Simulation Modelling Practice and Theory 132 (2024) 102897

Contents lists available at ScienceDirect

Simulation Modelling Practice and Theory

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

A multi-method simulation model to investigate the impact of sunflower seed segregation on silos

Louise Coetsee *, Wilna L. Bean

Department of Industrial and Systems Engineering, University of Pretoria, South Africa

ARTICLE INFO

Keywords:
Agent-based simulation
Discrete-event modelling
Mixed-method modelling
Bayesian networks
Agricultural supply chain management
Silo

ABSTRACT

The South African sunflower industry is considering transferring to a quality-based marketing system driven by an incentive. However, the ability of silos to offer necessary segregation services is critical in such a transition (Baker et al., 1997) and the silo industry is concerned about the negative impact segregation could have on operations and finances. This paper proposes a multi-method simulation approach to quantify the impact of quality-based segregation and sunflower farmer response to the incentive on silo bin utilisation and the ability of the silo to store contents of arriving trucks (service level). A combination of agent-based simulation, discrete event simulation and Bayesian network sampling is used to capture system behaviour where data is scarce. Therefore, in this study, a mixed methods ABM and DES model is implemented in a new environment: a grade-based segregation problem in the South African silo industry. Several scenarios are modelled to cross-validate methods and to tease out the impact of farmer response on the results. The model is applied to a case study silo complex to test the concept. Results obtained for the case study silo show a significant negative impact on costs due to lower service levels and bin utilisation, incurring relocation and opportunity costs. Overall, this study highlights that it is necessary to consider the impact that sunflower segregation could have on each unique silo complex and provides a method to quantify the stated impact.

1. Introduction

Key actors within the South African sunflower value chain must consider whether it can transition from a commodity-based to a quality-based marketing system to increase its overall competitiveness. In the quality-based system, farmers can earn a premium for sunflower seeds with high oil content (seed oil content above than 38%). The ability of silos to segregate grains and oil-seeds based on quality is critical in this transition [1]. The benefits of an incentive pricing structure for farmers and crushers have been proven [2], but there is uncertainty regarding the magnitude of the impact on the silo industry, and therefore its ability to perform the necessary services.

Until now the South African storage infrastructure was not set up to categorise and store different grades of sunflower seed [3]. South African silos could not segregate sunflower seeds based on oil quality mainly due to (1) the high cost of equipment and (2) the tedious and time-consuming procedure necessary to determine oil and protein content. However, new technologies emerge, making silo operations more efficient and allowing a competitive advantage at silos with the niche ability to identify these characteristics, possibly also allowing silo owners to minimise the risk of seed contamination and tap into financial incentives [4,5]. Some silos

* Corresponding author.
E-mail addresses: u17027226@tuks.co.za (L. Coetsee), wilna.bean@up.ac.za (W.L. Bean).

https://doi.org/10.1016/j.simpat.2024.102897
Received 11 June 2023; Received in revised form 22 December 2023; Accepted 8 January 2024
Available online 9 January 2024

in key sunflower production regions are currently equipped with near-infrared (NIR) spectroscopy equipment and can, therefore, rapidly grade each incoming truckload based on oil content.

Silo segregation problems, within the context of an Agricultural Supply Chain (ASC), have been solved with different modelling techniques [1,6,7]. In supply chain literature, simulation modelling has become the preferred tool to deal with high complexity and uncertainty [8]. Agent-based modelling (ABM) shows additional benefits for analysing the impact of policy intervention over traditional analytical methods [9]; however, in silo analysis, most segregation models have been focused on Discrete Event Simulation (DES) [1,6,10].

This paper proposes a mixed-method simulation approach in order to investigate the impact of sunflower seed segregation and farmer response to the incentive on silo service level, bin utilisation and costs within the South African sunflower and silo industries. The paper's main contributions include (1) the use of ABM to solve an ASC silo segregation problem in the South African silo context, (2) integrating an expert-driven Bayesian network in agent behaviour to overcome a data scarcity constraint, and (3) a method for silo segregation analysis within the sunflower seed and South African contexts, whereas the use of ABM is a novel approach to solve silo segregation problems, and, this study is the first to consider silo crop segregation within the South African context.

The remaining content of this paper is structured as follows: the following section discusses segregation problems and their solutions in literature, as well as theories for the behaviour of critical agents at silo complex in the context of silo segregation problems. In Section 3, the details of the simulation model method are presented, the model case study is discussed, and its implementation is validated. Next, Section 4 presents the results of the case study silo. Then, Section 5 discusses the model and its results. Finally, the work is concluded, and research opportunities are highlighted in Section 6.

2. Literature review

Quality characteristic based segregation at silos is an impotent concept, as it offers potential for financial growth at silo and supply chain level, however, at a set of risks and costs [6]. Key concerns and their related costs have been quantified by various methods, which will be covered in the first review section. Secondly, examples of farmers and silos in simulation models are reviewed to understand theories of these individual agents in the existing literature. Creating a digital twin for experimentation on a system in a virtual environment can be data intensive. Therefore, to address this potential constraint, Bayesian networks for governing agent behaviour is also reviewed.

2.1. Silo segregation problems in literature

Limited silo segregation problems are discussed in literature, with many papers restricted to logistics and Supply Chain Management (SCM), excluding physical and chemical segregation studies. Although some similarities are apparent in logistics and SCM silo segregation models, model variables and outputs vary to serve the unique requirements of each study [1,4,6,9]. Models developed to quantify the impact of segregation of a crop at a silo includes empirical process models to test the feasibility of isolating higher-quality soybeans at a grain silo to be sold at a premium later [4] and DES models to predict the impact of segregating grain at a grain silo [1,6].

Experimentation with real-life operations can be risky and costly as it can adversely impact a company's operations and profitability [11], therefore, the use of a simulation model is proffered. DES models for silo segregation are developed in various research contexts, however, to a lesser degree than some more popular application topics such as manufacturing. For example, a DES model was applied to a case in The United States of America, Kansas, to understand the cost of segregating wheat based on protein content to receive a price premium at a national level [1]. The two DES models developed specifically for quality based segregation impact analyses [1,6] have some shortcomings. Firstly, both model applications only analyse the impact on the receiving operations at a grain silo while ignoring the outflow of grain and, therefore, the storage capacity problem (specifically crucial for the research question at hand). Secondly, these model applications test the feasibility of silos to segregate grains such as soybeans and wheat, which are considered high-volume cash crops [12]. However, sunflower, which must be segregated in the current study, is a much smaller portion of deliveries throughout the harvesting season than crops such as soybeans or wheat. Thirdly, these model applications assumed truck arrival rates that do not account for the possible complex changes in the event of policy intervention. Fourthly, these model applications assume silo owners will benefit from selling the higher quality crop at a premium. However, in the South African silo context, silo owners often merely offer a storage solution service to farmers, traders and crushers. Finally, these models have not provided solutions to analyse service levels and bin utilisation as this research question requires.

Despite these shortcomings, the application of DES to silo segregation problems demonstrates that silo operations lends itself to DES. DES is described as computerised models consisting of entities that move through a system of queues and activities and resources that are shared between activities [13]. This system description applies to the operations at silos in the context of this problem, given that silo bins (resources) are limited, and the silo observes truck as a queue of arriving agents to be serviced with its resources. Therefore, the DES approach of [6] will be adapted to include incoming and out flowing activities at a specific silo complex of interest. Furthermore, the model will also run over multiple years and the flow of grains and oilseeds throughout the year, whereas existing segregation models only modelled peak seasons of the crop of interest.

To account especially for the direct impact that farmer response will have on the volumes and quality of sunflower arriving at the silo, ABMs are considered. An agent can be defined as a computer system capable of autonomous and flexible activities to meet its design objectives situated in an environment [14]. Intelligent multi-agent simulations consist of agents that can make decisions based on their environment and interaction with other agents. With emerging literature, agent intelligence is being improved by incorporating reinforcement learning, reasoning decision support, and Bayesian networks to improve adaptive agent behaviour. An ABM consists of two elements, namely, (1) a system of interacting agents and (2) emergent properties arising from agent interaction.

ABMs are used to analyse ASC activities [9]. It is specifically used to predict the impact of strategic industry decisions on supply chain nodes [15]. Even though the specific application of segregation within the oilseed industry has not been addressed with the specific methodology, relevant models have been developed for other agricultural value chain complexes within silo grain activities [7,9,15]. Agent frameworks for supply chain management support are proposed in agent-based literature reviews [9,14]. Farmers are the most frequently modelled agents in ASC-ABMs due to the significant impact the first echelon in an ASC has on other actors. Therefore, farmer behaviour must be recognised as critical when assessing policy, technology or environmental changes in agricultural frameworks [16]. Agent-based simulation has been used to test the long-term impact of policy change causing silo segregation activities [7]. For the problem at hand, ABM allows the model to incorporate the complex decisions of a farmer agent that will directly impact silo metrics.

A hybrid agent-based and discrete-event simulation model (multi-method simulation model) is fitting for this research study to quantify segregation cost, service levels and bin utilisation, given the complexity of farmers and the discrete nature of truck arrivals at silos. Although the definition of OR [17] includes the use of scientific tools and techniques to solve problems related to system operations, with a specific interest in optimal management solutions, there is consensus that simulation techniques also fall under OR [18,19].

2.2. Agent behavioural theories

Individual agent types are reviewed to identify methods and tools for model development. The problem at hand consists of the principal agent of interest, silo companies, and secondary agent: sunflower farmers and other crops competing for storage space, such as maize, soybeans and wheat. Therefore, theories of silos in literature are reviewed first.

2.2.1. Silo behavioural theories

In past literature, silos have been modelled with minimal consideration of farmer relationships. For example, silos report random truck arrivals without any reference to the source of the grain [1,6]. However, the interaction between farmer and silo has been modelled [7]. In these models, silo considerations are still purely operational, dealing with the flow of materials, but not with negotiations and influencing source agents.

Industry experts report that trucks randomly arrive at local silos [1,6]. In the rare case that sufficient storage is unavailable for the arriving truck, depending on the silo's policy, the grain is relocated, sometimes at the cost of the silo. Furthermore, depending on silo grading abilities, silos must decide whether or not to grade the seed [7]. After the grain is graded, the grain is stored in the appropriate storage bin until a request for grain dispatch is received. If a dispatch request is received, the silo must once again decide whether or not to grade the grain, depending on legislation and risks associated with the grain type [7].

Despite the high volatility in the agricultural industry, [20] state that annually, the grain received would typically be close to equal to the grain dispatched. Should less grain be dispatched than received for the year (as an odd case), it will balance out in the following year.

The discrete nature of the silo industry compels modellers to believe it lends itself to DES. In specific applications, some factors must be considered at silo level that can only be incorporated in the model using ABM, such as risk perception. However, for this study, the preferred DES method for the arrival of all crops, excluding sunflower seed, is sufficient.

