ruled out. The temporal resolution is year by year throughout the period of investigation (1999- 2009). By including information on slurry production and biogas potential by livestock type, estimates of slurry production and the potential of biogas production at the local level can be obtained.

2. Background

With the passing of the EU RES directive (Directive 2009/28/EC), all EU member states are obliged to promote the use of renewable energy. For Denmark, this means that by 2020, 30% of the gross energy consumption must come from renewable energy sources. This calls for an increase in the use of forestry and agricultural by-products and materials that today are regarded as waste. According to the governmental agreement on green growth in Denmark, 40% of the slurry from livestock farms, meaning pigs and dairy cows, should be utilised for biogas production by 2020 (Grøn vækst 2010). As transportation costs in Danish biogas production constitute app. 30% of the overall production costs (Nielsen et. al 2002), it is of great importance to be able to estimate the future local biogas production potential based on livestock slurry, and therefore also structural changes.

2.1 Historic development

Since the late 1940s, structural change, defined as the gradual consolidation of agricultural units into fewer and larger units, has been a core characteristic of modern agriculture (IFRE 2009). Within the EU, structural change is especially prevalent in the north, and hence also in Denmark, with continuous development towards a more centralised and specialised agricultural sector (Rasmussen 2011).

Rasmussen (2011) found that more than 90% of Danish full time farms are still far below the technically optimal scale of production. Continued structural change is therefore to be expected, although at some stage, structural change will level off as a lack of available farmland will limit the advantages connected to technology and size. Structural change and the development towards increasing size of production units is a locally expressed phenomenon, which is influenced by regional, national and international conditions.

Over the last 11 years, the Danish pig and dairy cattle sectors have been subject to substantial structural change. Table 1 shows the change in livestock numbers for five different sectors at the national level and the development in farm numbers between 1999 and 2009.

Year	1999	2002	2009
Sows (mio)	1.139	1.203	1.167
Finishers (mio)	7.107	7.808	6.808
Piglets (mio)	n.a.	3.221	5.416
Dairy cattle (mio)	0.632	0.567	0.559
Other cattle (mio)	0.871	0.701	0.573
Number of farms	32,670	27,134	14,093

Table 1 National production within five livestock sectors in Denmark. Before 2002, it was not a requirement to report the number of piglets, which explains the n.a. in 1999. The category "Other cattle" represent other cattle at dairy cattle farms, mainly young stock. Data: DJF Geodata (2010).

When looking at table 1, we see that there was an increase of more than 1 million animals across all sectors between 2002 and 2009, driven by the increase within piglet production. Meanwhile, during the same period the total number of farms has decreased by app. 50%, suggesting a significant production capacity expansion at many farms. Even though these national figures indicate substantial structural change, it is also clear in table 1 that there are large differences between the different sectors.

The spatial distribution of dairy cattle production in 1999 and 2009, as illustrated in figure 1 below, underlines the argument stated above, that structural change occurs on a local scale and with clear regional differences. Furthermore, figure 1 shows that the area without dairy cattle production is increasing in the eastern parts of Denmark, particularly around the major Danish cities such as Copenhagen, Århus and Odense. Dairy cattle production is concentrated in the south western parts of Jutland and the western parts of Himmerland (the orange and red areas).

In this study, we quantify the shifts in production location and predict and observe a continuous westerly movement regarding the location of dairy cattle production. This tendency has been observed for many years and the clear difference between the eastern and western parts of the country coincide with substantially lower soil prices in the west than in the east.

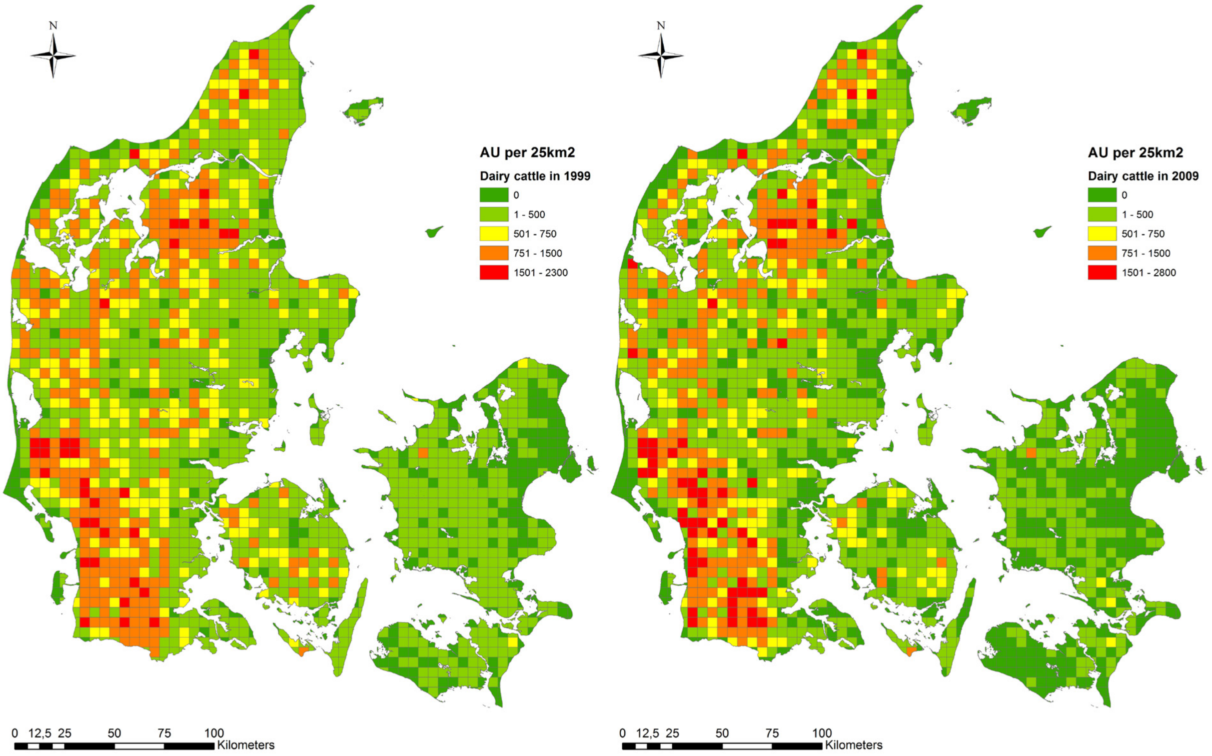

Figure 1 - Maps of the spatial distribution of structural changes within dairy cattle farming from 1999-2009 in Denmark. Data was obtained at 1km^2 but was aggregated to 5km^2 for visualization. Data source: DJF Geodata (2010).

