Developing a Digital Twin Model for Improved Pasture Management at Sheep Farm to Mitigate the Impact of Climate Change

Ntebaleng Junia Lemphane, Ben Kotze, Rangith Baby Kuriakose
Department of Electrical, Electronic and Computer Engineering
Central University of Technology, Bloemfontein, Free State, South Africa

Abstract—Small-scale livestock farmers experience significant losses because of decreased productivity caused by decline in pasture production brought on by climate change. Technology in livestock farming introduced the idea of "smart farming," which has simplified pasture management. Internet of Things (IoT), Artificial Intelligence (AI) and data analytics are just a few of the cutting-edge techniques that smart farming incorporates. Digital twin technology is proposed in this study to alleviate the challenge of changing weather patterns that affect pasture management. Digital twin model is developed to predict pasture height to ascertain the predicted amount of pasture and ensure that the sheep have access to enough food for sustainable production. Pasture growth is influenced by temperature, rainfall and soil moisture; thus, pasture height predictions depend on these factors. Digital twin is made of predictive models built on historical and real-time data collected from the IoT sensors and stored in ThingSpeak® cloud. Data analysis was performed in MATLAB® using the neural network algorithm and predictions of the system are modelled in SIMULINK® platform. Digital twin predicted the pasture height to be 52 cm while the observed reading was 56 cm. Therefore, with the prediction error of -4, the digital twin can serve to enhance pasture management through its capabilities and assist farmers in decision making.

Keywords—Artificial intelligence; artificial neural network; climate change; digital twin; Internet of Things; machine learning; pasture management; smart farming

I. INTRODUCTION

Livestock farming makes a substantial contribution to food security. Forage species or pastures are used in livestock production systems to feed the animals, and this has become a crucial aspect of managing the pastures also known as pasture-based livestock production [1]. In South Africa, there are many animals that depend on pastures for survival. Statistics indicate that sheep farming is most practised in the country [2] and the major feed source is utilization of pastures [3]. Hence, the focus of the study is on sheep farming.

Pasture is a grazing area for all ruminants, it also offers sheep with a nutritious diet [4]. To ensure profitable livestock production, pasture management is a crucial component of farm management. Pasture management involves maintaining healthy pasture and its companion plants to provide the animals with sustainable feed. A well-managed pasture has considerable advantages, such as increased forage yields and lower feed costs [4], which result in healthy sheep. For efficient grazing management, farmers who rely on grazing pasture as their main source of animal feed require accurate and timely observations of pasture height [5]. Since productivity is influenced by the extent of pasture utilization, which is a function of enhanced pasture growth, accurate measuring of pasture is crucial [6].

However, a scarcity of high-quality pasture is the main obstacle facing sheep farmers [7]. Poor management techniques, uncontrolled grazing systems, and a loss in pasture yield due climate variability can all contribute to this issue [8]. Nonetheless, smart farming tools are new approaches that are emerging to improve pasture-based systems and farming conditions which support farmer’s decision making and increase productivity [9, 10]. Smart farming is a technology that depends on the application of AI and IoT in the management of cyber-physical farms [11]. Despite the smartness provided by these technological advancements there are also negative environmental impacts that cannot be ignored. The use of modern technologies poses challenges to our environment and pastures are no exception to these problems. All these modern devices and machines raise concerns about waste, use of non-renewable materials and carbon footprint which contributes to climate change [12, 13]. However among these challenges, climate change is the main focus of this study as it affects the productivity of pasture-based systems [14].

In this study digital twin technology is the proposed solution to limit the impact of climate change in pasture management. A digital twin is a virtual version of a physical asset that is made possible by data and simulators to facilitate better decision-making, monitoring, controlling, and real-time prediction, optimization, and monitoring[15]. Although digital twin technology in livestock farming is still in its infancy, it has taken the advantage to use the current smart farming technology to improve farm management, animals welfare and production of animals products [16]. Digital twin technology promises to help farmers with better predictive models by combining big data, real-time data from the individual farm, and AI models trained by machine learning algorithms [16].