2.2.2. Farmer behavioural theories

Farmer agents are the most represented agent type in ABMs of ASCs [9] due to their importance in the ASC, as all activities are dependent on farmer decisions and actions. Farmer agents have been debated and tested in various model configurations [16,21–26]. Some models consider a farmer agent with simplified behavioural habits and study the relevant change in studied parameters [21,22,24,27], while others consider a farmer as a holistic person, considering net income [23,25,26].

Due to the critical but secondary role of the sunflower farmer in this study, a simplified behavioural model can be used for this research study. Farmer agents, with simplified behavioural habits related to a specific problem, can have fixed elements to an agent with a specific behavioural element evolving. Examples of farmer behavioural aspect models include risk control efforts [7], social networks influencing decisions on the affected farmer [7,16], and crop configuration for profit maximisation [16,22,28]. For this study, these constant parameters include the number of farmers surrounding the silo and the tonnage of a typical truckload delivered at the silo, as is done in an ABM of a silo and surrounding farmers [7]. At the same time, the evolving behavioural element is profit maximisation by changing crop configuration, as the farmer chooses between sunflower cultivars.

2.2.3. Expert driven Bayesian networks

For the farmer agent, the framework designed by Huber et al. [16] is an appropriate methodology to describe sensitive farmer behaviour. The model represents a detailed decision-making ability in farmer agents, who adapt and learn over time. However, the framework requires large data sets to train the algorithm (behavioural heuristics). Therefore, there are constraints to implementing the framework in this study. Firstly, to date, data on South African farmer behaviour is limited. Secondly, farmer behaviour will vary from country to country; therefore, the framework does not allow for significant input from other factors that could influence local farmers.

To address these challenges, the Bayesian-network approach to model agent behaviour is considered [29]. Various farmer agents have been modelled with Bayesian network methodology [24,25,30]. Bayes inference is a method to calculate posterior probabilities given prior likelihoods, based on Bayes theorem, as shown in Eq. (1) [31].

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)} \tag{1}$$

A Bayesian network is represented by an acyclic graph of a set of variables and their relationships. These links denote cause and effect from one node to the other [30]. Moreover, the links are accompanied by the probability of one node affecting the other in the form of a Conditional Probability Table (CPT) [32].

An expert-driven Bayesian network (EEBN) can be created entirely from domain expert knowledge with reasonable accuracy [29]. A domain expert determines both the network structure and the CPT. One of the advantages of using the BNN approach to construct behavioural logic is the knowledge-based approach to constructing expected behaviour when data is scarce and domain knowledge is crucial [32]. For this study, the expert-driven Bayesian network (EEBN) is ideal for the small amount of data available. An approach to create BNN for agents from expert knowledge is proposed by Nadkarni and Shenoy [32]. The causal map approach starts with semi-structured interviews with experts to identify causal relationships in a domain. Nadkarni and Shenoy [32] state that there are some crucial steps in the transformation of causal maps to Bayesian networks. Firstly, a clear differentiation must be made on the interdependence of map variables. In a Bayesian network, if two variables are not linked, they are assumed to be independent of one another, and if they are linked, they are dependent. In interviews, care must be taken to avoid redundant, circular reasoning due to confusion between abductive and deductive reasoning. Instead, the focus should be on the reasoning behind causal statements. Finally, circular relationships must be removed. Circularity involves splitting decisions from time dependency so that the acyclic graph reruns each decision time frame. In some cases, care must be taken to break variables into different time frames, as Bayesian networks cannot manage reciprocal relationships. The four-step procedure is as follows:

1. Elicit data
2. Derive causal maps
3. Modify causal maps to construct Bayesian maps
4. Derive parameters of the Bayesian causal map

A domain expert is interviewed and transcribed in the first step to form a research narrative text. Secondly, the content of the narrative is systematically analysed to represent the narrative in the form of a causal map. Thirdly, biases of textual analysis are removed to represent a Bayesian network. Finally, probability-encoding techniques are used to set the parameters of the Bayesian map.

Various methods to derive prior likelihoods from interviews are discussed [33]. Encoding methodology is based on the cumulative distribution function. Different methods include the P-method, V-method and PV-method, where P represents probabilities, and V represents values. Prior likelihoods are based on expert interviews by using verbal representations of a probability scale, such as very low (i.e. 0–0.2) and very high (i.e. 0.8–1) [31]. To decrease the computational cost of the Bayesian network and ease the data elicitation process, De Waal et al. [34] recommends removing factors which add complexity without adding value. [34] also proposed a 3D elicitation to decrease the number of values the expert needs to supply. In step four, coded concepts and causal connectors are used to develop a causal acyclic graph. Finally, prior (forward) sampling can be used to estimate model likelihoods for computational affordability instead of calculating the full joint probability table [35].

In conclusion, a mixed methods model consisting of ABM and DES components will be used to model the movement of grains and oilseed in and out of the silo. Sunflower farmer behaviour will be modelled with the aid of BNN to incorporate necessary factors. Theories on farmer agents are reviewed and summarised to be included in model behaviour. Data gathering for farmer agents will be done with the use of a causal mapping approach as developed by Nadkarni and Shenoy [32], the complexity of data elicitation can be lowered with the 3D elicitation method developed by Waal et al. [34], and sampling of the Bayesian network is best done by prior sampling [35]. The following section discusses the simulation model's development using the selected modelling methods.

3. Simulation model development and implementation

For experimentation at the silo complex, a multi-method simulation model consisting of ABM, DES and a Bayesian Network is designed by using interviews. The model is calibrated to a case study silo using primary and secondary data. Two critical agent types are modelled using AnyLogic simulation software, a silo and sunflower farmers, with the following functions:

Silo agent: The silo receives truckloads of grains and oil-seeds, as well as empty trucks for off-loading. Upon arrival, the silo determines whether it can service the truck by providing the required storage facilities (given the amount, type and grade of grain or oilseed arriving), or dispatch grain for the arriving truck. The silo is required by law to store grains and oil seeds of different grades separately. If the silo is unable to service a delivery truck due to insufficient storage capacity for the specific grain type and quality, a relocation cost is incurred by the silo. The silo also keeps score of average service levels and bin utilisation.

Sunflower farmer agent: Annually, sunflower farmers decide which sunflower seeds to plant, plant and harvest sunflower seeds and send truckloads to store seeds at the silo. The farmers goal is to maximise profit by selecting a cultivar (seed) with either a high yield potential (ton per hectare) or high oil content potential. These factors are inversely correlated, and is therefore a trade-off [2]. Sunflower seed quality is determined by various factors, but a decisive factor a sunflower farmer can change to increase seed oil content, is to select a cultivar (planting seed) with a high oil content potential.

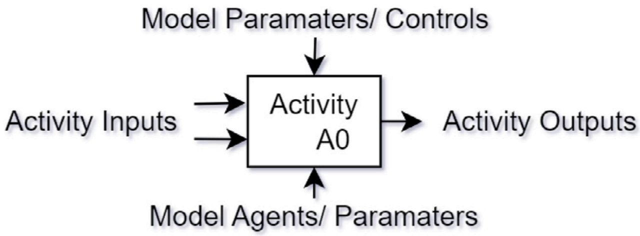


Fig. 1. IDEF concept.
Source: Adapted from Mahfouz et al. [36].

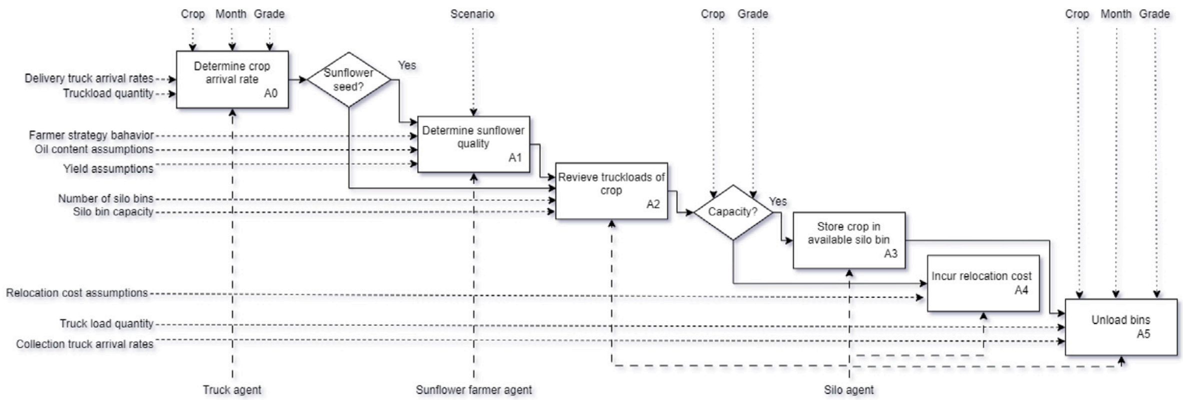


Fig. 2. Artifact IDEF concept.

An Integrated Definition Language (IDEF) model, as demonstrated in Fig. 1, is used to conceptually describe the multi-method simulation model in Fig. 2. Grain and oilseed arrive in truckloads at the silo. Crops compete for limited silo storage bins, and crops of different grades are stored separately. Therefore it is necessary to model the movement of all crops and crop grades delivered at the case study silo. The continuous process starts with simulating the production of crops, which is done by sampling the amounts of each crop to be delivered at the silo (A0, Fig. 2). Crop arrivals are determined depending on the model time, grade and specific crop. For each production season (off-season or on-season), crop and crop grade the model samples from arrival rates and typical truckload sizes. If the crop arrivals are being determined for is sunflower seed, an alternative process is used to determine the timing, amount and quality (determining the grade) of sunflower seed delivered to the silo (A1, Fig. 2). Sunflower seed quality and quantity are determined by the strategy a sunflower farmer follows, yield and oil content assumptions, however, the behaviour also depends on the model scenario (base case model or experimental model). Trucks are received by the silo (A2, Fig. 2), however, the model then determines whether it has the capacity to store the crop in the truck based on the number of silo bins, the capacity of each silo bin and the availability to store that specific grade of crop. If a silo bin is available, the crop is stored and takes up space in the silo bin, however, if capacity is not available the silo has to pay a relocation cost (A3, Fig. 2), based on relocation cost assumptions. Finally, trucks arrive at the silo agent to collect specific crops based on truck arrival rates and typical truckload assumptions, once again depending on system time and the crop at hand (A5, Fig. 2).

For the model, two scenarios are built. Firstly, ***the base case scenario***, where all grade one sunflower seeds can be stored in the same bin and sunflower farmers are only paid for yield. Secondly, the ***segregation scenario:*** a scenario in which sunflower seeds of different classes are split into different bins, and farmers are incentivised for oil content in sunflower seed.