In contrast to dairy cattle, we do not see the same profound trend around the larger cities within finisher farming (see figure 2). This is mainly due to the fact that pig production requires less land. A decrease in finisher production is found on Zealand as well as in large parts of Central Jutland. Areas with 750-1500 animal units per 25 km^2 in particular seem to be disappearing, the disappearing middle as it is termed by Huettel and Margarian (2009). However, in our data, the disappearing middle occurs at the regional level and not at the local level as argued by Huettel and Margarian (2009).

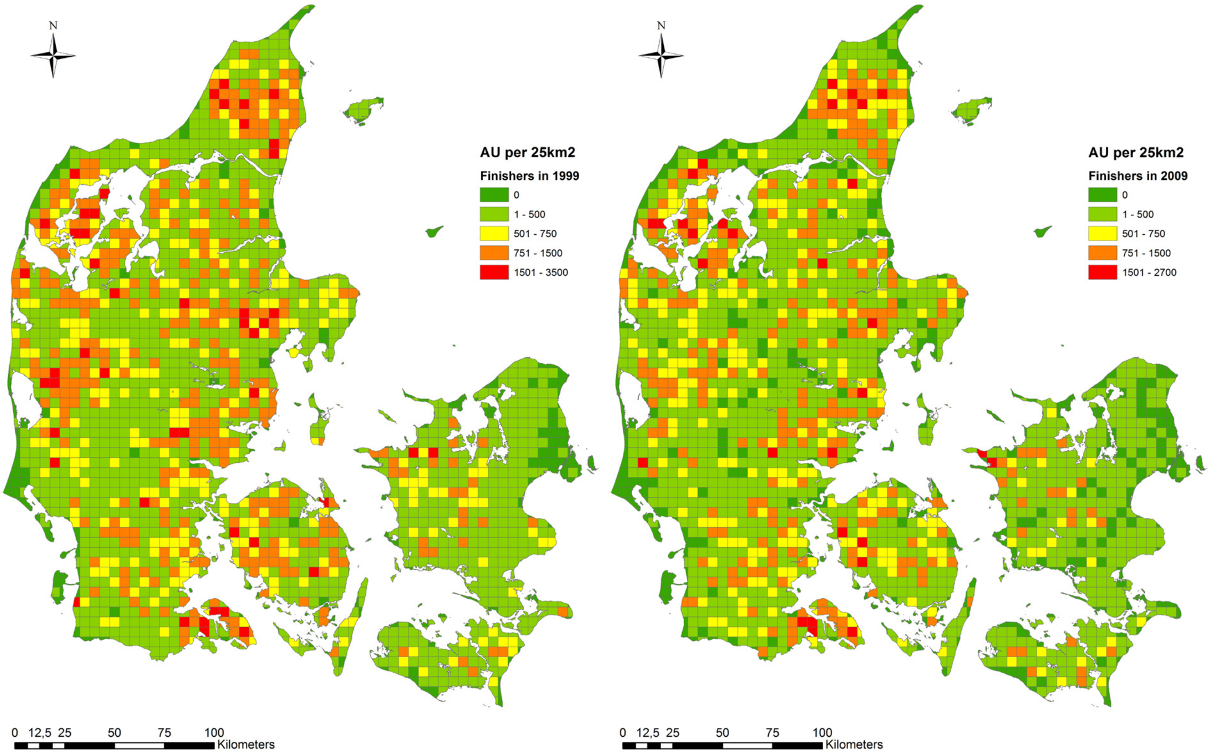

Figure 2 - Maps of the spatial distribution of structural changes within finisher farming from 1999-2009. Data was obtained at 1km^2 but was aggregated to 5km^2 for visualization. Data source: DJF Geodata (2010).

Figure 1 and 2 above illustrate structural change within dairy cattle and finisher production, two major Danish agricultural sectors, from 1999 to 2009. The two sectors exhibit different spatial development patterns both at the national and regional level. A number of national and international determined drivers, which are difficult to predict (e.g. future policy measures, legislative restrictions and market developments), are expected to influence these regional trends and by that the continuous structural changes.

3. Materials and methods

Fitting a Markov chain (multi-state) model to panel data generally relies on the Markov assumption that future evolution only depends on the current state. Before this can be attempted, decisions must be made concerning the data types, assumptions regarding stationary or non-stationary transition probabilities, the length of the time series to include in the analysis, the number of transition states and the explanatory variables. These model considerations are often determined by data availability, but they nevertheless dictate crucial model assumptions.

The general Markov Chain model, as described by, e.g. Lee et al. (1970) is given by:

$$n_{j(t)} = \sum_{i=1}^{J} n_{i\,(t-1)}\, P_{ij} \qquad (1)$$

Where:

$n_{j(t)}$ is a state vector counting the number of individuals, i.e. geocells in state class j (with j= {1…J}) at time t

$n_{i(t-1)}$ is a state vector counting the number of individuals, i.e. geocells in state class i (with i= {1…I}) at time t-1

P_{ij} is the probability of moving from state i to state j in one time period

The assumption of stationary transition probabilities is important, as it implies that the probability of moving from one state to another over one time period is constant between time periods. If this assumption holds, one can apply the same transition probability matrix several times in order to perform a forecast. In contrast, if non-stationary transition probabilities apply, the past does not reflect the future, and hence the value of historic data is diminished.

