AARHUS
BSS

2016-10
Reza Pourmoayed
PhD Dissertation

Optimization Methods in a Stochastic Production Environment

DEPARTMENT OF ECONOMICS AND BUSINESS ECONOMICS
AARHUS BSS • AARHUS UNIVERSITY • DENMARK

SCHOOL OF BUSINESS AND SOCIAL SCIENCES
AARHUS UNIVERSITY

OPTIMIZATION METHODS IN A STOCHASTIC PRODUCTION ENVIRONMENT

PhD dissertation

Reza Pourmoayed

Aarhus BSS, Aarhus University
Department of Economics and Business Economics

2016

Preface

This dissertation is the result of my PhD studies at Department of Economics and Business Economics, Aarhus University, during the period March 2013 to March 2016. My PhD project is a part of the PigIT project funded by the Danish Council for Strategic Research.

I have always been interested in working with real-world problems and I am very grateful that I could focus my PhD studies on a practical problem related to livestock production. Now that my PhD studies are coming to an end, I have a very good feeling and I think my knowledge has been improved in theoretical aspects as well as computer programming.

This PhD dissertation is the result of academic work on three research projects, each presented as an independent research paper in the dissertation. The first paper has already been published, the second paper has been submitted to a peer-reviewed operations research journal, and the last paper is in the final stage for submission to a peer-reviewed journal.

Acknowledgment

First of all, I would like to express my sincere gratitude to my supervisor Lars Relund Nielsen for continuous support during my PhD. His valuable advice and suggestions have extensively contributed to the content of this dissertation. I have learned a lot from his professional abilities in computer programming and his unique perspective on research. In addition to our academic collaboration, I greatly esteem him for his honesty and pleasant personality.

My sincere thanks also go to my first co-supervisor Anders Ringgaard Kristensen, who offered me a research stay for three months at Centre for Herd-oriented Education, Research and Development at University of Copenhagen. Anders comments and guidances regarding the research papers in this dissertation have always been valuable and useful for me. I would also thank my second co-supervisor Christian Larsen for his support and for reviewing the second and third papers in this dissertation.

Moreover, I would like to extend my gratitude to all members of the PigIT project. It has been a valuable experience for me to be in this project and to have collaborations with related specialists in academia and industry. In particular, I would like to thank Anders Ringgaard Kristensen, Cécile Cornou, Katarina Nielsen Dominiak, Dan Børge Jensen, Anna Helena Stygar, and Gang Jun Tu.

I would also express my thanks to the Department of Economics and Business Economics at Aarhus University and especially to the Cluster of Operations Research and Logistics (CORAL) for providing an excellent research environment. It has been an honor for me to be a member of the CORAL group and I always enjoyed participating in the interesting courses and seminars of this group. I am also thankful to the following administrative staff of the department for their various forms of support during my PhD studies: Ingrid Lautrup, Karin Vinding, Charlotte Sparrevohn, Christel Mortensen, and Susanne Christensen.

The friendship and support of my fellow PhD students have made the past three years an unforgettable period of my life. In particular I am thankful to my office mate Sune Lauth Gadegaard, I appreciate the help you gave me when I had to adapt to my new working environment and for answering a lot of questions during the past three years. In addition, I would like to thank Ata Jalili Marand, Samira Mirzaei, Maryam Ghoreishi, Parisa Bagheri Tookanlou, Viktoryia Buhayenko, and Agus Darmawan for a friendly work environment and many insightful discussions.

Most importantly, I am deeply indebted to my parents, sisters, and brother who have always been encouraging, supportive and shown belief in me and my work. Finally, last but not least, I want to give my special thanks to my wife Mahboobeh, for her love and patience during the past three years. Without her support and encouragement, finishing this dissertation would not have been possible.

Reza Pourmoayed
Aarhus, March 2016

Table of contents

Summary

This dissertation with an interdisciplinary approach applies techniques from Operations Research and Statistics in order to develop models that support decisions regarding feeding and marketing of growing/finishing pigs. Stochastic dynamic programming is used as the main optimization tool to model decisions, and state space models are used as the primary statistical technique to describe the stochastic nature of the system. Based on data streams from online farm sensors and market prices, the state space model transforms data into information which is embedded into the decision models using Bayesian updating.

In the production of finishing pigs, feeding is an important operation and has a direct influence on the cost and the quality of the meat. Another important operation is the timing of marketing. It refers to a sequence of culling decisions until the production unit is empty. As a result, the profit of the production unit is highly dependent on the feeding cost and on good timing of marketing, i.e. decisions about feeding and marketing have a high impact on the profitability. Hence, it is relevant to consider decision models that optimize feeding and marketing decisions. This dissertation focuses specifically on three challenges. 1) Marketing and feeding decisions could be optimized simultaneously. 2) Fluctuations in the pork, feed, and piglet prices may have an impact on the optimal marketing policy. 3) Cross-level constraints implied by considering decisions at different levels such as animal, pen, section, and herd may affect the marketing policy.

Besides an introduction, this dissertation consists of three papers. In the first paper, feeding and marketing decisions are taken into account simultaneously. Since the choice of feed-mix affects the pigs' growth, a specific feeding strategy has an impact on the marketing policy. That is, economic optimization of feeding and marketing decisions is interrelated and requires simultaneous analysis. It is therefore relevant to combine these decisions in the same model. In the second paper, marketing decisions are considered under fluctuating pork, feed, and piglet prices. Since the reward of marketing a pig depends on the pork, piglet, and feed prices, weekly fluctuations in these prices may have an impact on the farmer's decision on when to market the pigs and buy new piglets or feed stock for the farm. Hence, it is relevant to model stochastic market prices instead of assuming deterministic prices in the model. Finally, in the third paper, marketing decisions are considered at different levels simultaneously. The paper considers the challenge of integrating decisions at animal, pen, section, and herd level. That is, decisions at one level (e.g. animal or pen) may influence decisions at other levels (e.g. section or herd). An example could be that delivery to the abattoir must be coordinated with marketing at pen level and termination of a whole section at section level.

Resumé

Denne afhandling har en tværfaglig tilgang og anvender teori fra både operationsanalyse og statistik til at udvikle modeller, der understøtter beslutninger vedrørende fodring og levering af slagtesvin. Stokastisk dynamisk programmering anvendes som hovedoptimeringsværktøj til at finde optimale beslutninger, og state space modeller anvendes som det primære statistiske værktøj til at beskrive systemets stokastiske karakter. På basis af data fra online sensorer i bedriften og markedspriser, bruges de statistiske modeller til at omdanne data til relevant information, som inkorporeres i beslutningsmodellerne vha. Bayesiansk opdatering.

I produktionen af slagtesvin er fodring et vigtigt element, der har direkte indflydelse på produktionsomkostningerne og kødets kvalitet. Et andet vigtigt element er timingen af leveringer til slagteriet, dvs. vi må tage en række beslutninger om levering af de enkelte svin indtil produktionsenheden er tom. Som følge heraf er produktionsenhedens overskud i høj grad afhængig af omkostningerne til fodring og timingen af leveringer. Det er således vigtigt at betragte beslutningsmodeller, der optimerer beslutninger vedrørende fodring og levering. Denne afhandling koncentrerer sig om tre udfordringer: 1) at beslutninger vedr. levering og fodring i nogle tilfælde bør optimeres simultant; 2) at variationer i priserne for svinekød, foder og smågrise kan have indflydelse på den optimale leveringspolitik; og 3) at begrænsninger på tværs af forskellige niveauer, såsom dyr, sti, sektion og besætning, kan have indflydelse på leveringspolitikken.

Afhandlingen består af en introduktion samt tre selvstændige artikler. I den første artikel betragtes beslutninger vedrørende fodring og levering simultant. Eftersom valg af foderblanding har betydning for dyrenes vækst, så har en specifik fodringsstrategi også indflydelse på leveringsstrategien. Vi skal altså optimere beslutninger vedrørende fodring og levering simultant i den samme model. I den anden artikel ses der på leveringsstrategier under varierende priser på svinekød, foder og smågrise. Eftersom overskuddet afhænger af priserne på svinekød, smågrise og foder, kan ugentlige variationer i priserne have betydning for, hvornår grisene skal leveres, og hvornår der skal indkøbes nye smågrise eller foder. Det er således relevant, at modellere stokastiske markedspriser i stedet for at forudsætte deterministiske priser i modellen. Afhandlingens tredje artikel omhandler leveringsstrategier givet begrænsninger på tværs af forskellige niveauer, såsom dyr, sti, sektion og besætning. Beslutninger på ét niveau skal tages under disse begrænsninger. For eksempel skal levering til slagteriet betragtes givet en kapacitet af lastbilerne, og den begrænsning, at nye smågrise kun kan indsættes når en sektion er tom og rengjort.

Chapter 1

Introduction

This PhD dissertation considers optimization algorithms and statistical forecasting that support sequential marketing and feeding decisions in the production of growing/finishing pigs. More precisely, stochastic optimization models and statistical learning algorithms are applied to utilize information streams such as online information from farm data and market prices and make them an integral part of the decision process. As a result, this dissertation considers interdisciplinary research and uses techniques from both Operations Research and Statistics.

The dissertation is the outcome of a sub-project within the project PigIT[1] funded by the Danish Council for Strategic Research. PigIT focuses on integrating information and communication technology in the production process of growing/finishing pigs through statistical models for automatic monitoring and operations research methods for decision support systems. The objective of PigIT is to contribute significantly to the competitiveness of the Danish slaughter pig industry while still ensuring a satisfactory level of animal welfare. This dissertation focuses on developing decision models at farm level and the main scope is limited to marketing and feeding decisions of growing/finishing pigs in finisher production units.

The remainder of the introduction starts by giving an overview over finisher pig production with a focus on marketing and feeding decisions. Next, the main modeling techniques used in the dissertation are briefly discussed to familiarize the reader with the methodologies used in this dissertation. Finally, the structure and contributions of the dissertation are described.

1.1 Marketing and feeding decisions in finisher pig production

In Denmark approximately 30 million piglets are produced in every year. A significant number of these piglets are exported to other countries (around 11 mill) and the rest of them (approx. 19 million) are sent to fattening units in Denmark. The number of pig farms in Denmark are approximately 3600 where 50% are finishing farms, 30% are integrated farms (both sow and finishing pigs), and the remaining 20% are sow farms (Danish Agriculture and Food Council, 2015). Pork constitutes about 5% of the Danish export representing a profit of about 30 billion DKK per year (Landbrug & Fødevarer, 2015).

The different production processes within pig production can be classified as mating, gestation, farrowing, weaning, and finishing (Christensen, 2010). This PhD dissertation focuses on

[1] `http://pigit.net`

operational decisions in the production of finishing pigs, i.e. from inserting the piglets (with a weight of approx. 30 kg) into the finishing unit until *marketing/culling* the pigs for slaughter (with a weight of approx. 100-110 kg). In the finishing unit, animals are grouped at different levels: *herd*, *section*, *pen*, and *animal*. Herd is a group of sections, a section includes some pens, and a finisher pen involves some animals (usually 15-20). One of the important challenges in relation to decision support models in finishing units is how to integrate decisions at animal, pen, section, and herd level. That is, decisions at one level (e.g. animal or pen) may influence decisions at other levels (e.g. section or herd).

During the growth period, decisions should be taken such that the profit is maximized under a set of constraints such as transportation, weaner supply policy, welfare, disease strategy, and housing conditions. Two main factors affecting the profitability of the production process are feeding and marketing (Pourmoayed and Nielsen, 2014). The cost of feeding has a high impact on the production cost of a finisher pig and the reward of marketing a pig depends on the *pork price* of the carcass weight of the pig. The pork price per kg carcass weight is a piecewise linear function which is highest if the carcass weight lies in a specific interval. Hence, the farmer must find the best time to market the pigs for slaughter. Moreover, the farmer should choose the best feed-mix giving the optimum growth of the pigs at the lowest cost. As a result, feeding and marketing decisions are important in the production of finishing pigs.

Marketing of pigs refers to a sequence of culling decisions until the production unit is empty. In general pigs grow with different growth rates and obtain their slaughter weight at different times. Therefore, during the last weeks of the growing period, the decision maker should determine which pigs should be selected for slaughter (*individual marketing decision*). Next, after a sequence of individual marketings, the decision maker should decide when to terminate the whole section/pen, i.e. all the remaining pigs are sent to the abattoir (*termination decision*).

The feeding strategy of the pigs affects the growth of the pigs and hence the marketing time. Furthermore, a feed-mix resulting in a faster growth generally costs more compared to a feed-mix with a lower growth. Phase feeding is a common method used in the production of growing pigs. The whole growing period typically includes 3 or 4 phases and each phase (period) involves a predefined feed-mix which is a mixture of different ingredients (barley, soy, maize, etc). A relevant decision is when to change the current feed-mix and what type of feed-mix to use in the next phase (*feeding decision*). Since the choice of feed-mix affects the pigs' growth, a specific feeding strategy has an impact on the marketing strategy. That is, the economic optimization of feeding and marketing decisions is interrelated and requires a simultaneous analysis.

Remark that, there is a high degree of stochasticity in the production of finisher pigs. Animals in general do not grow at the same growth rate and hence there will be a high degree of uncertainty about the weight of the pigs during the growing period. Moreover, the reward of marketing a pig depends on the *pork price* of the carcass weight, the cost of buying the piglet on the market, i.e. the *piglet price*, and the cost of feeding which is dependent on the *feed price* at the time when the feed stock is bought. Weekly fluctuations in pork, feed, and piglet prices may impact farmer's decision on when to market the pigs and buy new feed stock for the farm. Hence, it is relevant to take into account stochastic elements such as growth and market price fluctuations in modeling feeding and marketing decisions.

1.2 Modeling techniques

This dissertation considers stochastic models from both Operations Research and Statistics. Stochastic dynamic programming and statistical learning algorithms based on Bayesian updating are the main tools used to model feeding and marketing decisions and to describe the stochastic nature of the system. More precisely, *state space models* (*SSMs*) are used as the main statistical tool for modeling weight, growth and market prices. Time-series of data retrieved from online monitoring (using e.g. sensor data or image processing) are transformed into information about weight and growth using Bayesian updating. Similarly, uncertainty about pork, feed, and piglet market prices can be described using SSMs to help the farmer in forecasting future prices.

In order to optimize marketing and feeding decisions at pen level, Hierarchical Markov decision processes are used for optimization and the uncertainty of weight and price information are embedded into the models using SSMs. Due to the importance of cross-level constraints (e.g. termination at section level and transportation at herd level), it is relevant to consider marketing decisions at herd level. However, when decisions are modeled at herd level, the size of the model increases dramatically and we cannot use regular techniques to find optimal marketing policies. That is, the *curse of dimensionality* becomes apparent due to the high number of states, actions and possible values of random information. Hence *approximate dynamic programming* is applied to optimize marketing decisions.

Below, a brief introduction to the different models is given. Part of the notation is taken from Pourmoayed et al. (2016) (Paper I) and in general it is kept as close as possible to the notation used in the other papers presented in this thesis.

1.2.1 Hierarchical Markov models

In order to have a frame of reference, first a finite time-horizon semi-Markov decision process (see e.g. Tijms, 2003, Chap. 7) is described and then an introduction to hierarchical Markov decision processes is given.

Finite time-horizon semi-Markov decision processes

A *finite-horizon semi-Markov decision process* (*semi-MDP*) models a sequential decision problem over N *stages*. At a (random) point in time (the start of a stage), the state of the system is observed and an action is chosen. The choice of action at the current state produces two results: an immediate reward is received, and the system evolves probabilistically to a new state at a subsequent (random) point in time.

Let I_n denote the finite set of system states at stage n. Given *state* $i \in I_n$ is observed, then an *action* a from the finite set of allowable actions $A_n(i)$ must be chosen, generating *reward* $r_n(i,a)$. Moreover, let $u_n(i,a)$ denote the *stage length*, i.e. the expected time until next decision epoch $(n+1)$ given action a and state i. Finally, let $\Pr(j \mid n,i,a)$ denote the *transition probability* of obtaining state $j \in I_{n+1}$ at stage $n+1$ given that action a is chosen in state i at stage n.

A *policy* R is a decision rule/function that assigns for each stage $n = 1,\dots,N-1$ and state $i \in I_n$ an action $R(i) \in A_n(i)$, i.e. a policy prescribes which action to take whenever the system is observed in state i at stage n. Note that no action is taken at stage N. Instead the *terminating reward* at this point of time is a function of the state $i \in I_N$ denoted by $r_N(i)$.

Given a policy and terminating rewards, the *expected reward until termination* when we start in state i at stage n can be found using the recursive equations:

$$v_n^R(i) = \begin{cases} r_N(i), & n = N, \\ r_n(i,R(i)) + \sum_{j \in I_{n+1}} \Pr(j \mid i,n,R(i))\, v_{n+1}^R(j), & n = 1,\dots,N-1. \end{cases} \tag{1.1}$$

Given an *initial distribution* $\Pr_0(\cdot)$ of the states at stage 1 and *expected initial reward* r_0, the *total expected reward* of the semi-MDP can be calculated as

$$v(R) = r_0 + \sum_{i \in I_1} \Pr_0(i) v_1^R(i). \tag{1.2}$$

Equations similar to (1.1) and (1.2) can be formulated to calculate the *expected time until termination* (when we start in state i at stage n) and the *total expected time* of the process, respectively.

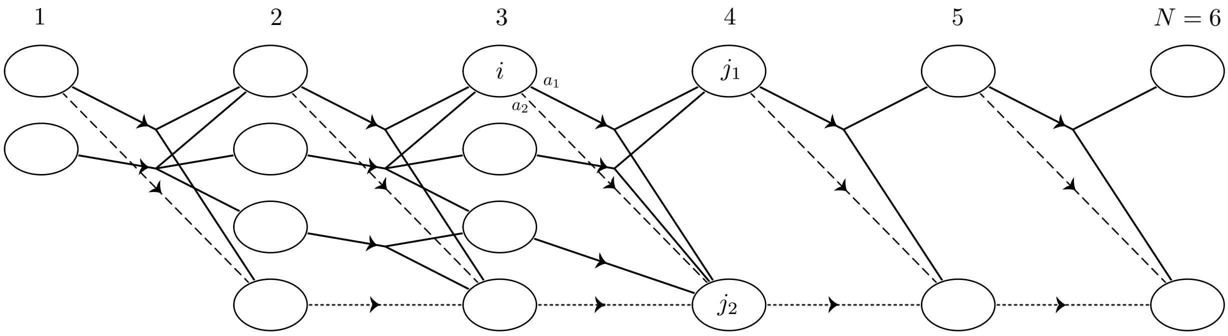


Figure 1.1: A semi-MDP with time-horizon $N = 6$ illustrated using a state-expanded hypergraph. At stage n each node corresponds to a state in I_n. The hyperarcs correspond to actions, e.g. if the system at stage 3 is in state i, then there are two possible actions. Action a^1 (solid line) results in a transition to either state j_1 or j_2 with a given probability.

A semi-MDP can be illustrated using a *state-expanded hypergraph* (Nielsen and Kristensen, 2006). An example is given in Figure 1.1 that illustrates a semi-MDP with time-horizon $N = 6$. At stage n each node corresponds to a state in I_n. For example, there are 2 states at stage 1 and four states at stage 3. Each *hyperarc* corresponds to an action, e.g. if the system at stage $n = 3$ is in state i, then there are two possible actions. Action a_1 (solid line) results in a transition to either state j_1 or j_2 with a probability $\Pr(j_1 \mid i, a_1, n)$ and $\Pr(j_2 \mid i, a_1, n)$, respectively. Action a_2 (dashed line) is a deterministic action with a transition from i to j_2. A policy corresponds to choosing a single hyperarc/action out of each node/state. Note that given a policy, the expected reward until termination can be found by assigning rewards and transition probabilities to all hyperarcs, terminating rewards to the nodes at stage 6, and then calculating (1.1) in the opposite direction of the hyperarcs.

Hierarchical Markov decision processes

A *hierarchical Markov decision process* (*HMDP*) is an extension of a semi-MDP where a series of finite-horizon semi-MDPs are combined into one infinite time-horizon process at the founder level called the founder process (Kristensen, 1988; Kristensen and Jørgensen, 2000). The idea is to expand stages of a process to so-called child processes, which again may expand stages further to new child processes leading to multiple levels. At the lowest level (Level 2 in Figure 1.2) the HMDP consists of a set of semi-MDPs.

Consider Figure 1.2 which illustrates a stage of a three-level HMDP. At the first level, a single *founder process* ρ^0 is defined. Index 0 indicates that the process has no ancestral processes. We

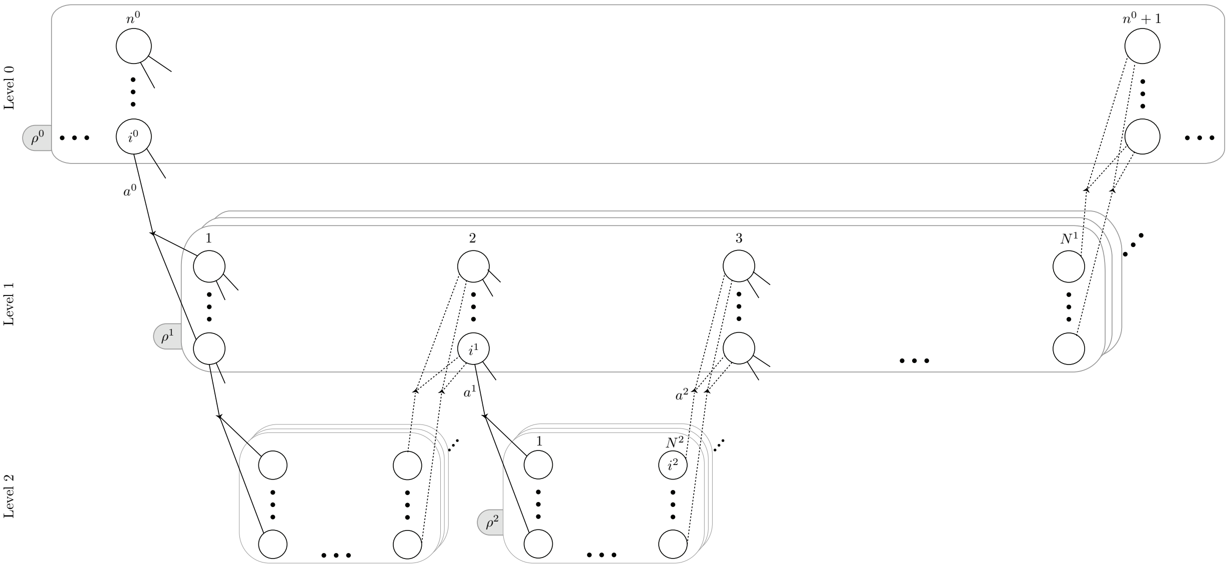


Figure 1.2: An illustration of a stage in an HMDP (Pourmoayed et al., 2016). At the founder level (Level 0) we have a single infinite-horizon founder process ρ^0. A child process, such as ρ^1 at Level 1 (oval box), is uniquely defined by a given stage, state (node), and action (hyperarc) of its parent process and linked with the parent process using its initial probability distribution (solid lines) and its terminating actions (dashed lines). Each process at level 2 is a semi-MDP (see Figure 1.1 for zoom).

assume that ρ^0 is running over an infinite number of stages and that all stages have identical state and action spaces and hence just a single stage is illustrated in Figure 1.2. Moreover, we may skip the argument referring to stage number when we define rewards, stage length, etc. of the founder process.

Let ρ^{l+1} denote a *child process* at level $l+1$, ρ^{l+1} is uniquely defined by a given stage n^l, state i^l and action a^l of its *parent process* ρ^l. For instance, the semi-MDP ρ^2 in Figure 1.2 is defined at stage $n^1 = 2$, state i^1 and action a^1 of the process ρ^1 symbolized by the notation $\rho^2 = (\rho^1 \parallel (n^1, i^1, a^1))$. Each process is connected to its parent and child processes using *jump actions* which can be divided into two groups, namely, a *child jump action* that represents an *initial probability distribution* of transitions to a child process or a *parent jump action* that represents a *terminating probability distribution* of transitions to a parent process. This is illustrated in Figure 1.2 for process ρ^1 where child jump action a_1 represents a transition to the child process ρ^2 with initial probabilities $\Pr\left(\cdot \mid n^1, i^1, a^1\right)$ and immediate expected reward $r_{n^1}(i^1, a^1)$. The expected length of the transition is $u_{n^1}(i^1, a^1)$ representing the expected time from the decision is taken until the child process starts. Moreover, parent jump action a^2 represents termination of the process ρ^2, with terminating probabilities $\Pr\left(\cdot \mid N^2, i^2, a^2\right)$ and expected reward $r_{N^2}(i^2, a^2)$, and a transition (of length $u_{N^2}(i^2, a^2)$) back to the next stage of its parent process ρ^1. The same holds for process $\rho^1 = (\rho^0 \parallel (n^0, i^0, a^0))$ defined for triple (n^0, i^0, a^0) with child jump action a^0. Moreover, note that a stage of process ρ^1 now is defined using its child process. Given the comments above and Figure 1.2, some observations can be made:

1) Jump actions are like traditional actions associated with an expected reward, action length and a set of transition probabilities.
2) A policy R is a decision rule that assigns to each state in a process a (jump) action which means that choosing a policy corresponds to choosing a single hyperarc out of each node in Figure 1.2.
3) Given a policy R and a semi-MDP, e.g. ρ^2 at the lowest level, the terminal reward of a state i^2 at stage N^2 equals the expected reward $r_{N^2}(i^2, a^2)$ of the chosen parent jump action a^2 given R. Hence, we can calculate the expected reward until termination of process ρ^2 using recursive equations (1.1) and, the total expected reward of the process can be found using the initial reward $r_{n^1}(i^1, a^1)$ and distribution $\Pr\left(\cdot \mid n^1, i^1, a^1\right)$ of child jump action a^1 and equation (1.2).
4) The reward at a stage of a parent process equals the total expected rewards of the corresponding child process under the policy chosen. For instance, the reward of choosing action a^1 in state i^1 at stage 2 in process ρ^1 equals the total expected reward (1.2) of process ρ^2. In fact, for the founder process ρ^0 the reward at stage n^0 of a given action a^0 can be calculated by assigning zero terminal rewards to the states/nodes at stage $n^0 + 1$ and calculating the expected reward of each state/node by processing the nodes in the opposite direction of the actions/hyperarcs in Figure 1.2.
5) The above arguments can also be used to calculate the transition probabilities and the stage length of an action at a stage of a parent process.

Let $Z_R(t)$ denote the total reward received up to time t of the founder process ρ^0 under policy R, and suppose the optimality criterion is to maximize the *expected reward per time unit*

$$z(R) = \lim_{t \to \infty} \frac{\mathbb{E}(Z_R(t))}{t}. \tag{1.3}$$

We assume that the Markov chain associated with ρ^0 under policy R has no two closed disjoint sets and hence $z(R)$ will be independent of the starting state of process ρ^0 (Tijms, 2003, Sec. 6.1).

Other criteria, such as total expected discounted reward, can easily be used instead of $z(R)$. The optimal policy maximizing (1.3) can be found using modified policy iteration. Given the current policy, the total expected reward per time unit z and the relative values v_i can be computed as the unique solution to the set of linear equations

$$v_i = \begin{cases} r(i,R(i)) - zu(i,R(i)) + \sum_{j \in I^0} \Pr(j \mid i,R(i))\, v_j, & i \in I^0 \setminus \{s\} \\ 0, & i = s, \end{cases} \tag{1.4}$$

where s is an arbitrary chosen state and I^0 denote the set of states at the founder level. For an HMDP the rewards $r(i,R(i))$, transition probabilities $\Pr(j \mid i,R(i))$, and stage length $u(i,R(i))$ at the founder process are not given explicitly. However, they can be calculated as observed above. As a result, the optimal policy of the HMDP (i.e. the founder process) can be found using a modified policy iteration algorithm which updates the current policy based on the relative values v_i using a value iteration approach (see observations above), and afterwards solving the equations in (1.4) given the new updated policy. These steps are repeated until the policy cannot be improved.

Since the number of states at the founder level is generally lower compared to modeling the problem as a semi-MDP, larger models can be solved because we do not have to solve such a large system of equations as if used policy iteration on the semi-MDP. For more details about the optimization technique, the interested reader can refer to Kristensen and Jørgensen (2000).

1.2.2 Approximate dynamic programming

Approximate dynamic programming (*ADP*) is a solution procedure for solving large Markov decision models. That is, ADP is used as a solution procedure to find an approximate optimal policy of a *Markov decision process* (*MDP*) with large state and action spaces (an MDP is a semi-MDP with equal stage lengths).

Consider an infinite-horizon MDP under the discounted reward criterion with stationary state and action spaces (see e.g. Puterman, 2005, Chap. 6). Here the maximum expected discounted reward can be found using the value function $\nu(i)$ satisfying the following optimality equations (Puterman, 2005, Sec. 6.2)

$$\nu(i) = \max_{a \in \mathbb{A}(i)} \left(r(i,a) + \gamma \mathbb{E}\left(\nu(j)\right) \right), \ \forall i \in I, \tag{1.5}$$

where γ is the discount factor, and the next random state j is obtained using the probabilistic transition function $\phi(i,a,\omega)$ given action $a \in \mathbb{A}(i)$ and random information ω received between the current and next decision epoch. In order to find optimal actions, the optimality equations in (1.5) should be solved for all states $i \in I$. However, the calculation of the value function $v(i)$ for all states may not always be easy. First, the size of the state and action space I and $\mathbb{A}(i)$ may be too large, i.e. computation of $v(i)$ for every possible state is impossible. Second, due to the large number of states and possible outcomes of random information ω, an exact computation of the expected value in (1.5) may be prohibitive. Finally, due to the expected value operator, the maximization problem in (1.5) is not deterministic and hence it may be difficult to be solved. These computational challenges are known as the three *curses of dimensionality* (Powell, 2007, Section 4.1) and prevent us from applying regular solution procedures of MDPs (e.g. value iteration) to solve the model.

ADP is an efficient method to deal with these computational problems and to find an approximate solution instead. The main idea is to approximate the value function $v(i)$ using a parametric function and use simulation to find the states that are most likely observed in the system. Based on the current estimation of the value function, a deterministic version of (1.5) is solved using e.g. linear or integer programming. This procedure is repeated until the parameters of the approximated value function are converged to a fixed set of values. For more details about ADP algorithms, the interested reader may refer to Powell (2007).

1.2.3 State space models

In animal production, time-series of data from online monitoring are often available. Online monitoring is a relevant method to obtain data for tracking changes and can be done regularly by sensors placed in the production unit.

A *state space model* (*SSM*) is a statistical model which may be used to transform large datasets obtained using online sensors into information about the production (West and Harrison, 1997). An SSM consists of a set of latent variables and a set of observed variables. At a specified point in time, the conditional distribution of the observed variables is a function of the latent variables specified via the observation equations. The latent variables change over time as described via the system equations. The observations are conditionally independent given the latent variables. Thus the value of the latent variables at a time point may be considered as the state of the system, and the SSM framework makes it possible to predict the latent variables/state of the system via the observed variables, both the current state and the future development in the state variables. In an SSM, when normality and linearity conditions of variables and equations are valid, *Bayesian* *updating* can be used to update the posterior distribution of the latent variables. That is, we use the Kalman filter to update our forecast when new data arrives (West and Harrison, 1997, page 103). Examples of SSMs applied to agricultural problems can be found in Cornou et al. (2008) and Bono et al. (2012, 2013).

In this dissertation, SSMs are applied to time-series of weight and feed intake data from the herd and market prices to describe the uncertainty of weight and price information in the models. These SSMs are categorized into different groups based on the dynamic nature of the considered system and the probability distribution assumed for the initial data. More precisely, two kinds of SSMs are considered and later embedded into the (H)MDPs. In the first type, the probability distribution of the observations, related to the weight and price data, is Gaussian and in the second type, these observations come from a non-Gaussian distribution. In both models, the dynamics of the system is modeled by linear equations.

1.3 Structure and contributions of the dissertation

The dissertation consists of three self-contained chapters, each one is presented as a journal paper. Each paper addresses the challenges mentioned in Section 1.1 from different points of view. In all papers, (H)MDPs with an infinite time horizon are the main modeling technique used to model marketing and feeding decisions. Moreover, SSMs based on Bayesian updating are the main statistical technique to describe the uncertainty of weight and price information in the models. A short description of each paper is given below.

Paper one focuses on combining marketing and feeding decisions at pen level. A three-level HMDP is formulated to represent the decision process. The model considers online measurements of pig weight and feed intake from a set of sensors in the pen. A Bayesian approach is used to update the state of the system such that it contains the relevant information based on the previous measurements. More precisely, two types of SSMs known as a Gaussian state space model (GSSM) and a non-Gaussian state space model (nGSSM) are applied to model weight, growth, and the inhomogeneity of the pigs in the pen. These models are embedded into the HMDP, i.e. transition probabilities of state variables in the three-level HMDP are calculated using Bayesian updating of the GSSM and nGSSM. A computational study is performed to show how the optimal policy adapts to different conditions at pen level. Moreover, a small sensitivity analysis is carried out to show the importance of the length of the growing period and the feed-mix cost.

Paper two focusses on fluctuating pork, feed, and piglet prices and shows how price fluctuations can affect the marketing decisions of finishing pigs. A two-level HMDP is used to model marketing decisions under price fluctuations of pork, feed, and piglet. Using available time-series of prices, three SSMs use Bayesian updating to forecast future prices in the model according to previous market prices. Moreover, we exploit a random regression model to model the evolution of pig weight during the growing period in the pen. Three scenarios with different patterns of price fluctuations are described to illustrate how marketing decisions are different when there is an increasing/decreasing trend in pork, feed, and piglet market prices. Furthermore, we calculate the value of including price information into the model by comparing the optimal policy of the two-level HMDP under both fluctuating and fixed prices.

Paper three considers marketing decisions at herd level and evaluates the effect of cross-level constraints (termination and transportation constraints) on the marketing policy at herd level. Marketing decisions are modeled using a discounted infinite-horizon MDP, and SSMs from the first paper are exploited to describe the uncertainty of weight information in the pens. Due to the curse of dimensionality, ADP is applied to solve the MDP model and find an approximate marketing policy at herd level. More precisely, we use the structure of the value function at pen level (found in the first paper) to approximate a parametric function for the value function at herd level. Parameters of the approximated value function are estimated using an approximate value iteration algorithm exploiting simulation and integer programming techniques. In order to validate the quality of the solution found by ADP, we first compare the ADP solution with the solution found using value iteration at pen level. We also provide an example at herd level to show how ADP finds the best marketing decisions under different conditions and why early termination and transportation costs are important at herd level. Finally, the marketing policy obtained using ADP is compared with other well-known marketing policies often applied at herd level.

The contributions of the three papers are significant. The first paper is the first published paper considering a sequential decision model taking both feeding and marketing decisions into account simultaneously at pen level. The SSMs developed in this paper consider the inhomogeneity of animals as regards growth and feed intake, which result in a more realistic estimation of weight information compared to other papers in the literature. The presented model can be used as a part of a decision support system with online data where the system state can be identified using Bayesian updating and the optimal policy of the HMDP can be used to execute the best feeding and marketing decision at pen level. One may argue that the model also has some limitations. For instance, possible constraints at section and herd levels are ignored (e.g. constraints on the transportation of culled pigs to the abattoir) and the price fluctuations of pork, feed, and piglet are not considered in the model. These limitations are covered in the second and third papers.

The second paper uses a novel method to embed price information into an optimization model using Bayesian updating, i.e. the price parameters are dynamically updated based on historical data. According to the related literature, this is the first paper taking into account fluctuations of pork, feed and piglet prices when considering marketing decisions given a high range of possible price values.

The third paper contributes significantly to the literature by considering sequential marketing decisions at herd level and to the best of my knowledge it is the first paper to simultaneously take into account decisions and constraints at pen, section, and herd levels. Due to the curse of dimensionality this problem has not be solved before. However the high-dimensional MDP can be addressed and solved by using ADP, i.e. approximating the value function.

The models developed in this dissertation can be extended to handle decisions about diseases by incorporating state variables related to the health of pigs. Moreover, factors such as the mortality rate of pigs and labor costs (if not considered fixed) can be modeled by modifying the transition probabilities and reward of an action.

References

C. Bono, C. Cornou, and A.R. Kristensen. Dynamic production monitoring in pig herds i: Modeling and monitoring litter size at herd and sow level. *Livestock Science*, 149(3):289–300, 2012. doi:10.1016/j.livsci.2012.07.023 .

C. Bono, C. Cornou, S. Lundbye-Christensen, and A.R. Kristensen. Dynamic production monitoring in pig herds ii. modeling and monitoring farrowing rate at herd level. *Livestock Science*, 155(1):92–102, 2013. doi:10.1016/j.livsci.2013.03.026 .

J.P. Christensen. *The basics of Pig Production*. Landbrugsforlaget, 2010. ISBN 978 87 91566 33 2.

C. Cornou, J. Vinther, and A.R. Kristensen. Automatic detection of oestrus and health disorders using data from electronic sow feeders. *Livestock Science*, 118(3):262–271, 2008.

Danish Agriculture and Food Council. Statistics 2014 pigmeat. Technical report, 2015. URL http://www.agricultureandfood.dk/Prices_Statistics/Annual_Statistics.aspx.

A.R. Kristensen. Hierarchic markov processes and their applications in replacement models. *European Journal of Operational Research*, 35(2):207–215, 1988. doi:10.1016/0377-2217(88)90031-8 .

A.R. Kristensen and E. Jørgensen. Multi-level hierarchic markov processes as a framework for herd management support. *Annals of Operations Research*, 94(1-4):69–89, 2000. doi:10.1023/A:1018921201113 .

Landbrug & Fødevarer. Udenrigshandel 2010 - 2014 fødevareklyngen – varer til hele verden. Technical report, 2015. URL https://www.lf.dk/Tal_og_Analyser/Fodevareklyngens_udenrigshandel.aspx.

L.R. Nielsen and A.R. Kristensen. Finding the K best policies in a finite-horizon Markov decision process. *European Journal of Operational Research*, 175(2):1164–1179, 2006. doi:10.1016/j.ejor.2005.06.011 .

R. Pourmoayed and L.R. Nielsen. An overview over pig production of fattening pigs with a focus on possible decisions in the production chain. Technical Report PigIT Report No. 4, Aarhus University, 2014. URL http://pigit.ku.dk/publications/PigIT-Report4.pdf.

R. Pourmoayed, L. R. Nielsen, and A. R. Kristensen. A hierarchical Markov decision process modeling feeding and marketing decisions of growing pigs. *European Journal of Operational Research*, 250(3):925–938, 2016. doi:10.1016/j.ejor.2015.09.038 .

W.B. Powell. *Approximate Dynamic Programming: Solving the curses of dimensionality*, volume 703. Wiley-Interscience, 2007. ISBN 978-0-470-60445-8.

M. L. Puterman. *Markov decision processes: discrete stochastic dynamic programming*. John Wiley & Sons, 2005. ISBN 978-0-471-72782-8.

H.C. Tijms. *A first course in stochastic models*. John Wiley & Sons Ltd, 2003. ISBN 978-0-471-49880-3.

M. West and J. Harrison. *Bayesian Forecasting and Dynamic Models (Springer Series in Statistics)*. Springer-Verlag, February 1997. ISBN 0387947256.

Chapter 2

Paper I: A hierarchical Markov decision process modeling feeding and marketing decisions of growing pigs

History: *This paper was prepared in collaboration with Lars Relund Nielsen and Anders Ringgaard Kristensen. It has been published in European Journal of Operational Research, 250(3):925-938, 2016. The paper has been presented at: IFORS 2014 - Conference of the International Federation of Operational Research Societies, July 2014, Barcelona, Spain; EAAP 2014 - Annual Meeting of the European Federation for Animal Science, August 2014, Copenhagen, Denmark and Workshop on OR in Agriculture and Forest Management, July 2014, Lleida, Spain.*

A hierarchical Markov decision process modeling feeding and marketing decisions of growing pigs

Reza Pourmoayed and Lars Relund Nielsen

CORAL, Department of Economics and Business Economics, Aarhus University, Fuglesangs Allé 4, DK-8210 Aarhus V, Denmark.

Anders Ringgaard Kristensen

HERD, Department of Large Animal Sciences, University of Copenhagen, Grønnegårdsvej 2, DK-1870 Frederiksberg C, Denmark.

Abstract: Feeding is the most important cost in the production of growing pigs and has a direct impact on the marketing decisions, growth and the final quality of the meat. In this paper, we address the sequential decision problem of when to change the feed-mix within a finisher pig pen and when to pick pigs for marketing. We formulate a hierarchical Markov decision process with three levels representing the decision process. The model considers decisions related to feeding and marketing and finds the optimal decision given the current state of the pen. The state of the system is based on information from on-line repeated measurements of pig weights and feeding and is updated using a Bayesian approach. Numerical examples are given to illustrate the features of the proposed optimization model.

Keywords: OR in agriculture; stochastic programming; hierarchical Markov decision process; herd management; Bayesian updating.

2.1 Introduction

In production systems of growing pigs, feeding is the most important operation and has a direct influence on the cost and the quality of the meat. Another important operation is the timing of *marketing*. It refers to a sequence of culling decisions until the production unit is empty. As a result the profit of the production unit is highly dependent on the feeding cost and on good timing of marketing, i.e. decisions about feeding and marketing have a direct impact on profit.

In a production system of growing/finishing pigs (Danish standards), the animals may be considered at different levels: herd, section, pen, or animal. The herd is a group of sections, a section includes some pens, and a finisher pen involves some animals (usually 15-20). New piglets are transferred to a weaner unit approx. four weeks after birth, and they stay for approx. eight weeks until they weigh approx. 30 kg. The pigs are then moved to a finisher pen where they grow until marketing (9-12 weeks). In the finisher pen, the farmer should determine which pigs should be selected for slaughter (individual marketing). The reward of marketing a pig depends on the unit meat price of the carcass weight and the leanness of the pig. The meat price is highest if the carcass weight of the pig lies in a specific interval. Next, after a sequence of individual marketings, the farmer must decide when to *terminate* (empty) the rest of the pen. Terminating a pen means that the remaining pigs in the pen are sent to the slaughterhouse (in one delivery) and after cleaning the pen, another group of piglets (each weighing approx. 30 kg) is inserted into the pen and *the production cycle* is repeated. That is, the farmer must time the marketing decisions while simultaneously considering the carcass weight in relation to the best interval, the leanness, and the length of the production cycle. For an extended overview over pig production of growing pigs, see Pourmoayed and Nielsen (2014a).

The growth and leanness of the pigs will be highly dependent on the feed given. Phase feeding is a common method used in the production of the growing pigs. In the finisher pen the growing period typically includes 3 or 4 phases and each phase involves a predefined *feed-mix* which is a mixture of different ingredients (barley, soy, maize, etc.). A relevant decision is when to change the current feed-mix (transition to a new phase) and what type of feed-mix to use in the next phase.

Since the choice of feed-mix affects the pigs' growth, a specific feeding strategy has an impact on the marketing strategy. That is, the economic optimization of feeding and marketing decisions is interrelated and requires a simultaneous analysis. Consequently, a sequential decision model is needed that considers both feeding and marketing decisions. To the best of our knowledge, there are only a few studies that take into account these decisions simultaneously (Niemi, 2006; Sirisatien et al., 2009). However, these studies consider the problem at animal level and do not take into account the inhomogeneity of animals in growth and feed intake. The aim of this paper is to close this gap and consider the problem at pen level instead.

In this paper we formulate a hierarchical Markov decision process which takes into account decisions related to feeding and marketing of growing pigs at pen level. We assume that the production is cyclic, i.e. when the pen is emptied, not only a regular state transition takes place, but rather the process (the current batch of pigs) is restarted.