3.1. Silo agent

The process of grains and oilseeds arriving and exiting the silo is developed following interviews with silo management. Trucks arrive at the silo throughout the year. Trucks either arrive to deliver a crop to be stored at the silo (arrival trucks) or arrive to collect grains and oilseed from a silo bin to be delivered to the next off-taker (discharge or collection trucks). The following inputs are required to study the impact at a specific silo complex, silo s:

$$n_s = \text{the number of silo bins at silo } s$$

$$C_{s,i} = \text{the capacity of bin } i \in 1 \ldots n_s \text{ at silo s}$$

$$Tar_{x,t,s} = \text{The historic arrival rate (trucks per month) of trucks containing crop } x \in X \text{ for season } t \in T$$

$$Taa_{x,s} = \text{The typical load (tonnes) of a truck containing crop } x \in X$$

$Tcr_{x,t,s}$ = The historic arrival rate of trucks collecting crop $x \in X$ for season $t \in T$

$Tca_{x,s}$ = The typical load (tonnes) of a truck collecting crop $x \in X$

R = The cost (R per tonne) of relocating crop $x \in X$

$$X = \begin{cases} 1 & = \text{White maize (Grade WM1)} \\ 2 & = \text{Yellow maize (Grade YM1)} \\ 3 & = \text{Wheat (Grade B1)} \\ 4 & = \text{Wheat (Grade B2)} \\ 5 & = \text{Wheat (Grade B3)} \\ 6 & = \text{Wheat (Grade BSG)} \\ 7 & = \text{Wheat (Grade BKA)} \\ 8 & = \text{Soybeans} \\ 9 & = \text{Sunflower (Grade 1, class FS)} \\ 10 & = \text{Sunflower (Grade 1, class FH)} \\ 11 & = \text{White maize (Grade WM2)} \\ 12 & = \text{White maize (Grade WM3)} \\ 13 & = \text{Yellow maize (Grade YM2)} \\ 14 & = \text{Yellow maize (Grade YM3)} \end{cases}$$

$$T = \begin{cases} 1 & = \text{Peak season} \\ 2 & = \text{Off peak season} \end{cases}$$

A silo is required by law to segregate grains of different grades [37]. However, the conditions that classify sunflower seed as class FS or FH (as stated in the array set *crop batch*) are vague. To date, the legislation only differentiates class FS as "high oil content" and class FH as "low oil content" without defining high or low oil content. For this research study, FS and FH can be defined according to the recommendations of Delport [2]. Sunflower seed is considered as high oil content (class FS) if it has an oil content of 38% or above, and low oil content (class FH) if its oil content is below 38%. It is assumed that all sunflower seed delivered at the silo satisfies the conditions to be classified as grade 1.

Most crop arrivals occur during a crop's peak season (usually determined by the regular harvest season). At the time of truck arrival, if any bin (tank) contains the incoming crop, and the remaining capacity of the tank is greater or equal to the incoming amount, the crop is routed to that specific bin. Alternatively, if the conditions mentioned above are not satisfied, the incoming batch is routed to the first empty bin if available. After that, relocation costs are incurred if no space is available, and the model increases the number of failed deliveries. Relocation cost is incurred due to fuel costs for each truckload to drive to the nearest silo with sufficient storage space. A typical relocation cost at the case study silo is based on a triangular distribution provided by silo management. For each collection truck arriving, the model searches for a random bin containing, at least, the tonne capacity of the truck or a bin containing the last tonnes of the crop and dispenses the capacity of the truck from the bin. At each delivery, throughout the model run-time, the silo service level is calculated as denoted by Eq. (2).

$$\text{s} = \frac{n_{tr}}{n_{ta}}$$

where

$$s \triangleq \text{Service level (\%)}$$

$$n_{tr} \triangleq \text{Number of truckloads successfully stored}$$

$$n_{ta} \triangleq \text{Number if truckloads arrived at silo door} \qquad (2)$$

Crops (excluding sunflower) arrive at a given monthly Poisson arrival rate, depending on the season. Each truck is assigned a crop, amount and grade once created within the appropriate crop arrival logic flow, determined by the state chart from which it originates. The same logic applies to both arrival and collecting (discharge) trucks. Once trucks arrive at the silo, they queue for service. Upon truck arrival, the silo is searched to determine whether storage space is available for the desired crop. Java functions and collections drive this logic. If sufficient space is available for a truck to dispense its contents, a specific bin is allocated to the truck, and the contents of the crop with accompanying information on the crop, grade and amount dispensed is handed on to the silo logic displayed in the bin logic section of Fig. 3. If a truck wants to collect contents, java functions and collections are used to identify the correct bin, and the desired amount (capacity of the truck) is dispensed from the corresponding bin.

3.2. Agent-based sunflower farmer model module

The farmer model module is used to capture a critical farmer decision; to plant or not to plant sunflower seed for high oil content. This decision ultimately determines which cultivar category to plant from, also referred to as the farmer's strategy in this study.

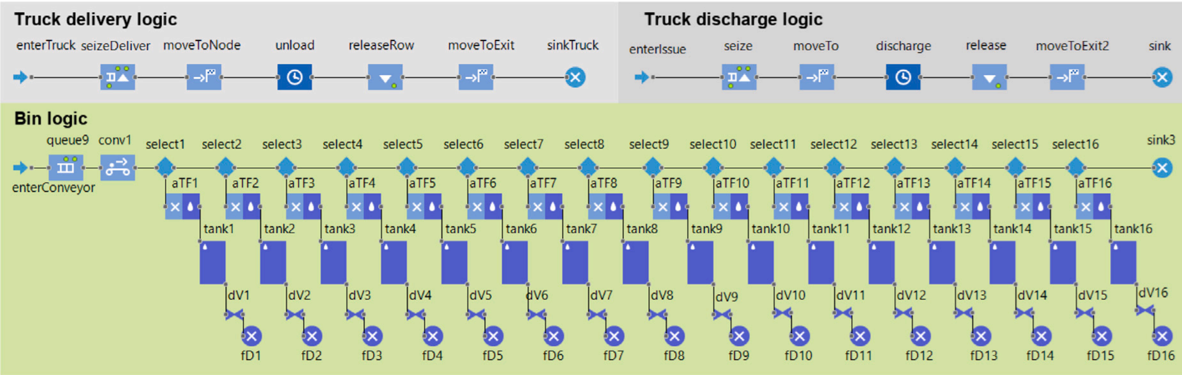


Fig. 3. Silo logic. Software: Anylogic.

Table 1
National cultivar trial performance.

Cultivar group	Yield (Tonne/Ha)	Oil content (%)
High yield	normal(2.37, 0.75)	normal(36.19,5.34)
High oil content	normal(1.91, 0.61)	normal(45.70, 1.95)

Farmers plant high-yielding cultivars during the base scenario and decide whether or not to deliver high-oil-content cultivars or high-yielding cultivars in the segregation scenarios. Farmer and industry experts are interviewed to identify critical agent behavioural patterns and support model process development. The average amount of sunflower seed delivered at the silo is divided by a typical sunflower farm size provided by the interview outcomes to estimate the number of surrounding farmers.

Annually, farmers choose which cultivar to plant before selecting seeds for the coming season. Cultivar choice is one of the few factors a farmer has control over when chasing quality parameters, alongside management practices such as planting date and fertilisation. Although cultivar selection is not the only factor a farmer is likely to change to grow for oil content, it has a significant impact. This is because a seed's oil content and yield potential are first and foremost determined by the genetics that preceded it (the cultivar). Traditionally, sunflower seed performance is defined by yield (tonne/hectare) and other health indicators, such as the absence of damage such as scarlatina and other diseases and pests. However, farmers also consider oil content (%) as a sunflower seed performance measure in this study. Therefore, cultivar selection in farmer behaviour is reduced to selecting a high oil content (%) cultivar or a high-yielding (tonne/hectare) cultivar.

Sunflower yield and oil content parameters for the different cultivar strategies are based on the outcome of the South African national cultivar trials. Oil content and yield for each strategy are sampled from normal distributions of the selected populations, as shown in Table 1.

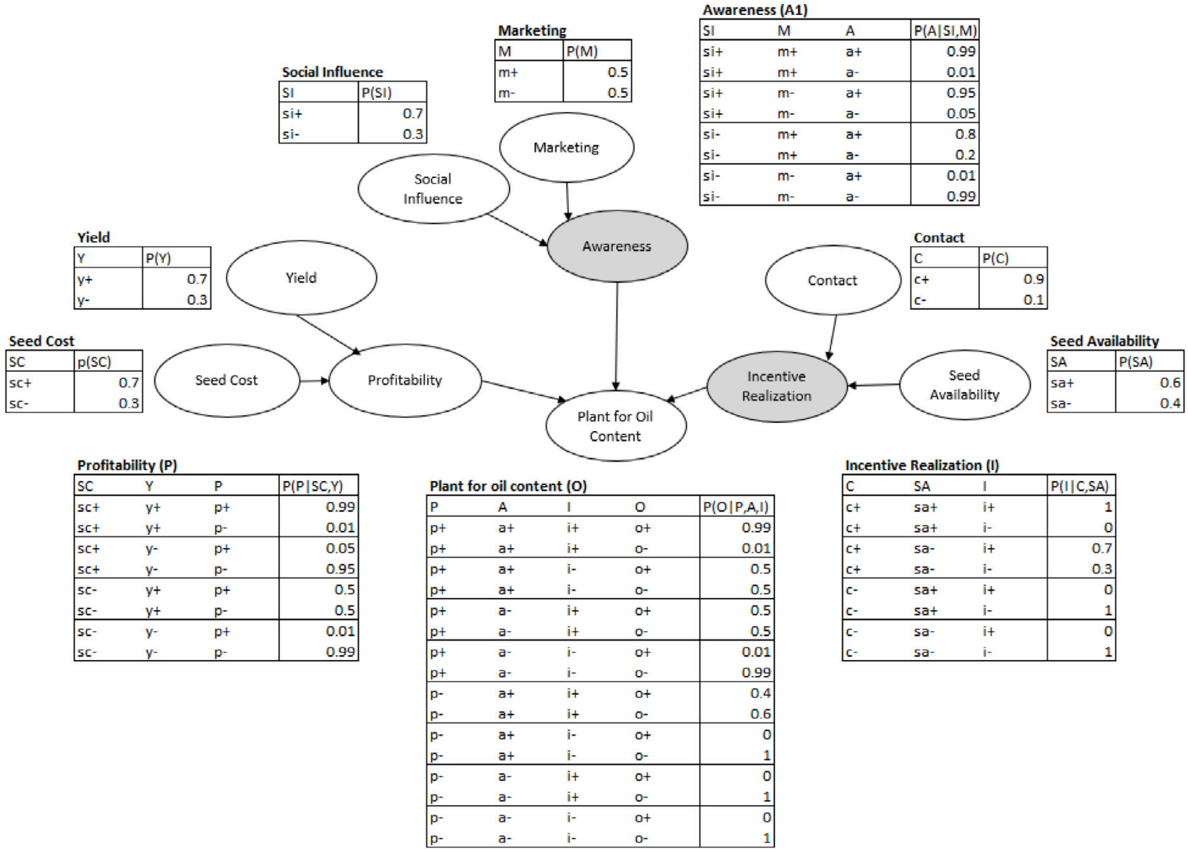
Five experienced sunflower farmers are interviewed to develop farmer agent behaviour. All sunflower farmers interviewed have been exposed to high oil-content cultivars and the possibility of an incentive. Farmers are asked about the processes and factors influencing their decision to pursue the incentive, along with technical questions about typical farm size, planting priorities and dates.