In a Markov Chain study concerned with farm sizes in the western part of Germany, Huettel and Margarian (2009) argue that it is very unlikely that the transition probabilities are stationary and that they expect them to vary over time. Nevertheless, Piet (2008) reported that, in more than 50% of 26 agricultural Markov Chain studies, stationary transition probabilities were applied. We apply stationary transition probabilities in this study and investigate the validity of this assumption.

Zimmermann et al. (2009) found that, in 66% of 29 agricultural studies, macro data were used, meaning aggregated data at the farm level and not at the individual animal level. They also found that the mean number of years included in the time series studied was 16, while the average number of included state classes was 6.

This study builds upon macro data, i.e. aggregated data in which it is not possible to follow individual animals or farms. Data is based on farmers' reports to the Danish central livestock register (CHR) which were obtained on a 1km^2 grid cell scale (DJF Geodata, 2010). The data contain information on 5 livestock types (sows, piglets, finishers, dairy cattle and “other cattle”, i.e. young stock and cows for breeding) and cover a time span of 11 years (1999-2009).

3.1MAUP

A fundamental concept in geography is that everything is related to everything else, but also that nearby entities often share more similarities than entities which are far apart. This idea is known as Tobler’s first law of geography (Miller 2004). In this study, the 1 km^2 data cells have been divided into 6 strata according to geo-regions based on prevailing soil types, as described by (Greve et al., 2007), see figure 3. The soil type is regarded as a proxy for soil prices, which is believed to be an important driver of structural change (Happe et al. 2008). The North of Zealand (the north eastern region) is not included in the forecast modelling because animal production is too small to estimate transition probabilities.

Since it first was documented by Gehlke and Biehl (1934), the modifiable areal unit problem (MAUP) has been an important source of uncertainty within spatial analysis. The problem has a scaling and zoning effect and it affects results when point-based measures of spatial phenomena, e.g. population density, are aggregated into districts. The resulting summary values, i.e. totals, rates and proportions, are influenced by the choice of district boundaries. The problem was first solved by Openshaw (1984). Since this study applies the smallest possible scale, i.e. 1 km^2, and applies natural boundaries based on soil types, we consider the MAUP effects to be ruled out and they therefore do not play any role in the forecasting analysis.

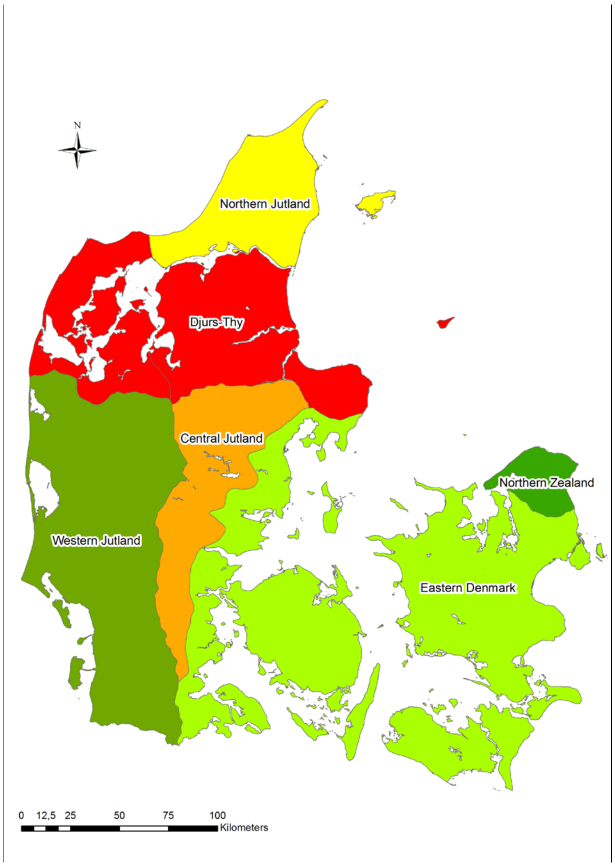

Figure 3 - Map showing soil type based geo-regions in Denmark, based on Greve et al. (2007)

According to the EU Nitrate Directive (91/676/EC), the application of N from slurry must not exceed 170 kg per hectare, the so-called harmony demand area. This is in order to prevent the leeching of nutrients to the aquatic environment. The EU Nitrate Directive stipulates how much land every farmer must have access to in order to safely dispose of the slurry from his animal production. Up until 2010, Danish agricultural regulations stipulated that, on top of the “harmony demand,” a set of ownership demands also needed to be fulfilled, which stated that given the size of your production, a certain percentage of the “harmony demand areas” should be owned by the producer (Retsinformation 2011). The rest of the “harmony demand areas” could be leased or slurry contracts with other farmers could be made in order to dispose surplus slurry. The ownership demand regime no longer applies. Since the data for the Markov Chain Models concerning the five Danish agricultural livestock branches, stems from a time period when the ownership demands were in action, the state class intervals will reflect these ownership demand threshold values and secure sufficient data in each state class interval to achieve solid estimates. Dairy cattle and finisher production is the main branch of production at many farms, whereas as sows, piglet and young stock (other cattle) are produced simultaneously or alongside with dairy cattle and finishers as subordinate production branches. Consequently the number of AU's within dairy cattle and finisher production is higher than for the other types of livestock, which is reflected in the state class intervals in table 2.

