



Optimized Parameter Estimation and Integrating Neural Network Forecasting of Dynamic Plant-Livestock Model For Early Warning in Agro-Environment Control Systems

Agolame M. S. Puoetsile^{1,*}, Mokaedi V. Lekgari¹, Semu M. Kassa¹, G. Mengistu Tsidu²

¹ Department of Mathematics and Statistical Sciences, Botswana International University of Science and Technology

² Department of Environmental and Earth Sciences, Botswana International University of Science and Technology,
mengistug@biust.ac.bw

Abstract

The research utilizes the Lotka-Volterra prey-predator model to study Plant-Herbivore dynamics, focusing on the relationship between traditional livestock farming and vegetation conditions. Advanced methods are developed to improve the precision and efficiency of parameter estimation in these models. Neural networks are incorporated to enhance forecasting abilities, and an extension of the Plant-Herbivore models includes Botswana's climate and livestock variables. Efficient parameter space exploration is achieved using the Runge-Kutta method along with Multistart and the local solver *fmincon* in MATLAB. This method improves parameter estimation accuracy. To address the impact of homogeneity assumptions in the data, estimate aggregation through weighting and time conversion is applied. Furthermore, the study investigates the use of nonlinear least squares to further refine the process, allowing for the identification of parameters that best fit observed livestock data, even with non-linearity. By using optimized parameter estimation techniques along with normalized nonlinear least squares, the cumulative error was reduced from an initial 1563.4521 to a final value of 0.0038, well within the specified thresholds (1.0, 0.1, and 0.01). Comparisons between Autoregressive Integrated Moving Average (ARIMA) and Neural Network Auto-Regressive (NNAR) models showed that NNAR models outperformed ARIMA models, with lower variance estimates (0.000004 - 0.000562) compared to ARIMA (0.103 - 0.155). NNAR models displayed Mean Error (ME) values ranging from -0.0012 to 0.0140, indicating a close match between forecasts and actual values with minor deviations. As a result, NNAR forecasting was used for predicting soil moisture, death, and harvest rates, which were integrated into the extended Plant-Herbivore model. This integration enabled the estimation of livestock production trajectories for 2021-2022, along with corresponding interpretations. The study also assessed the uncertainty propagation from NNAR forecasts onto the Plant-Herbivore dynamic model, revealing an increase in uncertainty with longer lead times.

Keywords Plant-Herbivore interaction model, Optimization, Parameter estimation, Neural Network Auto-Regressive model, Forecasting, Estimate weighting

AMS 2010 subject classifications 37N25, 62P12, 65D30, 62M10, 65K10

DOI: 10.19139/soic-2310-5070-1906

1. Introduction

1.1. Background

Herbivore-plant interactions offer insights into fundamental life processes observable in ecological patterns. Since their initial coexistence, plant and animal species have co-evolved. They have formed mutually advantageous

*Correspondence to: Agolame Puoetsile (agolame.puoetsile@studentmail.biust.ac.bw) Botswana International University of Science and Technology, Palapye, BW.

relationships at various levels, especially regarding feeding behavior [1]. Such relationships are based on multiple factors, ranging from mutual benefits to observed parasite interactions and the control of plant growth by wild ruminants. Due to the importance of Plant-Herbivore interactions, mathematical ecology has long concentrated on the dynamic interactions between these two groups of organisms [2]. Any event that disrupts herbivore activity equilibrium, such as environmental changes, species extinction, or bio-invasions, can lead to significant changes affecting more than just the impacted species [3]. Key policy areas like land management, environmental preservation, and animal husbandry rely on understanding the interactions between herbivores and plants. This understanding has led to the creation of dynamical system models of populations, which have garnered significant attention due to the remarkable diversity of dynamic behaviors exhibited by plant and herbivore species [4, 5, 6, 7, 8]. In traditional models, predator-prey systems are similar to plant-herbivore systems. Plant-herbivore dynamics do not usually follow a specific pattern. Instead, the demographic characteristics of plant and herbivore populations, along with the timing, type, and degree of density dependence they exhibit, significantly influence evolution [5]. According to [9], temperature is one of several crucial factors affecting herbivore performance, making it essential to study and document these changes as in [10], to establish applicable solutions and countermeasures. Strong ecological evidence suggests that plant population dynamics impact Plant-Herbivore interactions more significantly than herbivore populations [6].

However, as major components of the biosphere within the climate system, plants and herbivores interact with other components of the climate systems (e.g. atmosphere, hydrosphere, land surface). As a result, several studies have been conducted to include these interactions. For example, to explore the impact of climatic conditions in predator-prey models, several researchers have used periodic mathematical models. Scholars such as [11] investigated the periodic coefficients of the two-species Lotka-Volterra dispersal system, [12] generalized the logistic equation to include time-dependent periodic coefficients, [13] modified general systems of prey-predator dynamics by also assuming the coefficients are periodic functions of time, [14] investigated the existence of periodic solutions for some planar systems with application to Lotka-Volterra equations while [15] studied permanence, extinction, and periodic solution of periodic predator-prey system under the assumption of a Holling Type II functional response.

On the other hand, studies such as [16] which introduced specific density-dependent herbivore mortality rate function, assumed parameters as constant which is not representative of reality. [17] also did not consider periodic and time-dependent functions of the plant growth rate and feeding rates in their works. Recent work by [18] that used Soil Moisture Active Passive (SMAP) satellite data to simulate the spatial distribution of plants and the potential livestock production level in Botswana did not address several key areas. One aspect not considered in their work was the use of observations of state variables to calibrate the model parameters, instead, the parameters were either assumed, estimated or adopted from relevant literature. Lastly, none of these works have been able to forecast the yield from framing activities due to difficulties in predicting parameters. These shortcomings are thus the gaps identified and discussed as the justification of this paper.

Probabilistic and trajectory model forecasting are two methods employed to predict future events or outcomes using historical data and mathematical models. In trajectory forecasting of Plant-Herbivore dynamics, the system's initial conditions are defined, and ODEs are utilized to determine the system's evolution over time [19]. This approach is frequently used in recent mathematical modelling literature for climate data forecasting [20, 21, 22]. However, trajectory forecasting for the Plant-Herbivore dynamic model, as indicated in [18], is challenged by the periodic functional forms that include soil moisture data. Consequently, traditional trajectory forecasting has proven inadequate, making probabilistic forecasting techniques a viable alternative. Probabilistic forecasting accounts for the inherent uncertainty and variability in the data, offering a range of potential future scenarios along with their associated probabilities. Researchers heavily rely on time series data-based forecasts of these climatic variables, as highlighted by [2]. Their results are ultimately applied to agro-environmental control systems, which are then optimized. [23] asserts that the development of smart agricultural approaches to transform agricultural systems, ensuring ecological intensification and food security under a changing climate, is the primary motivator for using

modelling and forecasting techniques in ecology. One method under probabilistic forecasting is the Neural Network Autoregressive (NNAR) modelling, which is gaining popularity as a machine-learning technique. Several studies have shown the effectiveness of NNAR models in forecasting various time series data. For example, [24] proposed a NNAR model for financial time series forecasting, achieving superior performance compared to traditional methods such as ARIMA. Similarly, [25] used a NNAR model to forecast short-term electricity load, outperforming other forecasting techniques in terms of accuracy and flexibility.