3.2.1. Farmer model one: Bayesian network

Semi-structured interviews are used to develop an expert-driven Bayesian network by using the causal mapping approach [32], supplemented with the latent variable approach [34]. Farmers are asked: "Which factors drive the decision of planting a high yield or high oil content cultivar?".

Thereafter farmers are asked about the likelihood of factor states occurring based on their experience to elicit prior probabilities. For example, farmers are asked what the probability is that the seed cost factor is true (described in Table 2). This straightforward elicitation process is easy for outer factors; however, it becomes more complex for factors with two or three elements leading into it, and becomes exponentially more difficult for each additional factor state. For these factors, the elicitation method [34] is used to simplify and reduce the number of probabilities to be provided by farmers. The final, detailed directed acyclic graph (DAG) with factors, factor states and probabilities is shown in Fig. 4.

The logic that governs farmers driven by the Bayesian network is shown in Fig. 5. Each year, a farmer will cycle through the planting process modelled as a state chart, *stFarm*. Annually, a farmer will start planning by determining his planting strategy for the year, high-yielding sunflower or high oil content sunflower, with the primary concern of making a profit. For example, if it is planting season (November to March), the farmer will plant and transition to the harvesting state within a timeout of $normal(127.5, 2, 5)$ days, as sunflowers will grow for between 125 and 130 days before harvest. Thereafter, it can take three days to harvest and transport a harvest to the nearest silo, governing the transition to storage. Shortly thereafter, the farmer will review his strategy for the following season.



Marketing

M	P(M)
m+	0.5
m-	0.5

Social Influence

SI	P(SI)
si+	0.7
si-	0.3

Awareness (A1)

SI	M	A	P(A|SI,M)
si+	m+	a+	0.99
si+	m+	a-	0.01
si+	m-	a+	0.95
si+	m-	a-	0.05
si-	m+	a+	0.8
si-	m+	a-	0.2
si-	m-	a+	0.01
si-	m-	a-	0.99

Yield

Y	P(Y)
y+	0.7
y-	0.3

Contact

C	P(C)
c+	0.9
c-	0.1

Seed Cost

SC	p(SC)
sc+	0.7
sc-	0.3

Seed Availability

SA	P(SA)
sa+	0.6
sa-	0.4

Profitability (P)

SC	Y	P	P(P|SC,Y)
sc+	y+	p+	0.99
sc+	y+	p-	0.01
sc+	y-	p+	0.05
sc+	y-	p-	0.95
sc-	y+	p+	0.5
sc-	y+	p-	0.5
sc-	y-	p+	0.01
sc-	y-	p-	0.99

Plant for oil content (O)

P	A	I	O	P(O|P,A,I)
p+	a+	i+	o+	0.99
p+	a+	i+	o-	0.01
p+	a+	i-	o+	0.5
p+	a+	i-	o-	0.5
p+	a-	i+	o+	0.5
p+	a-	i+	o-	0.5
p+	a-	i-	o+	0.01
p+	a-	i-	o-	0.99
p-	a+	i+	o+	0.4
p-	a+	i+	o-	0.6
p-	a+	i-	o+	0
p-	a+	i-	o-	1
p-	a-	i+	o+	0
p-	a-	i+	o-	1
p-	a-	i-	o+	0
p-	a-	i-	o-	1

Incentive Realization (I)

C	SA	I	P(I|C,SA)
c+	sa+	i+	1
c+	sa+	i-	0
c+	sa-	i+	0.7
c+	sa-	i-	0.3
c-	sa+	i+	0
c-	sa+	i-	1
c-	sa-	i+	0
c-	sa-	i-	1

Fig. 4. Acyclical graph of factors influencing the decision to plant cultivars for oil content.
Source: Farmer research participants.

Table 2
Description of parameters of the farmers directed acyclic graph.

Parameter	Code	Description (for + parameter subscript)
Seed cost	SC	The seed does not cost significantly more than other seed
Yield	Y	Yield (t/ha) is at least 2 tonnes per ha
Profitability	P	Gross margin (R/Ha) is on par with previous cultivars
Social influence	SI	The farmer has engaged with another farmer who is aware of the incentive
Marketing	M	The farmer received marketing of an incentive for oil content
Awareness	A	The farmer is aware of a possible incentive for oil content in sunflower seed
Contract	C	A contract is secured stating a premium price for high oil content
Seed availability	SA	Seed for the specific cultivar the farmer wants to plant is available
Incentive realisation	IR	A farmer will receive compensation for oil content in sunflower seed
Plant for oil content	O	The farmer will plant a cultivar that is intentionally high in oil content

While the farmer is in the state *stPlan*, the integer, *varStrategy* is set by calling the function *fnSampleBN()*. The java function, *fnSampleBN()*, samples the Bayesian network shown in Fig. 4 by continuously sampling a random double variable $r = uniform(0, 1]$. Forward (prior) sampling is used to derive a value for *varStrategy*. If $varStategy = 1$, the farmer will plant a high oil content cultivar, and yield and oil content sampled from the high oil content cultivar's historical performance (Table 1). Else, if $varStategy = 2$, performance is sampled from the yield cultivars.

3.3. Farmer model two: Margin model

Farmers find it easier to indicate the resulting decision for specific events based on rules. These agents follow the same planting process as model one; however, they base their strategies on rules with an if-then structure instead of sampling it from the Bayesian network. Farmer model two allows for more detail to be modelled; for example, each farmer indicated that a cultivar would always be tested in a small block for at least one season before it is scaled to a larger portion of land. As with farmer model one, this model is based on interviews with the five sunflower farmers. Fig. 6 shows the state chart that governs the alternative farmer agents.

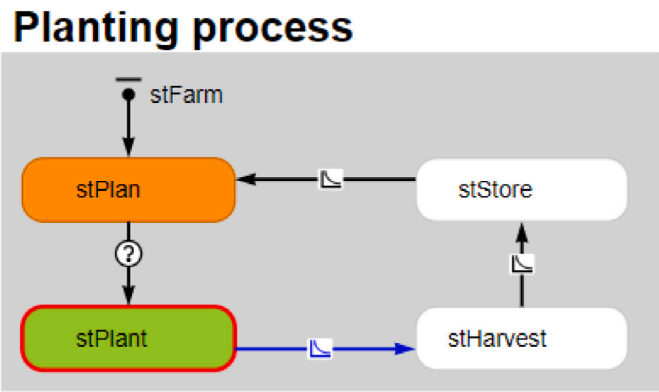


Fig. 5. Logic of farmers in model one. Software: Anylogic.

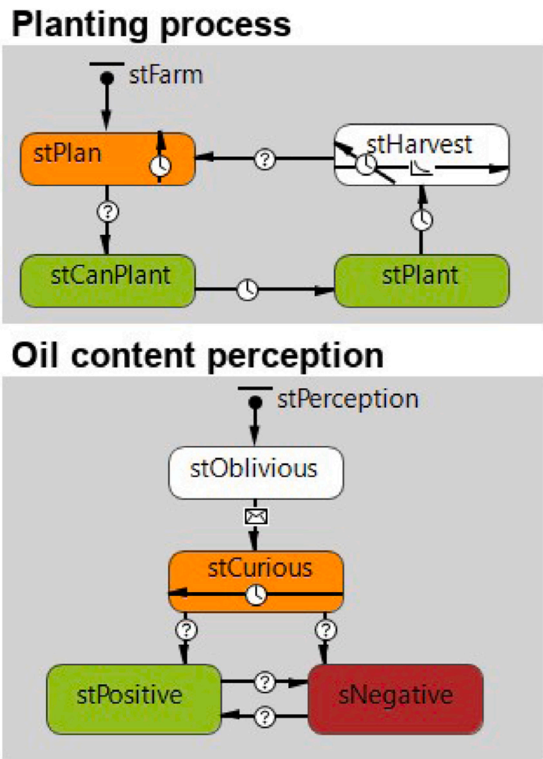


Fig. 6. Logic of farmers in model two. Software: Anylogic.

The conditional structure that drives the cultivar strategy for farmer model two is shown in Fig. 7. Farmers review strategies annually based on a collection of memories of economic performance linked to strategies. At first, a farmer is unaware of an incentive for oil content and will keep planting all hectares dedicated to the sunflower with high-yielding cultivars. Farmers become aware by either being in contact with a crusher that offers a contract based on oil content or being in contact with another neighbouring farmer who is aware of the incentive. To model the influence a contact from a crusher has on a farmer's decisions, one farmer is randomly selected to be aware at the start of the model run-time. To model the social dynamics between farmers, they are assigned random locations within a 100 km × 100 km matrix, as done by Ge et al. [7].

As a starting point to model the time it could take for information to travel from one farmer to another, as well as incorporate the possibility that information might not reach a farmer, sunflower farmers will interact with surrounding neighbours within 30 km (euclidean distance) at random rates and might become aware of the incentive or influence other neighbours. Once a farmer becomes aware, the farmer will test the new strategy on a small number of hectares (Trail block). If the trial performed well compared to the high-yield cultivars, the farmer would upscale the number of hectares dedicated to oil content. If the high-oil-content cultivars

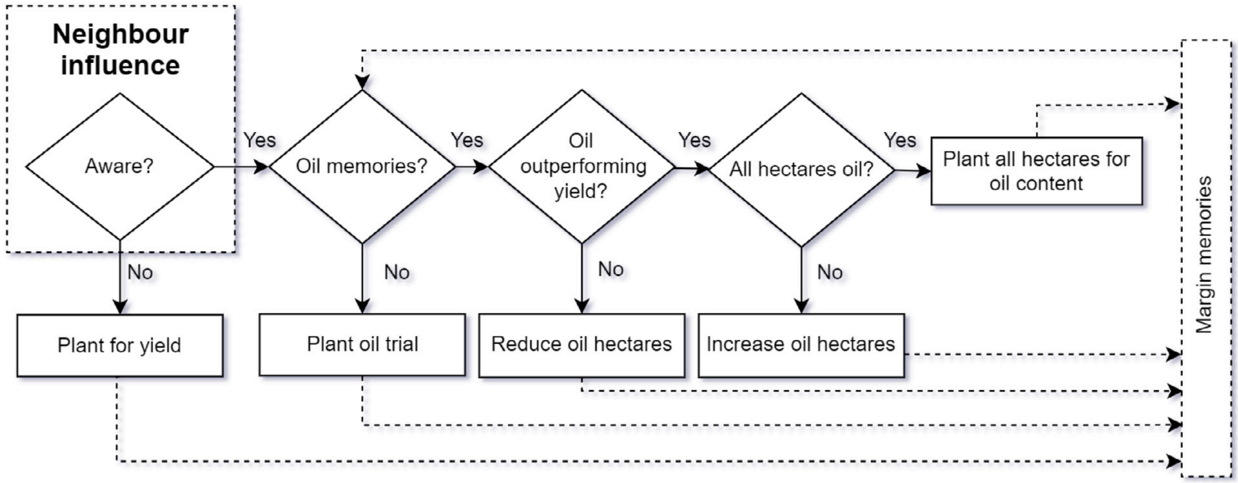


Fig. 7. Farmer model two strategy flow diagram.