State class	Animal units (AU) Dairy cattle and Finishers	Animal units (AU) Sows, Piglets and Other cattle	Verbal categories
1	0	0	No production
2	0-75	0-25	Hobby Farms
3	75-120	25-50	Small Farms
4	120-250	50-75	Large Farms
5	250-350	75-120	Very large Farms
6	>350	>120	Mega-farms

Table 2 - Overview of which state classes contain which number of animal units and their verbal categorization

Based on the above described characteristics, the Markov Chain model looks as follows:

$$n_{j(t),k,m} = \sum_{i=1}^{J} n_{i\,(t-1),k,m}\, P_{ij,k,m} \qquad (2)$$

Where:

$n_{j(t)}$	is a state vector counting the number of geocells in state class j (1,2,...,6) at time t, where class 1 is an absorbing state
$n_{i(t-1)}$	is a state vector counting the number of geocells in state class i(1,2,…,6) at time t-1
k	is georegion (1,2,..,4)
m	is livestock type (1,2,…, 5)
P_{ij}	is the probability of moving from state i to state j in one time period

$$P = \begin{vmatrix} p_{11} & p_{12} \cdots & p_{1j} \\ p_{21} & p_{22} \cdots & p_{2j} \\ \cdots & \cdots & \cdots \\ p_{i1} & p_{i2} \cdots & p_{ij} \end{vmatrix} \qquad (3)$$

Panel data comprising a time series of 11 years is applied and based on annual reports of livestock numbers. This gives us a discrete time Markov Chain model.

The model was built in the RStudio statistical software, version 0.95.262 by utilizing the add-on package msm 1.0.1 (Jackson 2011a). The msm package can be applied to fit continuous-time Markov models, where transitions can occur at any time by maximum likelihood, and discrete time models, where transitions are known in advance to only occur at multiples of a certain time unit. The discrete model is purely governed by the probability distributions of the state at the next time point, conditionally on the state at the current time. These transition probabilities are fitted in msm, assuming that a continuous time process underlies the data.

Continuous models are defined by intensities in the form of an RxR matrix Q, where the rows sum to zero. The movement on the discrete state space 1, …. , R is governed by transition intensities qrs(t, z(t)): r; s = 1, …, R. These may depend on time t, or, more generally, also on a set of individual-level or time-dependent explanatory variables z(t). The intensity represents the instantaneous risk of moving from state r to state s ≠ r. The likelihood for this discrete model, used in msm, is calculated from the transition probability matrix P(u, t + u). The (r, s) entry of P(u, t + u), p_{rs}(u, t + u), is the probability of being in state s at time t + u, given the state at time u is r. P(u, t + u) is calculated in terms of Q using the Kolmogorov differential equations (see, e.g. Cox and Miller 1965). If the transition intensity matrix Q is constant over the interval (u, t + u), as in a time homogeneous process, then P(u, t + u) = P(t) and the equations are solved by the matrix exponential of Q scaled by the time interval,

$$P(t) = Exp\ (tQ) \qquad (4)$$

For a discussion of the calibration of the matrix exponential, see e.g. Moler and van Loan (2003).

A Pearson’s Goodness of fit test was conducted, which compares the observed number of individuals occupying each state with forecasts from the fitted model.

The state vector $n_{i(t-1)}$ is calculated based on the observed mean value of each state class 1 to 6 in 2009. Whether these mean values change significantly during the observation period from 1999 to 2009 was investigated. If changes are significant, the change in mean value for the state class in question is forecasted assuming a linear trend with time as the sole explaining variable. Based on these models of state class mean values multiplied by the transition probability matrix, forecasts of the 2009 animal unit numbers are performed.

4. Results and discussion

In this section, the results for finishers in Western Jutland (WJ) and Eastern Denmark (ED) are presented in order to demonstrate the differences and similarities between regions. The final biogas production potential map includes all livestock types for the sake of completeness.

4.1 Estimation of transition probability matrices

In table 3, the transition probability matrix with corresponding confidence intervals for finishers in ED shows the probability of staying in the same state from one time period to the following. We see that the probability of staying in the same state is higher than changing from one state to another for state classes 1,2 and 4, whereas for state classes 3, 5 and 6 the highest probability is a shift to state class 2 or 4 respectively. This is seen from the numbers in bold as they constitute the matrix diagonal. The sum of each column provides an estimate of the general state growth, i.e. for values above 1, there is an increasing number of geocells in that particular state class. For finishers in ED, it is clear that state classes 1, 2 and 4 are expected to increase, i.e. an increase in geocells with no production and an increase in geocells with an annual production between 0-75AU and 120-250AU per year. For the two upper state classes 5 and 6, the column sums are quite low, implying that the number of geocells in these state classes is decreasing rapidly.

Transition probability matrix for state classes with confidence intervals for finishers in Eastern Denmark

	1		2		3	
1	**0.518**	(0.472,0.541)	0.357	(0.334,0.378)	0.052	(0.047,0.061)
2	0.280	(0.260,0.296)	**0.528**	(0.509,0.545)	0.080	(0.073,0.088)
3	0.189	(0.174,0.204)	0.435	(0.417,0.450)	**0.148**	(0.137,0.160)
4	0.125	(0.113,0.139)	0.314	(0.299,0.329)	0.164	(0.153,0.174)
5	0.106	(0.093,0.132)	0.266	(0.250,0.284)	0.155	(0.144,0.165)
6	0.113	(0.096,0.147)	0.261	(0.245,0.284)	0.147	(0.136,0.157)
sum	1.330		2.161		0.747	

	4		5		6	
1	0.060	(0.053,0.080)	0.008	(0.006,0.022)	0.006	(0.005,0.011)
2	0.090	(0.082,0.101)	0.012	(0.011,0.018)	0.010	(0.008,0.013)
3	0.184	(0.170,0.200)	0.025	(0.022,0.031)	0.019	(0.015,0.024)
4	**0.318**	(0.298,0.337)	0.047	(0.040,0.055)	0.033	(0.027,0.041)
5	0.346	(0.323,0.364)	**0.073**	(0.061,0.083)	0.055	(0.043,0.066)
6	0.322	(0.295,0.340)	0.081	(0.069,0.092)	**0.076**	(0.060,0.092)
sum	1.319		0.245		0.198	

Table 3 Maximum likelihood estimation of the transition probability matrix for finishers - Eastern Denmark, based on data from 1999-2009. Source: Own calculations.