The aim of this study is to investigate how a digital twin can predict pasture height and introduce soil moisture predictive model to form part of forecast models. Regression and artificial neural networks (ANN) machine learning algorithms were investigated to determine their performance.
for prediction. The digital shadow was developed to set the grounds and determine if the predictive models could be trusted on the digital twin. The digital twin is comprised of predictive models built with historical data and updated by real-time data from the IoT sensors set-up on the farm. The predictive models predicted temperature, rainfall, soil moisture and pasture height. The major goal of the study is to develop a better pasture management system with early detection of predicted scenarios for better risk management and decision-making processes.

II. OVERVIEW OF THE LITERATURE

In all of South Africa's provinces, sheep farming has a significant impact on socioeconomic and cultural life [17]. Small-scale businesses that practice sheep farming produce meat and wool as a source of revenue. Forage found on pastures is what sheep eat [3], hence pasture management techniques are used to create high-quality pasture to feed the sheep. However, pasture management is impacted by climate change, which leads to a decrease in the amount and quality of pasture production [18]. Farmers struggle to keep their animals alive as a result of the shortage of adequate pasture for the sheep to graze on [19]. According to survey data from commercial farmers, most farmers withdrew from sheep farming altogether over a three-year period, between 2012 and 2014, because the feeding conditions were too challenging owing to the drought [20]. This indicates the depth of climate change’s influence.

Climate change has an impact on pasture management and could make it even more difficult to manage grazing areas [21]. Pasture growth is directly impacted by the interaction of temperature, precipitation, and soil moisture. Production of pasture is negatively impacted by temperature increases, heavy rainfall, and high evaporation rates brought on by climate change [21] [22]. There are identified smart farming technologies that are already applied to overcome issues related to pasture management and utilization, animal monitoring and control at sheep farms [9]. Smart farming offers farmers with superior decision-making and management strategies [23] by merging information and communication technologies through AI and IoT sensors for use in agricultural and livestock production systems.

The Internet of Things (IoT) is the technology that allows sensors to be linked together and operated automatically using internet which includes various sensing methods for collecting, processing, analyzing and storing of real-time data [11]. Artificial intelligence (AI) is the science of making intelligent machines and programs by developing software’s and systems using machine learning and deep learning to solve problems and make decisions [24]. Just like the human brain, these software programs are provided with training data, and further, these intelligent devices provide the required result for every legitimate input [25].

Machine learning and deep learning are the fundamental building blocks of AI [24]. Machine learning is the capacity to learn something based on the data set without explicit programming [26] while deep learning is subcategory of machine learning which includes learning of neural networks made of neurons having various parameters and layers between the input and output [27]. Thus, both IoT and AI along with cloud-based technology play a critical role in farm management by collecting and analysing sensor’s data to perform temperature and rainfall predictions, monitor crop growth and soil management [28]. Smart farming has allowed new technologies to be implemented in pasture-based systems to improve efficiency. These pasture management systems include weather stations, capturing pasture measurements and soil conditions [5] [9] [10] [29]. However, using these technologies has their limitations.

Soil moisture as one of the key factors influencing pasture growth is not considered by forecast models [30]. Furthermore, the current systems are unable to accurately identify and monitor pasture growth [31]. Due to the impossibility to gain a comprehensive view of physical systems in real-time, this necessitates time-consuming remote monitoring and control [15]. Sensing technologies allow for real-time farm monitoring, which presents the chance to create farm-specific models that a specific farmer might use to schedule activities in response to changing conditions.

Literature reveals, however, that digital twin technology holds the promise of enhancing smart farming for better farm management and higher productivity[16]. A digital twin is a representation of a physical thing that replicates its behavior and states across time in a virtual environment [32]. Utilizing both current and past data, a digital twin provides analysis, forecasting, and operation optimization [33]. Large volumes of data must be able to be received, stored, and processed by the digital twin in real-time hence, significant computing, storage, and data processing resources are needed for this.

Using real-time data from numerous IoT sensors and devices, digital twins continuously adjust to operational changes to forecast the future of the physical system with the help of AI and machine learning [32]. Digital twins are able to predict systems’ future behaviours using models that incorporate machine learning [34], known as predictive modelling. In order to predict future outcomes, predictive models employ algorithms that learn and analyse both historical and present data. Algorithms for machine learning, deep learning and pattern recognition can be used to construct predictive analytics, which can help understand how operations are changing over time. Regression and Neural models are the most used predictive modelling methods [35].