1.2. Justification

The absence of real data for parameter estimation, as identified in previous studies such as [18, 4, 26], greatly hampers progress in modeling Plant-Herbivore interactions. Although simulation techniques have been used, their effectiveness is compromised by the lack of empirical data. This limitation poses a serious challenge when applying these models to real-world scenarios and decision-making processes. Consequently, there is an urgent need for research efforts dedicated to the acquisition and utilization of authentic field data, which would enable more accurate parameter estimation and subsequently enhance the predictive power of ecological models.

Another important research gap in the field of parameter estimation is the need to effectively bridge the variations in base data, often represented by the Plant-Herbivore model, and the observed data. This issue arises due to the inherent differences in the temporal sampling of collected data and model data. Resolving this gap is crucial for accurate parameter estimation and improving the reliability of modelling and analysis in various domains.

Furthermore, tackling the estimation of parameters for dynamic model systems that exhibit non-convex parameter relationships is a formidable challenge that demands innovative approaches. Traditional optimization techniques often struggle to identify global optima or may overlook multiple optimal solutions. Addressing this gap can contribute to the development of specialized optimization methods tailored to the intricacies of Plant-Herbivore systems.

Lastly, addressing the lack of forecasting capabilities for Plant-Herbivore models is an essential endeavor, especially considering the integration of climatic factors within these systems. The complexities arising from the interplay of ecological dynamics and model parameters necessitate novel forecasting techniques. Developing predictive models that capture the seasonality and patterns for soil moisture, death and harvest rates data within the Plant-Herbivore model can significantly aid in understanding the long-term implications of plant-herbivore interactions, informing conservation efforts and management strategies.

1.3. Contribution of Study

This study aims to address these research gaps by presenting a novel approach to optimized parameter estimation, focusing on the non-convexity in the optimization problem caused by the parameter relationships in the Plant-Herbivore dynamic model. We plan to utilize the mathematical model outlined by [4] and further developed by [26], to explore the behavior of the plant-herbivore system under specific conditions. This paper also seeks to incorporate machine learning techniques, such as Neural Network forecasting models, to generate reliable predictions of soil moisture, mortality, and harvest rates used within the dynamic model. By employing advanced global optimization algorithms, the proposed methodology strives to estimate system parameters with greater accuracy, capturing the dynamic interactions between plant and livestock components in Botswana. Additionally, by integrating neural networks with auto-regressive modeling, the combined NNAR models will improve forecasting accuracy by taking into account both historical data and the system's inherent dynamics. The contributions of this study are outlined below.

- Introduce more realistic parameter estimates using observed data for both state variables and some model parameters;
- Bridge the temporal mismatch between observed livestock data and model estimates for estimation of parameters;

- Leveraging optimized parameter estimation techniques to better estimate model parameters;
- Exploit neural network forecasting to forecast soil moisture, death and take-off rates; and
- Forecast livestock production using the Plant-Herbivore dynamic model.

The significance of this research lies in its potential to aid decision-making processes in plant-livestock systems, enabling stakeholders to make informed choices regarding expected harvest, resource allocation, production planning, and risk management. The proposed approach will provide more accurate and reliable forecasts, enabling farmers, policymakers, and agricultural practitioners to optimize their strategies, mitigate risks, and maximize productivity.

1.4. Organization of the Paper

This paper has been organized into four sections. Section 2 describes the methodology with specific focus on the proposed mathematical model for the study, then the objective function and temporal conversions which facilitate the parameter estimation for the model. Section 2 discusses optimization and forecasting techniques that are used in Section 3. In Section 3 the results of the parameter estimation and forecasting are discussed. Remarks and recommendations are given in Section 4.

2. Methodology

2.1. Proposed Model

An extension to the mathematical model given in [4] and further developed in [16] is considered. The proposed model is given as

$$\begin{aligned}\frac{dP}{dt} &= r(t)P \left[1 - \frac{P}{K} \right] - \left(\frac{a}{b+P} \right) PH \\ \frac{dH}{dt} &= H \left[c \left(\frac{a}{b+P} \right) P - D(H) - \mu(H) \right],\end{aligned}\quad (1)$$

where $P(0), H(0) > 0$ and $r(t)$ is the plant growth rate at time t , which is defined by:

$$r(t) = r(\theta(t), m) = m + [1 - m] U_r(\theta(t)) \beta_r(\theta(t)),$$

here $\theta(t)$ is the soil moisture at time t and m is the minimum plant growth rate, $U_r(\theta(t))$ and $\beta_r(\theta(t))$ are given by

$$U_r(\theta(t)) = \exp \left[-(\theta(t) - \theta_c)^2 / 2\pi\sigma(\theta)^2 \right],$$

and

$$\beta_r(\theta(t)) = \begin{cases} 1, & \theta(t) \geq \theta_c \\ \frac{\theta(t) - \theta_w}{\theta_c - \theta_w}, & \theta_w < \theta < \theta_c \\ 0, & \theta(t) \leq \theta_w, \end{cases}$$

respectively. θ_c is the critical soil moisture, θ_w is the wilting point soil moisture and $\sigma(\theta)^2$ is the variance of the soil moisture data.

The model depicts the interplay between a prey population (plant) and a predator population (livestock), whereby the prey population's growth rate is constrained by its carrying capacity (K) and predation by the livestock population. The availability of prey and the rate of harvesting constrain the rate of growth of the predator population. The model outlines a number of parameters, including the intrinsic growth rate of the prey population ($r(t)$), the constant livestock population's predation feeding rate (a), the carrying capacity (K), the harvesting rate ($\mu(H)$), the rate at which prey biomass is converted into predator biomass (c), and the livestock population's

mortality rate due to natural causes ($D(H)$).

The intrinsic plant growth rate ($r(t)$) has three embedded functions within it, these function are $U_r(\theta)$, $\beta_r(\theta)$ and m (which is assumed to be constant) as stated above. $U_r(\theta)$ and $\beta_r(\theta)$ are functions of the soil moisture (θ), that can be obtained from satellite. In this study the SMAP level 3 soil moisture data with spatial resolution of $36 \times 36 \text{ km}^2$ was used. The data taken from the SMAP level 3 is recorded as daily observations for 11 sampled spatial points running from April 2015 to December 2021. (Table 1)

Label	Spatial Point		District
	Latitude	Longitude	
A	-25.451523	24.617443	Borolong - Boteti sub-district
B	-21.996038	28.438864	Central Bobonong - Central District
C	-21.017070	24.497584	Central Boteti - Central District
D	-23.129488	26.883858	Central Mahalapye - Central District
E	-22.503600	26.995233	Central Serowe/Palapye - Central District
F	-24.338668	26.395879	Kgatlang - Kgatlang District
G	-23.633065	25.075264	Kweneng East - Kweneng District
H	-23.547486	23.355904	Kweneng West - Kweneng District
I	-24.828178	25.570083	Ngwaketse West - Ngwaketse District
J	-24.974965	25.205865	Ngwaketse - Ngwaketse District
K	-24.970862	25.796937	South East - Southern District

Table 1. Location of SMAP Soil Moisture.