Transition probability matrix for state classes with confidence intervals for finishers in Western Jutland

	1		2		3	
1	**0.771**	(0.749,0.790)	0.153	(0.138,0.168)	0.028	(0.024,0.033)
2	0.467	(0.446,0.490)	**0.345**	(0.324,0.363)	0.072	(0.065,0.080)
3	0.312	(0.291,0.341)	0.313	(0.296,0.328)	**0.140**	(0.124,0.152)
4	0.223	(0.204,0.256)	0.251	(0.235,0.266)	0.155	(0.142,0.166)
5	0.210	(0.190,0.271)	0.245	(0.225,0.265)	0.142	(0.127,0.153)
6	0.188	(0.168,0.274)	0.231	(0.213,0.251)	0.143	(0.126,0.154)
sum	2.172		1.537		0.680	

	4		5		6	
1	0.037	(0.031,0.044)	0.006	(0.005,0.010)	0.005	(0.004,0.007)
2	0.093	(0.083,0.104)	0.014	(0.012,0.018)	0.009	(0.007,0.012)
3	0.188	(0.168,0.207)	0.029	(0.023,0.036)	0.018	(0.014,0.023)
4	**0.295**	(0.267,0.318)	0.048	(0.038,0.059)	0.028	(0.021,0.036)
5	0.288	(0.253,0.311)	**0.078**	(0.058,0.095)	0.036	(0.026,0.048)
6	0.292	(0.250,0.312)	0.094	(0.071,0.110)	**0.052**	(0.036,0.067)
sum	1.193		0.270		0.149	

Table 4 – Maximum likelihood estimate of the transition probability matrix for finishers – Western Jutland, based on data from 1999-2009. Source: Own calculations.

The transition probability matrix for Finishers in WJ, Table 4, illustrates many of the same general trends as in ED, implying an increase in state classes 1, 2 and 4, but movements into state class 1, i.e. the rate that farms are going out of production is somewhat higher in WJ than in ED, which indicates that structural development is occurring at a faster pace in the western parts than in the eastern parts of the country.

Applying the same methodology as described above to the dairy cattle sectors, we find that, in the fitted transition probability matrices for ED, there is an increase in state classes 1 and 4 implying a rapid decrease in production intensity, while in WJ, there is an increase in state classes 1, 2, 4 and 6, while the summed ingrowth into state classes 3 and 5 is just below 1, which indicates a slow decrease in these two state classes. This also tells us that the number of mega-farms (>350 AU) and large farms (120-250 AU), state classes 4 and 6, will increase, although the number of very large farms (250-350 AU) will decrease. This could be explained by the hypothesis that once the decision to expand beyond full time farming is made, one aims to become a mega-farm.

4.2 Model fit

A Pearson's chi squared test is carried out to test the assumption of stationary transition probabilities. Table 5 and 6 below show that 90% of the model variation is found within the transitions into state class 6. For finishers in WJ in state class 6, the summed absolute deviation is 210.4 out of a total model deviation of 235.18 (see table 5). Similarly, in ED, the summed absolute deviation of state class 6 is 336.8 out of a total model deviation of 370 (see table 7). In addition, more than 75% of the model variation is found within the number of geocells staying in state class 6 between two time periods, both in WJ and in ED, which is caused by few observations. Therefore, the vast majority of the model only exhibits minor deviations from the observed data and the observed deviations imply under prediction which leads to conservative estimates regarding future husbandry production.

State class transition	**1**	**2**	**3**	**4**	**5**	**6**
1	0.0	0.8	0.0	0.0	0.0	-3.6
2	2.1	-0.4	2.8	0.3	0.0	-5.5
3	-0.3	4.4	-1.1	3.4	0.0	-5.0
4	-0.1	-0.1	3.8	-0.3	0.7	-12.3
5	-0.1	0.1	0.0	1.3	-0.3	-4.0
6	0.0	0.0	0.1	0.1	1.9	-180.0
Total	2.6	5.8	7.8	5.4	2.9	210.4

Table 5 - Pearson's chi squared model variation and total absolute variation, finishers Western Jutland. Source: Own calculations.

State class transition	1	2	3	4	5	6
1	-0.1	1.3	-0.1	0.1	-0.2	-1.8
2	1.8	-0.2	2.3	0.0	-0.1	-8.1
3	-0.1	4.5	-1.3	4.2	-0.2	-4.3
4	-0.0	-0.5	5.6	-0.5	1.9	-15.7
5	0.0	-0.2	-0.3	3.5	-0.9	-11.0
6	0.1	-0.1	0.3	0.1	3.2	-296.0
Total	2.0	6.7	9.8	8.3	6.4	336.8

Table 6 - Pearson's chi squared model variation and total absolute variation, finishers Eastern Denmark. Source: Own calculations.

When we compare the detailed findings concerning the model deviations in table 5 and 6 with the general model fit in table 7, we see that, despite the good model fit in all state classes except 6-6, the general model fit statistics suggest an overall rejection of the hypothesis that the fitted and observed numbers are from the same distribution.

Region	Total model deviation	DF lower	P-lower	DF-upper	P-upper
Eastern Denmark	370.0	0	0	30	0
Western Jutland	235.2	0	0	30	0

Table 7 - Pearson's chi squared statistics, finishers in Eastern Denmark and Western Jutland. Own calculations.

A similar result to that reported above was found for the other animal types considered in this study. The generally small discrepancies found for finishers in both WJ and ED provide a good argument for assuming time independent, i.e. stationary, transition probabilities. With regards to the model deviations in state class 6-6, this may indicate that the transition probabilities vary over time (Jackson 2011a). Another cause may be a failure of the Markov assumption, i.e. the transition intensities may depend on the time spent in the current state (a semi-Markov process) or other characteristics of the process history. Accounting for the process history is difficult as the process is only observed through a series of snapshots (Jackson 2011b).

4.3 Forecasting of animal units

We now consider the estimated transition probability matrices for all five livestock types in all five regions. For each of these, we calculate a vector of the mean number of animal units in each state class. We test whether or not there is significant development in the state class mean from 1999-2009, and based on this, we assume a linear development in the state class means and multiply this vector with the forecast transition probability matrix in order to obtain the number of animal units

(AU) in year 2015, 2020 and 2025, taking 2009 as the state class base year. Tables 8 and 9 show the forecasted numbers for each state class with regards to finishers in ED and WJ.

