Torcuato Di Tella university

Master in Management + Analytics Thesis

A combinatorial optimization approach for livestock planning

By

Joaquin Pablo Tempelsman

Advisor: **Javier Marenco**

January 2024

A combinatorial optimization approach for livestock planning

Master's Thesis in Data Management and Analysis

Joaquin Pablo Tempelsman

Abstract

This thesis explores the planning of a real livestock operation over a multi-period scheme. Our goal is to design and validate a prescriptive solution based on mixed integer linear programming. The applied strategy is based on optimizing revenue bound to restrictions that represent the complexities of a livestock breeding company.

To validate the proposed model, benchmarks are performed on multiple scenarios and settings against a contesting model that represents currently applied heuristics from the company managers.

The analysis concludes that the solution not only successfully adapts to complex and diverse proposed scenarios but also finds for the designed experiments, convincing and explainable solutions that can potentially be adopted by management to systematically improve the company operation.

In the concluding chapters, we discuss potential model enhancements and outline a roadmap for integration. The two-step integration proposal is designed to address the primary challenges identified during the assessment phase.

Contents

I. Context and justification for the study

In recent years, the intensive use of technology has been the main variable of change in multiple industries. Agriculture has also been a fertile ground for innovation, particularly if we consider the Argentine context. Since the 1990s, the increasing use of technology has made it possible to increase the yield per hectare, bringing it closer to the international production efficiency levels year after year.

Technology often manifests through proprietary software in distinct operational segments to solve a specific task, unlinked to strategic managerial decision making. This thesis aims to build a proposal for an open source prescriptive solution that can be designed, owned and integrated directly at the business core.

Industries like banking and healthcare with a longer history of harnessing and using data to truly and consistently direct their decision making process still face significant cultural challenges. When we shift to industries that are not traditionally digital, we face a double challenge both in the management of the data and the adoption of the solution. Despite this, the rewards of said integrations are yet abundant to be attained.

Thesis objective

The project goal is to build a prescriptive solution for a family-owned livestock farming company based in the Santa Fe Province, Argentina. The company management shared the need to utilize a data-driven method to guide them in optimizing central parts of their operation that currently are being handled by using a combination of past experience-based knowledge and future expectations.

Every year the company has to decide how much of the female stock is going to be assigned to reproduction (increasing future stock) or selling it in the present. Subsequently, they also need to plan how much stock will be sold at each weight category or stage among three main alternatives. Each decision prices and costs vary on a daily basis depending on diverse factors.

Interaction with business stakeholders revealed the need to build a quantitative method in order to enable them to strategically decide on how to plan two fundamental aspects of their operation. Discerning the **optimal balance between increasing reproductive stock or prioritizing present sales given different market conditions**. The second inquiry was to define the **optimal selling category** between weaning, middle weight and high weight.

By adopting a data-based decision-making process, the goal is not only to improve revenue but to also build a flexible tool that allows them to take better informed managerial decisions while doing so. Prescriptive integration allows one to quickly adapt and learn from potentially changing scenarios influenced by local and foreign market prices, trade regulations, and weather restrictions, among other variables.

Within a mixed integer linear programming (MILP) model we can consider multiple business variables and limit them by a series of restrictions in order to maximize productive performance in a given intertemporal horizon. We will seek to model most relevant business variables on an intertemporal decision scheme that maximizes profit given the available resources. We will consider local market prices for each category through time, local currency exchange rate and breeding cycles along other business dimensions. We will design a validation strategy, run benchmarks and analyze proposed solver solutions. Finally, we will outline a possible integration workflow into the company's current operation.

Literature revision

MILP models have proven valuable to represent multiple scenarios while being an extension of linear programming models where some or all decision variables are integers instead of real numbers. The use of these models in the field of agricultural production has a long history worldwide. Particularly in Argentina and Brazil, we can find some publications applying linear programming models to a diverse type of use cases.

Land allocation optimization between livestock and agriculture activities Ras, C. H., Marra, R. M. A., & Tettamanti (2020). Model allocates land to activities limited by resources available, maximizing profit obtained. Agroindustrial activities are defined at different sub-stages. The model solution proposes doubling allocated land to livestock production and adopting a new specialized activity related to weight gain that is not yet being considered. Restrictions consider costs, risk, sustainability among other considerations.

Multi-period transport and harvest operation optimization. Filho, A. A., Melo, T., & Pato, M. V. (2020). In the context of the Brazilian sugarcane industry, the model solves over a multi-period planning horizon, scheduling for machinery and transport vehicles deployment subject to constraints representing available resources, expected demand, crop yield, weather condition and work schedules. The objective is to minimize cost incurred by equipment utilized and time required to meet demand in a bi-objective function. The model output indicated shorter deployments at a higher cost, leading to reduced machine maintenance expenses and ultimately maximizing overall revenue.

Measuring and maximizing sustained population from resources provided by the Pampa húmeda. Frank, R. G. (2020). Evaluating per capita supply for 20 food products, this work studies how to allocate resources in order to maximize production. Results show that the Pampas region can sustain slightly over 113 million people following Argentine dietary guidelines.

Forest supply chain optimization while quantifying social and environmental impact. Campanella, S. R., Corsano, G., & Montagna, J. M. (2018). The model target is to optimize the operation to efficiently utilize wood residues generated by determining location, size and quantity produced while maximizing benefit. The solution was implemented considering the characteristics of the forest industry in the northeast region of Argentina.

Water resource optimization for the design of the Rio Negro basin. Gradowczyk, M. H., Jacovkis, P. M., Freisztav, A. M., Roussel, J.-M., & Tabak, E. G. (1990). This paper introduces OPER, a mathematical model composed of four submodels. The objective is to analyze and plan water resource systems composed of reservoirs, hydropower stations, navigation channels, and more, situated along a river. OPER4, a mixed integer programming model, optimizes investment sequences over the time horizon. This model was applied to design the water resources in the Rio Negro basin system in Argentina.

Thesis outline

Business case. In this chapter we will describe the livestock lifecycle and provide a general overview of the main business decisions regarding stock management. Commencing from birth and progressing to the weaning stage six months later, optionally transitioning to reproductive stock or continue until reaching medium or even high weight stages. Each decision is related to a specific point in time for each class that exists in the operation. We will categorize all livestock into three classes: male (Class 1), female (Class 2), or cattle designated for reproduction (Class 3). By the end of this section we will be presented with the main business variables that we will want to model in the future sections.

Data inputs and considerations. Here we will aim to describe how we will handle the data that we will use as input for our solution. We shall describe the processing done to cattle prices and future prices estimation for longer time window analysis. We will explain how maintenance costs are being accounted for while progressing through each step in the life cycle, which are the four possible sales moments or stages in which sales may occur, and we will provide measurements and approximations for animal aging and weight.

Solution proposal and model definition. In this section we will delve into how the MILP model will harbor all mentioned inputs and generate a solution we can comprehend and utilize in order to add business value. We will mention some general theoretical concepts related to combinatorial optimization and linear programming techniques. We will define revenue, our main metric that we will seek to optimize in our model and determine decisions for maximizing it. Finally we will specify the model in its mathematical form, providing the objective function, sets, parameters, variables, and constraints, and explaining their interactions.

Benchmark analysis. Within this segment we will outline two proposed strategies to validate the solution´s impact. We will mention some caveats like age translation and methods applied to shape and control the model solution into something the business can effectively use.

Moving forward, we will define 19 experiments with 6 variables that we will explore to assess how the solver performs on multiple varying scenarios. Examples of these variables are subsetting the experiment into different past historical moments in time impacting prices, different durations, varying pregnancy indexes, sales fix costs and a last variable that will use constant prices.

After running the model across the 19 experiments we will analyze the impact over the main business dimensions, including sales performance, transfers, stock, and generated revenue. Afterwards, we will select two representative experiments and run a deep analysis that leads to understanding the rationale on the solver's optimal solution and trying to determine the source of the generated value.

II. Business case

Description of the specific case to be addressed

In this section we will describe the livestock operation through one complete cycle, commencing in September when the newborns are bred yearly. Depending on the decisions taken, the cycle can last up to almost two years until completed if animals are selected to be sold at the high weight category. During this process, the cattle will mature in groups or batches, gaining weight and increasing its maintenance cost month to month. Typically at certain months of age, the animal will reach a weight category in which it can be sold on the local market for a price by kilo attained. There are three weight categories defined at a full weight of approximately 165 kgs, 300 kgs, and over 391 kgs.

Yearly breeding will be enabled by reproductive stock composed by previously separated female cows and bulls. Female cows for this purpose are selected at approximately 11 months of age. They will start breeding at two years old, every year, until they are sold in a special category for reproductive older stock that we will define as Stage 4. Reproductive female cows typically perform between one to seven reproductive cycles before being sold.

We will define Class 1 as all male stock, Class 2 as female stock and Class 3 as female stock selected for reproduction. Only Class 2 (female) cows can be transferred to Class 3 at 11 months of age. This process is described in Figure 1. Time advances left to right, starting at breeding.

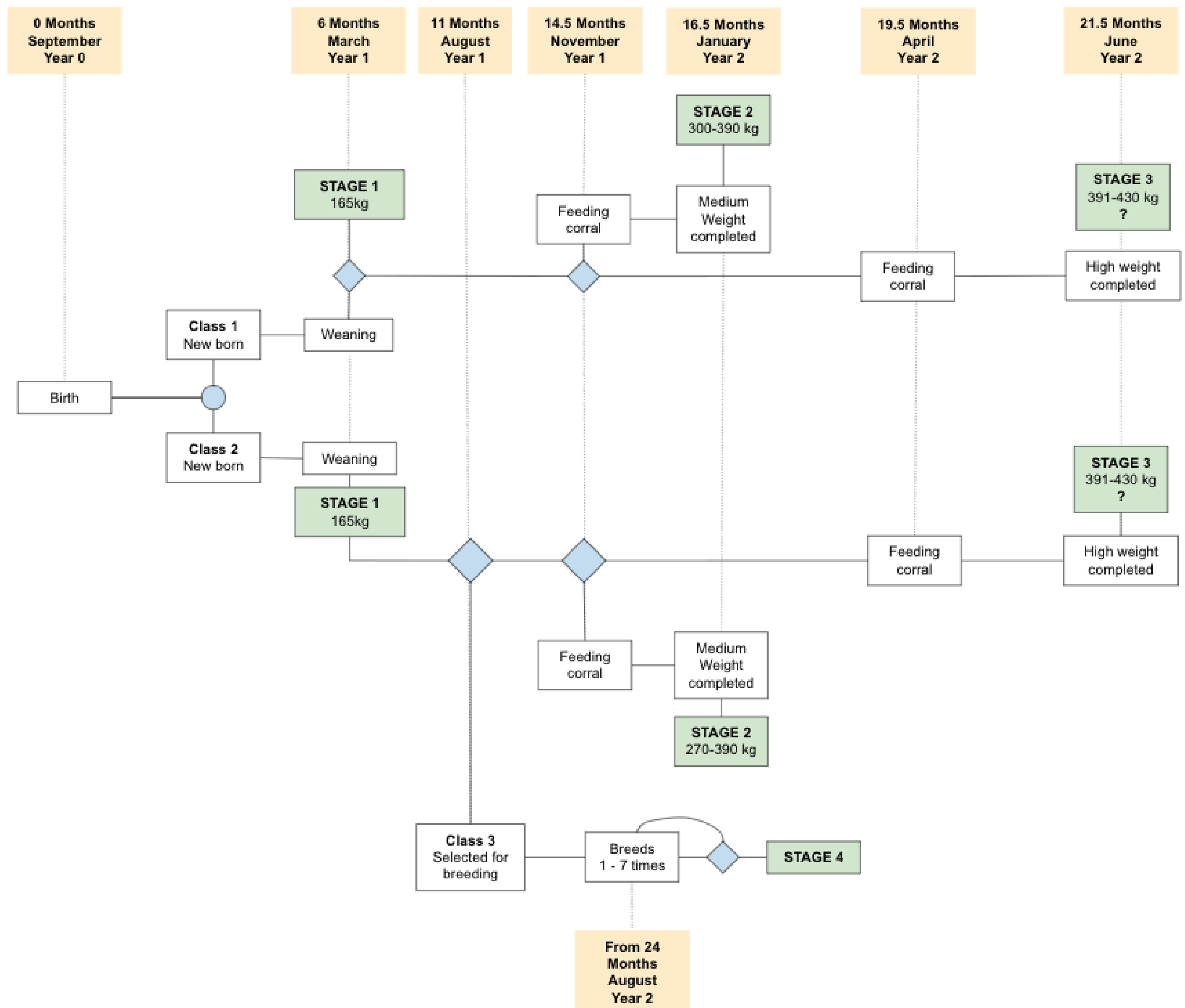

Figure 1: Livestock cycle business decision tree. Own elaboration

Female cows assigned to reproduction give birth during August, yearly. This is aligned with the beginning of spring as higher temperatures have a positive impact in the early development of the newborns. Six months later, both male and female young cattle are ready to be separated from their mother (weaning), the moment in which they can be sold (Stage 1) or raised in the local premises.