The model considers time series of pig weights and feeding obtained from online monitoring, e.g. from a set of sensors in the pen. A Bayesian approach is used to update the state of the system such that it contains the relevant information based on the previous measurements. More precisely, two state space models for *Bayesian forecasting* (West and Harrison, 1997) are used to update the estimates of live weights and feed intake on a weekly basis.

The structure of the paper is as follows. First, Section 2.2 gives a short literature review. Second, a detailed description of the optimization model is given in Section 2.3. Next, Section 2.4 presents the statistical models which are embedded into the model. In Section 2.5, numerical examples are considered to show the functionality of the proposed optimization model. Finally, conclusions and directions for further research are given in Section 2.6.

2.2 Literature review

Due to the dynamic nature of the production environment of growing pigs, the marketing and feeding decisions are sequential, complex and hard to optimize. Various models have been considered to deal with this complexity.

Some studies consider only the marketing decisions. Chavas et al. (1985) applied the concepts of optimal control theory to find the optimal time of marketing of individual animals. Jørgensen (1993) used a dynamic programming approach to optimize a given heuristic framework for delivering the pigs to the slaughterhouse. Boland et al. (1993) considered the optimal slaughter pig marketing problem under different pricing models and for each pricing system, they found the optimal slaughter weight. Kure (1997) considered the problem at batch level and used the replacement theory concepts and a recursive dynamic programming method to determine the optimal time of marketing the pigs. Toft et al. (2005) optimized both marketing and treatment decisions (e.g. regarding vaccination for disease problems) using a *hierarchical Markov decision process* (*HMDP*). Boys et al. (2007) implemented a simulation approach to determine the best marketing strategy to utilize full truck capacity for delivering the pigs to the packers. In the study by Ohlmann and Jones (2008), a mixed integer programming model was proposed to find the best marketing strategy under an annual profit criterion. Kristensen et al. (2012) suggested a two-level HMDP to find the best marketing strategy according to the data from an online monitoring system.

Other studies focus on sequential feeding decisions, i.e. finding the best strategy for choosing the appropriate feed-mix during the growing period of animals. One example is Glen (1983) who proposed a dynamic programming approach to determine the sequence of feed-mixes in the production unit. In the study by Boland et al. (1999), a linear programming approach was used to specify the optimal time of changing the feed-mix and also the optimal nutrient ingredients of the feed-mix. A genetic algorithm was applied by Alexander et al. (2006) to find the best nutrient components of each feed-mix.

Only a few studies take both marketing and feeding decisions into account. Niemi (2006) used a mechanistic function to model the animal growth trend during the growing period. Niemi (2006) further applied a stochastic dynamic programming method to find the best nutrient ingredients and also the best time of marketing. In the study by Sirisatien et al. (2009), a genetic algorithm was used. Each iteration resulted in a set of feeding schedules followed by the optimal values of the nutrient ingredients and feeding period. Both studies considered the problem at animal level and did not take into account the inhomogeneity of animals with respect to growth and feed conversion rate.

Markov decision models are a well-known modeling technique within animal science used to model livestock systems. See for instance Rodriguez et al. (2011) and Nielsen et al. (2010). For a recent survey see Nielsen and Kristensen (2014), which cites more than 100 papers using (hierarchical) Markov decision processes to model and optimize livestock systems. An HMDP is an extension of a semi Markov decision process (semi-MDP) where a series of finite-horizon semi-MDPs are combined into one process at the founder level called the main process (Kristensen, 1988; Kristensen and Jørgensen, 2000). As a result the state space at the founder level can be reduced and larger models can be solved using a modified policy iteration algorithm under different criteria (Nielsen and Kristensen, 2014). Modeling the problem using an HMDP compared to a semi-MDP contributes to reducing the *curse of dimensionality*, since the total number of state variables can be decreased. Moreover, the total number of states at the founder level is lower (i.e. the matrix which must be inverted in the modified policy iteration algorithm is much smaller).

A *state space model* (*SSM*) (West and Harrison, 1997) is a statistical model which may be used to transform large datasets obtained using online sensors into the required information about the production process. An SSM consists of a set of latent variables and a set of observed variables. At a specified point in time the conditional distribution of the observed variables is a function of the latent variables specified via the observation equations. The latent variables change over time as described via the system equations. The observations are conditionally independent given the latent variables. Thus the estimated value of the latent variables at a time point may be considered as the state of the system, and with Bayesian forecasting (the Kalman filter) we can estimate the latent variables/real state of the system via the observed variables. Examples of SSMs applied to agricultural problems are Cornou et al. (2008); Bono et al. (2012) and Bono et al. (2013). Moreover, an SSM can be discretized and embedded into an HMDP (Nielsen et al., 2011).

2.3 Model description

Our pig marketing and feeding problem is modeled using a *hierarchical Markov decision process* (*HMDP*) with three levels. A short introduction to HMDPs is given below. As techniques from both statistical forecasting and operations research are used, consistent notation can be hard to specify. To assist the reader, Appendix 2.A provides an overview.

An HMDP is an extension of a semi-MDP where a series of finite-horizon semi-MDPs are combined into one infinite time-horizon process at the founder level called the *founder process* (Kristensen and Jørgensen, 2000). The idea is to expand the stages of a process to so-called *child processes*, which again may expand stages further to new child processes leading to multiple levels. At the lowest level the HMDP consists of a set of finite-horizon semi-MDPs (see e.g. Tijms, 2003, Chap. 7). All processes are linked together using jump actions (see Figure 2.1).

A *finite-horizon semi-MDP* considers a sequential decision problem over $\mathcal{N}$ *stages*. Let $\mathbb{I}_n$ denote the finite set of system states at stage n. When *state* $i \in \mathbb{I}_n$ is observed, an *action* a from the finite set of allowable actions $\mathbb{A}_n(i)$ must be chosen, and this decision generates *reward* $r_n(i,a)$. Moreover, let $u_n(i,a)$ denote the *stage length* of action a, i.e. the expected time until the next decision epoch (stage $n+1$) given action a and state i. Finally, let $\Pr(j \mid n,i,a)$ denote the *transition probability* of obtaining state $j \in \mathbb{I}_{n+1}$ at stage $n+1$ given that action a is chosen in state i at stage n.

An HMDP with three levels is illustrated in Figure 2.1 using a *state-expanded hypergraph* (Nielsen and Kristensen, 2006). At the first level, a single *founder process* $\mathfrak{p}^0$ is defined. Index 0 indicates that the process has no ancestral processes. We assume that $\mathfrak{p}^0$ is running over an infinite number of stages and that all stages have identical state and action spaces and hence just a single stage is illustrated in Figure 2.1. Let $\mathfrak{p}^{l+1}$ denote a *child process* at level $l+1$. Process $\mathfrak{p}^{l+1}$ is uniquely defined by a given stage n^l, state i^l and action a^l of *parent process* $\mathfrak{p}^l$. For instance, the semi-MDP $\mathfrak{p}^2$ in Figure 2.1 is defined at stage $n^1 = 2$, state i^1 and action a^1 of the process $\mathfrak{p}^1$ symbolized by the notation $\mathfrak{p}^2 = (\mathfrak{p}^1 \parallel (n^1, i^1, a^1)\ )$. Each process is connected to its parent and child processes using *jump actions* which can be divided into two groups, namely, a *child*

Figure 2.1: An illustration of a stage in an HMDP. At the founder level (Level 0) we have a single infinite-horizon founder process $\mathfrak{p}^0$. A child process, such as $\mathfrak{p}^1$ at Level 1 (oval box), is uniquely defined by a given stage, state (node), and action (hyperarc) of its parent process and linked with the parent process using its initial probability distribution (solid lines) and its terminating actions (dashed lines). Each process at level 2 is a semi-MDP. Note that only a subset of the actions is drawn.

jump action that represents an *initial probability distribution* of transitions to a child process or a *parent jump action* that represents a *terminating probability distribution* of transitions to a parent process. This is illustrated in Figure 2.1 for process $\mathfrak{p}^1$ where child jump action a^1 (illustrated using a solid hyperarc) represents a transition to the child process $\mathfrak{p}^2$ and parent jump action a^2 (illustrated using a dashed hyperarc) represents termination of the process $\mathfrak{p}^2$. Like traditional actions, jump actions are associated with an expected reward, action length, and a set of transition probabilities. Each node in Figure 2.1 at a given stage n of a process $\mathfrak{p}^l$ corresponds to a state in $\mathbb{I}^l_n$. For example, there are 3 states at stage 1 in process $\mathfrak{p}^2$. Similarly each *hyperarc* corresponds to an action, e.g. action a (gray hyperarc) results in a transition to either state j_1 or j_2.

A policy is a decision rule/function that assigns to each state in a process a (jump) action. That is, choosing a policy corresponds to choosing a single hyperarc out of each node in Figure 2.1. Given a policy, the reward at a stage of a parent process equals the total expected rewards of the corresponding child processes. For instance, in Figure 2.1, the reward of choosing action a^1 in state i^1 at stage $n^1 = 2$ in process $\mathfrak{p}^1$ equals the total expected reward of process $\mathfrak{p}^2$. A similar approach can be used to calculate the transition probabilities and the stage length of an action at a stage of a parent process.

Different optimality criteria may be considered. In this paper, our optimality criterion is to maximize the *expected reward per time unit* and the optimal policy of the HMDP can be found using a modified policy iteration algorithm. For a detailed description of the algorithm, the interested reader may consult Nielsen and Kristensen (2014).

2.3.1 Assumptions

Consider the problem of optimizing feeding and marketing decisions in a finisher pig pen. The problem can be modeled as a three-level semi-HMDP under the following assumptions:

- $q^{\max}$ pigs are inserted into the finisher pen;
- a finite set of feed-mixes $\mathbb{F}$ is available and feed-mix $f \in \mathbb{F}$ cannot be changed before it has been used for at least $t_f^{\min}$ weeks (for simplicity, $t_f^{\min}$ is the same for all feed-mixes);
- at most $b^{\max}$ feeding phases can be used;
- marketing of pigs is started in week $t^{\min}$ at the earliest;
- the pen is terminated in week $t^{\max}$ at the latest, i.e. the maximum life time of a pig in the pen is $t^{\max}$;
- the growth of a pig is independent of the other pigs in the pen, i.e. the growth is not dependent on the number of pigs in the pen;
- weekly deliveries to the abattoir in the marketing period are based on a cooperative agreement where culled pigs from each pen are grouped in one transportation delivery at a fixed time each week, i.e. the transportation cost is fixed.

To give a complete description of the three-level HMDP with feeding and marketing decisions, each semi-MDP must be specified at all levels, i.e. stages, states, and (jump) actions including the corresponding rewards, stage lengths (measured in weeks), and transition probabilities.

2.3.2 Stages, states and actions

As illustrated in Figure 2.1, the founder process of the HMDP is an infinite time-horizon process where a stage represents a life of $q^{\max}$ pigs inserted into the pen (until termination). A stage of the process at the second level corresponds to a feeding phase in which the pigs are fed a specific feed-mix f. Finally, a stage at the third level is a week of the current production cycle in the pen under the specific feed-mix. The length, stage, states, and (jump) actions of each process at the different levels are described below. Whenever, the level is clear from the context, the superscript indicating the current level under consideration will be left out to avoid heavy notation.

Level 0 - Founder process $\mathfrak{p}^0$

Stage: A production cycle of $q^{\max}$ pigs, i.e. from inserting the piglets into the pen until terminating the pen.

Time horizon: Infinite (since the number of filling and emptying a pen is infinite).

States: A single state representing the start of a production cycle ($\mathbb{I} = \{i^0\}$).

Actions: One child jump action a^0 representing insertion of a new group of piglets ($\mathbb{A}(i^0) = \{a^0\}$).

Level 1 - Parent process $\mathfrak{p}^1 = (\mathfrak{p}^0 \parallel (n^0, i^0, a^0)\ )$

Stage: A feeding phase with a given feed-mix.

Time horizon: Given a maximum of $b^{\max}$ feeding phases, the maximum number of stages in process $\mathfrak{p}^1$ is $\mathscr{N} = b^{\max} + 1$ since a dummy stage is added at the end.

States: First, consider stage/feeding phase $2 \leq n \leq b^{\max}$. A state i is defined using the following state variables:

f_n: previous feed-mix (feed-mix in stage/phase $n-1$);

t_n: starting time of phase (week);

q_n: number of pigs in the pen at the beginning of stage/phase n;

$\mathbb{w}_n$: model information related to the weight of the pigs, obtained using Bayesian updating ($\mathbb{w}_n \in \mathbb{W}_n$). Section 2.4 provides details on the way the information is obtained.

Furthermore, at this level, a dummy state $\tilde{i}$ is added to represent pen termination. Note that due to the model assumptions, the earliest starting time of phase n is $(n-1)t_f^{\min} + 1$. Moreover, if $t_n \leq t^{\min}$ then $q_n = q^{\max}$. Hence the set of states becomes

$$\mathbb{I}_n = \{i = (f_n, t_n, q_n, \mathbb{w}_n) \mid f_n \in \mathbb{F}, t_n \in \{(n-1)t_f^{\min} + 1, \ldots, t^{\max} - 1\},$$
$$q_n \in \{q^{\max}\mathbf{I}_{\{t_n \leq t^{\min}\}} + \mathbf{I}_{\{t_n > t^{\min}\}}, \ldots, q^{\max}\}, \mathbb{w}_n \in \mathbb{W}_n\} \cup \{\tilde{i}\},$$

where $\mathbf{I}_{\{\cdot\}}$ denotes the indicator function.

Next, consider stage $n=1$. Here the number of states to $\mathbb{I}_n = \mathbb{W}_n$ can be reduced, since $t_n = 1$, $q_n = q^{\max}$, and there is no previous feed-mix.

Finally, at the dummy stage ($n = \mathcal{N}^1$), only the dummy state $\tilde{i}$ representing pen termination is defined.

Actions: At stage $n=1$, it is possible to choose a feed-mix $f \in \mathbb{F}$ at state $i = \mathbb{w}_n$, i.e. the set of child jump actions is $\mathbb{A}_n(i) = \{a_f \mid f \in \mathbb{F}\}$. At the subsequent stages ($1 < n < \mathcal{N}^1$), possible child jump actions at state $i = (f_n, t_n, q_n, \mathbb{w}_n)$ are $\mathbb{A}_n(i) = \{a_f \mid f \in \mathbb{F} \setminus \{f_n\}\}$. The length of all child jump actions choosing a feed-mix is zero. In the dummy state $\tilde{i}$ a single dummy parent jump action $\tilde{a}$ with length zero is considered which represents that the pen has been terminated.

Level 2 - MDP $\mathfrak{p}^2 = (\mathfrak{p}^1 \parallel (n^1, i^1, a^1))$

At the lowest level a semi-MDP is defined for each stage/feeding phase n^1, parent state $i^1 = (f_{n^1}, t_{n^1}, q_{n^1}, \mathbb{w}_{n^1})$, and action $a^1 = a_f$ corresponding to choosing feed-mix f.

Stage: A week in the current feeding phase.

Time horizon: A stage is defined for each week $t_{n^1}, \dots, t^{\max}$ and hence the time horizon becomes $\mathcal{N} = t^{\max} - t_{n^1} + 1$. That is, stage $n = 1, \dots, \mathcal{N}$ corresponds to week $t_{n^1} + n - 1$ ($n-1$ weeks since the feed-mix was changed).

States: Given stage n, a state i consists of the following state variables:

q_n: number of pigs in the pen at the beginning of the week;

$\mathbb{w}_n$: model information related to the weight of the pigs, obtained using Bayesian updating ($\mathbb{w}_n \in \mathbb{W}_n$);

$\mathbb{g}_n$: model information related to the growth of the pigs, obtained using Bayesian updating ($\mathbb{g}_n \in \mathbb{G}_n$). Further details on how $\mathbb{g}_n$ and $\mathbb{w}_n$ are obtained, are given in Section 2.4.

A dummy state $\tilde{i}$ is also added to represent pen termination. Therefore the set of states becomes:

$$\mathbb{I}_n = \{i = (q_n, \mathbb{w}_n, \mathbb{g}_n) \mid q_n \in \{q^{\max}\mathbf{I}_{\{t_{n^1}+n-1 \leq t^{\min}\}} + \mathbf{I}_{\{t_{n^1}+n-1 > t^{\min}\}}, \dots, q^{\max}\},$$
$$\mathbb{w}_n \in \mathbb{W}_n, \mathbb{g}_n \in \mathbb{G}_n\} \cup \{\tilde{i}\}.$$

Actions: Consider state $i = (q_n, \mathbb{w}_n, \mathbb{g}_n)$ at stage n. If marketing is not possible at this stage (since $t_{n^1} + n - 1 < t^{\min}$), then the production process continues for another week with the current feed-mix using action $a_{\texttt{cont}}$. If marketing is possible ($t_{n^1} + n - 1 \geq t^{\min}, n < \mathscr{N}^1$), then the set of actions can be expanded to the parent jump action $a_{\texttt{term}}$ where the pen is terminated and actions a_q, which implies that the q heaviest pigs are culled (individual marketing). If $n > t_f^{\min}$ the current feed-mix can be changed, which corresponds to parent jump action $a_{\texttt{newMix}}$. Finally, at the last stage $n = \mathscr{N}$, the pen must be terminated. Hence the set of actions becomes

$$\mathbb{A}_n(i) = \begin{cases} \{a_{\texttt{cont}}\}, & t_{n^1} + n - 1 < t^{\min}, n \leq t_f^{\min}, \\ \{a_{\texttt{cont}}, a_{\texttt{newMix}}\}, & t_{n^1} + n - 1 < t^{\min}, t_f^{\min} < n < \mathscr{N}, \\ \{a_{\texttt{cont}}, a_{\texttt{term}}\} \cup \{a_q \mid 1 \leq q < q_n\}, & t_{n^1} + n - 1 \geq t^{\min}, n \leq t_f^{\min}, \\ \{a_{\texttt{cont}}, a_{\texttt{newMix}}, a_{\texttt{term}}\} \cup \{a_q \mid 1 \leq q < q_n\}, & t_{n^1} + n - 1 \geq t^{\min}, t_f^{\min} < n < \mathscr{N}, \\ \{a_{\texttt{term}}\}, & n = \mathscr{N}. \end{cases}$$

The lengths of actions $a_{\texttt{cont}}$ and a_q are one week while the lengths of actions $a_{\texttt{term}}$ and $a_{\texttt{newMix}}$ are zero. State $\tilde{i}$ has a single dummy parent jump action $\tilde{a}$ of length zero.

2.3.3 Transition probabilities

To complete the formulation of the HMDP, transition probabilities must be specified for all (jump) actions.

Level 0 - Founder process $\mathfrak{p}^0$

Given state i^0 and child jump action a^0 (insertion of a new group of piglets), a transition to state $i^1 = \mathbb{w}_1$ at the first stage ($n^1 = 1$) of process $\mathfrak{p}^1$ happens with probability $\Pr\left(i^1 \mid i^0, a^0\right) = \Pr_0(\mathbb{w}_1)$, where $\Pr_0(\mathbb{w}_1)$ denotes the initial probability of weight information $\mathbb{w}_1$.

Level 1 - Parent process $\mathfrak{p}^1$

Consider state $i = (f_n, t_n, q_n, \mathbb{w}_n)$ and child jump action $a = a_f$ that corresponds to choosing a specific feed-mix $f \in \mathbb{F}$. A transition to state $i^2 = (\tilde{q}_1, \tilde{\mathbb{w}}_1, \tilde{\mathbb{g}}_1)$ at the first stage ($n^2 = 1$) of process $\mathfrak{p}^2$ happens with probability

$$\Pr\left(i^2 \mid n, i, a\right) = \begin{cases} \Pr_0(\tilde{\mathbb{g}}_1 \mid f), & \tilde{q}_1 = q_n,\ \tilde{\mathbb{w}}_1 = \mathbb{w}_n, \\ 0, & \text{otherwise}, \end{cases}$$

where $\Pr_0(\tilde{\mathbb{g}}_1 \mid f)$ denotes the initial probability of growth information for state $\tilde{\mathbb{g}}_1$ given feed-mix f. For dummy state $\tilde{i}$ and parent jump action $\tilde{a}$, a deterministic transition to state i^0 happens.

Level 2 - Semi-MDP $\mathfrak{p}^2 = (\mathfrak{p}^1 \parallel (n^1, i^1, a^1))$

First, consider state $i = (q_n, \mathbb{w}_n, \mathbb{g}_n)$ in process $\mathfrak{p}^2$ starting at week t_{n^1}, given $a^1 = a_f$, i.e. the process uses feed-mix f. At Level 2, two parent jump actions are considered. If the feed-mix is changed ($a = a_{\texttt{newMix}}$), then the process terminates and makes a deterministic transition to state $i^1 = (f, t_{n^1} + n - 1, q_n, \mathbb{w}_n)$ at stage $n^1 + 1$. If the process is terminated using parent jump action $a_{\texttt{term}}$, then the system makes a deterministic transition to state $i^1 = \tilde{i}$ in Level 1.

Next, consider states $i = (q_n, \mathbb{w}_n, \mathbb{g}_n)$ at stage n and $j = (q_{n+1}, \mathbb{w}_{n+1}, \mathbb{g}_{n+1})$ at stage $n+1$. Two types of actions are possible. If the current feed-mix is not changed, the transition probability equals

$$\Pr(j \mid i, a_{\texttt{cont}}) = \begin{cases} \Pr(\mathbb{w}_{n+1}, \mathbb{g}_{n+1} \mid \mathbb{w}_n, \mathbb{g}_n), & q_{n+1} = q_n, \\ 0, & \text{otherwise.} \end{cases} \tag{2.1}$$

and if q pigs are culled, the transition probability equals:

$$\Pr(j \mid i, a_q) = \begin{cases} \Pr(\mathbb{w}_{n+1}, \mathbb{g}_{n+1} \mid \mathbb{w}_n, \mathbb{g}_n), & q_{n+1} = q_n - q, \\ 0, & \text{otherwise.} \end{cases} \tag{2.2}$$

The probability $\Pr(\mathbb{w}_{n+1}, \mathbb{g}_{n+1} \mid \mathbb{w}_n, \mathbb{g}_n)$ depends on the statistical models used for Bayesian forecasting and will be given in Section 2.4.

Finally, if the dummy parent action in state $\tilde{i}$ is considered, a deterministic transition to state $i^1 = \tilde{i}$ in process $\mathfrak{p}^1$ occurs.

2.3.4 Expected rewards

To finalize the description of the model, the expected reward of each (jump) action must be specified.

Level 0 - Founder process $\mathfrak{p}^0$

Action a^0 represents the insertion of $q^{\max}$ piglets and hence the reward equals $r(i^0, a^0) = -c_{\texttt{pig}} q^{\max}$, where $c_{\texttt{pig}}$ denotes the unit cost of a piglet.

Level 1 - Parent process $\mathfrak{p}^1$

The reward of child jump action a_f (choose feed-mix f) is zero since the cost of reconfiguring the feeding system is added in Level 2. The same holds for the parent jump action $\tilde{a}$ where the reward is assumed to be zero.

Level 2 - MDP $\mathfrak{p}^2 = (\mathfrak{p}^1 \parallel (n^1, i^1, a^1))$

The reward of choosing a new feed-mix (parent jump action $a_{\texttt{newMix}}$) is $-c_{\texttt{newMix}}$ where $c_{\texttt{newMix}}$ denotes the fixed cost of changing from one feed-mix to another. The reward of the dummy parent jump action $\tilde{a}$ is zero.

For the remaining actions ($a_{\texttt{cont}}, a_{\texttt{term}}, a_q$) the expected reward equals the expected revenue from selling the pigs minus the expected cost of feeding the pigs conditioned on the values of the state variables and the action. Let $(w_{(1)}, z_{(1)}), \ldots, (w_{(q_n)}, z_{(q_n)})$ denote the weight and weekly feed intake of the pigs ordered such that $w_{(k)} \leq w_{(k+1)}$. That is, $w_{(k)}$ is the weight of kth pig, i.e. the kth order statistics. If the q heaviest pigs are culled, the revenue becomes

$$\sum_{j=q_n-q+1}^{q_n} \tilde{w}_{(j)} \cdot p(\tilde{w}_{(j)}, \breve{w}_{(j)}), \tag{2.3}$$

where $\tilde{w}_{(j)}$ and $\breve{w}_{(j)}$ denote the carcass weight (kg) and the leanness (non-fat percentage) of the jth pig in the pen, respectively. Price function $p(\cdot)$ is the unit price of the meat. Similarly, the cost of feeding the $q_n - q$ lightest pigs is

$$\sum_{j=1}^{q_n-q} z_{(j)} \cdot c_f, \tag{2.4}$$

where c_f denotes the unit cost of feed-mix f. The expected reward $r_n(i, a_q)$ can now be found as the difference between the expected value of Equations (2.3) and (2.4). Actions $a_{\texttt{cont}}$ and $a_{\texttt{term}}$ may be considered as extreme culling decisions ($q = 0$ and $q = q_n$), i.e. $r_n(i, a_{\texttt{cont}})$ equals the expectation of (2.4) with $q = 0$ and $r_n(i, a_{\texttt{term}})$ equals the expectation of (2.3) with $q = q_n$.

To evaluate the expected reward of (2.3) and (2.4), statistical models are needed to transform the repeated measurements of weight and feed intake into relevant information about weight and growth using Bayesian forecasting. This will be the focus in the next section.

2.4 Bayesian updating of weight and growth

In animal production, online monitoring is a relevant method to obtain data for tracking the changes and can be done regularly by sensors placed in the production units. Two types of online sensors are considered in the finisher pen which provide data about live weight and feed intake, respectively. To transform these data into information about weight and growth, we need a statistical model. In this paper state space models (SSMs) are used to estimate the mean weight μ_t and growth g_t of the pigs in the pen at time t. The same holds for the standard deviation σ_t of the pig weights in the pen.

The set $\mathbb{W}_n$ in the HMDP will therefore contain discretized estimates of the mean weight and standard deviation and the set $\mathbb{G}_n$ will include discretized estimates of the mean growth.

SSMs can be categorized into different groups based on the dynamic nature of the considered system and the probability distribution assumed. Two kinds of SSMs are considered and later embedded into the HMDP. In the first model, the probability distribution of the observations, related to the online sensors, is Gaussian (*GSSM*) and in the second model, these observations come from a non-Gaussian distribution (*nGSSM*). The dynamics of the system is modeled by linear equations in both models.

The next sections, first provide a description of the two models and afterwards use the models to calculate the reward and transition probabilities of the HMDP. For a short introduction to SSMs and the theorems used for Bayesian updating see Appendix 2.B.

2.4.1 A GSSM for average weight and growth estimations

Let $(\hat{w}_1, \dots, \hat{w}_d)_t$ denote d weight estimates obtained by an online weighting method (e.g. image processing) at time t. That is, an estimate of the average weight at time t is $\bar{w}_t = \sum_{k=1}^{d} \hat{w}_k / d$. Moreover, assume that the average feed intake per pig $\bar{z}_t$, given feed-mix f, is measured using an automatic feeding system. The following GSSM is used to model mean weight and growth:

$$\text{system equation } (\theta_t = G\theta_{t-1} + \omega_t): \quad \begin{pmatrix} \mu_t \\ g_t \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \mu_{t-1} \\ g_{t-1} \end{pmatrix} + \begin{pmatrix} \omega_{1t} \\ \omega_{2t} \end{pmatrix}, \tag{2.5}$$

$$\text{observation equation } (y_t = F_t'\theta_t + \nu_t): \quad \begin{pmatrix} \bar{w}_t \\ \bar{z}_t \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ k_{1t} & k_2 \end{pmatrix} \begin{pmatrix} \mu_t \\ g_t \end{pmatrix} + \begin{pmatrix} \nu_{1t} \\ \nu_{2t} \end{pmatrix}, \tag{2.6}$$

The system equation (2.5) models the relation between the latent variables $\theta_t = (\mu_t, g_t)'$ where the first equation in (2.5) states that the mean weight μ_t in the pen at time t equals the mean weight at time $t-1$ plus the mean growth and some noise. The second equation states that the mean growth g_t in the pen follows a random process. The system noise is $\omega_t \sim N(0,W)$ and the prior distribution is $\theta_0 \sim N(m_{0,f}, C_{0,f})$ given a fixed feed-mix f.

The observation equation (2.6) illustrates the relation between the observed variables $y_t = (\bar{w}_t, \bar{z}_t)'$ and the latent variables. That is, in the first equation, the observed average weight equals the mean weight plus the measurement error of the weighing method, and in the second equation, the observed average feed intake equals

$$\bar{z}_t = k_{1t}\mu_t + k_2 g_t + \nu_{2t},$$

where k_{1t} and k_2 are two known parameters. This relation is based on a linear approximation of the relation between feed intake and growth stated in Jørgensen (1993) where k_{1t} is a dynamic parameter to cover the non-linearity of the weight term. The observation error is assumed to be $\nu_t \sim N(0,V)$.

Let $\mathbb{D}_t = (y_1, ..., y_t, m_{0,f}, C_{0,f})$ denote the information available up to time t. When new values of the observable variable $y_t = (\bar{w}_t, \bar{z}_t)'$ are received, Bayesian updating (Theorem 1 in Appendix 2.B) can be used to update the posterior $(\theta_t \mid \mathbb{D}_t) \sim N(m_t, C_t)$ at time t. That is, the posterior mean and covariance given the current feed-mix f become

$$m_t = \begin{pmatrix} \hat{\mu}_t \\ \hat{g}_t \end{pmatrix}, \quad C_t = \begin{pmatrix} C_{t,1} & C_{t,12} \\ C_{t,12} & C_{t,2} \end{pmatrix}.$$

The estimated means $(\hat{\mu}_t, \hat{g}_t)'$ are the best estimate of the latent variables $(\mu_t, g_t)'$. The starting time of the GSSM is when the pigs are inserted into the pen, i.e. the prior mean of the latent variable is $m_{0,f} = (\hat{\mu}_0, \hat{g}_{0,f})$ where $\hat{\mu}_0$ denotes the average weight of the piglets at insertion and $\hat{g}_{0,f}$ is the estimated growth rate given feed-mix f (prior to receiving sensor data). The initial covariance $C_{0,f}$ contains the initial covariance components of live weight and growth rate at the time of insertion given feed-mix f.

If the feed-mix is changed at time t to a new feed-mix f, this change is interpreted as a system intervention (Kristensen et al., 2010, Section 8.2.5) and the posterior mean and covariance are modified to

$$m_t = \begin{pmatrix} \hat{\mu}_t \\ \hat{g}_{0,f} \end{pmatrix}, \quad C_t = \begin{pmatrix} C_{t,1} & C_{0,f,12} \\ C_{0,f,12} & C_{0,f,2} \end{pmatrix},$$

where $\hat{g}_{0,f}$ denotes the initial estimate of the growth rate of the new feed-mix (prior to receiving sensor data for feed-mix f) and $C_{0,f,\cdot}$ denotes the initial covariances for the feed-mix f.

2.4.2 An nGSSM to estimate the variance of weights in the pen

Assuming d weight estimates $(\hat{w}_1,\ldots,\hat{w}_d)_t$ at time t, the unbiased sample variance at time t is $s_t^2 = \sum_{k=1}^{d}(\hat{w}_k - \bar{w}_t)^2/(d-1)$. It is well known that if s_t^2 is based on observations from a normal distribution with true variance $(\sigma_t)^2$, then $(d-1)s_t^2/(\sigma_t)^2$ follows a chi-square distribution with $d-1$ degrees of freedom (Wackerly et al., 2008, p357). Hence the sample variance s_t^2 follows a gamma distribution with shape $\mathfrak{a}_t$ and scale $\mathfrak{b}_t$ given as

$$\mathfrak{a}_t = \frac{d-1}{2}, \quad \mathfrak{b}_t = \frac{2(\sigma_t)^2}{d-1}.$$

Note that since d is constant, $\mathfrak{a}_t$ is constant and known for all $t > 1$.

An nGSSM can now be defined with observation $y_t = s_t^2$ and latent variable $\theta_t = (\sigma_t)^2$ at time t where y_t follows a gamma distribution with shape $\mathfrak{a}_t$ and scale $\mathfrak{b}_t$. The natural parameter becomes $\eta_t = -1/(\sigma_t)^2$ and the impact on the latent variable $(g(\eta_t) = F_t'\theta_t)$ is defined as $g(\eta_t) = (\sigma_t)^2$, i.e. $g(\eta_t) = -1/\eta_t, F_t' = 1$ (see Appendix 2.B). The system equation is:

$$(\sigma_t)^2 = G_t(\sigma_{t-1})^2,$$

where $G_t = (\frac{t}{t-1})^2$ for $t > 1$ ($G_1 = 1$). That is, it is assumed that the true variance in the pen increases by coefficient $(\frac{t}{t-1})^2$ (Kristensen et al., 2012).

It should be noted that the conjugate prior distribution of $(\sigma_t)^2$ is an inverse-gamma distribution (Gelman et al., 2004, p50). Hence, when the piglets are inserted into the pen ($t = 0$), the initial (prior) distribution of the variance is

$$\theta_0 \sim \text{Inv-Gamma}\left(\mathfrak{c}_0 = \frac{d-1}{2}, \mathfrak{d}_0 = \frac{(d-1)s_0^2}{2}\right), \tag{2.7}$$

with shape $\mathfrak{c}_0$ and scale $\mathfrak{d}_0$ where s_0^2 is the initial estimated sample variance of the live weight with sample size d. Given the nGSSM and the initial distribution (2.7), the estimate of the variance can now be updated when a new observation s_t^2 is obtained from the pen by using Theorem 3 and Corollary 1 in Appendix 2.B.

2.4.3 Embedding the SSMs into the HMDP

The two SSMs described in the previous sections provide information about the mean weight and growth (μ_t, g_t) and the standard deviation σ_t of the weights in the pen. To embed this information into the HMDP these values have to be discretized (Nielsen et al., 2011).

Let $\mathbb{U}_{x_n} = \{\Pi_1, \ldots, \Pi_{|\mathbb{U}_{x_n}|}\}$ be a set of disjoint intervals representing the partitioning of possible values of the continuous state variable x at stage n (e.g. $x = \hat{\mu}_n$). Moreover, given interval Π, let *centre point* π denote the centre of the interval. The set of possible values of the state variables in the HMDP related to information about weight is $\mathbb{W}_n = \mathbb{U}_{\hat{\mu}_n} \times \mathbb{U}_{\hat{\sigma}_n}$ and hence a state $\mathbb{w}_n$ corresponds to area $\Pi_{\hat{\mu}_n} \times \Pi_{\hat{\sigma}_n}$ and is represented using the centre point $\mathbb{w}_n = (\pi_{\hat{\mu}_n}, \pi_{\hat{\sigma}_n})$. Similarly the corresponding set of possible values of the state variable related to information about growth is $\mathbb{G}_n = \mathbb{U}_{\hat{g}_n}$.

Transition probabilities

It is now possible to compute the transition probability $\Pr\left(\mathbb{w}_{n+1}, \mathbb{g}_{n+1} \mid \mathbb{w}_n, \mathbb{g}_n\right)$ used in (2.1) and (2.2). Since the mean and variance estimations are treated separately in different SSMs, this transition probability can be split into two parts:

$$\begin{aligned}\Pr\left(\mathbb{w}_{n+1}, \mathbb{g}_{n+1} \mid \mathbb{w}_n, \mathbb{g}_n\right) = &\Pr\left(m_{n+1} = (\hat{\mu}_{n+1}, \hat{g}_{n+1}) \in \Pi_{\hat{\mu}_{n+1}} \times \Pi_{\hat{g}_{n+1}} \mid m_n = (\pi_{\hat{\mu}_n}, \pi_{\hat{g}_n})\right) \\ &\cdot \Pr\left(m'_{n+1} = \hat{\sigma}_{n+1} \in \Pi_{\hat{\sigma}_{n+1}} \mid m'_n = \pi_{\hat{\sigma}_n}\right).\end{aligned}$$

The first part can be calculated using the GSSM as

$$\Pr\left(m_{n+1} \in \Pi_{\hat{\mu}_{n+1}} \times \Pi_{\hat{g}_{n+1}} \mid m_n\right) = \int_{\Pi_{\hat{\mu}_{n+1}}} \int_{\Pi_{\hat{g}_{n+1}}} f_{(m_{n+1}|m_n)}(x,y)\,dy\,dx,$$

where the distribution of $(m_{n+1} \mid m_n)$ can be found using Theorem 2 in Appendix 2.B. The second part can be calculated using the nGSSM as

$$\Pr\left(m'_{n+1} \in \Pi_{\hat{\sigma}_{n+1}} \mid m'_n\right) = \int_{\Pi_{\hat{\sigma}_{n+1}}} f_{(m'_{n+1}|m'_n)}(x)\,dx,$$

where the distribution of $(m'_{n+1} \mid m'_n)$ can be found using Theorem 4 in Appendix 2.B.

Expected rewards

The expected reward given stage n and state $(q_n, \mathbb{w}_n, \mathbb{g}_n) = (q_n, (\pi_{\hat{\mu}_n}, \pi_{\hat{\sigma}_n}), \pi_{\hat{g}_n})$ in process $\mathfrak{p}^2$ can be calculated as the expected value of (2.3) minus the expected value of (2.4). The expected revenue of (2.3) can be written as

$$\sum_{k=q_n-q+1}^{q_n} \mathbb{E}\left(\tilde{w}_{(k)} \cdot p(\tilde{w}_{(k)}, \check{w}_{(k)})\right), \tag{2.8}$$

where $\tilde{w}_{(k)}$ and $\breve{w}_{(k)}$ denote the carcass weight and leanness of the jth pig (based on the order statistics $w_{(k)}$, see Section 2.3.4). The carcass weight, $\tilde{w}_{(k)}$, of the jth pig is a fraction of live weight (Andersen et al., 1999):

$$\tilde{w}_{(k)} = 0.84 w_{(k)} - 5.89 + \varepsilon, \tag{2.9}$$

where $\varepsilon \sim N(0, \sigma_\varepsilon^2)$ is a random error. Furthermore, Kristensen et al. (2012) proposed a rule of thumb for use in production units, which is used to compute the lean meat percentage $\breve{w}_{(k)}$ at marketing:

$$\breve{w}_{(k)} = \frac{-30(\bar{g}_{(k)} - \bar{g})}{4} + \bar{\breve{w}},$$

where $\bar{g}_{(k)}$ denotes the average daily growth/gain of the kth pig until marketing, $\bar{g}$ is the average daily growth in the herd, and $\bar{\breve{w}}$ is the average herd leanness percentage at marketing. The average daily growth t days after insertion into the pen is $\bar{g}_{(k)} = (w_{(k)} - \hat{\mu}_0)/t$, where $\hat{\mu}_0$ denotes the average weight at time of insertion into the pen.

The expected cost of (2.4) is:

$$c_f \sum_{k=1}^{q_n - q} \mathbb{E}\left(z_{(k)}\right). \tag{2.10}$$

and from (2.6), the ordered random variable $z_{(k)}$ equals:

$$z_{(k)} = k_{1t} w_{(k)} + k_2 g_{(k)}.$$

Note that the evaluation of (2.8) and (2.10) is rather complex since it involves calculating the mean of a piecewise reward function and the truncated normal distribution. However, the values of (2.8) and (2.10) can be simulated using a simple sorting procedure and given the fact that $w \sim N(\pi_{\hat{\mu}}, \pi_{\hat{\sigma}})$ where w denotes the weight of a pig randomly selected in the pen at the current stage.

2.5 Numerical example

To illustrate the functionality of the proposed optimization model, the HMDP is applied on three pens with different properties (average weekly gain). The average weekly gain of Pen 2 is assumed to be “normal” (an initial growth of 5.8 kg/week using Feed-mix 1), and Pens 1 and 3 grow 20 percent slower and faster, respectively, than Pen 2. Moreover, to initialize the three pens with the same conditions, the pigs are fed by the same feed-mix (Feed-mix 1) at the time of insertion into the pen.

2.5.1 Model parameters and observation data

To obtain time series of observations $(\bar{w}_t, \bar{z}_t)$ and s_t^2 used by the GSSM and nGSSM a simulation model was developed. The model is based on the biological growth formulas in Jørgensen (1993). The simulation model and the generated data are available online for reproducibility (see Pourmoayed and Nielsen (2014b)).

An example is given in Figure 2.2 that shows the observed values of average live weight $\bar{w}_t$, average feed intake $\bar{z}_t$, and the standard deviation s_t (resulted from the simulation). It also gives the estimated information of live weight and growth (calculated using Bayesian updating with the GSSM and nGSSM) in the three pens. These values together with the other state variables are used to identify the current state in the HMDP. Note that the simulation is started with Feed-mix 1 and each time the feed-mix is changed, we continue the simulation using the new feed-mix.

The parameters used for the HMDP are given in Table 2.1. The parameter values have been obtained using information about finisher pig production (Danish conditions) and related literature (see the footnotes in Table 2.1).

Table 2.2 contains parameter values related to the GSSM and nGSSM. The values have been estimated with time series generated using the simulation model. More specifically, we used the expectation-maximization algorithm (Dethlefsen, 2001) to find V and W, and the initial posterior parameters $m_{0,1}$ and $C_{0,1}$ were estimated using the weight data at the time of inserting the piglets into the pen. For the nGSSM, the initial sample variance s_0^2 was calculated using the time series data and $d = 35$ is used as the number of weight estimates per day. Finally, note that each feed-mix implies a special growth rate in the pen ($\hat{g}_{0,f}$) and that feed-mixes with higher growth rates are more expensive in comparison with other feed-mixes (c_f).

To calculate the revenue of marketing in (2.3), the unit price function $p(\tilde{w}_{(j)}, \breve{w}_{(j)})$ should be specified, which under Danish conditions is the sum of two piecewise linear functions $\tilde{p}(\tilde{w})$ and $\breve{p}(\breve{w})$ related to the price of carcass and a bonus for the leanness percentage per kg meat, respectively. We define $\tilde{p}(\tilde{w})$ and $\breve{p}(\breve{w})$ based on the meat prices used in Kristensen et al. (2012)[1] as

[1]For current prices see `http://www.danishcrown.dk/Ejer/Noteringer/Aktuel-svinenotering.aspx`

Table 2.1: Parameter values (HMDP).

Parameter	Value	Description
$q^{\max}$	15	Number of pigs inserted into the pen.[a]
$b^{\max}$	4	Maximum number of feeding phases.[a]
$\lvert\mathbb{F}\rvert$	3	Number of available feed-mixes.[a]
$t_f^{\min}$	3	Minimum number of weeks using feed-mix f.[a]
$t^{\max}$	12	Maximum number of weeks in a growing period.[a]
$t^{\min}$	9	First possible week of marketing decisions.[a]
$c_{\texttt{newMix}}$	0	Cost of changing the feed-mix (DKK).[a]
$c_{\texttt{pig}}$	375	Cost of a piglet (DKK).[bc]
c_f	$\{1.8, 1.88, 1.96\}$	Unit cost of feed-mix $f = 1,\dots,3$ (DKK/FEsv).[ad]
$\bar{g}$	6	Average weekly gain (kg) in the herd.[c]
$\bar{\breve{w}}$	61	Average leanness percentage in the herd.[c]
σ_ε	1.4	Standard deviation of ε.[c]

[a] Value based on discussions with experts in Danish pig production. [b] Time series of Danish prices can be seen at `http://www.notering.dk/WebFrontend/`. [c] Value taken from Kristensen et al. (2012). [d] FEsv is the energy unit used for feeding the pigs in Denmark. One FEsv is equivalent to 7.72 MJ.

$$\tilde{p}(\tilde{w}) = \begin{cases} 0 & \tilde{w} < 50 \\ 0.2\tilde{w} - 2.7 & 50 \leq \tilde{w} < 60 \\ 0.1\tilde{w} + 3.3 & 60 \leq \tilde{w} < 70 \\ 10.3 & 70 \leq \tilde{w} < 86 \\ -0.1\tilde{w} + 18.9 & 86 \leq \tilde{w} < 95 \\ 9.3 & 95 \leq \tilde{w} < 100 \\ 9.1 & \tilde{w} \geq 100, \end{cases} \tag{2.11}$$

and

$$\breve{p}(\breve{w}) = 0.1(\breve{w} - 61).$$

A plot of the two functions is given in Figure 2.3.