State class in 2009	AU in 2015	AU in 2020	AU in 2025
1	13	22	29
2	35	40	43
3	85	78	72
4	146	125	107
5	207	160	128
6	263	180	136

Table 8 - Forecast of number of animal units within finisher production in Eastern Denmark. Source: Own calculations.

State class in 2009	AU in 2015	AU in 2020	AU in 2025
1	8	13	17
2	35	39	40
3	88	79	71
4	143	120	100
5	188	140	110
6	256	168	124

Table 9 - Forecast of number of animal units within finisher production in Western Jutland. Source: Own calculations.

Because we apply state class mean values for the state vectors, these forecasted numbers of animal units in each state class do not capture the extreme values, and consequently, the forecasted figures for the number of animal units are rather conservative. In the further application, the obtained figures should be seen as minimum values of future animal husbandry production in the five regions.

4.4 Spatial representation

Table 8 and 9 are joined with a table in ArcMAP 10.0 with geocell number and the state class in 2009. By applying table 8 and 9 to every geocell in the corresponding region and adding information regarding the livestock type’s specific slurry production and biogas potential for these slurry types (see table 10), maps were produced with the spatial location of biogas potential in 2009, 2015,2020 and 2025 (see figure 4).

Livestock type	Slurry production, ton per AU	m3 methane per ton
Sows	24	6
Finishers	17.5	11
Piglets	24	6
Dairy cattle	19.89	13
Other cattle	25.77	13

Table 10 – Slurry production and methane potential for different livestock types. Data: Hjort-Gregersen 2011.

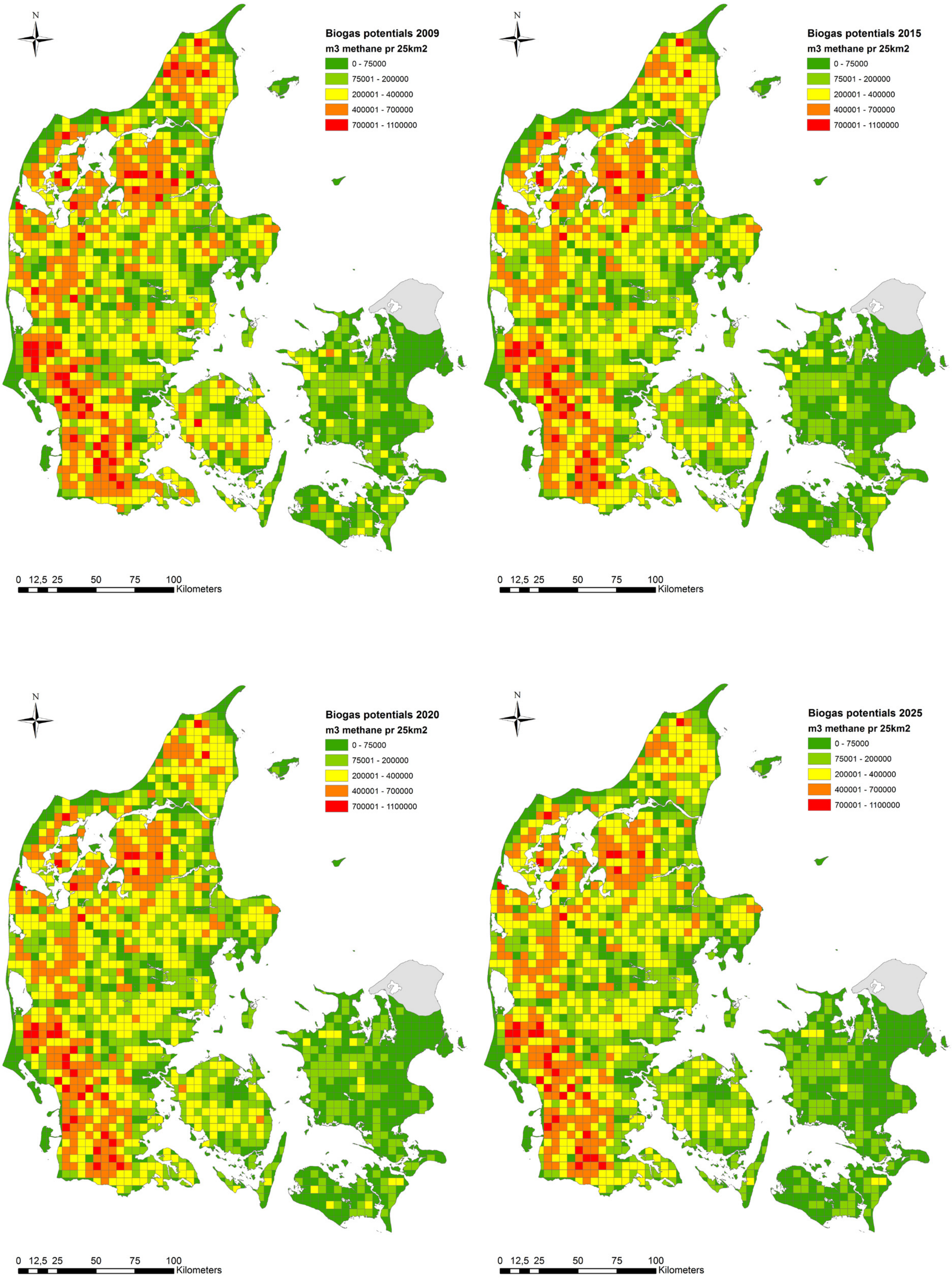

Figure 4 - Spatial representation of forecasted biogas potential 2009, 2015, 2020 and 2025 in 5 Danish georegions. Numbers were modeled at a 1km2 scale, but were aggregated to 5km2 for visualization. Data source: DJF Geodata (2010)

Year	2009	2015	2020	2025
Number of 25km^2 Geocells :	2242	2242	2242	2242
Minimum potential (mill. m^{3}methane):	0	0	0	0
Maximum potential (mill. m^3 methane):	1.062	0.987	0.983	0.985
Sum of potential (mill. m^3 methane):	424.362	387.129	380.159	370.897
Mean potential (mill. m^3 methane):	0.189	0.173	0.17	0.165
Standard Deviation (mill.m^3 methane):	0.204	0.190	0.187	0.184

Table 11 - Summary statistics of production potential from 2009-2025 measured in m3 methane. Own calculations.

