A model for prediction of spatial farm structure

Øyvind Hoveid and Grete Stokstad
oyvind.hoveid@nilf.no

Paper prepared for presentation at the EAAE 2011 Congress
Change and Uncertainty
Challenges for Agriculture,
Food and Natural Resources

August 30 to September 2, 2011
ETH Zurich, Zurich, Switzerland

A model for prediction of spatial farm structure

Øyvind Hoveid*and Grete Stokstad†
oyvind.hoveid@nilf.no

February 17, 2011

Abstract

Spatial micro structure and its change over time is recorded for Norwegian farm firms. Relative strong correlations between geographically close neighbors are expected, either because growing farms swallow the smaller ones, or because they are affected by some spatially related unobserved factors. Strong correlations over time are also expected because of prevalent family farming.

The paper proposes a state-of-the-art Markov chain model in order to predict the spatial and temporal micro structure taking account of both non-stationarity and spatio/temporal correlations by means of techniques from non-linear state space modeling and Gaussian Markov random fields.

The model and the complete data set is then a device with which one can investigate the consequences of ignoring spatial and/or temporal correlations, both with complete data and with more sparsely sampled data, like FADN panels or USDA's repeated cross-sections (ARMS).

1 Introduction

For Norwegian agriculture are almost complete single farm geo-referenced time series of crop allocation and livestock recorded. These are gathered from the application of direct support to Norwegian farms. Only farms which are too small to qualify for support are exempt from this register. An illustration of farm sizes from the municipality Trøgstad in 1999 and 2009 is given in figure 1. The distribution of crop allocation and livestock heads are basic aspects of farm structure, and these data then constitute an exceptional frame for prediction of farm structure over time and space. The relevance of such predictions stems from micro-based policy analysis which captures the policy effects on typical farms, but ignores the policy effects on farm development and distribution.

In the literature has Markov chains played a dominating role in the modeling of structural change of firms. Zimmermann, Heckelei, and Domínguez (2009) give a review with regard to farms. In this tradition (1) the space of farm observations is divided in K different subsets according to farm size and type. Each set is considered a

*Norwegian Ag. Econ. Research Institute
†The Norwegian Forest and Landscape Institute

state. (2) $K \times K$-transition matrices, $\mathbf{R}(t, Z)$, between states are modeled as functions of time t and other co-variates Z. [1] And (3) alternative equilibrium structures are found as discrete distributions $p^*(t, Z)$ over states, solving $(I - \mathbf{R}(t, Z))p^*(t, Z) = 0$.

Markov equilibrium structures $p^*(t, Z)$ are steady states of the stochastic process given by the transition matrix $\mathbf{R}(t, Z)$. It is important to observe — and often misunderstood — that the $p^*(t, Z)$ is not a prediction for the future. Since Markov chains are stationary stochastic processes, this might be a valid interpretation only in stationary economies. Rather, it should be interpreted as a counterfactual prediction for time t. From this insight, the usual estimation of the transition matrix from firm changes between two points in time, is dubious in non-stationary contexts.

Several additional short-comings of the Markov chain approach, relevant even in stationary economies, can easily be pointed out: (A) Computational issues force K to be kept relatively low compared to the number of observations. There is a loss of information by the reduction of an almost continuous and multi-dimensional observation of a farm to a K-nomial state. It is the information on heterogeneity within each class which is lost.[2] (B) Even if states were relatively homogeneous, one should not expect transitions to obey the Markov property. Actually, the theory of survival analysis and competing risks — which is the backdrop of such transition data — points out the time spent in each state as a decisive determinant for transitions. There is thus inter-temporal correlation between probabilities that a farm is in a certain state at two different points in time. Unless variables carrying information from other points in time are taken into account, there is little scope for Markov models in this context. (C) Probabilities may also be correlated in space. Farms with favorable natural and social conditions for being active, may also have neighbors with similar conditions. Probabilities for being active are then positively correlated. On the other hand, the expansion of one farm will often come at the expense of neighbor farms. Probabilities for being active are then negatively correlated.

Apart from these short-comings of the Markov chain approach to farm structure, predictions of the equilibrium farm structure at t, $\widehat{p}^*(t, Z)$, as a result of a path of exogenous policy variables, Z, is policy relevant in several situations. This is further discussed in section 2. Moreover, Markov chain methods are indeed relevant for the estimation of such predictions. First, a Markov model of hidden continuous states should be utilized to model the continuous aspects of firm data. The well-known technique of state space modeling presented in section 3 takes care of correlations over time by means of unobserved Gaussian state variables. The related Gaussian Markov random fields are convenient for modeling of spatial correlations. This technique is easily incorporated in the state space frame section 4. For the distinction between active farms and inactive farms, discrete states are still needed. The simultaneous modeling of observed finite states and unobserved continuous states is treated in section 5. Section 6 concludes and give perspective for further research.