The density dependent herbivore death rate, is given as a function stated in [4] and it is bounded as follows $d_1 \leq D(F(P)) \leq d_2$, where d_1, d_2 are the minimum and maximum death rates respectfully. We propose the modification of this parameter to be a variable rather than a function as the death rate data for livestock (cattle, goats and sheep) exists for dates from as 2017 to 2019 which is taken from [27]. This change would reduce the number of parameters to be estimated by dropping both d_1 and d_2 thus reducing the complexity of the Plant-Herbivore model stated in [4]. When formulating the death rate for herbivores, it is important to note that there are groups of animals that make up herbivores. Thus, the death rate should account for the death rate of each of these groups, which is defined as follows.

Let $S(t), C(t), G(t), H_{or}(t), D(t)$ be the sheep, cattle, goat, horse and donkey population densities from our data for a particular year such that $H(t) = [S(t) + C(t) + G(t) + H_{or}(t) + D(t)]$.

$$D(t) = [\alpha_1 d(t)_{cattle} + \alpha_2 d(t)_{goats} + \alpha_3 d(t)_{sheep} + \alpha_4 d(t)_{horse} + \alpha_5 d(t)_{donkey}],$$

where $\alpha_1 = \frac{C(t)}{H(t)}$, $\alpha_2 = \frac{G(t)}{H(t)}$, $\alpha_3 = \frac{S(t)}{H(t)}$, $\alpha_4 = \frac{H_{or}(t)}{H(t)}$ and $\alpha_5 = \frac{D(t)}{H(t)}$ are the weights assigned to each animal type such that $\sum_{r=1}^5 \alpha_r = 1$. The individual death rates for each of these groups are represented by $d(t)_{cattle}$, $d(t)_{goats}$ and $d(t)_{sheep}$.

The harvest rate is given as a function of two parameters μ and μ_0 in [4]. Applying the same weighting method used above to this parameter and using the take-off values found in the data set, we can change the harvest rate into a variable as follows:

$$\mu(t) = [\alpha_1 \mu^*(t)_{cattle} + \alpha_2 \mu^*(t)_{goats} + \alpha_3 \mu^*(t)_{sheep} + \alpha_4 \mu^*(t)_{horse} + \alpha_5 \mu^*(t)_{donkey}],$$

with each individual take-off rates for each of these groups represented by $\mu^*(t)_{cattle}$, $\mu^*(t)_{goats}$, $\mu^*(t)_{sheep}$, $\mu^*(t)_{horse}$ and $\mu^*(t)_{donkey}$. The weights α_r are the same as the ones used for the death rates.

Livestock was assumed to reproduce naturally without human intervention [4], and was accounted for by the term $(\frac{H}{h+H})$, where h is the allee constant. In our case, we assume farmer intervention is present in the forms of artificial

insemination and bull swapping, therefore h is taken to be 0. Setting the allee effect to 0 results in $(\frac{H}{h+H}) = 1$. Thus further reducing the complexity by eliminating a parameter. Thus the parameters to be estimated are m , a , c and b with K being assumed to be the same as in [4].

2.2. Objective Function

The objective function is the nonlinear least squares equation given by

$$e = \sum_{t=1}^n \left[\frac{(P(t)_{obs} - P(t)_{est})^2}{P(t)_{obs}^2} + \frac{(H(t)_{obs} - H(t)_{est})^2}{H(t)_{obs}^2} \right], \tag{2}$$

where n is the total number of observed data points, $P(t)_{obs}$ and $P(t)_{est}$ are the observed and estimated plant biomass, respectively. Similarly $H(t)_{obs}$ and $H(t)_{est}$ are the observed livestock totals from [27] and estimated population totals respectively. The nonlinear least squares equation under (2) has been normalized by dividing each contributing group by the square of its observed value.

According to the Food and Agriculture Organization (FAO), an agency of the United Nations [28], the average live weight of mature cattle in Botswana is around 400-450 kg, with some breeds such as the Tswana and Ngamiland being larger and heavier. For goats the average live weight of mature goat in Botswana is around 25-30 kg, although this can vary depending on the breed and gender of the goat. Whereas the average live weight of mature sheep, donkey and horse in Botswana is 35-40, 200-400 and 700-1000 kg respectfully. It should be noted that these weights are averages, and actual weights can vary depending on factors such as age, gender, breed, and feeding practices. Additionally, there may be regional or local variations in weight due to differences in environmental conditions and management practices.

In order to convert the different yearly animal counts into yearly livestock data with the units in (kg), we simply multiply each animal groups yearly count by their corresponding lower bound average weight. The sum of these values then gives the observed yearly livestock weights which are referred to as $H(t)_{obs}$.

2.3. Time Conversions

The solutions of our modified Plant-Herbivore model are 388 weekly plant biomass and livestock densities for a 36×36 square km spatial segment in each of the 11 spatial points. This presents a problem when comparing this data to the observed livestock densities which are yearly and are country wide values. As such, it is necessary to convert the solutions from weekly regional estimates to yearly country wide estimates. This task has been achieved using the following procedures.

Each of the 11 spatial points has an area segment of 36×36 square km, therefore to get a livestock density estimate for the entire region we calculate the multiplying coefficient (a_i) of each region as the stated below:

$$a_i = \frac{A_i}{n_i}, \quad \text{for } i = 1, 2, \dots, 11,$$

where A_i is the land mass of the district containing the i^{th} spatial point and n_i the number of spatial points that fall within the same district as the i^{th} spatial point . Therefore the estimated i^{th} region's t^{th} weekly livestock weight in kg is given by

$$H_i(t)_{esti} = a_i H_i(t),$$

where $H_i(t)$ is the i^{th} region's weekly model livestock density estimate for the t^{th} week measured in (kg/m²). The above equation can be formulated to obtain yearly weighted livestock estimates for the country.