Table 3
Example sunflower margin.
Source: Mcpherson [38]

Income		
Yield	T/HA	2
Price per tonne	R/T	9050
Travel differential	R/T	300
Price per tonne	R/T	8750
Total income	R/HA	17 500
Variable costs		
Seed cost	R/HA	650
Fertiliser	R/HA	1800
Weed control	R/HA	1200
Diesel	R/HA	1250
Crop insurance	R/HA	620
Harvesting	R/HA	750
Marketing and Transport	R/HA	400
Other costs	R/HA	1000
Total variable cost	R/HA	7670
Gross margin	R/HA	9830
Gross margin	R/TON	4915

perform well for three consecutive years, the farmer will upscale his entire sunflower operation to high-oil-content cultivars. The farmer will reduce the hectares dedicated to high-oil-content cultivars if the high-oil-content cultivars do not perform well.

Economic performance is measured by gross margin. The sunflower gross margin is calculated as shown in Eq. (3). A breakdown if typical input costs leading to an example gross margin is shown in Table 3. To tease out the effect specifically of yield and oil content on margins in this study, input costs are assumed to remain constant, while yield and oil content adjust the gross margin annually. As depicted by Delport [2], the incentive for oil content is added to the gross margin (Eq. (3)).

$$\text{Gross margin} = \begin{cases} p * \left(y + \frac{(o-38\%)}{100} \right) - v & \text{if } o > 38\% \\ y * p - v & \text{otherwise} \end{cases}$$

where

$$o \triangleq \text{Oil content (\%)}$$
$$y \triangleq \text{Yield (Tonne/Hectare)}$$
$$p \triangleq \text{Price per tonne (R/Tonne)}$$
$$v \triangleq \text{Variable cost (R/Ha)} \qquad (3)$$

With the model boundaries and behaviour of the leading agents, the sunflower farmers and silo is established, the outputs of the designed simulation model are discussed in the following subsection.

Table 4
Crop batch arrival rates (trucks per month).

Crop	Jan	Feb	March	April	May	June	July	Aug	Sep	Oct	Nov	Dec
Maize (YM1)	0.25	0.25	7.75	28.75	167.5	303.25	243.75	88	8.25	9.33	0.33	0
Maize (YM2)	0	0	0	0	1.5	6.5	11.75	16.5	0	0	0	0
Maize (WM1)	0	0	0	0.5	0.75	33.25	72.5	31	1.5	0	0.33	0
Maize (WM2)	0	0	0	0	0	1.25	15.25	12	3.5	0	0	0
Maize (WM3)	0	0	0	0	0	0.25	1.75	2.75	0.5	0	0	0
Wheat (B1)	19	5.5	4.75	7.25	0	2.25	0.5	1	0	0	10.67	46
Wheat (B2)	0.75	0	0.5	0	0	0.75	0.25	0.25	0	0	7	17.67
Wheat (B3)	3.75	3.25	0	0	0.25	0	0	0	0	0	23.7	38.33
Wheat (BKA)	2.5	0.25	0.25	1.25	0	0	0	0	0	0	17	26
Wheat (BSG)	29.25	12	0.25	0.25	1	0	0.25	0	0	0.67	57	130
Soybeans	0	0	0	1.25	7.5	1.25	0	0	0	0	0	0

3.4. Model outputs

The model keeps track of variables either updated upon an event or when an action triggers a function that calculates its corresponding output. The model ouputs include:

Number of deliver trucks arrived: the amount of trucks arriving at the silo to deliver grains and oilseed, including sunflower and other crops (number of trucks/annum).

Number of failed deliveries: The amount of truckloads that could not be stored away due to insufficient storage bin capacity (number of trucks/annum).

Crops that could not be delivered: The corresponding crop within a truckload that could not be stored due to insufficient storage space.

Service level: A fraction indicating the percentage of truckloads that could be successfully stored out of all the arrivals (%).

Total relocation cost: The total annual cost incurred by the silo to compensate for failed deliveries (R/annum).

Amount of each crop delivered annually: The total amount of each crop delivered annually (tonnes/annum).

Bin utilisation: The percentage (%) of total bin capacity used per month.

These outputs are all critical to understanding sunflower segregation's impact on bin utilisation and service levels, with supporting context such as relocation costs and the specific crops impacted. They also support the consideration of model validity, discussed in the following subsection. Finally, outputs are captured after a short model run-time of 2 s per model run (20 virtual years, run in weekly increments).

3.5. Case study

The model as developed throughout Sections 3.1–3.4 is applied at a case study silo. Secondary quantitative data from a case study silo is used to determine the number of storage bins, each bin's storage capacity, and the types of grain usually stored at the silo. This is a data-centric section focused on all information derived from the secondary data provided by the case study silo.

Of all crops delivered at the case study silo, yellow maize is the highest volume crop. Arrival rates are based on secondary data from the case study silo, as shown in Table 4 for arrival trucks and Table 5 for collection trucks. The average amount of each crop delivered per truck for each crop is shown in Table 6.

The case study silo consists of bins with capacities listed in Table 7. Bin capacity is given in tonnes of maize; however, not all grain and oilseed have the same density as maize. Therefore, conversion factors shown in Table 8 are assumed based on industry engagement. For example, although most bins can hold up to 4700 tonnes of maize on average, the same bin will only be able to hold 2200 tonnes of sunflower due to the higher density of maize compared to sunflower.

The number of sunflower farmers is determined based on the typical farm size, yield, and silo sunflower volumes gathered from interviews. For the case study silo, farmers reported an average sunflower farm size of 450 ha per farmer per year. An average yield of 2 t/ha is used to determine the number of farmers required to meet the average sunflower yield of approximately 9000 tonnes achieved for the case study silo. Therefore, the case study silo assumed 10 sunflower farmers in the surrounding area.

A typical relocation cost at the case study silo is based on a triangular distribution provided by silo management. Different factors influence relocation costs, such as truck size and load, customer relationship and bargaining. When considering the maximum, minimum and average historic relocation costs easily provided by silo management, the grain relocation cost (R/tonne) can be sampled from the distribution $\sim Triangular(65, 380, 237.17)$.

Table 5
Crop collection truck arrival rates (2019–2021 average).

Crop	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Yellow maize	50.25	51	131	63	65	34.75	29.25	20	21	42	45	84.33
White maize	1.75	3.25	6	8.25	3.25	0.25	0.75	9.25	9.5	10	12.33	25
Wheat	67.5	81.25	35.25	26	16.5	22.75	28	46.5	17	6	4.67	14
Soybeans	0	0.5	1.25	1	1.5	2.25	1	0	7.25	3	7.33	0
Sunflower	5.5	1.25	5	10	9.25	19.5	46.5	57.25	42	41	22.33	12.67

Table 6
Typical truck amounts per crop (2019–2021 average).

Crop	Peak season	Delivery truck amount (Tonnes)	Collection truck amount (Tonnes)
Yellow maize (YM1)	May–August	22.4	29.4
Yellow maize (YM2)	May–August	22.4	29.4
White maize (WM1)	May–August	21	36.6
White maize (WM2)	June–September	19.7	36.6
White maize (WM3)	June–September	19.2	36.6
Wheat (B1)	November–January	28.8	34.1
Wheat (B2)	November–December	25.1	34.1
Wheat (B3)	November–February	21.4	34.1
Wheat (BKA)	November–January	19.9	34.1
Wheat (BSG)	November–February	27.7	34.1
Soybeans	April–June	13.43	36.6
Sunflower	April–July	12.2	27.4

Table 7
Case study silo bin parameters.

Parameter	Value	Source
Storage bins	16	Silo management
	4 × 1250 tonnes	Secondary data
Bin capacity	1 × 2450 tonnes	Secondary data
	11 × 4790 tonnes	Secondary data

Table 8
Grain density conversion factors.

Crop	Bin capacity (Tonne)	Crop bulk density (t/m^3)
Maize	4700	0.72
Sunflower seed	2200	0.42
Soybeans	4700	0.72
Wheat	5200	0.8

3.6. Model validation

ABMs are highly volatile and complex systems, lending themselves more to soft calibration, such as stakeholder knowledge for validation [39]. Therefore, expert validation combined with validation with historical data is used to validate the DES and ABM components of the model [39,40].

The discrepancy in bin usage throughout the year is shown in Fig. 9 with a mean absolute percentage error (MAPE) of −1.95% and absolute percentage error never exceeding 25%, indicating acceptable accuracy. The model reasonably accurately mimics the seasonality of bin volumes in the silo (Figs. 8 and 9). However, it slightly underestimates volumes during peak months (September and October). The volume of sunflower seed arriving also falls within an acceptable range (Fig. 10) since the resulting volumes never exceed the minimum and maximum amount of sunflower seen in reality, validating sunflower farmer agent behaviour. The tonnes delivered for each crop compared to empirical averages are shown in Fig. 11. Crop volumes will vary from year to year especially in terms of the grades delivered for crops graded on quality. The graph shows that the modelled output is within an acceptable range (the difference between model output and secondary data does not exceed 250 tonnes) for each crop when compared with the secondary data. Therefore it can be concluded that the model results in acceptable crop amounts in a year. The service level output of the model was discussed with silo management and is considered acceptable (Estimated to be around 97%). The farmer's behavioural logic regarding sunflower quality and quantity was validated by reviewing the final Bayesian network with a farmer research participant, who acknowledged that it is challenging to validate but that the results are acceptable.

The model outputs on service levels are discussed with silo management for final validation. Management estimates that currently, high service levels of around 97% are observed, as most trucks can be serviced, therefore, validating the service level output of the as-is model. Although service levels over multiple runs follow a downward trend, resulting in an average service

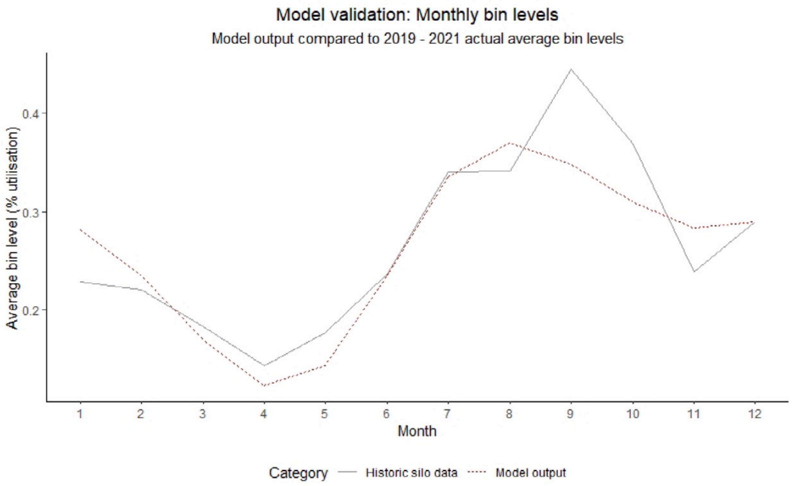


Fig. 8. Silo logic validation: Comparison of model outputs to secondary data.