From the maps in figure 4, which illustrate biogas production potential as a function of structural changes until 2025, it can be seen that, with a 25km^2 aggregation, the spatial location of production potential is fairly stable over time. Table 11 presents the overall figures for biogas production potential and one can see an estimated decrease of app. 10% in total production between 2009 and 2015, but very little difference between 2015 and 2025. These findings show that, despite the conservative estimates obtained by using the Markov Chain model, state class mean values and weighted averages for the expected number of animal units in future state classes, the biogas production potential of Danish livestock production does not seem to change dramatically over time, neither in quantity nor in location. One could argue that a forecast until 2025 may be on the edge of what the data can support, due to the converging effects of the Markov Chain model in combination with the additional effects described above.

5. Conclusion

5.1 Academic importance and further use of the model

In this study, we fill the gap in the agricultural Markov Chain literature by building a spatially disaggregated Markov Chain Model and applying it spatially. By following the application described in this paper, we obtain estimates of the geographical location of Danish livestock production until 2025. By adding information regarding production and the methane content of livestock slurry, we end up with estimates for biogas production potential at a 1km^2 scale, which we then aggregate to a 5km^2 scale for better visualization.

Due to little model variation concerning all state classes (except state class 6-6, i.e. mega-farms), the assumption of stationary transition probabilities is well supported. This finding challenges the findings of other recent studies, which advocate time-dependant transition probabilities.

One of the major differences between this study and other recent studies is the coherent period of observed data. This study builds on ten years of annually observed data and the proposed model exhibits high stability regarding predictions about the structural results of the economic decisions within agricultural production. The implication of stable transition probabilities is that the dynamics of the structural changes within the observed timeframe capture the dynamics of structural changes in the years to come.

The results of this paper can serve as inputs to a number of analyses concerning Danish agriculture-based production, including decision support systems concerning the future development of the Danish biogas sector.

5.2 Limitations and restrictions

The implication of the model deviations regarding the farms which remained mega-farms from one time period to another is that estimates for livestock, and consequently slurry production, should be considered conservative when used in any analysis of the future available resources for the evolving Danish biogas sector.

5.2 Future research

Due to the limitations of the proposed model, i.e. the model variation concerning farms with more than 350 animal units, future research should focus on how to capture these extremes. One way of doing this might be to consider the neighbouring effects between cells. The parameter estimation of the MCM approach suggested here could be extended to include, not only the value of individual cells at preceding times, but also the values of neighboring or nearby cells (considering given distance-decay functions). As suggested by Hansen (2008), such parameters could be applied to the modeling/simulation of future situations by means of Cellular Automata models (CA) where the value of a cell at time t, whether it be addressed discretely or stochastically, is based on the value of the cell, and the cells around it, at time t-1. By adopting such an approach, one would be able to use the model from this study and include decision variables in order to enhance the precision of the transition probabilities regarding mega-farms.

Acknowledgements

The authors would like to thank Inge T. Kristensen from the University of Aarhus, Foulum research station, for her assistance with acquiring and preparing the unique dataset for this analysis without which this study would not have been possible.

References

1. Allen, P. G. (1994). Economic forecasting in agriculture. International Journal of Forecasting, 10 (1994), 81-135.
2. Balmann, A. (1997). Farm-Based modelling of regional structural change: A cellular automata approach. European Review of Agricultural Economics, 24 (1997), 85- 108.
3. Bonabeau, E. (2002). Agent-based modeling: Methods and techniques for simulating human systems. Proceedings of the National Academy of Sciences of the United States of America (PNAS), May 14, 2002, vol. 99, suppl. 3, 7280–7287.
4. Cox D.R. and Miller, H.D. (1965). The Theory of Stochastic Processes. Chapman and Hall, London.
5. Directive 2009/28/EC. (2009). Available online: [http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2009:140:0016:0062:en:PDF], cited: 25.11.2012.
6. Directive 91/676/EC. (1991). Available online: [http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:1991:375:0001:0008:DA:PDF], cited: 08.10.2013
7. DJF Geodata (2010). Generelle Landbrugs Register, Centrale Husdyr Register - Aarhus University, Department of Agroecology, PO Box 50, Blichers Allé, DK-8830 Tjele. Data achieved 4. November 2010.
8. Gehlke, C. and Biehl, H. (1934). Certain effects of grouping upon the size of the correlation

coefficient in census tract material. Journal of the American Statistical Association Supplement, 29, 169–170.