For the general case this includes 2016, 2017 and 2018:

$$\bar{H}(k)_{esti} = \frac{1}{52} \sum_{i=1}^{11} \sum_{t=(k-1)52+1}^{k \times 52} a_i H_i(t).$$

Since the satellite data runs from March 2015, we have adjusted for the year to be averaged by 35 weeks that make up 2015. Therefore for $k = 1$ we have :

$$\bar{H}(1)_{esti} = \frac{1}{35} \sum_{i=1}^{11} \sum_{t=1}^{35} a_i H_i(t).$$

Another factor to consider is the relative contribution factor of each region to the overall country livestock population density. This factor can be calculated using information from the 2019 annual livestock census, in which animal counts for each region was given for the year 2019. Thus the formulation of the contribution factor is

$$w_i = \frac{H(i)_{observed}}{H_{2019}}, \quad \text{for } i = 1, 2, \dots, 11, \quad (3)$$

where $H(i)_{observed}$ is the observed livestock density of the i^{th} region in 2019 and H_{2019} is the total country wide livestock density for 2019. For the 2019 ($k=4$) case:

$$\bar{H}(5)_{esti} = \frac{1}{52} \sum_{i=1}^{11} \sum_{t=209}^{260} a_i w_i H_i(t).$$

2.4. Optimization Algorithm

We have chosen the **function minimization with constraints** also known as *fmincon* as the local solver for our optimization problem. It is a popular optimization algorithm in MATLAB that performs constrained optimization using nonlinear programming and is widely used for solving a variety of optimization problems. It is based on the interior-point method, which is an iterative optimization method that works by solving a sequence of linearized sub-problems. It can handle both linear and nonlinear constraints, as well as nonlinear objective functions which makes it ideal for handling our nonlinear least squares objective function.

fmincon is subject to the restriction of only identifying a local minimum solution to the optimization function, therefore Multistart optimization technique is used to find the global minimum for the nonlinear least squares function that has multiple local minima. The algorithm then returns the solution with the lowest objective function value among all the runs. By running the optimization algorithm from different starting points, the chances of finding the global minimum are increased. To insure that the random sample is spread out evenly across the domain of the objective function then we can set a loop of re-sampling for multiple Multistart runs and compare the lowest values from each iteration of multistart sample.

2.5. Forecasting Technique

The proposed methodology in this study for time series forecasting is based on a Neural Network Autoregressive (NNAR) model. The NNAR model combines the power of artificial neural networks (ANNs) with the autoregressive concept. The autoregressive component allows the model to capture the temporal dependencies in the time series data, while the neural network component enables the model to learn complex patterns and non-linear relationships. Thus, we will be applying this technique to forecast both the death and harvest rates as well as the soil moisture at chosen spatial-temporal grids.

2.5.1. Network Architecture: A feed-forward neural network will be fitted with lagged values of each specific variable as inputs and a single hidden layer with a varying number of nodes. The inputs are for lags 1 to p (Number of seasonal/nonseasonal lags used as inputs), and lags m to (mp) where m is the frequency of the variable

[29]. A total of repeats networks are fitted, each with random starting weights. These are then averaged when computing forecasts. The network is trained for one-step forecasting. Multi-step forecasts are computed recursively. For non-seasonal data, the fitted model is denoted as an NNAR(p, k) model, where k is the number of hidden nodes.

2.5.2. Formulation of Forecast Confidence Intervals: The NNAR forecast function generates prediction intervals using simulations since NNAR models are nonlinear autoregressive models and prediction intervals cannot be analytically derived [29]. The error series is assumed by the function to be homoscedastic (weakly normally distributed as well). When this assumption was not met, log transformations and the Box-Cox transformation (with λ set to 0.5) are used to coerce the residuals to be approximately homoscedastic. It then iteratively simulates the model's future sample pathways by resampling from the historical values or randomly creating a value from a normal distribution. It did this by repeatedly simulating a sample path for each variable in the future. By repeatedly simulating sample paths, it builds up knowledge of the distribution for all future values based on the fitted neural network. This is then repeated extensively (default 1000), thus being able to visualize the forecast distributions. The extremes of these simulations are considered to reflect a 95% confidence intervals of the each of the parameter forecasts. These are then used in the Plant-Herbivore model to propagate new model forecasts.

2.6. Models Performance Metrics

The accuracy of the forecasts can be evaluated using various metrics, these are $ME = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$ (Mean Error), $RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$ (Root Mean Squared Error), $MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$ (Mean Absolute Error), $MPE = \frac{1}{n} \sum_{i=1}^n \left(\frac{y_i - \hat{y}_i}{y_i} \right) \times 100\%$ (Mean Percentage Error), and $MAPE = \frac{100}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right|$ (Mean Absolute Percentage Error). These measures provide insights into the forecast errors and the overall accuracy of the models and performance of the NNAR model can be compared with Autoregressive Intergrated Moving Average (ARIMA) forecasting methods.

3. Results & Discussion

3.1. Parameter Estimation

In this section, we present a comparative analysis of two classes (initial and adjusted iterations) of MultiStart optimization results. The objective of the optimization was to minimize an initial objective value through multiple runs of a local solver algorithm. The focus of the analysis was on the convergence rates, final objective values, and the optimized variable values. Table 2 shows a summary of the initial iteration results.

Table 2. Initial Iteration Results.

Iteration	N° of Solvers	Initial (e^2)	Final (e^2)	Convergence (%)
1	10	1563.4521	81.3626	60
2	100	1563.4521	5.8666	47
3	500	1563.4521	5.7372	64.5

The first set of results from the initial iterations class comprised of 10 local solver runs, out of which 6 converged successfully with a positive local solver exit flag. The final objective function error (81.3626) achieved after optimization was significantly lower than the initial objective (1563.4521), indicating a successful optimization process. The optimized variable values of m , a , b and c were 0.79109, 258.9894, 190.1992, and 0.62262,

respectively.

The second set of results included 100 local solver runs, with 47 converging successfully. The final objective function value (5.8666) achieved after optimization was even lower than the previous set, demonstrating further improvement. The optimized variable values of m , a , b and c were 0.42314, 227.491, 286.4273, and 0.48007, respectively. When the local solver *fmincon* is initialized 500 times, we see an increase in convergence rate from the second run but a minimal change in the final objective function. This set of results indicates that increasing the number of local solver runs can enhance the optimization process and yield more favorable outcomes.

The total percentage error for each of the final objective function is 5.20%, 0.38 % and 0.37 % of the initial error. The percentage-based analysis reveals a clear trend of continuous improvement in error reduction rates, signifying that the optimization process was progressively refining the objective function with each iteration. Under the the hypothesis of plant density simulations amplifying the true error in estimation in our initial iteration results, we tested this as shown in Table 3.

Table 3. Adjusted Iteration Results.

Iteration	N° of Solvers	Initial (e^2)	Final (e^2)	Convergence (%)
1	100	0.93729	0.83915	49
2	500	0.84709	0.83216	99.6

After adjustments in the objective function, a significant decrease in the initial objective functions from Table 2 and Table 3 can be noted. This confirms that the plant simulations amplified the error in estimation therefore removing it was justified. Further adjustments in the form of testing multiple population start points, narrowing of parameter intervals and increases in local solvers were made. The initial objective value was reduced to 0.84709 from 1563.4521. The numerous runs initiated by MultiStart show that 498 out of 500 local solver runs successfully converged to optimal solutions. The convergence of such a high number of runs with positive local solver exit flags indicates that narrowing of the estimation intervals was significant. The high number of convergence also reveal how robust and reliable Multistart with *fmincon* and non-linear least squares are in the optimization process.

Table 4. Refined Iteration Results.

