Attention-Guided Bidirectional Temporal Modelling with Graph-Based Regional Spatial Context for Bajra Crop Yield Prediction

Mamta Kumari[1]*, Suman[2], Devendra Prasad[3]

Deenbandhu Chhotu Ram University of Science and Technology, Sonepat, India[1,2]

PIET Samalkha, India[3]

Abstract—**Bajra (pearl millet) is a very important crop in Rajasthan, India, since it is drought-resistant, nutritious, and culturally important. But its productivity is becoming vulnerable to changes in climate, such as erratic rain and temperature changes, and thus precise estimation of yield is vital. Crop Yield Prediction (CYP) indicators like soil decomposition, rainfall and meteorological patterns are slowly evolving, exhibiting long-term temporal dependency and propagating over time. Conventional cropping prediction algorithms based on artificial intelligence process the historical data and these indicators in a unidirectional manner. While mapping the temporal dependencies, these algorithms consider each year independently and do not capture the delayed effect, like salt degradation. To address this issue, the study proposes a region-based spatiotemporal model with an attention-guided Bidirectional LSTM (Long-Short Term Memory) framework for CYP, termed as G-BiLSTM. The proposed model reproduces the spatial relationships between districts via GCN (Graph Convolution Network) -based immediate neighbour extraction. Further, a Bidirectional LSTM is used to model multi-year CYP temporal features, allowing each annual observation to be encoded using both past and future temporal context. A variance-reduced and comprehensible representation is produced by integrating an attention mechanism to adaptively highlight the most informative years within a temporal window. Using 15 agro-environmental characteristics, including understudied elements like saline and alkaline soil composition, the framework is assessed on a dataset that includes 32 districts in Rajasthan over 13 years (2007–2019). The suggested attention-enhanced BiLSTM consistently outperforms traditional temporal models, achieving lower prediction error and better generalisation, according to experimental results analysis using a three-year sliding temporal window. For regional crop yield forecasting, the suggested method offers a scalable solution.**

Keywords—Bajra crop yield prediction; regional crop yield forecasting; attention-guided Bidirectional LSTM; saline and alkaline soil composition

I. Introduction

It is anticipated that by the year 2100, the global population will hit 9.6 billion. The need for farm goods increases [1]. By the year 2100, India's population is projected to increase from 1.4 billion to 1.7 billion. Agriculture covers 1.6 million square kilometres (about half the area of India), and the agricultural land ranks second behind China [1]. The government must always ensure food to protect its inhabitants. Modern agriculture relies on crop yield prediction for better resource allocation, to decrease risks, and to ensure food security. However, it has become challenging with the complex aspects affecting Crop Yield Prediction (CYP). To enhance agricultural and rural statistics, researchers in the field are working on crop yield prediction methodologies and investigating ways to improve agricultural data collection for yield prediction. Considering the recent survey, yield is mainly dependent on data collected based on factors like climatic conditions, soil quality, and water quality [2-4]. However, the study in papers mainly differs based on regions, crop types, weather conditions, and irrigation techniques. Thus, there is no standard dataset that is fully accessible for agricultural study in a particular region [5]. Advancements in technology and changes in climate significantly influence agricultural production levels. The impact of weather on crop yield can be explored by considering time (measured in years) or time-related functions [6]. Various elements are expected to gradually increase yield over time, yet weather is recognised for its unpredictable impact. Therefore, to forecast agricultural outputs across different years or within the same year, a majority of research utilizes time-series algorithms that leverage past data and are regarded as more suitable than various machine learning and deep learning methods [7-9]. This study centres on the CYP of Bajra (pearl millet), which is widely grown in India, particularly in the states of Rajasthan, Maharashtra, Haryana, Uttar Pradesh, and Gujarat. The crop is capable of enduring conditions such as intense heat, lack of moisture, low soil fertility, and minimal rainfall. Bajra thrives in areas where other cereal crops struggle due to its specific growing needs. This study examines the precise CYP related to the Bajra crop in Rajasthan.

A. Related Work

There is an increasing body of literature that analyses machine learning and deep learning in predicting crop yield in different regions and crops. Among the studies, [7] optimises crop yield prediction in Rajasthan state for 10 districts and 7 crops using random forest (RF) and decision trees with gradient boosting. From 1997 to 2018, the research analysed acreage, productivity, yield, and rainfall. Gradient-boosting outperformed all other methods with 92.7% accuracy, 21.58 MAE, and 15.01 RMSE. Only a few of Rajasthan's traits and districts were studied. Computer-intensive gradient boosting is subject to noise. Overfitting can occur with too many decision trees. Therefore, many papers avoid these methods. The study [8] predicts Jaipur's agricultural productivity for five crops. The analysis focuses on the data spanning from 1991 to 2020. The variables that were measured included rainfall,

hours of sunshine, temperature, and humidity levels. The writers employed Random Forest, Support Vector Machines, and Artificial Neural Networks.

The authors in [8] discovered that the Random Forest model achieved an accuracy rate of 92.3%, along with MAE and RMSE figures of 1.68 and 2.19, respectively. The study only covered one Rajasthan region and utilized a few criteria. The study [9] employs random forest, SVM, gradient descent, LSTM, and Lasso regression algorithms. The dataset comprises five crops for 33 districts in Rajasthan. All algorithms were examined, and the findings revealed that RF, SVM, and Lasso regression best-predicted yield. In the study [9], the limitations include the possibility of using a richer and larger dataset for each crop to enhance forecasts. An Improved Optimization Function (IOF) is proposed in [6] and implemented with the LSTM model to enhance prediction accuracy. Temperature, precipitation, humidity, solar radiance and various crop related data attributes were used in the study. The article [10] predicts mustard production in 5 Rajasthan districts using long-term weather and yield data. This study uses random forest, SVM, and ANN deep learning. Feature selection in [5] utilizes Principal Component Analysis (PCA) and Stepwise Multiple Linear Regression (SMLR). Assessment was conducted on agricultural area, soil type, temperature, rainfall, and climate. The research assesses six different model combinations and concludes that PCA-SVM stands out as the most effective yield prediction model. A hybrid model driven by data was created by the authors in [11] to accurately predict paddy yield through the use of Multiple Linear Regression (MLR) intercepts. The Artificial Neural Network (ANN) utilizing Feed forward Back Propagation is trained with characteristics like the area, count of open wells, tank volumes, canal lengths, and peak temperatures throughout the season examined in the study.