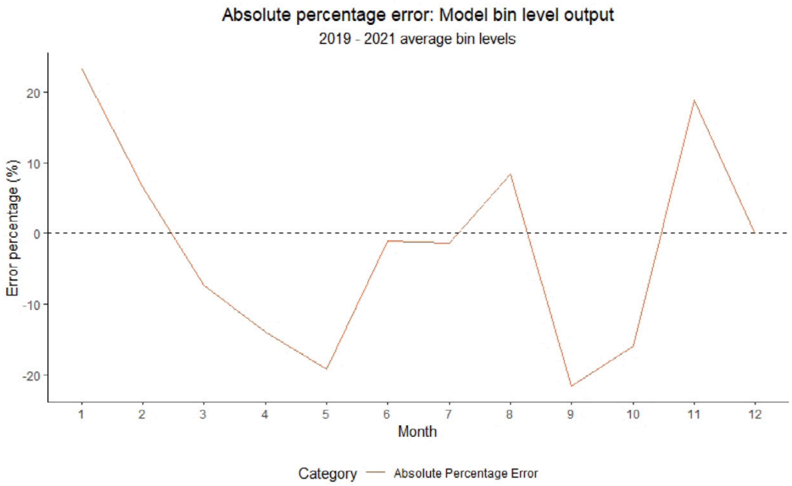


Fig. 9. Absolute percentage error (APE) of the base case model over time.

level of 95.8%, the model stays within an acceptable range of the current service level. This trend results from years in which the farmers have a big harvest, and collection trucks are not adjusted. When calibrating the model by increasing collection trucks, the bin volume output is no longer accurate; therefore, the service level output is acceptable at this level. Due to the yield and hectare input models, sunflower tonnes will experience higher variability than other arrivals based on arrival rates. The occurrence of an unmanageable sunflower harvest is, however, rare and, therefore, not a significant concern for this paper.

Fig. 12 demonstrates that the Bayesian network sampling yields that the fraction of sunflower farmers growing for oil content remains constant over the virtual modelled period. However, in the margin model, farmers respond over time, but the final result converges within 10% of farmers. Since there is a discrepancy in model performance, both models will be analysed in the results section to understand the impact of different levels of farmers responding to the incentive.

Although the model is not a perfect representation of reality, model validation demonstrates that the model performed within reasonable accuracy using quantitative and soft validation methods. Therefore, since it has been verified that the model performs according to its intended purpose and the base case results are valid compared to reality, the following section discusses the simulated scenarios' results.

4. Results

The results for the case study silo are structured according key output metrics: service level, bin utilisation and economic implication. The experiment is performed in Anylogic with 300 replications (simulation runs) for each of the four scenarios. Each

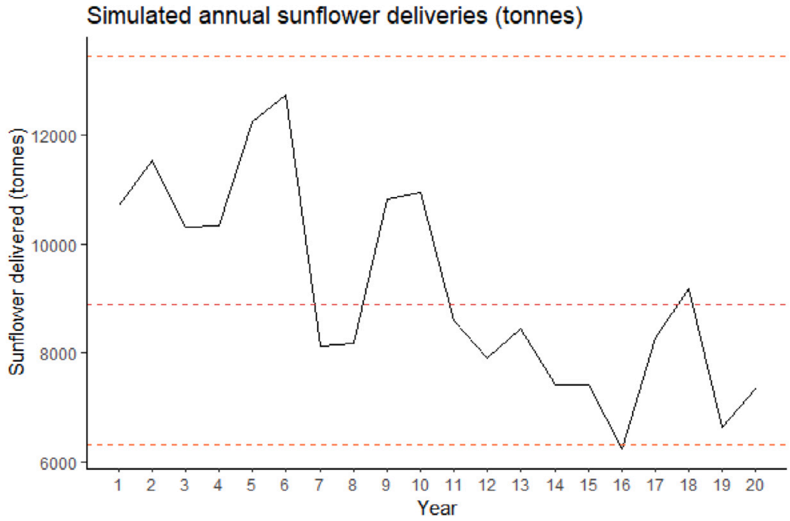


Fig. 10. Simulated sunflower delivery volumes compared to the actual acceptable range (6050 - 13 500 tonnes).

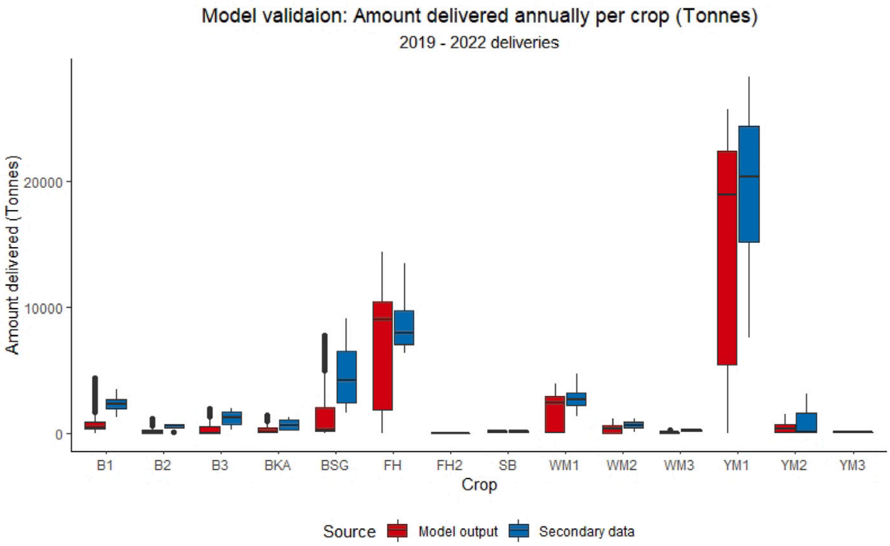


Fig. 11. Validation of tonnes delivered per annum per crop.

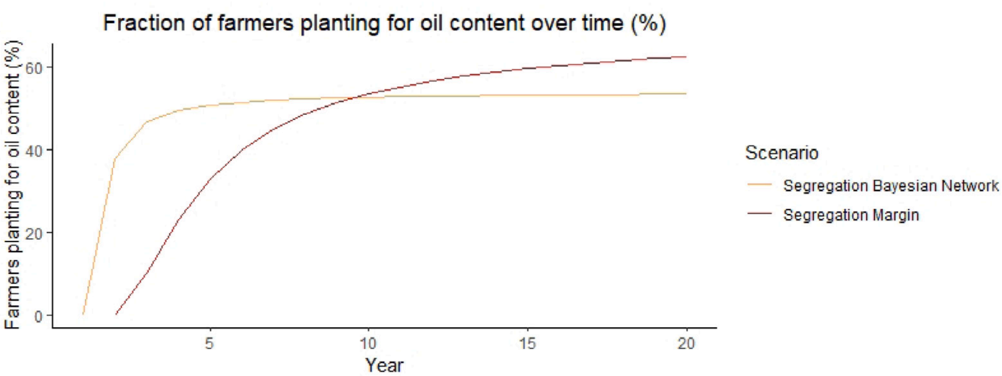


Fig. 12. Farmers planting for oil content (%) over time.

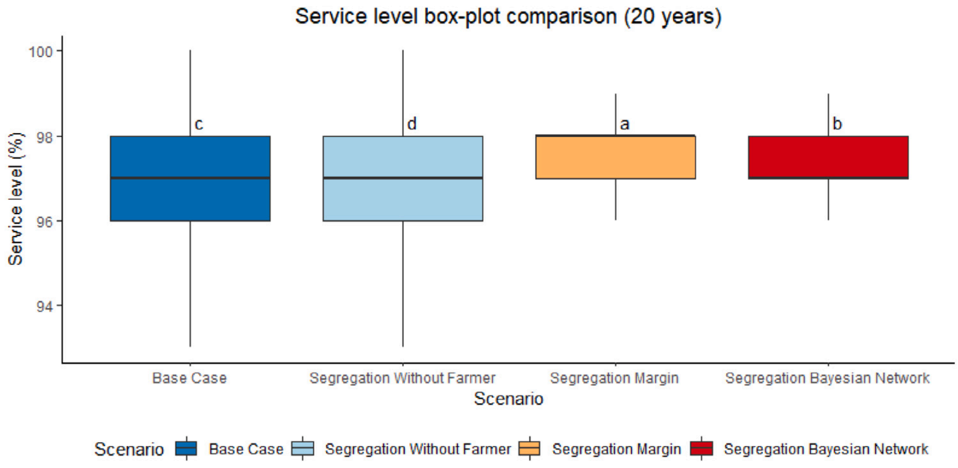


Fig. 13. Service level box-plot comparison with Tukey's HSD letters.

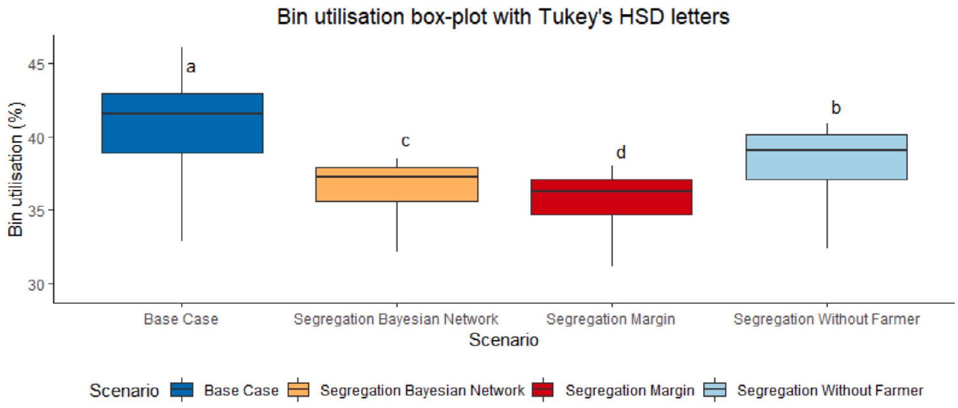


Fig. 14. Bin utilisation box plot comparison with Tukey's HSD letters.

simulation run is conducted with a different random seed and runs for 20 virtual years (1040 weeks). Observations are recorded either weekly or on an occurrence basis, depending on the variable. Model outputs are analysed with RStudio statistical software.

4.1. Service level

The average service level for the base case, segregation case without farmer strategy, Bayesian network farmer, and margin farmer were 96.3%, 95.94%, 97.36%, and 97.60%, respectively (see Fig. 13). Analysis of Variance (ANOVA) results in a test statistic of $p < 2e^{-10}$; therefore, it can be concluded that the difference between group means is significant. Although statistically significant, segregation without farmer response resulted in a 0.36% reduction in service level compared to the base case. On the other hand, farmer-influenced segregation scenarios show an improvement in service level by 2%. This is due to the lower amounts of sunflower delivered if sunflower farmers grow high-oil-content cultivars.