Iteration	N° of Solvers	Initial (e^2)	Final (e^2)	Convergence (%)
3	100	0.20624	0.12908	99
3	500	0.20624	0.12850	99.6

To investigate whether the weighting parameter w_i was increasing the error of estimation, we tested whether its presence had a significant impact on the initial and final the objective functions values. The final objective value achieved after the optimization process was significantly improved to 0.12908 achieving the threshold set at 1.0 (Table 4). This improvement can be attributed to the change in weighting parameter w_i for the year 2019. It is important to note that the error function is normalized and thus the thresholds were assigned with this in mind. The 3 thresholds used in the estimation are 1.0, 0.1 and 0.01, all of which are applied in the order stated throughout the optimization process.

The last iteration yielded the best results compared to previous iterations as shown in Table 5. The objective function had values that were lower than the adjusted threshold (0.01). Increasing the number of solvers did not produce much of a difference between the runs thus the results of optimization led to the minimum plant growth rate (m) = 0.80656, consumption/feeding rate of herbivores (a) = 0.98903, half saturation constant of herbivore feeding on the plant biomass (b) = 185.7086, and Conversion rate of plant biomass to herbivore offspring (c) =

Table 5. Final Iteration Results.

Iteration	N° of Solvers	Initial (e^2)	Final (e^2)	Convergence (%)
4	10	0.012349	0.0038387	100
4	100	0.012349	0.0038326	97

0.99009 from the first run. Figure 1 shows graphs of livestock population weights and livestock population using the optimized parameters for the period 2015-2019 along with their corresponding observed data values.

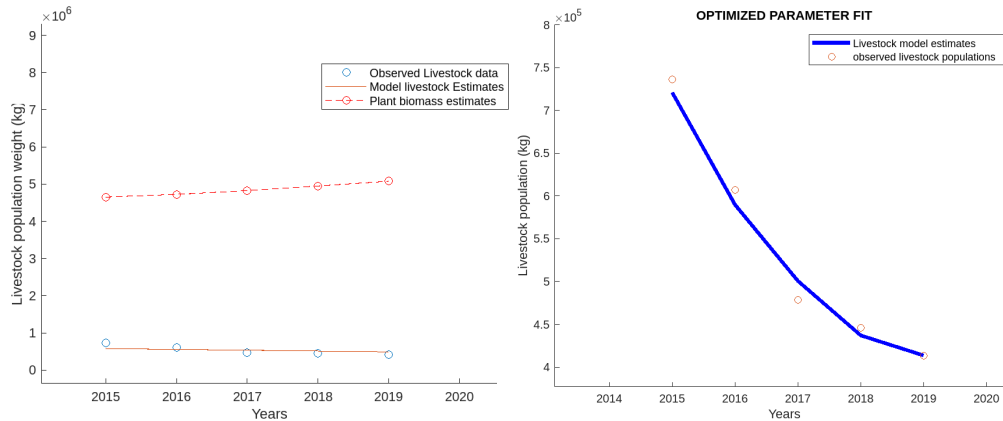


Figure 1. Optimal Parameter Plot.

Model runs based on the final optimized parameters values have produced results that have well captured the observations. Refining the death rates actually yielded better fits for our model. The final objective value falling within the set thresholds (1.0, 0.1 & 0.01) and the high convergence rate achieved by the MultiStart optimization method all demonstrate the effectiveness and reliability of the methods. The optimized parameter values offer valuable insights into the parameter configurations associated with optimal solutions despite limited observational data points. In regards to the sensitivity analysis of parameters, the authors in [18] have investigated the impact of any change on these parameter values on the output of the model using a global sensitivity analysis procedure. In the reference it is reported that the model is highly sensitive to the variation in the parameter values b (half saturation constant of the herbivores), a (animal feeding rate), and m (the minimum plant growth rate). Overall, while the improvement in the objective value are significant, it is still does not account for the models relationship with plant biomass (vegetative coverage); this means that the parameters do not offer an accurate representation of the plant and livestock interaction.

3.2. Forecasting

In constructing the dynamic model, we utilize data on soil moisture from 11 regions, as well as livestock harvest and death rates in Botswana. To effectively use the dynamic model for forecasting, it is crucial to have these variables available for the desired forecast period. This requires integrating additional models within the dynamic model to accurately predict soil moisture, harvest, and death rates. In this research, we develop NNAR models to forecast these variables, as detailed in the following sections.

3.2.1. Harvest and Death Rates Forecasting

From Figure 2, we can see that livestock death rates have always been higher than harvest rates. This is a cause for concern since it suggests that farmers are losing more cattle to disease and other causes of death than they sell or feed off. However, these high death rates can be explained by the criteria used by Statistics Botswana (SB) to classify a livestock death. According to SB cattle lost to thieves, gone missing, culled and killed by car accidents

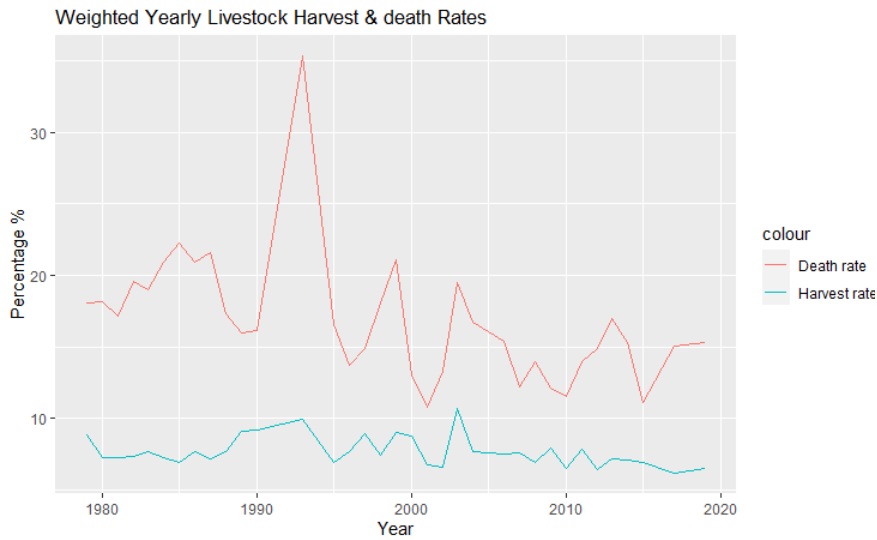


Figure 2. Weighted Livestock Harvest and Death Rates : Statistics Botswana.

are all included in these rates. This classification amplifies the rates with the inclusion of human intervention rather than natural death rates.

Table 6. NNAR Output Summaries for Regions A,B,C & D.

Model- NNAR(,)	harvest Rate - (1,1)	Death rate -(1,1)
σ^2	6.233	0.09744
Transformed data σ^2	0.8067	0.02499
Weights	4	4

The NNAR model for death and harvest rate are relatively simple with each having only 4 weights, 1 lagged input, 1 neuron in the hidden layer with a linear model structure. This can be attributed to the relative size of the data which has only 41 observations. The forecast confidence intervals for the harvest rates are narrower than that of the death rate as the data is less erratic in nature as shown in Figure 2 . Square root transformation of both time series had significant improvements in σ^2 estimates, with initial estimates of 6.233 and 0.8067 which later reduced to 0.09744 and 0.02499 for death and harvest rates respectively.