Digital twins have the ability to address smart farming challenges. This includes the performing of future predictions [36] using real-time monitoring systems, control and analysis. Digital twins make use of models that accurately depict an object’s behavior throughout time [37]. To project unforeseen events, the models can also be tested with what-if scenarios [16]. Additionally, digital twins allow for the merging of models by setting up a shared model space that defines the correlation between models, enables data flow between models, and establishes a connection between the digital twin and the physical asset [38]. Moreover, creating connected physical objects and a digital twin, the digital twins can address the problem of seamless integration between IoT and data analytics [39].
Although digital twin technology is still in its early stages, literature shows that digital twins have potential to shift the animal husbandry’s future by integrating smart farming technology which use real-time data manipulated by AI analyses, which can then fuel better business decisions, improve animal health and well-being, and maximize the return from farming resources.

III. MATERIALS AND METHODS

Sheep farmers are having a challenge of providing healthy pasture for their animals due to climate change. Pasture growth depends on climatic conditions therefore the focus of this study is to develop a digital twin that predicts pasture growth based on predicted temperature, rainfall and soil moisture. The research methodology is implemented into three phases:

- Phase A – Developing predictive algorithms through data analysis, structure, and processing, and use the selected algorithm to develop a digital shadow.
- Phase B – Monitoring farm conditions on the farm to manage variables that affect pasture’s growth.
- Phase C – Creating a digital twin with predictive models that are updated by real-time data to forecast pasture height and identify problems in advance, thus assisting a farmer to come up with quick solutions.

A. Phase A – Developing Predictive Algorithms

Predictive modeling is essential to the study as it forecasts future results. Due to unpredictable weather patterns brought on by climate change, farming has become more challenging. Predictive models are helpful in this situation. To make predictions, predictive models are built using previous data and new data sets. Predictive models were created in this study using regression and ANN machine learning algorithms.

1) Selecting suitable algorithm. For predicting climate, regression and ANN models are frequently employed [40], hence these were the only models investigated in the study. Developing a prediction model to track climate change requires a development of dataset which contains historical data of weather information. Climate change is monitored based on decades of the earth’s atmospheric observation. Hence, the ten-year historical data of temperature and rainfall was used to develop the prediction models.

The study uses historical data from 2011 to 2020 gathered from [41] source, at the area where the farm is located. The same set of data was used on both algorithms and analyzed in MATLAB® to get insight into the data collected. After learning and being trained on data, both algorithms were used to create prediction models, and the results were compared to see which algorithm could be a better fit. Two predictive models on each algorithm were developed to predict average temperatures and rainfall for the year 2021.

Root Mean square Error (RMSE) is a measure of accuracy to compare forecasting errors of different models [42] and is commonly used as an error metric for numeral predictions. The lower the RMSE, the higher the accuracy of the model. The model performance is also determined by comparison of true and predicted response. Fig. 1 demonstrates the RMSE values obtained while training temperature predictive models using regression and ANN algorithm with the same data set.

After training the models, the predictions were performed for the first nine weeks of the year 2021. Nine-week duration was selected because pasture is ready for grazing at six to nine weeks after sowing. Therefore, having models that could precisely predict climatic conditions up to forage life cycle would be ideal.

2) Digital shadow development. The selected algorithm was then used to develop a digital shadow. Developing a digital shadow is the first stage in creating a digital twin [43]. Digital shadow is used to visualize operating, status, or process data that is gathered while the product is in use or during an ongoing process. Then, the digital shadow replicates the digital model, which is made up of all the data from the design and production phases and serves as an intelligent link to the digital twin [44]. The digital shadow development sets the foundation for a digital twin. A digital shadow was made of predictive models that predicted temperature, soil moisture, rainfall and pasture height.

The digital shadow was developed using historical data and collected data from the farm. The pasture is planted twice a year – in Autumn (March to May) and Spring (September – November). Thus, digital shadow was trained with data of both seasons of the year 2021 and the predictions were performed for Autumn season in 2022. The objective was to ascertain how temperature, soil moisture, and rainfall influence pasture growth. The readings were recorded weekly and prepared in spread sheets. Fig. 2 shows sample data collected for the month of April.