The creators of [12] devised a deep learning approach using a Recurrent Neural Network to enhance the Q-Learning reinforcement learning technique. The study predicts paddy yield in Vellore with an accuracy of 93.7% by utilizing properties of soil, weather, and groundwater, all while maintaining the original distribution of the data. The study in [13] forecasted the yield of sugarcane in Karnataka by utilizing Long Term-Time-Series (LTTS), weather and soil characteristics, Normalized Vegetation Index (NDVI), along with supervised machine learning (SML) techniques. Support vector regression (SVR) was employed by researchers. The scientist's analysed data related to climate and soil, along with satellite images from the years 2008 to 2018, to assess remote sensing information. A study based on empirical data focused on forecasting agricultural production by employing regression and time-series models to predict the crop known as 'Bajra' or pearl millet [14]. Models for prediction comprised ARIMA and ARIMAX, incorporating an external variable. When evaluating 'Bajra', the ARIMAX model outperformed the regression time series model. The study [13] proposed a regression model utilizing ANN and SVM to forecast Kharif crop yields based on rainfall information from Visakhapatnam. This research focuses on four types of crops: Maize, Ragi, Rice, and Bajra. This research employs a modular artificial neural network to predict monsoon precipitation, subsequently utilizing support vector regression to assess the primary Kharif crop yields in relation to rainfall and cultivated area. Ensembles of CNN and

DNN machine learning predict corn production in the US corn belt, utilizing corn yield data from 1980 to 2019 [15]. Data on meteorology, soil, and management at the county level for 12 states in the US Corn Belt is provided. Constructed using deep neural networks and CNNs, ensembles are formed. This model accurately predicted maize grain yields for 2019, achieving an 8.5% relative root mean square error. The study in [16] introduces a framework utilizing deep learning for predicting agricultural output, employing CNNs and RNNs that rely on environmental data and management strategies. The CNN-RNN framework, in conjunction with RF, DFNN, and LASSO, predicts corn and soybean yields in the US Corn Belt utilizing data from 2016 to 2018. The new model surpassed all prior methods with an RMSE of 9% and 8% of usual yields. However, computationally expensive ANNs require careful design and hyper parameter adjustments. Too many trees can over fit and make random forests computationally expensive, and SVR is susceptible to outliers and requires careful kernel function selection [17]. With the rise of deep learning, advanced deep network topologies like CNNs [18] and RNNs [19-20] have been developed. Neural network-based agricultural production forecast models have increased due to deep learning. Out of the surveyed literature, 69% of papers employed deep learning approaches, with just 20% using a CNN-based model. There have been very few or negligible studies that use the graph-based model to harness the spatial features from the crop data for CYP. Like in [21], the Graph Convolution Network (GCN) is used for learning the weed plant features from labeled and unlabeled images, which are already extracted from CNN. The authors also work towards using limited labeled data and learning from graph structure to improve recognition accuracy. However, this study does not directly relate to predicting the yield.

The overall survey for dataset attributes and characteristics used for CYP in the studies is majorly divided into meteorological [2], soil [3], phenotype [22], fertilizer, diseases, and pests [4] related attributes, as well as other attributes, as shown in Table I. All of these characteristics are significant for determining crop output. As a result, in the present manuscript, a summary is created of the data attributes utilized in earlier work to determine which attributes are least studied and which ones need to be worked on. It is observed from Table I, that phenotypic, irrigation accessibility, fertilization, and undesirable features (diseases, weeds, and pests) are the least investigated qualities (3%, 3%,3%, and 14%). However, all these characteristics are investigated within the Precise Agriculture application, which makes use of IoT [3]. But, because the work focuses on data specifically for yield prediction, the study will select the other attributes of the yield prediction process. The other attribute that is highly used for CYP is soil characteristics. The sub-categories in Table I, demonstrate that groundwater levels and soil moisture characteristics are also understudied, but the current study included rainfall because it directly affects yield [23]. Red, black, desert, clay, or a mix of soil types has been understudied, and data is collected in only a few studies; therefore, soil type is considered in the present study. The kind of soil determines which plants thrive and which struggle. Cotton, for example, thrives in dark soil [24]. Also, infertile or barren soil, which includes excessive salt levels, is not considered. This study also assessed saline and alkaline soil in numerous state districts. The research

TABLE I. SUMMARY OF ATTRIBUTES FROM PREVIOUS STUDIES

Attributes Category (%)	Attributes	Attributes Category (%)	Attributes
Meteorological (40%)	Rainfall Temperature Humidity Sunshine duration Evapotranspiration Dew points temp	Phenotype (3%)	Lodging rate Inversion Rate Plant Height Ear Height Empty Stalk Rate
	Wind speed	Undesired Traits (14%)	Weed Determination Pest
Fertilizer (7%)	Fertilizer Strength		Disease
			Groundwater
Area (3%)	Area under Cultivation		Soil Water
		Soil (23%)	Soil Moisture
Irrigation Accessibilities (3%)	No. of Tanks Canal Length Open Wells		Soil Temperatures Nutrients Soil pH

also evaluated the local climate. The local climate impacts agricultural growth. Some crops thrive in arid climates, while others favour humidity [25].

Numerous works are conducted in the field of temporal modelling; the model in [26] uses an enhanced Gate Recurrent Unit through an LSTM model for the prediction of CYP. However, spatial connections are considered static by implementing a Grey Wolf Optimisation (GWO). Similar to the proposed model, the authors in [27] have proposed an attention mechanism for a CNN that is a convolutional neural network-based BiLSTM model for CYP. However, different from the proposed model, the spatial connections are considered independent through a convolutional neural network. The proposed model focuses towards a unified mapping of the spatial connection among districts affecting yield parameters through the GCN model and maps the enhanced bidirectional temporal dependencies. An enhanced deep learning model is presented in [28] through a Bayesian optimised CNN-BiLSTM model similar to the model in [26]. However, the parameters are considered independent, and the delayed effect of rainfall is not considered in [27-28]. Similar to the proposed model in the study and in [26], the authors in [29] applied channel attention to BiLSTM for CYP on image-based data. Moreover, unlike existing temporal models, the proposed model treats the spatial representations dependant as a graph and joint modelling of these with long-term bidirectional temporal dependencies, contributes to regional-aware temporal modelling.

Table I shows that phenotype characteristics, irrigation accessibility, fertilisation, and undesirable features (such as diseases, weeds, and pests) are the least studied attributes, with percentages of 3%, 3%,7%, and 14%, respectively. These attributes are integral to Precision Agriculture applications leveraging IoT [3]. However, focusing specifically on yield prediction, this study will emphasise other relevant attributes in the prediction process.