[1] Early studies have transition matrices independent of any co-variate. These are useless for policy modeling.

[2] See Piet (2008) for a more detailed critique.

2 Policy consequences for farm structure and structural change

The links between policy, farm structure and structural change are complex. To avoid confusion over different methodologies, some theoretical concepts and assumptions are needed. First an assumption — since farms tend to be fewer and fewer — all potentially active farms at any point in time are assumed to belong to some finite set of farms which at some time have been active. These individual farms are indexed, $n = 1, \ldots, N$. The assumption is greatly simplifying our modeling. Without it we would have to deal with distributions of farm entry in the continuous space.

The object of study is some positive variable, $y = y(z)$, affected by policy variables, z, and aggregated over farms into, $Y(Z) = \sum_n y_n(z_n)$. Sometimes is Y observable and $z_n = z$ for all n. Then, can policy analysis be conducted in the aggregate. In all other cases one need to approach policy analysis with a micro foundation.

Since, $y_n(z_n)$ necessarily is a stochastic variable, the aggregate need to be modeled as a sum of expectations of individual y_n relative to some probability distribution, $\pi(y_n|z_n)$.[3] Each distribution is in turn decomposed in a discrete part, the probability that farm n is active ($y_n > 0$), and a continuous part, the probability distribution of y_n given that n is active:

$$\mathrm{E}Y(Z) = \sum_n \mathrm{E}_{\pi(y_n|z_n)} y_n = \sum_n \int \pi(y_n > 0|z_n)\ \pi(y_n|y_n > 0, z_n)\ y_n\ dy_n$$

This decomposition is favorable partly because discrete and continuous probability distribution need different specifications, and partly because distinct aspects of policy effects on farm structure are mirrored. The policy effect can now be modeled as:

$$\begin{aligned}\partial_Z \mathrm{E}Y(Z) = \sum_n \int \left[\partial_{z_n} \pi(y_n|y_n > 0, z_n)\ \frac{\partial_{z_n}}{\partial Z}\right]\ \pi(y_n > 0|z_n)\ y_n\ dy_n \\ + \sum_n \int \left[\partial_{z_n} \pi(y_n > 0|z_n)\ \frac{\partial_{z_n}}{\partial Z}\right]\ \pi(y_n|z_n, y_n > 0)\ y_n dy_n \quad (1)\end{aligned}$$

where the first sum is the effect via change on the active farms, *the farm size effect*, and the second sum is the effect via the number of active farms, *the farm number effect*. Depending on the policy, the one can be zero while the other is non-zero, both can be non-zero of same sign or of opposite signs. In the latter case it may happen that the two effects cancel out. Price support is a case for a strong farm size change effect on the quantity produced, while the farm number effect is expected to be vague. Direct differentiated payments to active farms will be different with a negative farm size effect due to splitting and a positive farm number effect. Possibly, these effects will cancel out, but that depends on policy details.

3 State space models

Presentation of state space models are typically made in terms of time series which are followed in real time. This corresponds to the framework where they were first

[3] $\pi(\cdot)$ is utilized as generic notation for probability distributions.

employed: the control of spacecrafts based on noisy signals of their position. The application is based on work of Kalman (1960) and Kalman and Bucy (1961) introducing the *Kalman filter*. Applications in economics are in particular guided by Harvey (1989).

A standard presentation of a linear state space model consists of a state equation and a measurement equation. The former can be specified as:

$$x_t = F_t x_{t-1} + \xi_t \tag{2}$$

where x_t is a J-dimensional vector of unobserved state variables, F_t is a transition matrix, and ξ_t is a vector of $\mathcal{N}(0, Q_t)$ stochastic variables mutually independent over time, where Q_t is a symmetric positive semi-definite matrix. The state equation and its assumptions implies a Markov property: The distribution of x_t conditional on all previous realizations is identical to the distribution of x_t conditional on the latest previous realization, generically $\pi(x_t|x_{t-1}, \ldots, x_0) = \pi(x_t|x_{t-1})$.

In turn the measurement equation

$$y_t = Z_t x_t + \epsilon_t \tag{3}$$

links the vector of observations, y_t, with the state variables by means of a matrix of exogenous variables Z_t, and measurement errors, ϵ_t, which are mutually independent over time, also independent of the ξ_t-s, and distributed $\mathcal{N}(0, \Sigma_t)$, where Σ_t is a positive definite diagonal matrix. One does not need off-diagonal elements of Σ since covariance between different elements of y_t can be expressed by off-diagonal elements of Q_t. The measurement equation and its assumptions inhibits conditional independence of $y_t|x_t$. Moreover, the Markov property of the state equation is carried over to the complete model, which consequently is a Markov model with hidden continuous states.