Figure 3 shows the 3 years forecasts for the death and harvest rates. There seems to be a decreasing trend and a non-constant variance in the time series data. A large spike in the death rate during the early 90’s is seen and linked to historical events of lung disease (and foot & mouth). The highest harvest rates can be seen in the period between the 1990 and 2005, with the largest being in 2003.

3.2.2. Soil Moisture Forecasting

This section includes a comparison of estimated model variances and forecast accuracy results for NNAR and ARIMA models for two levels of training and testing data sets. Table 7 below shows the estimated variances by both forecasting techniques. The ETS models tend to have higher variance estimates than NNAR models as shown in Table 7. Higher variance estimates are due to relatively large smoothing parameters, resulting in more weight being assigned to recent observations and less to older ones. ETS models generally have a simpler structure compared to the flexibility of NNAR models.

Table 8-9 present the accuracy measures obtained from the ARIMA (Error-Trend-Seasonality) and NNAR forecasting models for various spatial points. The accuracy measures serve as quantitative indicators to assess the performance and reliability of the forecasting models. The accuracy measures are for each spatial point.

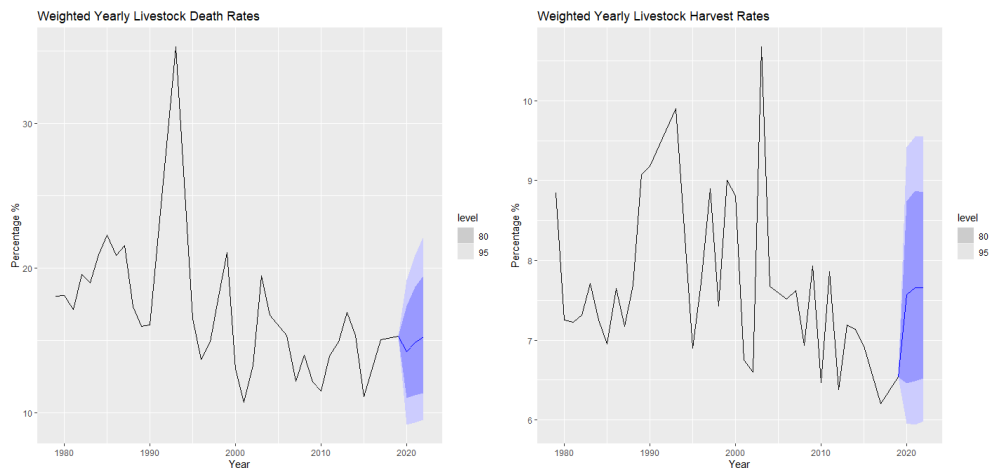


Figure 3. Harvest & Death Rate Forecast.

Table 7. Comparison of Estimated Variances

KEY	$\hat{\sigma}^2$ ETS Model	$\hat{\sigma}^2$ NNAR Model
A	0.155	0.00025000
B	0.103	0.00040800
C	0.150	0.00008930
D	0.176	0.00010100
E	0.136	0.00047300
F	0.150	0.00052100
G	0.155	0.00000429
H	0.145	0.00056200
I	0.123	0.00018000
J	0.109	0.00039500
K	0.149	0.00037900

The spatial points, labeled A to K, represent specific locations or regions for which the forecasting models were applied. Each spatial point is associated with its respective accuracy measures.

Spatial Point	ME	RMSE	MAE	MPE	MAPE
A	-0.0012	0.0363	0.0294	-19.0546	39.4099
B	-0.0038	0.0292	0.0227	-16.3578	30.7445
C	0.0003	0.0348	0.0269	-16.0818	35.9904
D	0.0032	0.0380	0.0266	-11.4256	32.0214
E	0.0063	0.0342	0.0269	-5.0689	30.1192
F	0.0041	0.0337	0.0261	-10.9046	34.2076
G	-0.0026	0.0349	0.0285	-17.2919	36.7096
H	-0.0027	0.0332	0.0269	-20.2358	38.2159
I	-0.0018	0.0299	0.0244	-14.1769	31.9589
J	-0.0140	0.0272	0.0235	-28.8870	36.5995
K	-0.0017	0.0304	0.0246	-13.8183	30.8865

Table 8. Accuracy Metrics for NNAR Forecasting Models.

Spatial Point	ME	RMSE	MAE	MPE	MAPE
A	0.0022	0.0363	0.0288	-14.3299	37.1059
B	-0.0027	0.0300	0.0230	-15.2511	30.9511
C	-0.0021	0.0357	0.0282	-19.6111	38.3375
D	0.0010	0.0381	0.0270	-14.3997	33.3792
E	-0.0024	0.0337	0.0283	-16.2875	35.0846
F	-0.0033	0.0344	0.0275	-21.4659	39.1589
G	-0.0056	0.0338	0.0279	-21.0717	37.1336
H	-0.0079	0.0346	0.0290	-27.9470	43.1502
I	-0.0074	0.0295	0.0252	-21.3041	34.8094
J	-0.0107	0.0256	0.0218	-23.9857	33.2574
K	-0.0047	0.0306	0.0252	-17.6019	32.7280

Table 9. Accuracy measures for ARIMA forecasting models.

For the NNAR forecasting model (Table 8), it can be observed that the ME (Mean Error) values range from -0.0012 to 0.0140, indicating that the model's forecasts are generally close to the actual values with small deviations. The RMSE (Root Mean Squared Error) values vary between 0.0272 to 0.0380, suggesting that the NNAR model has moderate to low error dispersion. Similarly, the MAE (Mean Absolute Error) values range from 0.0174 to 0.0356, representing the absolute magnitude of the errors, which are relatively small.

The MPE (Mean Percentage Error) values range from -5.0689 to -28.8870, signifying that the NNAR model tend to underestimates (negative MPE) the forecasts, but the overall average percentage error is reasonably close to zero. The MAPE (Mean Absolute Percentage Error) values span from 30.1192 to 39.4099, indicating that, on average, the NNAR model's percentage errors can be up to 39.41.

In contrast, the ARIMA forecasting model (Table 9) displays similar patterns for the ME, RMSE, and MAE values compared to the NNAR model. The MAPE values for the ARIMA model range from 30.9511 to 43.1502, which is higher than the NNAR model's MAPE values. This suggests that the ARIMA model has a higher percentage error, making it less accurate for forecasting. The NNAR models outperform the ARIMA models in 9 of the spatial points with the exception of only two spatial points (A & J).

In conclusion, based on the accuracy metrics, the NNAR model is more skillful than the ARIMA model at most spatial points, as it generally shows lower MAPE values, indicating better overall forecasting accuracy.