Finally, in order to initialize the HMDP, possible values of the state variables should be determined for each stage. For the discrete state variables (q_n, t_n, f_n), the possible values are set according to the set of states defined in Section 2.3.2. Moreover, based on the discretization method in the beginning of Section 2.4.3, the continuous state variables ($\hat{\mu}_n$, $\hat{\sigma}_n$, $\hat{g}_n$) are divided

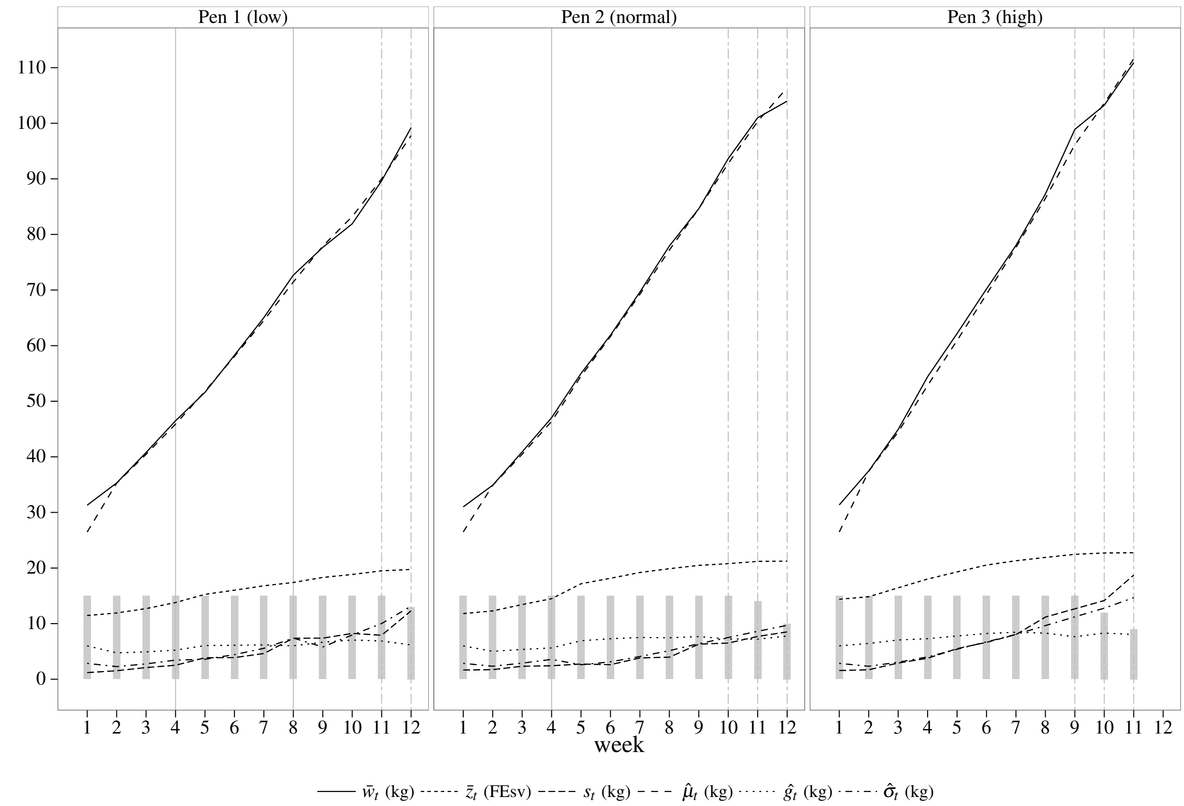
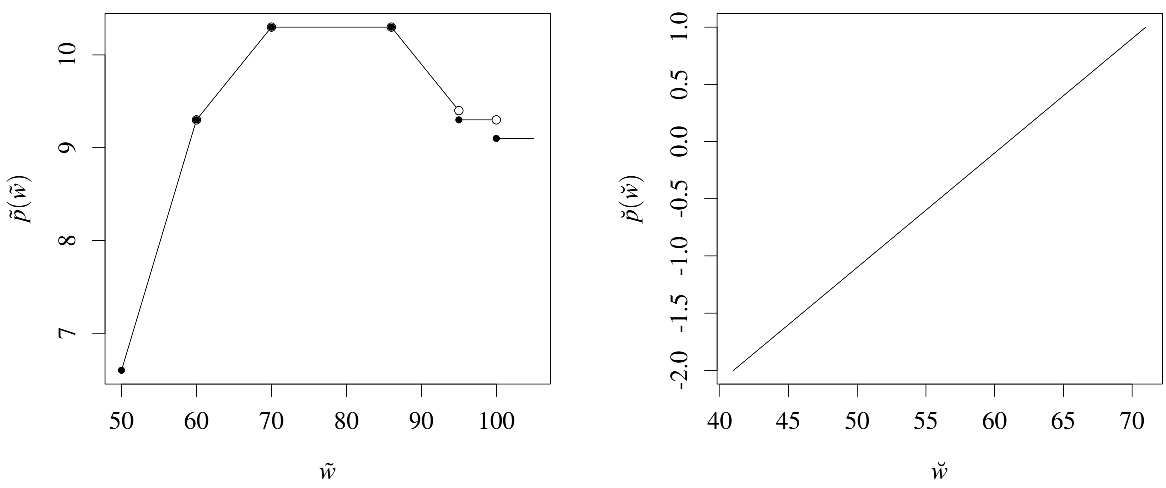


Figure 2.2: Observed and estimated information of live weight and growth rate in the three pens. Observed information are average live weight $\bar{w}_t$, average feed intake $\bar{z}_t$ and standard deviation of live weight s_t per week (resulted from simulation). Estimated information are estimated means of live weight and growth rate, $\hat{\mu}_t$ and $\hat{g}_t$ (computed using the GSSM), and estimated standard deviation of live weight $\hat{\sigma}_t$ (computed using the nGSSM). Bars show the number of pigs q_n in the pen before the optimal action is carried out. The vertical dotted and solid lines show the times when the marketing and feeding decisions occur in the system based on the optimal policy, respectively.

into the 11, 7 and 5 intervals, respectively. The centre points of these intervals are specified such that they represent the possible values of the weight and growth information in the system. An overview over the values of each state variable is given in Table 2.3.

2.5.2 Model results

The HMDP was coded using C++ (gcc compiler) and R (R Core Team, 2015). The source code is available online (Pourmoayed and Nielsen, 2014b). After the model was built, the optimal policy was calculated with the modified policy iteration algorithm using the R package “MDP”



(a) Carcass weight. (b) Leanness.

Figure 2.3: Price functions (DKK/kg) given carcass weight ($\tilde{p}(\tilde{w}) = 0$ for $\tilde{w} < 50$) and leanness percentage.

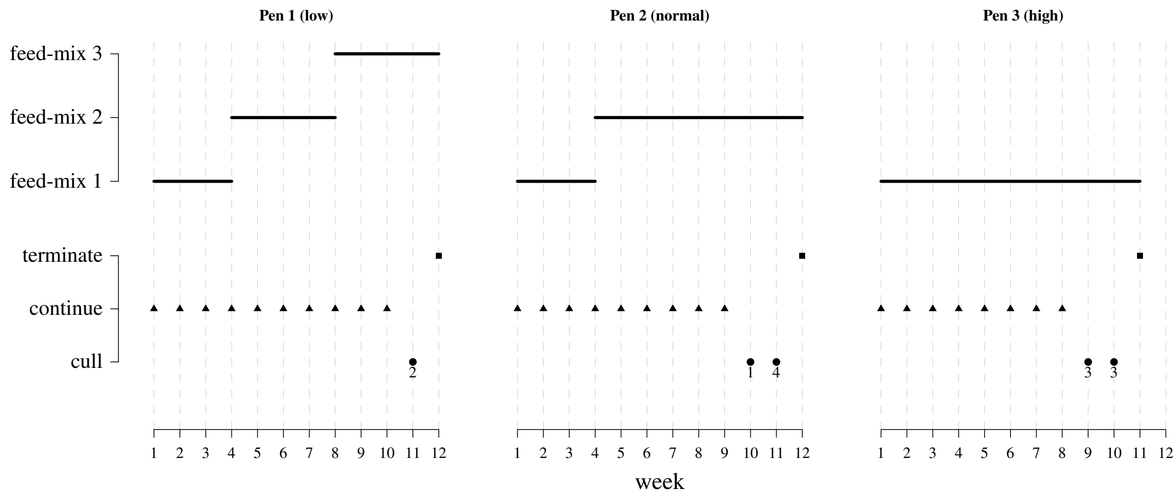


Figure 2.4: The optimal feeding and marketing decisions for the three pens. Upper part of each plot illustrates the optimal feed-mix (solid line) and the lower part shows the optimal marketing decision. Numbers close to cull actions (a_q) are the number of pigs culled.

Table 2.2: Parameter values (GSSM and nGSSM).

Parameter	Value	Explanation
GSSM		
V	$\begin{pmatrix} 0.066 & 0.027 \\ 0.027 & 0.012 \end{pmatrix}$	Observation variance.[a]
W	$\begin{pmatrix} 2.1 & -0.124 \\ -0.124 & 0.112 \end{pmatrix}$	System variance.[a]
$m_{0,1}$	$\begin{pmatrix} 26.49 \\ 5.8 \end{pmatrix}$	Initial prior mean weight and growth $m_{0,1} = (\hat{\mu}_0, \hat{g}_{0,1})$ for Feed-mix 1.[a]
$C_{0,1}$	$\begin{pmatrix} 4.26 & 0.32 \\ 0.32 & 0.53 \end{pmatrix}$	Initial prior covariance matrix for Feed-mix 1.[a]
k_{1t}	$\{0.134 - 0.004i + 0.0001i^2 : i = 1, \dots, 12\}$	Energy requirements (FEsv) per kg live weight.[b]
k_2	1.549	Energy requirement (FEsv) per kg gain.[c]
$\hat{g}_{0,f}$	$\{5.8, 6.3, 6.8\}$	Initial growth rate estimate of feed-mix $f = 1, \dots, 3$.[d]
nGSSM		
s_0^2	7.65	Initial sample variance (kg^2).[a]
d	35	Sample size (observations per day).[e]

[a] Estimated based on time series generated using the simulation model. [b] Based on a linear approximation of the relation between feed intake and growth stated in Jørgensen (1993). [c] Value taken from Jørgensen (1993). [d] Value based on discussions with experts in Danish pig production. [e] Value used in the simulation model.

Table 2.3: Possible values of the state variables and the range of the centre points in the HMDP.

State / Week (n)	1	2	3	4	5	6	7	8	9	10	11	12
q_n	15	15	15	15	15	15	15	15	1-15	1-15	1-15	1-15
t_n (week)	1	1	1	1, 4	1, 4-5	1, 4-6	1, 4-7	1, 4-8	1, 4-9	1, 4-10	1, 4-11	1, 4-11
f_n	1	1	1	1-3	1-3	1-3	1-3	1-3	1-3	1-3	1-3	1-3
$\pi_{\hat{\mu}_n}$ (kg)[a]	7-47	14-54	20-61	28-68	35-75	42-82	49-88	56-96	63-103	70-110	77-116	84-124
$\pi_{\hat{\sigma}_n}$ (kg)[a]	1.6-6.4	2.1-6.9	2.6-7.4	3.1-7.9	3.6-8.4	4.1-8.9	4.6-9.4	5.1-9.9	5.6-10.4	6.1-10.9	6.6-11.4	7.1-11.9
$\pi_{\hat{g}_n}$ (kg)[a]	4.4-8.2	4.4-8.2	4.4-8.2	4.4-8.2	4.4-8.2	4.4-8.2	4.4-8.2	4.4-8.2	4.4-8.2	4.4-8.2	4.4-8.2	4.4-8.2

[a] Variables $\hat{\mu}_n$, $\hat{\sigma}_n$, $\hat{g}_n$ are discretized into 11, 7, and 5 intervals, respectively. Rows $\pi_{\hat{\mu}_n}, \pi_{\hat{\sigma}_n}$, and $\pi_{\hat{g}_n}$ show the range of the possible values of $\hat{\mu}_n$, $\hat{\sigma}_n$ and $\hat{g}_n$.

(Nielsen, 2009). The resulting model consists of 802581 states and 5050446 actions (one stage of the founder process including states and actions of sub-processes).

The CPU time for building and solving the model was 268 and 94 seconds, respectively (Fujitsu laptop with i7-4800MQ CPU and 32 GB of memory running on a Windows 7, 64 bit OS). Note that the model has only to be resolved when the model parameters change, e.g. a new estimation of V and W which might be re-estimated monthly. Therefore, a fast solution time is not the primary focus.

The information from each pen, i.e. the values $q_n, t_n, f_n, \hat{\mu}_n, \hat{\sigma}_n$ and $\hat{g}_n$, is used to find the relevant state in the HMDP. Next, the optimal action is found using the calculated optimal policy. The resulting optimal feeding and marketing decisions (i.e. the sample path of the MDP) are illustrated in Figures 2.2 and 2.4 for each pen. In Figure 2.2 the vertical dotted and solid lines show weeks where marketing and feeding decisions are taken in the system. The bars show the number of pigs left in the pen before a (possible) marketing decision. For instance, in week 9, three pigs in Pen 3 are marketed. A detailed overview of the optimal decisions is given in Figure 2.4. Here the plot of each pen is separated into two parts. The solid line in the upper part shows the optimal feed-mix. A jump indicates that the optimal decision is to change the current feed-mix. In the lower part of each plot the optimal marketing decision is illustrated by means of symbols. For instance, the black dots indicate a culling action.

A closer look at Figure 2.4 shows that all pens start with Feed-mix 1. After 3 weeks, the feed-mix in Pen 1 (with the lowest weekly gain) changes to Feed-mix 2, resulting in a better growth rate compared to Feed-mix 1. Pen 1 (low growth) uses this feed-mix until week 8 and after that Feed-mix 3 is chosen for the remaining weeks because a higher growth is obtained (compared to using Feed-mix 2), and hence the appropriate live weight is reached at the end of the growing period. In Pen 2 (normal growth), we change the feed-mix in week 4 from Feed-mix 1 to Feed-mix 2 and until week 12 this feed-mix is used in the pen. In this pen, the average growth rate is appropriate and there is no need to use a more expensive feed-mix (Feed-mix 3) with a faster expected growth rate. Finally, in Pen 3 (high growth), the feed-mix remains unchanged since the pigs genetically grow fast in this pen using the cheapest feed-mix (Feed-mix 1) and they will have an appropriate live weight in the last weeks of the growing period.

The length of the growing period, i.e. the number of weeks before terminating the pen, differs between pens. Pens 1 and 2 are terminated at the maximum growth length (week 12) since a good slaughter weight is reached for the remaining pigs. However, Pen 3 is terminated in week 11. Due to the high growth rate in this pen, the average weight in week 11 is appropriate and the pen is terminated such that a new batch of piglets can be inserted into the pen earlier (new production cycle).

Individual marketing decisions are made in all pens to select the heaviest pigs for marketing. Usually pigs grow with different growth values in the pen and hence in the last weeks of the growing period (from week 9 to 12) they obtain different live weights. Hence, these decisions are made in order to market the pigs that are in the best slaughter weight interval (with a live weight approximately between 89 and 109 kg due to (2.9) and (2.11)). For instance, in Pen 2, the

Table 2.4: Three groups of scenarios to show the impact of changing model parameters. The basic scenario is based on the parameters in Tables 2.1 and 2.2 where $t^{\min} = 9$ and $t^{\max} = 12$ weeks.

Scenario group	Parameter change	Gross margin per week (DKK)
Basic	none	71.349
Group 1 - starting time of marketing period	$t^{\min} = 8$	71.355
	$t^{\min} = 10$	71.249
	$t^{\min} = 11$	70.428
Group 2 - maximum length of marketing period	$t^{\max} = 11$	34.155
	$t^{\max} = 13$	92.618
	$t^{\max} = 14$	104.644
	$t^{\max} = 15$	110.884
Group 3 - feed-mix unit cost	10% increase	27.444
	10% decrease	116.064

4 heaviest pigs are culled in week 11. As a result, these individual marketing decisions lead to a decrease in the inhomogeneity between the remaining pigs in the pen, which implies a more consistent growth among the remaining pigs.

Changing the parameters of the model influences on the optimal policy. To make a small sensitivity analysis, three groups of scenarios are considered and compared with the basic scenario based on the parameters in Tables 2.1 and 2.2. In the first group of scenarios the starting time of the marketing period is changed by considering different values of $t^{\min}$ under a fixed growing period ($t^{\max} = 12$). In the second group, the maximum length of the marketing period is altered by changing $t^{\max}$ under a fixed starting time for marketing decisions ($t^{\min} = 9$) and in the third group different feed-mix unit costs are taken into account. Under each scenario, the optimal policy of the HMDP and the gross margin per week are calculated for comparison.

The results are shown in Table 2.4. In Group 1, a change in the starting time of possible marketing decisions ($t^{\min}$) has a small impact on the gross margin while in Group 2 the maximum length of the marketing period ($t^{\max}$) has a much higher impact on the gross margin per week. Therefore, it is better to increase the marketing length by extending the growing period $t^{\max}$ compared to lowering $t^{\min}$. This illustrates the importance of the length of the growing period in the pen. Finally, in Group 3, a decrease/increase in the feed-mix unit cost gives a higher/lower gross margin. The effect is relatively high which shows that the profit of the production unit is extremely dependent on the feeding costs.

2.6 Conclusions and further research

In the production of growing pigs, the decision maker must consider feeding and marketing decisions simultaneously. In this paper, we presented a three-level HMDP which considers both feeding and marketing decisions at pen level.

We used a Bayesian approach to update the state of the system such that it contains the updated information based on previous measurements. More specifically, a GSSM is used to forecast mean weight and growth information based on online measurements and an nGSSM is used to forecast the weight variance within the pen. By embedding the SSMs into the HMDP, the model takes into account new online measurements. Both SSMs are embedded into the HMDP using a general discretization method.

A numerical example shows that the optimal policy adapts to different pen conditions (we used three pens with different genetic properties) and chooses actions which maximize the expected reward per time unit. Furthermore, a marginal sensitivity analysis illustrated the importance of the length of the growing period and feed-mix cost.

The model presented in this paper can be used as a part of a decision support system with online data such that the system state can be found using Bayesian updating and the optimal policy of the HMDP can determine the best feeding and marketing decisions at pen level. For simplicity we have assumed that the three alternative feed mixes available for the pigs are the same throughout the production period. In practice it would be natural to adjust the feed mixes to the various growth phases so that the alternatives taken into account depend on the age of the pigs. It would be straightforward to implement such a more realistic setup so it is not a limitation for a practical use. Moreover, in practice there is a risk of death in the pen such that about 4% of the pigs die during the growing period in the finisher unit (Vinther, 2011). The mortality rate can easily be considered in the model by adding a fixed probability of death to the transition probabilities of an action. There are, however, some other limitations of the model that require more thorough consideration.

First, the model considers feeding and marketing decisions at pen level and ignores possible constraints at section or herd level. For instance, limitations in the feeding management system and the transportation strategy to the abattoir are currently ignored. That is, we assume weekly deliveries to the abattoir in the marketing period based on a cooperative agreement where culled pigs from each pen are grouped in one delivery. Hence, the transportation cost is fixed and can be ignored. This is the situation in many Danish herds since the majority of farmers in Denmark use a single abattoir which also handles the transport. To handle constraints and decisions about transportation costs (e.g. truck capacity), we need to extend the model from pen level to section or herd level. Given the current modeling framework, this extension may be difficult due to the curse of dimensionality since the number of states will grow dramatically. As a result there is a need for an approximation method to approximate the value function of the HMDP and find the best marketing policy in the herd. This can be done by using an approximate dynamic programming approach (Powell, 2007) and is a possible direction of future research.

Second, we may have weekly variations in the carcass price (2.11) and piglet cost in practice. This fact may have an influence on the marketing decisions but has been ignored in this study and in previous papers using HMDP models (Nielsen and Kristensen, 2014). Considering price variations in a model with marketing and feeding decisions is difficult since state variables related to price information have to be introduced into the model which will result in an exponential increase in the number of state variables. Two directions are possible. Either approximate dynamic programming methods are applied or other state variables are excluded from the model. Price variations can be analyzed using an SSM and Bayesian updating and embedded into an HMDP which is a subject of future research.

Finally, the model may be extended to handle information and decisions about diseases such as diarrhea and tail biting.

Acknowledgment

We thank the three anonymous referees for their useful comments and suggestions. This article has been written with support from the Danish Council for Strategic Research (The PigIT project, Grant number 11-116191).

References

D.L.J. Alexander, P.C.H. Morel, and G.R. Wood. Feeding strategies for maximising gross margin in pig production. In János D. Pintér, editor, *Global Optimization*, volume 85 of *Nonconvex Optimization and Its Applications*, chapter 2, pages 33–43. Springer, 2006. URL http://link.springer.com/chapter/10.1007%2F0-387-30927-6_2.

S. Andersen, B. Pedersen, and M. Ogannisian. Slagtesvindets sammensætning. meddelelse 429. Technical report, Landsudvalget for Svin og Danske Slagterier, 1999. URL http://vsp.lf.dk/Publikationer/Kilder/lu_medd/medd/429.aspx.

M.A. Boland, P.V. Preckel, and A.P. Schinckel. Optimal hog slaughter weights under alternative pricing systems. *Journal of agricultural and applied economics*, 25:148–148, 1993. URL http://purl.umn.edu/15033.

M.A. Boland, K.A. Foster, and P.V. Preckel. Nutrition and the economics of swine management. *Journal of Agricultural and Applied Economics*, 31:83–96, 1999. URL http://ageconsearch.umn.edu/bitstream/15131/1/31010083.pdf.

C. Bono, C. Cornou, and A.R. Kristensen. Dynamic production monitoring in pig herds i: Modeling and monitoring litter size at herd and sow level. *Livestock Science*, 149(3):289–300, 2012. doi:10.1016/j.livsci.2012.07.023 .

C. Bono, C. Cornou, S. Lundbye-Christensen, and A.R. Kristensen. Dynamic production monitoring in pig herds ii. modeling and monitoring farrowing rate at herd level. *Livestock Science*, 155(1):92–102, 2013. doi:10.1016/j.livsci.2013.03.026 .

K.A. Boys, N. Li, P.V. Preckel, A.P. Schinckel, and K.A. Foster. Economic replacement of a heterogeneous herd. *American Journal of Agricultural Economics*, 89(1):24–35, 2007. doi:10.1111/j.1467-8276.2007.00960.x .

J.P. Chavas, J. Kliebenstein, and T.D. Crenshaw. Modeling dynamic agricultural production response: The case of swine production. *American Journal of Agricultural Economics*, 67(3): 636–646, 1985. doi:10.2307/1241087 .

C. Cornou, J. Vinther, and A.R. Kristensen. Automatic detection of oestrus and health disorders using data from electronic sow feeders. *Livestock Science*, 118(3):262–271, 2008.

G.E. Crooks. Survey of simple, continuous, univariate probability distributions. Technical report, Lawrence Berkeley National Lab, 2013. URL http://threeplusone.com/Crooks-GUDv5.pdf.

C. Dethlefsen. *Space Time Problems and Applications*. PhD thesis, Aalborg University, Denmark, 2001.

A. Gelman, J.B. Carlin, H.S. Stern, D.B. Dunson, A. Vehtari, and D.B. Rubin. *Bayesian data analysis*. CRC press, 2004. ISBN 978-1584883883.

J.J. Glen. A dynamic programming model for pig production. *Journal of the Operational Research Society*, 34:511–519, 1983. doi:10.1057/jors.1983.118 .

E. Jørgensen. The influence of weighing precision on delivery decisions in slaughter pig production. *Acta Agriculturae Scandinavica, Section A - Animal Science*, 43(3):181–189, August 1993. doi:10.1080/09064709309410163 .

A. R. Kristensen, E. Jørgensen, and N. Toft. *Herd Management Science. II. Advanced topics*. Academic books, Copenhagen, 2010. ISBN 8779970796.

A.R. Kristensen. Hierarchic markov processes and their applications in replacement models. *European Journal of Operational Research*, 35(2):207–215, 1988. doi:10.1016/0377-2217(88)90031-8 .

A.R. Kristensen and E. Jørgensen. Multi-level hierarchic markov processes as a framework for herd management support. *Annals of Operations Research*, 94(1-4):69–89, 2000. doi:10.1023/A:1018921201113 .

A.R. Kristensen, L. Nielsen, and M.S. Nielsen. Optimal slaughter pig marketing with emphasis on information from on-line live weight assessment. *Livestock Science*, 145(1-3):95–108, May 2012. doi:10.1016/j.livsci.2012.01.003 .

H. Kure. *Marketing Management Support in Slaughter Pig Production*. PhD thesis, The Royal Veterinary and Agricultural University, 1997. URL `http://www.prodstyr.ihh.kvl.dk/pub/phd/kure_thesis.pdf`.

L.R Nielsen. Mdp: Markov decision processes in R. R package v1.1., 2009. URL `http://r-forge.r-project.org/projects/mdp/`.

L.R. Nielsen and A.R. Kristensen. Finding the K best policies in a finite-horizon Markov decision process. *European Journal of Operational Research*, 175(2):1164–1179, 2006. doi:10.1016/j.ejor.2005.06.011 .

L.R. Nielsen and A.R. Kristensen. Markov decision processes to model livestock systems. In Lluis M. Plà-Aragonés, editor, *Handbook of Operations Research in Agriculture and the Agri-Food Industry*, volume 224 of *International Series in Operations Research & Management Science*, pages 419–454. Springer, 2014. doi:10.1007/978-1-4939-2483-7_19 .

L.R Nielsen, E. Jørgensen, A.R. Kristensen, and S. Østergaard. Optimal replacement policies for dairy cows based on daily yield measurements. *Journal of Dairy Science*, 93(1):77–92, 2010. doi:10.3168/jds.2009-2209 .

L.R. Nielsen, E. Jørgensen, and S. Højsgaard. Embedding a state space model into a markov decision process. *Annals of Operations Research*, 190(1):289–309, 2011. doi:10.1007/s10479-010-0688-z .

J.K. Niemi. *A dynamic programming model for optimising feeding and slaughter decisions regarding fattening pigs | NIEMI | Agricultural and Food Science*. PhD thesis, MTT Agrifood research, 2006. URL `http://ojs.tsv.fi/index.php/AFS/article/view/5855`.

J.W. Ohlmann and P.C. Jones. An integer programming model for optimal pork marketing. *Annals of Operations Research*, 190(1):271–287, November 2008. doi:10.1007/s10479-008-0466-3 .

R. Pourmoayed and L.R. Nielsen. An overview over pig production of fattening pigs with a focus on possible decisions in the production chain. Technical Report PigIT Report No. 4, Aarhus University, 2014a. URL `http://pigit.ku.dk/publications/PigIT-Report4.pdf`.

R. Pourmoayed and L.R. Nielsen. Github repository: A hierarchical Markov decision process modeling feeding and marketing decisions of growing pigs, 2014b. URL `https://github.com/relund/hmdp-feed-pigit`.

W.B. Powell. *Approximate Dynamic Programming: Solving the curses of dimensionality*, volume 703. Wiley-Interscience, 2007. ISBN 978-0-470-60445-8.

R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2015. URL `http://www.R-project.org/`.

S.V. Rodriguez, T.B. Jensen, L.M. Pla, and A.R. Kristensen. Optimal replacement policies and economic value of clinical observations in sow herds. *Livestock Science*, 138(1-3):207–219, June 2011. doi:10.1016/j.livsci.2010.12.026 .

D. Sirisatien, G.R. Wood, M. Dong, and P.C.H. Morel. Two aspects of optimal diet determination for pig production: efficiency of solution and incorporation of cost variation. *Journal of Global Optimization*, 43(2-3):249–261, 2009. doi:10.1007/s10898-007-9262-x .

H.C. Tijms. *A first course in stochastic models*. John Wiley & Sons Ltd, 2003. ISBN 978-0-471-49880-3.

N. Toft, A.R. Kristensen, and E. Jørgensen. A framework for decision support related to infectious diseases in slaughter pig fattening units. *Agricultural Systems*, 85(2):120–137, 2005. doi:

10.1016/j.agsy.2004.07.017 .

J. Vinther. Landsgennemsnit for produktivitet i svineproduktionen 2010. *Videncenter for Svineproduktion. Notat*, (1114), 2011.

D. Wackerly, W. Mendenhall, and R. Scheaffer. *Mathematical statistics with applications*. Thomson Learning, 2008. ISBN 0495385069.

M. West and J. Harrison. *Bayesian Forecasting and Dynamic Models (Springer Series in Statistics)*. Springer-Verlag, February 1997. ISBN 0387947256.

2.A Notations

Since the paper uses techniques from both statistical forecasting and operations research, we had to make some choices with respect to notation. In general, we use capital letters for matrices and let A' denote the transpose of A. Capital blackboard bold letters are used for sets (e.g. $\mathbb{W}_n$ and $\mathbb{F}$). Subscript indices indicate e.g. stage, week, and feed-mix and are separated using a comma. Superscript is only used to indicate the level in the HMDP except when lower and upper limits on ranges (e.g. $t_f^{\min}$ and $t^{\max}$) are considered. Greek letters are used for some stochastic variables and their mean and variance such as θ_t and μ_t. Finally, accent $\hat{x}$ (hat) is used to denote an estimate of x and accent $\bar{x}$ (bar) the average of a group of x-variables. A description of the notation introduced in Section 2.3 and Section 2.4 is given in Tables 2.5 and 2.6, respectively.

Table 2.5: Notation - HMDP model (Section 2.3). Given in approx. the order introduced.

Symbol	Description
$\mathbb{I}_n$	Set of states at stage n.
$\mathbb{A}_n(i)$	Set of actions given stage n and state i.
$r_n(i,a)$	Reward at stage n given state i and action a.
$u_n(i,a)$	Expected length until the next decision epoch at stage n given state i and action a.
$\Pr(j \mid n,i,a)$	Transition probability from state i at stage n to state j at the next stage under action a.
$\Pr_0(i)$	Initial probability of state i.
$\mathfrak{p}^l$	A process at level l (superscript is used to indicate level).
$\mathscr{N}^l$	Time horizon of process $\mathfrak{p}^l$ at level l.
n^l, i^l, a^l	A stage, state, and action in process $\mathfrak{p}^l$.
$q^{\max}$	Number of pigs inserted into the pen.
$b^{\max}$	Maximum number of feeding phases.
$t^{\max}$	Latest week of pen termination.
$t_f^{\min}$	Minimum number of weeks for using feed-mix f.
$t^{\min}$	First possible week of marketing decisions.
f_n	Previous feed-mix used at stage/phase $n-1$, $f_n \in \mathbb{F}$ (set of possible feed-mixes).
t_n	Starting time of phase/stage n (week number), $1 \leq t_n \leq t^{\max}-1$.
q_n	Remaining pigs in the pen at stage n, $1 \leq q_n \leq q^{\max}$.
$\mathbb{w}_n$	Model information related to the weight of the pigs, $\mathbb{w}_n \in \mathbb{W}_n$ (set of possible weight information).
$\mathbb{g}_n$	Model information related to the growth of the pigs, $\mathbb{g}_n \in \mathbb{G}_n$ (set of possible growth information).
a_f	Child jump action for choosing feed-mix $f \in \mathbb{F}$.
$a_{\texttt{newMix}}$	Parent jump action related to changing the current feed-mix.
$a_{\texttt{term}}$	Parent jump action related to terminating a pen.
$a_{\texttt{cont}}$	Action related to continuing the production process without any marketing.
a_q	Action related to marketing the q heaviest pigs in the pen ($1 \leq q < q_n$).
$c_{\texttt{pig}}$	Unit cost of a piglet (DKK).
$c_{\texttt{newMix}}$	Fixed cost of changing the feed-mix (DKK).
c_f	Unit cost of feed-mix f (DKK/FEsv).
$w_{(k)}$	Weight of the kth pig in the pen (kg).
$z_{(k)}$	Weekly feed intake of the kth pig in the pen (FEsv).
$\tilde{w}_{(k)}$	Carcass weight of the kth pig in the pen (kg).
$\breve{w}_{(k)}$	Lean meat percentage of the kth pig in the pen (%).
$\tilde{p}(\tilde{w})$	Unit price of carcass meat (DKK/kg).
$\breve{p}(\breve{w})$	Leanness bonus for 1 kg meat (DKK/kg).
$p(\tilde{w},\breve{w})$	Unit price of meat, $p(\tilde{w},\breve{w}) = \tilde{p}(\tilde{w}) + \breve{p}(\breve{w})$ (DKK/kg).

Table 2.6: Notation - GSSM and nGSSM models (Section 2.4). Given in approx. the order introduced.

Symbol	Description
GSSM (Section 2.4.1)	
μ_t	Mean weight in the pen at time t.
g_t	Mean growth in the pen at time t.
$\bar{w}_t$	Average weight estimate at time t, $\bar{w}_t = \sum_{k=1}^{d} \hat{w}_k/d$ where $\hat{w}_k$ denotes the kth weight estimate and d is the number of weight observations per time unit (sample size).
$\bar{z}_t$	Average feed intake per pig at time t.
θ_t	Latent/unobservable variable(s).
y_t	Observable variable(s).
G	Design matrix of system equation.
F_t	Design matrix of observation equation.
ω_t	System noise, $\omega_t \sim N(0,W)$ where W denotes the system covariance matrix.
v_t	Observation error, $v_t \sim N(0,V)$ where V denotes the observation covariance matrix.
(m_0,C_0)	Mean and covariance matrix of prior given feed-mix f, $\theta_0 \sim N(m_0,C_0$.
$\mathbb{D}_t$	Set of information available up to time t in the system, $\mathbb{D}_t = (y_1,...,y_t,m_0,C_0)$.
(m_t,C_t)	Mean and covariance matrix of posterior at time t, $(\theta_t \mid \mathbb{D}_t) \sim N(m_t,C_t)$.
k_{1t}	Energy requirement per kg live weight (FEsv/kg).
k_2	Energy requirement per kg gain (FEsv/kg).
nGSSM (Section 2.4.2)	
σ_t	Standard deviation of the weights in the pen at time t.
s_t^2	Sample variance of weights in the pen at time t, $s_t^2 = \sum_{j=1}^{d} (\hat{w}_j - \bar{w}_t)^2/(d-1)$.
η_t	Natural parameter of the exponential family distribution.
$(\mathfrak{a}_t,\mathfrak{b}_t)$	Shape and scale parameter of the Gamma distribution, $s_t^2 \sim \text{Gamma}(\mathfrak{a}_t,\mathfrak{b}_t)$.
$(\mathfrak{c}_t,\mathfrak{d}_t)$	Shape and scale parameter of the inverse-Gamma distribution, $(\sigma_0)^2 \sim \text{Inv-Gamma}(\mathfrak{c}_t,\mathfrak{d}_t)$.
Embedding into the HMDP (Section 2.4.3)	
$\mathbb{U}_{x_n}$	Set of disjoint intervals representing the partitioning of possible values of estimate x at stage n, $\mathbb{U}_{x_n} = \{\Pi_1,\ldots,\Pi_{\lvert\mathbb{U}_{x_n}\rvert}\}$ where Π_k denotes interval k.
π_k	Centre point of Π_k.
ε	Random error in estimation of carcass weight given live weight, $\varepsilon \sim N(0,\sigma_\varepsilon^2)$ where σ_ε denotes the standard deviation.
$\overline{g}$	Average weekly gain (kg) in the herd.
$\bar{g}_{(k)}$	Average daily growth/gain of the kth pig until marketing.
$\bar{\bar{w}}$	Average leanness percentage in the herd.
$\hat{\mu}_0$	Initial average weight (kg) at insertion time into the pen.

2.B Statistical models for Bayesian updating

Gaussian state space model (GSSM)

A GSSM includes a set of observable and latent/unobservable continuous variables. The set of latent variables $\theta_{\{t=0,1,...\}}$ evolves over time using *system equation* (written using matrix notation)

$$\theta_t = G_t \theta_{t-1} + \omega_t,$$

where $\omega_t \sim N(0, W_t)$ is a random term and G_t is a matrix of known values. We assume that the prior $\theta_0 \sim N(m_0, C_0)$ is given. Moreover, we have a set of observable variables $y_{\{t=1,2,...\}}$ (the data acquired from the online sensors) which are dependent on the latent variable using *observation equation*

$$y_t = F_t' \theta_t + \nu_t,$$

with $\nu_t \sim N(0, V_t)$. Here F is the design matrix of system equations with known values and F' denotes the transpose to matrix F.

The error sequences ω_t and ν_t are internally and mutually independent. Hence given θ_t we have that y_t is independent of all other observations and in general the past and the future are independent given the present.

Let $\mathbb{D}_{t-1} = (y_1, ..., y_{t-1}, m_0, C_0)$ denote the information available up to time $t-1$. Given the posterior of the latent variable at time $t-1$, we can use Bayesian updating (the Kalman filter) to update the distributions at time t (West and Harrison, 1997, Theorem 4.1).

Theorem 1 *Suppose that at time $t-1$ we have*

$$(\theta_{t-1} \mid \mathbb{D}_{t-1}) \sim N(m_{t-1}, C_{t-1}), \quad \textit{(posterior at time } t-1\textit{)}.$$

then

$$\begin{aligned}
(\theta_t \mid \mathbb{D}_{t-1}) &\sim N(b_t, R_t), \quad \textit{(prior at time } t\textit{)} \\
(y_t \mid \mathbb{D}_{t-1}) &\sim N(f_t, Q_t), \quad \textit{(one-step forecast at time } t-1\textit{)} \\
(\theta_t \mid \mathbb{D}_t) &\sim N(m_t, C_t), \quad \textit{(posterior at time } t\textit{)}
\end{aligned}$$

where

$$
\begin{aligned}
b_t &= G_t m_{t-1}, & R_t &= G_t C_{t-1} G_t' + W_t \\
f_t &= F_t' b_t, & Q_t &= F_t' R_t F_t + V_t \\
e_t &= y_t - f_t, & B_t &= R_t F_t Q_t^{-1} \\
m_t &= b_t + B_t e_t, & C_t &= R_t - B_t Q_t B_t'.
\end{aligned}
$$

Note that the one-step forecast mean f_t only depends on m_{t-1}, i.e. we only need to keep the most recent conditional mean of θ_{t-1} to forecast the next value. Hence when making a prediction based on $\mathbb{D}_{t-1}$, we need only to store m_{t-1}. Similarly, the variance Q_t only depends on the number of observations made, i.e. we can calculate a sequence $Q_1, ..., Q_t$ without knowing the observations $y_1, ..., y_t$.

The distribution of $(m_{t+1} \mid m_t)$ can also be found (Nielsen et al., 2011, page 303).

Theorem 2 *The conditional random variable* $(m_{t+1} \mid m_t)$ *follows a multivariate normal distribution*

$$(m_{t+1} \mid m_t) \sim N(G_{t+1} m_t, B_{t+1} Q_{t+1} B_{t+1}').$$

Non-Gaussian state space model (nGSSM)

An nGSSM relaxes the Gaussian assumption of the observed values, i.e. observations are not conditional Gaussian given the values of the latent variable θ_t. Instead the probability distribution of the observable variable y_t belongs to the exponential family, i.e. the density function is:

$$f(y_t | \eta_t, \mathrm{u}_t) = \exp\left(\frac{x(y_t)\eta_t - a(\eta_t)}{\mathrm{u}_t}\right) q(y_t, \mathrm{u}_t), \tag{2.12}$$

with *natural parameter* η_t and *scale parameter* u_t. Functions $a(\eta_t)$, $x(y_t)$, and $q(y_t, \mathrm{u}_t)$ are assumed known. The equation

$$g(\eta_t) = F_t \theta_t, \tag{2.13}$$

defines the impact of the latent variable θ_t on the natural parameter η_t. Here, $g(\eta_t)$ is a known function. Finally, to specify the full nGSSM model, a *system equation* has to be specified:

$$\theta_t = G_t \theta_{t-1} + \kappa_t,$$

with $\kappa_t \sim [0, H_t]$, meaning that κ_t has zero mean and a covariance matrix H_t. There is no assumption about a normal distribution. In other words, the distribution is only partially specified through its mean and variance (we use the notation $\kappa_t \sim [m_t, H_t]$).

As for the GSSM, the purpose of Bayesian updating is to estimate the latent variable θ_t using previous information $\mathbb{D}_{t-1} = (y_1,...,y_{t-1},m_0,C_0)$ available up to time $t-1$. However, due to (2.13) we also estimate the parameter η_t. An updating procedure was presented by Kristensen et al. (2010, Section 8.5.4). Since there is no normality assumption, only an approximate analysis can be conducted. Moreover, the conjugate family of η_t must be known.

In our application a gamma distribution with shape parameter $\mathfrak{a}_t$ and scale parameter $\mathfrak{b}_t$ is used, i.e. $\eta_t = -1/\mathfrak{a}_t\mathfrak{b}_t$, $V_t = 1/\mathfrak{a}_t$, $a(\eta_t) = \ln(\frac{-1}{\eta_t})$, $x(y_t) = y_t$ and $q(y_t,\mathfrak{a}_t) = y_t^{\mathfrak{a}_t-1}\mathfrak{a}_t^{\mathfrak{a}_t}/\Gamma(\mathfrak{a}_t)$ and the density becomes

$$f(y_t|\mathfrak{a}_t,\mathfrak{b}_t) = \frac{\exp(-y_t/\mathfrak{b}_t)y_t^{\mathfrak{a}_t-1}}{\mathfrak{b}_t^{\mathfrak{a}_t}\Gamma(\mathfrak{a}_t)}. \tag{2.14}$$

Moreover, the conjugate prior of $g(\eta_t)$ is an *Inverse-Gamma* distribution. As a result, the updating procedure (Kristensen et al., 2010, Section 8.5.4) reduces to the theorem below.