Sales are possible at four different stages. The price for each animal is defined by its weight multiplied by the price per kilo at the category in which the animal is being sold at. Stage 1 for selling newborns between six and eight months old as soon as they are separated from their mother. Stage 2 in the middleweight category between 16 and 18 months old. Stage 3 in the high weight category between 30 to 36 months of age. Stage 4, only for cows assigned for reproduction that can be sold at any age and time.

For the medium and high weight category, cattle is raised and fed on an open grass-feed field until 30 days before being sold, when it is moved into the feeding corral. Weight gain per day increases resulting also in a higher maintenance cost during this month. If the cattle is sold at medium weight category they reach complete weight between 16 to 17 months. High weight category is reached between the months 30 to 36 depending on when it's moved to the feeding corral. Once the cattle is moved into the feeding corral, it wont go back to open-grass feeding again.

Only for the female branch, during the 11th month of life the operation decides how much is permanently assigned to reproductive stock. The main difference is that reproductive stock will remain in open-grass feeding until sold, having less overall maintenance cost than the other classes. All reproductive stock females will give birth every year from two years of age and older, with an average 86% success rate (the rate is a simplification that we will expand later on). Years later at the end of their life cycle, they can be sold at Stage 4, for a specific market price, composed of half the price per kilo of the female middleweight category.

On collaboration, assessment and business discovery stage

The first step into designing the solution is to determine which business dimensions are relevant to include into the model. Appropriate abstraction is key as it is not possible nor necessary to represent every moving part of the operation. Avoiding this way to design a model that is either too complex to manipulate and replicate or computationally exceedingly demanding to solve. This requires the analyst and company representatives to work together, gathering key business definitions in a continuous refinement process that involves analyzing data available, building assumptions and managing expectations.

Particularly, this project benefited from consistent support from the company stakeholders during no less than two and a half years. Central to this interaction is that the objective is not only to try a theoretical approach circumscribed to this paper but to improve the operation processes empirically. Aligned incentives enabled good will and a high level of collaboration that is required for these projects to have real impact, propelling the analyst to deeply learn the specificities of the business and the company to be open minded to explore and improve into new ways of running their business.

Strategic business needs and aspirations

Exploratory discussions with key stakeholders within the organization revealed two crucial needs. One being aimed at determining the optimal balance between expanding reproductive stock against prioritizing current sales. The second motivation being to find the optimal selling category among weaning, middleweight, and high weight options.

Increasing present reproductive stock would mean higher future returns considering scenarios where prices and costs stay constant over time. In future sections we will see that this is not the case as both prices and cost vary significantly over time, Influenced by a variety of factors both domestic and foreign. While expenses have a mix component on both local currency and US dollars, sales are quoted exclusively in US dollars. For this reason, changes in the currency exchange rate impact directly on revenue. Increasing inhouse maintained stock over a certain limit would exceed forage local production,requiring sourcing it from external producers at a higher cost.

Complexity increases when we consider the number of intertwined intertemporal variables. A robust business plan implementation is daunting to design and assess over time. The challenge is higher for small to medium sized businesses as they dont have a team that can focus on every facet, specially for Argentina based companies due to the elevated macroeconomic fluctuations as it is in our case.

Data Inputs and considerations

Prices. Daily average price per kilo for each selling stage obtained from mercadodeliniers.com.ar. Cattle prices are set at the local Argentinian market, capturing the month price at the first workday of that month. The values are converted to US dollars using the informal exchange rate (dollar blue), taking the previous 7 day average value. The informal rate was considered to normalize price fluctuation as the official exchange rate was being controlled by the local government as a means to control inflation. Utilizing the local inflation index IPC to deflate prices could be another valid strategy, but the company opted to use the dollar as its primary reference instead. Weaning stage prices are specified as 20% more than female Stage 2 prices as a business definition.

Prices are captured from Jan 2017. From Jan 2023 and afterwards, prices are forecasted. Prices seasonal effects are modeled using a Holts Winter method as shown on Figure 2. This is a naive approach where the idea is to represent seasonal prices during the year and be able to run benchmarks on how the solution performs for long periods of time.

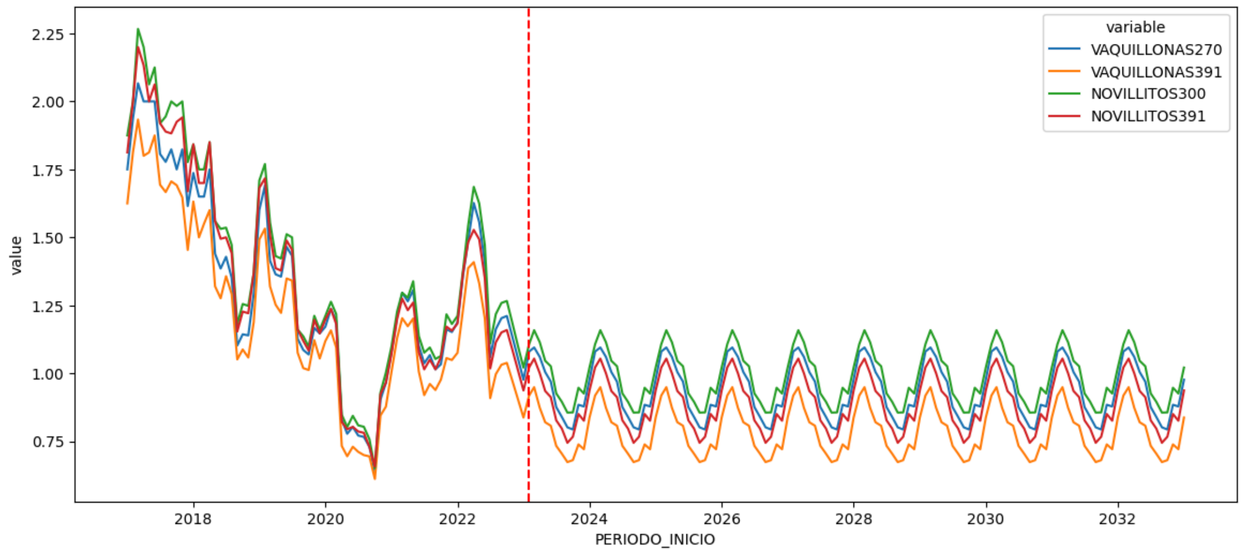

Figure 2: Median price per kilo in US dollars from January 2017 to January 2023 and forecasted until January 2033. Own elaboration. Source: www.mercadodeliniers.com.ar www.dolarhoy.com

Maintenance Cost. Costs disclosed by the business were in the form of accumulated costs at each possible selling point for each class from Stage 1 to Stage 4. The model calculates the cost per period incurred as it is more realistic, specially when solving an intertemporal period to period structure. To solve this transformation we used B-spline interpolation of order three to build a smooth curve and then calculated the derivative of that function. Evaluating the derivative for all months required allows us to obtain the partial cost per period for all the considered periods.

Male cattle cost 10% more from month 15 and beyond, this is visible on Figure 3 where the male class shows this increase. The cost of the reproductive class is fixed as their food intake stays constant during time.

In Figure 3 we can see the cost curves per class overlapped by the three selling stages. Reproductive classes can be sold at any time. Selling Stage 1 takes place from three to eight months of age. The second selling stage spans from 16 to 18 months, while the third selling stage is performed between months 30 to 36.

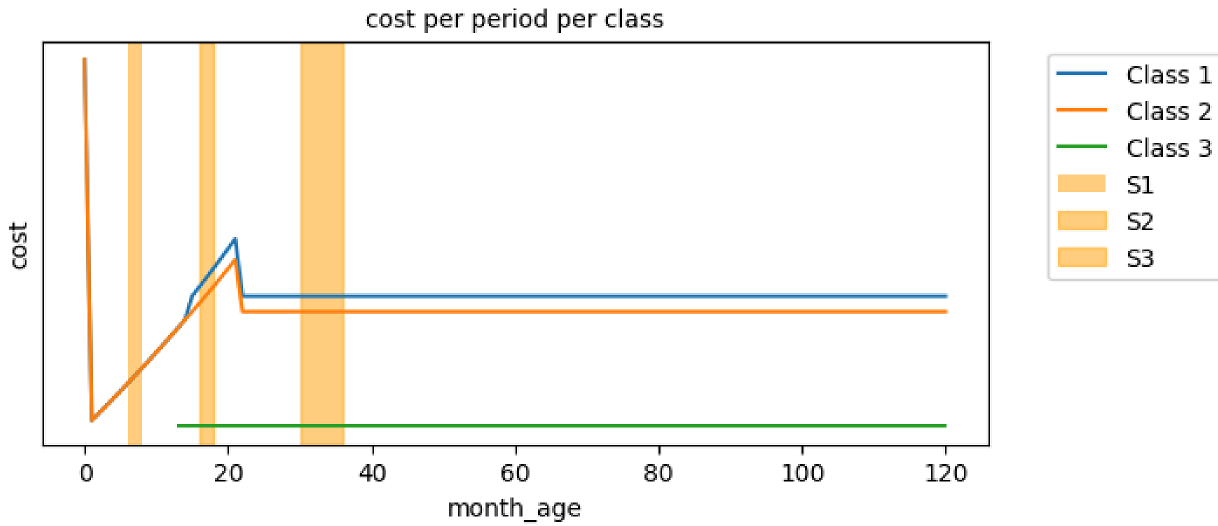

Figure 3: Cost per period per class in US dollars with selling stages. Source: Company data. Y axis values removed for data privacy. Own elaboration

Selling Class 1 or Class 2 at medium or high weight implies two different cost curves given that we move the cattle into the feeding corral at different moments in time (two months before selling). As the model requires only one cost curve per class, one presented caveat is that we are modeling that all animals are transitioned to the corral at medium weight periods and then apply normal maintenance cost for the rest of the cost function. This is represented in the upward slope and then downward, which in a real scenario would not have that downward slope. To resume, we are shifting the cost increase for Stage 3 feeding corral stock 15 periods before. This can be refined by adding a fourth class in the future.

Sales. Sales are restricted to occur only on Stage 1 to Stage 3 during the lifecycle of the animal. Stage 4 is again only for reproductive stock. Fix sales costs represent cattle transportation, paperwork generation and group weight-ins among other required steps during selling stock. To avoid low quantity unrealistic movements, a minimum sale quantity is set at ten per period for Class 1 and 2. We expect to reward higher quantity group sales that diminish the fixed cost per period. Reproduction stock can be sold at any moment in time by half of the price per kilo of the medium weight female category. This could be improved in the future, by applying a decreasing price the older the animal which is indeed how prices are set.