There are more general specifications of state space models than (2) and (3) (de Jong 1991), but this one suffices for the current context. Typically, the context will also suggest some structure for the involved matrices. State variables can be grouped in different independent blocks, leading to a block-diagonal structure of F_t and Q_t. A group of state variables can exhibit deterministic change from $t-1$ to t in terms of the relevant block of F_t when the corresponding block of Q_t is equal to zero. Random walks are obtained with a block of F_t equal to the identity matrix. Independence between time points $t-1$ and t is obtained with a zero block of F_t. Cycles can be modeled with a rotation matrix block of F_t.

Grouping of state variables are also relevant for the column groups of the matrix Z_t. Some groups will simply have columns with a certain pattern of 1-s and 0-s for all t, meaning that state variables are summed in a certain way for the explanation of observations y. Other groups can have time-varying exogenous data in their columns. The state variable will then represent a random (and possibly time-varying) regression effect of the exogenous data on y.[4]

Economic time series are somewhat different from spacecraft positions. In the spacecraft case the involved matrices are known or estimated in advance. The system is known, the focus is on estimated states. For economic series the system specified by the matrices is not known in advance. Hence, F_t, Q_t and Σ_t are parameters or functions of parameters that need to be estimated. Actually, the revelation of the

[4]Within agricultural economics Chavas (1985) contributed early along this line with an analysis of the demand for meat over time.

system parameters is most often the main focus. However, if one succeeds in finding a stable and broadly accepted model, then can this model eventually be exploited in monitoring and governing the economy by spacecraft analogy.

Economic time series have always some unobserved previous history of which little or nothing is known. Again, this is a contrast to spacecraft navigation. However, such lack of information can be handled within state space frame because of the inherently Bayesian nature (Meinhold and Singpurwalla 1983). A normal prior for the initial state variables, x_0, can be specified with $\mathrm{E}x_0 = \overline{x}_0$ and $\mathrm{Var}\ ^{-1}x_0 = \overline{Q}_0^{-1}$. The prior has an impact on the remaining state variable estimates, in particular when the time series are short. Consequently, it is desirable to estimate $\overline{x}_0$ and $\overline{Q}_0$, when these are not given from some other procedure of inference.

At last, economic time series tend to be incomplete, in the sense that the involved array of dependent and independent variables tend to be too small to provide the correct model, or the data set has to few observations to identify all effects in that model. In such cases the state variables mimic the heterogeneity caused by all the omitted variables, and will reduce the problems of mis-specification.

A nice and convenient feature of state space models is that they contain both autoregressive (AR) and moving average (MA) time series models as special cases. $\Sigma_t = 0$ for all t means an AR model, whereas $Q_t = 0$ for all t means an MA model (Durbin and Koopman 2001). The order of the model depend on the structure of the matrices. Thus, because every ARMA-model is a special case of a state space model, the Markov structure of the state variables means little loss of generality, potentially. The statistician/economist is responsible for choosing a structure of F_t, Z_t, Q_t and Σ_t which makes the Markov property likely. If he succeeds, he has arrived at a relatively sparse parametrization of a dynamic system where everything depends on everything. The sparseness combined with generality is the virtue of this Markov model.

Since state-variables are unobserved, one need to predict them conditional on the prior distribution, the system matrices and the observations. Estimation of parameters has to rely on such predictions. Recursive prediction of state variables over time given observations up to that point in time is known as *filtering*[5], and this is of course the core of spacecraft navigation. When using a diffuse prior, the filter predictions at early points in time will be poor being based on little information. The technique of *smoothing* gives relief for this problem by running a backwards recursion so that estimates of state variables at all points in time can be stated, given priors and all observations. Details of filtering and smoothing can be found in textbooks like Durbin and Koopman (2001) and will not be presented here. Instead we turn to Fahrmeir and Tutz (2001) (pp. 340-42) who takes a different approach based on **?**.