9. Gillespie, J. and Fulton, J. (2001). Entry, Exit and Changes in the Size of Hog Production Firms in the U.S.: A Markov Chain Analysis. Agribusiness: An International Journal, 17 (2001), 557-570.
10. Greve, M.H. Greve, M.B, Bøcher, P.K, Balstrøm. T, Madsen H. B, Krogh, L. (2007). Generating a Danish raster-based topsoil property map combining choropleth maps and point information. Geografisk Tidskrift 107 (2).
11. Grøn vækst (2010). Ministeriet for Fødevarer, Landbrug og Fiskeri, available online: [http://www.fvm.dk/Admin/Public/DWSDownload.aspx?File=%2fFiles%2fFiler%2fLandbrug%2fAftaletekstlandbrug.pdf], cited 25.11.2012.
12. Hansen, H.S. (2008). Quantifying and Analysing Neighbourhood Characteristics Supporting Urban Land-use Modelling. In Bernard, L., Friis-Christensen, A. & Pundt, H. (Eds.) The European Information Society, Lecture Notes in Geoinformation and cartography pp. 283-299. Springer-Verlag Berlin Heidelberg, 2008.
13. Happe, K. Balmann, A. Kellermann, K. Sahrbacher, C. (2008). Does structure matter? The impact if switching the agricultural policy regime on farm structures. Journal of Economic behaviour and organization, 67, 431-444.
14. Happe, K., Hutchings, N. J., Dalgaard, T., Kellerman, K. (2011). Modelling the interactions between regional farming structure, nitrogen losses and environmental regulation. Agricultural Systems, 104 (2011) 281–291.
15. Hjort-Gregersen (2011). Økonomiske betragtninger ved planlægning, etablering og drift af biogasanlæg. Presentation held at Bioenergy seminars at Sørup Herregård, June 9th 2011, Ringsted, Denmark. Khg@agrotech.dk
16. Huettel, S. and Jongeneel, R., (2008). Structural change in the dairy sectors of Germany and the Netherlands—a Markov chain analysis. Paper presented at the 12th Congress of the European Association of Agricultural Economists—EAAE 2008, Ghent, Belgium.
17. Huettel, S. and Margarian, M. (2009). Structural change in the West German agricultural sector. Agricultural Economics, vol. 40 (2009) supplement, pp. 759–772.
18. IFRE (2009). Landbrugets økonomi 2009, ISSN 1902-0813 (online, Landbrugets økonomi). Available online: [http://www.foi.life.ku.dk/Publikationer/FOI_serier/~/media/Foi/docs/Publikationer/Rapporter/Landbrugets%20%C3%98konomi/2009.pdf.ashx].
19. Jackson, C. (2011a). R statistical package, msm, Multi-state Markov and hidden Markov

models in continuous time, version: 1.0.1, Date: 2011-05-26, chris.jackson@mrc-bsu.cam.ac.uk

20. Jackson, C. (2011b). Multi-State Models for Panel Data: The msm Package for R. Journal of Statistical Software, January 2011, Volume 38, Issue 8.
21. Jongeneel, R. (2002). An Analysis of the Impact of Alternative EU Dairy Policies on the Size Distribution of Dutch Dairy Farms: An Information Based Approach to the Non-Stationary Markov Chain Model. Paper presented at the 10th EAAE Congress, August 28–31, Zaragoza, Spain.
22. Jongeneel, R., Longworth, N., Huettel, S. (2005). Dairy farm size distribution in east and west: evolution and sensitivity to structural and policy variables: case-studies of the Netherlands, Germany, Poland and Hungary. Paper presented at the 11th EAAE Congress, August 24–27, Copenhagen, Denmark.
23. Karantininis, K. (2002). Information-based estimators for the non-stationary transition probability matrix: an application to the Danish pork industry, Journal of Econometrics, 107, 275–90.
24. Lee, T., Judge, G., and Zellner, A. (1970). Estimating the Parameters of the Markov Probability model from Aggregate Time Series Data. Amsterdam: North-Holland.
25. Margarian, A. (2009). Farmers' conservative behaviour and adapted strategies: Economic foundation and implications for structural change. Paper submitted for the EAEPE Conference 2009 in Amarsterdam, Research area [M] Social and cultural economics.
26. Matis, J. H., Saito, T., Grant, W. E., Iwig, W. C., Ritchie, J. T. (1985). A markov chain approach to crop yield forecasting. Agricultural Systems, 18 (1985), 171-187.
27. Miller, H. J., (2004). Tobler's first law and spatial analysis. Annals of the Association of American Geographers, 94(2), 2004, pp. 284–289.
28. Moler C, van Loan C (2003). Nineteen Dubious Ways to Compute the Exponential of a Matrix, Twenty-Five Years Later. SIAM Review, 45(1), 3-49.
29. Nielsen, L. H.; Hjort-Gregersen, K.; Thygesen, P.; Christensen, J. (2002). Socio-economic analysis of centralised Biogas Plants - with technical and corporate economic analysis. Institute of Food and Resource Economics. Report 136, Copenhagen 2002.
30. Openshaw, S. (1984). The Modifiable Areal Unit Problem. Norwich: Geo Books. ISBN 0-86094-134-5.
31. Piet, L. (2008). The evolution of farm size distribution: revisiting the Markov chain model. Paper sumitted for the 12th Congress of the European Association of Agricultural

Economists – EAAE 2008, Ghent Belgium.

32. Rasmussen, S. (2011). Estimating the technical optimal scale of production in Danish agriculture. Food Economics - Acta Agriculturae Scandinavica, Section C, April 2011, vol 8, Issue 1, 1-19.
33. Retsinformation. (2011). BEK nr 1152 af 23/11/2006 - Bekendtgørelse om husdyrhold og arealkrav mv. Available online: [https://www.retsinformation.dk/Forms/R0710.aspx?id=8523], cited 09.09.2011.
34. Shastri, Y.; Rodríguez, L.; Hansen, A.; Ting, K. C. (2011). Agent-based Analysis of biomass Feedstock production dynamics. Bioenergy Resource, 4, 258-275.
35. Stokes, J. R. (2006). Entry, exit, and structural change in Pennsylvania's dairy sector, Agricultural and Resource Economics Review, 35, 357–73.
36. Skinner, S. P. (1981). A markov analysis of structural change and output prediction in the New England egg industry. Life sciences and agriculture experiment station, University of Maineat Orono, Technical Bulletin 102 July 1981.
37. St-Pierre, N. R. and Jones, L. R. (2001). Forecasting herd structure and milk production for production risk management. Journal of Dairy Science, 84,1805–1813
38. Tonini, A. and Jongeneel, R. (2009). The distribution of dairy farm size in Poland: a markov approach based on information theory. Applied Economics, 41 (1), 55 - 69
39. Zimmermann et al 2009. Modelling farm structural change for integrated ex-ante assessment: review of methods and determinants. Environmental science and policy, 12(2009), 601-618.