4.2. Bin utilisation

Bin utilisation refers to the annual usage of the bins and is a measure of revenue as farmers pay for storage per tonne per day. The average bin utilisation of the base case, segregation case without farmer strategy, Bayesian network farmer and margin farmer are 40.5%, 38.2%, 36.5% and 35.7%, respectively (see Fig. 14). ANOVA shows that the difference between average bin utilisation is significant with a test statistic of $p < 2e^{-10}$. Therefore, segregation has a significant negative impact on bin utilisation.

4.3. Combined financial impact

Bin utilisation and service level can be conflicting objectives. High bin utilisation results from bins containing more grain; however, the fuller bins are, the higher the chances that the silo will no longer have the capacity to store truckloads of grain.

Table 9
Scenario comparison: Annual opportunity cost loss/savings.

Scenario	Mean annual service level	Trucks not serviced	Opportunity cost	Relative opportunity cost	Total relative loss (Savings)
Base case	96.3%	123	R236,423	–	–
Segregation without farmer	95.95%	128	R275,133	R9,611	+4.1%
Segregation Bayesian network	97.26%	67	R128,783	(R107,640)	−45.5%
Segregation margin	97.50%	56	R107,640	(R128,783)	−53.5%

Table 10
Scenario comparison: Total annual service level related financial loss/savings.

Scenario	Mean annual service level	Trucks not serviced	Relocation cost	Opportunity cost	Total loss per annum
Base case	96.3%	123	R27,859	R236,423	R264,282
Segregation without farmer	95.95%	128	R29,100	R275,133	R275,133
Segregation Bayesian network	97.26%	67	R15,327	R128,783	R144,110
Segregation margin	97.50%	56	R12,727	R107,640	R120,367

Table 11
Scenario comparison: Financial implication of bin utilisation (Annual).

Scenario	Mean bin utilisation	Storage revenue	Relative change in revenue	Change (%)
Base case	40.5%	R8.9 million	–	–
Segregation without farmer	38.2%	R8.3 million	R0.5 million	−5.98%
Segregation Bayesian network	36.5%	R8.0 million	R0.9 million	−8.79%
Segregation margin	35.7%	R7.8 million	R1.1 million	−10.55%

Table 12
Scenario comparison: Total relative financial loss/savings (annual).

Scenario	Relative relocation cost	Relative opportunity cost	Relative storage revenue	Total loss (Savings)	Percentage gain (loss)
Base case	–	–	–	–	–
Segregation without farmer	R1,240	R9,611	R504,875	R515,726	(6%)
Segregation Bayesian network	(R12,532)	(R107,640)	R878,044	R757,872	(9%)
Segregation margin	(R15,132)	(R128,783)	R1,053,653	R909,738	(11%)

When bins are full, bin utilisation is high; however, trucks cannot be serviced, lowering service levels. Only when some contents are collected from bins (lowering bin utilisation) can trucks be serviced again (increasing service levels). There is, however, a delicate balance, as service levels can increase without bins being filled to the brim, therefore not affecting service levels. It depends on the seasonal demand for storage space.

The results show a reduction in bin utilisation and service levels for the segregation case without farmer response. In contrast, a significant reduction in bin utilisation but a significant improvement in service levels is seen for the segregation cases with farmer response. Therefore, to understand the combined impact of the two measures, the financial implications of the changes in bin utilisation and service levels are discussed. Financially, reducing service levels increases relocation costs but also incurs opportunity costs at the current silo, as business is moved to a different silo. Furthermore, lower bin utilisation results in lower storage revenue annually.

Reducing service levels results in an increase in relocation costs, with annual relocation costs ranging from R12,730 to R29,100 (Table 10). The opportunity cost is also estimated for each scenario, with the Bayesian network and margin scenarios achieving a decrease of 45.5% and 53.5%, respectively, while the segregation scenario without farmer response resulted in a 4.1% increase (Table 9). Table 11 shows the estimated change in revenue as a result of bin utilisation, ranging between R504,900 for the segregation without farmer scenario, and R1,054,000 for the margin scenarios.

As a result, all segregation scenarios resulted in financial losses for the silo, with a 6%–11% reduction in financial performance relative to the base case (Table 12). Overall, small changes in service level and bin utilisation result in significant changes in financial performance at the case study silo. The combined financial impact of all factors shows that the slightest change in bin utilisation will have a significant impact on revenue generation, outweighing the gains from increased service levels in scenarios with farmer response.

4.4. Crops included

Silo management is also concerned about the inability to service yellow maize trucks. Yellow maize is the highest volume crop at the silo and generates the most revenue. Therefore being unable to serve maize farmers during peak season could lead to significant

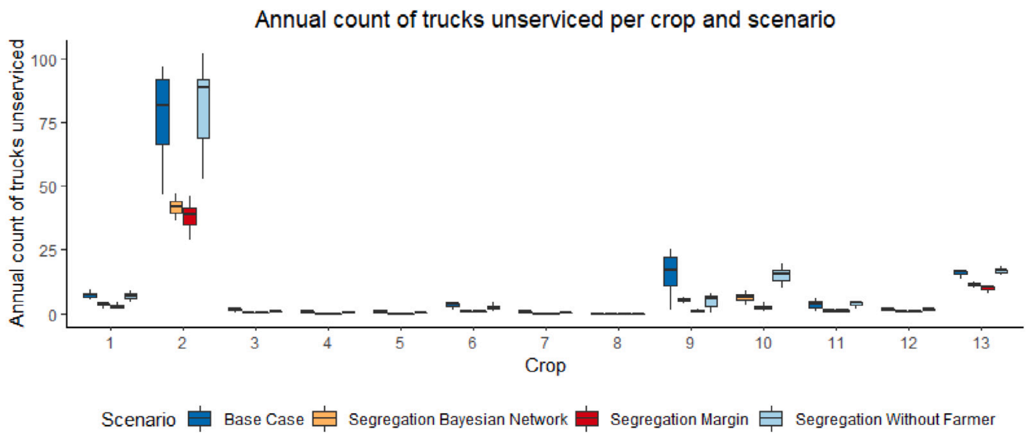


Fig. 15. Annual number of trucks unserviced grouped by crop and scenario.

revenue losses or high relocation costs. Fig. 15 shows the number of truckloads of each crop that could not be serviced due to insufficient storage space. Crop code descriptions are available under Section 3.1. Most that need to be relocated are crop 1, Yellow maize grade YM1, followed by crops 9 and 10, sunflower seed grades FS and FH, respectively (Fig. 15).

5. Discussion

The multi-method simulation model is applied to a case study silo, chosen explicitly due to the high volumes of sunflowers stored annually. At the case study silo, segregation has a significant negative impact on service level (0,36% reduction) if sunflower seed is segregated without any change in farmer behaviour. This can be assumed to be the outcome if the silo starts segregating sunflower seeds with the current state of cultivar decisions and performance in the industry. This also remains the case if sunflower yields remain relatively stable over time. Farmer-influenced segregation scenarios resulted in an average significant 1.2% increase in service level. As farmers justify lower yields with an incentive for oil content, the tonnes of sunflower delivered at the silo also reduce, resulting in a positive impact on service levels.

On the other hand, all segregation scenarios result in lower bin utilisation. The 2.3% reduction in bin utilisation results from segregation for the as-is cultivar state in the industry (segregation without farmer). Segregation reduces service level by allocating another bin to a minor crop and decreasing bin utilisation. However, these results suggest that lower bin utilisation in the segregation Bayesian network and margin case is not due to only segregation, as service levels increased in these scenarios. The segregation will lead to small amounts of grain stored in a bin, leaving no room for another high-volume crop, as service levels and crops unable to be serviced are reduced. It is also known that the scenarios influence farmers to choose high-oil-content cultivars, which usually result in lower yields. As a result, fewer tonnes of sunflower seed are delivered annually. Therefore, it can be concluded that the reduction in bin utilisation in these scenarios will be from the smaller amount of sunflower seeds received annually.

A reduction of 4.4% in average bin utilisation for the farmer-influenced segregation scenarios results in a 10%–11% (R500,000 - R900,000) reduction in revenue created by tonnes stored annually. This implies that an improvement in the sunflower industry has the unintended consequence of periodically lowering the amount of sunflower produced, which has a negative financial impact on silos. Nevertheless, it is essential to acknowledge that demand and supply will consistently achieve equilibrium and that national supply would be expected to increase over time as farmers plant more hectares of sunflower and cultivars and management practices evolve and mature. However, further research is required to model this complex economic phenomenon.

Given the estimated negative financial impact of between R500,000 and R900,00 for different segregation scenarios, farmers will not be able to expect segregation services without negotiation and collaboration of the sunflower industry. Therefore, financial investment is required to make up for the losses of the silo. Albeit a higher tariff for farmers, subsidies, or payment from crushers, collaboration is required to make up the estimated 6%–11% loss in relocation, opportunity and storage costs annually. For the time being, until the silo and sunflower industry has devised a solution for the loss on the part of silos where segregation is implemented, farmers will need to use alternative storage solutions to preserve the identity of grain to be sold at a premium.

In conclusion, segregation will have a significant negative impact on service levels and relative costs when sunflower is segregated with the current farmer yields (segregation without farmer scenario). The farmer-influenced segregation scenarios predict that the medium to long-term impact that segregation will have is an increase in service levels but a significant decrease in bin utilisation. The positive impact on service levels is due to the lower yield in high oil content cultivars, resulting in fewer tons of sunflower delivered with fewer trucks and resulting in fewer bins required by sunflower seed. More bins can then be allocated to maize that is also competing for storage space and would previously need to be relocated in the segregation scenarios. However, as cultivars evolve and yields increase, the impact on volumes will adapt accordingly. Therefore, as high oil content cultivars are developed to higher yields, in the long run, this positive impact on service level will reduce as volumes increase due to higher yields.

Finally, for decision-makers, this study proves that the specific case study will suffer significant losses from offering segregation services to sunflower farmers. For key value chain actors, this study is a starting point for negotiating the segregation of sunflower seeds at strategically positioned silos. Industry-wide cooperation is required to develop a solution for the sunflower industry's needs so that each value chain actor benefits. The model also serves as a proof of concept to be applied to other case studies and to understand the impact of segregation in other silos.

6. Conclusion

This article discusses the impact of segregating sunflower seeds based on oil content at a case study silo, as crop industries that switch to a quality-based marketing system rely on silos to offer necessary segregation services. The multi-method simulation model developed in the study aims to predict the impact of segregation on the silo by considering service level, relocation cost, bin utilisation, and combined financial loss due to these measures. At the case study silo, segregating sunflower seeds based on oil content has a significant negative impact on service level and bin utilisation in the short term, with a combined negative financial implication of between 6% and 11% for all segregation scenarios. However, it could have a positive impact on service levels in the medium to long term if sunflower farmers change their strategy to qualify for the incentive. The study suggests that alternative storage methods will be the status quo until a compromise can be reached by the silo and sunflower industry due to the significant negative financial impact on silos and the potential lower utilisation of silo bins with lower-yielding cultivars. Results from this case study silo should not be extrapolated as the case for all silos, and the model serves as a starting point for discussions between silo and sunflower value chain actors to increase the competitiveness of the overall sunflower value chain.