Theorem 3 *Suppose that at time $t-1$ we have*

$$(\theta_{t-1} \mid \mathbb{D}_{t-1}) \sim [m_{t-1},C_{t-1}] \quad \textit{(posterior at time } t-1\textit{)},$$

Moreover, assume that $g(\eta_t) \sim$ Inv-Gamma $(\mathfrak{c}_t,\mathfrak{d}_t)$, $g(\eta_t) = -1/\eta_t$, and that the density $f(y_t|\mathfrak{a}_t,\mathfrak{b}_t)$ equals (2.14). *Then*

$$\begin{aligned}
(\theta_t \mid \mathbb{D}_{t-1}) &\sim [b_t,R_t] \quad \textit{(prior at time } t\textit{)},\\
(g(\eta_t) \mid \mathbb{D}_{t-1}) &\sim [f_t,Q_t] \quad \textit{(prior of } g(\eta_t) \textit{ at time } t\textit{)},\\
(\theta_t \mid \mathbb{D}_t) &\sim [m_t,C_t] \quad \textit{(posterior at time } t\textit{)},
\end{aligned}$$

where

$$\begin{aligned}
b_t &= G_t m_{t-1}, & R_t &= G_t C_{t-1} G_t' + H_t,\\
f_t &= F_t' b_t, & Q_t &= F_t' R_t F_t,\\
m_t &= b_t + R_t F_t (f_t^* - f_t)/Q_t, & C_t &= R_t - R_t F_t F_t' R_t (1 - Q_t^*/Q_t)/Q_t,\\
f_t^* &= \frac{\alpha_t^*}{\beta_t^*}, & Q_t^* &= \frac{\alpha_t^{*2}}{(\beta_t^*)^2(\beta_t^*-1)}\\
\alpha_t^* &= \alpha_t + \mathfrak{a}_t y_t, & \beta_t^* &= \beta_t + \mathfrak{a}_t,\\
\alpha_t &= \frac{f_t^3}{Q_t} + f_t, & \beta_t &= \frac{f_t^2}{Q_t} + 1.
\end{aligned}$$

PROOF Consider the updating procedure by Kristensen et al. (2010, Section 8.5.4) which consists of seven steps. The first three steps are the same, but repeated below for readability.

a) Posterior information for θ_{t-1} at time $t-1$:

$$(\theta_{t-1} \mid \mathbb{D}_{t-1}) \sim [m_{t-1}, C_{t-1}],$$

b) Prior for θ_t at time t:

$$(\theta_t \mid \mathbb{D}_{t-1}) \sim [b_t, R_t], \quad b_t = G_t m_{t-1},\, R_t = G_t C_{t-1} G_t' + H_t.$$

c) Prior for $g(\eta_t)$ at time t:

$$(g(\eta_t) \mid \mathbb{D}_{t-1}) \sim [f_t, Q_t], \quad f_t = F_t' b_t,\, Q_t = F_t' R_t F_t.$$

d) Approximate full prior for η_t at time t: According to our assumptions we have that $(g(\eta_t) \mid \mathbb{D}_{t-1}) \sim \text{Inv-Gamma}(\mathfrak{c}_t, \mathfrak{d}_t)$ where $\mathfrak{c}_t$ and $\mathfrak{d}_t$ are the shape and scale parameters, $g(\eta_t) = -1/\eta_t$, and that the density of y_t is (2.14).

In this step we need to identify the conjugate prior of $\eta_t = -1/g(\eta_t)$ using the general form of the conjugate prior with two parameters α_t and β_t (Kristensen et al., 2010):

$$f(\eta_t|\mathbb{D}_{t-1}) = c(\alpha_t, \beta_t)\exp(\alpha_t \eta_t - \beta_t a(\eta_t)),$$

where $c(\alpha_t, \beta_t)$ is a known function and $a(\eta_t) = \ln(-1/\eta_t)$ as defined in (2.12). If we suppose $y = \eta_t$ and $x = g(\eta_t)$, i.e. $y = h(x) = \frac{-1}{x}$, then by applying the transformation rule, the density function of η_t is

$$\begin{aligned}
f(\eta_t|\mathbb{D}_{t-1}) &= f_x(h^{-1}(y)|\mathbb{D}_{t-1})\frac{\partial h^{-1}(y)}{\partial y} \\
&= \frac{\mathfrak{d}_t^{\mathfrak{c}_t}}{\Gamma(\mathfrak{c}_t)}(-1/\eta_t)^{-\mathfrak{c}_t-1}\exp\left(-\frac{\mathfrak{d}_t}{-1/\eta_t}\right)\frac{1}{\eta_t^2} \\
&= \frac{\mathfrak{d}_t^{\mathfrak{c}_t}}{\Gamma(\mathfrak{c}_t)}(-\eta_t)^{\mathfrak{c}_t-1}\exp(\mathfrak{d}_t\eta_t) \\
&= \frac{\mathfrak{d}_t^{\mathfrak{c}_t}}{\Gamma(\mathfrak{c}_t)}\exp(\mathfrak{d}_t\eta_t - (\mathfrak{c}_t - 1)\ln(\frac{-1}{\eta_t})).
\end{aligned}$$

Hence the parameters α_t and β_t in the conjugate prior of η_t become:

$$\alpha_t = \mathfrak{d}_t, \;\; \beta_t = \mathfrak{c}_t - 1. \tag{2.15}$$

Finally, we fit α_t and β_t such that

$$\begin{aligned}
\mathbb{E}(g(\eta_t) \mid \mathbb{D}_{t-1}) &= \frac{\mathfrak{d}_t}{\mathfrak{c}_t - 1} = \frac{\alpha_t}{\beta_t} = f_t, &(2.16)\\
\text{Var}(g(\eta_t) \mid \mathbb{D}_{t-1}) &= \frac{\mathfrak{d}_t^2}{(\mathfrak{c}_t-1)^2(\mathfrak{c}_t-2)} = \frac{\alpha_t^2}{(\beta_t)^2(\beta_t-1)} = Q_t,
\end{aligned}$$

implying that

$$\alpha_t = \frac{f_t^3}{Q_t} + f_t, \quad \beta_t = \frac{f_t^2}{Q_t} + 1.$$

e) One step forecast of y_t: In this step we need to find the forecast distribution $f(y_t|\mathbb{D}_{t-1})$. According to the concepts of the nGSSM models, the general form of this distribution with two parameters α_t and β_t is (Kristensen et al., 2010):

$$f(y_t|\mathbb{D}_{t-1}) = \frac{c(\alpha_t, \beta_t) q(y_t, \mathfrak{u}_t)}{c(\alpha_t + \phi_t x(y_t), \beta_t + \phi_t)},$$

where $q(y_t, \mathfrak{u}_t)$ and $x(y_t)$ have been defined in (2.12) and $\phi_t = \frac{1}{\mathfrak{u}_t}$. Using the values of α_t and β_t found in Step d, the forecast distribution $f(y_t|\mathbb{D}_{t-1})$ equals

$$f(y_t|\mathbb{D}_{t-1}) = \frac{1}{B(\mathfrak{a}_t, \mathfrak{c}_t)} \cdot \frac{1}{\mathfrak{d}_t/\mathfrak{a}_t} \cdot \left(\frac{y_t - 0}{\mathfrak{d}_t/\mathfrak{a}_t}\right)^{\mathfrak{a}_t - 1} \cdot \left(1 + \frac{y_t - 0}{\mathfrak{d}_t/\mathfrak{a}_t}\right)^{-\mathfrak{a}_t - \mathfrak{c}_t},$$

where $B(\mathfrak{a}_t, \mathfrak{c}_t) = \frac{\Gamma(\mathfrak{a}_t)\Gamma(\mathfrak{c}_t)}{\Gamma(\mathfrak{a}_t + \mathfrak{c}_t)}$. That is a *generalized beta prime* distribution denoted by β' (Crooks, 2013, page 50) and hence

$$(y_t|\mathbb{D}_{t-1}) \sim \beta'(\psi_1, \psi_2, \psi_3, \psi_4, \psi_5), \tag{2.17}$$

with parameters: $\psi_1 = 0$ (location); $\psi_2 = \mathfrak{d}_t/\mathfrak{a}_t$ (scale); $\psi_3 = \mathfrak{a}_t$ (first shape); $\psi_4 = \mathfrak{c}_t$ (second shape); and $\psi_5 = 1$ (Weibull power parameter).

f) Posterior distributions for η_t and $g(\eta_t)$ at time t: The general form of posterior density function of η_t is:

$$f(\eta_t|\mathbb{D}_t) = c(\alpha_t^*, \beta_t^*) \exp(\alpha_t^* \eta_t - \beta_t^* a(\eta_t)),$$

with

$$\alpha_t^* = \alpha_t + \phi_t y_t = \alpha_t + \mathfrak{a}_t y_t, \quad \beta_t^* = \beta_t + \phi_t = \beta_t + \mathfrak{a}_t. \tag{2.18}$$

The last relation follows from $\phi_t = \frac{1}{\mathfrak{u}_t} = \mathfrak{a}_t$. If we suppose $x = \eta_t$ and $y = g(\eta_t)$ then, based on relation $g(\eta_t) = -1/\eta_t$, we have $y = h(x) = \frac{-1}{x}$ and using the transformation rule, the posterior distribution of $g(\eta_t)$ is:

$$\begin{aligned} f_{g(\eta_t)}(g(\eta_t)|\mathbb{D}_t) &= f_x(h^{-1}(y)|\mathbb{D}_t)\frac{\partial h^{-1}(y)}{\partial y} \\ &= \frac{(\alpha_t^*)^{\beta_t^*+1}}{\Gamma(\beta_t^*+1)}\exp(\frac{-\alpha_t^*}{g(\eta_t)} - \beta_t^*\ln(g(\eta_t)))\frac{1}{g(\eta_t)^2} \\ &= \frac{(\alpha_t^*)^{\beta_t^*+1}}{\Gamma(\beta_t^*+1)}(g(\eta_t))^{-(\beta_t^*+1)-1}\exp(-\frac{\alpha_t^*}{g(\eta_t)}). \end{aligned} \tag{2.19}$$

It follows from (2.19) that $(g(\eta_t) \mid \mathbb{D}_t) \sim$ Inv-Gamma$(\mathfrak{c}_t^*, \mathfrak{d}_t^*)$ with

$$\mathfrak{c}_t^* = \beta_t^* + 1, \quad \mathfrak{d}_t^* = \alpha_t^*.$$

Next, we fit α_t^* and β_t^* such that

$$\begin{aligned} f_t^* &= \mathbb{E}(g(\eta_t) \mid \mathbb{D}_t) = \frac{\mathfrak{d}_t^*}{\mathfrak{c}_t^* - 1} = \frac{\alpha_t^*}{\beta_t^*}, \\ Q_t^* &= \mathrm{Var}(g(\eta_t) \mid \mathbb{D}_t) = \frac{(\mathfrak{d}_t^*)^2}{(\mathfrak{c}_t^*-1)^2(\mathfrak{c}_t^*-2)} = \frac{(\alpha_t^*)^2}{(\beta_t^*)^2(\beta_t^*-1)}, \end{aligned}$$

g) Posterior of θ_t at t: The posterior parameters m_t and C_t are

$$m_t = b_t + R_tF_t(f_t^* - f_t)/Q_t, \quad C_t = R_t - R_tF_tF_t'R_t(1 - Q_t^*/Q_t)/Q_t.$$

Since the above steps calculate the values stated in Theorem 3, this finishes the proof.

Corollary 1 *Given Theorem 3 and $F_t = 1$ we have that*

$$\begin{aligned} f_t &= b_t, & Q_t &= R_t, \\ m_t &= f_t^*, & C_t &= Q_t^*. \end{aligned}$$

Finally, the probability distribution of $(m_{t+1} \mid m_t)$ can be found.

Theorem 4 *Under Theorem 3 and Corollary 1 and assuming $H_t = 0$, the conditional random variable $(m_{t+1} \mid m_t)$ follows a generalized beta prime distribution. That is,*

$$(m_{t+1} \mid m_t) \sim \beta'(\psi_1, \psi_2, \psi_3, \psi_4, \psi_5),$$

with parameters: $\psi_1 = G_{t+1}m_t\beta_t^*/(\beta_t^* + \mathfrak{a}_{t+1})$ (location), $\psi_2 = \psi_1$ (scale), $\psi_3 = \mathfrak{a}_{t+1}$ (first shape), $\psi_4 = \beta_t^* + 1$ (second shape), and $\psi_5 = 1$ (Weibull power) where $\mathfrak{a}_{t+1}$ is the shape parameter of the exponential family distribution of y_{t+1}.

PROOF Assume that Theorem 3 and Corollary 1 hold. Then

$$(m_{t+1} \mid m_t) = (f^*_{t+1} \mid m_t) = (\frac{\alpha^*_{t+1}}{\beta^*_{t+1}} \mid m_t) = A_{t+1} + B_{t+1}(y_{t+1}|\mathbb{D}_t),$$

since based on (2.18) we have that

$$\frac{\alpha^*_{t+1}}{\beta^*_{t+1}} = A_{t+1} + B_{t+1}y_{t+1},$$

where

$$A_{t+1} = \frac{\alpha_{t+1}}{\beta_{t+1} + \mathfrak{a}_{t+1}}, \quad B_{t+1} = \frac{\mathfrak{a}_{t+1}}{\beta_{t+1} + \mathfrak{a}_{t+1}}.$$

From (2.16) and since $H_t = 0$, we have that

$$\beta_{t+1} = \frac{\alpha_{t+1}}{f_{t+1}} = \frac{f^2_{t+1}}{Q_{t+1}} + 1 = \frac{(G_{t+1}f^*_t)^2}{G_{t+1}Q^*_t G'_{t+1}} + 1 = \frac{f^{*2}_t}{Q^*_t} + 1 = \left(\frac{\alpha^*_t}{\beta^*_t}\right)^2 \frac{\alpha^{*2}_t}{\beta^{*2}_t(\beta^*_t - 1)} + 1 = \beta^*_t,$$

which implies that

$$\alpha_{t+1} = f_{t+1}\beta_{t+1} = f_{t+1}\beta^*_t = G_{t+1}m_t\beta^*_t.$$

As a result we can compute A_{t+1} and B_{t+1} as

$$A_{t+1} = \frac{G_{t+1}m_t\beta^*_t}{\beta^*_t + \mathfrak{a}_{t+1}}, \quad B_{t+1} = \frac{\mathfrak{a}_{t+1}}{\beta^*_t + \mathfrak{a}_{t+1}},$$

which are two scalars given m_t (since parameters $\mathfrak{a}_{t+1}$ and β^*_t are known values given t).

Recall from (2.17), we have that $(y_{t+1} \mid \mathbb{D}_t) \sim \beta'(\breve{\psi}_1, \breve{\psi}_2, \breve{\psi}_3, \breve{\psi}_4, \breve{\psi}_5)$, with parameters: $\breve{\psi}_1 = 0$ (location), $\breve{\psi}_2 = \mathfrak{d}_{t+1}/\mathfrak{a}_{t+1}$ (scale), $\breve{\psi}_3 = \mathfrak{a}_{t+1}$ (first shape), $\breve{\psi}_4 = \mathfrak{c}_{t+1}$ (second shape), and $\breve{\psi}_5 = 1$ (Weibull power parameter). Hence we have that

$$(m_{t+1} \mid m_t) \sim \beta'(\psi_1, \psi_2, \psi_3, \psi_4, \psi_5),$$

with parameters: $\psi_1 = A_{t+1}$, $\psi_2 = B_{t+1}\mathfrak{d}_{t+1}/\mathfrak{a}_{t+1}$, $\psi_3 = \mathfrak{a}_{t+1}$, $\psi_4 = \mathfrak{c}_{t+1}$, and $\psi_5 = 1$. Here we have used the property that if

$$X \sim \beta'(\psi_1, \psi_2, \psi_3, \psi_4, \psi_5),$$

then (based on the transformation rule)

$$a + bX \sim \beta'(a + \psi_1, b\psi_2, \psi_3, \psi_4, \psi_5).$$

Note that due to (2.15), we have that

$$\psi_2 = B_{t+1}\mathfrak{d}_{t+1}/\mathfrak{a}_{t+1} = \frac{\mathfrak{d}_{t+1}}{\beta_t^* + \mathfrak{a}_{t+1}} = \frac{\alpha_{t+1}}{\beta_t^* + \mathfrak{a}_{t+1}} = A_{t+1},$$

and

$$\psi_4 = \mathfrak{c}_{t+1} = \beta_{t+1} + 1 = \beta_t^* + 1,$$

which finishes the proof.

Chapter 3

Paper II: Slaughter pig marketing under price fluctuations

History: *This paper was prepared in collaboration with Lars Relund Nielsen. It has been submitted to Annals of Operations Research. The paper has been presented at INFORMS annual meeting 2015, November 2015, Philadelphia, USA.*

Slaughter pig marketing under price fluctuations

Reza Pourmoayed and Lars Relund Nielsen

CORAL, Department of Economics and Business Economics, Aarhus University, Fuglesangs Allé 4, DK-8210 Aarhus V, Denmark.

Abstract: In the production of fattening pigs, pig marketing refers to a sequence of culling decisions until the production unit is empty. The profit of a production unit is highly dependent on the price of pork, the cost of feeding and the cost of buying piglets. Price fluctuations in the market consequently influence the profit, and the optimal marketing decisions may change under different price conditions. In this paper we formulate a hierarchical Markov decision process with two levels which model sequential marketing decisions under price fluctuations in a pig pen. The state of the system is based on information about pork, piglet and feed prices. Moreover, the information is updated using a Bayesian approach and embedded into the hierarchical Markov decision process. The optimal policy is analyzed under different patterns of price fluctuations. We also assess the value of including price information into the model.

Keywords: OR in agriculture; Markov decision process; herd management; Bayesian updating; price fluctuations.

3.1 Introduction

In the production of fattening pigs, one of the main managerial decisions is *pig marketing* (Kure, 1997). It refers to a sequence of culling decisions until the production unit is empty. The profit at marketing depends on endogenous factors such as growth, housing conditions and management policy as well as exogenous factors such as market prices. Prices of pork, piglets and feed may fluctuate in the market on a weekly basis and hence the farmer should take into account the influence of price fluctuations when he decides when to send animals to the abattoir.

In a production system of growing/finishing pigs (Danish standards), animals may be considered at different levels: herd, section, pen, or animal. The herd is a group of sections, a section includes some pens, and a pen involves some animals (usually 15-20). When the production process of growing/finishing (fattening) pigs is started, the farmer either buys piglets on the market or transfers them from another production unit when they weigh approx. 30 kg. The piglets are inserted into a *finisher pen* where they grow until marketing (9-15 weeks). Since pigs in general grow with different growth rates, they obtain their slaughter weight at different times in the last weeks of the growing period. At the end of the growing period the farmer should therefore determine which pigs should be selected for slaughter (individual marketing). The reward of marketing a pig depends on the *pork price* of the carcass weight, the cost of buying the piglet on the market, i.e. the *piglet price* and the cost of feeding which is dependent on the *feed price* at the time when the feed stock is bought (e.g. at the start of the production cycle). Next, after a sequence of individual marketings, the farmer must decide when to *terminate* (empty) the rest of the pen. Terminating a pen means that the remaining pigs in the pen are sent to the slaughterhouse (in one delivery) and after cleaning the pen, another group of piglets (each weighing approx. 30 kg) is inserted into the pen and *the production cycle* is repeated. That is, the farmer must time the marketing decisions while simultaneously considering the carcass weight, the length of the production cycle and exogenous price conditions. For an extended overview over pig production of growing pigs, see Pourmoayed and Nielsen (2014).

Various studies have considered pig marketing (see e.g. Ohlmann and Jones (2008); Kristensen et al. (2012), and Khamjan et al. (2013)). However, in these studies, the marketing policy has been investigated under constant price conditions. Only a few studies take price fluctuations into account. Broekmans (1992) analyzed the effect of price fluctuations on the marketing policy of fattening pigs by using a two-level hierarchical Markov decision process. He analysed price fluctuations by a first order autoregressive model proposed by Jørgensen (1992). Due to the curse of dimensionality, in this study a limited range of possible price values was considered in the problem such that the state variables related to the price information were divided into a limited number of groups. Moreover, learning aspects of price parameters from the historical data were not taken into account in this research. In the study by Roemen and de Klein (2000), only a fluctuating pork price was considered and the piglet price was modeled as a constant factor of the pork price. They used a Markov decision process to model the sequential marketing decisions under pork price fluctuations but no numerical example was given to show the efficiency of the proposed model.

In order to close this gap in the literature, we consider pig marketing at pen level under three price fluctuations, namely, the pork, piglet and feed prices. A *hierarchical Markov decision process* (HMDP) with two levels is used to model the sequential decisions of marketing at pen level. The state of the system is based on information about pork, piglet and feed prices. The model considers time series of pork, piglet and feed prices obtained from the market and a learning approach based on Bayesian updating is applied to update price information using the historical data which is embedded into the HMDP. More precisely, we use three state space models for *Bayesian forecasting* (West and Harrison, 1997) to update the future estimates of pork, piglet and feed prices on a weekly basis. Numerical examples are given to analyze the optimal decisions under different patterns of price fluctuations and to evaluate the value of including price information into the model.

The paper is organized as follows. First, Section 3.2 gives a short literature review. In Section 3.3, the optimization model is explained in detail. Section 3.4 gives an overview on how Bayesian updating is used to update price information and describes a procedure for embedding the statistical models into the HMDP. Next, in Section 3.5 we test the model under different scenarios and evaluate the value of including price information into the HMDP, and finally in Section 3.6, we conclude the paper.

3.2 Literature review

The problem of finding the optimal pig marketing policy, i.e. an optimal sequence of culling decisions until the production unit is empty, has been studied by a variety of researchers under different conditions.

Jørgensen (1993) proposed a stochastic dynamic programming model with a probabilistic growth function to find the best marketing policy of fattening pigs and examined the value of weighing precision and its relation with the marketing decisions. Chavas et al. (1985) showed the importance of the animal growth on the marketing decisions by using the concepts of optimal control theory. Kure (1997) applied the principles of replacement theory to find the best timing of marketing decisions in a batch of animals. Toft et al. (2005) used a multi-level hierarchical Markov decision process to optimize the delivery strategy of pigs to the abattoir and to control epidemic diseases simultaneously. Boys et al. (2007) determined the best marketing policy of pigs using a simulation approach to utilize the maximum capacity of trucks for delivering the pigs to the abattoir. Ohlmann and Jones (2008) considered the effect of stocking space and shipping on the problem and found the best timing of delivery to the packers using a mixed-integer linear programming model. Kristensen et al. (2012) proposed a two-level hierarchical Markov decision process and a state space model to optimize the marketing policy of the farm under online information acquired from sensor data. In the study by Plà-Aragonés et al. (2013), the optimal marketing policy was found by a mixed integer linear programming method under an all-in all-out strategy. Khamjan et al. (2013) considered a two level supply chain of fattening units (as supplier) and slaughterhouse (as buyer) to find the best procurement plan of buying the pigs from a zone of farms. They formulated the problem by a mathematical programming model and solved their model using a heuristic approach under different pig size distributions and pig growth rates.

The above mentioned studies investigated the marketing policy under constant price conditions. Only a few studies take price fluctuations into account. Broekmans (1992) analyzed the effect of price fluctuations on the marketing policy of fattening pigs using a first order autoregressive model proposed by Jørgensen (1992) with a limited range of possible price values. In the study by Roemen and de Klein (2000), a Markov model with fluctuating pork prices was suggested.

Few studies take both marketing and feeding decisions into account. Niemi (2006) used a mechanistic function to model the animal growth trend during the growing period. Niemi (2006) further applied a stochastic dynamic programming method to find the best nutrient ingredients and also the best time of marketing. In the study by Sirisatien et al. (2009), a genetic algorithm was used to find a set of feeding schedules followed by the optimal values of the nutrient ingredients and feeding period. Both studies considered the problem at animal level and did not take into account the inhomogeneity of animals with respect to growth and feed conversion rate. In a recent study, Pourmoayed et al. (2016) have considered optimal marketing and feeding strategies in a finisher pen.

Markov decision models are a well-known modeling technique within animal science used to model livestock systems. See for instance Rodriguez et al. (2011) and Nielsen et al. (2010). For a recent survey see Nielsen and Kristensen (2014), which cites more than 100 papers using (hierarchical) Markov decision processes to model and optimize livestock systems. An HMDP is an extension of a semi Markov decision process (semi-MDP) where a series of finite-horizon semi-MDPs are combined into one process at the founder level called the main/founder process (Kristensen, 1988; Kristensen and Jørgensen, 2000). As a result, the state space at the founder level can be reduced, and larger models can be solved using a modified policy iteration algorithm under different criteria (Nielsen and Kristensen, 2014).

In this paper we will model price fluctuations using a *state space model* (*SSM*) (West and Harrison, 1997). An SSM is a statistical model which may be used to transform time-series obtained from the market or via online sensors into the required information needed by e.g. the HMDP. Applications of SSMs in time-series analysis of e.g. price data can be found in Durbin

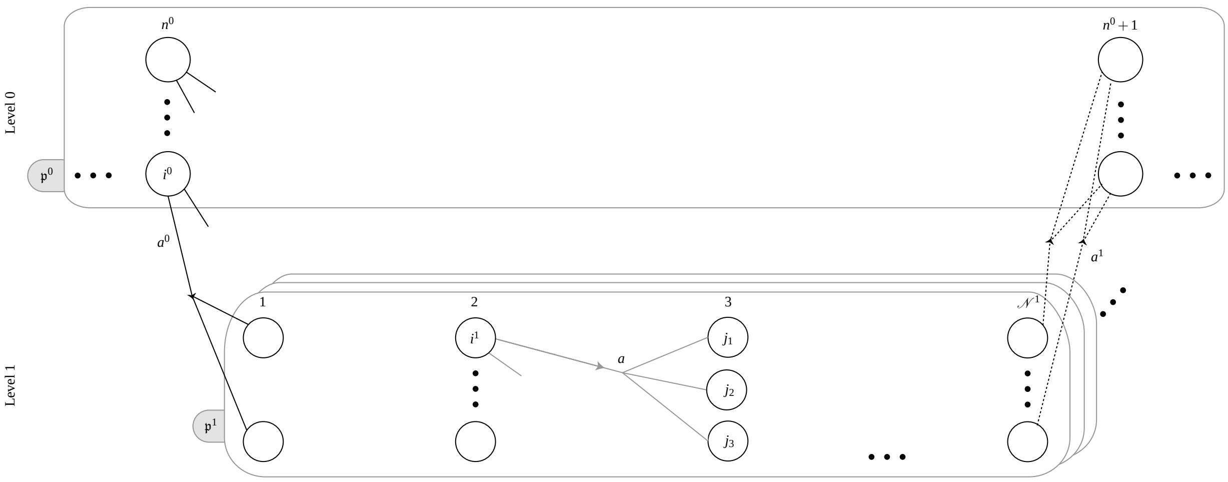


Figure 3.1: An illustration of a stage in an HMDP. At the founder level (Level 0) there is a single infinite-horizon founder process $\mathfrak{p}^0$. A child process, such as $\mathfrak{p}^1$ at Level 1 (oval box), is uniquely defined by a given stage, state (node), and action (hyperarc) of its parent process and linked with the parent process using its initial probability distribution (solid lines) and its terminating actions (dashed lines). Each process at level 2 is a semi-MDP. Note that only a subset of the actions have been shown in the figure.

and Koopman (2012). Examples of SSMs applied to agricultural problems are Bono et al. (2012); Cornou et al. (2008) and Bono et al. (2013). Moreover, an SSM can be discretized and embedded into an HMDP (Nielsen et al., 2011).

3.3 Optimization model

In this study, sequential pig marketing decisions are modeled using a *hierarchical Markov decision process* (*HMDP*). A short introduction to HMDPs is given below. Since techniques from both statistical forecasting and operations research are used, a consistent notation can be hard to specify. To assist the reader, Appendix 3.A provides an overview over the notation.

An HMDP is an extension of a *semi-Markov decision process* (*semi-MDP*) where a series of finite-horizon semi-MDPs are combined into one infinite time-horizon process at the founder level called the *founder process* (Kristensen and Jørgensen, 2000). The idea is to expand stages of a process to so-called *child processes*, which again may expand stages further to new child processes leading to multiple levels. At the lowest level, the HMDP consists of a set of finite-horizon semi-MDPs (see e.g. Tijms, 2003, Chap. 7). All processes are linked together using jump actions (see Figure 3.1).

In order to have a frame of reference, we exploit the notation used for a semi-MDP and extend it to an HMDP. A finite-horizon semi-MDP models a sequential decision problem over $\mathscr{N}$ stages. Let $\mathbb{I}_n$ denote the finite set of system states at stage n. Given system state $i \in \mathbb{I}_n$ at stage n, an action a from the finite set of allowable actions $\mathbb{A}_n(i)$ is chosen generating two outcomes: an immediate reward $r_n(i,a)$ and a probabilistic transition to state $j \in \mathbb{I}_{n+1}$ at stage $n+1$ with *transition probability* $\Pr(j \mid n,i,a)$. Moreover, let $u_n(i,a)$ denote the *stage length* of action a, i.e. the expected time until the next decision epoch (stage $n+1$) given action a and state i.

An HMDP with two levels is illustrated in Figure 3.1 using a *state-expanded hypergraph* (Nielsen and Kristensen, 2006). At the first level, a single *founder process* $\mathfrak{p}^0$ is defined. Index 0 indicates that the process has no ancestral processes. Process $\mathfrak{p}^0$ is running over an infinite number of stages and all stages have identical state and action spaces and hence just a single stage is illustrated in Figure 3.1. Let $\mathfrak{p}^{l+1}$ denote a *child process* at level $l+1$. Process $\mathfrak{p}^{l+1}$ is uniquely defined by a given stage n^l, state i^l and action a^l of *parent process* $\mathfrak{p}^l$. For instance, the semi-MDP $\mathfrak{p}^1$ in Figure 3.1 is defined at stage n^0, state i^0 and action a^0 of the founder process $\mathfrak{p}^0$ symbolized by the notation $\mathfrak{p}^1 = (\mathfrak{p}^0 \parallel (n^0, i^0, a^0)\ )$. Each process is connected to its parent and child processes using *jump actions* which can be divided into two groups, namely, a *child jump action* that represents an *initial probability distribution* of transitions to a child process or a *parent jump action* that represents a *terminating probability distribution* of transitions to a parent process. This is illustrated in Figure 3.1 where child jump action a^0 (illustrated using a solid *hyperarc*) represents a transition to the child process $\mathfrak{p}^1$ and parent jump action a^1 (illustrated using a dashed hyperarc) represents termination of the process $\mathfrak{p}^1$. Jump actions are like the traditional actions associated with an expected reward, action length, and a set of transition probabilities. Each node in Figure 3.1 at a given stage n of a process $\mathfrak{p}^l$ corresponds to a state in $\mathbb{I}_n$. For example, there are three states at stage 3 in process $\mathfrak{p}^1$. Similarly, each gray hyperarc corresponds to an action, e.g. action a results in a transition from state i^1 to either state j_1, j_2 or j_3.

A policy is a decision rule/function that assigns to each state in a process a (jump) action. That is, choosing a policy corresponds to choosing a single hyperarc out of each node in Figure 3.1. Given a policy, the reward at a stage of a parent process equals the total expected rewards of the corresponding child process. For instance, in Figure 3.1, the reward of choosing action a^0 in state i^0 at stage n^0 in process $\mathfrak{p}^0$ equals the total expected reward of process $\mathfrak{p}^1$. With a similar approach, the transition probabilities and the stage length of an action can be calculated at a stage of a parent process.

Different optimality criteria may be considered. In this paper, our optimality criterion is to maximize the *expected reward per time unit* and the optimal policy of the HMDP is found using a modified policy iteration algorithm. For a detailed description of the algorithm, the interested reader may consult Kristensen and Jørgensen (2000) and Nielsen and Kristensen (2014).

3.3.1 Assumptions

The HMDP which models marketing decisions in a finisher pen is formulated under the following assumptions:

1. The fixed number of pigs inserted into the pen at the beginning of each production cycle is $q^{\max}$.
2. Marketing of pigs is started in week $t^{\min}$ at the earliest;
3. The pen is terminated in week $t^{\max}$ at the latest, i.e. the maximum life time of a pig in the pen is $t^{\max}$.
4. The sequence of feed-mixes used during the production cycle (feeding strategy) is known and fixed.
5. When a marketing decision happens, the preparation time for delivering the pigs to the abattoir is b.
6. Weekly deliveries to the abattoir in the marketing period are based on a cooperative agreement where culled pigs from each pen are grouped into one delivery, i.e. the transportation cost is fixed.
7. Marketing decisions are taken on a weekly basis, and a decision must be taken b days before each delivery.
8. After terminating the pen, the length of the period for cleaning the pen is h.
9. A new batch of piglets and the required feed stock are bought using market prices at the start of each production cycle.
10. The growth of a pig is independent of the other pigs in the pen, i.e. the growth does not depend on the number of pigs in the pen.
11. Pigs are sold to the abattoir using the market pork price.

To give a complete description of the two-level HMDP with marketing decisions, the characteristics of each semi-MDP should be specified at all levels, i.e. stages, states, and (jump) actions including the corresponding rewards, stage lengths (measured in weeks), and transition probabilities.

3.3.2 Stages, states and actions

As illustrated in Figure 3.1, the founder process of the HMDP is an infinite time-horizon process where a stage represents a lifetime of $q^{\max}$ pigs inserted into the pen (until termination). A stage of a process at the second level corresponds to either the period from insertion of the piglets until the marketing starts or a week in the marketing period (weeks $t^{\min}$ to $t^{\max}$). The length, stage, states, and (jump) actions of each process at levels 0 and 1 are described below. To avoid heavy notation, the superscript indicating the current level under consideration is left out whenever the level is clear from the context.

Level 0 - Founder process $\mathfrak{p}^0$

Stage: A production cycle of $q^{\max}$ pigs, i.e. from inserting the piglets into the pen until terminating the pen.

Time horizon: Infinite (the pen is filled and emptied an infinite number of times).

States: Due to the infinite time horizon, the state space is homogeneous and hence the stage index can be ignored when a state is defined at the founder process. A state $i^0 = \mathbb{p} \in \mathbb{P}$ represents our information about the pork, piglet and feed prices (i.e. $\mathbb{I} = \mathbb{P}$). The price information is obtained from the market. Definition of $\mathbb{p}$ is given in Section 3.4.

Actions: For each state, a single child jump action a^0 (insertion of the piglets into the pen) is defined representing the initial probability distribution of transitions to the child process. The length of this action is zero.

Since the stage index can be ignored and there is only a single action, a child process is uniquely defined for each state $i^0 = \mathbb{p}$. That is, child process $\mathfrak{p}^1 = (\mathfrak{p}^0||n^0, i^0, a^0)$ is equivalent to $\mathfrak{p}^1 = (\mathfrak{p}^0||\mathbb{p})$.

Level 1 - Child process $\mathfrak{p}^1 = (\mathfrak{p}^0||n^0, i^0, a^0)$

Stage: The first stage ($n = 1$) represents the period from insertion of the piglets (week 1) until the start of marketing decisions (week $t^{\min}$). The remaining stages ($n > 1$) represent a week in the marketing period (weeks $t^{\min}$ to $t^{\max}$). That is, stage $n = 1$ corresponds to the start of week 1 and stage $n > 1$ the start of week $n + t^{\min} - 2$.

Time horizon: Due to the definition of stages, the maximum number of stages is $\mathscr{N} = t^{\max} - t^{\min} + 2$.

States: Given stage n, state i is defined using state variables:

$\mathbb{d}_n$: information related to the deviations from the pork, piglet and feed price information given in state i^0, acquired using Bayesian updating ($\mathbb{d}_n \in \mathbb{D}_n$). This information is obtained using the SSMs explained in Section 3.4;

q_n: number of pigs in the pen at the beginning of stage n.

Note that if $n \leq 2$ then $q_n = q^{\max}$. Hence the set of states becomes

$$\mathbb{I}_n = \left\{ i = (\mathbb{d}_n, q_n) \mid \mathbb{d}_n \in \mathbb{D}_n, q_n \in \{1 \cdot \mathbf{I}_{\{n>2\}} + q^{\max} \cdot \mathbf{I}_{\{n\leq 2\}}, \dots, q^{\max}\} \right\},$$

where $\mathbf{I}_{\{\cdot\}}$ denotes the indicator function.

Actions: Consider state $i = (\mathbb{d}_n, q_n)$ at stage n. If $n = 1$, then marketing is not possible and the production process continues using action $a_{\texttt{cont}}$. If $1 < n < \mathscr{N}$, then the set of actions are $a_{\texttt{cont}}$, the parent jump action $a_{\texttt{term}}$ showing the pen is terminated, and actions a_q implying that the q heaviest pigs are culled (individual marketing). Finally, at the last stage $n = \mathscr{N}$, the pen must be terminated ($a_{\texttt{term}}$). Hence the set of actions becomes:

$$\mathbb{A}_n(i) = \begin{cases} \{a_{\texttt{cont}}\} & n = 1 \\ \{a_{\texttt{term}}, a_{\texttt{cont}}\} \cup \{a_q \mid 1 \leq q < q_n\} & 1 < n < \mathscr{N}, \\ \{a_{\texttt{term}}\} & n = \mathscr{N}. \end{cases} \tag{3.1}$$

The length of action $a_{\texttt{cont}}$ at stage 1 is $t^{\min} - 1$ weeks while for stage $n > 1$ the lengths of actions $a_{\texttt{cont}}$ and a_q are one week. For action $a_{\texttt{term}}$ the length is $h + b$ days.

3.3.3 Transition probabilities

Founder process $\mathfrak{p}^0$

Given stage n^0 and state $i^0 = \mathbb{p}$, a single child jump action a^0 was defined with a transition between the levels of the HMDP from state i^0 to state $j^1 = (\mathbb{d}_1, q_1)$ at the first stage of process $\mathfrak{p}^1$ ($n^1 = 1$). Since $q_1 = q^{\max}$, the transition probability becomes

$$\Pr\left(j^1 \mid n^0, i^0, a^0\right) = \Pr\left(\mathbb{d}_1 \mid \mathbb{p}\right), \tag{3.2}$$

where $\Pr(\mathbb{d}_1 \mid \mathbb{p})$ is the initial probability of price deviations $\mathbb{d}_1$ given price information $\mathbb{p}$. The probability $\Pr(\mathbb{d}_1 \mid \mathbb{p})$ depends on the statistical models used for Bayesian updating of the price information and will be explained in Section 3.4.

Child process $\mathfrak{p}^1 = (\mathfrak{p}^0 || n^0, i^0, a^0)$

As described in (3.1), for a given state $i = (\mathbb{d}_n, q_n)$ at stage n of process $\mathfrak{p}^1$, there are three possible actions $a_{\texttt{cont}}$, a_q and $a_{\texttt{term}}$.

Given actions $a_{\texttt{cont}}$ or a_q, a transition occurs to state $j = (\mathbb{d}_{n+1}, q_{n+1})$ at the next stage of process $\mathfrak{p}^1$. If the process continues without marketing decisions (action $a_{\texttt{cont}}$), the only change of the system is related to the state variable describing price deviations. Hence, the transition probability is

$$\Pr(j \mid n, i, a_{\texttt{cont}}) = \begin{cases} \Pr(\mathbb{d}_{n+1} \mid \mathbb{d}_n) & q_{n+1} = q_n, \\ 0 & \text{otherwise.} \end{cases} \tag{3.3}$$

If the q heaviest pigs are culled from the pen (action a_q), the transition probability becomes

$$\Pr\left(j \mid n, i, a_q\right) = \begin{cases} \Pr(\mathbb{d}_{n+1} \mid \mathbb{d}_n) & q_{n+1} = q_n - q, \\ 0 & \text{otherwise.} \end{cases} \tag{3.4}$$

If the pen is terminated (action $a_{\texttt{term}}$), a new production cycle is started with a transition to state $j^0 = \tilde{\mathbb{p}}$ at the next stage in process $\mathfrak{p}^0$. Hence, the transition probability becomes

$$\Pr\left(j^0 \mid n, i, a_{\texttt{term}}\right) = \Pr(\tilde{\mathbb{p}} \mid \mathbb{d}_n), \tag{3.5}$$

where $\Pr(\tilde{\mathbb{p}} \mid \mathbb{d}_n)$ denotes the terminating probability of parent jump action $a_{\texttt{term}}$.

Probabilities $\Pr(\mathbb{d}_{n+1} \mid \mathbb{d}_n)$ and $\Pr(\tilde{\mathbb{p}} \mid \mathbb{d}_n)$ depend on the statistical models used for Bayesian updating and will be described in Section 3.4.

3.3.4 Expected rewards

Founder process $\mathfrak{p}^0$

At the beginning of a production cycle, $q^{\max}$ piglets are inserted into the pen, i.e. the reward equals the cost of buying new piglets. That is, given state $i = \mathbb{p}$ and action a^0, the expected reward becomes

$$r_n(i, a^0) = -\mathbb{E}\left(p^{\texttt{piglet}} q^{\max}\right), \tag{3.6}$$

where $p^{\texttt{piglet}}$ is the price of one piglet at the beginning of the current production cycle.

Child process $\mathfrak{p}^1 = (\mathfrak{p}^0 || n^0, i^0, a^0)$

At this level, the expected reward equals the expected revenue from selling the pigs minus the expected cost of feeding the remaining pigs conditioned on the values of the state variables and actions.

Consider state $i = (\mathbb{d}_n, q_n)$ at stage n and let $(w_{(1)}, ..w_{(k)}, .., w_{(q_n)})_n$ denote the weight distribution of the pigs in the pen such that $w_{(1)}$, $w_{(k)}$, and $w_{(q_n)}$ are ordered random variables (order statistics) related to the weight of the lightest, kth and the heaviest pigs in the pen at stage n, respectively.

If the process continues without marketing decisions, the reward equals the expected feeding cost of q_n pigs until the next decision epoch

$$r_n(i, a_{\mathtt{cont}}) = -\mathbb{E}\left(p^{\mathtt{feed}} \sum_{k=1}^{q_n} f^{\mathtt{feed}}_{(k),n}(t)\right), \tag{3.7}$$

where $p^{\mathtt{feed}}$ is the feed price of one feed unit (FEsv[1]) at the beginning of the current production cycle and $f^{\mathtt{feed}}_{(k),n}(t)$ denotes the expected feed intake of the kth pig from the start of stage n and the next t days ahead. Note that when $n = 1$ and $n > 1$, t will be equal to $7(t^{\min} - 1)$ and 7, respectively (see Section 3.3.2).

If the q heaviest pigs are culled and the remaining $q_n - q$ pigs are kept in the pen, the expected reward of action a_q becomes

$$\begin{aligned} r_n(i, a_q) = \mathbb{E}\left(\sum_{k=q_n-q+1}^{q_n} \tilde{w}_{(k)} \cdot p^{\mathtt{pork}}_{(k),n}(\tilde{w}_{(k)}, \breve{w}_{(k)})\right) - \mathbb{E}\left(p^{\mathtt{feed}} \sum_{k=q_n-q+1}^{q_n} f^{\mathtt{feed}}_{(k),n}(b)\right) \\ - \mathbb{E}\left(p^{\mathtt{feed}} \sum_{k=1}^{q_n-q} f^{\mathtt{feed}}_{(k),n}(7)\right), \end{aligned} \tag{3.8}$$

where $\tilde{w}_{(k)}$ and $\breve{w}_{(k)}$ denote the carcass weight (kg) and the leanness (non-fat percentage) of the kth pig in the pen at delivery, respectively. The price function $p^{\mathtt{pork}}_{(k),n}(\cdot)$ is the *settlement pork price* of one kg of meat at delivery to the abattoir. This price may be different than the market pork price which is the price given if the pigs are in perfect conditions. In (3.8), the first term is the reward of culling the pigs, the second term is the feeding cost of the culled pigs until they are sent to the abattoir, and the last term is the feeding cost of the remaining pigs.

[1]Danish pig feed unit (1 FEsv = 7.72 MJ)

Finally, if the pen is terminated, the expected reward becomes

$$r_n(i, a_{\mathtt{term}}) = \mathbb{E}\left(\sum_{k=1}^{q_n} \tilde{w}_{(k)} \cdot p^{\mathtt{pork}}_{(k),n}(\tilde{w}_{(k)}, \breve{w}_{(k)})\right) - \mathbb{E}\left(p^{\mathtt{feed}} \sum_{k=1}^{q_n} f^{\mathtt{feed}}_{(k),n}(b)\right). \tag{3.9}$$

To calculate the expected values in equations (3.6) to (3.9), more information is needed: the order statistics of the weights in the pen; transformation of weight to carcass weight and leanness; the feed intake and settlement pork price functions; the pork, feed and piglet prices. A random regression model is used to estimate the mean and standard deviation of weight in the pen at a given week and hence the probability distribution of the ordered weights can be found using well-known formulas. The carcass weight and leanness of a pig can be calculated using biological formulas from the literature. The feed intake function is based on biological relations between weight, growth and feed intake, while the settlement pork price function is a piecewise linear function depending on the carcass weight, leanness and the market pork price. Due to the limited space, further details are given in Appendix 3.B. Finally, information about the pork, feed and piglet prices is embedded into the HMDP using state space models based on Bayesian updating. The state space models are described in the next section.