Age. Measurement of age at an individual level is not currently performed by the business. Weekly stock reports are used to map stock units assigning animals to different age bins that represent their age through the cattle life cycle. The model´s input requires specific input at the months of age level. This requires a “translation” that is done by randomly assigning the age for a given interval using uniform distributions for all classes and a right skewed normal distribution for Class 3 that better represents the age distribution among that class.

Weight. Same as age, precise measurement of each unit weight is not feasible. Periodic measurements for each bin/group are performed to keep track of the average weight for each group. This is a costly and high risk operation that requires manipulation of hundreds of large size animals and getting them on an industrial scale that gets the average group weight. For our solution, we will use the average weight per group based on the latest measurements available.

Births. Every August, for each reproductive class animal in stock that is 24 months or older, 0.43 newborns are added to the stock for both classes. This is calculated by multiplying 93% successful pregnancy rate with 93% of successful weaning rate. Then, splitting it in half represents the fifty-fifty percent chance of getting a male and female newborn. This is a simplification as 86% is the real pregnancy rate and 6 months afterwards 86% survive the weaning stage. We are applying 86% successful rate directly at breeding and skipping to represent failed weaning stages cases.

Data transformations done to conceal private data. The data utilized for this experiment is real. To protect sensitive information related to stock and costs per category, we have done a linear scaling to all the data inputs by a certain number. The results regarding impact of the solution and decisions made by each solver are exactly the same as if the transformation has not been made.

Proposal for a solution using mixed linear optimization

Introduction

Before diving into the model itself, let's break down its key components:

Model: The model is the mathematical representation of our business problem to solve. It is an abstraction that represents relationships between decision variables and constraints given a specific data input, which we are going to define later in this section. The model is written in the ZIMPL modeling language.

Solver: The solver refers to the computational implementation of the algorithm or algorithms applied to find the solution that maximizes our objective function (defined in the model). Among all feasible solutions in the solutions space, we seek to maximize the revenue defined in the objective function. The specific solver we are using is SCIP, a non commercial optimization tool for both mixed integer programming and mixed integer nonlinear programming. It is widely used and open source.

The solver or optimization algorithm is going to assign values to decision variables and obtain a certain revenue as output. Decision variables are limited to deciding quantity to **sell**, **maintain,** and **transfer** for each class, age, and period while fulfilling all constraints defined in the model. Among all feasible solutions that satisfy defined restrictions, the solver efficiently explores the feasible search space trying to find the optimal solution that maximizes revenue.

Revenue is calculated and represented in the objective function for each period as **stock sold** multiplied by its price minus **stock maintained** in that period multiplied by its maintenance cost. Maintenance cost can exist with or without sales in any particular period. An independent term representing sales fix cost weights in for periods where sales are carried out for each age and class sold, representing handling and administrative operation costs. The sum for all periods of said objective function summarizes for **total revenue**.

When we **sell**, there will be a specific **price** for that age, period and class. When we **maintain stock**, we will pay a maintenance **cost** specific for the period, age, and class.

Price, costs, and **initial stock** are data inputs we give to the model with other parameters like max age allowed, minimum selling quantity allowed per month, fixed sales cost per month, pregnancy index, month selling enabled months per class (Stages 1 to 4), among others.

Model

Sets

$T(\text{Periods})$ Represents discrete time periods, from 1 to a maximum period, in which decisions about cattle management are made. Throughout the model definition, we shall use the subscript “t” to refer to elements from this set.

$T0(\text{Periods})$ Represents discrete time periods, from 0 to a maximum period, we shall use the subscript “t” to refer to elements from this set. Differs from the previous set by including the 0 as it will be used for variable domain definitions instead of restrictions as the previous set.

$E(\text{Ages})$ Encompasses the possible ages of animals from -1 to a certain maximum age (animal_max_age). We shall use the subscript “e” to refer to elements from this set. We need to define e from -1 because the starting period is 0 and the objective function establishes a relationship with e-1, which is period -1.

$C(\text{Class})$ Represents the considered classes of cattle, with values 1, 2, and 3. Class 1 represents male stock, Class 2 denotes female stock, and Class 3 characterizes female stock selected for reproduction. Similarly, we shall use the subscript “c” to refer to elements from this set.

Parameters

$\text{cost}[t, e, c]$ Defines the cost associated with animals of a particular age e, class c, and period t.

$\text{price}[t, e, c]$ Indicates the selling price of animals for each age e, class c, and period t.

$\text{initial\_stock}[e, c]$ Specifies the initial stock of animals for each age e and class c at the initial period.

$sell\_1$ as the ages in which Stage 1 cattle can be sold which in our particular case is between 6 and 8 months of age.

$sell\_2\_3$ as the ages in which Stage 2 and Stage 3 cattle can be sold, in our case are months 16 to 18 and 30 to 36.

$preg\_rate$ as the mean pregnancy success rate for female breeding.

$min\_sell\_c1\_c2$ as the minimum selling required quantity for Class 1 and Class 2 per period. In case sales are performed, this sets a minimum quantity for that operation.

$max\_sell\_c1\_c2$ as the maximum selling quantity allowed in that period. Defined as an auxiliar to establish the functioning of "minimum sales" 8th restriction.

$max\_periods$ as the total number of periods and last period available in the model.

$not\_august\_periods$ as a list of periods that are not August, in which births won't happen.

$august\_periods$ as the list of periods that correspond to the month of August, in which births will occur.

$sell\_fix\_cost$ as a fixed sales cost for each period and age and class with existing sales. This represents the administrative overhead, handling and transport costs of a sales operation. This cost is higher the more distinct types of stock are sold within that period.

Variables

$x[t,e,c]$ as available stock in each period t in T, at each age e in E, for each class c in C.

$y[t,e,c]$ as sales in each period t in T, at each age e in E, for each class c in C.

$w[t,e]$ as the transfers from Class 2 to Class 3 in each period t in T, at each age e in E.

$n[t,c]$ as the number of births in each period t in T for each class c in C.

$s[t]$ as a binary variable denoting whether sales occur in that period, for each period t in T.

$k[t]$ as a binary variable allowing sales when there are no births for constraints 26 and 27 (defined in section "*Benchmark Strategy 2 - business baseline vs optimized solution*").

$l[t]$ as a binary variable allowing transfers when there are no births for constraint 28 (defined in section "*Benchmark Strategy 2 - business baseline vs optimized solution*").

Objective function

$$\sum_{t,e,c \in T,E,C} y[t,e,c] * \mathrm{price}[t,e,c] - x[t,e,c] * \mathrm{cost}[t,e,c] - s[t] * sell\_fix\_cost$$

The goal of the model is to maximize profit, calculated as the sum of revenue from sales minus the maintenance cost and fixed costs associated with sales.

Non-negativity constraints

$$\forall \langle t,e,c \rangle \ \in T0 \times E \times C: \quad x[t,e,c] \ \geqq \ 0 \tag{1}$$

$$\forall \langle t,e,c\rangle \ \in T0 \times E \times C: \quad y[t,e,c] \ \geqq \ 0 \tag{2}$$
$$\forall \langle t,e\rangle \ \in T0 \times E: \quad w[t,e] \ \geqq \ 0 \tag{3}$$
$$\forall \langle t\rangle \ \in T0: \quad n[t] \ \geqq \ 0 \tag{4}$$

Flow constraints

Conservation of Flow Equation - Class 1

$$\forall \langle t,e\rangle \in T \times E \text{ with } e > 0 : x[t,e,1] = x[t-1,e-1,1] - y[t-1,e-1,1] \tag{5}$$

Conservation of Flow Equation - Class 2

$$\forall \langle t,e\rangle \in T \times E \text{ with } e > 0 : x[t,e,2] = x[t-1,e-1,2] - y[t-1,e-1,2] - w[t-1,e-1] \tag{6}$$

Conservation of Flow Equation - Class 3

$$\forall \langle t,e\rangle \in T \times E \text{ with } e > 0 : x[t,e,3] = x[t-1,e-1,3] - y[t-1,e-1,3] + w[t-1,e-1] \tag{7}$$

Constraints 5, 6, and 7 set up the logical relationship between present, past, and future for all three classes, giving our model a sense of time. We are binding current stock to be defined as available stock on previous period minus previous period sales and transfers.

Sales constraints

Minimum Sales

$$\forall t \in T: \sum_{e,c \ \in \ E \times C \text{ with } c \neq 3} y[t,e,c] \geq \text{min\_sell\_c1\_c2} \cdot s[t] \tag{8}$$

Maximum Sales

$$\forall t \in T: \sum_{e,c \ \in E \times C \text{ with } c \neq 3} y[t,e,c] \leq \text{max\_sell\_c1\_c2} \cdot s[t] \tag{9}$$

No sales on final period

$$\forall \langle e, c \rangle \in E \times C : y[max\_periods, e, c] = 0 \tag{10}$$

Connect sales to available stock

$$\forall \langle t, e, c \rangle \in T \times E \times C : y[t, e, c] \leq x[t, e, c] \tag{11}$$

No sales with zero price

$$\forall \langle t, e, c \rangle \in T \times E \times C : \text{if } (\text{price}[t, e, c] \leq 0) \text{ then } y[t, e, c] = 0 \tag{12}$$

Constraint 8 demands that if sales occur, they must be at least for a certain minimum quantity $min\_sell\_c1\_c2$ for both Class 1 and Class 2. Constraint 9 complements Constraint 8 based on a virtual maximum selling limit defined as three times the total initial stock.

Constraint 10 forbids sales in the final period, this avoids last period sales that could comply with end stock requirements like constraints 15, 16, 26, 27, and 28. Constraint 11 binds sales to be less than the available existing stock. Finally, constraint 12 negates zero price sales (model input generation related) only enabling sales during specific animal ages that are related to selling stages.

Stock constraints

Initial Stock Set

$$\forall \langle e, c \rangle \in E \times C : x[0, e, c] = \text{initial\_stock}[e, c] \tag{13}$$

Forbid stock with negative age

$$\forall \langle t, c \rangle \in T \times C : x[t, -1, c] = 0 \tag{14}$$

Finish with at least as many Class 3 cattle as in the first period

$$\sum_{e \in E} x[0, e, 3] \leq \sum_{e \in E} x[\text{max\_periods}, e, 3] \tag{15}$$

Finish with at least as many young Class 3 cattle as in the first period

$$\sum_{e \in E \text{ with } e<30} x[0,e,3] \leq \sum_{e \in E} x[\text{max\_periods},e,3] \tag{16}$$

Constraint 13 sets initial stock from loaded inputs. Constraint 14 forbids negative age stock as a particular case that is not covered by restrictions 5,6 and 7. Constraint 15 restricts that the end stock of animals in Class 3 are at least the same levels as the initial stock of Class 3 for all ages. Constraint 16 repeats 15 for a subset of less than 30 month-old breeding cows, ensuring that the business cycle is possible in the future while the proportion of younger cows in Class 3 remains relatively stable.

Transfers constraints

Control transfers

$$\forall \langle t,e \rangle \in T \times E : w[t,e] \leq x[t,e,2] - y[t,e,2] \tag{17}$$

Forbid transfers in the initial period

$$\forall e \in E : w[0,e] = 0 \tag{18}$$

Allow transfers only at 11 months of age

$$\forall \langle t,e \rangle \in T \times E \text{ with } e \neq 11 : w[t,e] = 0 \tag{19}$$

Constraint 17 binds transfers to be equal or less than the available stock minus sales in that period. Constraint 18 forbids transfers on the initial period 0 as a special case controlling unwanted behaviors. Constraint 19 forbids transfers for all ages except 11.