3.3. The integrated and Optimized dynamical Model forecasts

Figure 4(a) displays the optimized model fit from 2015-2019 (solid blue line) along with the NNAR forecasts for the livestock data from 2019-2022 (red dash line) and the model forecasts that are divided into two, those formulated using observed soil moisture (2019-2021, green dash line) and those that integrated NNAR soil moisture forecasts (2021-2022, blue dash line). The two forecasting methods used above differ on the basis that the NNAR livestock forecasts are generated solely from the the observed livestock data while the model forecasts are trajectories based on parameter such as soil moisture, death and harvest rate forecasts which are integrated within the model. Between the years of 2019-2020 the NNAR livestock forecasts and the model-parameter estimates forecasts were predicting similar population declines which confirms the continuity of the historical livestock data trends.

This agreement is expected as NNAR has low uncertainty at shorter lead time and the Plant-Herbivore model used observed soil moisture data. However, the NNAR forecasts project a recovery in livestock populations from 2020 - 2022, while the model forecasts indicate a further decline. This discrepancy reflects the fundamental difference between the two methodologies. NNAR forecasts are myopic, relying mainly on historical data, and may not

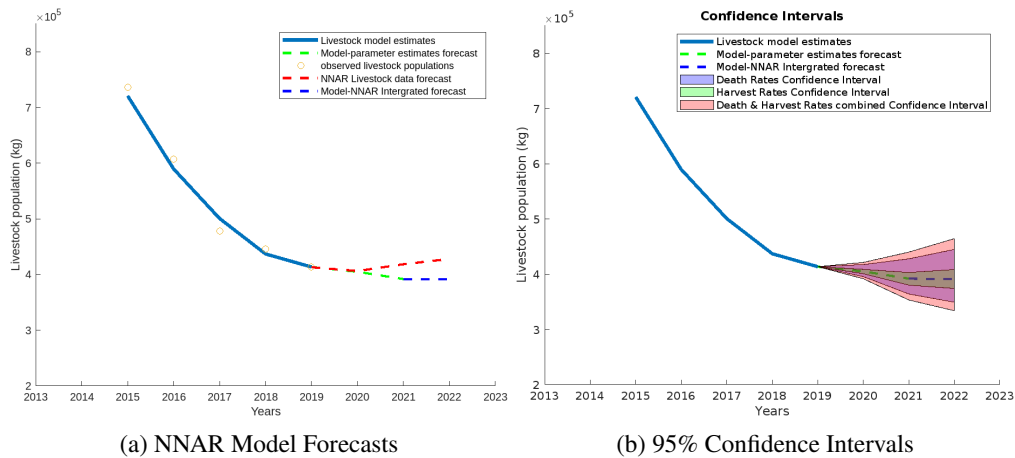


Figure 4. NNAR and Optimized Parameter Plant-Herbivore Model Forecasts and Confidence intervals.

consider external factors like soil moisture or other ecological variables. The NNAR models generate forecasts that are unaffected by climatic and environmental variations thus do not take into account the dynamic nature livestock populations. As a result, the NNAR models are prone to high uncertainty with longer lead time. On the other hand, the Plant-Herbivore model is constrained by observed SMAP soil moisture until 2021 and by soil moisture forecast from NNAR between 2021 to 2022. As noted earlier for the NNAR livestock forecast, soil moisture forecasts from NNAR is also expected to have low uncertainty at shorter lead time. Therefore, the integration of SMAP observation and NNAR soil moisture into the Plant-Herbivore model lead to consistent simulation of livestock population. In addition, the Plant-Herbivore model given by Eqn (1) is an initial value problem whose forecasts depend on the initial values. Therefore, smooth evolution of livestock population that retains historical trend shows physical consistence inherent in the model.

Figure 4(b) shows the possible forecast range with changes arising from NNAR inaccuracies in each of the three parameters mentioned. The corresponding 95% confidence intervals are derived from death rates, harvest rates, soil moisture forecasts and a combination of the rates. These confidence intervals were derived from the extreme values obtained from the 95% confidence intervals of the each of the parameter forecasts. The forecast are weakly sensitive to changes in soil moisture due to the compounding nature of the estimates from weekly to yearly values, as such the model forecast show very slight deviations. This does not apply to the harvest rate changes which result in wider range of livestock estimates than soil moisture but less than that of death rate changes. This is due to the harvest rate being smaller (on average) than the death rates and consequently having a smaller standard deviation. The widest of the confidence intervals is the harvest and death rate combined intervals, which use both extremes of both parameters. This results in higher forecast ranges due to the direct additive relationship these parameters have with the livestock estimate. For example, the livestock estimate by the end of 2022 by the dynamic model can accumulate $\pm 6 \times 10^4$ kg of error from nearly zero in 2019. This figure is about a 17% deviation for lead time of three years and may not affect the use of such a model for early warning systems to prevent related hazards.

4. Conclusions

The Plant-Herbivore models proposed by [26, 18] served as a good reference in studying the relationship between traditional livestock production and vegetative conditions. The models were found to effectively model the complex interplay of parameters in Plant-Herbivore interactions. These models were extended in order to incorporate the existing information for Botswana’s climate and livestock factors thus addressing the absence of real data for parameter estimation. Estimate aggregation through weighting and time conversion are incorporated

to reduce the error brought about by the assumption of homogeneity within the data. This approach effectively bridges the variations in base data, often represented by the Plant-Herbivore model, and the observed data.

The utilization of the Runge-Kutta method with Multistart and local solver *fmincon* enabled efficient exploration of parameter spaces, improved the accuracy of parameter estimation, with objective functions achieving set thresholds in each of the runs. The combination of these optimized parameter estimation techniques, with the normalized nonlinear least squares, presented very robust results which provide a promising avenues for dynamic modelling in agro-environmental control systems. The incorporation of nonlinear least squares is explored to further enhance the optimization process, allowing for the identification of parameters that best fit the observed livestock data, even in the presence of non-linearity. The combination of these optimized parameter estimation techniques with the normalized nonlinear least squares, achieved a final cumulative error of 0.0038 from initial error of 1563.4521, falling within the set thresholds (1.0, 0.1 & 0.01). These techniques offered a systematic approach to estimate parameters of the extended Plant-Herbivore models accurately tackling the issue of estimation of parameters for dynamic model systems that exhibit non-convex parameter relationships despite the low number of observation points.

Neural Network forecasting models were integrated into the parameter estimation process by incorporating them as a component of the optimization framework. Soil moisture is a dynamic variable within the model that can exhibit complex temporal and spatial patterns. Neural Networks effectively captured these dynamics by utilizing their inherent ability to model sequential data. The ARIMA models yielded significantly higher variance estimates (0.103 - 0.155) than NNAR models (0.000004 - 0.000562), NNAR models had ME (Mean Error) values ranging from -0.0012 to 0.0140, indicating that the model's forecasts are generally close to the actual values with small deviations. The MAPE values for the ARIMA model range from 30.9511 to 43.1502, which is higher than the NNAR model's MAPE values (30.1192 - 39.4099). The NNAR models were found to generally perform better for smaller numbers of predictions than ARIMA models. NNAR forecasts were found to have narrower confidence intervals than ETS for soil moisture forecasts, the NNAR models were found to better capture the evolving nature of soil moisture and make more informed predictions. NNAR forecasting was used to obtain soil moisture, death and harvest rates forecasts, which were then integrated within the extended Plant-Herbivore model to estimate 2021 - 2022 livestock production trajectories This enables accurate and adaptable forecasting addressing the lack of forecasting capabilities for Plant-Herbivore models.