Soil characteristics are another key attribute widely utilised in CYP. Table I also indicates that groundwater levels and soil moisture characteristics are areas with limited research. However, the current study incorporates rainfall due to its direct impact on crop yield [23]. Soil types such as red, black, desert, clay, or a mixed soil have received less attention in data collection across studies, making soil type a focal point in this study. Soil type plays a critical role in determining which crops flourish and which struggle; for instance, cotton thrives in dark soil [24]. Furthermore, infertile or barren soils, including those with excessive salt levels, are not considered. This study also examines saline and alkaline soils across various state districts. Additionally, local climate conditions are evaluated, as they significantly influence agricultural productivity. Some crops thrive in arid climates, while others prefer humid environments [25]. These deep learning models have become increasingly popular in agricultural production forecasting. According to recent surveys, 69% of studies have adopted deep learning techniques, with CNN-based models being utilised in only 20% of cases.

1) Challenges in the existing studies need to be addressed: Existing studies in CYP focus on using data (images, new features, and focused data for yield prediction), spatial and temporal modelling of the data. However, the existing studies exhibit challenges that need to be addressed. These are discussed through a Table II, mapping the challenges on how the contribution of the study is connected.

B. Contributions of the Paper

As discussed in Table I, the study differs and addresses the challenges identified in existing studies. However, the main contributions of the study are:

- The suggested model performs CYP for Bajra (pearl millet) over 13 years (2007–2019) in 32 districts in the Indian state of Rajasthan. By adding a wide range of agro-environmental variables, such as weather patterns and underutilised soil characteristics, the study considerably broadens the scope of data collection, building on previous research.

- The proposed model suggests a three-phase framework for spatiotemporal learning that includes: 1) Data Collection, preprocessing and capturing Annual Regional-immediate neighbour information; 2) Joint Spatio-temporal feature extraction through attention-guided G-BiLSTM; and 3) yield prediction.

- The suggested framework uses an attention-guided Bidirectional LSTM (BiLSTM), in contrast to traditional temporal models, to encode each year from CYP data representation using both past accumulation and future-aware agro-environmental context, and the attention mechanism selectively highlights the most significant temporal points, resulting in a spatiotemporal embedding that is interpretable and reduces variance. The proposed framework minimises variance in the predictions across various regions or temporal windows by down-weighting uninformative years, which prevents the model from reacting excessively to inflated shifts in the training data (such as an anomalous rainfall year).

- Evaluation through RMSE, R2 score, correlation coefficient, and other common performance metrics is conducted to assess the suggested Attention-guided G-BiLSTM framework.

Study Organisation: the introduction, related work and contributions are explained in Section I, the detailed proposed

TABLE II. CONNECTING THE CHALLENGES WITH THE CONTRIBUTIONS OF THE PAPER

Aspect	Challenges in the Existing Studies	Contribution of the Paper to Address the Challenges
Data Reliability and Underutilization of Soil-related Attribution in Data for CYP	Soil attributes were often understudied or included in only a few analyses. Given the significant influence of soil quality on crop growth and productivity, overlooking these factors undermines the robustness and accuracy of yield prediction models [7], [8], [9], [10], [3], [24].	The paper introduces saline and alkaline soil parameters collected for Rajasthan through an open-source government website.
Regional Modelling of Districts in the Dataset	Existing studies like [7], [8], [10], [13], [15], [16], [17], [18], [19], [23], [24], [25] treat the annual yield data of regions independently, ignoring the temporal and changing dependency of features like rainfall and soil nutrients.	The study reproduces the spatial relationships between districts via GCN (Graph Convolution Network)-based immediate neighbour extraction. Different from models in [21], [24], [29], where GCN and Graphical Neural Network (GNN) models are used, the proposed model in the paper performs joint spatio-temporal modelling. This is done by considering year-wise embedding learnt through graphical deep learning and directly mapping it for temporal modelling.
Joint Spatio-temporal Dependency Handling	Existing temporal models like RNN, LSTM [9], [10], [11] do not account for delayed and cumulative agro-environmental effects of rainfall and such parameters because they process historical data in a unidirectional manner.	Different from models in [21], [24], [29], where GCN and Graphical Neural Network (GNN) models are used, the proposed model in the paper performs joint spatio-temporal modelling. This is done by considering year-wise embedding learnt through graphical deep learning and directly mapping it for temporal modelling. The proposed model introduces an attention mechanism that highlights the most influential temporal points from the joint future-aware and past information of the CYP data.

model is explained in Section II. The results and experimentation are explained in Section III. Section IV presents the conclusion and Section V details the limitations and future scope of the study.

II. PROPOSED METHODOLOGY

In this section, the overall flow of the proposed model is explained. The contributions stated in Section I are detailed and achieved in this section through Data Collection, Preprocessing, and capturing Annual Regional-immediate neighbor information given in Section II B and finally the Joint Spatio-temporal feature extraction through attention-guided G-BiLSTM for Yield Prediction in Section II C. However, the motivation of deriving the contributions and formulation of the study is explained in Section II A.

A. Motivation and Proposed Formulation of the Study

Let's represent the attributes/characteristics/features of each district as $x_{d,t}$ in the final data and the actual crop output as $y_{d,t}$, where d, t represent the district and the year, respectively. Here $x_{d,t} = \{x_{d,t}^w, x_{d,t}^s, x_{d,t}^l\}$ where meteorological attributes (annual rainfall, climate type) are represented as $x_{d,t}^w$, soil attributes (soil type 1, soil type 2, saline soil, alkaline soil, phosphorus, potassium) $x_{d,t}^s$, and yield-related attributes (area, production, yield) as $x_{d,t}^l$. From these features, yield (area, production, yield, soil characteristics) and meteorological attributes (annual rainfall) change both spatially and temporally. Whereas, soil attributes (like soil type-1, soil type-2), and climate type, remain stable over time, district-wise. Attributes like $x_{d,t}^w$ varies regionally and temporally, whereas attributes in $x_{d,t}^s$ are some of the slowly changing factors that affect yield in real-world agricultural datasets. Yield prediction is a challenging spatiotemporal learning problem due to its long-term period relationships and the delayed consequences [30][31]. Each district in the dataset considered in this study is represented by a series of 13-year annual observations that include a wide range of agro-environmental characteristics $x_{d,t}$. Understanding how these characteristics build up, interact, and spread over time to affect yield in subsequent years is the real-challenge to be addressed.