Births constraints

Forbid births on initial period

$$\forall c \in C : n[0,c] = 0 \tag{20}$$

Class 1 births for Class 3 animals older than two years

$$\forall t \in \text{august\_periods} : n[t,1] = \sum_{e\ \in E \text{ with } e \geq 24} x[t,e,3]\ *\ preg\_rate/2 \tag{21}$$

Class 2 births for Class 3 animals older than two years

$$\forall t \in \text{august\_periods} : n[t,2] = \sum_{e \in E \text{ with } e \geq 24} x[t,e,3] * preg\_rate/2 \tag{22}$$

Limit births for Class 1 and Class 2 to certain months

$$\forall \langle t,c \rangle \in \text{not\_august\_periods} \times C : n[t,c] = 0 \tag{23}$$

Forbid Class 3 births

$$\forall t \in T : n[t,3] = 0 \tag{24}$$

Births define age zero stock except in the initial period

$$\forall \langle t,c \rangle \in T \times C \text{ with } t > 0 \text{ and } c \neq 3 : x[t,0,c] = n[t,c] \tag{25}$$

Constraint 20 forbids births in the initial period, in order to avoid unwanted behavior. Constraints 21 and 22 enable births for each existent breeding cow over 23 months age, preg_rate animals from Class 1 and Class 2 will be added to the stock in that period for each birth-ready cow. Constraint 23 forbids births in months different from August. Constraint 24 forbids births of Class 3, thus avoiding unwanted behavior. Constraint 25 binds age zero stock to be defined only by births except on the initial period.

III. Benchmark analysis

In this section we will describe two strategies designed to analyze and validate the solver's solutions. After this analysis we must be able to understand if our solution adds business value or not, in which scenarios it works, and how it adapts to changes in settings.

As we are using real data inputs, the main metric we are going to study is revenue as it is convenient while also considering other dimensions like changes in stock allocation, sales distribution, reproduction, and transfers.

Benchmark strategy 1 - Business past operation vs solver

This first benchmark strategy attempt was to map real past business decisions and replicate the same input settings for the model, get revenue obtained from both, and compare for multiple time windows. This strategy failed to ensure comparability between business and model as assumptions had to be implemented and noisy data translations carried out that made the analysis more complex than valuable. However, there were three noteworthy lessons obtained from the process, which are interesting from a modeling standpoint.

End of the world. As the solver optimizes given the modeled time period, it will tend to sell all possible stock (including reproduction) stock before or during the last period. This can lead to noise and unwanted conclusions. To deal with this, buffer extra periods were added to the model experiment to delay the end of the world effect and have sound conclusions. This was complemented with manually calculating the experiment revenue before the buffer periods are reached. Even with buffer periods, the end-of-the-world effect is minimized but still relevant and must still be accounted for.

Age translation. In the last period, all business side stocks exist in a specific category which has a selling price and an associated cost, there are no “in between” quantities in the business stock report. Differently, the model will usually end with some stock units at an age that has no price associated and thus cannot be sold directly (specially when adding buffer periods). The simple strategy applied to solve this was to “move” it to the nearest selling stage and use that price and cost, sometimes upward the lifecycle and sometimes downward depending on which one is closer.

Solver-informed prices. Future prices are unknown to the business owners so decisions are made based on expected future prices. The solver instead optimizes knowing the prices for all periods, which is an advantage when comparing against human past choices. Possible solutions were to forecast future prices to serve as input for the solver or add some noise to the solver input prices and calculate revenue on real prices instead.

The aforementioned issues led to another strategy described in the following section.

Benchmark Strategy 2 - business baseline vs optimized solution

If we could synthesize day to day business decision making into general heuristics that we could abstract into model constraints, then we would be able to compare this setting with an unrestricted solver in a more elegant benchmark with less to none comparability issues.

After discussing with the business owners, they could describe three rules that represent how they managed their stock and sales flows. Shared heuristics were stated as follows:

“For all female newborns during the exercise, 30% is sold at early stage (Stage 1), 30% is sold between Stages 2 and 3, and the remaining 40% is saved for reproduction”.

These rules are represented by the following constraints and will only be applied to the baseline variant model that we will benchmark against the free model.

Births heuristic constraints

30% births to sales at stage 1

$$\forall t \in \text{august\_periods with } t + 6 \leq \text{max\_periods}: \\ n[t,2] \cdot 0.3 = \sum_{v \in \text{ sell\_1}} y[t+v, v, 2] \cdot k[t] \tag{26}$$

30% births to sales at Stages 2 and 3

$$\forall t \in \text{august\_periods with } t + 16 \leq \text{max\_periods}: \\ n[t,2] \cdot 0.3 = \sum_{v \in \text{ sell\_2\_3}} y[t+v, v, 2] \cdot k[t] \tag{27}$$

40% births assigned to Class 3

$$\forall t \in \text{august\_periods with } t + 11 \leq \text{max\_periods} : \\ n[t, 2] \cdot 0.4 = w[t + 11, 11] \cdot l[t] \tag{28}$$

Remember that we will restrict both models to finish the exercise with the same amount or more of reproductive class as received during the initial period. This enforces both models to avoid a solution that won't allow the business to keep running in the future. This strategy controls the end of the world effect, limiting sales on final periods to no less than initial stock levels.

As one model is more restricted than the other, we will always expect that the less restricted objective function value to be higher. The key will be in understanding both the magnitude of the difference and how that difference is obtained regarding business calls, on multiple settings.

Analyzing benchmark results

We defined a series of 19 instances based on a grid of settings that allows us to evaluate the model impact in multiple scenarios. Instances are given by specific settings on which we are going to compare the free variant results to the baseline variant results, thus each is composed of two variants.

The settings that we explored are the following and are represented at Figure 4:

- **experiment:** instance label that declares the settings and differentiates that specific instance from the other 18 instances.
- **start date**: establishes the date where the instance will begin.
- **end date**: ending date of the instance.
- **periods**: number of periods (months) between start and end date.
- **fix prices**: Given the high prices variability from period to period, exploring this setting allows us to assess impact without the influence of price variations. This is done by

keeping the selling prices constant at the initial period through the whole experiment.

- **MILP free objective func:** objective function solution's value for the model that is not restricted with the business heuristic constraints.

- **lp heur objective func:** objective function solution's value for the baseline model.

- **relative impact:** relative change or percentage difference between the free solution compared to the baseline restricted solution. This shows how much relative economic value the unrestricted solution has obtained compared to the other solution. The formula is defined as:

$$Relative\ Impact = \frac{\sum_{t \in \text{periods}} Objective\ Free_t - \sum_{t \in \text{periods}} Objective\ baseline_t}{\sum_{t \in \text{periods}} Objective\ baseline}$$

Minor settings: the following three minor settings apply to some of the instances and are specified on the instance name. This settings are:

- **Forecasted prices**: instances named with 'fcst' utilize forecasted prices as part of the data inputs from January and later. To forecast prices we used a detrended Holts Winter model (no trend). This allows us to run experiments on more periods beyond our existing price data.

- **fix_sales_cost_100** refers to instances where sales fix cost is increased from 10 to 100. This scenario could represent an increase in cattle handling and administrative costs.

- **P_index** modifies the pregnancy index from 86% to 70% representing a case where breeding cows is harder. This would impact required transfers to comply with business sustainability restrictions among other factors.

Remember that the only difference between both models is that the baseline model is applying the three extra constraints that reflect the way that decisions are manually taken by the business owners, namely from all the female newborns:

- 30% is sold at Stage 1,
- 30% is sold between Stages 2 and 3,
- 40% is transferred to reproductive stock when old enough.

Benchmark results

experiment	start_date	end_date	periods	fix_prices	lp_free_objective_func	lp_baseline_objective_func	relative impact
2019_24periods	18/01/2019	08/01/2021	24	FALSE	138,960	135,881	0.023
m2019_24periods	07/06/2019	03/06/2021	24	FALSE	214,763	174,783	0.229
2020_24periods	03/01/2020	06/01/2022	24	FALSE	155,077	147,267	0.053
2019_36periods	18/01/2019	06/01/2022	36	FALSE	180,379	157,766	0.143
2019_42periods	18/01/2019	02/06/2022	42	FALSE	275,843	201,437	0.369
2019_48periods	18/01/2019	05/01/2023	48	FALSE	239,279	206,043	0.161
2019_24periods_fix_prices	18/01/2019	08/01/2021	24	TRUE	276,975	245,771	0.127
m2019_24periods_fix_prices	07/06/2019	03/06/2021	24	TRUE	419,392	341,765	0.227
2020_24periods_fix_prices	03/01/2020	06/01/2022	24	TRUE	386,107	358,127	0.078
2019_36periods_fix_prices	18/01/2019	06/01/2022	36	TRUE	383,903	313,586	0.224
2019_42periods_fix_prices	18/01/2019	02/06/2022	42	TRUE	492,248	367,824	0.338
2019_48periods_fix_prices	18/01/2019	05/01/2023	48	TRUE	506,185	395,788	0.279
2019_120periods_fcst	18/01/2019	08/01/2029	120	FALSE	442,013	313,385	0.41
2019_120periods_fix_prices_fcst	18/01/2019	08/01/2029	120	TRUE	2,377,258	1,191,302	0.996
2023_24periods_fcst_no_trend	05/01/2023	05/01/2025	24	FALSE	49,024	46,298	0.059
2023_48periods_fcst_no_trend	05/01/2023	05/01/2027	48	FALSE	93,886	73,026	0.286
2023_72periods_fcst_no_trend	05/01/2023	05/01/2029	72	FALSE	134,147	94,211	0.424
2019_24periods_p_index_70	18/01/2019	08/01/2021	24	FALSE	134,479	131,268	0.024
2019_24periods_fix_costs_100	18/01/2019	08/01/2021	24	FALSE	136,680	133,469	0.024

Figure 4: Instance settings with the obtained results. Own elaboration, final objective function values rounded to thousands.

In the next sections we will explore the relationships between existing decision variables and described experiment settings. The goal is to gain a general understanding on how these diverse scenarios are impacting obtained solutions. During the optimization, the solver explores a search space ranging from 21.497 variables in the 24 period experiments to 104.057 variables in the 120 period experiments.

The plan is first to confirm that the model is behaving reasonably according to business criteria and not obtaining revenue by exploiting a flaw in the model definition. During the course of this project, considerable effort was invested in adjusting and fine-tuning the model to rightfully represent the business problem.

Afterwards, to validate the model and gain deeper knowledge on the inner workings of the proposed solutions we would want to examine decision variables separately and then investigate associations and correlations between them. Identifying the underlying logic guiding the solver into capturing theoretical revenue while also detecting the main source or sources is key. This will allow us to establish model competence to theoretically generalize and adapt to multiple settings. This is the first step into building a flexible tool that business decision makers can use to assist them to improve their operation.

To facilitate the reading of the analysis, unless otherwise specified, we maintain the comparison relationship from the standpoint of the free variant. For example if we state a 20% increase in a particular variable, this means that the free variant is proposing a 20% increase compared to the baseline one.

Objective function analysis

The value of the objective function describes the revenue obtained in the exercise, expressed in US dollars. Within each one of the 19 instances, each variant (baseline and free) will output one objective function value. In this section we will explore the relative difference between them.

The results reveal a minimum positive impact of 2% and a maximum of 96% with a median value of 22.4%. Excluding the outlier exhibited on the “2019_120periods_fix_prices_fcst” instance, we can interpret that the positive impact revolves around 0% to 40% as described in Figure 5-1.