The paper primarily focuses on the dynamics of Plant-Herbivore interactions and their relationship with traditional livestock production. However, there might be other external factors, such as land use changes, human interventions, or disease outbreaks, that can significantly influence the vegetative conditions and livestock dynamics. These factors are not explicitly considered in the models. The accuracy and reliability of the results heavily depend on the availability and quality of data, the low number of livestock observation points used within the optimization, may limit of the generality of the findings. The increase in NNAR uncertainty with increasing lead time might affect the accuracy of forecast from the Plant-Herbivore model. Therefore, future work should take this limitation into account and explore other robust data-based forecast such as Long Short-term Memory Network (LSTM) used in the field of Deep learning for integration with the Plant-Herbivore dynamic model.

REFERENCES

1. M.R. Haferkamp. Environmental factors affecting plant productivity. In *Fort Keogh Research Symposium*, 1987.
2. N.A. Barber and R.J. Marquis. Leaf quality, predators, and stochastic processes in the assembly of a diverse herbivore community. *Ecology*, 92:699–708, 2011.
3. A. Buonocore, L. Caputo, E. Pirozzi, and A. Nobile. A non-autonomous stochastic predator–prey model. *Math Biosci Eng*, 11(2):167–188, 2014.
4. M.D. Asfaw. *Mathematical analysis of plant–herbivore interactions: with temporal and spatial variations in the Ethiopian climate*. PhD thesis, Department of Mathematics, Addis Ababa University, 2018.

5. G.J. Butler and H.I. Freedman. Periodic solutions of a predator–prey system with periodic coefficients. *Math. Biosci.*, 55:27–38, 1981.
6. D. Choquenot and J. Parkes. Setting thresholds for pest control: how does pest density affect resource viability? *Biol. Conserv.*, 99:29–46, 2001.
7. P.D. Coley. Possible effects of climate change on plant/herbivore interaction in moist tropical forests. *Climate Change*, 39:455–472, 1998.
8. M.B. Coughenour. Spatial component of plant–herbivore interactions in pastoral, ranching and native ungulate ecosystem. *J. Range Manage.*, 44:530–542, 1991.
9. C.K. Ho and S.C. Pennings. Preference and performance in plant–herbivore interactions across latitude—a study in us atlantic salt marshes. *PLoS one*, 8(3):e59829, 2013.
10. R.P. Bekker, P.V De Wit, K.S Fernando, and D.M. Radcliffe. Vegetation map of the republic of Botswana. Chief Land Utilization Officer, Land Utilization Division, Ministry of Agriculture, P/Bag 003, Gaborone, Botswana, 1991.
11. T. Zhidong. Uniform persistence of the periodic predator–prey lotka–volterra systems. *Appl. Anal.: Int. J.*, 72(3–4):339–352, 1999.
12. J.M. Cushing. Stable positive periodic solutions of the time-dependent logistic equation under possible hereditary influences. *J. math. Anal. Appl.*, 60:747–754, 1977.
13. J.M. Cushing. Periodic time-dependent predator–prey systems. *SIAM J. Appl. Math.*, 32(1):82–95, 1977.
14. T. Ding, H. Huang, and F. Zanolin. A priori bounds and periodic solutions for a class of planar systems with applications to lotka–volterra equations. *Discrete Continuous Dyn. Syst.*, 1:103–117, 1995.
15. J. Cui. Permanence of predator–prey system with periodic coefficients. *Math. Comput. Modell.*, 42:87–98, 2005.
16. M.D. Asfaw, S.M. Kassa, and E.M. Lungu. Co-existence thresholds in the dynamics of the plant–herbivore interaction with allee effect and harvest. *International Journal of Biomathematics*, 11(4), 2018.
17. M.D. Asfaw, S.M. Kassa, and E.M. Lungu. Threshold dynamics of the plant–herbivore interaction with allee effect. *BIUST Research and Innovation Symposium 2017 (RDAIS 2017)*, 1(1):66–71, 2017.
18. A. Ejigu and S. Asfaw, M. D. and Kassa. Estimating the spatial distribution of possible livestock production level using a mathematical model and the smap soil moisture data: The case of Botswana. *Communications in Mathematical Biology and Neuroscience*, 2022, 12 2022.
19. K. W. Morton and D. F. Mayers. *Numerical solution of partial differential equations: An introduction*. Cambridge University Press, 2nd edition, 2005.
20. A.C. Algar, H.M. Kharouba, E.R. Young, and J.T. Kerr. Predicting the future of species diversity: macroecological theory, climate change, and direct tests of alternative forecasting methods. *Ecography*, 32(1):22–33, 2009.
21. B. Helmuth, N. Mieszkowska, P. Moore, and S.J. Hawkins. Living on the edge of two changing worlds: forecasting the responses of rocky intertidal ecosystems to climate change. *Annual review of ecology, evolution, and systematics*, pages 373–404, 2006.
22. B.A. Yerlikaya, S. Ömezli, and N. Aydoğan. Climate change forecasting and modeling for the year of 2050. In *Environment, climate, plant and vegetation growth*, pages 109–122. Springer, Cham, 2020.
23. J. A. Santos, C. Yang, H. Fraga, A. C. Malheiro, J. Moutinho-Pereira, L.T. Dinis, C. Correia, M. Moriondo, M. Bindi, L. Leolini, C. Dibari, S. Costafreda-Aumedes, N. Bartoloni, T. Kartschall, C. Menz, D. Molitor, J. Junk, M. Beyer, and H.R. Schultz. Long-term adaptation of European viticulture to climate change: an overview from the h2020 clim4vitis action. *IVIES Technical Reviews, vine and wine*, 2021, 2021. Original language of the article: English.
24. G. P. Zhang and M. Qi. Time series forecasting using a hybrid Arima and neural network model. *Neurocomputing*, 64:175–185, 2005.
25. C. Bennett, R. A. Stewart, and J. Lu. Autoregressive with exogenous variables and neural network short-term load forecast models for residential low voltage distribution networks. *Energies*, 7(5):2938–2960, 2014.
26. M.D. Asfaw, S.M. Kassa, E.M. Lungu, and W. Bewket. Effects of temperature and rainfall in plant–herbivore interactions at different altitude. *Ecological Modelling*, 406:50–59, 2019.
27. Statistics Botswana. Annual agricultural survey report 2019–Statistics Botswana. <https://www.statsbots.org/bw/annual-agricultural-survey-report-2019>, 2020.
28. Food and Agriculture Organization (FAO) of the United Nations. Botswana: Country pasture/forage resource profiles, 2021. Retrieved on August 30, 2021.
29. R. J. Hyndman, Md. U. Akram, and B. Koo. Automatic time series forecasting: The forecast package for r. *Journal of Statistical Software*, 27(3):1–22, 2008.