3.4 Bayesian updating of prices

The revenue of the pigs in a production cycle depends on the pork, piglet, and feed prices which fluctuate on the market every week. Figure 3.2 shows weekly changes of these prices in Denmark in the period of 2006 to the end of 2014[2].

To transform these price data into the information for the HMDP, we need a statistical model for time-series analysis. In our case, due to the non-stationary behavior of price data, we use a *state space model* (SSM) (West and Harrison, 1997). An SSM consists of a set of latent variables and a set of observed variables. At a specified point in time the conditional distribution of the observed variables is a function of the latent variables specified via the observation equations. The latent variables change over time as described via the system equations. The observations are conditionally independent given the latent variables. Thus the estimated value of the latent variables at a time point may be considered as the state of the system, and Bayesian updating (the Kalman filter) can be applied to estimate the latent variables/state of the system via the observed

[2]Time series of pig, piglet and feed prices in Denmark can be found on `http://www.notering.dk/WebFrontend/`. Pork and feed prices are for finisher pigs, and the piglet price is the "30 kg basic" price.

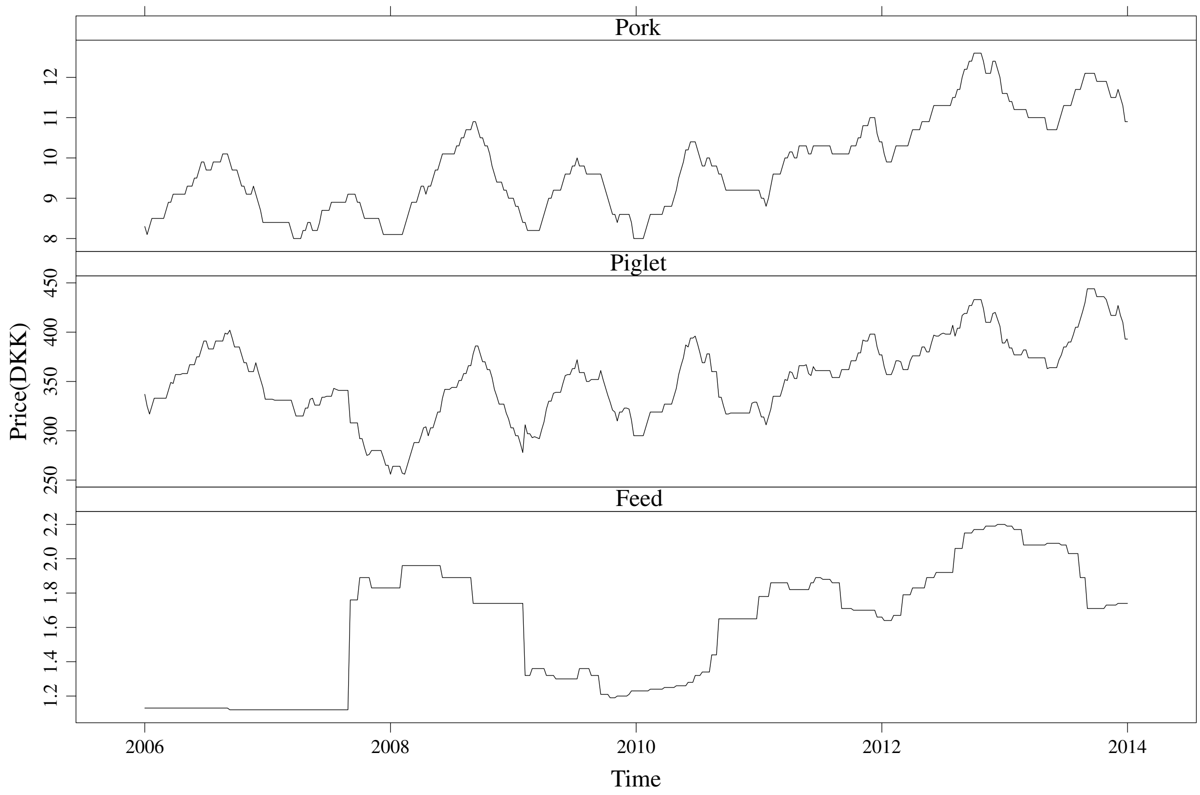


Figure 3.2: Weekly price data of pork, piglet and feed prices in Denmark (years 2006-2014) in DKK. The pork price is the price of the carcass at the abattoir per kilogram when the total carcass weight is between 70 and 95 kg. The piglet price is the price of one piglet with a weight of approx. 30 kg. The feed price is the price per feed unit (FEsv - equivalent to 7.72 MJ).

variables. SSMs can be categorized into different groups based on the dynamic nature of the considered system and the probability distribution assumed for the observed data. In this paper, the probability distribution of the prices is Gaussian and the dynamics of the system is modeled by linear equations. For a short introduction to SSMs and the theorems used for Bayesian updating, see Appendix 3.C.

In the next subsections, we first give a description of the SSMs and afterwards explain how they can be embedded into the HMDP.

3.4.1 SSMs for price prediction

We formulate three SSMs for the pork, piglet and feed prices to be embedded into the HMDP in Section 3.4.2.

Pork price

In order to estimate weekly price deviations and to forecast future pork prices, a *local linear trend* SSM (Durbin and Koopman, 2012, Sec. 3.2.1) is used:

$$\text{Observation equation } (y_t = F'\theta_t + \nu_t): \qquad p_t^{\texttt{pork}} = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} \mu_t^{\texttt{pork}} \\ \lambda_t^{\texttt{pork}} \end{pmatrix}, \tag{3.10}$$

$$\text{System equation } (\theta_t = G\theta_{t-1} + \omega_t): \qquad \begin{pmatrix} \mu_t^{\texttt{pork}} \\ \lambda_t^{\texttt{pork}} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \mu_{t-1}^{\texttt{pork}} \\ \lambda_{t-1}^{\texttt{pork}} \end{pmatrix} + \begin{pmatrix} 0 \\ \omega_t^{\texttt{pork}} \end{pmatrix},$$

where $p_t^{\texttt{pork}}$ is the observed pork price at time t, the latent variable $\lambda_t^{\texttt{pork}}$ represents the deviation of pork price at time t from the price at time $t-1$, $\mu_t^{\texttt{pork}} = p_t^{\texttt{pork}}$ is a supplementary latent variable, and $\omega_t^{\texttt{pork}} \sim N(0, W^{\texttt{pork}})$ is a random term. The initial prior distribution is $\theta_0 \sim N(m_0^{\texttt{pork}}, C_0^{\texttt{pork}})$.

Feed price

A *local level* SSM (Durbin and Koopman, 2012, Pages 9-10) is used to model the feed price:

$$\text{Observation equation } (y_t = F'\theta_t + \nu_t): \quad p_t^{\texttt{feed}} - p^{\texttt{feed}} = \lambda_t^{\texttt{feed}} + \nu_t^{\texttt{feed}}, \tag{3.11}$$

$$\text{System equation } (\theta_t = G\theta_{t-1} + \omega_t): \quad \lambda_t^{\texttt{feed}} = \lambda_{t-1}^{\texttt{feed}} + \omega_t^{\texttt{feed}},$$

where the observed variable $y_t = p_t^{\texttt{feed}} - p^{\texttt{feed}}$ at time t denotes the difference between the current feed price $p_t^{\texttt{feed}}$ and the observed feed price $p^{\texttt{feed}}$ at the start of the current production cycle. The latent variable $\lambda_t^{\texttt{feed}}$ shows the deviation of feed price from $p^{\texttt{feed}}$. $\omega_t^{\texttt{feed}} \sim N(0, W^{\texttt{feed}})$ and $\nu_t^{\texttt{feed}} \sim N(0, V^{\texttt{feed}})$ are two random terms. The initial prior distribution is $\theta_0 \sim N(m_0^{\texttt{feed}}, C_0^{\texttt{feed}})$.

Piglet price

According to Figure 3.2, when the pork price is high, the piglet price $p_t^{\texttt{piglet}}$ is also high and generally they follow each other (e.g. during year 2014 their correlation is 93%). This is a known relation, see e.g. Roemen and de Klein (2000) and Broekmans (1992). Hence, the fraction $p_t^{\texttt{piglet}}/p_t^{\texttt{pork}}$ is approximately constant given t. Therefore, the piglet price can be estimated given the pork price. However, to increase our precision, we may apply a logarithmic transformation and update the deviation between the logarithms of piglet and pig prices using a local level SSM:

$$\text{Observation equation } (y_t = F_t'\theta_t + \nu_t): \quad d_t^{\texttt{piglet}} = \lambda_t^{\texttt{piglet}} + \nu_t^{\texttt{piglet}}, \tag{3.12}$$

$$\text{System equation } (\theta_t = G_t\theta_{t-1} + \omega_t): \quad \lambda_t^{\texttt{piglet}} = \lambda_{t-1}^{\texttt{piglet}} + \omega_t^{\texttt{piglet}},$$

where $d_t^{\texttt{piglet}} = log(p_t^{\texttt{piglet}}) - log(p_t^{\texttt{pork}})$ is the *log transformed observed piglet price ratio*, and $\lambda_t^{\texttt{piglet}}$ is a latent variable for the deviation between these logarithms. $\omega_t^{\texttt{piglet}} \sim N(0, W^{\texttt{piglet}})$ and $v_t^{\texttt{piglet}} \sim N(0, V^{\texttt{piglet}})$ are two random terms. The initial prior is $\theta_0 \sim N(m_0^{\texttt{piglet}}, C_0^{\texttt{piglet}})$.

3.4.2 Embedding the SSMs into the HMDP

The three SSMs described in the previous section provide information about the current prices. Given one of the SSMs, let $\mathbb{D}_{t-1} = (y_1, ..., y_{t-1}, m_0, C_0)$ denote the information available up to time $t-1$. Each time new information is received about a price, Bayesian updating (Theorem 5 in Appendix 3.C) can be used to update the posterior distribution $(\theta_t \mid \mathbb{D}_t) \sim N(m_t, C_t)$ at time t. That is, m_t is our best estimate of the latent variable, i.e. our best estimate of price deviations in addition to the observed market prices.

Hence, in order to embed this information into the HMDP, the state variables $\mathbb{p}$ and $\mathbb{d}_n$ are defined to represent price information at Levels 0 and 1 as

$$\mathbb{p} = \left(p^{\texttt{pork}}, p^{\texttt{feed}}, d^{\texttt{piglet}}\right), \tag{3.13}$$

$$\mathbb{d}_n = \left(m_n^{\texttt{pork}}, m_n^{\texttt{feed}}, m_n^{\texttt{piglet}}\right) = \left((\hat{\mu}_n^{\texttt{pork}}, \hat{\lambda}_n^{\texttt{pork}}), \hat{\lambda}_n^{\texttt{feed}}, \hat{\lambda}_n^{\texttt{piglet}}\right), \tag{3.14}$$

where $p^{\texttt{pork}}$ and $p^{\texttt{feed}}$ denote the observed pork and feed prices at the start of a production cycle and $d^{\texttt{piglet}}$ is the log transformed observed piglet price ratio. Similarly, $(\hat{\mu}_n^{\texttt{pork}}, \hat{\lambda}_n^{\texttt{pork}})$, $\hat{\lambda}_n^{\texttt{feed}}$ and $\hat{\lambda}_n^{\texttt{piglet}}$ denote the posterior mean values of the latent variables in the SSMs for pork, feed and piglet prices, respectively.

States $\mathbb{p}$ and $\mathbb{d}_n$ are used to calculate our expected rewards in Section 3.3.4. The piglet price used in (3.6) is $p^{\texttt{piglet}} = p^{\texttt{pork}} \exp(d^{\texttt{piglet}})$, the feed price used in (3.7) is $p^{\texttt{feed}}$, and the settlement pork price function (3.23) used in (3.8) is based on the market pork price $\hat{\mu}_n^{\texttt{pork}}$.

Moreover, we need states $\mathbb{p}$ and $\mathbb{d}_n$ for the calculation of transition probabilities in the HMDP. Before calculating the transition probabilities, a discretization approach should be specified for the continuous state variables in (3.13) and (3.14) since states in an HMDP must be discrete (Nielsen et al., 2011). Let $\mathbb{U}_{x_n} = \{\Pi_1, \ldots, \Pi_{|\mathbb{U}_{x_n}|}\}$ is a set of disjoint intervals that represent the partitioning of possible values for the continuous state variable x_n at stage n (e.g. $x_n = \hat{\mu}_n^{\texttt{pork}}$). Moreover, given interval Π, let *centre point* π denote the centre of the interval. That is, a possible value of state variable x_n can be represented by centre point π_{x_n} in interval Π_{x_n}. As a result, the state sets at Levels 0 and 1 (see Section 3.3.2) become

$$\mathbb{P} = \mathbb{U}_{p^{\texttt{pork}}} \times \mathbb{U}_{p^{\texttt{feed}}} \times \mathbb{U}_{d^{\texttt{piglet}}},$$
$$\mathbb{D}_n = \mathbb{U}_{\hat{\mu}_n^{\texttt{pork}}} \times \mathbb{U}_{\hat{\lambda}_n^{\texttt{pork}}} \times \mathbb{U}_{\hat{\lambda}_n^{\texttt{feed}}} \times \mathbb{U}_{\hat{\lambda}_n^{\texttt{piglet}}}.$$

Now the transition probabilities (3.2)-(3.5) in Section 3.3.3 can be calculated. First, consider child jump probability (3.2) with a transition to stage 1 at Level 1. This transition is deterministic

$$\Pr(\mathbb{d}_1 \mid \mathbb{p}) = \begin{cases} 1 & \mathbb{d}_1 = ((p^{\texttt{pork}}, 0), 0, d^{\texttt{piglet}}), \\ 0 & \text{otherwise}, \end{cases}$$

since $\hat{\mu}_1^{\texttt{pork}} = p^{\texttt{pork}}$ due to (3.10), $\hat{\lambda}_1^{\texttt{pork}}$ is assumed zero, $\hat{\lambda}_1^{\texttt{feed}} = 0$ due to (3.11) and $\hat{\lambda}_1^{\texttt{piglet}} = d^{\texttt{piglet}}$ due to (3.12).

Next, consider the transition probability $\Pr(\mathbb{d}_{n+1} \mid \mathbb{d}_n)$ for the actions $a_{\texttt{cont}}$ and a_q used in (3.3) and (3.4). Since $\mathbb{d}_n$ includes state variables related to the three independent SSMs of pork, feed, and piglet prices, this probability equals to

$$\begin{aligned}
\Pr(\mathbb{d}_{n+1} \mid \mathbb{d}_n) =& \Pr\left(m_{n+1}^{\texttt{pork}} \mid m_n^{\texttt{pork}}\right) \cdot \Pr\left(m_{n+1}^{\texttt{feed}} \mid m_n^{\texttt{feed}}\right) \cdot \Pr\left(m_{n+1}^{\texttt{piglet}} \mid m_n^{\texttt{piglet}}\right) \\
=& \Pr\left((\mu_{n+1}^{\texttt{pork}}, \lambda_{n+1}^{\texttt{pork}}) \in \Pi_{\mu_{n+1}^{\texttt{pork}}} \times \Pi_{\lambda_{n+1}^{\texttt{pork}}} \mid (\pi_{\mu_n^{\texttt{pork}}}, \pi_{\lambda_n^{\texttt{pork}}})\right) \\
& \cdot \Pr\left(\lambda_{n+1}^{\texttt{feed}} \in \Pi_{\lambda_{n+1}^{\texttt{feed}}} \mid \pi_{\lambda_n^{\texttt{feed}}}\right) \cdot \Pr\left(\lambda_{n+1}^{\texttt{piglet}} \in \Pi_{\lambda_{n+1}^{\texttt{piglet}}} \mid \pi_{\lambda_n^{\texttt{piglet}}}\right).
\end{aligned}$$

Notice that due to our discretization approach, the probabilities are calculated over intervals given previous centre points. Moreover, the probability distribution of $(m_{n+1}|m_n)$ can be obtained using the *k-step posterior mean distribution* defined in Theorem 6 in Appendix 3.C where k denotes the length of the current stage.

Finally, for parent jump probability (3.5) to state $\tilde{\mathbb{p}} = (p^{\texttt{pork}}, p^{\texttt{feed}}, d^{\texttt{piglet}})$ used under action $a_{\texttt{term}}$, the probability becomes

$$\begin{aligned}
\Pr(\tilde{\mathbb{p}} \mid \mathbb{d}_n) =& \Pr\left(p^{\texttt{pork}} \mid m_n^{\texttt{pork}}\right) \cdot \Pr\left(p^{\texttt{feed}} \mid m_n^{\texttt{feed}}\right) \cdot \Pr\left(d^{\texttt{piglet}} \mid m_n^{\texttt{piglet}}\right) \\
=& \Pr\left(p^{\texttt{pork}} \in \Pi_{p^{\texttt{pork}}} \mid (\pi_{\mu_n^{\texttt{pork}}}, \pi_{\lambda_n^{\texttt{pork}}})\right) \\
& \cdot \Pr\left(p^{\texttt{feed}} \in \Pi_{p^{\texttt{feed}}} \mid \pi_{\lambda_n^{\texttt{feed}}}\right) \cdot \Pr\left(d^{\texttt{piglet}} \in \Pi_{d^{\texttt{piglet}}} \mid \pi_{\lambda_n^{\texttt{piglet}}}\right).
\end{aligned}$$

Note that the conditional distribution of $(p|m_n)$ can be obtained using the *k-step forecast distribution* defined in Theorem 6 in Appendix 3.C where k is the expected length of action $a_{\texttt{term}}$ ($k = h + b$ days).

3.5 Optimal policy and value of information

In this section, we calculate the optimal policy of the HMDP to investigate the influence of price fluctuations on the marketing decisions. We consider three scenarios with different patterns of price fluctuations and comment on the optimal marketing decisions. Moreover, we compare the optimal marketing decisions in a model with and without price fluctuations and calculate the value of price information.

3.5.1 Model parameters

In order to initialize the model, we need the parameter values of the HMDP and the statistical models embedded into the HMDP. The parameter values are given in Table 3.1. They have been obtained using historical pork, piglet and feed market prices, information about finisher pig production (Danish conditions) and related literature (see the footnotes in Table 3.1).

More precisely, the parameter values of the HMDP were set based on discussions with Danish experts in pig production, standard Danish herd conditions and related literature. The system and observational variances of each SSM modeling the pork, piglet and feed market prices were estimated using maximum likelihood estimation (MLE) applied to historical prices in Denmark from 2006-2014. To calculate the expected revenue of each state and action in the HMDP, we need to specify the settlement pork price $p^{\text{pork}}(\tilde{w}, \breve{w})$ which is a piecewise linear function under current Danish conditions and is specified in Appendix 3.B. Moreover, to estimate parameters in the random regression model (RRM) for finding the weight distribution in the pen, we used the restricted maximum likelihood method (RMLE) applied to a set of weight data acquired from a standard Danish herd. Finally, in order to formulate the HMDP, we need to specify possible values of the discrete state variable and the range of centre points for the continuous state variables in the HMDP. Possible values for the discrete state variable q_n are 1 to $q^{\max}$ and, based on our discretization method in Section 3.4.2, possible values of the continuous state variables $p^{\texttt{pork}}$, $p^{\texttt{feed}}$, $d^{\texttt{piglet}}$, $\hat{\mu}_n^{\texttt{pork}}$, $\hat{\lambda}_n^{\texttt{pork}}$, $\hat{\lambda}_n^{\texttt{feed}}$, and $\hat{\lambda}_n^{\texttt{piglet}}$ are divided into intervals with given centre points. An overview over the values of each state variable is given in Table 3.2.

3.5.2 Optimal marketing decisions under different scenarios

To see the behavior of the optimal policy under different patterns of price fluctuations we consider three scenarios, illustrated in Figure 3.3, over a period of 15 weeks assuming that the production

Table 3.1: Parameter values.

Parameter	Value	Explanation
HMDP (Section 3.3)		
$q^{\max}$	15	Number of pigs inserted into the pen.[a]
$t^{\max}$	14	Maximum number of weeks in a production cycle.[a]
$t^{\min}$	9	First possible week of marketing decisions.[a]
h	4	Days used for cleaning the pen after termination.[a]
b	3	Days before delivery to abattoir after a marketing decision.[a]
SSMs (Section 3.4.1)		
$W^{\texttt{pork}}$	$\begin{pmatrix} 0 & 0 \\ 0 & 0.173^2 \end{pmatrix}$	System variance (pork price).[b]
$W^{\texttt{feed}}$	0.044^2	System variance (feed price).[b]
$W^{\texttt{piglet}}$	0.0108^2	System variance (piglet price).[b]
$V^{\texttt{feed}}$	0	Observation variance (feed price).[b]
$V^{\texttt{piglet}}$	0	Observation variance (piglet price).[b]
$m_0^{\texttt{pork}}$	$\begin{pmatrix} 9.85 & 0 \end{pmatrix}$	Prior mean (pork price).[b]
$C_0^{\texttt{pork}}$	$\begin{pmatrix} 0 & 0 \\ 0 & 0.139^2 \end{pmatrix}$	Prior variance (pork price).[b]
$m_0^{\texttt{feed}}$	0	Prior mean (feed price).[b]
$C_0^{\texttt{feed}}$	0.336^2	Prior variance (feed price).[b]
$m_0^{\texttt{piglet}}$	3.55	Prior mean (piglet price).[b]
$C_0^{\texttt{piglet}}$	0.057^2	Prior variance (piglet price).[b]
Calculation of expected reward (Appendix 3.B)		
β	$\begin{pmatrix} 21.767 & 4.914 & 0.149 \end{pmatrix}'$	Fixed parameters (RRM).[c]
V	$\begin{pmatrix} 2.072 & 0.828 & 0.01 \\ 0.828 & 1.753 & -0.142 \\ 0.01 & -0.142 & 0.015 \end{pmatrix}$	Covariance matrix for α_j (RRM).[c]
R	2.04	Standard deviation of residual error (RRM).[c]
$\bar{g}$	6	Average weekly gain (kg) in the herd.[d]
$\bar{\bar{w}}$	61	Average leanness percentage in the herd.[d]
σ_c^2	1.4	Standard deviation of conversion rate c_s.[d]
k_2	0.044	Energy requirements (FEsv) per kg metabolic weight.[d]
k_1	1.549	Energy requirement (FEsv) per kg gain.[d]

[a] Value based on discussions with experts in Danish pig production. [b] Estimated based on time series of pig, feed and piglet prices that can be found on `http://www.notering.dk/WebFrontend/`. [c] Estimated using the weight data in a standard Danish herd. [d] Value taken from Kristensen et al. (2012).

Table 3.2: Cardinality of the discrete state variable and range of the center points for the continuous state variables.

Process level	0			1^a				
State variable	$p^{\mathtt{pork}}$	$p^{\mathtt{feed}}$	$d^{\mathtt{piglet}}$	q_n	$\hat{\mu}_n^{\mathtt{pork}}$	$\hat{\lambda}_n^{\mathtt{pork}}$	$\hat{\lambda}_n^{\mathtt{feed}}$	$\hat{\lambda}_n^{\mathtt{piglet}}$
Intervals/cardinality	16	15	5	15	16	5	5	5
Range of centre points	9.2-12.2	1.5-2.2	3.4-3.6	1-15	9.2-12.2	-0.4-0.4	-0.1-0.1	3.5-3.7

[a] At stage $n = 1$ the only possible values are $q_n = 15$ and $\hat{\lambda}_n^{\mathtt{pork}} = \hat{\lambda}_n^{\mathtt{feed}} = 0$.

cycle starts at week one and ends at the start of week 15 at the latest:

Scenario 1: Favorable trend of pork price and unfavorable trends of feed and piglet prices. Pork price increases from 10.3 to 11.3 DKK, feed price increases from 1.79 to 1.92 DKK and piglet price increases from 336 to 396 DKK. This scenario is based on the historical data from weeks 11-25 in 2012.

Scenario 2: Favorable trends of pork and feed prices and unfavorable trend of piglet price. Pork price increases from 10.3 to 11.3 DKK, feed price decreases from 1.79 to 1.66 DKK and piglet price increases from 336 to 396 DKK.

Scenario 3: Unfavorable trends of pork and feed prices and favorable trend of piglet price. Pork price decreases from 10.3 to 9.3 DKK, feed price increases from 1.79 to 1.92 DKK and piglet price decreases from 362 to 328 DKK.

During the 15 weeks period, the average weight in the pen increases from 26.8 to 128.9 kg with a standard deviation increasing from 3 to 15.4 kg (see Equation (3.18)). Notice that the growth of the pigs is the same in the three scenarios and hence the only factor affecting the marketing policy is the price information.

To find the optimal policy of the HMDP, the model was coded using the C++ programming language (gcc compiler) and R (R Core Team, 2015), and the optimal policy of the HMDP was calculated using the modified policy iteration algorithm[3] using the R package "MDP" (Nielsen, 2009). The source code is available on-line (Pourmoayed and Nielsen, 2015). Given the parameters in Table 3.1 and the discretization of state variables in Table 3.2, the number of states and actions at the founder level are both 1,200. Moreover, each child process at Level 1

[3]Using shared and external processes, i.e. the memory used for child processes may be shared and loaded when needed. For further information see the documentation in Nielsen (2009).

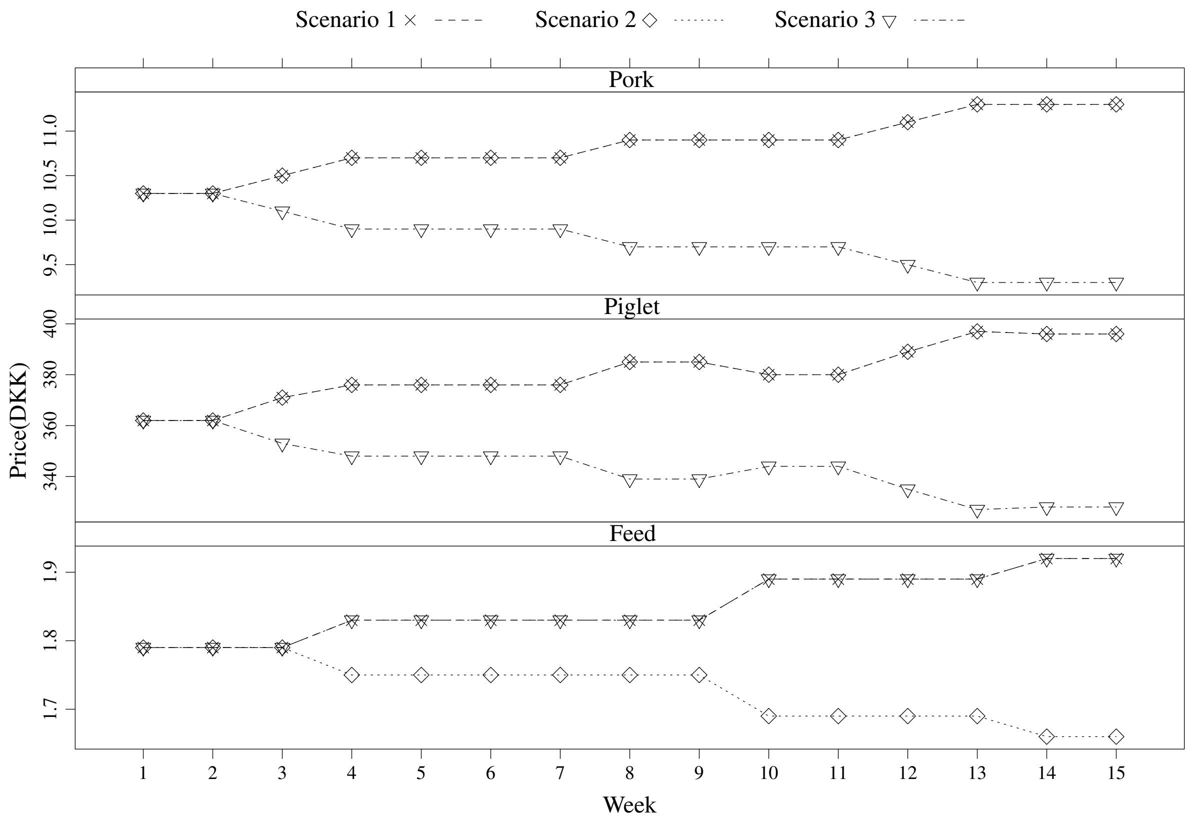


Figure 3.3: Price fluctuations in the three scenarios. In Scenario 1, the trends of feed and piglet prices are unfavorable and the trend of pork price is favorable. In Scenario 2, the trends of pork and feed prices are favorable and the trend of piglet price is unfavorable. In Scenario 3, the trends of pork and feed prices are unfavorable and the trend of piglet price is favorable.

contains 194,080 states and 1,412,080 actions. That is, the total numbers of states and actions of the model are 232,897,200 and 1,694,497,200, respectively.

For each scenario we use the SSMs to find the values of the state variables related to the price information in the HMDP. That is, for each scenario we identify the relevant state and the corresponding optimal action. The results for each scenario are illustrated in Figure 3.4 which include estimations of posterior mean parameters in the SSMs and the number of remaining pigs in the pen in each week (bars). The optimal decision a^* is shown just above the x-axis where the numbers denote the number of the heaviest pigs culled from the pen (a_q), the letter "T" indicates the termination decision ($a_{\texttt{term}}$), and the letter "C" corresponds to continuing the production process without marketing decisions ($a_{\texttt{cont}}$).

In Scenarios 1 and 2, fluctuations of pork and piglet prices are the same while fluctuations of

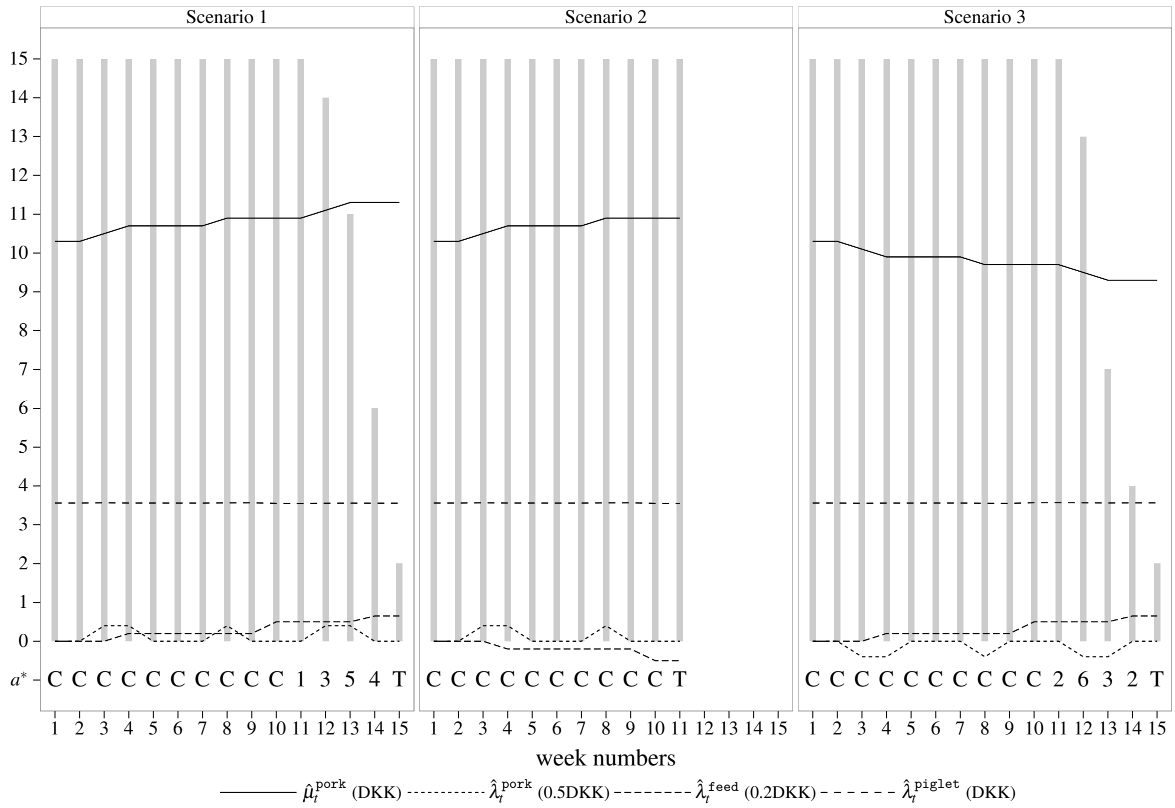


Figure 3.4: Estimated means of posterior parameters in the SSMs and the optimal decisions of the HMDP for the three scenarios. $\hat{\lambda}_t^{\texttt{pork}}, \hat{\lambda}_t^{\texttt{feed}}$, and $\hat{\lambda}_t^{\texttt{piglet}}$ are the mean estimates of pork, feed and piglet price deviations, respectively, and $\hat{\mu}_t^{\texttt{pork}}$ is the estimated mean of pork price. The optimal decision is shown just above the x-axis where the numbers denote the number of the heaviest pigs culled from the pen (a_q), the letter "T" indicates the termination decision ($a_{\texttt{term}}$), and the letter "C" corresponds to continuing the production process without marketing decisions ($a_{\texttt{cont}}$). The bars show the number of remaining pigs in the pen before making a decision. In the plot, the values of $\hat{\lambda}_t^{\texttt{pork}}$ and $\hat{\lambda}_t^{\texttt{feed}}$ have been scaled with factors 2 and 5, respectively.

feed price are different. By comparing the two scenarios, we observe that the different trends of feed price have a significant impact on the optimal policy. In Scenario 2, a decreasing feed price leads to an earlier termination (at week 11) compared to Scenario 1 with an increasing feed price (termination at week 15). Note that due to Assumption 9 on page 66, when the pen is terminated, a low feed price affects the feeding cost of the next production cycle and hence when the feed price is low, it may be beneficial to terminate the pen earlier and start a new production cycle. On the other hand, an increasing feed price in Scenario 1 during the marketing period (an increase from 1.83 to 1.92 in weeks 9-15) results in a longer production cycle and individual marketings in weeks 11 to 14.

In Scenario 3, we have an increasing trend in feed price (similar to Scenario 1) but unlike Scenarios 1 and 2, the trends of pork and piglet prices are decreasing in this scenario (see Figure 3.3). Here a decreasing piglet price does not result in an earlier termination as we had in Scenario 2. Like in Scenario 1, the termination occurs at week $t^{\max} = 15$ in this scenario too, which is due to the increasing trend of feed price. That is, the feed price compared to the piglet price has a higher impact on the optimal policy and reward. This observation was also supported in Pourmoayed et al. (2016). We also see that in Scenario 3 the fraction of remaining pigs in the pen in every week of the marketing period is lower than Scenario 1. This is because of the increasing trend of pork price in Scenario 1 that makes it more beneficial to keep more pigs in the pen and sell them in the next weeks while in Scenario 3 it is better to sell the pigs earlier since the pork price decreases in the next weeks.

3.5.3 Value of price information

Is it relevant to embed a statistical model with fluctuating prices into the HMDP? To answer this question, we compare the optimal policy of the HMDP against the policy considering marketing decisions under fixed prices, i.e. the decision maker has no information about the market prices and the price deviations and therefore follows a policy taking the actions specified by the state corresponding to the fixed prices. The extra reward per time unit gained by using fluctuating prices compared to fixed prices may then be considered as the value of information about prices.

To be more specific, consider a fixed price setting, i.e. a specific state at the founder level:

$$\tilde{\mathbb{p}} = \left(\tilde{p}^{\texttt{pork}}, \tilde{p}^{\texttt{feed}}, \tilde{d}^{\texttt{piglet}}\right)$$

According to the structure of the HMDP in Section 3.3.2, having no information about the price deviations in the child process $\tilde{\mathfrak{p}}^1 = (\mathfrak{p}^0 || \tilde{\mathbb{p}})$, implies to use state variable $\tilde{\mathbb{d}}_n$ in $\tilde{\mathfrak{p}}^1$ equal to

$$\tilde{\mathbb{d}}_n = \left((\hat{\mu}_n^{\texttt{pork}}, \hat{\lambda}_n^{\texttt{pork}}), \hat{\lambda}_n^{\texttt{feed}}, \hat{\lambda}_n^{\texttt{piglet}}\right) = \left((\tilde{p}^{\texttt{pork}}, 0), 0, m_0^{\texttt{piglet}}\right),$$

where $m_0^{\texttt{piglet}}$ is the predefined prior mean of $\hat{\lambda}_n^{\texttt{piglet}}$ defined in Table 3.1. Moreover, define action $a_n^*(\tilde{\mathbb{p}}, \tilde{\mathbb{d}}_n, q_n)$ as the optimal action to state $(\tilde{\mathbb{d}}_n, q_n) = \left((\tilde{p}_n^{\texttt{pork}}, 0), 0, m_0^{\texttt{piglet}}, q_n\right)$ at stage n of process $\tilde{\mathfrak{p}}^1$. Now, the *no information policy* is defined such that for each child process $\mathfrak{p}^1 = (\mathfrak{p}^0 || \mathbb{p})$ and state $i = (\mathbb{d}_n, q_n)$ at stage n, we use action

$$a(i) = a_n(\mathbb{d}_n, q_n) = a_n^*(\tilde{\mathbb{p}}, \tilde{\mathbb{d}}_n, q_n).$$

Table 3.3: VOI values for comparing the optimal policy of the HMDP (under price fluctuations) with the policies considering marketing decisions under fixed prices in 5 different price settings. In the first group of price settings, price values are related to the average prices of pork, feed, and piglet in the period of 2011 to the end of 2014 in Denmark. Other groups are defined by changing the price values of the first group to make a small sensitivity analysis on VOI.

Price setting (PS)	Pork price	Feed price	Piglet price	VOI (DKK)
1	10.8	1.85	376	41.07
2	10.8	1.55	376	41.29
3	10.8	2.15	376	43.91
4	9.4	1.85	327	41.26
5	12	1.85	418	40.25
6	12	2.15	418	43.75
7	9.4	1.55	327	55.21

That is, the actions used under the *no information policy* are the optimal actions related to the fixed price setting $\tilde{\mathbb{p}}$.

Under the expected reward per time unit criterion and a given price setting $\tilde{\mathbb{p}}$, the *value of information* (*VOI*) is defined as the difference between the expected reward per time unit under the optimal policy and under the no information policy. That is, VOI is the extra reward per time unit gained by embedding the SSMs into the HMDP for predicting the future market prices.

To evaluate the benefit of price information under fluctuating prices, we consider five different price settings (*PS*) and calculate VOI for each setting. The results are given in Table 3.3. Under PS 1, the price values equal the average prices of pork, feed and piglet in the period of 2011 to the end of 2014 in Denmark. VOI in this group represents the average expected loss per week by assuming average prices instead of using a model with fluctuating prices. The other price settings are defined with the purpose of making a small sensitivity analysis on VOI for different price settings. Under PS 2 and PS 3, feed prices are low (1.55 DKK/FEsv) and high (2.15 DKK/FEsv), respectively, and pork and piglet prices equal PS 1. Similarly, in PS 4 and PS 5, pork and piglet prices are low and high and the feed prices are similar to PS 1. In PS 6 and PS 7, all prices are different from PS 1 and are set to high and low values, respectively.

In PS 2 to PS 6, the values of VOI are not considerably different than the value of VOI in PS 1. However, it seems that assuming a high fixed feed price (PS 3 and PS 6) has a higher impact on VOI. Moreover, the effect of assuming a high pork price (PS 5) is lower compared to the average prices (PS 1). Finally, in PS 7 where all prices are assumed low, VOI is higher. This shows that assuming combined low pork, feed and piglet prices has a higher impact on VOI compared to, for instance, a low feed price only.

3.6 Conclusions

In the production of fattening pigs, price fluctuations in the market have an effect on marketing decisions. In this paper we used a two-level HMDP to model marketing decisions under fluctuating pork, piglet and feed prices.

We used a Bayesian approach to update the state of the system such that it contains updated information based on previous market prices. That is, three SSMs were formulated to forecast future prices and each SSM was embedded into the HMDP such that the model takes into account new market prices using a general discretization method.

Numerical examples show that price fluctuations have an impact on marketing decisions; the effect of a fluctuating feed price was especially noticeable. Moreover, we analyzed the value of including information about fluctuating prices into the HMDP compared to using fixed prices. The results showed that the long-term average reward per time unit of the production unit can be improved by including price fluctuations into the model.

Acknowledgment

The authors would like to thank Professors Anders Ringgaard Kristensen and Christian Larsen for useful comments on an earlier draft of this paper. This article has been written with support from The Danish Council for Strategic Research (The PigIT project, Grant number 11-116191).

References

S. Andersen, B. Pedersen, and M. Ogannisian. Slagtesvindets sammensætning. meddelelse 429. Technical report, Landsudvalget for Svin og Danske Slagterier, 1999. URL `http://vsp.lf.dk/Publikationer/Kilder/lu_medd/medd/429.aspx`.

C. Bono, C. Cornou, and A.R. Kristensen. Dynamic production monitoring in pig herds i: Modeling and monitoring litter size at herd and sow level. *Livestock Science*, 149(3):289–300, 2012. doi:10.1016/j.livsci.2012.07.023 .

C. Bono, C. Cornou, S. Lundbye-Christensen, and A.R. Kristensen. Dynamic production monitoring in pig herds ii. modeling and monitoring farrowing rate at herd level. *Livestock Science*, 155(1):92–102, 2013. doi:10.1016/j.livsci.2013.03.026 .

K.A. Boys, N. Li, P.V. Preckel, A.P. Schinckel, and K.A. Foster. Economic replacement of a heterogeneous herd. *American Journal of Agricultural Economics*, 89(1):24–35, 2007. doi:10.1111/j.1467-8276.2007.00960.x .

J.E. Broekmans. Influence of price fluctuations on delivery strategies for slaughter pigs. Technical Report Dina Notat 7, Research Centre Foulum, 1992.

W. Cai, H. Wu, and J. Dekkers. Longitudinal analysis of body weight and feed intake in selection lines for residual feed intake in pigs. *Asian Australasian Journal of Animal Science*, 24(1):17, 2011. URL `http://www.ajas.info/editor/manuscript/upload/24-3.pdf`.

J.P. Chavas, J. Kliebenstein, and T.D. Crenshaw. Modeling dynamic agricultural production response: The case of swine production. *American Journal of Agricultural Economics*, 67(3): 636–646, 1985. doi:10.2307/1241087 .

C. Cornou, J. Vinther, and A.R. Kristensen. Automatic detection of oestrus and health disorders using data from electronic sow feeders. *Livestock Science*, 118(3):262–271, 2008.

J. Durbin and S.J. Koopman. *Time series analysis by state space methods*. Oxford University Press, 2 edition, 2012. ISBN 978–0–19–964117–8.

E. Jørgensen. Price fluctuations described by first order autoregressive process. Technical report, Department of Research in Pigs and Horses, National Institute of Animal Science., 1992.

E. Jørgensen. The influence of weighing precision on delivery decisions in slaughter pig production. *Acta Agriculturae Scandinavica, Section A - Animal Science*, 43(3):181–189, August 1993. doi:10.1080/09064709309410163 .

E Jørgensen. Foderforbrug pr kg tilvækst hos slagtesvin. fordeling mellem forbrug til vedligehold og til produktion i besætninger under den rullende afprøvning. Technical report, Danish Institute of Agricultural Sciences, Biometry Research Unit, 2003.

S. Khamjan, K. Piewthongngam, and S. Pathumnakul. Pig procurement plan considering pig growth and size distribution. *Computers & Industrial Engineering*, 64:886–894, 2013. doi:10.1016/j.cie.2012.12.022 .

A.R. Kristensen. Hierarchic markov processes and their applications in replacement models. *European Journal of Operational Research*, 35(2):207–215, 1988. doi:10.1016/0377-2217(88)90031-8 .