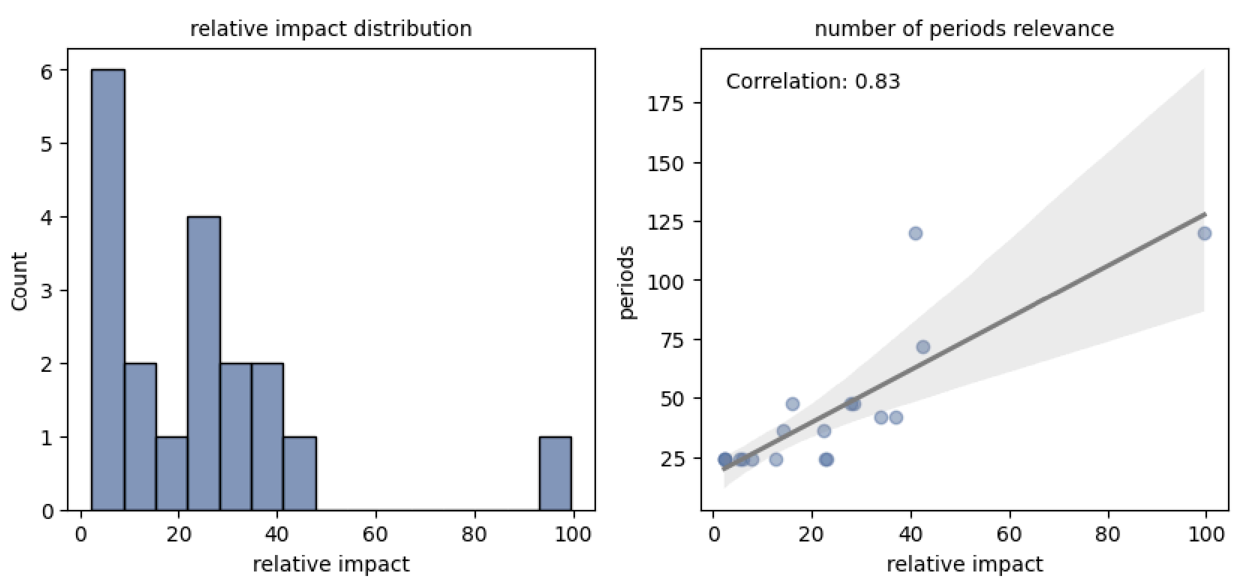

Figure 5-1: Relative impact distribution. Figure 5-2: Relative impact distribution versus instance total periods. Own elaboration

The outlier case "2019_120periods_fix_prices_fcst" achieves 99% relative impact because settings fix prices in January 2019, which maintains higher relative prices compared to the variable price trend shown in Figure 2. This allows optimization over a longer duration (compared to other experiments) and at a higher price, leading to increased revenue. This increase is not reflected in the other 120 period variant with variable prices, which only reaches a 41% relative impact.

Evidence shows that longer runs appear to establish bigger impact as shown on figure 5-2, with a Pearson correlation coefficient between relative impact and number of experiment periods of 0.83. To understand how this difference operates over time, Figure 6 contains one line plot per instance showing the objective function absolute difference cumulative sum between variants. The line over zero means that the free variant is accumulating a higher value at that specific period on the horizontal axis. It can give us a sense on how each variant competes period to period to maximize revenue.

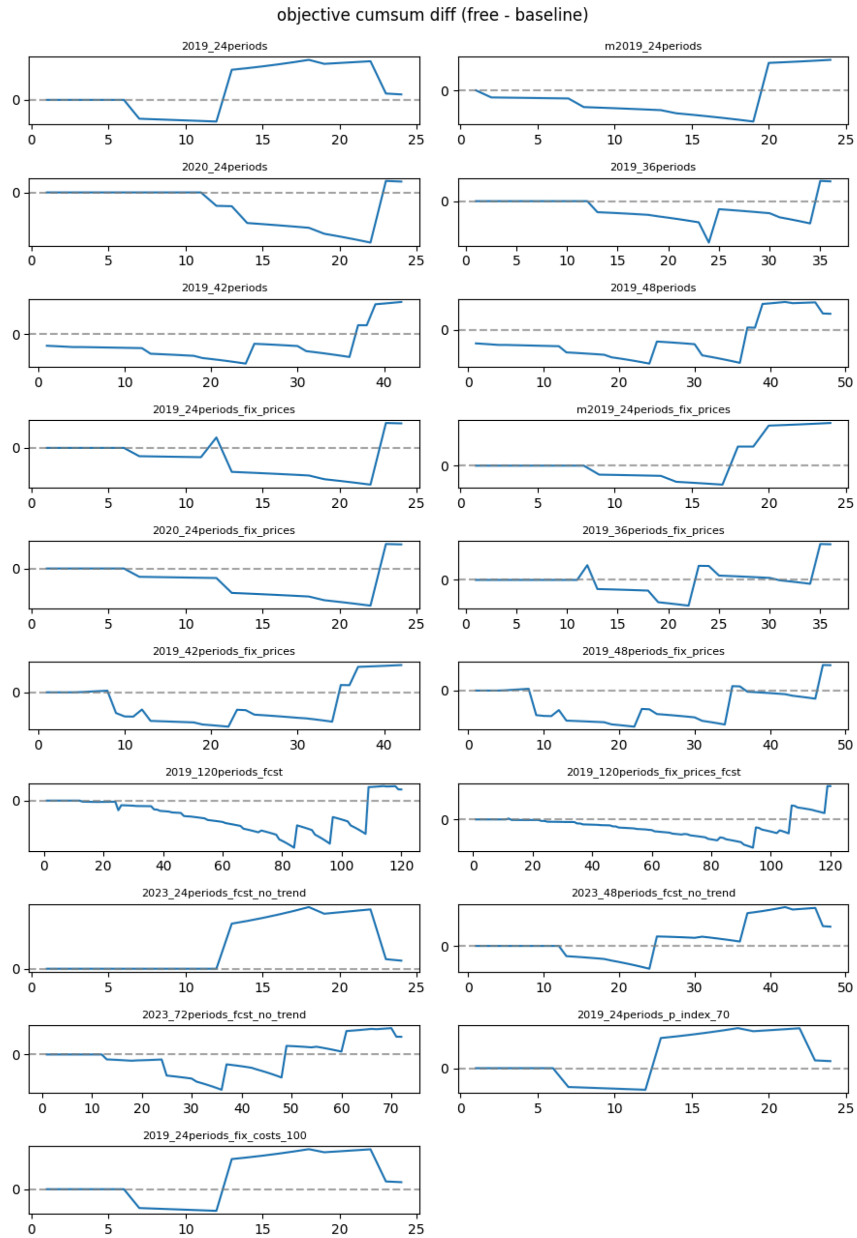

Figure 6: Objective function cumulative absolute difference between variants per instance through periods. Own elaboration

Except for the instance "2023_24periods__fcst_no_trend", there is always a subset of periods where the baseline variant accumulates more revenue, to be later surpassed by the free variant. This indicates a pattern of delaying present revenue into later periods.

Revenue difference emerges during the second half of all instances. The transition from negative to positive occurs mostly on the last 20% of the periods, as Figure 7 shows. This highlights that these periods are highly relevant in shaping the global impact within each instance. We can conclude that free variants are accumulating less revenue during the first half of the experiments, which is then surpassed by the revenue they make in the second half, making a positive difference over the baseline variants. This is probably related to higher upfront costs related to increasing reproductive stock towards a higher return rate in future periods.

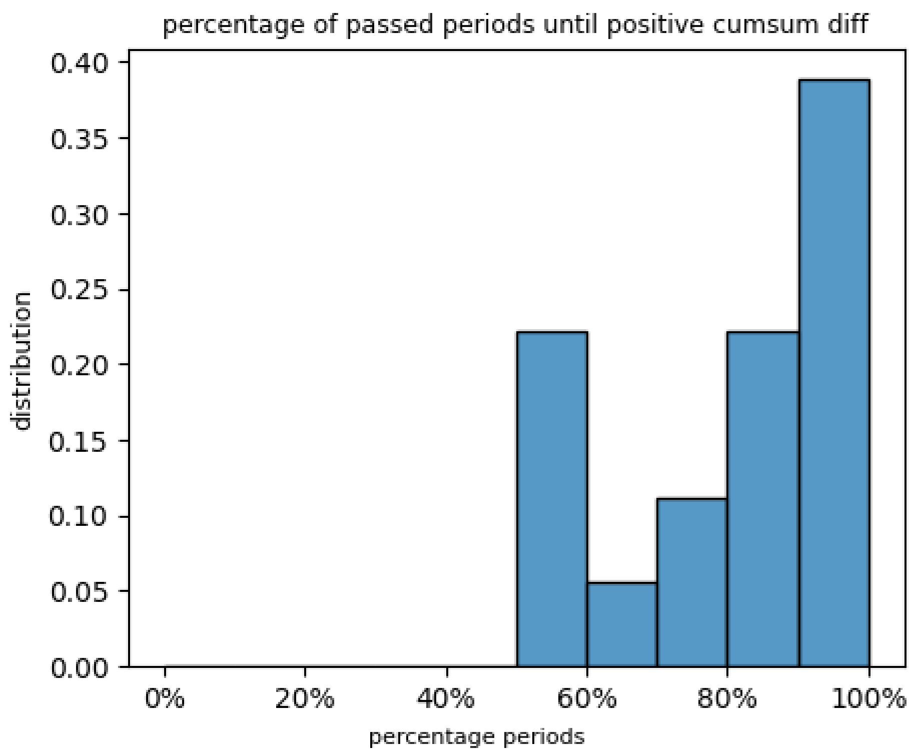

Figure 7: Percentage of periods passed until positive advantage obtained from free variant. Own elaboration

Sales analysis

In this section we will explore how sales are solved for. Considering we can sell at four different stages across the experiment duration, on Figure 8a and 8b we can examine which stage was preferred between variants. The second Figure isolates the 120 periods instances to better represent the dimensions. We calculate the difference between each selling stage for each instance. Bars towards the negative x axis values, represent more sales of the free against the baseline for that stage.

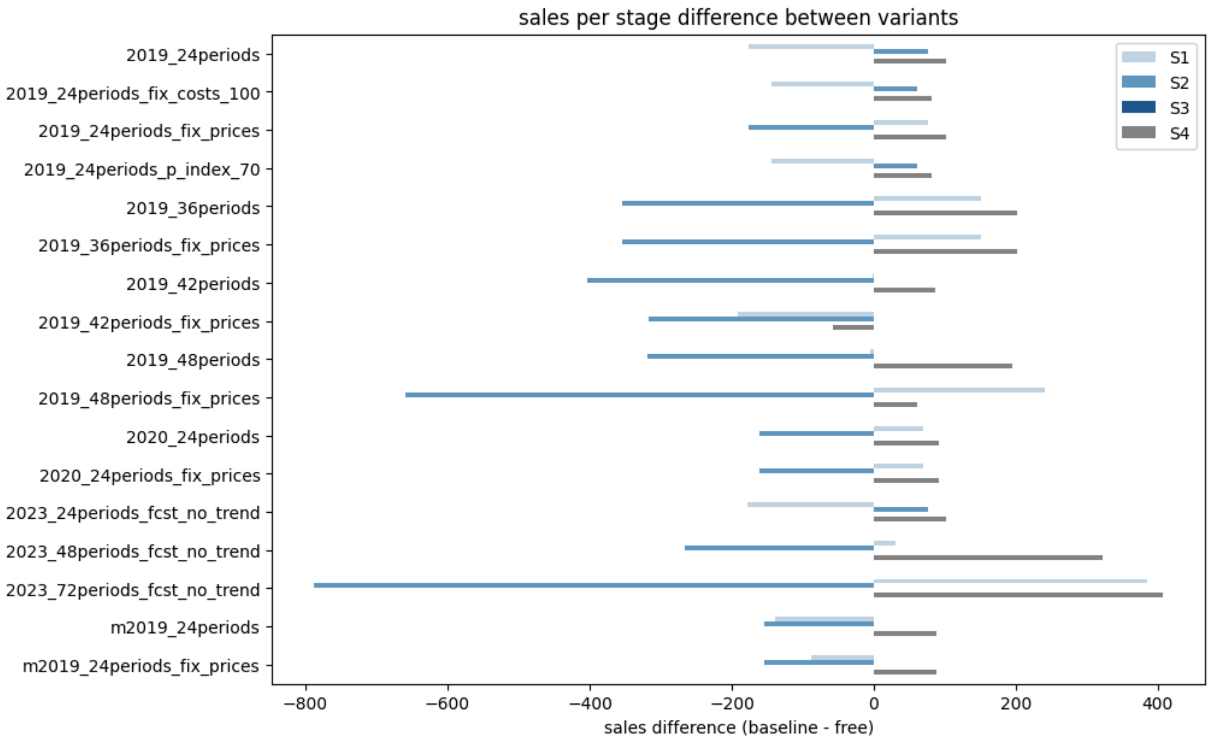

Figure 8a: Sales change per stage between variants for each instance. Own elaboration

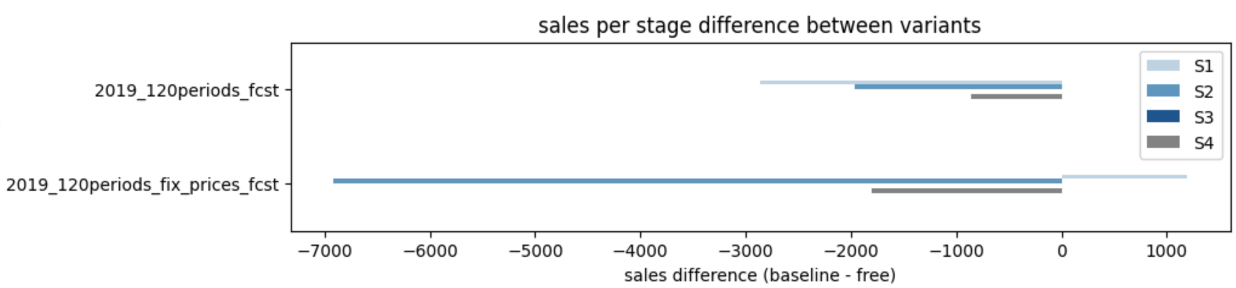

Figure 8b: Sales change per stage between variants for each instance. Own elaboration

There is a general exchange between selling more at Stages 2, and 3 with a counterpart of selling less at Stage 4. Stage 1 sales vary depending on the instance. This means that the baseline variant is selling more reproductive stock (at a lesser price) than the free variants.