The digital shadow was composed of four different prediction models. Temperature and rainfall models were developed using historical data and were updated by data collect in the farm. Soil moisture and pasture height were developed by farm data. Fig. 3 shows how the models were structured. The soil moisture model is dependent on temperature and rainfall models and pasture height model is dependent on other three models.
Fig. 2. Farm data collected for the month of April.

Digital shadow sets the foundation of the digital twin. The results of the prediction models of the digital shadow determine whether the models could be trusted when implemented in the digital twin.

B. Phase B – Monitoring Farm Conditions

For a digital twin to exist there must be a physical system. The "smart farm" served as the study’s representation of the physical system. IoT sensors were identified to monitor pasture growth and were integrated to create a smart farm as a link between the physical system and the digital counterpart. Sensors are also crucial to the development of digital twins as they gather information in real time that is utilized to create and update the prediction models. In the process of limiting factors that cause climate change, the selection of IoT sensors was based on devices which do not emit carbon footprint and waste. Cloud platform, IoT sensors and a gateway are the building blocks of the physical system.

1) Cloud platform: IoT sensors gather data and transmit it to cloud storage. ThingSpeak® was selected for use in this study because of its capabilities to transmit sensor data in real-time to the cloud [45]. The ability to pre-process and analyze the data using MATLAB® is another advantage of using ThingSpeak® [46].

2) IoT sensors: Sensors that monitor pasture growth were set up on the farm as part of the physical system. These include sensors that measure temperature, rainfall, soil moisture and pasture height. The sensors used are as follows:

- A weather station - measures temperatures and rainfall and to log the results in ThingSpeak®.
- The ESP32 CAM - captures the images of the pasture to be interpreted in MATLAB® to calculate pasture height.
- Raspberry Pi – runs the image processing script to calculate the height of the pasture based on the captured images and the result was logged into ThingSpeak®.

3) IoT gateway: The connection between IoT sensors and the cloud must be made through an IoT gateway. IoT gateway serves as a network router that directs data between IoT sensors and the cloud which enable internet connectivity. 4G LTE Wi-Fi router was selected as a gateway.

The complete physical system was the integration of IoT sensors, gateway and ThingSpeak®. The design of the physical system is shown in Fig. 4.

C. Phase C – Creating a Digital Twin Model

To test the chosen algorithm on the physical system and determine if it can be trusted when deployed on the physical system, creating a digital shadow was the first stage in creating a digital twin. The design of the digital shadow clarified the scenarios of the physical system. Set of digital shadow models were helpful to understand the structure and the behaviour of the system in the physical world as these models were used to create a digital twin to perform predictions for the future.

The design of the digital twin facilitates the activities, monitoring and digital control of operations in all the connections of the system [47]. The digital twin development is comprised of two platforms [47] - a physical system and the digital platform. Physical system is basically the setup of IoT sensors that collect data and store the result on a cloud platform. Running systems in real-time distinguishes the digital twin from the digital shadow [48]. The digital twin is developed in three phases, namely [49]:

- data collection and monitoring;
- data storage,
• data analytics and predictive modelling.

IoT sensors allow real-time monitoring and store data in ThingSpeak® which enable the link between the physical system and digital platform. The real-time data is used to update the prediction models in MATLAB® and the updated models are used to perform the predictions in SIMULINK® platform. The process in demonstrated in Fig. 5.

![Digital twin development process](image)

Fig. 5. Digital twin development process.

The digital twin is made up of four predictive models which are updated by sensor data which is retrieved in real-time. The models are built and trained in MATLAB® platform. The link between MATLAB® and SIMULINK® allow smooth transfer of data to update models and enable predictions. Fig. 6 summarizes the process of analysing sensor data and using the data to keep prediction models up to date.

![Data analysis and predictive modelling in the digital twin](image)

Fig. 6. Data analysis and predictive modelling in the digital twin.

The digital twin model is developed with models from the digital shadow, the difference is that digital twin models are kept updated with real time data from IoT sensors. Fig. 6 shows that sensors gather live data and store it in ThingSpeak® cloud. This data is then analyzed to get weekly statistics because pasture height is measured weekly. SUMULINK® platform makes it easy to transfer data from ThingSpeak® and update the predictive models. The digital twin model portrays the dependencies between models.