Deep learning models like LSTM, RNN, and CNN may assume independence among the districts even though attributes like soil type 1, soil type 2, and climate type remain stable over time, district-wise, are different, and while processing these models, they may not fully use the spatial structure of the original data. To address this issue, the goal of the study is to employ a graph-based model to extract the immediate neighbour information of the districts ($A_{d,t}$) over the years $h_{d,t}$ which is leveraged by the spatio-temporal model for historical time series yield prediction. This process is illustrated in Fig. 3. Hence, the task of CYP is formulated as $y_{d,t}$ with $x_{d,t}$ where $t = \Delta t$ and Δt is taken as 3 years (that is, training the proposed model over 3 years of historical data). The main aim of the proposed model is to address these issues through:

$$y_{d,t}(t + 1) =$$
$$\text{Prediction}\Big(\text{GBiLSTM}(\alpha\,\text{BiLSTM}(\text{GCN}(x_{d,t}, A_{d,t})))\Big)$$
(1)

In Eq. (1), the prediction is based on the MSE loss function. In the first stage the Annual Regional-immediate neighbour, information is learnt through CYP features and $A_{d,t}$ which carries the regional information. This data which is year wise, is learnt through BiLSTM model to get the spatiotemporal embedding. Further, these are learnt over attention model to highlight the most significant temporal points towards yeild prediction. In the upcoming sections, 1) Data Collection, preprocessing and capturing Annual Regional-immediate neighbour information; 2) Joint Spatio-temporal feature extraction through attention-guided G-BiLSTM; and 3) yield prediction are detailed. The overall flow of the proposed model is illustrated in Fig. 1.

B. Data Collection, Preprocessing, and Capturing Annual Regional-Immediate Neighbour Information

The proposed model is evaluated on the primary regions for Bajra cultivation in India, specifically Rajasthan. Among the regions like Maharashtra, Gujarat, Haryana, and Uttar Pradesh, Rajasthan leads as the foremost state for Bajra production.
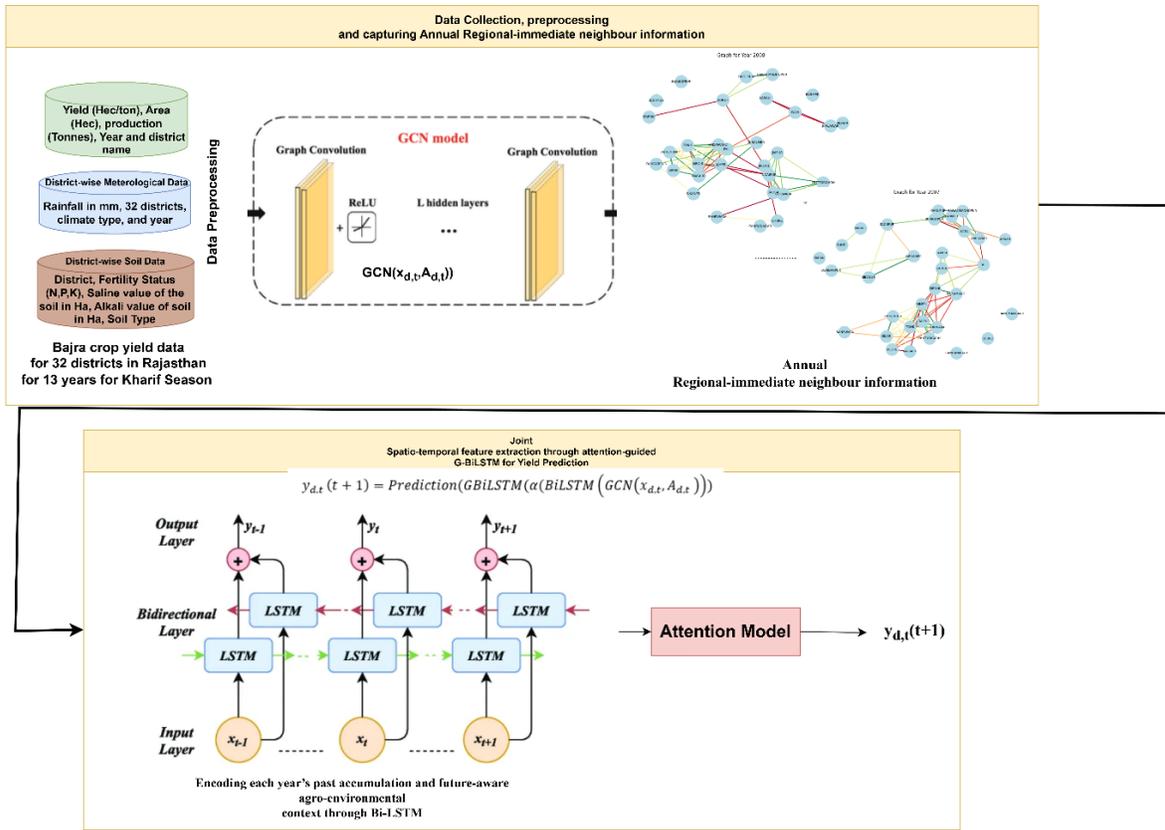
Fig. 1. Overall flow of the proposed model.

1) Geographical study of Rajasthan: In Rajasthan, the weather poses difficulties for growing various crops, particularly cereals. However, due to its resilience to harsh climatic conditions, millet farming has taken up a significant portion of the land in the region. Rajasthan accounts for 56.23% (4.43 million hectares) of the entire area dedicated to Bajra, contributing 41.40% (3.80 million tons) of India's total output. This fulfils the essential food needs of the state as well. From 2016 to 2017, Alwar district recorded the peak Bajra yield at 2149 kg/ha, surpassing all other Bajra-producing districts in Rajasthan. Traditionally, farmers depend on techniques like specific seed types, irrigation, and chemical fertilisers to sustain soil fertility and agricultural productivity. Nonetheless, these conventional methods are increasingly unreliable because of their vulnerability to the effects of climate change, further exacerbated by alterations in soil quality [14]. Therefore, examining the yield is crucial, considering elements such as soil characteristics and climatic conditions. This study collects data on soil and weather, which have influenced the Bajra crop in Rajasthan over the years, to effectively predict the yield. This study aims to analyse the yield fluctuations of Bajra and forecast its performance across the 32 districts of Rajasthan, as illustrated in Fig. 2.

This study focuses on capturing the yield variability of Bajra, as well as its prediction in Rajasthan and its 32 districts, as shown in Fig, 2.