A.R. Kristensen and E. Jørgensen. Multi-level hierarchic markov processes as a framework for herd management support. *Annals of Operations Research*, 94(1-4):69–89, 2000. doi:10.1023/A:1018921201113 .

A.R. Kristensen, L. Nielsen, and M.S. Nielsen. Optimal slaughter pig marketing with emphasis on information from on-line live weight assessment. *Livestock Science*, 145(1-3):95–108, May 2012. doi:10.1016/j.livsci.2012.01.003 .

H. Kure. *Marketing Management Support in Slaughter Pig Production*. PhD thesis, The Royal Veterinary and Agricultural University, 1997. URL `http://www.prodstyr.ihh.kvl.dk/pub/phd/kure_thesis.pdf`.

L.R Nielsen. Mdp: Markov decision processes in R. R package v1.1., 2009. URL `http://r-forge.r-project.org/projects/mdp/`.

L.R. Nielsen and A.R. Kristensen. Finding the K best policies in a finite-horizon Markov decision process. *European Journal of Operational Research*, 175(2):1164–1179, 2006. doi:10.1016/j.ejor.2005.06.011 .

L.R. Nielsen and A.R. Kristensen. Markov decision processes to model livestock systems. In Lluis M. Plà-Aragonés, editor, *Handbook of Operations Research in Agriculture and the Agri-Food Industry*, volume 224 of *International Series in Operations Research & Management Science*, pages 419–454. Springer, 2014. doi:10.1007/978-1-4939-2483-7_19 .

L.R Nielsen, E. Jørgensen, A.R. Kristensen, and S. Østergaard. Optimal replacement policies for dairy cows based on daily yield measurements. *Journal of Dairy Science*, 93(1):77–92, 2010. doi:10.3168/jds.2009-2209 .

L.R. Nielsen, E. Jørgensen, and S. Højsgaard. Embedding a state space model into a markov decision process. *Annals of Operations Research*, 190(1):289–309, 2011. doi:10.1007/s10479-010-0688-z .

J.K. Niemi. *A dynamic programming model for optimising feeding and slaughter decisions regarding fattening pigs | NIEMI | Agricultural and Food Science*. PhD thesis, MTT Agrifood research, 2006. URL http://ojs.tsv.fi/index.php/AFS/article/view/5855.

J.W. Ohlmann and P.C. Jones. An integer programming model for optimal pork marketing. *Annals of Operations Research*, 190(1):271–287, November 2008. doi:10.1007/s10479-008-0466-3 .

H Desmond Patterson and Robin Thompson. Recovery of inter-block information when block sizes are unequal. *Biometrika*, 58(3):545–554, 1971.

J. Pitmand. *Probability (Springer Texts in Statistics)*. Springer, 1993. ISBN 978-0-387-94594-1.

Lluís M Plà-Aragonés, S Rodriguez-Sanchez, and Victoria Rebillas-Loredo. A mixed integer linear programming model for optimal delivery of fattened pigs to the abattoir. *J Appl Oper Res*, 5:164–175, 2013.

R. Pourmoayed and L.R. Nielsen. An overview over pig production of fattening pigs with a focus on possible decisions in the production chain. Technical Report PigIT Report No. 4, Aarhus University, 2014. URL http://pigit.ku.dk/publications/PigIT-Report4.pdf.

R. Pourmoayed and L.R. Nielsen. Github repository: Slaughter pig marketing under price fluctuations, 2015. URL https://github.com/pourmoayed/hmdpPricePigIT.git.

R. Pourmoayed, L. R. Nielsen, and A. R. Kristensen. A hierarchical Markov decision process modeling feeding and marketing decisions of growing pigs. *European Journal of Operational Research*, 250(3):925–938, 2016. doi:10.1016/j.ejor.2015.09.038 .

R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2015. URL http://www.R-project.org/.

S.V. Rodriguez, T.B. Jensen, L.M. Pla, and A.R. Kristensen. Optimal replacement policies and economic value of clinical observations in sow herds. *Livestock Science*, 138(1-3):207–219, June 2011. doi:10.1016/j.livsci.2010.12.026 .

J. Roemen and J. de Klein. An optimal marketing strategy for porkers with differences in growth rates and dependent prices. In *Proceedings of the International Symposium on Pig Herd Management Modeling and Information Technologies Related*, pages 107–116, IRTA & Universitat de Lleida, Lleida, 2000.

L.R. Schaeffer. Application of random regression models in animal breeding. *Livestock Production Science*, 86(1):35–45, 2004. doi:10.1016/S0301-6226(03)00151-9 .

D. Sirisatien, G.R. Wood, M. Dong, and P.C.H. Morel. Two aspects of optimal diet determination for pig production: efficiency of solution and incorporation of cost variation. *Journal of Global Optimization*, 43(2-3):249–261, 2009. doi:10.1007/s10898-007-9262-x .

H.C. Tijms. *A first course in stochastic models*. John Wiley & Sons Ltd, 2003. ISBN 978-0-471-49880-3.

N. Toft, A.R. Kristensen, and E. Jørgensen. A framework for decision support related to infectious diseases in slaughter pig fattening units. *Agricultural Systems*, 85(2):120–137, 2005. doi: 10.1016/j.agsy.2004.07.017 .

M. West and J. Harrison. *Bayesian Forecasting and Dynamic Models (Springer Series in Statistics)*. Springer-Verlag, February 1997. ISBN 0387947256.

3.A Notation

Since the paper uses techniques from both statistical forecasting and operations research, we have to make some choices with respect to notation. In general, we use capital letters for matrices and let A' denote the transpose of A. Capital blackboard bold letters are used for sets (e.g. $\mathbb{P}$ and $\mathbb{D}_n$). Finally, accent $\hat{x}$ (hat) is used to denote an estimate of x. A description of the notation introduced in Section 3.3 and Section 3.4 is given in Tables 3.4 and 3.5, respectively.

Table 3.4: Notation - HMDP model (Section 3.3).

Symbol	Description
$\mathbb{I}_n$	Set of states at stage n.
$\mathbb{A}_n(i)$	Set of actions given stage n and state i.
$r_n(i,a)$	Reward at stage n given state i and action a.
$u_n(i,a)$	Expected length until the next decision epoch at stage n given state i and action a.
$\Pr(j \mid n,i,a)$	Transition probability from state i at stage n to state j at the next stage under action a.
$\mathfrak{p}^l$	A process at level l (superscript is used to indicate level).
$\mathscr{N}l$	Time horizon of process $\mathfrak{p}^l$ at level l.
n^l, i^l, a^l	A stage, state, and action in process $\mathfrak{p}^l$.
$q^{\max}$	Number of pigs inserted into the pen.
$t^{\max}$	Latest possible week of pen termination.
$t^{\min}$	First possible week of marketing decisions.
h	Number of days for cleaning the pen after termination.
b	Number of days of preparation for delivery to the abattoir.
q_n	Remaining pigs in the pen at stage n, $1 \leq q_n \leq q^{\max}$.
$\mathbb{p}$	Model information related to the price information in the first level of HMDP, $\mathbb{p} \in \mathbb{P}$.
$\mathbb{d}_n$	Model information related to the price deviations in the second level of HMDP, $\mathbb{d}_n \in \mathbb{D}_n$.
$a_{\texttt{term}}$	Action related to pen termination.
$a_{\texttt{cont}}$	Action related to continuing the production process without marketing.
a_q	Action related to marketing the q heaviest pigs in the pen ($1 \leq q < q_n$).
$p^{\texttt{feed}}$	Market feed price at the beginning of a production cycle (DKK).
$p^{\texttt{piglet}}$	Market piglet price at the beginning of a production cycle (DKK).
$w_{(k)}$	Weight of the kth pig in the pen (kg).
$f^{\texttt{feed}}_{(k),n}(t)$	Expected feed intake of the kth pig from the start of stage n and the next t days ahead (FEsv).
$\tilde{w}_{(k)}$	Carcass weight of the kth pig at delivery to the abattoir (kg).
$\breve{w}_{(k)}$	Leanness (non-fat percentage) of the kth pig at delivery to the abattoir.
$p^{\texttt{pork}}_{(k),n}(\cdot)$	Settlement pork price of the kth pig of one kg of meat at delivery to the abattoir.

Table 3.5: Notation - Bayesian updating of prices (Section 3.4).

Symbol	Description
θ_t	Latent/unobservable variable(s).
y_t	Observable variable(s).
G_t	Design matrix of system equation.
F_t	Design matrix of observation equation.
ω_t	System noise, $\omega_t \sim N(0, W_t)$ where W_t denotes the system covariance matrix.
ν_t	Observation error, $\nu_t \sim N(0, V_t)$ where V_t denotes the observation covariance matrix.
$\mathbb{D}_t$	Set of information available up to time t in the system.
(m_0, C_0)	Mean and covariance matrix of the prior, $\theta_0 \sim N(m_0, C_0)$.
(m_t, C_t)	Mean and covariance matrix of the posterior at time t, $(\theta_t \vert \mathbb{D}_t) \sim N(m_t, C_t)$.
$p_t^{\texttt{pork}}$	Observed market pork price at time t (DKK).
$\mu_t^{\texttt{pork}}$	A supplementary latent variable in the SSM of pork price ($\mu_t^{\texttt{pork}} = p_t^{\texttt{pork}}$).
$\lambda_t^{\texttt{pork}}$	Price deviation related to pork price at time t.
$p_t^{\texttt{feed}}$	Observed market feed price at time t (DKK).
$\lambda_t^{\texttt{feed}}$	Price deviation related to feed price at time t.
$p_t^{\texttt{piglet}}$	Observed market piglet price at time t (DKK).
$d_t^{\texttt{piglet}}$	Log transformed observed piglet ratio.
$\lambda_t^{\texttt{piglet}}$	Price deviation related to piglet price at time t.
$\mathbb{U}_{x_n}$	Set of disjoint intervals representing the partitioning of state variable x_n at stage n, $\mathbb{U}_{x_n} = \{\Pi_1, .., \Pi_k, .., \Pi_{\vert\mathbb{U}_{x_n}\vert}\}$ where Π_k denotes interval k.
π_k	Centre point of interval Π_k.

3.B Calculating expected reward

Modeling weights in the pen

During the growing period in the pen, pigs grow with different growth rates; that is, given a certain week in the production cycle, there is a variation between the weights of the individual animals in the pen. Moreover, as the pigs grow, this variation increases and our uncertainty about the average weight of the pen increases.

Let $(w_{(1)},..w_{(k)},..,w_{(q)})_t$ denote the weight distribution of the q pigs in the pen at week t such that $w_{(1)}$, $w_{(k)}$, and $w_{(q)}$ are ordered random variables (order statistics) related to the weight of the lightest, kth and the heaviest pig in the pen at week t, respectively. To find the probability distribution of the ordered random variable $w_{(k)}$, first the weight distribution of a randomly selected pig should be determined in the pen. To specify this distribution, a *random regression model* (*RRM*) is used that is usually applied in the animal breeding models (Schaeffer, 2004).

Let $w_{j,t}$ denotes the weight of pig j at week t, randomly selected in the pen. $w_{j,t}$ can be described using an RRM:

$$w_{j,t} = X_t\beta + Z_t\alpha_j + \varepsilon_{j,t}, \tag{3.15}$$

where X_t and Z_t are time covariate vectors, β is the vector of fixed parameters, α_j is the vector of random parameters and $\varepsilon_{j,t}$ is a residual error. In this RRM, $X_t\beta$ is the fixed effect of the model representing the average weight of the pen and $Z_t\alpha_j$ is the random effect showing a deviation between the weight of pig j and the average weight of the pen. A quadratic RRM is used suggested by Cai et al. (2011) where $X_t = Z_t = \begin{pmatrix}1 & t & t^2\end{pmatrix}$, $\alpha_j = \begin{pmatrix}\alpha_{0j} & \alpha_{1j} & \alpha_{2j}\end{pmatrix}'$ and $\beta = \begin{pmatrix}\beta_0 & \beta_1 & \beta_2\end{pmatrix}'$:

$$w_{j,t} = \beta_0 + \beta_1 t + \beta_2 t^2 + \alpha_{0j} + \alpha_{1j}t + \alpha_{2j}t^2 + \varepsilon_{j,t}. \tag{3.16}$$

Random parameter α_j follows a normal distribution with parameters

$$\alpha_j = \begin{pmatrix}\alpha_{0j}\\ \alpha_{1j}\\ \alpha_{2j}\end{pmatrix} \sim N(0, V = \begin{pmatrix}\sigma_0^2 & \sigma_{01} & \sigma_{02}\\ \sigma_{01} & \sigma_1^2 & \sigma_{12}\\ \sigma_{02} & \sigma_{12} & \sigma_2^2\end{pmatrix}), \tag{3.17}$$

where V is independent of pig j and time t. Moreover, the residual errors $\varepsilon_{j,t} \sim N(0,R)$ are independent random variables. Since (3.15) is linear with respect to random parameters α_j and

$\varepsilon_{j,t}$, we can conclude

$$w_{j,t} \sim N(\mu_t = X_t\beta, \sigma_t^2 = Z_t V Z_t' + R). \tag{3.18}$$

The parameters β, V and R can be estimated using the restricted maximum likelihood (REML) method (Patterson and Thompson, 1971).

Since the probability distribution of $w_{j,t}$ is independent of pig j, the weight distribution of all q pigs in the pen are i.i.d at time t. Hence, the probability density function of the ordered random variable $w_{(k)}$ becomes (Pitmand, 1993, page 326)

$$\phi_{(k)}(w) = \frac{q!}{(k-1)!(q-k)!}\Phi^{k-1}(w)[1-\Phi(w)]^{q-k}\phi(w),$$

where $\Phi(w)$ and $\phi(w)$ are the cumulative and density functions of the normal distribution defined in (3.18).

Carcass weight, leanness, feed intake and growth

Consider the kth ordered pig at stage n with weight w and daily growth g. The carcass weight $\tilde{w}$ can be approximated as (Andersen et al., 1999)

$$\tilde{w} = c_s w - 5.89 + e_c, \tag{3.19}$$

where $e_c \sim N(0, \sigma_c^2)$ is a normal distributed term. The relation between growth rate, leanness (lean meat percentage) and feed conversion ratio varies widely between herds. Hence, these formulas must be herd specific. The leanness $\breve{w}$ can be found as (Kristensen et al., 2012)

$$\breve{w} = \frac{-30(g-\bar{g})}{4} + \bar{\breve{w}}, \tag{3.20}$$

where, $\bar{g}$ is the average daily growth in the herd, $\bar{\breve{w}}$ is the average herd leanness percentage.

The feed intake (energy intake) is modelled as the sum of feed for maintenance and feed for growth. The basic relation between daily feed intake f (FEsv), live weight and daily gain is (Jørgensen, 2003)

$$f = k_1 g + k_2 w^{0.75}, \tag{3.21}$$

where k_1 and k_2 are constants describing the use of feed per kg gain and per kg metabolic weight, respectively. As a result the expected feed intake of a pig over the next $\hat{t}$ days equals

$$f^{\texttt{feed}}_{(k),n}(\hat{t}) = \mathbb{E}\left(\sum_{t=1}^{\hat{t}} f_t\right) = \mathbb{E}\left(\sum_{t=1}^{\hat{t}} \left(k_1 g + k_2(w+(t-1)g)^{0.75}\right)\right) = \mathbb{E}\left(\hat{t}k_1 g + k_2 \sum_{t=1}^{\hat{t}} (w+(t-1)g)^{0.75}\right), \quad (3.22)$$

where f_t denote the feed intake at day t calculated recursively using (3.21).

Settlement pork price

Consider the kth ordered pig at stage n with carcass weight $\tilde{w}$ and leanness $\breve{w}$ at delivery. The settlement pork price, under Danish conditions, is the sum of two linear piecewise functions related to the price of the carcass and a bonus of the leanness:

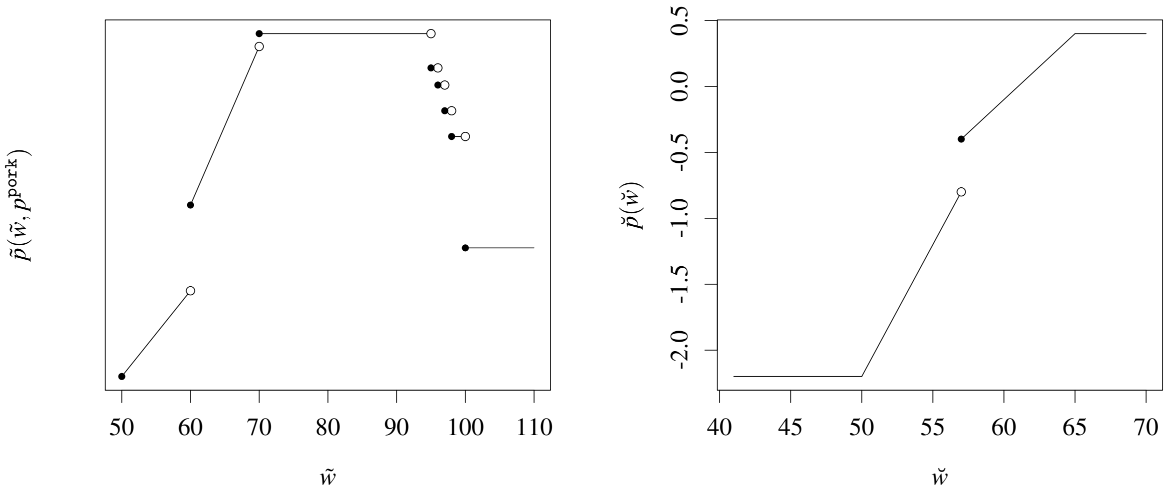
$$p^{\texttt{pork}}_{(k),n}(\tilde{w},\breve{w}) = \tilde{p}(\tilde{w}, p^{\texttt{pork}}) + \breve{p}(\breve{w}), \quad (3.23)$$

where $p^{\texttt{pork}}$ is the current pork price at the market. Functions $\tilde{p}(\tilde{w}, p^{\texttt{pork}})$ and $\breve{p}(\breve{w})$ correspond to the unit price of carcass and the bonus of leanness for 1 kg meat, respectively. A plot of each function is given in Figure 3.5.

Given the price structure, based on the Danish slaughter pig market[4], the unit price of 1 kg carcass is

$$\tilde{p}(\tilde{w}, p^{\texttt{pork}}) = \begin{cases} 0 & \tilde{w} < 50 \\ \frac{1}{9.9}(\tilde{w}-50) + p^{\texttt{pork}} - 4 & 50 \leq \tilde{w} < 60 \\ \frac{1.85}{9.9}(\tilde{w}-60) + p^{\texttt{pork}} - 2 & 60 \leq \tilde{w} < 70 \\ p^{\texttt{pork}} & 70 \leq \tilde{w} < 95 \\ p^{\texttt{pork}} - 0.2 & 95 \leq \tilde{w} < 96 \\ p^{\texttt{pork}} - 0.6 & 96 \leq \tilde{w} < 97 \\ p^{\texttt{pork}} - 0.9 & 97 \leq \tilde{w} < 98 \\ p^{\texttt{pork}} - 1.2 & 98 \leq \tilde{w} < 100 \\ p^{\texttt{pork}} - 2.5 & \tilde{w} \geq 100. \end{cases}$$

[4] `http://www.danishcrown.dk/Ejer/Noteringer/Aktuel-svinenotering.aspx` (October 2015)



(a) Carcass weight. (b) Leanness.

Figure 3.5: Price functions (DKK/kg) given carcass weight and leanness.

That is, the market pork price $p^{\mathtt{pork}}$ may be interpreted as the maximum price of 1 kg carcass that can be obtained (when the carcass weight lies between 70 and 95 kg).

The bonus of leanness is calculated as

$$\breve{p}(\breve{w}) = \begin{cases} -2.2 & \breve{w} < 50 \\ 0.2(\breve{w} - 61) & 50 \leq \breve{w} < 57 \\ 0.1(\breve{w} - 61) & 57 \leq \breve{w} < 65 \\ 0.4 & \breve{w} \geq 65. \end{cases}$$

Calculation of expected values

The calculations of the expected values (3.6)-(3.9) is rather complex due to the ordered random variables and the non-continuous functions $\tilde{p}(\tilde{w}, p^{\mathtt{pork}})$ and $\breve{p}(\breve{w})$. However, the expectations can be calculated using simulation with a simple sorting procedure as described below.

Step 0 For each pig $j = 1, \dots q^{\max}$, use (3.17) and draw a sample of random vector $\alpha_j \sim N(0, V)$.

Step 1 For each week t and pig j, according to the RRM model in (3.15) and (3.16), draw a sample of random residual $\varepsilon_{j,t} \sim N(0,R)$ and find weight

$$w_{j,t} = X_t\beta + Z_t\alpha_j + \varepsilon_{j,t}.$$

Moreover, use the weights to find the daily growth g during a week.

Step 2 For each week t and pig j, use (3.19) and (3.20) to find the carcass weight and leanness (b days ahead), respectively. Moreover, use (3.21) to find the feed intake for the next $t = 7$ and b days, i.e. (3.22) is calculated.

Step 3 For each week t, pig j and possible centre point of pork price, calculate the settlement pork price (3.23).

Step 4 For each week t, sort the obtained values of feed intake and settlement pork price in non-decreasing order of weight.

We run the simulation 10000 times to calculate average values of the feed intake and settlement pork price and next use the values to calculate the expected values (3.6)-(3.9).

3.C Bayesian updating of SSMs

An SSM includes a set of observable and latent/unobservable continuous variables. The set of latent variables $\theta_{\{t=0,1,\dots\}}$ evolves over time using *system equation* (written using matrix notation)

$$\theta_t = G_t\theta_{t-1} + \omega_t, \tag{3.24}$$

where $\omega_t \sim N(0,W_t)$ is a random term and G_t is a matrix of known values. We assume that the prior $\theta_0 \sim N(m_0,C_0)$ is given. Moreover, we have a set of observable variables $y_{\{t=1,2,\dots\}}$ (time-series data of prices) which are dependent on the latent variable using *observation equation*

$$y_t = F_t'\theta_t + \nu_t, \tag{3.25}$$

with $\nu_t \sim N(0,V_t)$. Here F_t is the design matrix of system equations with known values and F' denote the transpose of matrix F. The error sequences ω_t and ν_t are internally and mutually independent. Hence given θ_t we have that y_t is independent of all other observations and in general the past and the future are independent given the present.

Let $\mathbb{D}_{t-1} = (y_1,\dots,y_{t-1},m_0,C_0)$ denote the information available up to time $t-1$. Given the posterior of the latent variable at time $t-1$, we can use Bayesian updating (the Kalman filter) to update the distributions at time t (West and Harrison, 1997, Thm 4.1).

Theorem 5 *Suppose that at time $t-1$ we have*

$$(\theta_{t-1} \mid \mathbb{D}_{t-1}) \sim N(m_{t-1}, C_{t-1}), \quad \textit{(posterior at time } t-1\textit{).}$$

then

$$\begin{aligned}
(\theta_t \mid \mathbb{D}_{t-1}) &\sim N(a_t, R_t), && \textit{(one-step state distribution)} \\
(y_t \mid \mathbb{D}_{t-1}) &\sim N(f_t, Q_t), && \textit{(one-step forecast distribution)} \\
(\theta_t \mid \mathbb{D}_t) &\sim N(m_t, C_t), && \textit{(posterior at time t)}
\end{aligned}$$

where

$$\begin{aligned}
a_t &= G_t m_{t-1}, & R_t &= G_t C_{t-1} G_t' + W_t \\
f_t &= F_t' a_t, & Q_t &= F_t' R_t F_t + V_t \\
e_t &= y_t - f_t, & A_t &= R_t F_t Q_t^{-1} \\
m_t &= a_t + A_t e_t, & C_t &= R_t - A_t Q_t A_t'.
\end{aligned}$$

Note that the mean of the one-step state or forecast distribution, a_t or f_t, only depends on m_{t-1}. Moreover variance C_t only depends on the number of observations made, i.e. we can calculate it without knowing the observations $y_1, \ldots, y_t$. Similarly, we can find k-step conditional distributions.

Theorem 6 *Suppose that at time t we have*

$$(\theta_t \mid \mathbb{D}_t) \sim N(m_t, C_t), \quad \textit{(posterior at time } t\textit{).}$$

then

$$\begin{aligned}
(y_{t+k}|m_t) = (y_{t+k} \mid \mathbb{D}_t) &\sim N(f_t(k), Q_t(k)), && \textit{(k-step forecast distribution)} \\
(m_{t+k}|m_t) = (m_{t+k}|\mathbb{D}_t) &\sim N\left(a_t(k), A_t(k) Q_t(k) A_t'(k)\right), && \textit{(k-step posterior mean distribution)}
\end{aligned}$$

where $f_t(k) = F_{t+k}' a_t(k)$, $Q_t(k) = F_{t+k}' R_t(k) F_{t+k} + V_{t+k}$ and $A_t(k) = R_t(k) F_{t+k} Q_t(k)^{-1}$ which can be recursively calculated using

$$\begin{aligned}
a_t(k) &= G_{t+k} a_t(k-1), \\
R_t(k) &= G_{t+k} R_t(k-1) G_{t+k}' + W_{t+k},
\end{aligned}$$

with starting values $a_t(0) = m_t$ and $R_t(0) = C_t$.

PROOF First, note that the probability distribution of $(y_{t+k}|\mathbb{D}_t)$ and the related proof have been given in (West and Harrison, 1997, Thm 4.2). Moreover, since $f_t(k)$ is a function of m_t, we have that $(y_{t+k} \mid \mathbb{D}_t) = (y_{t+k}|m_t)$.

Next, to fine the probability distribution of $(m_{t+k}|\mathbb{D}_t)$, we use the similar procedure given in the proof of Theorem 4.2 in (West and Harrison, 1997, page 107-108). According to the repeated application of system equation in an SSM (West and Harrison, 1997, page 107), the k-step evolution of latent variable θ_t can be formulated as

$$\theta_{t+k} = G_{t+k}(k)\theta_t + \sum_{r=1}^{k} G_{t+k}(k-r)\omega_{t+r}, \tag{3.26}$$

where $G_{t+k}(r) = G_{t+k}G_{t+k-1}...G_{t+k-r+1}$ for $r < k$, with $G_{t+k}(0) = I$. Now, using (3.24), (3.25) and (3.26), we can generate an SSM modelling the k-step evolution of θ_t:

Observation equation: $y_{t+k} = F'_{t+k}\theta_{t+k} + \nu_{t+k}$

System equation: $\theta_{t+k} = G_{t+k}(k)\theta_t + \sum_{r=1}^{k} G_{t+k}(k-r)\omega_{t+r}$.

For this SSM we can use the general properties of Theorem 5 with $t-1$, t, G_t and ω_t replaced with t, $t+k$, $G_{t+k}(k)$ and $\sum_{r=1}^{k} G_{t+k}(k-r)\omega_{t+r}$, respectively. Hence

$$m_{t+k} = a_t(k) + A_t(k)e_t(k),$$

where

$$a_t(k) = G_{t+k}(k)m_t, \quad e_t(k) = y_{t+k} - f_t(k).$$

Based on these equations and $(y_{t+k}|m_t) \sim N(f_t(k), Q_t(k))$, we have that

$$(m_{t+k}|m_t) \sim N(a_t(k), A_t(k)Q_t(k)A_t(k)'),$$

where based on the recursive equation for $G_{t+k}(r)$ and Theorem 5, we have that

$$\begin{aligned}
a_t(k) &= G_{t+k}(k)m_t = G_{t+k}a_t(k-1),\\
R_t(k) &= G_{t+k}(k)C_tG'_{t+k} + \sum_{r=1}^{k} G_{t+k}(k-r)W_{t+r}G_{t+k}(k-r)'\\
&= G_{t+k}R_t(k-1)G'_{t+k} + W_{t+k},\\
A_t(k) &= R_t(k)F_{t+k}Q_t(k)^{-1},\\
Q_t(k) &= F'_{t+k}R_t(k)F_{t+k} + V_{t+k},
\end{aligned}$$

which finishes the proof.

Chapter 4

Paper III: An approximate dynamic programming approach for sequential pig marketing decisions at herd level

History: *This paper has been written for submission to a peer-reviewed operations research journal. The paper has been presented at PigIT annual meeting, February 2016, Copenhagen, Denmark.*

An approximate dynamic programming approach for sequential pig marketing decisions at herd level

Reza Pourmoayed

CORAL, Department of Economics and Business Economics, Aarhus University, Fuglesangs Allé 4, DK-8210 Aarhus V, Denmark.

Abstract:

One of the most important operations in the production of growing/finishing pigs is the marketing of pigs for slaughter. The production may be considered at different levels such as animal, pen, section, and herd imposed by sectioning the production facility. Moreover, cross-level constraints may have an impact on the optimal marketing policy. As a result, it is beneficial to optimize all levels simultaneously at herd level. In this paper, we consider sequential marketing decisions at herd level taking into account other levels. A high-dimensional infinite-horizon Markov decision process (MDP) is formulated which, due to the curse of dimensionality, cannot be solved using standard MDP optimization techniques. Instead approximate dynamic programming (ADP) is applied to solve the model and find the best marketing policy at herd level. Under the total expected discounted reward criterion, the proposed ADP approach is first compared with a standard solution algorithm of MDP at pen level (where the optimal policy can be found) to show the accuracy of the solution procedure. Next, numerical experiments at herd level are given to confirm how the marketing policy adapts itself to different costs and cross-level constraints (e.g. transportation cost). Finally, the marketing policy found by ADP is compared with other well-known marketing polices, often applied at herd level.

Keywords: Approximate dynamic programming; Markov decision process; herd management.

4.1 Introduction

One of the most important operations in the production of growing/finishing pigs is the marketing of pigs for slaughter. Each week, the farm manager should decide which pigs should be delivered to the abattoir and when the pen (or section) should be emptied (Kure, 1997). In the production system, animals may be considered at different levels: herd, section, pen, or animal. The herd is a group of sections, a section includes a number of pens, and a pen involves some animals (usually 15-20 animals). Marketing decisions can be considered at different levels, e.g. animal (Glen, 1983) or pen (Kristensen et al., 2012; Kure, 1997). The complexity of the marketing decisions depends on the number of levels that are taken into account simultaneously and on how decisions at different levels are linked together. In other words, cross-level constraints between pen, section, and herd level can affect marketing decisions and they should be considered in the problem. In this paper, we focus on marketing decisions at herd level. In the following, different aspects of the problem are described.

The production process of growing/finishing pigs is started by buying piglets on the market or transferring them from another production unit when they weigh approx. 30 kg. Next, the piglets are moved to the sections of a finisher unit where they grow until marketing (9-12 weeks). In the finisher unit, pigs in general grow at different growth rates and hence they obtain their slaughter weight at different times in the last weeks of the growing period. At the end of the growing period, the decision maker should therefore determine which pigs should be selected for slaughter (individual marketing) in each pen. After a sequence of individual marketings at pen level, the decision maker must decide when to terminate (empty) each section. Terminating a section means that the remaining pigs in the section are sent to the slaughterhouse (in one delivery) and when the section has been cleaned, another group of piglets (with a weight of approx. 30 kg) is inserted into the pens of the section and the production is repeated in the section.

During the marketing period, marketed pigs from the pens are grouped in one weekly delivery and are transported to the abattoir by means of a number of trucks. Depending on the number of marketed pigs at herd level and the capacity of each truck, the decision maker should also determine the number of trucks needed to transport the pigs to the abattoir. That is, transportation costs may have an effect on the marketing policy of the fattening pigs. In most Danish herds, transportation of culled pigs is handled by a single abattoir and hence the cost of transportation is fixed in the system. Notice that the reward of marketing depends on the price of the carcass weight in the abattoir, the cost of buying the piglets, feeding the pigs in the production system, and the cost of transporting the culled pigs to the abattoir. The best meat price is obtained if the carcass weight lies within a specific interval. Therefore, the farmer must time the marketing decisions while simultaneously considering the carcass weight in relation to the best interval, the transportation cost of trucks, and the length of the production cycle for feeding the rest of the pigs. For an extended overview over pig production of growing/finishing pigs, see Pourmoayed and Nielsen (2014).

Marketing decisions have been studied by a number of researchers. Kure (1997) proposed a recursive dynamic programming method and used replacement theory concepts to find the best marketing strategy. In the study by Jørgensen (1993), a *hierarchical Markov decision process (HMDP)* was applied to analyze the precision of the weighing methods on the marketing policy of fattening pigs. Toft et al. (2005) combined decisions related to the delivery strategy of pigs to the abattoir and epidemic diseases using a multi-level HMDP. In the study by Kristensen et al. (2012) an HMDP was employed to model marketing decisions under online weight information and the focus was on the definition of the state space of the HMDP acquired by dynamic linear models and Bayesian updating. In the study by Plà-Aragonés et al. (2013), the optimal marketing policy was found by a mixed integer linear programming method under an all-in all-out strategy. They formulated the problem by a mathematical programming model and solved their model using a heuristic approach under different pig size distributions and pig growth rates. Niemi (2006) applied a stochastic dynamic programming method to find the best time of marketing for an individual pig and the best nutrient ingredients in the feed-mix, simultaneously. In a recent study, Pourmoayed et al. (2016) have considered optimal marketing and feeding strategies at pen level.

To the best of our knowledge, there are few studies considering the problem at herd level and taking into account the effect of transportation costs on marketing decisions (Ohlmann and Jones, 2008; Boys et al., 2007). Ohlmann and Jones (2008) used a mixed integer programming model to find the best marketing strategy in a barn of pigs. Boys et al. (2007) analyzed the effect of single and multiple shipping decisions on the marketing strategy of a heterogeneous herd using a simulation method. In both studies marketing decisions are examined under an annual profit criterion and the impact of physical sectioning of the production facility with respect to the pen, section and herd levels is not considered in the models. Moreover, in these studies termination decisions are not considered in the models and therefore they do not take into account the effect of terminating the production cycle at an earlier point in time, even though such earlier termination would increase the number of production cycles and possibly lead to higher profits.

In this paper we consider marketing decisions in a herd composed of sections and pens. We formulate the sequential marketing decisions using a discounted infinite-horizon *Markov decision process (MDP)* and assume that the production process is cyclic at section level, i.e. when a section is terminated, a new batch of piglets is inserted into the pens of this section and a new production cycle is started. The model is stochastic because of the uncertainty of the weight of pigs in the pens. This uncertainty is described by a stochastic process relying on the *state space models* formulated in Pourmoayed et al. (2016). Due to the large number of states and actions in the model, the *curse of dimensionality* becomes apparent and the usual solution procedures of MDPs (e.g. policy iteration) can not be used to solve the model. Therefore, we use an approximation strategy to give an approximate solution to the problem and find the best approximate marketing policy at herd level. More precisely, we first use the properties of the value function in a discounted infinite-horizon MDP at pen level to find an approximation architecture for the value function at herd level, and next we apply an *approximate dynamic programming (ADP)* approach with *post-decision states* to find the best marketing policy at herd level (Powell, 2007).

Examples of approximation strategies used to optimize livestock systems are Kristensen (1992) and Ben-Ari and Gal (1986) that exploit a parameter iteration algorithm developed by Gal (1989) to find the best replacement policy in a dairy herd. In these studies, however, the estimation of transition probabilities and the calculation of expected value operators in the solution procedure are computationally challenging/require much computational effort, due to complicated transition functions and large state and action spaces. In the present study, reformulating the Bellman equations in the form of *post-decision states* can significantly improve the computational efficiency of the solution procedure which is a novel approach for optimizing livestock systems modeled by high-dimensional MDPs.

The rest of the paper is organized as follows. In Section 4.2, sequential marketing decisions are modeled using a discounted infinite-horizon MDP model. Section 4.3 describes an approximate dynamic programming approach for solving the MDP model. In Section 4.4, numerical examples are given and the policy resulting from the ADP is compared with other marketing policies, and finally in Section 4.5, we conclude the paper.

4.2 Model description

We use a discounted infinite-horizon MDP to model marketing decisions at herd level. A short description of the model is given below.

A discounted infinite-horizon MDP models a sequential decision problem over an infinite time horizon. Assume that a decision epoch is the first day of a week when marketing decisions are made and let $\mathbb{S}$ denote the finite set of system states at an arbitrary decision epoch. Given system state $s \in \mathbb{S}$ at the current decision epoch, an action a from the finite set of allowable actions $\mathbb{A}(s)$ is chosen resulting in an immediate reward $r(s,a)$, and a probabilistic transition to state $s_+ \in \mathbb{S}$ at the next decision epoch. This transition is based on the transition function $\phi(s,a,\omega)$ where ω denotes random information received between the current and next decision epochs. Random information ω might depend on a stochastic process affecting the system state.

We consider the following assumptions when modeling sequential marketing decisions at herd level:

1. A herd consists of $|I|$ sections and each section $i \in I$ includes $|J|$ pens.
2. Each pen $j \in J$ involves a maximum of $q^{\max}$ pigs. In the beginning of the production cycle in the section, each pen is filled with $q^{\max}$ pigs.
3. The marketing decisions are taken on a weekly basis and the culled pigs are transferred to the abattoir after few days.
4. Individual marketing at pen level is started in week $t^{\min}$ at the earliest.
5. A section is terminated in week $t^{\max}$ at the latest, i.e. the maximum life time of a pig is $t^{\max}$ weeks.
6. Weekly deliveries to the abattoir are based on a cooperative agreement where culled pigs from each section in the herd are grouped and transferred to the abattoir by trucks. Transportation of culled pigs is handled by the abattoir and hence the cost of transportation per truck is fixed in the model. Variable costs of transportation, e.g. costs of loading a pig into the truck, are not considered in the model of this paper.
7. When a section is terminated and cleaned, a new batch of piglets is inserted into the pens of this section immediately, i.e. the piglets are always available. As a consequence, information on the growth and weight of the piglets is known at insertion time.
8. The production process in a section is independent of other sections, i.e. the piglets can be inserted into the sections at different times, and a section can be terminated earlier compared to other sections.
9. The sequence of feed-mixes used during the production cycle (feeding strategy) is known and fixed.
10. A new batch of piglets and the required feed stock are bought using known and fixed prices.
11. The pigs are sold to the abattoir using a known settlement pork price function.

12. The growth of a pig is independent of the other pigs in the pen, i.e. the growth is not dependent on the number of pigs in the pen.

Under these assumptions, we describe the state space, action space, transition function, reward, and the optimality criterion of the discounted infinite-horizon MDP model in the following.

4.2.1 State space

In a given decision epoch, state s is defined using state variables:

t_i week number in a production cycle of section i, $\forall\, i \in I$;

q_{ij} number of remaining pigs in pen j of section i, $\forall\, i \in I$ & $j \in J$;

w_{ij} model information related to the weight of pigs in pen j of section i, $\forall\, i \in I$ & $j \in J$. $w_{ij} \in W$ is a vector of the three state variables μ_{ij}, σ_{ij}, and g_{ij} corresponding to the mean and standard deviation of weight and the average growth of pigs in pen j of section i, respectively, i.e. $w_{ij} = (\mu_{ij}, \sigma_{ij}, g_{ij})$. The values of these variables can be obtained from repeated measurements of weight and feed intake data in the farm using state space models based on Bayesian updating (see Pourmoayed et al. (2016)).

In general state s takes the form

$$s = (\vec{t}, \vec{q}, \vec{w}),$$

where

$$\begin{aligned}
\vec{t} &= (t_1, .., t_i, .., t_{|I|}),\\
\vec{q} &= (q_{11}, .., q_{ij}, .., q_{|I||J|}),\\
\vec{w} &= (w_{11}, .., w_{ij}, .., w_{|I||J|}).
\end{aligned}$$

Hence the set of states becomes

$$\begin{aligned}
\mathbb{S} = \{s = (\vec{t}, \vec{q}, \vec{w}) \mid i \in I,\ j \in J,\ t_i \in \{1, \ldots, t^{\max}\},\\
q_{ij} \in \{0 \cdot \mathbf{I}_{\{t_i > t^{\min}\}} + q^{\max}\mathbf{I}_{\{t_i \leq t^{\min}\}}, \ldots, q^{\max}\}, w_{ij} \in W\},
\end{aligned}$$

where $\mathbf{I}_{\{\cdot\}}$ denotes the indicator function.

4.2.2 Action space

Consider state s and assume that the pigs in pen j of section i are sorted in ascending order based on their live weight such that index k denotes the kth heaviest pig in this pen. We consider the following decision variables for defining action a:

x_{ijk} a binary variable equal 1 if the kth pig in pen j of section i is culled, $\forall\, i \in I,\ j \in J,\ \&\ k \in \{1, 2, \dots, q_{ij}\}$;

y_i a binary variable equal 1 if section i is terminated, $\forall i \in I$.

Using decision variables x_{ijk} and y_i, action a is defined as

$$a = (\vec{x}, \vec{y}),$$

where

$$\begin{aligned}
\vec{x} =& (x_{111}, .., x_{i,j,k}, .., x_{|I||J|q_{|I||J|}}), \\
\vec{y} =& (y_1, .., y_j, .., y_{|I|}).
\end{aligned}$$

Possible actions for state s must satisfy the following constraints:

$$x_{ijk} \le x_{ijk+1}, \quad \forall\, i \in I,\ j \in J,\ \&\ k \in \{1, 2, \dots, q_{ij} - 1\}, \tag{4.1}$$

$$\sum_{j \in J} q_{ij} - \sum_{j \in J} \sum_{k=1}^{q_{ij}} x_{ijk} \le M(1 - y_i), \quad \forall\, i \in I, \tag{4.2}$$

$$\sum_{j \in J} q_{ij} - \sum_{j \in J} \sum_{k=1}^{q_{ij}} x_{ijk} \ge (1 - y_i), \quad \forall\, i \in I, \tag{4.3}$$

$$\sum_{i \in I} \sum_{j \in J} \sum_{k=1}^{q_{ij}} x_{ijk} \le k^{\texttt{truck}} z, \tag{4.4}$$

$$x_{ijk} = 0, \qquad \text{if } 1 \le t_i < t^{\min}, \quad \forall\, i \in I,\ j \in J,\ \&\ k \in \{1, 2, \dots, q_{ij}\}, \tag{4.5}$$

$$y_i = 1, \qquad \text{if } t_i = t^{\max}, \quad \forall\, i \in I, \tag{4.6}$$

$$x_{ijk},\ y_i, \qquad \text{binary}, \forall\, i \in I,\ j \in J,\ \&\ k \in \{1, 2, \dots, q_{ij}\},\ z \text{ integer}. \tag{4.7}$$

Based on the order of the sorted pigs in the pen, constraint (4.1) enforces that a lighter pig cannot be marketed earlier than the heavier pigs. When a termination happens in section i ($y_i = 1$), constraint (4.2) implies that all the remaining pigs in the pens of this section must be marketed. In this constraint, M is a predefined large number. Moreover, when all pigs are marketed from section

i, constraint (4.3) ensures that this section must be terminated. Constraint (4.4) expresses that the total number of marketed pigs in the herd must be less than the capacity of the trucks called from the abattoir. In this constraint, $k^{\texttt{truck}}$ is the capacity of one truck and z is a supplementary integer variable denoting the number of trucks needed to transport the marketed pigs to the abattoir. In constraint (4.5), marketing is not allowed when $1 \leq t_i < t^{\min}$ in a production cycle at section i. Constraint (4.6) ensures that termination in section i occurs in the latest week of a production cycle ($t_i = t^{\max}$). Finally, constraint (4.7) defines the type of decision variables x_{ijk}, y_i and z.

