Computers and Electronics in Agriculture 143 (2017) 273–285

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

Original papers

Improvement of feedlot operations through statistical learning and business analytics tools

Hector Flores[a], Cesar Meneses[b], J. Rene Villalobos[a,b,*], Octavio Sanchez[b]

[a] International Logistics and Productivity Improvement Laboratory, Arizona State University, Tempe, AZ, United States
[b] TIS Consulting Group, Mesa, AZ, United States

ARTICLE INFO

Keywords:
Growth patterns
Feedlot
Cattle industry
Statistical machine learning

ABSTRACT

A decision-support, modeling tool is developed that can project future cattle growth patterns in a feedlot based on a low dimensionality dataset available at the start of the feeding process. This work adapts the predictive performance of two well-known statistical machine modeling tools, gradient boosting and random forest regression, to predict future cattle growth. Time series analysis techniques are then used to create an ensemble method that further improves prediction accuracy from individual modeling outputs. Hierarchical clustering techniques are used to leverage projected growth patterns to increase group homogeneity when assigning cattle to different feeding pens. Finally, a profit maximization method is developed that estimates the optimal time each individual cattle should remain in the system under different revenue and cost estimates.

The purpose of this work is to incentivize the implementation of modern statistical learning tools in cattle management operations, especially within low-to-mid scale operations that traditionally rely on the expertise of its workers and have limited cattle and process information. Access to 'off-the-shelf' statistical learning tools, requiring minimal user-interaction, not only enhances prediction accuracy but helps automate operational decisions. This results in higher process efficiencies and improved standardization practices, while also helping identify profit opportunities. Finally, integrating these components into a single operating framework allows the tool to adapt to changes in data characteristics, which is especially important within non-standardized processes. We show the application of this tool through a case study implementation on a mid-scale operation in the northwestern state of Sonora, Mexico. From our case-study results, it was found that the modeling tool can satisfactorily predict growth patterns based on a low-dimensional set. It also can also capture historic decision-making when segmenting cattle into homogenous groups during their feeding process. Furthermore, it can help identify profit opportunities when estimating optimal cattle system times under varying market conditions.

1. Introduction

Livestock management in developing countries have traditionally relied on the expertise of its workers for tactical and operational decisions, especially in places that lack access to sophisticated technologies. In this environment of work, the lack of analytics systems capable of transforming data into actionable items limits planning capabilities. Access to better analytics systems, such as predictive and decision-making tools, can help livestock management systems improve standardization practices, which in turn increases process efficiencies and enhances the ability to capitalizable profitable opportunities. For example, in 2010, it was estimated that in the United States there were about 14 million cattle heads on over 75,000 feedlots (USDA, 2011). In recent years, the use of advanced analytics within agricultural processes has risen due in part to better access to information technologies and increased focus on data-driven decisions. The same trends are observable in the cattle industry as access to cheaper data collection system has opened new questions on how to transform this data into useful information and effective decision-making (Gebbers and Adamchuk, 2010). In this paper, we introduce an integrated framework that incorporates machine learning and forecasting tools into tactical level planning for commercial feedlots for beef cattle. As part of this work, we present a decision-support modeling tool that can forecast individual cattle growth based on a low dimensional dataset at the start of the feeding process. Access to this tool allows operators to create homogenous cattle groups based on individual predicted growth patterns, which in turn enables the implementation of more efficient batch-based systems. In this paper, we discuss a decision framework that allows feedlot operators identify optimal time periods for slaughtering individual animals based on defined revenue and cost functions.

* Corresponding author at: International Logistics and Productivity Improvement Laboratory, Arizona State University, Tempe, AZ 85287, United States.
E-mail address: rene.villalobos@asu.edu (J.R. Villalobos).

https://doi.org/10.1016/j.compag.2017.10.026
Received 22 December 2016; Received in revised form 18 July 2017; Accepted 28 October 2017
Available online 09 November 2017

Overall, our goal is to promote the use of modern analytical tools in cattle management systems that operate with limited access to sophisticated information technologies. In order to put the problem at hand within the proper perspective we give a general description of the beef-cattle feedlot operations next.

1.1. Context of the problem

This section presents a brief introduction to the set of activities performed within commercial feedlots for beef cattle and gives an overview of operational aspects relevant to this paper. For a more detailed description of a feedlot operations the reader is referred to other sources such the USDA's Beef Feedlot Industry Manual (USDA, 2011). This paper is based on intensive cattle feedlot operations. Typically, each feedlot is divided into multiple pens to which cattle are confined and given diets tailored to each pen. The number of animals in these pens may vary but a typical number is approximately 60. Thus, a feedlot with 10,000 cattle head would be comprised of an estimated 150 pens.