This study fills a gap in scientific literature by presenting a silo model that successfully tracks the movement of grain into and out of a silo, as well as proposing a method for testing the impact of segregation on service level and bin utilisation. It also addresses the segregation of sunflower seed based on quality. Furthermore, it utilises a method for modelling farmers as agents in the event of the scarcity of data on agents. This study contributes to the academic literature by applying a combination of agent-based modelling, discrete event modelling and Bayesian network modelling to a new environment; the South African silo industry with consideration of grade-based sunflower segregation.

In conclusion, the research study conducted in this paper addressed a new problem in the South African sunflower and silo industries by developing a novel modelling approach for the field of silo segregation problems in the ASC context. It offered an initial solution at a case study silo to consider silos in expanding the value chain to a quality-based system. This study is the first application of a silo segregation problem in the South African context and the first consideration of the capability of the sunflower industry to segregate sunflower seed based on oil content in South Africa since initial remarks were made by Nel [3]. Therefore, this research should support the growth of the South African sunflower industry.

This research study intersects multiple topics with future research opportunities. Firstly, data for case studies were limited as an independent researcher with little engagement in the silo industry. There is, therefore, an opportunity to apply the model to a representative sample of sunflower silos for a conclusion of the resulting impact at a national level. Secondly, the risk of transporting a harvest for hundreds of kilometres might be a constraint in farmers' response to taking up the incentive. Therefore, segregation at silo level might have another response that was not accounted for in this model. Silos with this offering might experience a further increase of sunflower delivered as farmers choose to store high oil content sunflower seed at a silo between the farmer and crusher to reduce risk. Therefore, sunflower farmers, crushers and producers would have to cooperate to find a mutually beneficial setup for the incentive to work in the long run. It might be beneficial for certain players in the industry to consider the spatial implication of testing points, sources and off-takers to minimise risk on the part of the farmer and minimise capital expenditure on the part of the silo industry. Thirdly, it will benefit the silo industry if further research is done on silo behaviour in an agent-based behavioural model. The aim will be to model the response of variable supply and demand at the silo level and, in turn, the impact of risk management due to more grades of grain combined with this variability. Fourthly, the performance of sunflower cultivars is highly variable and dependent on various factors such as locality, weather, fertiliser and planting dates. Continuous improvement and breeding of new cultivar hybrids to increase crop performance and engagement with seed development companies are critical. Farmers will always try to maximise yield, as yield acts as a multiplier in the gross margin, driving up profitability. Therefore, seed companies play a crucial role in finding the suitable trade-off between these two agronomic traits for customers, which has been shown to have a negative correlation [2]. Finally, sunflower farmers are not aware of the oil content of their product, except for the select few which has had cultivars tested at independent labs. Therefore the state of data gathering and sharing within the sunflower industry requires attention before such an incentive can be formally implemented as a policy.

Funding

This work was supported by the Oilseeds Advisory Committee (OAC).

Data availability

The data that has been used is confidential.

References

[1] S. Baker, T. Herrman, F. Fairchild, Capability of Kansas Grain Elevators to segregate wheat during harvest, Kans. Agric. Exp. Stn. 1 (1) (1997).
[2] M. Delport, Key drivers that affect sunflower seed quality, Oilseeds Focus 6 (2020) 48–49.
[3] A.A. Nel, Determinants of Sunflower Seed Quality for Processing, University of Pretoria, 2001.
[4] Hurburgh, Identification and segregation of high-value soybeans at a country elevator, JAOGS 71 (10) (1994) 1073–1078.
[5] K. Reichard, H. Vachal, Identity Preserved Grain - Logistical Overview, Upper Great Plains Transportation Institute, 2000.
[6] B. Remigio, D. Maier, Using modeling techniques to test the feasibility of segregating non-GMO soybeans at commercial elevators, in: IFAC Conference on Modelling and Control in Agriculture, Horticulture and Post Harvest Industry, 2000, pp. 245–250.
[7] H. Ge, R. Gray, J. Nolan, Agricultural supply chain optimization and complexity: A comparison of analytic vs simulated solutions and policies, Int. J. Prod. Econ. 159 (2014) 208–220.
[8] M.C. Caniels, H.A. Romijn, Supply chain development insights from strategic niche management, Learn. Organ. 15 (2008) 336–353.
[9] D.S. Utomo, B.S. Onggo, S. Eldridge, Applications of agent-based modelling and simulation in the agri-food supply chains, European J. Oper. Res. 269 (2017) 794–805.
[10] R. Berruto, D.E. Maier, Analuzing the receiving operation of different grain types in a single–pit country elevator, Am. Soc. Agric. Eng. 44 (3) (2001) 631–638.
[11] L.M. Sonderegger-Wakolbinger, C. Stummer, An agent-based simulation of customer multi-channel choice behavior, CEJOR Cent. Eur. J. Oper. Res. 23 (2015) 459–477.
[12] F. Meyer, G. van der Burgh, The competitiveness of the South African sunflower value chain, Oilseeds Focus 149 (2015) 28–31.
[13] S. Brailsford, L. Churilov, B. Dangerfield, Discrete-Event Simulation and System Dynamics for Management Decision Making, John Wiley & Sons, Incorporated, 2014.
[14] J. Li, Z. Sheng, A multi-agent model for the reasoning of uncertainty information in supply chains, Int. J. Prod. Res. 49 (19) (2011) 5737–5753.
[15] S. Lemieux, S. D'Amours, J. Gaudreault, J.-M. Frayret, Agent-based simulation to anticipate impacts of tactical supply chain decision-making in the lumber industry, CIRRELT (51) (2009) 1–30.
[16] R. Huber, H. Xiong, K. Keller, R. Finger, Bridging behavioural factors and standard bio- economic modelling in an agent- based modelling framework, J. Agric. Econ. (2021) 1–29.
[17] S.R. Yadav, A.K. Malik, Operation Research, first ed., Oxford University Press, ISBN: 9780198096184, 2014.
[18] S.C. Brailsford, T. Eldabi, M. Kunc, N. Mustafee, A. F.Osorio, Hybrid simulation modelling in operational research: A state-of-the-art review, European J. Oper. Res. 278 (2018) 721–737.
[19] O.M. Ülgen, E.J. Williams, Simulation methodology, tools, and application, in: Maynard's Industrial Engineering Handbook, fifth ed., McGraw-Hill, 2001, Chapter 90.
[20] C. Winckler, D. Boshoff, The valuation of commercial grain silos, second ed., Acta Structilia 23 (2016) 25–56.
[21] K. Kristensen, I.A. Rasmussen, The use of a Bayesian network in the design of a decision support system for growing malting barley without use of pesticides, Comput. Electron. Agric. 33 (2002) 197–217.
[22] D. Maes, S.V. Passel, An agent-based model of farmer behaviour to explain the limited adaptability of Flemish agriculture, Environ. Innov. Soc. Transit. 22 (2016) 63–77.
[23] A. Malawska, C.J. Topping, Evaluating the role of behavioral factors and practical constraints in the performance of an agent-based model of farmer decision making, Agric. Syst. 143 (2016) 136–146.
[24] M. Mogliaa, K.S. Alexander, M. Thephavanhc, P. Thammavongd, V. Sodahake, B. Khounsyc, S.V.S. Larson, J. Connell, P. Case, TA Bayesian network model to explore practice change by smallholder rice farmers in Lao PDR, Agric. Syst. 164 (2018) 84–94.
[25] P. Poppenborg, T. Koellner, A Bayesian network approach to model farmers' crop choice using socio-psychological measurements of expected benefits of ecosystem services, Environ. Model. Softw. 57 (2014) 227–234.
[26] T.G. Williams, D.G. Brown, A. Agrawal, S.D. Guikema, Let the farmer decide: examining smallholder autonomy in large-scale land acquisitions with an agent-based model, Environ. Res. Lett. 16 (2021) 2–16.
[27] N. Singh, N. Gupta, Decision making in integrated pest managements and Bayesian network, Int. J. Comput. Sci. Inf. Technol. 9 (2) (2017) 31–37.
[28] K. Happe, K. Kellermann, A. Balmann, Aagent-based analysis of agricultural policies: an illustration of the agricultural policy simulator AgriPoliS, its adaptation and behavior, Ecol. Soc. 11 (2006) 1–27.
[29] S.A. Abdulkareem, Y.T. Mustafa, E.-W. Augustijn, T. Filatova, Bayesian networks for spatial learning: a workflow on using limited survey data for intelligent learning in spatial agent-based models, GeoInformatica 23 (2019) 243—268.
[30] J.A. Verstegen, A. Ligtenberg, M. Wachowicz, Learning actors in spatial planning: Incorporating Bayesian networks in an agent based model, in: AGILE International Conference on Geographic Information Science, Vol. 13, 2010, pp. 1–10.
[31] A.C. Constantinou, N. Fentona, M. W, L. Radlinski, From complex questionnaire and interviewing data to intelligent Bayesian Network models for medical decision support, Artif. Intell. Med. (2016).
[32] S. Nadkarni, P.P. Shenoy, A causal mapping approach to constructing Bayesian networks, Decis. Support Syst. 38 (2004) 259–281.
[33] C.S. Spetzler, C.S.S.V. Holstein, Probability encoding in decision analysis, Manage. Sci. 22 (3) (1975) 340–358.
[34] A.D. Waal, P. de Villiers, H. Roodt, N. Moorosi, G. Pavlin, Construction and evaluation of Bayesian networks with expert-defined latent variables, in: Proceedings of the Conference of Information Fusion, Vol. 19, 2016.
[35] C. Li, S. Mahadevan, Sensitivity analysis of a Bayesian network, ASCE-ASME J. Risk Uncertain. Eng. Syst. 4 (2018) 1–10.
[36] A. Mahfouz, D. Allen, A. Arisha, R. Elbert, M. Gleser, A post-brexit transportation scenario analysis for an agri-fresh produce supply chain, in: 2019 Winter Simulation Conference (WSC), 2019, pp. 1789–1800, http://dx.doi.org/10.1109/WSC40007.2019.9004790.
[37] Agricultural Product Standards Act of 1990. Act No. 119 of 1990.
[38] R. Mcpherson, Budget; plan ahead and ensure success, 2021, https://www.grainsa.co.za/budget;-plan-ahead-and-ensure-success. [Accessed: 28 July 2022].
[39] S. Heckbert, T. Baynes, A. Reeson, Agent-based modeling in ecological economics, Ann. New York Acad. Sci. 1185 (2010) 39–53.
[40] H.S. Sætra, Exploring the use of agent-based modeling (ABM) in mixed methods research, Rev. Castell.-Manchega Cienc. Soc. 22 (2011) 15–31.