2) Regional data collection for Rajasthan: To support the proposed approach, we curated a novel dataset that com-

Fig. 2. Geographical area of the study.

bines district-level soil properties, nutrient concentrations, and weather indicators over multiple cropping seasons. Data were sourced from verified government repositories and harmonised to ensure temporal and spatial consistency. This rich dataset serves as the foundation for the spatial modelling process and enables more accurate learning of Bajra yield patterns than the limited datasets used in prior research. In order to tackle the issues outlined in Section I, the process of data collection is aimed at obtaining thorough meteorological and soil information. This includes meteorological factors such as rainfall and climate patterns, particularly during the Kharif season, which is the primary growing season for crops like Bajra in regions across Asia and Africa. Additionally, soil data is being collected, including levels of potassium and phosphorus, as well as saline and alkaline content in the soil. Moreover, soil types and their sub-categories are being categorised to capture the diverse soil compositions in the study areas [35]. The main goal of gathering data on soil and weather is to comprehend the complex connection between these environmental elements and agricultural results. For example, changes in precipitation have a direct effect on the levels of water in the soil, which subsequently affects soil characteristics like the concentrations of nutrients and salts. By comprehensively analyzing soil and weather data, insights are aimed at being gained into how these factors interact and affect crop growth and yield. A more profound comprehension will facilitate the creation of precise predictive models and offer essential insights for agricultural choices, ultimately aiding farmers in adjusting to evolving environmental circumstances and enhancing crop yield. The data collection process involved sourcing various attributes from official Rajasthan state government websites.

3) Description of data sources: Firstly, the yield data is retrieved from the DACNET website of the Ministry of Agriculture and Farmers' Welfare [33]. This yields data for 32 districts of Rajasthan, over 23 years from 1997. The attributes collected are yield, area, production, year, and district name. To collect information about meteorological data, annual rainfall over the year [32] and climate type are collected [31]. Annual rainfall data is obtained from the Department of Water Resources, Rajasthan Government's official website [32]. This data was distributed across multiple PDF files, with each file containing daily rainfall records for different districts over the years. The data is present in the form of PDF (Portable Document Format) files, with different files for each year and each file containing daily rainfall for each district. Each one of these files had to be parsed in Python using libraries like PyPDF2 and regex. Whereas the climate data for yield prediction for each district in Rajasthan is obtained from [31]. According to [31], Rajasthan's territory is categorised into five climate types: arid, semi-arid, semi-humid, humid, and Mediterranean. Each of the 32 districts was assigned its respective climate classification based on this categorisation. This is achieved using regex functions, where each district is assigned the respective climate.

Now, the district-wise soil characteristics, including sodium, phosphorus, and potassium levels, as well as the extent of problematic soils such as saline and alkaline soils, were obtained from the Rajasthan Agriculture Statistics at a Glance report available on the official Rajasthan Government website [34]. However, a challenge arose as this data was presented in image formats. To overcome this, Microsoft's Optical Char-

TABLE III. AGRO-ENVIRONMENTAL ATTRIBUTES DESCRIPTION

Column Name	Description
District	Names of the various districts of Rajasthan used in the present study
Year	The year for which the data was collected
Area (Hectare)	Total area in hectares of the district under cultivation for that particular year
Production (Tons)	Total production in tons of Bajra for that particular year
Yield (Tons/Hectare)	Amount of production in tons per unit hectare of area under cultivation for a particular year
Soil Type 1	One type of soil found in that district of Rajasthan
Soil Type 2	Another type of soil found in that district of Rajasthan
Phosphorous	Amount of phosphorous content in the soil for the given district
Potassium	Amount of potassium content in the soil for the given district
Saline Soil (Ha)	Total area of saline soil in hectares
Sodic/Alkaline Soil (Ha)	Total area of alkaline soil in hectares
Annual Rainfall (mm)	Rainfall for the given district in mm each year
Climate Type	The climate type of each district of Rajasthan like arid, humid, wet, etc.

acter Recognition (OCR) tool was utilized to extract the data from images and store it in CSV files. Additionally, soil type data was sourced from another reference [35], which provided information on 16 different categories of soil [35]. Each region within Rajasthan was found to have soils falling into two distinct categories. Finally, the final data obtained will consist of meteorological data, soil-related data, and yield data, which is 12 attributes for 32 districts over 13 years. The data is shown in Table III, and the overall steps are given as follows:

- Yield Data Retrieval (DACNET Website): Attributes collected: yield, area, production, year, and district name for 23 years from 1997 [33].

- Then the annual rainfall data is collected from [32], and for each year, this is parsed and collected from PDF files using PyPDF2 and regex. Then it is incorporated with yield data, which leaves data for 13 years from 2007 to 2019 for 32 districts.

- Further, the climate type data is recognized with the data collected until step 2, where each district is explained with a certain type of climate as per [31]. Here, using regex, the climate type is added to the climate type column for each district in the data collected till now.

- Now, the data for soil type from [35] is added district-wise to the data collected until step 3. Then the district-wise low, medium, and high labels of Sodium, Phosphorus and Potassium in soil are merged with the data from [34]. This data was earlier present in the form of images, from which the data is extracted using OCR tools and merged with the data till now.

- The final data will have 12 attributes such as annual rainfall, district name, year, yield, area, production, soil type 1, soil type 2, phosphorus, potassium, saline soil, and alkaline soil, as given in Table III.

4) Regional data pre-processing: Initially, the data pre-processing involves verifying the presence of null values, and upon loading the dataset, it is observed that no null values

TABLE IV. DATA TYPE OF THE AGRO-ENVIRONMENTAL ATTRIBUTES IN THE CYP DATASET.

Attribute	Data Type
District	object
Year	object
Area (Hectare)	int64
Production (Tons)	int64
Yield (Tons/Hectare)	float64
Soil Type 1	object
Soil Type 2	object
Phosphorus	object
Potassium	object
Saline Soil (Ha)	float64
Alkaline Soil (Ha)	float64
Annual Normal Rainfall (mm)	float64
Climate Type	object

exist. The second step is to check for the object type of attributes and encode them. As per Table IV, a label encoder from the sklearn library is applied for the attributes with object type.

Some attributes labeled as "unnamed: 0", in the dataset, which contained serial numbers, were dropped during processing. Additionally, certain attribute columns such as potassium, climate type, saline and alkaline soil were named as unnamed and subsequently renamed to appropriate names. For instance, "unnamed: 8" representing potassium was renamed using Panda's library and regex expressions. The attribute "year" contained value in the format "2007-2008, 2008-2009." To standardize these values, a new list of individual years like ["2007", "2008" ….'"2019"] was created and merged with the existing dataset. Following this pre-processing, the initial step involves constructing graphs based on the years. Further details on these processes will be provided in the subsequent section.

The statistical summary of the attributes considered in this study is given in Section II. Rainfall is the attribute that is related to the climatic/meteorological type of data, as shown in Table IV, along with the climate type in Rajasthan for the 32 districts. While the climatic conditions affect the yield of the Bajra crop, soil attributes will also help framers take actions to improve yield. In Rajasthan, the rainfall varies from 0.82mm to 8734mm with a remarkable deviation from average, as seen in Table IV. This explains the change in climate from humid to semi-humid, which in turn affects the yield of Bajra [31]. Most of the districts in Rajasthan experience arid and humid climates [31]. The saline values in the soil also vary from 21.64 to 56369493. High saline levels in salts will affect the crops' ability to hold water content and remain hydrated. However, poor climatic conditions affect the soil conditions and lead to less yield. Thus, these conditions force farmers to use pesticides to boost the yield. Hence, a systematic process is carried out to collect the data from the respective sources.