If $\mathbb{X}_s$ and $\mathbb{Y}_s$ are the set of decision variables x_{ijk} and y_i satisfying the above constraints, then the set of possible actions $\mathbb{A}(s)$ for the given state s is defined as

$$\mathbb{A}(s) = \{a = (\vec{x}, \vec{y}) | \vec{x} \in \mathbb{X}_s, \vec{y} \in \mathbb{Y}_s\}.$$

4.2.3 Rewards

The reward of action $a \in \mathbb{A}(s)$ is calculated as the revenue of selling the marketed pigs to the abattoir minus the cost of feeding the pigs that we decide to keep in the pens until the next decision epoch, the cost of transferring the marketed pigs to the abattoir by trucks, and the cost of buying a new batch of piglets when a section has been terminated. Hence, the reward associated with state s and action a is formulated as

$$r(s,a) = \sum_{i \in I} \sum_{j \in J} \sum_{k=1}^{q_{ij}} c_k^{\texttt{cull}}(w_{ij}) x_{ijk} - \sum_{i \in I} \sum_{j \in J} \sum_{k=1}^{q_{ij}} c_k^{\texttt{feed}}(w_{ij})(1 - x_{ijk}) - c^{\texttt{truck}} z - \sum_{i \in I} c^{\texttt{piglet}} |J| q^{\max} y_i, \tag{4.8}$$

where $c_k^{\texttt{cull}}(w_{ij})$ is the unit reward of selling the kth pig in pen j of section i to the abattoir given weight information w_{ij}. Similarly, $c_k^{\texttt{feed}}(w_{ij})$ denotes the feeding cost of the kth pig in pen j of section i kept in the herd until the next decision epoch. Note that when marketing decisions are made, culled pigs are sent to the abattoir after few days and therefore the additional feeding cost and reward, resulting from the weight gain of culled pigs in this period, are considered in the calculation of $c_k^{\texttt{cull}}(w_{ij})$. We use a simulation method to calculate functions $c_k^{\texttt{cull}}(w_{ij})$ and $c_k^{\texttt{feed}}(w_{ij})$ given weight information w_{ij} (see Appendix 4.B). The coefficient $c^{\texttt{truck}}$ is the fixed cost of a truck to transfer the culled pigs to the abattoir, and $c^{\texttt{piglet}}$ is the cost of buying a new piglet in the beginning of a production cycle.

4.2.4 Transition function

Given state $s = (\vec{t}, \vec{q}, \vec{w})$ and action a, the transition function $\phi(s, a, \omega)$ describes how the system evolves from state s to state s_+:

$$s_+ = \phi(s, a, \omega) = (\vec{t}_+, \vec{q}_+, \vec{w}_+), \tag{4.9}$$

such that for every $i \in I$ and $j \in J$:

$$\begin{aligned}
t_{i_+} &= (1 - y_i)(t_i + 1) + y_i, \\
q_{ij_+} &= (1 - y_i)(q_{ij} - \sum_{k=1}^{q_{ij}} x_{ijk}) + y_i q^{\max}, \\
w_{ij_+} &= (1 - y_i)\Gamma(w_{ij}, \omega) + y_i w_0,
\end{aligned}$$

where w_0 is the weight information at the start of a production cycle in the section and $\Gamma(w_{ij}, \omega)$ describes a stochastic transition between w_{ij} and w_{ij_+} given random information ω. In order to describe this stochastic transition, we define a stochastic process relying on two state space models with Bayesian updating suggested in Pourmoayed et al. (2016), i.e. given weight information $w_{ij} = (\mu_{ij}, \sigma_{ij}, g_{ij})$, the transition probabilities for the state variables μ_{ij}, σ_{ij}, and g_{ij} are obtained using state space models based on Bayesian updating. For more details about this stochastic process see Appendix 4.A.

4.2.5 Optimality criterion

A *policy* π is a decision rule or function that assigns for each state $s \in \mathbb{S}$ an action $a = \pi(s) \in \mathbb{A}(s)$, i.e. a policy prescribes which action to take whenever the system is observed in state s. Under the optimality criterion *maximization of total expected discounted reward*, the objective of the model is to find the best policy that maximizes the total expected discounted reward over an infinite time horizon:

$$\max_{\pi \in \Pi} \mathbb{E}\left(\sum_{n=1}^{\infty} \gamma^{n-1} r(s_n, \pi(s_n)) \right), \tag{4.10}$$

where γ is the discount factor used to calculate the present value of a future cash flow, subscript n indicates stage number, and Π is the set of all deterministic stationary policies. In order to find the optimal actions $a^* = \pi^*(s)$, the following optimality equations should be satisfied for all states $s \in \mathbb{S}$ (Puterman, 2005, Section 6.2)

$$\nu(s) = \max_{a \in \mathbb{A}(s)} \left(r(s, a) + \gamma \mathbb{E}\left(\nu(s_+)\right) \right), \quad \forall s \in \mathbb{S}, \tag{4.11}$$

where $s_+ = \phi(s,a,\omega)$, and the value function $v(s)$ denotes the maximum expected discounted reward of being in state s to the end of the time horizon. Notice that when $0 < \gamma < 1$, $\mathbb{S}$ and $\mathbb{A}(s)$ are finite and $r(s,a)$ is bounded, a unique solution to the optimality equations (4.11) will exist and the resulting policy is a deterministic stationary optimal policy (Puterman, 2005, Theorems 6.2.5-6.2.10).

4.3 Approximate dynamic programming

In the calculation of the value function $v(s)$ for all states $s \in \mathbb{S}$ using the optimality equations in (4.11), we face some computational obstacles. First, the size of the state and action spaces in the model are $t^{\max}q^{\max}|W|^{|I||J|}$ and $2^{q^{\max}|I||J|+|I|}$, respectively, which means that reasonable values of $t^{\max}, q^{\max}, |W|, |I|$, and $|J|$ (e.g. 15, 18, 1134, 3, and 20, respectively) make computation of $v(s)$ for every possible state unmanageable. Second, due to the large number of states and possible outcomes for random information ω, an exact computation of expected value $\mathbb{E}\left(v(s_+)\right)$ is prohibitive. Finally, due to the expected value operator, the maximization problem in (4.11) is not deterministic and hence it may be difficult to solve it and find the optimal actions. These computational challenges are known as three curses of dimensionality (Powell, 2007, Section 4.1) that prevent us from applying the regular solution procedures of MDPs (e.g. policy and value iteration methods) to solve the discounted infinite-horizon MDP model.

Approximate dynamic programming is an efficient way to deal with these computational problems and to find an approximate solution for the high-dimensional MDPs. The main idea is to approximate the value function $v(s)$ and find the best actions for the states that are most likely observed in the system. The approximation architecture of the value function is often described by a parametric function and there are well-known algorithms exploiting simulation and linear programming techniques to estimate the parameters of the approximated value function (see e.g. Powell (2010), Topaloglu and Powell (2006), Toriello et al. (2010), de Farias and Van Roy (2003), and Patrick et al. (2008)). For more details about ADP algorithms, the interested reader may refer to Powell (2007). Examples of approximation methods in optimizing livestock systems are Kristensen (1992) and Ben-Ari and Gal (1986) that use the parameter iteration method suggested by Gal (1989).

In this section, we apply the ADP approach and use post-decision states to give an approximate solution to the discounted infinite-horizon MDP model. First, in Section 4.3.1, we design a parametric approximation architecture for the value function $v(s)$. Second, in Section 4.3.2, the optimality equations are reformulated in terms of post-decision states to find a deterministic version of the maximization problem in (4.11), and finally in order to estimate the parameters of the approximated value function, an approximate value iteration algorithm is employed and presented in Section 4.3.3.

4.3.1 Approximation architecture of the value function

When the sizes of the state and action spaces are too large, it is often not possible to update the value function for every possible state. An alternative option is to find an approximation architecture for the value function $v(s)$. When choosing the approximation architecture, we should consider a balance between the computational efficiency and the performance of the resulting policy. For instance, a second or third polynomial approximation for the value function $v(s)$ may change the maximization problem in (4.11) to a very difficult non-linear model. In this study, we use the properties of the value function of the marketing decisions at pen level to find an approximation architecture for the value function at herd level.

Based on the hierarchy structure of the marketing decisions at herd level, we will assume that the value function at herd level is additive over the value functions at pen level and hence an approximation of the value function $v(s)$ is given as:

$$v(s) \approx \sum_{i \in I} \sum_{j \in J} v_{ij}(t_i, q_{ij}, w_{ij}), \tag{4.12}$$

where $v_{ij}(t_i, q_{ij}, w_{ij})$ is the value function related to pen j of section i with weight information w_{ij} including q_{ij} pigs in week number t_i. According to the form of the value function of marketing decisions at pen level in Pourmoayed et al. (2016), we found that a good functional form of $v_{ij}(t_i, q_{ij}, w_{ij})$ is

$$v_{ij}(t_i, q_{ij}, w_{ij}) \approx q_{ij} b(t_i, w_{ij}), \tag{4.13}$$

where $b(t_i, w_{ij})$ is a parameter describing the effect of weight information w_{ij} and week number t_i on the marketing decisions at pen level. In other words, $b(t_i, w_{ij})$ is the value of having one pig more in pen j of section i with weight information w_{ij} at week number t_i. We may infer that this parameter is the gradient of the value function with respect to the number of remaining pigs, q_{ij}, and hence we name it a slope parameter in the approximation architecture.

By substituting (4.13) in (4.12), the final form of the approximation architecture for the value function at herd level is as

$$v(s) \approx \sum_{i \in I} \sum_{j \in J} q_{ij} b(t_i, w_{ij}). \tag{4.14}$$

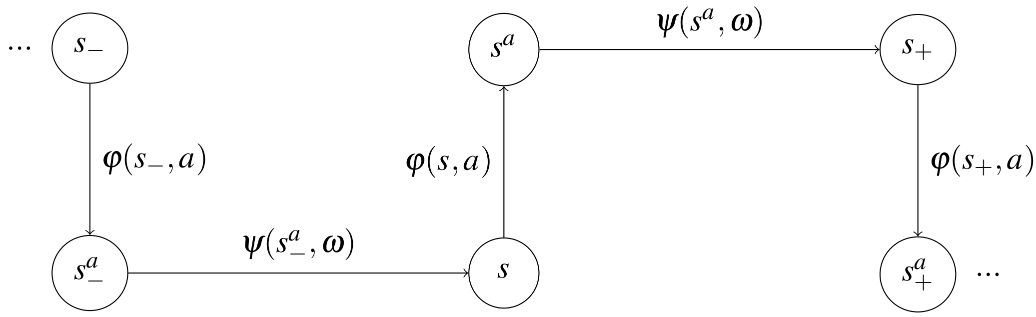


Figure 4.1: Relation between pre- and post-decision states in three decision epochs. $\varphi(.)$ indicates a deterministic transition from pre-decision to post-decision states, and $\psi(.)$ shows a stochastic transition from post-decision to pre-decision states. Subscripts '-' and '+' clarify a state in the previous and the next decision epochs, respectively.

Now instead of updating the value function for every possible state, we only need to find a very good estimation for the slope parameters $b(t_i,w_{ij})$ ($t_i \in \{1,...,t^{\max}\}$ and $w_{ij} \in W$). Note that the number of slope parameters, $t^{\max}|W|$, needed to be estimated is much lower than the number of possible states, $t^{\max}q^{\max}|W|^{|I||J|}$, and therefore the computational effort of the solution procedure is noticeably reduced. In section 4.3.3, we apply an approximate value iteration algorithm to estimate the slope parameters.

4.3.2 Post-decision state

In the transition function $\phi(s,a,\omega)$ defined in (4.9), we can separate the effect of action a and the random information ω. Therefore, for a given state $s=(\vec{t},\vec{q},\vec{w})$ and action a, transition to $s_+=\phi(s,a,\omega)$ is divided into two steps:

$$\begin{aligned} s^a =& \varphi(s,a) = (\vec{t^a},\vec{q^a},\vec{w^a}) \\ s_+ =& \psi(s^a,\omega) = (\vec{t}_+,\vec{q}_+,\vec{w}_+) \end{aligned} \tag{4.15}$$

where for every $i \in I$ and $j \in J$:

$$\begin{aligned} t_i^a &= (1-y_i)(t_i+1)+y_i, \\ q_{ij}^a &= (1-y_i)(q_{ij}-\sum_{k=1}^{q_{ij}} x_{ijk})+y_i q^{\max}, \\ w_{ij}^a &= (1-y_i)w_{ij}+y_i w_0. \end{aligned}$$

Here, s^a is known as the *post-decision state*, and s_+ as the *pre-decision state* (Powell, 2007). $\varphi(s,a)$ is a deterministic transition that only considers the effect of action a on the pre-decision state s while $\psi(s^a,\omega)$ is a stochastic transition taking into account the effect of random information ω on the post-decision state s^a. We may infer that the post-decision state s^a is the state of the system immediately after making a decision while the pre-decision state s refers to the state of the system just before we make a decision. Figure 4.1 shows the relation between pre- and post-decision states in three successive decision epochs.

Now suppose $s \in \mathbb{S}$ is the pre-decision state of the system at the current decision epoch. Moreover, assume that $s^a \in \mathbb{S}$ and $s^a_- \in \mathbb{S}$ are two post-decision states in the current and the previous decision epochs, receptively. According to the transition function $\psi(s^a_-,\omega)$, the pre-decision state s is the result of a stochastic transition from post-decision state s^a_- (see Figure 4.1) and hence we can conclude that

$$\nu(s^a_-) = \mathbb{E}\left(\nu(s)|s^a_-\right). \tag{4.16}$$

Furthermore, by reformulating the optimality equations in (4.11), the value of being in pre-decision state s can be obtained using $\nu(s^a)$:

$$\nu(s) = \max_{a\in\mathbb{A}(s)} (\, r(s,a) + \gamma\nu(s^a)\,), \tag{4.17}$$

where $s^a = \varphi(s,a)$. Now by substituting (4.17) in (4.16), the new optimality equation for the value function in post-decision state s^a_- will be

$$\nu(s^a_-) = \mathbb{E}\left(\max_{a\in\mathbb{A}(s)} (\, r(s,a) + \gamma\nu(s^a))|s^a_-\right). \tag{4.18}$$

Note that in (4.18), the expectation is now outside the max operator and the main benefit of using this new form of optimality equation, compared to Equation (4.11), is that the optimization problem inside the expectation is deterministic and hence it can be solved by well-known optimization techniques. However, to calculate $\nu(s^a_-)$, the expectation in equation (4.18) must be computed. Calculation of this expected value is often computationally intractable and is usually approximated using simulation techniques. For more details about post-decision states see (Powell, 2007, Chapter 4).

4.3.3 Approximate value iteration algorithm

In this section, we present an *approximate value iteration (AVI)* algorithm, the purpose of which is to estimate the value of the slope parameters in the approximation architecture of

Algorithm 1 AVI

Initialization:
  Estimate initial values of slope parameters $\hat{b}(t_i^a, w_{ij}^a)$ for every $t_i^a \in \{1, ..., t^{\max}\}$ & $w_{ij}^a \in W$.
  Set h=1 (iteration counter).
while $h \leq \mathscr{H}$ or $||\vec{b}_h - \vec{b}_{h-1}|| \leq \lambda$ **do**
  Initialize s_0 for iteration h and set $s := s_0$ and $\vec{b}_- := \vec{b}$
  for $n = 1, 2, ..., \mathscr{N}$ **do**
    Solve maximization problem:
      $a^* = arg\ max_{a \in \mathbb{A}(s)}(r(s,a) + \gamma \tilde{v}_{\vec{b}_-}(s^a)),$
    if n>1 **then**
      Run Algorithm 2 given pre- and post-decision states s and $s^{\tilde{a}}$ and action $a^* = (\vec{x}^*, \vec{g}^*)$ to update the slope parameters.
    end if
    Set $\tilde{a} = a^*$ and find the post-decision state:
      $s^{\tilde{a}} = \varphi(s, \tilde{a}).$
    Generate sample $\hat{\omega}$ and find the next pre-decision state:
      $s_+ = \psi(s^{\tilde{a}}, \hat{\omega}).$
    Set $n := n+1$ and $s := s_+$
  end for
  $h = h+1$
end while
Return $\hat{b}(t_i^a, w_{ij}^a)$ for every $t_i^a \in \{1, ..., t^{\max}\}$ & $w_{ij}^a \in W$.

the value function defined in (4.14). In this algorithm, the parametric approximation architecture $\tilde{v}_{\vec{b}}(.) : s^a \rightarrow \mathbb{R}$ represents the value function in terms of post-decision states where vector $\vec{b}$ shows the set of estimated slope parameters in the approximate value function, $\vec{b} = (\hat{b}(t_1^a, w_{11}^a), .., \hat{b}(t_i^a, w_{ij}^a), .., \hat{b}(t_{|I|}^a, w_{|I||J|}^a))$. Algorithm 1 illustrates the pseudo code of the algorithm. In this algorithm, first the slope parameters are initialized to known values, and then several state trajectories are simulated in two loops (outer and interior loops) and the slope parameters are updated. The main parts of the algorithm are described below.

Initialization. The algorithm is started by estimating the initial values for the slope parameters $\hat{b}(t_i^a, w_{ij}^a)$ ($t_i^a \in \{1, ..., t^{\max}\}$ and $w_{ij}^a \in W$). The initial values are estimated according to the behavior of the value function of marketing decisions at pen level in Pourmoayed et al. (2016).

Maximization problem. In each iteration of the algorithm, a maximization problem should be

Algorithm 2 Updating of slope parameters

In state $s = (\vec{t}, \vec{q}, \vec{w})$, for every $i \in I$ and $j \in J$ such that $t_i >= t^{\min}$ do:
if $q_{ij} < q^{\max}$ **then**
$q'_{ij} = q_{ij} + 1 \ \rightarrow \ s' = (\vec{t}, \vec{q}', \vec{w})$,
$\hat{\nabla} \nu = max_{a \in \mathbb{A}(s')}(r(s', a) + \gamma \tilde{\nu}_{\vec{b}_-}(s'^a)) - max_{a \in \mathbb{A}(s)}(r(s, a) + \gamma \tilde{\nu}_{\vec{b}_-}(s^a))$,
else
$q'_{ij} = q_{ij} - 1 \ \rightarrow \ s' = (\vec{t}, \vec{q}', \vec{w})$,
$\hat{\nabla} \nu = max_{a \in \mathbb{A}(s)}(r(s, a) + \gamma \tilde{\nu}_{\vec{b}_-}(s^a)) - max_{a \in \mathbb{A}(s')}(r(s', a) + \gamma \tilde{\nu}_{\vec{b}_-}(s'^a))$,
end if

$$\hat{b}(t_i^{\tilde{a}}, w_{ij}^{\tilde{a}}) = (1 - \alpha_h) \hat{b}(t_i^{\tilde{a}}, w_{ij}^{\tilde{a}})_- + \alpha_h \hat{\nabla} \nu.$$

In state $s = (\vec{t}, \vec{q}, \vec{w})$, for every $i \in I$ if $y_i^* = 1$ do:
$\hat{\nu} = \sum_{j \in J} q_{ij} \hat{b}(t_i, w_{ij}) - (\sum_{j \in J} \sum_{k=1}^{q_{ij}} c_k^{\texttt{cull}}(w_{ij}) - c^{\texttt{piglet}} q^{\max} |J|)$,

$$\hat{b}(t_i^a, w_{ij}^a) = \hat{b}(1, w_0) = (1 - \alpha_h) \hat{b}(t_i^a, w_{ij}^a)_- + \alpha_h \frac{\hat{\nu}}{q^{\max} |J| \gamma}$$

solved. This maximization problem in a given pre-decision state $s = (\vec{t}, \vec{q}, \vec{w})$ is defined as

$$max_{a \in \mathbb{A}(s)}(r(s, a) + \gamma \tilde{\nu}_{\vec{b}_-}(s^a)), \tag{4.19}$$

where the reward function $r(s, a)$ was given in (4.8) and the value function $\tilde{\nu}_{\vec{b}_-}(s^a)$ can be obtained using the approximation architecture of the value function in (4.14) and the deterministic transition of post-decision states defined in (4.15). Vector $\vec{b}_-$ in the value function includes the last estimates of the slope parameters obtained in the previous iteration of the algorithm. Moreover, the set of possible actions $\mathbb{A}(s)$ is defined using constraints (4.1), (4.2), (4.3), (4.6), (4.5), (4.4), and (4.7). Therefore the final form of the maximization problem is defined as follows:

$$\begin{aligned}
\max_{\vec{x}\,\vec{g}\,z} \ & \sum_{i \in I} \sum_{j \in J} \sum_{k=1}^{q_{ij}} c_k^{\texttt{cull}}(w_{ij}) x_{ijk} - \sum_{i \in I} \sum_{j \in J} \sum_{k=1}^{q_{ij}} c_k^{\texttt{feed}}(w_{ij})(1 - x_{ijk}) - c^{\texttt{truck}} z - \\
& \sum_{i \in I} c^{\texttt{piglet}} |J| q^{\max} y_i + \gamma \sum_{i \in I} \sum_{j \in J} \Big(\big(q_{ij} - \sum_{k=1}^{q_{ij}} x_{ijk}\big) \hat{b}(t_i^a, w_{ij}^a)_- + y_i q^{\max} \hat{b}(1, w_0)_- \Big), \\
& \text{s.t. } (4.1) \ - \ (4.7).
\end{aligned} \tag{4.20}$$

This integer programming model is solved to find the best action $a^* = (\vec{x}^*, \vec{g}^*)$ for the pre-decision state s that is used in the next steps of the algorithm for generating the post-decision state of the system and for updating the slope parameters of the value function.

Simulating the state trajectories. In the outer loop of the algorithm, in each iteration h, first an initial state s_0 is randomly selected and then from this state, a state trajectory is simulated for decision epochs $1, 2, \ldots, \mathscr{N}$ using an interior loop. Note that since we cannot simulate an infinite trajectory, parameter $\mathscr{N}$ is chosen so as to be large enough to have a good estimation of the average discounted reward of the model defined in (4.10). That is, $\mathscr{N}$ is estimated such that the error criterion $\gamma^{\mathscr{N}+1}\bar{r}/(\gamma - 1)$ gets sufficiently small. This criterion shows the difference between the average discounted reward of the model with an infinite time horizon and with a long time horizon $\mathscr{N}$, where $\bar{r}$ is a pre-estimation of the average reward per time unit in the model (for more details see (Powell, 2007, page 340)). Moreover, in order to improve the performance of the algorithm the discount factor γ is set to a smaller value than would be needed to capture the time value of money (Powell, 2007, pages 343, 597).

Next, in iteration n of the interior loop (related to decision epoch n), when pre-decision state $s = (\vec{t}, \vec{q}, \vec{w})$ is observed, the maximization problem (4.19) is solved and the value of $a = (\vec{x}, \vec{g})$ that solves this problem is stored in $\tilde{a}$ (after slope parameters have been updated). Given $\tilde{a}$, the post-decision state $s^{\tilde{a}}$ is found using the transition function $\varphi(s, \tilde{a})$. Finally, a sample of random information ω is generated (denoted by $\hat{\omega}$) and the next pre-decision state of the system is found using the transition function $\psi(s^{\tilde{a}}, \hat{\omega})$. Note that using sample $\hat{\omega}$, the transition function $\psi(s^{\tilde{a}}, \hat{\omega})$ results in a deterministic transition from $s^{\tilde{a}}$ to the next pre-decision state s_{+}.

Updating the slope parameters. The updating procedure of the slope parameters is given in Algorithm 2. Note that according to equations (4.16) and (4.18), an estimation of the value function $v(s)$ in pre-decision state s gives an estimation of the value function $v(s^a_-)$ in the previous post-decision state s^a_-. In Algorithm 2, we use the same logic to update the value of slope parameter $b(t^a_i, w^a_{ij})$ (when $t_i >= t^{\min}$). More precisely, in a given decision epoch n ($n > 1$), we first estimate the gradients of the approximate value function with respect to the number of remaining pigs and we then use them to update the slope parameters of the value function in post-decision state $s^{\tilde{a}}$ observed in decision epoch $n-1$. Note that the gradients of the value function are estimated according to the weight information observed in all the pens. In order to calculate the gradient $\hat{\nabla} v$ in a pen with information w_{ij} and t_i, first we add one average weighted resource into this pen (when there are $q^{\max}$ pigs in the pen, one average weighted pig is removed from this pen) and then we calculate the change in the objective function value of the maximization problem (4.20) in which, the coefficients $c_k^{\texttt{cull}}(w_{ij})$ and $c_k^{\texttt{feed}}(w_{ij})$ are replaced by the reward and cost of an average weighted pig.

Now using the estimated gradient $\hat{\nabla}v$, the value of the slope parameter $b(t_i^a, w_{ij}^a)$ is updated as follows

$$\hat{b}(t_i^{\tilde{a}}, w_{ij}^{\tilde{a}}) = (1-\alpha_h)\hat{b}(t_i^{\tilde{a}}, w_{ij}^{\tilde{a}})_- + \alpha_h \hat{\nabla} v,$$

where $\hat{b}(t_i^{\tilde{a}}, w_{ij}^{\tilde{a}})_-$ is an estimated slope of the value function in post-decision state $s^{\tilde{a}}$ obtained in the previous decision epoch. The step size α_h specifies how much weight we should consider for the estimated gradient $\hat{\nabla}v$ to update slope $b(t_i^a, w_{ij}^a)$ at iteration h. The step size α_h has an important effect on the convergence rate of the algorithm. We use the *generalized harmonic step size rule* to determine the value of α_h in each iteration of the algorithm (Powell, 2007, page 430). That is, $\alpha_h = \bar{\alpha}/(\bar{\alpha}+h-1)$ where $\bar{\alpha}$ is a parameter specifying the reduction rate of the step size to zero while the iteration counter h is increasing in the algorithm.

Finally, when a termination occurs in the section ($y_i^* = 1$), the value of the slope parameter $b(t_i^a, w_{ij}^a) = b(1, w_0)$ is updated in the algorithm. More precisely, using the structure of equation (4.17) and the value of the value function at termination time, we first estimate a sample of the value function at section level at the beginning time of the production cycle ($t_i = 1$) and next we use it to update the value of the slope parameter $b(1, w_0)$ in the current iteration of the algorithm.

Stopping criteria. The algorithm is stopped when the norm ($||.||$) of difference between two consecutive sets of parameter values (in an iteration of the outer loop) is less than or equal to a predefined error criterion λ, $||\vec{b}_h - \vec{b}_{h-1}|| \le \lambda$, or when we reach a maximum number of iterations $\mathscr{H}$.

The final values of the slope parameters, estimated at the end of the algorithm, can be used to find the best actions in the system, i.e. given the slope vector $\vec{b}$ including the final estimates of the slope parameters, the following maximization problem is solved to find the best action a^* for an arbitrary observed state s:

$$a^* = arg\ max_{a \in \mathbb{A}(s)}(r(s,a) + \tilde{v}_{\vec{b}}(s^a)). \tag{4.21}$$

where $s^a = \varphi_n(s,a)$. That is, in contrast to the usual solution procedures of MDPs, we do not generate a policy containing the optimal actions for every possible state. Instead, we solve the above maximization problem to find the best actions for the observed states in the system.

Table 4.1: Parameter values used in the discounted infinite-horizon MDP.

Parameter	Value	Description
$q^{\max}$	18	Maximum number of pigs in a pen.[a]
$t^{\max}$	15	Maximum number of weeks in a growing period (week).[a]
$t^{\min}$	9	First possible week of marketing decisions.[a]
$\lvert I\rvert$	3	Number of sections.[a]
$\lvert J\rvert$	20	Number of pens in a section.[a]
$k^{\texttt{truck}}$	205	Capacity of truck transferring the culled pigs to the abattoir (pig).[b]
$c^{\texttt{truck}}$	400	Fixed cost of the truck (DKK).[b]
$c^{\texttt{piglet}}$	375	Cost of buying a piglet (DKK).[c]
γ	0.95	Discount factor.
w_0	$(26.4, 3, 6.2)$	Initial weight information, $w_0 = (\mu_0, \sigma_0, g_0)$, at the start of a production cycle (kg).[d]

[a] Value based on discussions with experts in Danish pig production. [b] Value taken from information in DANISH CROWN (`http://www.danishcrown.dk/Ejer/Svineleverandoer/DC-Afregning/DC-Logistik-svin.aspx`). [c] Value taken from Kristensen et al. (2012). [d] Estimated based on time series generated using simulation.

Table 4.2: Range of centre points for continuous state variables μ_{ij}, σ_{ij}, and g_{ij}. Given centre points, the continuous state variables μ_{ij}, σ_{ij}, and g_{ij} are discretized into 21, 9, and 6 intervals, respectively. The unit of values in the table is kg.

Week (t)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
u_t^{μ}	6.4-46.4	13.4-53.4	20.4-60.4	27.4-67.4	34.4-74.4	41.4-81.4	48.4-87.4	55.4-95.4	62.4-102.4	69.4-109.4	76.4-115.4	83.4-123.4	90.4-130.4	97.4-137.4	104.4-144.4
u_t^{σ}	2-10	2.5-10.5	3-11	3.5-11.5	4-12	4.5-12.5	5-13	5.5-13.5	6-14	6.5-14.5	7-15	7.5-15.5	8-16	8.5-16.9	9-17
u_t^{g}	4.4-8.2	4.4-8.2	4.4-8.2	4.4-8.2	4.4-8.2	4.4-8.2	4.2-8.2	4.2-8.2	4.2-8.2	4.2-8.2	4.2-8.2	4.2-8.2	4.2-8.2	4.2-8.2	4.2-8.2

4.4 Computational results

In order to show the functionality of the proposed model, we use it in three numerical experiments. In Experiment 1, we apply the AVI algorithm at pen level and investigate the accuracy of this algorithm. Since the size of the model is small at pen level, we first solve it by the *value iteration* algorithm and compare the results with the AVI algorithm. Next, in Experiment 2, we give an example at herd level and show how the ADP can find the best marketing decisions when there are different conditions in the average growth of pigs in sections. Finally, in Experiment 3, the marketing policy obtained by the ADP is compared with other marketing policies that may be applied in the production unit.

4.4.1 Parameters

In order to initialize the ADP and use it for the marketing decisions, we need to specify the parameter values of the discounted infinite-horizon MDP model and the AVI algorithm.

The values of the parameters used in the discounted infinite-horizon MDP model are given in Table 4.1. These values have been obtained from information in finisher pig production units (Danish conditions) and related literature. In the MDP model, we also need to calculate functions/coefficients $c_k^{\texttt{cull}}(w_{ij})$ and $c_k^{\texttt{feed}}(w_{ij})$ used in the reward function of the model in (4.8). These functions depend on the carcass weight, leanness, feeding cost, and the settlement pork price of individual pigs in pens. The method of acquiring this information and the related parameter values are given in Appendix 4.B.

Moreover, since the framework of the model is based on a discrete Markov process, the possible values of the continuous state variables μ_{ij}, σ_{ij}, and g_{ij} related to weight information w_{ij} (defined in Section 4.2.1) must be split up into the discrete intervals. We use a simple method to discretize state variables μ_{ij}, σ_{ij}, and g_{ij} such that in every week of the growing period they are divided into 21, 9, and 6 intervals of equal length, respectively. That is, the size of set W including the weight information in the model is $|W| = 21 \times 9 \times 6 = 1,134$. This discretization of the state space splits the domain of each continuous state variable into intervals of sufficiently small length, e.g. by splitting the possible range of state variable μ_{ij} into 21 intervals per week, the length of each weight interval will be 2 kg. In a given week number t, the centre points of these intervals are denoted by u_t^{μ}, u_t^{σ}, and u_t^{g} and their ranges are specified in a way such that they represent possible values of weight and growth information in the system. That is, the range of the centre points for the average and standard deviation of weight, u_t^{μ} and u_t^{σ}, increases linearly since it is not possible to have, for instance, large values of μ_{ij} and σ_{ij} (e.g. 100 and 15, respectively) in the first weeks of production. Table 4.2 shows the possible values of these centre points during the growing period in pens.

The AVI algorithm was coded in C++ (MS VS2010 compiler), and CPLEX 12.6.2 (C++ API by Concert Tecnology) was used as the optimization solver to solve the maximization problem of the algorithm. In the algorithm, parameters $\mathscr{N}$, $\mathscr{H}$, λ, $\bar{\alpha}$ are set to 120, 300, 5, and 100, respectively. The best setting has been obtained by testing the algorithm with different values of these parameters. Note that under these parameter values, 36,000 decision epochs are simulated in the algorithm and in each simulation, slope parameters are updated according to the weight information in 60 pens, i.e. the maximization problem of the algorithm is solved 2,160,000 times to update the value of slope parameters in the model (at herd level). When the slope parameters have been estimated, they are used in the maximization problem in (4.21) to find the best actions for the observed states in the test instances.

The test instances used in the computational experiments are randomly generated from the stochastic process of weight information ($w_{ij} = (\mu_{ij}, \sigma_{ij}, g_{ij}),\ j \in J\ \&\ i \in I$) described in Appendix 4.A. More precisely, we simulate sample paths for state variables μ_{ij}, σ_{ij}, and g_{ij} during the growing period in the pens. These samples are obtained using the conditional probability distributions of variables (μ_{ij}, g_{ij}) and σ_{ij} acquired by the *state space models* formulated in Pourmoayed et al. (2016).

4.4.2 Experiment 1: Accuracy of the AVI algorithm

In this section, we investigate the accuracy of the AVI algorithm and compare it with the *value iteration* (VI) algorithm. For the purpose of comparison, we consider the sequential marketing decisions at pen level and first we use the VI algorithm to solve the discounted infinite-horizon MDP model described in Section 4.2. After solving the model, the value function is estimated for every possible state in the model. Figure 4.2 shows the behavior of the value function with respect to the number of remaining pigs in the pen (denoted by q). Each black line in this figure corresponds to different conditions in the pen that may happen according to week number t and weight information w. As we mentioned in Section 4.3.1, a good functional form for the value function at pen level, in state $s = (t, q, w)$, is $\nu(s) \approx qb(t, w)$ where the parameter $b(t, w)$ is approximately equal to the slope of the line related to week number $t \in \{1, ..., t^{\max}\}$ and weight information $w \in W$ in Figure 4.2. In order to compare the performance of the AVI and VI algorithms, we compare these slopes (resulting from the VI algorithm) with the estimated slopes obtained by the AVI algorithm. The AVI algorithm is implemented in a single section with one pen such that all slope parameters in the algorithm are initialized to one fixed number (e.g. 450). Based on the values of parameters $\mathscr{N}$, $\mathscr{H}$, λ, $\bar{\alpha}$ in the algorithm, the final values of the slope parameters are estimated after approx. 45 minutes.

Figure 4.3 illustrates how the slope parameters of the approximate value function (obtained by the AVI algorithm) converge to the constant values. Since the algorithm is based on simulating the states, some of the states are observed more often and hence the related slope parameters converge earlier. Figure 4.4 shows how the estimation of slope parameters by AVI is close to the related values of slope parameters obtained by the VI algorithm. In this figure, each point corresponds to a specific pair (t, w) such that the x-coordinate value of a point is the estimated value of slope $b(t, w)$ obtained by the AVI algorithm and the y-coordinate is the slope value estimated by the VI algorithm. The dashed line is a 45-degree line considered as a marker in the plot when the values of x and y coordinates are equal. As seen in the figure, the data points are close to the marker

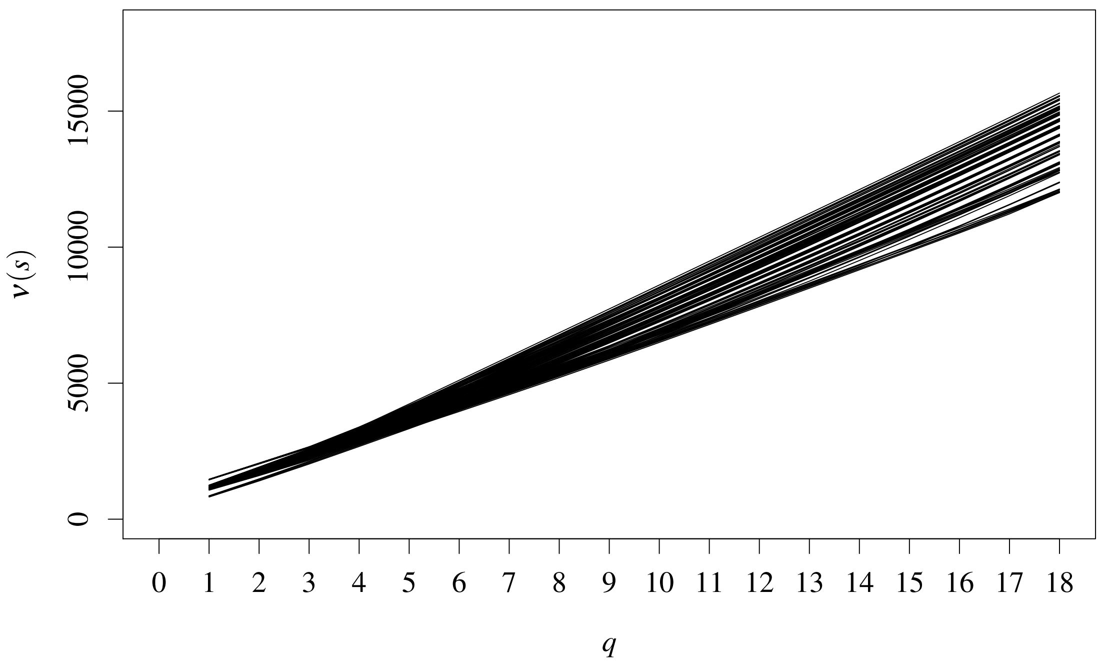


Figure 4.2: Behaviour of the value function with respect to the number of remaining pigs in the pen (q) resulting from solving the discounted infinite-horizon MDP model using the VI algorithm. Horizontal axis shows the number of remaining pigs in the pen and the vertical axis the value of value function $v(s)$ in state $s = (t, q, w)$. Each black line shows different conditions in the pen that may happen according to week number t and weight information w.

line indicating that there is a little difference between the slope values acquired by the AVI and VI algorithms. Over-estimation and under-estimation of slope parameters, estimated by the AVI algorithm, happens in the first and last weeks of the growing period when the pen is almost full and empty, respectively. The *root mean square error* (RSME) between the slopes obtained by the AVI and VI is 22.7 which is an acceptable error according to the range of the estimated values for the slope parameters (between 500 and 900). This shows that the AVI algorithm gives a reasonable estimation for the slope parameters of the approximate value function. Note that if the values of parameters $\mathscr{N}$ and $\mathscr{H}$ are increased in the algorithm, the estimation error will be lower. However, the parameters of the algorithm should be adjusted in such a way to give a good balance between the computational efficiency of the algorithm and the quality of the results. Notice that, when the algorithm is run at herd level, the quality of the results will be much better

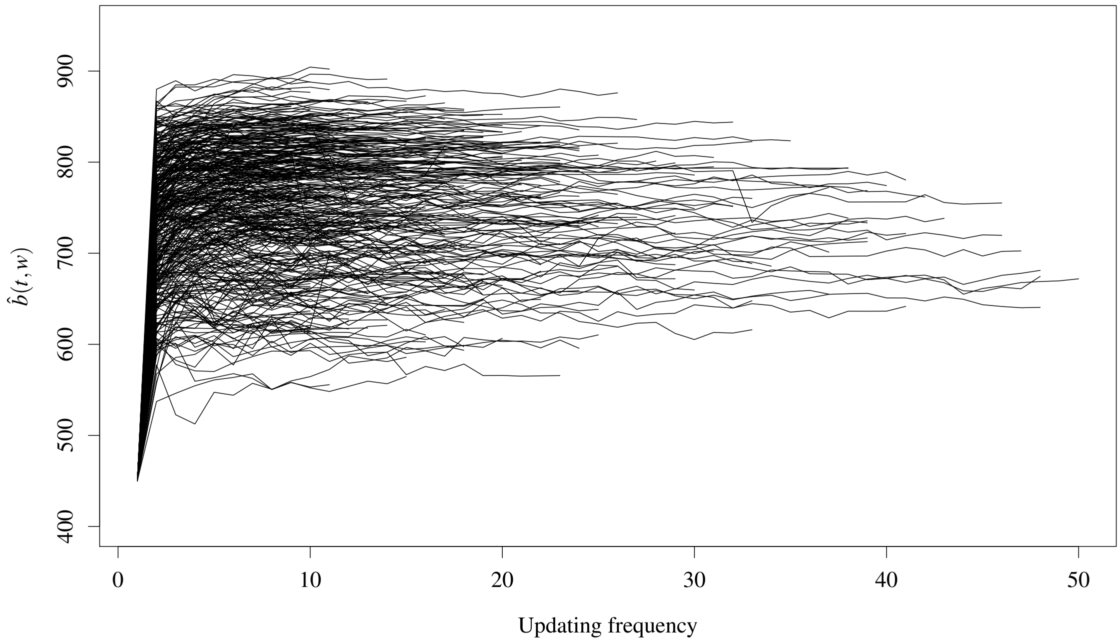


Figure 4.3: Converging trends of slope parameters in the AVI algorithm. The horizontal axis shows the number of updates of an observed slope and vertical axis indicates the estimated value of observed slope parameters. Black lines show the converging trend of observed slope parameters in the algorithm.

since the slope parameters are estimated using the simulated weight information in 60 pens at the herd instead of one pen, i.e. each loop of the algorithm for updating the slope parameters will contain 60 interior loops, each for one pen (see Algorithm 2).

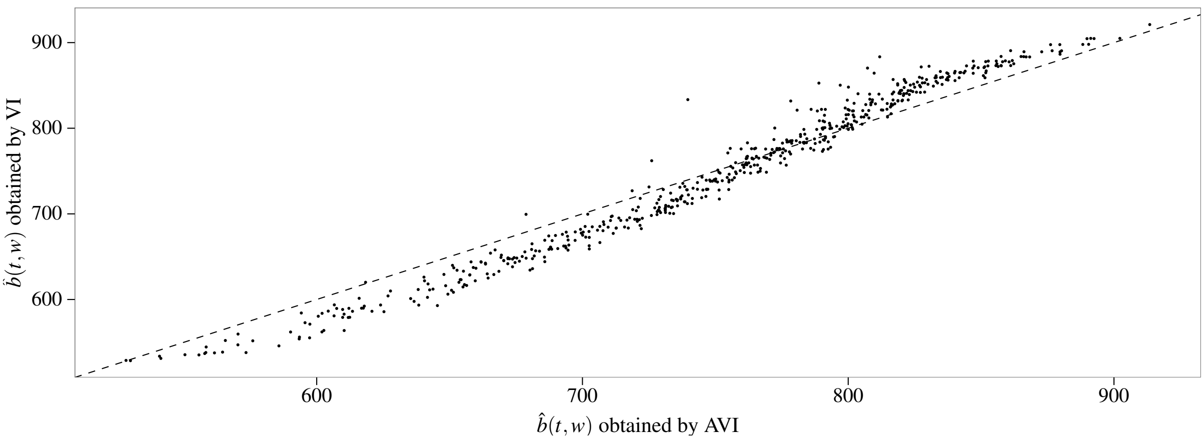


Figure 4.4: Comparison between the estimated slope values obtained by the AVI and VI algorithms. Each point corresponds to a specific pair (t, w). The x-coordinate value of each point is the estimated value of slope $b(t, w)$ obtained by the AVI algorithm, and the y-coordinate value is the slope value estimated by the VI algorithm. The dashed line is a 45-degree line considered as a marker in the plot when the values of the x and y coordinates are equal.

4.4.3 Experiment 2: Marketing decisions at herd level

In this section, we use the ADP to find the best marketing decisions in a herd with 3 sections. In order to show how the marketing decisions change according to the conditions of the herd, we assume that there are different environmental effects (e.g. temperature, housing conditions, and humidity) for each section resulting in different average growth rates of pigs in the three sections. We therefore set the average growth rate of pigs in Section 2 as normal (6 kg per week) while in Sections 1 and 3 pigs grow 10 percent faster and slower than Section 2, respectively.