In simple terms, the main purpose of the feedlot is to efficiently transform cattle feed to weight gain. In a typical operation, feedlot cattle are received from different suppliers where they often have been raised in open range conditions and fed with grass and other forage. Once cattle arrive to the feedlot, they are temporarily assigned to a pen where they will stay until it is deemed necessary to move them to a homogenous group. During this time, feedlot operators decide how to homogenize cattle population based on weight and rate of development, from which then a feeding plan is constructed. Reassigning cattle to different pens is usually an expensive process, hence feedlot operators try to minimize this procedure. Thus, the initial and subsequent assignments of cattle to feed pens is a key decision in any feedlot operation. Fig. 1 presents a flow chart with the main sorting/pen assignment decisions made during a cattle's feedlot process, which shows the points of the process in which cattle are assigned and reassigned to different pens.

The importance of pen assignment/transportation decisions in a feedlot has been previously documented and addressed from different perspectives (e.g. Dror and Leung, 1998; Basarab et al., 1999; Kononoff et al. 2015). One of the key factors in minimizing cattle pen movement and maximizing profit is the ability to construct growth-homogenous groups within individual pens. This allows operators to have similar feed requirements and meat quality per pen, as well as reduce the need for cattle reassignments during the weight gain process. The main factor used for this set of decisions is the animal's weight measurement at different times during the feedlot process. However, other factors used in sorting the cattle into different groups include backfat and marbling estimates, genotype, and hide color (Kononoff et al., 2015).

1.2. Planning problem description

Within sophisticated cattle management systems, feedlot operators have access to specific information regarding the breed and type of cattle that enter the system. With this information, operators can use widely available cattle growth models to accurately estimate future weight patterns. However, when this level of data granularity does not exist, the use of these cattle growth models become less adequate, and the system becomes more reliant on the expertise of its feedlot operators. Although in some cases these systems still may perform well, lack of accuracy on growth pattern predictions increases process variability and hinders efforts to standardize the feeding process. For example, inaccurate growth projections lead to non-homogenous cattle groupings and prevents the implementation of more efficient batch systems. Moreover, cattle in this type of livestock operations is usually acquired from breeders with different production practices and feeding systems (Kilpatrick and Steen, 1999), which adds further variability to initial growth predictions. Thus, the development of growth prediction models tailored to operating environments with limited background information on cattle is key to efficient and standardized feeding processes.

Another problem in feedlot operations within low information environments is a lack of access to continuous weight monitoring systems. In these cases, weight information is typically collected at three distinct points during its feedlot process: (1) an initial measurement when cattle is received, (2) measurements after a short observation period during which initial growth is assessed, and (3) final weight measurements before being sacrificed. Based on these three pieces of information, as well as their own previous work experience, operators estimate each animal's future growth and segment individual animals into homogenous groups. For example, cattle expected to show higher gain weight during its feeding process, or better performance, will be grouped with other high-performing animals. This enables batch operational planning rather than working on a per-animal basis. However, as cattle operations expand and the number of breeders increase, cattle characteristics become more heterogeneous making this type of decision-making more difficult, especially when access to additional experienced feedlot operators is limited. Moreover, even under the availability of reliable human expertise, sudden changes in cattle characteristics may be difficult to identify without the proper decision support tools in place.

1.3. Study objectives

In this study, we address these problems by developing an integrated decision-making framework that focuses on three different aspects of the cattle feeding process. The first component of this work is the development of a predictive modeling tool that can project future cattle growth by leveraging on information typically available to low-to-mid scale operations. The second component details a clustering method that groups cattle according to individual forecasted growth curves at the beginning of the process. Lastly, the third component formulates a method of identifying suggested optimal times at which each animal should be slaughtered with the goal of maximizing expected profit margins. It should be noted that these tools do not seek to replace feedlot operator expertise but would serve as a decision-support tool aimed at facilitating decision-making tasks.

Overall, the objective of this framework is the development of an automated decision-making tool implementable within low-information environments. Thus, the integrated framework is developed under an open-source platform able to interact with different technology environments. To validate the different techniques within this framework, a case-study is implemented for a mid-scale, operation in the state of Sonora, in Mexico. Available datasets are used to create a fully-integrated decision support tool for feedlot operators to provide individual sets of decision suggestions for each animal at the start of the feeding process. Finally, given that the framework is fully automated, the decision-making system has the flexibility to adapt to changing data characteristics.

In the following few sections, a review of literature is presented that highlights current work related to cattle growth modeling and applications of statistical tools in feedlot operations. The third section details our methodology framework, including brief theoretical background into its different components. In the fourth and fifth sections, the results of our case-study are presented and future work is discussed, respectively.

2. Review of literature

In this section, we present a review of works that seek to apply analytical tools in beef production management. In our review, focus is given to works related to prediction modeling techniques for projecting future cattle growth. As part of this section, traditional metrics used to gauge animal performance are presented. Also, this section outlines the different types of statistical tools currently used to predict cattle growth, including random regression, machine learning, and