This can be explained because heuristic rules consistently assign 30% of female newborns to reproductive stock. This stock exceeds the requirement to at least maintain the reproductive stock given in the initial period by the end of the instance. This excess enables sales of reproductive stock to increase revenue (even at a reduced price compared to the other stages).

There are two instances that don't follow this rationale and make more Stage 4 sales, both being the 120 period instances described on Figure 8b. Apparently with enough reproduction cycles (one per year) available during the instance duration, a positive impact in revenue can be obtained by increasing reproductive stock as it increases future sales. This potentially compensates revenue loss for reduced price for reproductive stock at Stage 4.

In summary, there is a clear difference between how sales are conducted. First conclusion being that Stage 4 sales are lower except on the two longer variants. The second conclusion is then that instance duration potentially may impact sales composition.

Transfer analysis

In this section we will explore transfers. Meaning, the quantity of Class 2 (female) stock that is transferred into Class 3 stock (reproductive). This change can be done at the 11th month of age only and is irreversible.

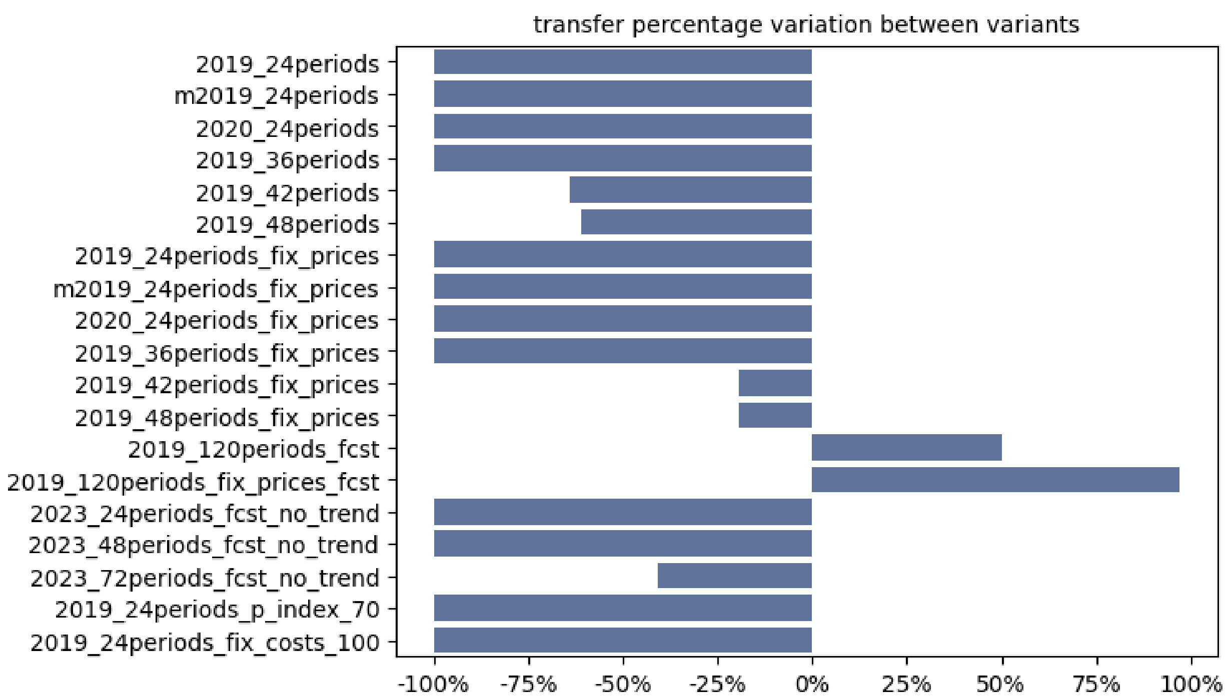

Figure 9: Transfers change between variants per instance. Own elaboration

As described on Figure 9, We can see that most variants show a decrease in transfers ranging between 25% to 100%. In the later cases where no transfers are performed, this must be aligned with no Stage 4 sales to fulfill the sustainability restriction. In correlation with the sales analysis, the only two instances that show increased transfers are the longer 120 period variants, from a 50% increase to doubling the number of transfers.

In Figure 10 we explore the relationship between transfers percentage variation and number of periods. We can establish that more periods are correlated to more transfers. The Pearson correlation coefficient between both variables is 0.9263. As mentioned before, the outlier 120 periods instance “2019_120periods_fix_prices_fcst” that is proposing a 97% transfer increase from baseline is being motivated by higher fixed selling prices, proposing even more transfers than the non fixed variant that proposes a 50% increase.

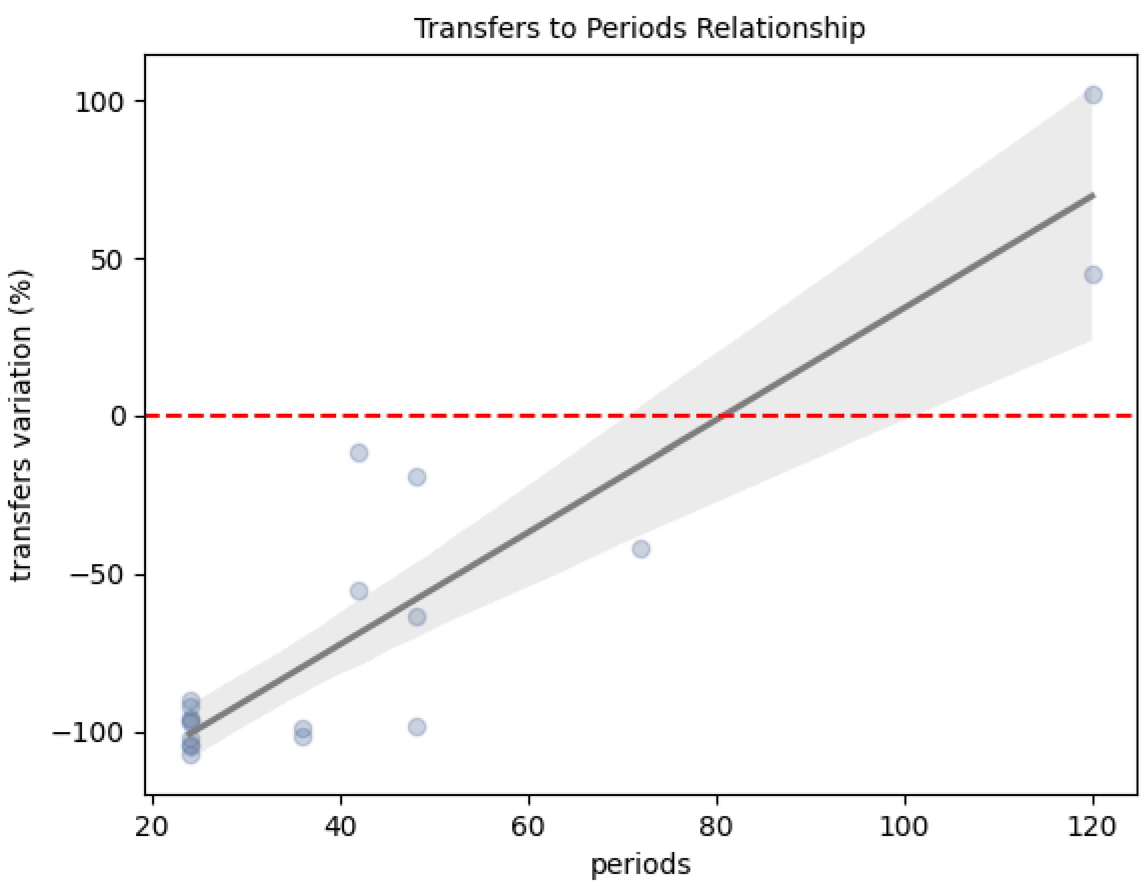

Figure 10: Transfer variation at different instance duration. Jitter added on the vertical axis in order to improve visualization.

We start to see a clear pattern and one of the key findings. Transfers and sales are strictly correlated and are both affected by the variant length.

In the next section we will make a deep dive on one instance from each group to better understand the logic behind proposed solutions. One group being shorter period duration instances where the free model did not increase reproductive stock (17 out of 19) and one of the other two instances where this does take place.

Analysis of specific instances

Instance m2019_24periods analysis:

In this section we will review a 24 period duration instance that starts in mid 2019. The idea is to visualize how each variant was solved for and detect key differences. Particularly we are selecting one of the shorter period set of instances (17 out of 19) where the free variant is making fewer transfers than the baseline one.

Figure 11 compares the free variant above to the baseline variant below. Barplots height represents stock maintained per period, horizontally divided in three blocks, one per class. The color darkness represents stock age. This plot gives us a perspective on how stock decisions are made period to period for each class.

Figure 11: Stock per period, class and age per variant for instance “m2019_24periods”.

This instance shows two interesting aspects that are worth mentioning. The first one being that the baseline variant is ending with unsold Class 2 stock, which is a loss of potential revenue. This can be explained by the newborn sales composition restrictions that the baseline variant enforces.

In this instance we have two breeding windows at period 2 and 14. All generated newborns from the second breeding period are sold at Stage 1 at the free variant as there is not enough time to reach any other selling stage, regardless of the profit that each stage could generate. The baseline variant is only allowed to sell 30% of newborns at Stage 1, ending with 155 units of unsold Class 2 stock. This generates a profit loss to the baseline variant compared to the other variant.

The second relevant aspect that stands out is the baseline variant selling 88.58 units of reproductive Class 3 stock at the second period. Why not wait until the second enabled breeding window at period 14 to make those sales and generate more newborn stock?

We can perform the following marginal analysis for increasing one unit of Class 3 between periods 3 to 14 to reach the breeding window:

- $C1n\_m$ Class 1 newborn maintenance cost between periods 14 to 20 for ages 0 to 6.
- $C2n\_m$ Class 2 newborn maintenance cost between periods 20 to 24 (unsold stock)
- $C3\_m$ maintenance cost for Class 3 between periods 13 to 20.
- $C1n\_p1$ selling price for Class 1 newborn at Stage 1.
- $C2n\_p1$ selling price for Class 2 newborn at Stage 1.
- P_ind pregnancy index at 86%
- 50% are born female, 50% male.

$$Maintenance\_costs = C1n\_m \times 50\% \times P_{\text{ind}} + C2n\_m \times 50\% \times P_{\text{ind}} + C3\_m$$

$$Selling\_income = C1n\_p1 \times 50\% \times P_{\text{ind}} + C2n\_p1 \times 50\% \times P_{\text{ind}} \times 30\%$$

$$Selling\_income < \text{Maintenance\_costs}$$

The cost to maintain potentially generated newborn stock and higher Class 3 maintenance costs until period 14 is higher than newborn’s selling value at Stage 1, for 30% Class 2 and 100%

Class 1. In conclusion, it is better to not breed stock at all if the variant is only allowed to sell 30% of female stock, as the maintenance costs are higher than the revenue obtained.