Considering the preprocessed agro-environmental attributes data, the Annual Regional-immediate neighbour information is learnt as $GCN(x_{d,t}, A_{d,t})$, as in Eq. (1). The scenario of information considering the preprocessed agro-environmental attributes data, the Annual Regional-immediate neighbour information is learnt as regional information in $x_{d,t}$, which is explained through Fig. 3 . It displays the graph for the year 2008, and the agro-environmental attribute similarity across different districts/regions in Rajasthan.
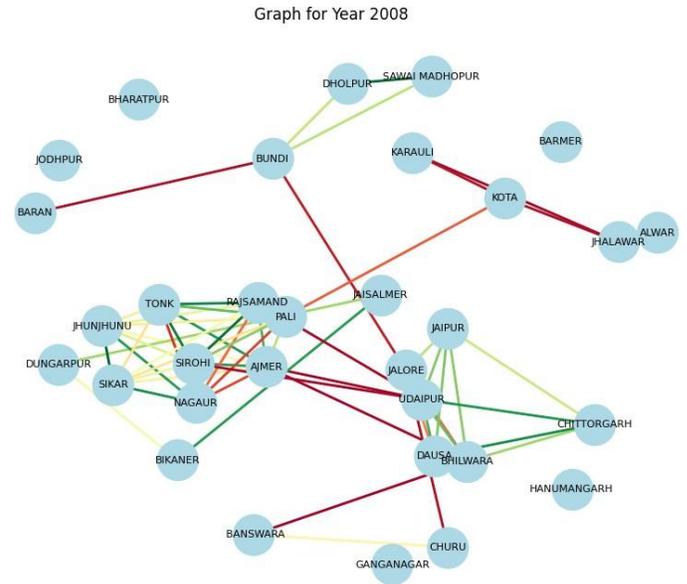

Fig. 3. Annual regional information as graph for the year 2008.

The lack of an edge linking Jodhpur and Bharatpur indicates a significant difference in the attributes described in Table II. This intriguing observation suggests that Jodhpur and Bharatpur differ from each other in certain ways. The striking differences, such as the lack of boundaries between these districts, highlight the importance of some characteristics and may indicate the distinct soil and climate types affecting their productivity. Such Regional immediate information is captured annually based on the data.

The scenario explained as similarity among the regions is captured as $A_{d,t}$. This is captured using Immediate Neighbor Extraction strategy. Let the spatial adjacency matrix $A_{d,t} \in \mathbb{R}^{N \times N}$, N is number of samples, it is be defined as:

$$A_{d,t} = \begin{cases} 1 & \text{if districts } d_i \text{ and } d_j \text{ share a boundary at year t} \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

For Eq. (2), we consider the property of cosine similarity that yields of neighboring districts that often exhibit correlated variations due to shared rainfall patterns and soil structures.

For each annual-wise graph, as shown in Fig. 3, the graph's edges are converted into edge indices. This results in a tensor of shape [2, num_{edges}], where num_{edges} is the number of the graph's edges. This tensor's columns signify edges, where the initial row holds the indices of source nodes and the subsequent row features the indices of target nodes. Therefore, this tensor acts as a compact representation of the $A_{d,t}$ in Eq. (3). The propagation rule for a l-layer GCN is followed here as per in [20]:

$$\begin{aligned} GCN(x_{d,t}, A_{d,t}) &= H^{(l+1)} \\ &= \sigma\left(D^{-\frac{1}{2}} A_{d,t} D^{-\frac{1}{2}} H^{(l)} W^{(l)} x_{d,t}\right) \end{aligned} \quad (3)$$

where, $H^{(l)}$ represents the Annual Regional-immediate Neighbour Information at layer l , $A = R + I$ is the adjacency matrix with self-loops added, as shown in [20] and the notations are similar to those in [20].

The final output of the GCN model after passing through all l-layers, explained in Eq. (3), is stored in gcn_output. This output represents the node embeddings or features learned by the GCN model for the given input graph. Each row of gcn_output corresponds to the feature representation of a node (district), and each column represents a feature dimension. **D** is the degree matrix of $A_{d,t}$, $\mathbf{W}^{(l)}$ denotes the learnable weight matrix at layer l, and σ is the activation function, ReLU. Further, the embedding's array is generated after processing of GCN. Thus, this is substituted in Eq. (1), where this is processed through BiLSTM model. This process is detailed in next section.

C. Joint Spatio-Temporal Feature Extraction Through Attention-Guided G-BiLSTM for Yield Prediction

Joint spatio-temporal is explained in Eq. (1), for each district d the Annual Regional-immediate Neighbour Information enriched embeddings $H^{(l+1)}$ is learnt for each year t that is, $H_t^{(l+1)}$ is fed into the BiLSTM.

The base LSTM units is employed here from [6], this unit maintain a memory cell c_t and hidden state h_t, defined by the classical LSTM equations. The forget, input, and output gates are [see Eq. (4)]:

$$
\begin{aligned}
f_t &= \sigma\left(W_f H_t^{(l+1)} + U_f h_{t-1} + b_f\right) \\
i_t &= \sigma\left(W_i H_t^{(l+1)} + U_i h_{t-1} + b_i\right) \\
o_t &= \sigma\left(W_o H_t^{(l+1)} + U_o h_{t-1} + b_o\right)
\end{aligned}
\tag{4}
$$

whereas, here the goal is to encode each year's past accumulation and future-aware agro-environmental context. This is achieved by updating the candidate cell state as Eq. (5):

$$
\widetilde{c}_t = tanh\left(W_c H_t^{(l+1)} + U_c h_{t-1} + b_c\right)
\tag{5}
$$

and the actual cell state updates as Eq. (6):

$$
c_t = f_t \odot c_{t-1} + i_t \odot \widetilde{c}_t
\tag{6}
$$

The hidden state is Eq. (7):

$$
h_t = o_t \odot tanh\left(c_t\right)
\tag{7}
$$

BiLSTM is employed here to extend this by processing the year's past accumulation and future-aware agro-environmental context. If the forward states are $h_t^{(f)}$ and backward states $h_t^{(b)}$, then the combined representation is Eq. (8):

$$
H_t = \left[h_t^{(f)} \| h_t^{(b)}\right]
\tag{8}
$$

Further, to get more comprehensive spatio-temporal embeddings, the attention mechanism is introduced here highlighting the most influential seasonal time points.