To find the best marketing decisions, we first run the AVI algorithm under the parameters given in Table 4.1 and Section 4.4.1 to estimate the slope parameters in the approximation architecture of the value function at herd level. After running the algorithm, the slope parameters are converged after approximately 21 hours. These slope values are then used in the maximization problem (4.21) to find the best marketing decisions in the simulated states of the herd. The simulated states are dependent on the weight information of the pens which is randomly generated using the stochastic process of weight information described in Appendix 4.A.

Figure 4.5 shows the range of the simulated weight data during $t^{\max} = 15$ weeks of the growing period in the three sections. The distribution of weight values in each week is shown by a box plot where the horizontal line inside the box indicates the median of the simulated data. As seen in the figure, in each week (e.g. week 10), the weight values of pigs in Section 3 are higher

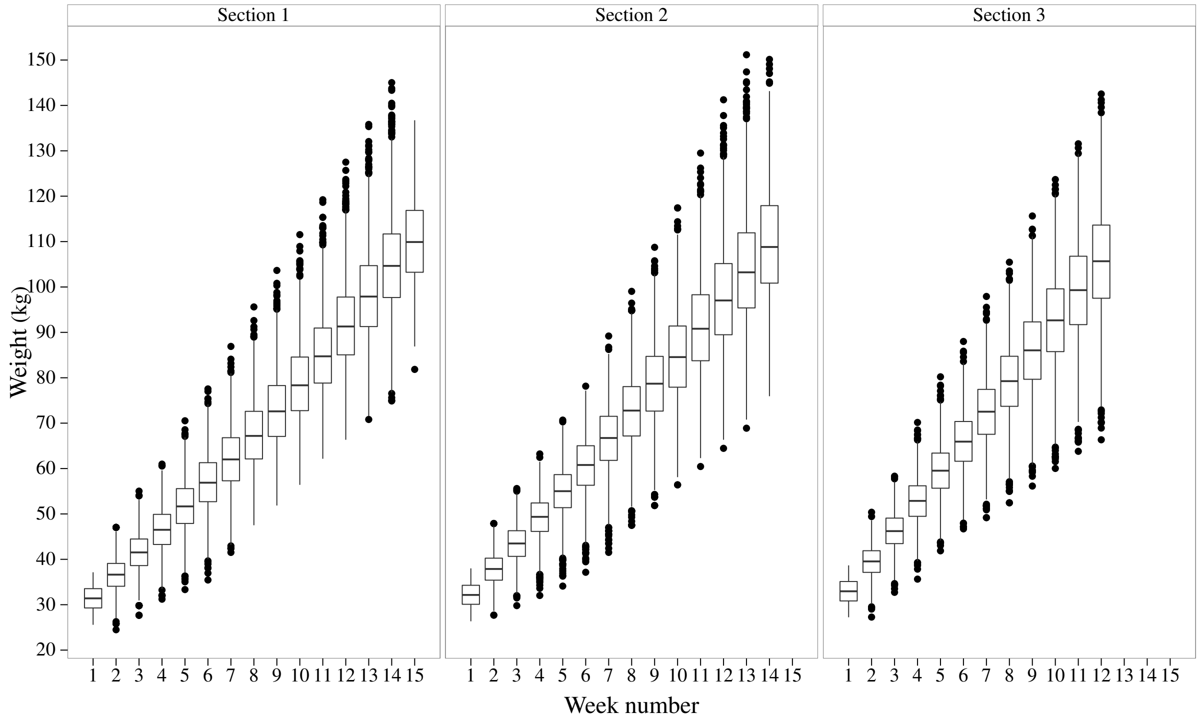


Figure 4.5: The range of the simulated weight data in the three sections during $t^{\max} = 15$ weeks of growing period in the production cycles. Each box-plot contains 360 weight values split into 4 quartiles. The body of the box includes 50 % of the data (second and third quartiles) and within the box the horizontal line shows the median of the simulated data. The up and down whisker lines show the range of the data in the first and last quartiles. If a data point diverges considerably from the overall pattern, it is plotted as a point.

than Sections 1 and 2 and the pigs in Section 2 grow faster compared to Section 1. Moreover, when the pigs become older, the inhomogeneity of weight distribution increases between the pigs with the same age in all sections (see the length of the box plots from week numbers 1 to 15).

Figure 4.6 illustrates the best marketing decisions during 52 weeks of production in the herd (from week 46 to 94). We have randomly selected this period to have a plot of marketing decisions in each section. That is, weeks 1 to 45 can be considered as the warm-up period of the simulation for reaching steady state. In this figure, the bar lines show the number of remaining pigs before a decision is made, the numbers below the bars denote the number of heaviest pigs culled from the section, the letter "T" indicates the termination of a production cycle in the section, and the letter "C" corresponds to continuing the production process without marketing. Note that when termination occurs in a section (letter "T"), a new production cycle is started in this section. Moreover, the time period between two successive terminations shows the length of a production

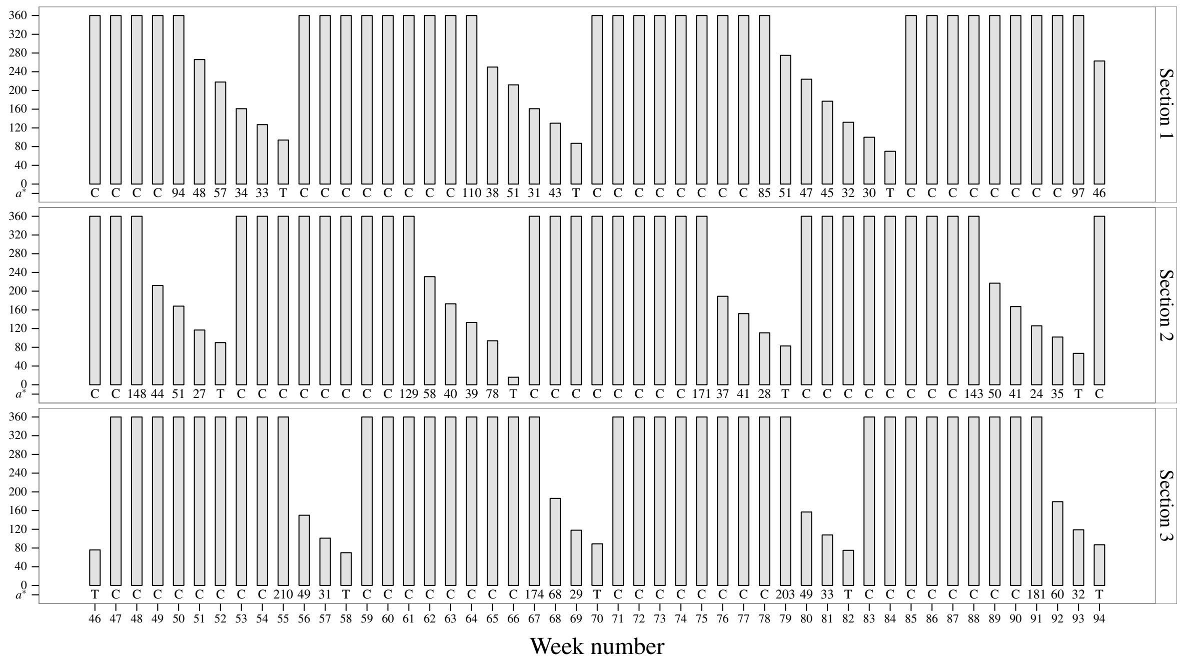


Figure 4.6: Results of marketing decisions in a herd with three sections during 52 weeks of production (from week 46 to 94). Bar lines show the number of remaining pigs in the sections before a decision is made. Numbers below the bars denote the number of heaviest pigs culled from the section, the letter “T” indicates the termination of a production cycle in the section, and the letter “C” corresponds to continuing the production process without marketing.

cycle in the section.

As seen in Figure 4.6, the length of the production cycles in Section 3 is shorter than other sections since according to Figure 4.5 the average weight of pigs in this section is higher than Sections 1 and 2 and hence the pigs obtain their optimal slaughter weight earlier. It is therefore beneficial that after a few weeks of individual marketing in Section 3, an early termination occurs in this section resulting in a higher number of production cycles in this section compared to other sections. In Sections 1 and 2, the pigs grow slower than Section 3 and hence it is better to keep them in the pens for a longer period compared to Section 3, i.e. the length of the production cycles in Sections 1 and 2 is longer than Section 3. However, after starting the marketing decisions in Sections 1 and 2, we observe that in each production cycle the fraction of pigs remaining in Section 1 is more than Section 2 (see the height of the bar lines for Sections 1 and 2 in Figure 4.6). This happens since the average growth of pigs in Section 2 is better than Section 1 and hence more pigs are culled from Section 2 compared to Section 1 resulting in a lower number of pigs in Section 2.

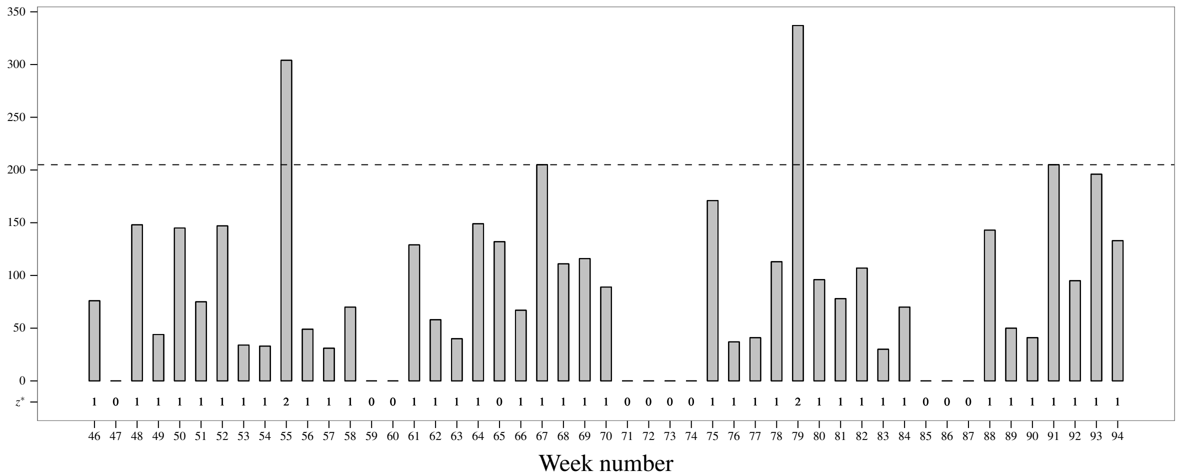


Figure 4.7: The total number of culled pigs from the herd in weeks 46 to 94. The numbers below the bars show the number of trucks needed to transfer the culled pigs to the abattoir. The horizontal dashed line shows the full capacity of a truck.

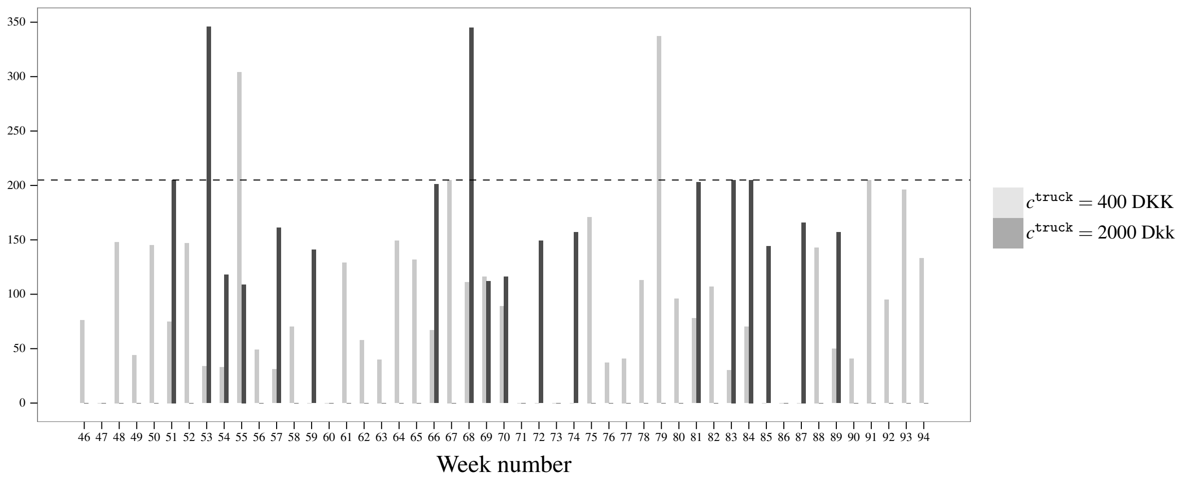


Figure 4.8: The effect of the transportation cost on the marketing policy of the herd. Black and gray bars show the number of pigs culled from the herd when the fixed costs of transportation are 400 DKK and 2000 DKK, respectively. The horizontal dashed line shows the full capacity of a truck (205 pigs).

Figure 4.7 illustrates the total number of culled pigs in the herd from weeks 46 to 94. The numbers below the bars show the number of trucks needed to transport the culled pigs to the abattoir and the horizontal dashed line shows the full capacity of a truck . As can be seen, the full capacity of the trucks is not used in most cases and only the pigs with appropriate live weight are sent to the abattoir (the capacity of a full truck is $k^{\mathtt{truck}} = 205$). This happens since the fixed cost

Table 4.3: Comparison between the performance of different marketing polices. Since each policy is applied in 100 sample paths, the results are reported as 95% confidence interval $m \pm 1.96s/\sqrt{100}$, where m and s are the mean and standard deviation of the values of comparison criteria in 100 sample paths.

Policy	ADP	M	FTC	All-In All-Out						
				9	10	11	12	13	14	15
Av. discounted reward per week	2072 ± 45.5	1078 ± 11.1	1830 ± 103	470 ± 11.9	1398 ± 10.3	1831 ± 7.6	1881 ± 6.7	1679 ± 4.6	1365 ± 4	1014 ± 3.9
Av. length of production cycle	11.8 ± 0.01	10 ± 0	11.6 ± 0.06	9 ± 0	10 ± 0	11 ± 0	12 ± 0	13 ± 0	14 ± 0	15 ± 0
Number of production cycles	30.5 ± 0.14	36 ± 0	31.3 ± 0.28	39 ± 0	36 ± 0	33 ± 0	30 ± 0	27 ± 0	27 ± 0	24 ± 0
% of truckload capacity utilized	76 ± 1	75 ± 0	88 ± 0	87 ± 0	87 ± 0	87 ± 0	87 ± 0	87 ± 0	87 ± 0	87 ± 0
Number of trucks sent to abattoir	73.7 ± 1	84 ± 0	63.6 ± 0.5	84 ± 0	72 ± 0	66 ± 0	60 ± 0	54 ± 0	54 ± 0	48 ± 0

of calling a truck (see Table 4.1) is lower than the profit obtained by culling a specific number of pigs that are ready for slaughter. For instance, when 40 pigs are ready for slaughter, it is beneficial to call a truck and transport them to the abattoir. However, this happens in few cases and there are not many deliveries with a small number of culled pigs. Moreover, note that in some cases a compulsory termination at the latest week of the growing period (week $t^{\max} = 15$) leads to a small number of pigs being culled and sent to the abattoir by one truck (e.g. see week 84 in Figure 4.6).

Transportation costs can affect the marketing policy of the farm. Figure 4.8 shows the change in the number of marketed pigs in the herd when the fixed costs of transportation ($c^{\texttt{truck}}$) are 400 and 2,000 DKK. As we see in this figure, a high transportation cost leads to a better utilization of trucks and fewer deliveries to the abattoir. More precisely, with the fixed transportation cost $c^{\texttt{truck}} = 400$, the number of deliveries to the abattoir is 39 while with $c^{\texttt{truck}} = 2,000$ this number is decreased to 18 that may result in culling some pigs that are not ready for slaughter yet. Note that the variable cost of transportation per pig (e.g. cost of loading a pig into the truck) is not considered in the model since this variable cost is much smaller than the value of a pig in the abattoir and hence it would not have a noticeable impact on the marketing policy and can be ignored.

4.4.4 Experiment 3: ADP compared to other marketing policies

In order to evaluate the performance of the marketing policy obtained by the ADP, we compare it with other well-known marketing polices often applied at herd level. In order to have a valid comparison, we use same test instances for all the policies in which the simulated weight information in the pens shows normal conditions at the herd. Since we cannot simulate an infinite trajectory, marketing decisions are considered over a long time horizon ($\mathscr{N} = 120$ weeks approximately equal to 2.5 years) and 100 sample paths are generated in this horizon for comparing the policies. The *average discounted reward of the production unit per week* is chosen as the main criterion to compare the policies. This criterion is defined as $\sum_{n=1}^{\mathcal{N}} \gamma^{n-1} r_n(s,\pi(s))/\mathcal{N}$ for a given policy π. Moreover, the average length of the production cycle, the number of production cycles, the percentage of truckload capacity utilized, and the number of trucks sent to the abattoir during this horizon are reported for each policy. We compare the marketing policy resulting from ADP with the following policies:

Myopic policy *(M)*. This policy does not consider the impact of the present decisions on the future conditions of the system, i.e. only the immediate reward of the decisions is taken into account in the model. To find the best marketing decisions under this policy, in each decision epoch, we solve the following maximization problem without considering the future reward of decisions:

$$a^* = arg\ max_{a \in \mathbb{A}(s)}(r(s,a)).$$

This policy was chosen for comparison in order to stress the importance of using dynamic programming that considers the future value of decisions.

All-In All-Out policy. When the length of the growing period in the section equals a specified value, all pigs are marketed from the section in one delivery, i.e. individual marketing decisions are not considered in the model. We evaluate this policy under the assumptions that the section must be terminated at weeks 9, 10, 11, 12, 13, 14, and 15. This policy is commonly used in the industry as the main delivery strategy to the abattoir.

Full truck capacity policy *(FTC)*. When a delivery to the abattoir takes place, the farmer prefers to use the maximum capacity of the trucks. In order to evaluate this policy, we change the inequality transportation constraint (4.4) in the model to an equality constraint, and assume that a compulsory termination at herd level occurs when the number of pigs at the herd is less than the full capacity of one truck ($k^{\texttt{truck}}$). This policy is usually followed by farmers who want to use the maximum capacity of the truck when sending the pigs to the abattoir.

Table 4.3 shows the results for the ADP, M, FTC, and All-In All-Out policies. Since each policy is used in 100 sample paths, results are reported as 95% confidence interval $m \pm 1.96s/\sqrt{100}$, where m and s are the mean and standard deviation of the values of comparison criteria in 100 sample paths.

As we see in the table, the ADP policy outperforms other policies in terms of average discounted reward per week, and the All-In All-Out policy with a length of 9 weeks results in the lowest reward. The difference between the ADP and FTC policies is noticeable. Using the maximum capacity of trucks in the FTC policy results in marketing some pigs that are not ready for slaughter yet and hence the farmer loses the possible profit that could be earned by keeping these pigs for a longer time in the herd. Moreover in this policy, as expected, the utilization of truckload capacity is better than other policies (see the percentage of truckload capacity utilized). It seems that All-In All-Out policies with lengths 11 and 12 weeks perform well in the test instances. This is because of the lengths of the growing period in these policies that are close to the average length of the production cycle in the ADP policy resulting in the best marketing time.

In the All-In All-Out policies, as expected, when the length of the growing period is increasing (from 9 to 15 weeks), the number of production cycles decreases during 120 weeks of production and hence we need to have fewer trucks to transfer the culled pigs to the abattoir. Furthermore in these policies utilization of truckload capacity is high and close to the FTC policy, since in one delivery all the pigs are sent to the abattoir and hence the full capacity of most trucks are utilized. Finally note that when the length of the production cycle is high (e.g. 14 or 15 weeks), the average discounted reward per week is reduced noticeably. This happens since the feeding cost of the pigs increases considerably during a longer growing period and the revenue that can be obtained by selling the heavy pigs in this period cannot compensate well the feeding cost of the system and therefore the average reward of the production unit will be reduced. This shows the importance of the feeding costs in the herd.

4.5 Conclusions

In the production of growing/finishing pigs, cross-level constraints (e.g. termination at section level or transportation at herd level) resulting from marketing decisions at different levels (e.g. animal, pen, section, and herd level) have an effect on the marketing policy of the farm.

In this paper, we considered marketing decisions of growing/finishing at herd level modeled by a discounted infinite-horizon MDP model. The state of the system was based on weight information in the pens described by a stochastic process relying on state space models. Due to the curse of dimensionality problem, an ADP approach with post-decision states was used to find the best marketing policy.

We suggested a parametric function to approximate the value function of the MDP model and estimated the slope parameters of the approximation architecture using the AVI algorithm. A numerical experiment at pen level demonstrated the accuracy of the AVI algorithm in estimating the slope parameters of the approximated value function.

Numerical experiments at herd level showed that the ADP can find the best marketing decisions under different conditions in the average growth rate of pigs in the herd. Moreover, the effect of the fixed transportation cost was noticeable on the marketing decisions. Finally, the marketing policy obtained by the ADP outperformed other marketing polices in terms of average discounted reward per week and the effect of feeding cost on the discounted reward of the production unit was important.

Acknowledgment

The author would like to thank Associate Professor Lars Relund Nielsen, and Professors Anders Ringgaard Kristensen and Christian Larsen for useful comments on an earlier draft of this paper. This article has been written with support from The Danish Council for Strategic Research (The PigIT project, Grant number 11-116191).

References

S. Andersen, B. Pedersen, and M. Ogannisian. Slagtesvindets sammensætning. meddelelse 429. Technical report, Landsudvalget for Svin og Danske Slagterier, 1999. URL `http://vsp.lf.dk/Publikationer/Kilder/lu_medd/medd/429.aspx`.

Y. Ben-Ari and Sh. Gal. Optimal replacement policy for multicomponent systems: An application to a dairy herd. *European Journal of Operational Research*, 23(2):213–221, 1986. doi: 10.1016/0377-2217(86)90240-7 .

K.A. Boys, N. Li, P.V. Preckel, A.P. Schinckel, and K.A. Foster. Economic replacement of a heterogeneous herd. *American Journal of Agricultural Economics*, 89(1):24–35, 2007. doi:10.1111/j.1467-8276.2007.00960.x .

D.P. de Farias and B. Van Roy. The linear programming approach to approximate dynamic programming. *Operations Research*, 51(6):850–865, 2003. doi:10.1287/opre.51.6.850.24925 .

Sh. Gal. The parameter iteration method in dynamic programming. *Management Science*, 35(6): 675–684, 1989. doi:10.1287/mnsc.35.6.675 .

J.J. Glen. A dynamic programming model for pig production. *Journal of the Operational Research Society*, 34:511–519, 1983. doi:10.1057/jors.1983.118 .

E. Jørgensen. The influence of weighing precision on delivery decisions in slaughter pig production. *Acta Agriculturae Scandinavica, Section A - Animal Science*, 43(3):181–189, August 1993. doi:10.1080/09064709309410163 .

E Jørgensen. Foderforbrug pr kg tilvækst hos slagtesvin. fordeling mellem forbrug til vedligehold og til produktion i besætninger under den rullende afprøvning. Technical report, Danish Institute of Agricultural Sciences, Biometry Research Unit, 2003.

A. R. Kristensen. Optimal replacement in the dairy herd: A multi-component system. *Agricultural Systems*, 39(1):1–24, 1992. doi:10.1016/0377-2217(86)90240-7 .

A.R. Kristensen, L. Nielsen, and M.S. Nielsen. Optimal slaughter pig marketing with emphasis on information from on-line live weight assessment. *Livestock Science*, 145(1-3):95–108, May 2012. doi:10.1016/j.livsci.2012.01.003 .

H. Kure. *Marketing Management Support in Slaughter Pig Production*. PhD thesis, The Royal Veterinary and Agricultural University, 1997. URL `http://www.prodstyr.ihh.kvl.dk/pub/phd/kure_thesis.pdf`.

J.K. Niemi. *A dynamic programming model for optimising feeding and slaughter decisions regarding fattening pigs | NIEMI | Agricultural and Food Science*. PhD thesis, MTT Agrifood research, 2006. URL `http://ojs.tsv.fi/index.php/AFS/article/view/5855`.

J.W. Ohlmann and P.C. Jones. An integer programming model for optimal pork marketing. *Annals of Operations Research*, 190(1):271–287, November 2008. doi:10.1007/s10479-008-0466-3 .

J. Patrick, M.L. Puterman, and M. Queyranne. Dynamic multipriority patient scheduling for a diagnostic resource. *Operations research*, 56(6):1507–1525, 2008. doi:10.1287/opre.1080.0590 .

Lluís M Plà-Aragonés, S Rodriguez-Sanchez, and Victoria Rebillas-Loredo. A mixed integer linear programming model for optimal delivery of fattened pigs to the abattoir. *J Appl Oper Res*, 5:164–175, 2013.

R. Pourmoayed and L.R. Nielsen. An overview over pig production of fattening pigs with a focus on possible decisions in the production chain. Technical Report PigIT Report No. 4, Aarhus University, 2014. URL `http://pigit.ku.dk/publications/PigIT-Report4.pdf`.

R. Pourmoayed, L. R. Nielsen, and A. R. Kristensen. A hierarchical Markov decision process modeling feeding and marketing decisions of growing pigs. *European Journal of Operational Research*, 250(3):925–938, 2016. doi:10.1016/j.ejor.2015.09.038 .

W.B. Powell. *Approximate Dynamic Programming: Solving the curses of dimensionality*, volume 703. Wiley-Interscience, 2007. ISBN 978-0-470-60445-8.

W.B. Powell. Merging AI and OR to solve high-dimensional stochastic optimization problems using approximate dynamic programming. *INFORMS Journal on Computing*, 22(1):2–17, 2010. doi:10.1287/ijoc.1090.0349 .

M. L. Puterman. *Markov decision processes: discrete stochastic dynamic programming*. John Wiley & Sons, 2005. ISBN 978-0-471-72782-8.

H.C. Tijms. *A first course in stochastic models*. John Wiley & Sons Ltd, 2003. ISBN 978-0-471-49880-3.

N. Toft, A.R. Kristensen, and E. Jørgensen. A framework for decision support related to infectious diseases in slaughter pig fattening units. *Agricultural Systems*, 85(2):120–137, 2005. doi:10.1016/j.agsy.2004.07.017 .

H. Topaloglu and W.B. Powell. Dynamic-programming approximations for stochastic time-staged integer multicommodity-flow problems. *INFORMS Journal on Computing*, 18(1):31–42, 2006.

doi:10.1287/ijoc.1040.0079 .

A. Toriello, G. Nemhauser, and M. Savelsbergh. Decomposing inventory routing problems with approximate value functions. *Naval Research Logistics (NRL)*, 57(8):718–727, 2010. doi:10.1002/nav.20433 .

4.A Modeling stochastic process of weight information in the pen

Suppose (μ, g) and σ denote the average weight and growth, and the standard deviation of the weight of pigs, respectively, in the pen. In order to describe the stochastic nature of the weight information in the pen, we apply a *discrete time stochastic process* modeling the dynamics of independent random variables (μ, g) and σ during the growing period ($t^{\max}$ weeks).

Suppose $\Omega_{(\mu,g)}$ and Ω_σ are the sets of all possible outcomes for the random variables (μ, g) and σ. Moreover, assume that $X = ((\mu, g), \sigma)$ is a supplementary random variable denoting the random information of weight in the pen. A realization (sample) of this random variable is denoted by $\hat{X} \in \Omega_{(\mu,g)} \times \Omega_\sigma$.

A *discrete time stochastic process* modeling weight information in the pen is defined as a collection of random variables X indexed by week number t:

$$\{X_t, t = 1, 2, ..., t^{\max}\}.$$

In this process, in a given week number t, the random variable X_t depends on the earlier values observed in the process, $\hat{X}_{t-1}, \hat{X}_{t-2}, .., \hat{X}_1$. Therefore, in order to analyze the process, we need to find the probability distribution of the conditional random variables in the form of

$$\Pr(X_t | \hat{X}_{t-1}, \hat{X}_{t-2}, .., \hat{X}_1).$$

In order to calculate this probability, we use the properties of two *state space models* (SSMs) formulated in Pourmoayed et al. (2016) for modeling the dynamics of the state variables (μ_t, g_t) and σ_t in a finisher pen.

First, according to the Markovian property of these SSMs, we have

$$\Pr(X_t | \hat{X}_{t-1}, \hat{X}_{t-2}, .., \hat{X}_1) = \Pr(X_t | \hat{X}_{t-1})$$

and based on the properties of the independent random variables (μ_t, g_t) and σ_t in X_t we can conclude

$$\Pr(X_t | \hat{X}_{t-1}) = \Pr(\ (\mu_t, g_t) \mid (\hat{\mu}_{t-1}, \hat{g}_{t-1})\ ). \Pr(\sigma_t | \hat{\sigma}_{t-1}).$$

Now we only need to find the probability distributions of the conditional random variables $(\ (\mu_t, g_t) \mid (\hat{\mu}_{t-1}, \hat{g}_{t-1})\ )$ and $(\sigma_t | \hat{\sigma}_{t-1})$. The probability distributions of these random variables can be found in Theorems 2 and 4 in Pourmoayed et al. (2016).

Therefore, in the discounted infinite-horizon MDP model, the random information ω in the stochastic transition function $\Gamma(.)$ (defined in Section 4.2.4) is described using the probability distributions of the conditional random variables $(\ (\mu_t, g_t) \mid (\hat{\mu}_{t-1}, \hat{g}_{t-1})\ )$ and $(\sigma_t|\hat{\sigma}_{t-1})$. Moreover, in order to generate sample $\hat{\omega}$ in the AVI algorithm (see Algorithm 1), we need to draw a sample from the probability distributions of these random variables.

4.B Calculation of $c_k^{\texttt{cull}}(w_{ij})$ and $c_k^{\texttt{feed}}(w_{ij})$

Before calculating $c_k^{\texttt{cull}}(w_{ij})$ and $c_k^{\texttt{feed}}(w_{ij})$, we need to describe the methods of obtaining the carcass weight, leanness, feeding cost, and the settlement pork price function.

Carcass weight, leanness, feeding cost, and settlement pork price

Consider the kth ordered pig with live weight w^l and daily growth g in the pen. The carcass weight $\tilde{w}$ can be approximated as (Andersen et al., 1999)

$$\tilde{w} = 0.84w^l - 5.89 + e_{\texttt{c}}, \tag{4.22}$$

where $e_{\texttt{c}} \sim N(0, 1.96)$ is a normal distributed term. The relation between growth rate, leanness (lean meat percentage), and feed conversion ratio varies widely between herds. Hence, these formulas must be herd specific. The leanness $\breve{w}$ can be found as (Kristensen et al., 2012)

$$\breve{w} = \frac{-30(g-6)}{4} + 61. \tag{4.23}$$

The feed intake (energy intake) is modeled as the sum of feed for maintenance and feed for growth. The basic relation between daily feed intake f (FEsv[1]), live weight and daily gain is (Jørgensen, 2003)

$$f = k_1 g + k_2 w^{0.75}, \tag{4.24}$$

where $k_1 = 1.549$ and $k_2 = 0.044$ are constants describing the use of feed per kg gain and per kg metabolic weight, respectively. As a result the expected feed intake of the kth pig over the next $\tilde{d}$

[1]FEsv is the energy unit used for feeding the pigs in Denmark. One FEsv is equivalent to 7.72 MJ.

days equals

$$f_{(k)}^{\texttt{feed}}(\tilde{d}) = \mathbb{E}\left(\sum_{d=1}^{\tilde{d}} f_d\right) = \mathbb{E}\left(\sum_{d=1}^{\tilde{d}} \left(k_1 g + k_2(w^l + (d-1)g)^{0.75}\right)\right) = \mathbb{E}\left(\tilde{d}k_1 g + k_2 \sum_{d=1}^{\tilde{d}} (w^l + (d-1)g)^{0.75}\right),$$

where f_d denotes the feed intake at day d calculated recursively using (4.24). The feeding cost for the kth pig during $\tilde{d}$ days can be calculated by multiplying $f_{(k)}^{\texttt{feed}}(\tilde{d})$ to the unit feed cost per FEsv (1.8 DKK[2]):

$$p_{(k)}^{\texttt{feed}}(\tilde{d}) = 1.8 f_{(k)}^{\texttt{feed}}(\tilde{d}). \tag{4.25}$$

Consider the kth ordered pig with carcass weight $\tilde{w}$ and leanness $\breve{w}$ at delivery. Under the Danish system, the settlement pork price is the sum of two linear piecewise functions related to the price of the carcass and a bonus of the leanness:

$$p_{(k)}^{\texttt{pork}}(\tilde{w}, \breve{w}) = \tilde{p}(\tilde{w}) + \breve{p}(\breve{w}). \tag{4.26}$$

Functions $\tilde{p}(\tilde{w})$ and $\breve{p}(\breve{w})$ correspond to the unit price of carcass and the bonus of leanness for 1 kg meat, respectively. We define $\tilde{p}(\tilde{w})$ and $\breve{p}(\breve{w})$ based on the meat prices used in (Kristensen et al., 2012) [3] as

$$\tilde{p}(\tilde{w}) = \begin{cases} 0 & \tilde{w} < 50 \\ 0.2\tilde{w} - 2.7 & 50 \le \tilde{w} < 60 \\ 0.1\tilde{w} + 3.3 & 60 \le \tilde{w} < 70 \\ 10.3 & 70 \le \tilde{w} < 86 \\ -0.1\tilde{w} + 18.9 & 86 \le \tilde{w} < 95 \\ 9.3 & 95 \le \tilde{w} < 100 \\ 9.1 & \tilde{w} \ge 100. \end{cases}$$

$$\breve{p}(\breve{w}) = 0.1(\breve{w} - 61).$$

[2]Current feed price can be found at `http://www.notering.dk/WebFrontend/`.

[3]For the current prices see `http://www.danishcrown.dk/Ejer/Noteringer/Aktuel-svinenotering.aspx`

Note that when marketing decisions are made, the culled pigs are sent to the abattoir after $b = 3$ days. The additional feeding cost and reward, resulting from keeping the culled pigs in this period ($b = 3$ days), can be calculated using equations (4.25) and (4.26), respectively.

Calculation of reward and feed cost

The calculations of $c_k^{\texttt{cull}}(w_{ij})$ and $c_k^{\texttt{feed}}(w_{ij})$ are rather complex due to the ordered random variables and the non-continuous function $\tilde{p}(\tilde{w})$. However, these values can be calculated using simulation with a simple sorting procedure as described below.

In pen j of section i given the weight information $w_{ij} = (\mu_{ij}, \sigma_{ij}, g_{ij})$, do the following steps:

Step 0 Use the value of state variables μ_{ij} and σ_{ij} to find the probability distribution of live weight in the pen. i.e. $w_{ij}^l \sim N(\hat{\mu}_{ij}, \hat{\sigma}_{ij})$.

Step 1 Draw $q^{\max}$ random weights $\hat{w}^l$ from the probability distribution $w_{ij}^l \sim N(\hat{\mu}_{ij}, \hat{g}_{ij})$. Moreover, find the daily growth of the pen using the value of state variable g_{ij}, i.e. $g = \hat{g}_{ij}/7$

Step 2 For each week t and each random weight $\hat{w}^l$, use (4.22) and (4.23) to find the carcass weight and leanness (b days ahead), respectively. Moreover, use (4.25) to find the feeding cost for the next $d = 7$ and $b = 3$ days.

Step 3 For each week t and each random weight $\hat{w}^l$, calculate the settlement pork price (4.26) and deduct the feeding cost for b days from it, i.e. the reward of selling a pig with weight $\hat{w}^l$ to the abattoir is calculated.

Step 4 For each week t, sort the obtained values of feed cost and reward of selling in non-decreasing order of weights $\hat{w}^l$.

We run the simulation 1,000 times and calculate the average values for the feed cost and the reward of selling to the abattoir for the sorted weights, i.e. $c_k^{\texttt{cull}}(w_{ij})$ and $c_k^{\texttt{feed}}(w_{ij})$ are calculated.

DEPARTMENT OF ECONOMICS AND BUSINESS ECONOMICS

AARHUS UNIVERSITY
SCHOOL OF BUSINESS AND SOCIAL SCIENCES
www.econ.au.dk

PhD dissertations since 1 July 2011

2011-4 Anders Bredahl Kock: Forecasting and Oracle Efficient Econometrics

2011-5 Christian Bach: The Game of Risk

2011-6 Stefan Holst Bache: Quantile Regression: Three Econometric Studies

2011:12 Bisheng Du: Essays on Advance Demand Information, Prioritization and Real Options in Inventory Management

2011:13 Christian Gormsen Schmidt: Exploring the Barriers to Globalization

2011:16 Dewi Fitriasari: Analyses of Social and Environmental Reporting as a Practice of Accountability to Stakeholders

2011:22 Sanne Hiller: Essays on International Trade and Migration: Firm Behavior, Networks and Barriers to Trade

2012-1 Johannes Tang Kristensen: From Determinants of Low Birthweight to Factor-Based Macroeconomic Forecasting

2012-2 Karina Hjortshøj Kjeldsen: Routing and Scheduling in Liner Shipping

2012-3 Soheil Abginehchi: Essays on Inventory Control in Presence of Multiple Sourcing

2012-4 Zhenjiang Qin: Essays on Heterogeneous Beliefs, Public Information, and Asset Pricing

2012-5 Lasse Frisgaard Gunnersen: Income Redistribution Policies

2012-6 Miriam Wüst: Essays on early investments in child health

2012-7 Yukai Yang: Modelling Nonlinear Vector Economic Time Series

2012-8 Lene Kjærsgaard: Empirical Essays of Active Labor Market Policy on Employment

2012-9 Henrik Nørholm: Structured Retail Products and Return Predictability

2012-10 Signe Frederiksen: Empirical Essays on Placements in Outside Home Care

2012-11 Mateusz P. Dziubinski: Essays on Financial Econometrics and Derivatives Pricing

2012-12 Jens Riis Andersen: Option Games under Incomplete Information

2012-13 Margit Malmmose: The Role of Management Accounting in New Public Management Reforms: Implications in a Socio-Political Health Care Context

2012-14 Laurent Callot: Large Panels and High-dimensional VAR

2012-15 Christian Rix-Nielsen: Strategic Investment

2013-1 Kenneth Lykke Sørensen: Essays on Wage Determination

2013-2 Tue Rauff Lind Christensen: Network Design Problems with Piecewise Linear Cost Functions

2013-3 Dominyka Sakalauskaite: A Challenge for Experts: Auditors, Forensic Specialists and the Detection of Fraud

2013-4 Rune Bysted: Essays on Innovative Work Behavior

2013-5 Mikkel Nørlem Hermansen: Longer Human Lifespan and the Retirement Decision

2013-6 Jannie H.G. Kristoffersen: Empirical Essays on Economics of Education

2013-7 Mark Strøm Kristoffersen: Essays on Economic Policies over the Business Cycle

2013-8 Philipp Meinen: Essays on Firms in International Trade

2013-9 Cédric Gorinas: Essays on Marginalization and Integration of Immigrants and Young Criminals – A Labour Economics Perspective

2013-10 Ina Charlotte Jäkel: Product Quality, Trade Policy, and Voter Preferences: Essays on International Trade

2013-11 Anna Gerstrøm: World Disruption - How Bankers Reconstruct the Financial Crisis: Essays on Interpretation

2013-12 Paola Andrea Barrientos Quiroga: Essays on Development Economics

2013-13 Peter Bodnar: Essays on Warehouse Operations

2013-14 Rune Vammen Lesner: Essays on Determinants of Inequality

2013-15 Peter Arendorf Bache: Firms and International Trade

2013-16 Anders Laugesen: On Complementarities, Heterogeneous Firms, and International Trade

2013-17 Anders Bruun Jonassen: Regression Discontinuity Analyses of the Disincentive Effects of Increasing Social Assistance

2014-1 David Sloth Pedersen: A Journey into the Dark Arts of Quantitative Finance

2014-2 Martin Schultz-Nielsen: Optimal Corporate Investments and Capital Structure

2014-3 Lukas Bach: Routing and Scheduling Problems - Optimization using Exact and Heuristic Methods

2014-4 Tanja Groth: Regulatory impacts in relation to a renewable fuel CHP technology: A financial and socioeconomic analysis

2014-5 Niels Strange Hansen: Forecasting Based on Unobserved Variables

2014-6 Ritwik Banerjee: Economics of Misbehavior

2014-7 Christina Annette Gravert: Giving and Taking – Essays in Experimental Economics

2014-8 Astrid Hanghøj: Papers in purchasing and supply management: A capability-based perspective

2014-9 Nima Nonejad: Essays in Applied Bayesian Particle and Markov Chain Monte Carlo Techniques in Time Series Econometrics

2014-10 Tine L. Mundbjerg Eriksen: Essays on Bullying: an Economist's Perspective

2014-11 Sashka Dimova: Essays on Job Search Assistance

2014-12 Rasmus Tangsgaard Varneskov: Econometric Analysis of Volatility in Financial Additive Noise Models

2015-1 Anne Floor Brix: Estimation of Continuous Time Models Driven by Lévy Processes

2015-2 Kasper Vinther Olesen: Realizing Conditional Distributions and Coherence Across Financial Asset Classes

2015-3 Manuel Sebastian Lukas: Estimation and Model Specification for Econometric Forecasting

2015-4 Sofie Theilade Nyland Brodersen: Essays on Job Search Assistance and Labor Market Outcomes

2015-5 Jesper Nydam Wulff: Empirical Research in Foreign Market Entry Mode

2015-6	Sanni Nørgaard Breining: The Sibling Relationship Dynamics and Spillovers
2015-7	Marie Herly: Empirical Studies of Earnings Quality
2015-8	Stine Ludvig Bech: The Relationship between Caseworkers and Unemployed Workers
2015-9	Kaleb Girma Abreha: Empirical Essays on Heterogeneous Firms and International Trade
2015-10	Jeanne Andersen: Modelling and Optimisation of Renewable Energy Systems
2015-11	Rasmus Landersø: Essays in the Economics of Crime
2015-12	Juan Carlos Parra-Alvarez: Solution Methods and Inference in Continuous-Time Dynamic Equilibrium Economies (with Applications in Asset Pricing and Income Fluctuation Models)
2015-13	Sakshi Girdhar: The Internationalization of Big Accounting Firms and the Implications on their Practices and Structures: An Institutional Analysis
2015-14	Wenjing Wang: Corporate Innovation, R&D Personnel and External Knowledge Utilization
2015-15	Lene Gilje Justesen: Empirical Banking
2015-16	Jonas Maibom: Structural and Empirical Analysis of the Labour Market
2015-17	Sylvanus Kwaku Afesorgbor: Essays on International Economics and Development
2015-18	Orimar Sauri: Lévy Semistationary Models with Applications in Energy Markets
2015-19	Kristine Vasiljeva: Essays on Immigration in a Generous Welfare State
2015-20	Jonas Nygaard Eriksen: Business Cycles and Expected Returns
2015-21	Simon Juul Hviid: Dynamic Models of the Housing Market
2016-1	Silvia Migali: Essays on International Migration: Institutions, Skill Recognition, and the Welfare State
2016-2	Lorenzo Boldrini: Essays on Forecasting with Linear State-Space Systems
2016-3	Palle Sørensen: Financial Frictions, Price Rigidities, and the Business Cycle
2016-4	Camilla Pisani: Volatility and Correlation in Financial Markets: Theoretical Developments and Numerical Analysis

2016-5 Anders Kronborg: Methods and Applications to DSGE Models

2016-6 Morten Visby Krægpøth: Empirical Studies in Economics of Education

2016-7 Anne Odile Peschel: Essays on Implicit Information Processing at the Point of Sale: Evidence from Experiments and Scanner Data Analysis

2016-8 Girum Dagnachew Abate: Essays in Spatial Econometrics

2016-9 Kai Rehwald: Essays in Public Policy Evaluation

2016-10 Reza Pourmoayed: Optimization Methods in a Stochastic Production Environment

ISBN: 9788793195394