Fahrmeir treats filtering and smoothing simultaneously by estimating the posterior mode of state variables given all observations. With some change of notation from Fahrmeir's exposition, let:

$$G = \begin{pmatrix} -F_1 & I & 0 & & 0 \\ 0 & -F_2 & I & \ddots & \\ & \ddots & \ddots & \ddots & 0 \\ 0 & & 0 & -F_T & I \end{pmatrix}, \quad Z = \begin{pmatrix} Z_1 & & 0 \\ & \ddots & \\ 0 & & Z_T \end{pmatrix}$$

[5]The Kalman filter is the first and most well-known version, but several other filters exist.

$$x = (x_0', x_1', \ldots, x_T')', \quad \xi = (\xi_1', \ldots, \xi_T')', \quad y = (\overline{x}_0', y_1', \ldots, y_T')', \quad \epsilon = (x_0 - \overline{x}_0, \epsilon_1', \ldots, \epsilon_T')'$$

so that the state equation is written:

$$Gx = \xi \tag{4}$$

with $\xi \sim \mathcal{N}(0, Q)$ where $Q = \text{diag}(Q_1, \ldots, Q_T)$, while the initial condition and the measurement equation are joined as:

$$y = Zx + \epsilon \tag{5}$$

with $\epsilon \sim \mathcal{N}(0, \Sigma)$ where $\Sigma = \text{diag}(\overline{Q}_0, \Sigma_1, \ldots, \Sigma_T)$. This leads to the following complete log-likelihood of both observed and unobserved data in terms of unknown but fixed parameters θ_0:

$$\mathcal{L}(x, y|\theta_0) = -\frac{1}{2}\Bigg(\log|Q| + \log|\Sigma| + x'G'Q^{-1}Gx + (y - Zx)'\Sigma^{-1}(y - Zx)\Bigg) \tag{6}$$

This log-likelihood can for given parameters, θ_0, be maximized with respect to x to obtain an estimate of the posterior mode of x, $\widehat{x} = \widehat{x}(y, \theta_0)$. The estimate maximizes the density of the distribution of x. Since the distribution is normal and hence symmetric, this is also an estimate of the expectation of x. Fahrmeir shows that this is equivalent to that obtained from Kalman filtering and smoothing (Fahrmeir and Tutz 2001) pp. 342. We prefer this form instead of the explicit Kalman recursions partly because it is conceptually simpler, partly because it is easily generalized to models with non-linear measurements which will be introduced later.

The likelihood function (6) constitutes also a core for parameter estimation with the EM-algorithm (Dempster, Laird, and Rubin 1977), which again is essential for non-linear measurement models. This is based on the second order Taylor expansion of $\mathcal{L}(x, y|\theta_0)$ around $\widehat{x}$:

$$\mathcal{L}(x, y|\theta_0) = \mathcal{L}(\widehat{x}, y|\theta_0) + 1/2(x - \widehat{x})'\partial^2_{xx'}\mathcal{L}(\widehat{x}, y|\theta_0)(x - \widehat{x})$$

This Taylor expansion is exact because $\mathcal{L}$ is a quadratic function in x. Moreover, the first order term is zero because $\widehat{x}$ satisfies the first order condition, $\partial_x\mathcal{L} = 0$.

To approach the likelihood $\mathcal{L}(y|\theta)$ of the true parameters, we apply Bayes theorem to write:

$$\mathcal{L}(y|\theta) = \mathcal{L}(x, y|\theta) - \log\pi(x|y, \theta)$$

and take expectations over $\pi(x|y, \theta_0)$ to obtain

$$\mathcal{L}(y|\theta) = \int \mathcal{L}(y|\theta)\pi(x|y, \theta_0)dx = \int \left(\mathcal{L}(x, y|\theta) - \log\pi(x|y, \theta_0)\right)\pi(x|y, \theta_0)dx$$

When $\mathcal{L}(x, y|\theta)$ is a quadratic function in x and $\pi(x|y, \theta_0)$ is normal, the first term can be integrated analytically. The second term does not depend on θ, hence maximizing the first term lead to an improved estimate of θ. When it comes to the normal densities $\pi(x|y, \theta_0)$, their expectations and variances are usually be found by the Kalman filter and smoother. However, the complete log-likelihood (6) points to another way of deriving them. When the maximization of this log-likelihood with respect to x, gives the expectations of $\pi(x|y, \theta_0)$, the inverse variance matrices will turn out from the negative Hessian evaluated in the optimal point.

4 State space models with spatial interactions

There is a symmetry between neighbors in space and neighbors in time which is exploited in spatial econometrics. By and large spatial econometrics is mimicking time-series techniques in spatial contexts. Both AR and MA models have their counterparts in spatial models. This is to some extent also the case with state space models, although a slight contrast between time and space need to be emphasized. Time has a natural ordering, whereas space in general has not.[6] The statistician is responsible for stating the spatial relationships. These should both be justified and lead to computable models in reasonable time.