In conclusion, two sources of increased revenue are found that give the free variant means to reach a 22.9% relative revenue increase. One being the higher flexibility regarding how to allocate newborns sales on different selling stages. The second factor being flexibility on transfer management and business sustainability, represented in maintaining reproductive stock overtime while optimizing revenue.

We could make heuristic restrictions smarter to avoid decisions like holding up stock for a selling stage that is non-existent within the time period of the experiment to comply with one of the heuristic restrictions. This is unrequired as this will be addressed in the longer instances analysis. It is also unneeded as enhancing the baseline variant is not a project goal but rather a step towards validating and improving the main free variant model. In summary, the free variant final solution is described as follows:

- zero transfers
- Stage 1 sales 13,17% increase
- Stage 2 sales 53,37% decrease
- Stage 3 same amount (46 units)
- Stage 4 zero sales (88.58 units on the baseline variant)
- births 9.4% increase
- revenue increase 22.9%

Instance 2019_120periods_fcst analysis:

This longer duration instance is composed of 120 periods, 10 of them being breeding periods. Contrasting to the other group of 17 shorter period instances, this one represents one of the two cases where transfers are increased. Same as before, we explore in Figure 12, stock per period, for each class.

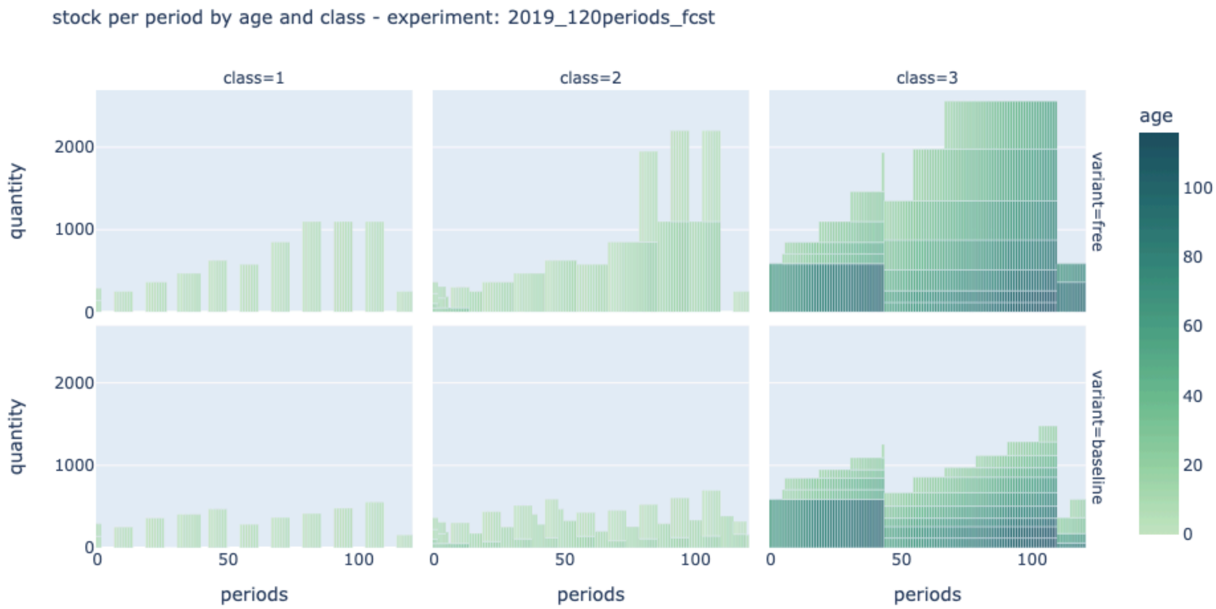

Figure 12: Stock per period, class and age per variant for instance "2019_120periods_fcst".

Both variants exhibit an increase of Class 3 stock over time. Eleven periods prior to the instance ending at period 109, both variants sell Class 3 stock, falling back to initial stock levels to comply with the sustainability restriction. The free variant solution obtained a 41.1% revenue increase and proposes an increase in transfers of 50.42%.

A primary distinction is at the phase on which transfers are allocated. baseline variant steadily increases reproductive stock by following up with the heuristic rules, assigning 40% of newborn stock to Class 3 every year. The free variant surpasses this ratio by rapidly increasing Class 3 allocation in detriment of Class 2 earlier sales.

The free variant continues transferring stock from Class 2 to Class 3 until period 66. Class 3 quantities remain constant until period 109 at which point, 61% of total stock is sold. We could ask why it is optimal to sell Class 3 stock on period 109 if there is one more last breeding window at period 115. The answer lies in the fact that newborn generated stock for that breeding window hits age of 5 by the last period. This is one period away to reach selling Stage 1 which occurs between the ages of six to eight. This makes breeding more, not optimal on the last window. Some amount is proposed by both solutions as a certain amount of Class 3 is required to match initial Class 3 stock to comply with sustainability restrictions. The baseline solution performs more breedings on this last window provoked by a more restricted model regarding transfers.

In Figure 13 we report sales per period for both variants. Barplots height represent quantity of sold stock per period, horizontally divided in three blocks, one per class. The darkness represents age.

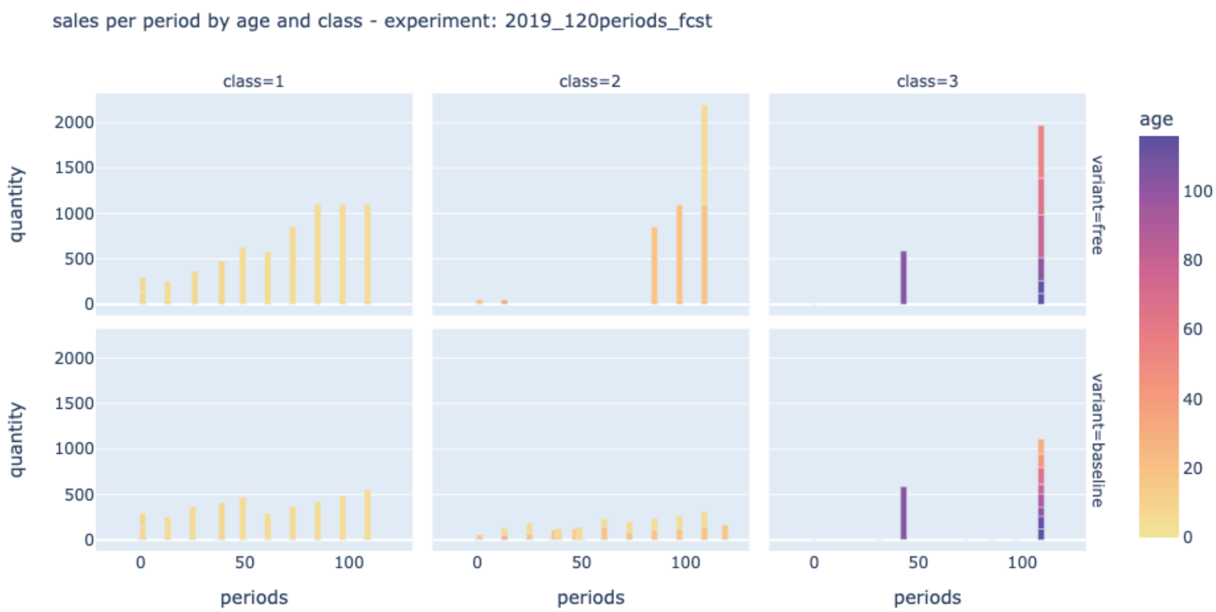

Figure 13: Sales per period, class and age per variant for experiment “2019_120periods_fcst”.

Class 2 free variant sales are deferred towards earlier transfers, proposing no sales between period 13 and 85, this choice accelerates the rate on which reproductive stock increases. During the last periods the free variant sales outperforms and yields a significant revenue difference compared to the baseline variant.

Class 1 sales quantities increase year to year for both variants. Higher quantities of maintained reproductive stock also leads to higher male newborns, subsequently translating into more sales.

In summary, contrasting the analysis from shorter duration instances, there is a shift in the solution proposed strategy. Longer instances are giving enough time to redeem the investment made on reproductive stock and capture long term revenue by performing later higher sales. This is one example of the model's capacity to adapt effectively in response to various settings and input variables, showcasing its flexibility to respond to diverse proposed scenarios.

The turning point into making this shift could be generated by a combination of factors. Increasing selling prices, cost reduction, a higher pregnancy rate or a combination of all of them with other inputs can directly reduce or increase the number of periods into reaching this strategy shift. Is worth mentioning that stock accumulation could be potentially restricted by infrastructure or labor availability. This and other aspects will be mentioned in the conclusions section.

The free solver solution proposes:

- transfers 50.42% increase
- Stage 1 sales 57,56% increase
- Stage 2 sales 173,62% increase
- Stage 3 same amount sold
- Stage 4 sales 50,42% increase
- births 78,19% increase
- revenue 41.1% increase

Variant duration analysis

In this section we aim to explore the relationship between instance duration and transfers, seeking to identify the occurrence of this strategy shift that lies between short and long period instances. To perform this analysis we will re-run the same instance multiple times, adding each run 10 extra periods, increasing the time horizon and exploring how solutions are shaped.

We selected the starting period in January 2019. To remove the impact of price fluctuation over time, we will also run one set of instances with fixed prices. The search space generated is defined from 24 to 120 periods with steps of 10 periods, totalling 18 instances, nine with fixed prices and nine with variable prices.

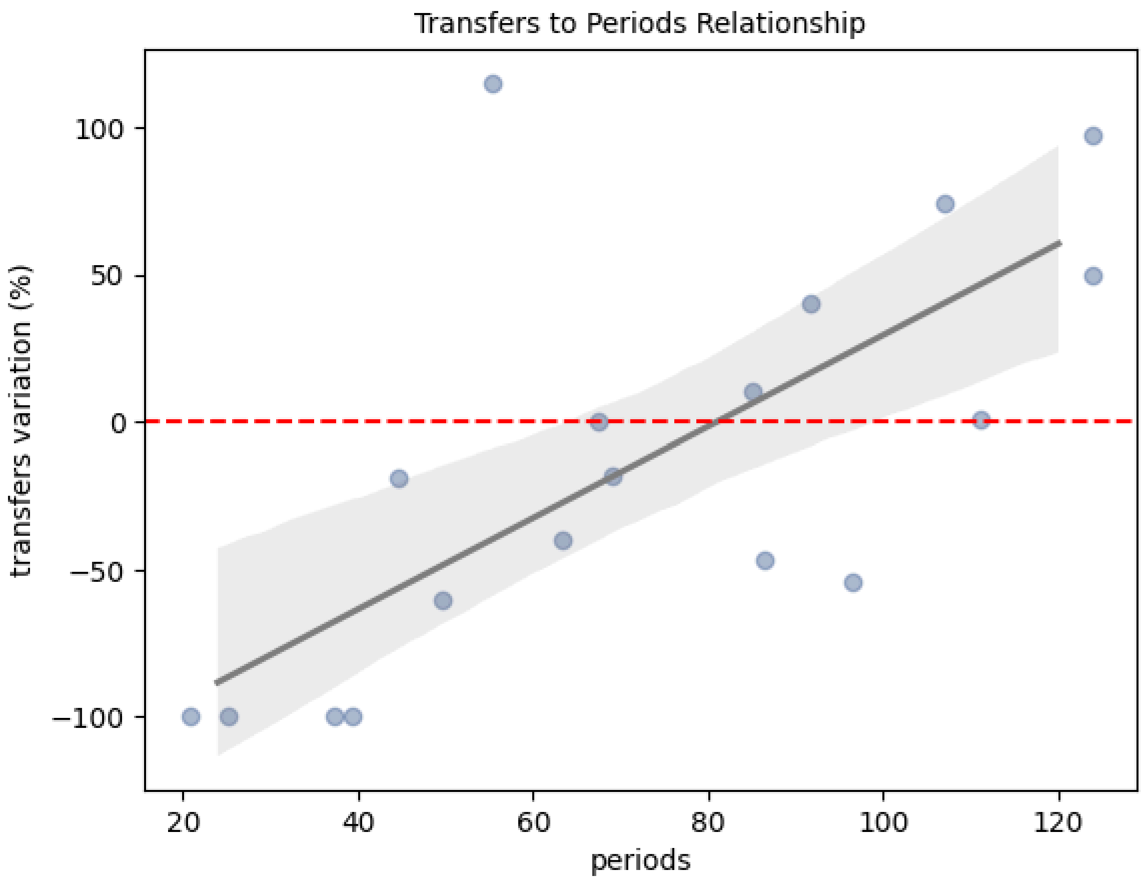

Figure 14: Transfer variation between free and baseline variant per instance number of periods. Jitter added on the X axis to avoid overlap between points.