To perform the predictions, the digital twin accepts the duration of the prediction as the input (year and the week). The temperature, rainfall and soil moisture prediction models will then perform predictions for the specified duration. Pasture height prediction model will make predictions based on the output of the three models (temperature, rainfall and soil moisture). Fig. 7 demonstrates how digital twin is modelled in SIMULINK®.

![Digital twin modelled in SIMULINK®](image)

Fig. 7. Digital twin modelled in SIMULINK®.

IV. RESULTS AND DISCUSSIONS

This section shows the results obtained in the process of developing the digital twin to predict the pasture height, starting from algorithm section, digital shadow development until the digital twin development. The analysis of the results will also be discussed.

A. Comparison of the Algorithms

Both regression and ANN algorithms were investigated to determine which one performs predictions better. RMSE was used as a measure of accuracy to compare prediction errors on temperature and rainfall predictive models. After obtaining the predicted values from the prediction models, these values were compared with the observed values to evaluate the performance of the model.

It is important to perform model evaluation because it helps to assess the efficiency of the model during the initial research phases. The model evaluation was therefore conducted using two evaluation metrics, namely: correlation coefficient and estimated error. Correlation coefficient measures the relationship between two variables [50]. The values range is between -1 and 1. The closer the calculated value moves to 1, the stronger the relationship between two values and vice versa. Prediction error is the difference between the observed value and the predicted value [51]. The smaller the difference, the better. The results are shown in Table I.
Table I shows the results obtained while developing temperature and rainfall prediction models using both prediction algorithms. The outcome demonstrates ANN algorithm performing better with lower RMSE and lower estimation error. It further proves itself with a higher correlation value which symbolizes a strong relationship between the predicted and observed values. Therefore, ANN algorithm was selected as an algorithm to develop prediction models on the study.

### B. Digital Shadow Development

Developing a digital shadow is an important stage towards developing a digital twin. The effectiveness of the digital twin to be created is determined by the results from the digital shadow, as the digital twin will employ the same algorithm as the digital shadow. Thus, ANN algorithm was used to develop a digital shadow. The digital shadow was made up of four prediction models – temperature, rainfall, soil moisture, and pasture height prediction models which were developed using gathered farm data. The digital shadow model predicted the expected pasture height based on the predicted temperature, rainfall and soil moisture. Table II demonstrates the results obtained.

### C. Digital Twin Development

Digital twin is made up of a physical system and a digital platform. The digital twin development was composed of three phases which are: data collection, data storage, and predictive modelling.

1) **Data collection.** The physical system was made up of IoT sensors which were set up on the farm to gather data. The structure of the physical system in shown on Fig. 8.

![Fig. 8. Structure of the physical system on the farm.](image)

Table II demonstrates the results obtained in the predictions of the digital shadow. Pasture height prediction model depends on the temperature, rainfall and soil moisture models to perform predictions. The ANN algorithm proved to perform better than regression for predictions, however there are uncertainties related ANN prediction models. The problem with ANNs is that is no clear understanding on how they analyze patterns of data to give the output on the predictions; they give final result [52]. This is proved with the outcome of the digital shadow model. Temperature and rainfall prediction model are expected to perform better than the other models as they were trained with more data than others. Nonetheless, soil moisture prediction model still had a lower prediction error than the rainfall prediction model. Hence this could be justified about changing rainfall patterns which makes the training of the models difficult to analyze data.

The other problem is that ANNs need a lot of data to train the models for them to work efficiently. Pasture height prediction model in the digital shadow anticipated the final pasture height of 59cm. The observed pasture height was 72 cm, and the highest prediction error was 13 cm. The reason of a higher prediction error in pasture height prediction model is that, the model was only trained with two seasons worth data. Thus more data was needed to improve its efficiency, including the soil moisture prediction model. Even though the model predicted a lower pasture height but the overall data seems promising. Therefore these models could be trusted in the digital twin, more data is to improve the models efficiency.