Cases with a large time dimension and a small spatial dimension can in principle be treated as a multivariate time series where the observations from all locations, y_{nt}, $n = (1, \ldots, N)$ are stacked into one large time-dependent vector, $y_t = (y'_{1t}, \ldots, y'_{Nt})'$. No particular attention is then paid to spatial structure. The spatial structure is then estimated. If this does not work computationally, an argued spatial structure is required.

In extending the state space model to spatial dimensions the diagonal structure of Σ^{-1} is retained to have contingent independence among observations, $y_{nt}|x_{nt}$, for all (nt). All the spatial and temporal relationships are then taken care of by the normally distributed state variables. Corresponding to the Markov structure of the time series state space model, one can impose a similar structure in the spatial dimension with specified feedbacks only between spatial neighbors. This is the core of a Gaussian Markov Random Field (GMRF) (Rue and Held 2005). With a GMRF it is ensured that the spatial dependence of everything upon everything is parameterized with a sparse structure involving only neighbors.

To make this more explicit, let B be a $JN \times JN$ matrix of eventual neighbor feedbacks. This matrix is sparse as all elements which are not associated with neighbors are zero. The neighbors are thus given by some index set, $N^* \subset (N \times N)$. Let C be a block diagonal JN-dimensional variance matrix of unit-specific stochastic variables, u_n, independent between units. Spatially dependent state variables can now be specified as x satisfying: $x = Bx + u$. The variance matrix of x is given most conveniently by its inverse, the precision matrix:

$$Q^{-1} = (I^{JN} - B)'C^{-1}(I^{JN} - B)$$

which also is sparse. Actually, the precision matrix shows the contingencies in the spatial structure with non-zero entries.

Of particular relevance to economics, there are also other ways to model spatial interactions among neighbors. Economic time series of a population of single firms will exhibit transactions between firms, even though it is not observed who is buying from who. When these transactions happens mostly between neighbors, a specified spatial structure can account for them. This is immediately relevant for structural change in agriculture.

Let S be a $(N \times (N-1))$ matrix, and let u be a stochastic $J^{-}N(N-1)$ vector with distribution $\mathcal{N}(0, \psi \otimes I^{N(N-1)})$ where ψ is scalar. The possibility of a transaction

[6] E.g. the Key West archipelago, Chile and local parts of Norway, can at a certain level of spatial aggregation be considered ordered.

between neighbor firms n and m is specified with

$$S^-_{l,n\times m} = \begin{cases} d^{-1}_{nm} & \text{for } l = n, (n,m) \in N^* \\ -d^{-1}_{nm} & \text{for } l = m, (n,m) \in N^* \\ 0 & \text{otherwise} \end{cases}$$

where d_{nm} is a measure of distance between neighbors n and m. Transactions of a certain commodity arise according to u. The expression Su — which is a block of mutually dependent state variables — will now be the stochastic outcome of transactions between neighbor firms. The distribution is normal, expectation is zero, and variance matrix is $\psi SS'$, which by construction has a spatial structure. This variance matrix is not necessarily full rank. Actually, whenever some a firm is located to have only a single neighbor, the variance matrix will be singular.

In multivariate settings the probability of transactions will vary with the commodity, but the spatial structure of neighborhoods and distances is anyway invariant. One would then assume that the variance of the outcome of transactions takes the form $\Psi \otimes SS'$, where Ψ is some symmetric positive definite matrix.

The variance matrix SS' is sparse, but its inverse or generalized inverse — the precision matrix — is not. The precision matrix is constant, however, and need only be computed once. The lack of sparseness is then not a serious computational problem. The generality of the GMRF-s with only sparse precision matrices is thus challenged.

Another likely possibility in such economic series is positive correlation between neighbors due to unobserved common shocks. Such effects can be specified with corresponding shocks in much the same way as transactions. Now a matrix S^+ is defined with

$$S^+_{l,n\times m} = \begin{cases} d^{-1}_{nm} & \text{for } l = n \text{ or } l = n, (n,m) \in N^* \\ 0 \text{ else} \end{cases}$$

and a stochastic u^+ is distributed $\mathcal{N}(0, \psi^+ I^{N(N-1)})$. S^+u^+ is then the stochastic outcome of shocks common to neighbors.

Apart from such structuring of the variance matrices, Q_t, no essential change to the state space model is needed when adapting to a spatio-temporal context. However, spatial models tend to have a smaller time-dimension than purely temporal models. A common case is a pure cross-section with no time at all. Will the time-series modeling then have any relevance? In any case are the initial assumptions, $\overline{x}_0$ and $\overline{Q}_0$, relevant, but can the EM algorithm be utilized in their estimation? In principle the answer is yes. A single data point is sufficient to estimate a posterior distribution, $\pi(x_0|y)$, to be applied in an EM-algorithm.