Given the sequence output $\{H_1, \ldots, H_T\}$, attention weights α_t as given in Eq. (1) is captured as Eq.(9):

$$
\begin{aligned}
e_t &= v^\top tanh\left(W_a H_t + b_a\right) \\
\alpha_t &= \frac{exp(e_t)}{\sum_{\tau=1}^T exp(e_\tau)}
\end{aligned}
\tag{9}
$$

The context vector becomes a weighted sum of hidden states, as in Eq. (10):

$$
c = \sum_{t=1}^T \alpha_t H_t
\tag{10}
$$

This context vector encodes the spatio-temporal essence of each district's agricultural dynamics [36][37]. Thus, the final yield prediction is captured through substituting the information in Eq. (1). Thus, the overall process detailed as a pseudocode and the proposed model is given in Algorithm 1.

Algorithm 1 Input: Agro-environmental Attributes, model parameters (hidden size, learning rate, epochs), $x_{d,t}$ (features in dataset), N (number of samples) Output: predicted yield $y_{d,t}$

Input: Agro-environmental Attributes, model parameters (hidden size, learning rate, epochs), $x_{d,t}$ (features in dataset), N (number of samples)
Output: predicted yield $y_{d,t}$
for all $d \in \{1, \ldots, 9\}$ **do**
    for all $t \in \{1, \ldots, 12\}$ **do**
        for all $N \in \{1, \ldots, 32\}$ **do**
            a. Initialize $GCN\left(x_{d,t}, A_{d,t}\right)$ #Section 2.2.3
            b. Calculate similarity among the regions captured as $A_{d,t}$ as per Eq. (2) to get the Immediate Neighbor Extraction.
            c. Construct the Annual Regional information as graph using $A_{d,t}$ # Fig. 2
        end for
    end for
end for
for all $epoches \in \{1, \ldots, 500\}$ **do**
    To capture the Annual Regional-immediate Neighbour Information at layer l by training the GCN model using Eq. (3)
end for
Initialize the attention-guided G-BiLSTM for Yield Prediction for Joint Spatio-temporal feature extraction. # Section II C
for all $epoches \in \{1, \ldots, 60\}$ **do**
    a. Input the embeddings from step 4f
    b. Encode each year's past accumulation and future-aware agro-environmental context by updating the candidate cell as shown in Eq. (5).
    c. Generate the yield prediction of the crop as output $(Y_{d,t})$ from Eq. (1).
end for

TABLE V. SIMULATION PARAMETERS FOR PROPOSED MODEL

Model Name	Parameters	Value
GCN (GCN Conv)	Number of Hidden Channels	32
	Dropout Probability	0.3
	Weight Decay	1.00E-04
	Learning Rate	0.01
	Number of Epochs	500
	Input Size	32
	Hidden Size	64
BiLSTM Model with LSTM Set As	Output Size	32
	Learning Rate	0.001
	Number of Epochs	60
	Window Size	3

III. RESULTS AND DISCUSSION

The proposed G-BiLSTM model was evaluated to assess its ability to predict the yield of Bajra in 32 districts and across 12 multi-year agricultural cycles.

A. Experiment Scenario and Settings

The experiment involved the three combined sets of data including soil nutrient, soil type profiles and high-resolution weather records that were utilized in 12 years, given in Section II A. The train-validation-test split was on five randomized folds to make it robust. The given section is aimed at providing the quantitative outcomes, statistical analysis of significance, error distributions, performance tables, and comparative interpretation to make it clear that the given method is superior to current work.This part covers the evaluation of the proposed model's training and testing, alongside proposed model (excluding soil data), baseline LSTM from [9], baseline RNN model from [38], and CNN-RNN from [16]. The suggested model's performance is evaluated against these baseline models, where the GCN is trained using a cross entropy loss function [20], and LSTM is utilised for the final prediction. Parameters like the learning rate and the count of hidden layers are adjusted to enhance RMSE, R2 score, and correlation coefficient, which act as metrics for assessing model performance. These calculations originate from the libraries Sklearn and SciPy.

The proposed model is implemented on the Google Colab platform using Python and its libraries. The parameters set are given in Table V, where GCN provides the learned spatial embedding's for span of 13 years for the 32 districts. Graph construction utilizes the NetworkX library, the GCN employs torch-geometric package and the BiLSTM model is built using PyTorch.

The model is evaluated based on, RMSE, R^2 and correlation coefficients are detailed as, RMSE is a frequently used metrics in [7,9] to assess how well regression models function, RMSE is defined as Eq. (11):

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \widehat{y_i})^2} \qquad (11)$$

R^2 is another metric for evaluating a regression model's quality of fit is the coefficient of determination. It calculates the percentage of the dependent variable's variation that can be predicted based on the independent variables (features) (goal (yield).R^2ranges from 0 to 1, where 1 indicates the perfect fit. Mathematically, $R^2$is calculated as Eq. (12):

$$R^2 = 1 - \frac{SSRes}{SSTot} \qquad (12)$$

where, $SSRes$ and $SSTot$ is the overall sum of squares and the sum of squares of the residuals, respectively. R^2 provides insight into how effectively the model accounts for data variability. The total sum of squares indicates the overall variability in the observed yields, whereas residuals show the discrepancies between the observed (actual) yields and the yields predicted by the regression model. A higher R^2 value suggests that the model fits the data better since it accounts for a greater percentage of the variability in yields.

The degree and direction of the linear link between two variables are measured by the correlation coefficient, which is frequently represented by the symbol "Pearson". Regarding yield forecasting, the correlation coefficient measures how well the predicted yields from the proposed model align with the actual observed yields. Mathematically, it is given as Eq. (13):

$$pearsonr = \frac{\sum (x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum (x_i - \overline{x})^2 \ \sum (y_i - \overline{y})^2}} \qquad (13)$$

where, x_i and y_i represent the two variables' individual data points, and $\overline{x}$ and $\overline{y}$ represent their respective means. These metrics are computed for each of the four models in this study for both training and testing.

The errors in prediction were examined in the areas that received rainfall extensively, those experiencing droughts and those that were a combination of the two. The given structure was particularly strong when it comes to the capture of deviations of monsoons. Mathematically, it is explained by the attention mechanism which places substantial weights on sudden changes in rainfalls:

$$\alpha_t = \frac{exp(e_t)}{\sum e_\tau}, e_t \propto tanh(W_a H_t),$$

resulting in high sensitivity to time steps, where climatic volatility is highest. Only in extreme anomaly years when the rainfall varied more than 40% of the 15-year mean, error peaks were realized but the model still provided an RMSE of less than 60, which was better than LSTM and GNN models.