### TABLE II. Digital Shadow Prediction Results

<table>
<thead>
<tr>
<th>Prediction model</th>
<th>Prediction error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temperature</td>
<td>±1.67 °C</td>
</tr>
<tr>
<td>Rainfall</td>
<td>±5.9mm</td>
</tr>
<tr>
<td>Soil moisture</td>
<td>≤4si</td>
</tr>
<tr>
<td>Pasture height</td>
<td>13cm (max)</td>
</tr>
</tbody>
</table>

### TABLE I. Comparison on Regression and ANN Algorithms While Developing Temperature and Rainfall Prediction Models

<table>
<thead>
<tr>
<th>Metric</th>
<th>Temperature model</th>
<th>Rainfall model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Regression</td>
<td>ANN</td>
</tr>
<tr>
<td>RSME</td>
<td>1.98</td>
<td>5.557</td>
</tr>
<tr>
<td>Correlation</td>
<td>0.645</td>
<td>0.945</td>
</tr>
<tr>
<td>Estimation Error</td>
<td>±5.03</td>
<td>±22.21</td>
</tr>
</tbody>
</table>

www.ijacsa.thesai.org
3) **Predictive modelling.** The digital twin was composed of four prediction models from the digital shadow which are now updated by sensor data. Pasture height prediction model depends on three parameters which are temperature, rainfall and soil moisture. Thus, the pasture height prediction model was provided with new data of predicted average temperatures, predicted total rainfall, and predicted average soil moisture to perform the predictions of the next season. The predictions were done for the year 2022 for Spring season. The prediction started from the 38th week and looped through until the 46th week, since the pasture growth takes nine weeks. The prediction results are shown in Fig. 10.

On the presented outcomes, temperature and soil moisture prediction models showed good performance with a prediction error of ±1.62 and -1.7, respectively. Rainfall prediction model had a prediction error of ± 6.03, which was slightly higher as compared to other prediction models. The results on the rainfall prediction model signify the impact of changing weather patterns. However, the results of this model demonstrated a good outcome as predicted values and observed results showed a close correlation.

Based on the predicted temperature, rainfall, and soil moisture, the prediction of the pasture height was conducted for the duration of pasture growth. The aim was to predict the expected pasture height which was tracked from the plantation date. The analysis is demonstrated on Table III.

Table III shows that the prediction error of the model ranging between 0 and -4. Even though the observed pasture height was higher than the predicted outcome, the predicted pasture height represents a good outcome from the model, meaning there will be enough feed for the sheep. Thus the digital twin successfully predicted pasture height by integrating temperature, rainfall and soil moisture prediction models.

<table>
<thead>
<tr>
<th>Week</th>
<th>Predicted pasture height (cm)</th>
<th>Observed pasture height (cm)</th>
<th>Prediction error</th>
</tr>
</thead>
<tbody>
<tr>
<td>38</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>39</td>
<td>0</td>
<td>2</td>
<td>-2</td>
</tr>
<tr>
<td>40</td>
<td>12</td>
<td>15</td>
<td>-3</td>
</tr>
<tr>
<td>41</td>
<td>22</td>
<td>24</td>
<td>-2</td>
</tr>
<tr>
<td>42</td>
<td>29</td>
<td>32</td>
<td>-3</td>
</tr>
<tr>
<td>43</td>
<td>42</td>
<td>45</td>
<td>-3</td>
</tr>
<tr>
<td>44</td>
<td>48</td>
<td>52</td>
<td>-4</td>
</tr>
<tr>
<td>45</td>
<td>51</td>
<td>54</td>
<td>-3</td>
</tr>
<tr>
<td>46</td>
<td>52</td>
<td>56</td>
<td>-4</td>
</tr>
</tbody>
</table>

V. **Conclusion**

The aim of the digital twin development was to predict pasture height for the future to determine if there will be enough feed for the sheep. The digital twin model was proposed after determining that sheep farmers struggle to keep their animals alive due to improper pasture management caused by changing farming seasons due to climate change. Soil moisture predictive model was successfully implemented and integrated with temperature and rainfall predictive models to acquire anticipated pasture height. ANN machine learning was helpful in developing predictive models for forecasting. The study also shows how AI and IoT technologies are collaborated to develop real-time systems with predictive models. Thus, the results show a digital twin made of real-time monitoring of the pasture growth with predictive models can assist the farmer in taking proper decision on time thus improving management strategies.

**REFERENCES**