5 State space model mixed with observed finite states

Entries and exits are always associated with panels of farm data, and are closely linked to structural change. Thus, entry and exit need to be modeled simultaneously with the continuous aspects of the data. However, entry and exit are not single dimensional. A farm may be active in various activities, $k = (1, \ldots, K)$. We write observed activity of n at t w.r.t. activity k as $c_{ntk} = 1$, and inactivity as $c_{ntk} = 0$. Since the conditional probability of activity, $\pi(c_{ntk} = 1|x_{nt})$, not necessarily is associated with a high expected scale, $\mathrm{E}_{\pi(y_{ntk}|c_{ntk}=1,x_{nt})}y_{ntk}$, the two contingent probabilities, $\pi(c_{ntk} = 1|x_{nt})$ and $\pi(y_{ntk}|c_{ntk} = 1, x_{nt})$, are specified independently. The first is specified as a logit:

$$\log \pi(c_{ntk} = 1|x_{nt}) - \log \pi(c_{ntk} = 0|x_{nt}) = (Z^{0k}x)_{nt}$$

$$\log \pi(c_{ntk} = 0|x_{nt}) = -log(1 + \exp(Z^{0k}x)_{nt})$$

where Z^{0k}_{nt} is some exogenous matrix summing a set of state-variables. For the k-active farm we observe not only activity, but also the level, $y_{ntk} > 0$. The positivity of the observation suggests that some other distribution than the normal should be utilized. In any case each element y_{ntk} of y_{nt} will be modeled separately with a distribution, f_{ntk}, of its own. This is in line with the linear state space models where errors of the measurement equation are assumed independent between elements. This modeling strategy opens for the use of univariate parametric distributions like the gamma distribution. To make the distribution contingent on state variables in line with generalized linear modeling, a response function, h_k, which has argument an aggregate of state variables $Z^{1k}x)_{nt}$, is chosen. $h(Z^{1k}_{nt})$ is then utilized as the expectation parameter of the distribution, f_{ntk}.

As a response function we propose, $h_k(\eta) = (\eta + \sqrt{\eta^2 + \eta_k})/2$, where η_k is a positive parameter. This response function has the desirable property of being almost linear in η.

One additional complication needs to be dealt with. Our data is subject to censoring. Farms with activities in small scale only will not be observed. Censoring can be modeled relative to linear combination of state variables which may vary over time, $\gamma_t Z^* x \leq \gamma_0$, where γ_t is a parameter vector, $Z^* = Z^{11}, \dots, Z^{1I}$, and γ_0 is a scalar. The probability that a farm with state x is not censored at t can be also specified as a logit probability,

$$\log \pi(y_{nt} \neq 0|x_{nt}) - \log \pi(y_{nt} = 0) = \gamma_t(Z^*x)_{nt} - \gamma_0$$

$$\log \pi(y_{nt} = 0|x_{nt}) = \log(1 + \exp(\gamma_t(Z^*x)_{nt} - \gamma_0))$$

The interpretation is that when state variables are relatively large indicating relatively large expectations of observations, the probability of censoring is small.

These modifications of the measurement model have consequences for the complete log-likelihood function:

$$\begin{aligned}\mathcal{L}(x; y, \theta) = \sum_{nt} \log \pi(y_{nt} = 0|x_{nt}) + \sum_{y_{nt}\neq 0} \left[\gamma_t(Z^*x)_{nt} - \gamma_0 + \sum_k \log \pi(c_{ntk} = 0|x_{nt}) \right] \\ + \sum_{y_{ntk}=1} \left[(Z^{0k}x)_{nt} + \log f_{ntk}(y_{ntk}; (Z^{0k}x)_{nt})) \right] \\ - \frac{1}{2} \left(\log|Q_0| + \log|Q| + (x_0 - \overline{x}_0)'\Sigma^{-1}(x_0 - \overline{x}_0) + x'G'Q^{-1}Gx \right) \quad (7)\end{aligned}$$

Estimation with the EM-algorithm can proceed almost as in section 3. As for the linear model, expectations and variances of $\pi(x|y, \theta)$ can be found from the maximization of (7). But since the measurement model no longer is quadratic in x and exact analytic integration is impossible, some numeric integration method is required.

6 Conclusions and further research

An estimable model of farm structure is proposed above. The model unifies both the continuous change on existing farms and the change of farm numbers. An estimated model will make it possible to predict both aspects both over time and space. Although not pronounced in the model development, exogenous policy variables z can be introduced in the model either as suggested as coefficients for random effects, Zx, or as fixed effects subtracted from the measurements, $y_{nt} - \beta z_{nt}$.