In Figure 14 there is a noticeable trend where longer instances correlate to more transfers, with a positive Pearson correlation coefficient of 0.72. Based on given settings and inputs, it appears that the strategy shift tends to occur around the 80-period mark.

The Pearson correlation within groups of fixed and variable prices over each period is 0.9721 for fixed price instances and 0.4846 for variable prices instances. This indicates that our reasoning holds a strong and direct correlation while being representative for simpler instances. However, when introducing complexity of real-world scenarios with changing prices, costs or any other setting, the influence of instance duration diminishes over a broader relevance of a bigger picture regarding inputs.

We conclude that even when instance duration is a key parameter, it is not the only relevant parameter to shift the solution towards increasing earlier reproductive stock levels. This also supports the notion that operation research approaches, as analytical solutions, possess the advantage of consuming large amounts of data inputs. The intricacies of the solutions proposed can escape the scope of human made analysis achievable through single or multi variate exploration.

IV. Conclusions

We conducted a benchmark analysis for a set of 19 instances, exploring how solutions proposed solve for a diverse grid of configurations. Afterwards, we ran a deep case analysis over two instances that represent two different kinds of solutions. The following conclusions were obtained:

- Positive impact obtained is between 2% and 40% with a median value of 22.4%, excluding an outlier instance of significant higher revenue.

- Prioritized selling stages between variants exhibit high variability based on inputs showing model flexibility upon given inputs.

- As the number of periods increases, greater revenue impact can be achieved.

- baseline levels of reproductive stock allocation is exceeded only in longer instances as time is required for the investment to yield benefits.

- Instance duration plays a crucial role towards optimal solution being to exceed baseline levels, however other relevant model inputs such as price fluctuations directly influence when this shift occurs.

Evidence indicates that the analytical solution based on optimization techniques can successfully abstract the main dimensions of the business problem presented. Solutions obtained show high adaptability to changing input variables, demonstrating versatility while utilizing explainable strategies that can be used to provide insights to the decision maker into improving their operation. Complemented with data visualization tools the solutions can be interpretable improving the comprehension and explainability of the outputs.

Upon these findings, spending resources and time to refine the model and build an integration as a prescriptive solution emerges as a promising business opportunity and satisfies the objective of this thesis.

Valuable business dimensions for next iterations

Due to time constraints, some valuable dimensions or caveats worth considering were left for future iterations.

Grass feed mixture utilized to feed cattle is produced locally. When depleted, it is then bought from other producers. Outsourcing is generaly more expensive and the price can increase more if the market supply is low due to scarce rains or hail falling during the previous harvesting season. This can impact breeding costs and is usually associated with selling at earlier stages of the life cycle (Stage 1) rather than later stages. Opting to reduce production and utilize only inhouse grown grass feed reduces price uncertainty but could also negatively impact revenue. Including grass consumption, availability and outsourcing prices would provide a more realistic cost dynamic for the breeding operation.

Decisions and restrictions related to how many **bulls** are required and their cost to satisfy breeding for all the female class is not being modeled. We consider that modeling bulls wont pay off the increase in complexity for adding a fourth class to the model or adding more restrictions to describe this. This could change if we want to consider multiple species where this variable could gain relevance.

Costs related to infrastructure and hand labor adapting according to the stock being maintained in different age groups. Data related to how costs and labor increases per operation size was not available but would be relevant to gather in future research.

Local and foreign government policy is constantly changing, affecting how agents decide and project business expected revenue, both short and long term. Examples of government policies implemented in the last ten years range from variable taxes, enforcing a selling quota to a specific selling stage. Depending on the political orientation of the ruling government, drastically different scenarios are expected to happen given the polarization in both parties related to commercial policies. These can be a lead factor and one of the central factors in the decision making. Incorporating these externalities into a framework that describe how different scenarios could impact the operation theoretically could be valuable to consider.

Design **simulations** representing theoretical scenarios can provide valuable insights when validating strategies or measuring impact on possible outcomes. “Experimentation is possible with a model, whereas it is often not possible or desirable to experiment with the object being modeled” (Williams, 2013). For instance, one could consider probabilistic scenarios such as low vs high expected export prices, combined with a possible export quota restriction policy during a particular season or an increase in costs for a specific category to give a few examples. This opens a window of opportunity that is not being attained in this thesis.

Incorporating a money discount rate would represent that present revenue is more valuable than deferred revenue. Currently, we are calculating the sum of the objective function period to period, giving the same economic value to money earned in the earlier periods to the one obtained in the final periods. Adding a money discount rate could refine the model into increasing the relative value of earlier earnings by either a fixed or variable discount rate per period. This way shaping the model into a more business realistic setting.

We could include a **sensitivity analysis** by incorporating the methodology proposed by Hazell (1971) in his work, “A Linear Alternative to Quadratic and Semivariance Programming for Farm Planning under Uncertainty.”

Next steps into business implementation

In this section we discuss proposed steps and considerations around integrating the solution into the company in a way that it can generate a positive lasting impact.

To ensure a proper delivery of our optimization tool and transform it into a real prescriptive solution, it is essential to design an integration plan. Ideally this plan should be made conjunctively between the technical team and the company's main decision makers. This approach helps prevent one-sided implementations and emphasizes collaborative efforts to build trust and visibility from the ground up. It is key for the technical staff to share both potential value and opportunity as well as assumptions made, caveats, and current limitations to align expectations and efforts with objectives.

Considering that the first issue the company raised concerned finding the optimal selling category between stages one to four on a given period of time. While the second one being the need to find optimal balance between increasing reproductive stock vs sales prioritization aligned with different potential scenarios. We can split the workflow into two parts as *sales optimization* and *breeding optimization*.

Sales optimization

A first target would be to solve the simpler question of deciding which is the optimal selling category to aim for, given a set of expected future prices, costs and present stock. Considering at first a short time window of one year. The focus is not on how to optimize breeding but sales in an environment of high price and cost fluctuations, that make it difficult to target the right category during a breeding cycle.

A quick proof of value may be a good idea, showing how the solution optimizes sales for a given period and comparing them to the ones performed by the company owners during the same period. We could explore if they match or not and validate how the optimization captures the higher revenue category or categories for each class per period. Keeping things simple, this may show how we can provide a data driven response to a complex problem.

Afterwards, we could repeat the same experiment but with forecasted prices utilizing a simple model, similar as the one applied during our previous benchmark, repeat the comparison adding the nuances on how the forecasting error impacts the revenue obtained. Emphasizing that model inputs are as important as the model itself and investing time on improving them is part of future efforts.

Breeding optimization

This analysis could be focused on a more extended time frame than the previous one. The final goal being to give data oriented insights towards planning a longer term breeding strategy. One way to do this is by simulating different scenarios, calculating potential revenue given different strategies when applied to different economic settings.

Solving how to plan the operation between economic cycles where commodities prices and manufacturing costs change is complex. By combining a set of pre-defined expected scenarios and studying how the solution optimizes for those scenarios can provide valuable insights.

The first step would be to design a group of possible scenarios as model inputs. This could range between pessimistic and optimistic or design. Also complement with possible government policies that could give either positive or negative impact in different business dimensions. Afterwards, apply the optimization to the different settings and capture insights and theoretical revenue on each scenario. Deciding to go for a specific strategy could work very well in Scenarios A and B but very bad in Scenario C. How possible is for Scenario C to occur given our current knowledge? Such a discussion could be quite valuable.

Performing sensibility analysis, like exploring how much certain variables impact optimal presented strategies, could also be interesting. We could consider, e.g., ¿how much cost increase can we tolerate until the solution takes a specific form or targets another selling stage? ¿How low can the breeding index be while keeping the operation profitable in this particular setting?

Simulating different probabilistic scenarios could lead to calculating expected revenue for different strategies. Choosing the optimal strategy evolves into a combination of business knowledge, economic expectations and the solver solution itself.

Other considerations

During the introduction of the solution a good practice would be to draft a document that describes in a very succinct non-technical manner, main aspects of the solution that we could use as an initial stepping stone towards presenting the model to the company management.

First describing model data inputs requirements, most important model and business assumptions, main data transformations and limitations. Then, it is recommended to align a list of future improvements and validations that can be made. This document would be a user-friendly guide to allow the non technical user to build ownership and increase adoption.

Future improvements should always consider added value against complexity and effort required to implement. First features to improve should be the “low hanging fruit” that are easy to build and highly valuable. After such a first step, we could decide then if the best path is to go for the big bets that have high value and high complexity, or to solve small features that are less complex while being also less valuable instead.

It could be beneficial to find ways to quickly utilize and validate the model suggestions and apply them on a near future business problem. In particular, it is crucial to avoid lingering into focusing only on model improvement and not working on the model utilization and adoption.

V. Bibliographic references

1. Ras, C. H., Marra, R. M. A., & Tettamanti, M. F. A. (2020). Cambios en el uso del suelo entre agricultura y ganadería en un establecimiento típico en General Pinto, Pradera Pampeana, Argentina, analizado con programación lineal. Agronomía & Ambiente, 40(2), 151-163.
2. Aliano Filho, A., Melo, T., & Vaz Pato, M. (2020). Tactical planning of sugarcane harvest and transport operations. Schriftenreihe Logistik der Fakultät für Wirtschaftswissenschaften der htw saar (Technical reports on Logistics of the Saarland Business School No. 18).
3. Frank, R. G. (2021). ¿Cuántas personas puede alimentar la región pampeana? Anales de la ANAV, LXXII, 72-83. ISSN: 0327-8093.
4. Campanella, S. R., Corsano, G., & Montagna, J. M. (2018). A modeling framework for the optimal forest supply chain design considering residues reuse. Sustainable Production and Consumption. Elsevier B.V. ISSN: 2352-5509.
5. Gradowczyk, M. H., Jacovkis, P. M., Freisztav, A. M., Roussel, J.-M., & Tabak, E. G. (1990). Water resources optimization—simulation in Argentina. European Journal of Operational Research, 49(2), 247-253.
6. Williams, H. P. (2013). Model Building in Mathematical Programming. John Wiley & Sons.
7. Paul Goodwin, 2010. "The Holt-Winters Approach to Exponential Smoothing: 50 Years Old and Going Strong," Foresight: The International Journal of Applied Forecasting, International Institute of Forecasters, issue 19, pages 30-33, Fall.
8. Hazell, P.. 1971. A linear alternative to Quadratic and Semivariance Programming for Farm Planning under Uncertainty. American Journal of Agricultural Economy. 53: 53-62.