Markov methods are central for the feasibility of this modeling. First, with a Markov property on the unobserved (hidden) state variables, the dependency of everything upon everything can be modeled with relatively sparse structures. Secondly, Markov chains are required in the estimation algorithms. Apart from the keyword "Markov", there is virtually nothing in common with the traditional Markov chains of farm/firm structural change. Actually, the transition matrix, which is the starting point of those models, has played no role here. For sake of comparison one might state a transition probability with the established notation and methodology:

$$\pi(y_{ntk} = y_{1k}|y_{n,t-1} = y_0) \\ = \mathrm{E}_{\pi(x_{nt}|x_{n,t-1})\pi(x_{n,t-1})}\pi(y_{ntk} = y_{1k}|x_{nt})/\mathrm{E}_{\pi(x_{n,t-1})}\pi(y_{n,t-1} = y_0|x_{nt})$$

All elements of the transition are involved in the estimated model. The contingent probability of y_{nt} given x_{nt}, $\pi(y_{ntk} = y_{1k}|x_{nt})$, comes from the measurement model. The contingent probability of x_{nt} given $x_{n,t-1}$ comes from the state equation. And a last $pi(x_{nt})$ is a prior distribution of the state variable consisting with previous observations. Since every entity is contingent on state variables, and these are appropriately integrated, there is no heterogeneity problem.

There is no point in computing transition matrices in this way in order to look for Markov chain equilibria. As pointed out in the introduction, the transition matrix is a record of the structural process in time. The equilibrium structure at a certain point in time is given by the estimated distributions at that point in time. These are derived taken time series information into account in general terms without forcing it into the straight-jacket of finite-state transition matrices.

An estimated model combined with the complete data set of Norwegian farms, can be utilized in various ways. First, it can give micro-based spatial and temporal predictions of farm structure based on policy variables, hopefully as intended. By the separation of the continuous and discrete aspects of farm structure there will be a much higher level of detail in these predictions than what can be obtained with traditional Markov chain models. Since the micro-spatial aspect of structural change is completely novel, an evaluation of different spatial specifications is required.

Secondly, the model can be utilized as a benchmark for various alternative models based on less informative data. FADN (European Commission 2011) has time series of individual farms sampled with regional balance but with no regard to spatial neighbors. USDA has repeated cross-sections of individual farms with the ARMS data (United States Departement of Agriculture 2011). This set pays little regard both to spatial and temporal neighbors. Apart from such relatively structured data sets, aggregate data of various types is the mostly used source for prediction of micro structure. The ability of such data sets to predict micro structure is hypothesized to be ranked in the order they are mentioned.

At last, various alternative sampling schemes can be evaluated according to their best model results. Possibly, one should systematically sample close neighbors both in time and space to each randomly selected farm. Such a sampling procedure may be a serious contender to complete sampling — which for most practical purposes will be too costly.

References

Chavas, J.-P. (1985): "Structural change in the demand for meat," *American Journal of Agricultural Economics*, 65(1).

de Jong, P. (1991): "The diffuse Kalman filter," *The Annals of Statistics*, 19(2), 1073–1083.

Dempster, A. P., N. M. Laird, and D. B. Rubin (1977): "Maximum Likelihood from Incomplete Data via the EM Algorithm," *Journal of the Royal Statistical Society. Series B (Methodological)*, 39, 1–38.

Durbin, J., and S. J. Koopman (2001): *Time Series Analysis by State Space Methods*. Oxford Statistical Science Series.

European Commission (2011): "The Farm Accountancy Data Network," `http://ec.europa.eu/agriculture/rica/index_en.cfm`.

Fahrmeir, L., and G. Tutz (2001): *Multivariate Statistical Modelling Based on Generalized Linear Models*. Springer, 2nd edition.

Harvey, A. C. (1989): *Forecasting, structural time series models and the Kalman filter*. Cambridge Univesity Press.

Kalman, R. (1960): "A new approach to linear filtering and prediction problems," *Journal of Basic Engineering*, 82(1), 35–45.

Kalman, R., and R. Bucy (1961): "New Results in Linear Filtering and Prediction Theory," *Journal of Basic Engineering*, 83, 95–108.

Meinhold, R. J., and N. D. Singpurwalla (1983): "Understanding the Kalman Filter," *The American Statistician*, 37(2), 123–127.

Piet, L. (2008): "The evolution of farm size distribution: revisiting the Markov chain model," in *12th EAAE Congress, August 26-29, Ghent, Belgium*.

Rue, H., and L. Held (2005): *Gaussian Markov Random Fields: Theory and Applications*, vol. 104 of *Monographs on Statistics and Applied Probability*. Chapman & Hall, London.

United States Departement of Agriculture (2011): "Agricultural Resource Management Survey," `http://www.ers.usda.gov/Data/ARMS/`.

Zimmermann, A., T. Heckelei, and I. P. Domínguez (2009): "Modelling farm structural change for integrated ex-ante assessment: review of methods and determinants," *Environmental science and policy*, 12, 601–618.

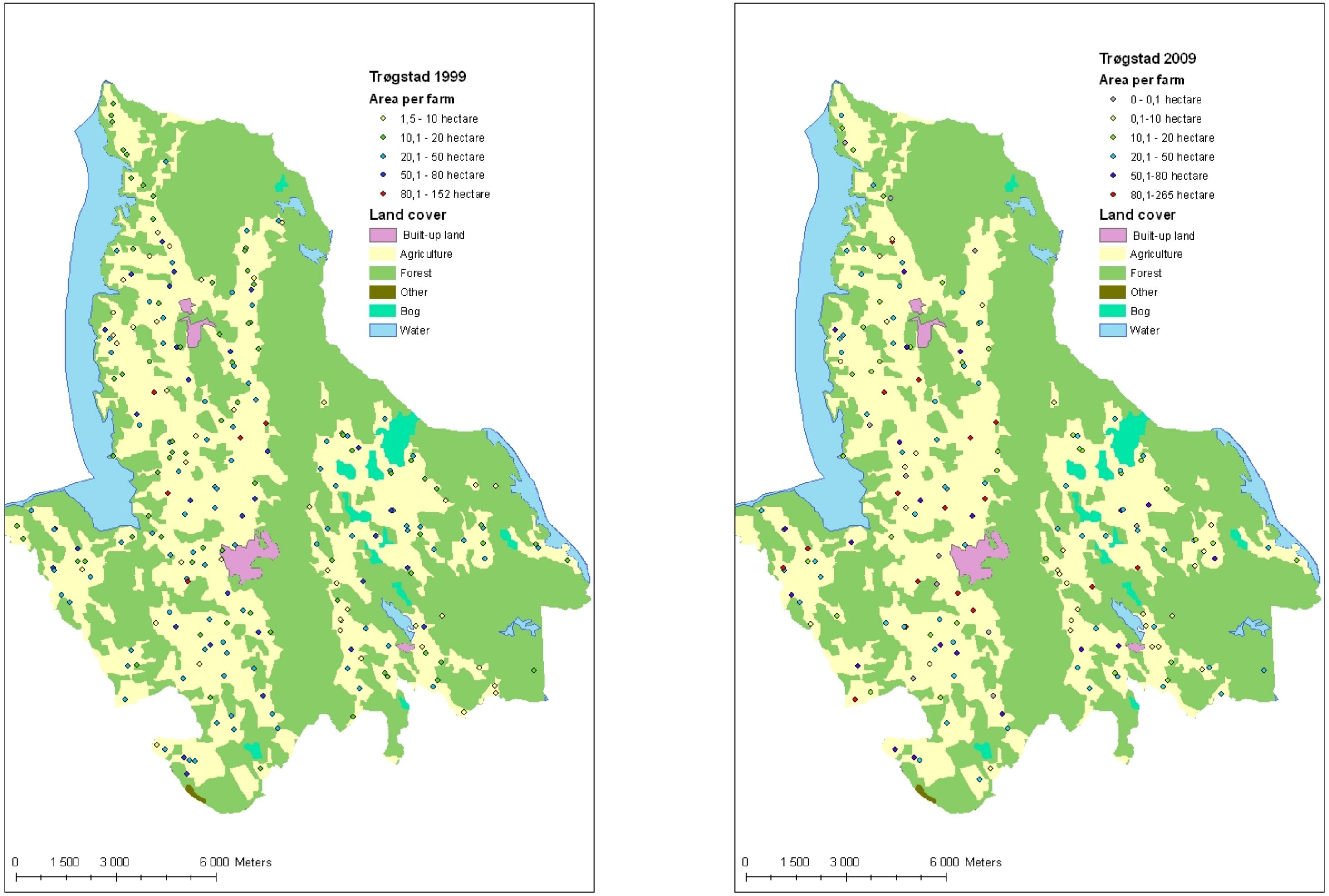

Figure 1: Farm locations and size in municipality Trøgstad, 1999 and 2009. Complied from The Norwegian Agricultural Authority (SLF), Application for agricultural support, and AR250 from Skog og landskap