B. Analysis of the Proposed Model

When the model is evaluated for heatmap representations of Agro-environmental Attributes in Fig. 4, it is seen that the attributes related to meteorological type are mostly used for yield prediction. To understand the effect of the attributes in the dataset considered here over the yield of Bajra Pearson's correlation is conducted, and the heatmap is illustrated in Fig. 4. Before conducting this, the data is encoded into numbers using the label encoder library from the Sklearn package. Further, the Pearson correlation is also imported from Sklearn. The colour legend in the heatmap shows the high and low correlation between the attributes and yield based on the values varying

from -1 to 1. The analysis of the heatmap reveals notable patterns in the correlation between various attributes and the yield of the Bajra crop. Specifically, soil attributes, such as soil type 2, exhibit a high positive correlation with the yield attribute. Rainfall attribute also shows positive trend with yield, the values exhibits variations along the year as seen in Fig. 5.
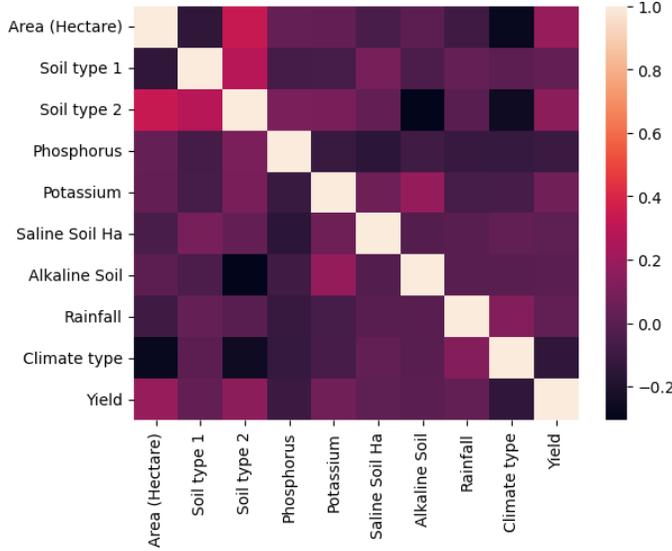

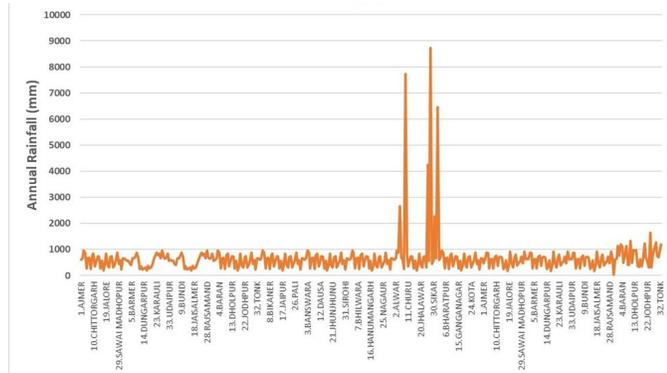Fig. 4. Heatmap representation of the correlation matrix of the agro-environmental attributes.

Fig. 5. Annual average normal rainfall (mm) over the districts from 2007 to 2020.

Fig. 6 shows minimal deviation between actual and predicted crop yields, highlighting the model's accuracy. Incorporating soil data, which significantly impacts crop yield allows the proposed model to effectively capture the influence of soil characteristics on Bajra crop yield.

Fig. 7 shows the scatter plot analysis of predictions after using LSTM on left and after using Bi-LSTM on right. The proposed model, when encodes the agro-environment data through normal data, the data points are more scattered and away from actual values showing high errors. When the candidate cell is updated to capture both past and future trends if rainfalls, the delayed trend is reduced causing the predicted points to be more near to actual values. This leads to high correlation coefficients and low RMSE. The detailed analysis
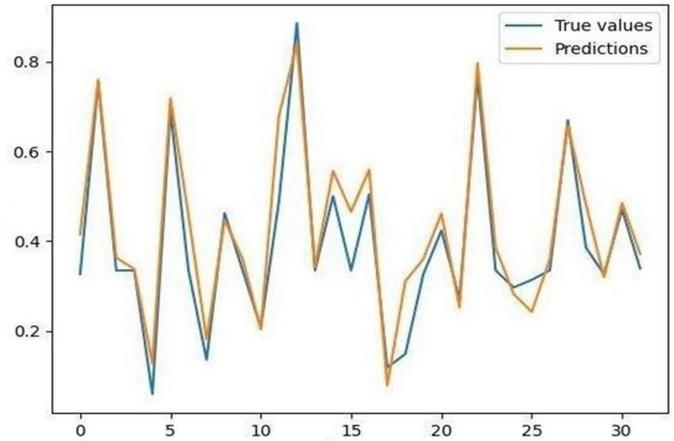

Fig. 6. Comparison of true values versus predicted yield values using immediate neighbor extraction and temporal analysis model.
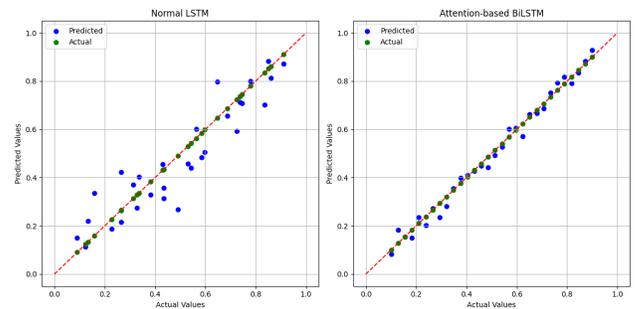

Fig. 7. Comparison of actual values versus predicted yield values when only temporal values are captured (normal LSTM (left)) and when each year's past accumulation and future-aware agro-environmental context is encoded (attention-based BiLSTM (right)).

of performance of proposed model compared to baselines is shown in next section.

1) Statistical analysis: To analyze the generalization of the proposed model, the model is run for 5-fold experiments with different year train-test splits. This is shown in Fig. 8, where the central line shows the median range, and the top line of the box shows the IQR (inter quartile range). The analysis shows that not much variance is exhibited as the mean RMSE is around 10.89. the values over 5-fold doesn't vary much as upper and lower quartiles show 10.9 and 10.7 RMSE.

Considering this, the proposed model shows low RMSE compared to baseline and shows reliability for future real-time CYP task.

2) State-of-the-art comparison: The current study compares a number of baseline models, including CNN-RNN, RNN, and LSTM, as shown in Table VI. These models were implemented by keeping the training and testing data same as proposed to maintain fairness. The default LSTM model's architecture largely adheres to the guidelines given in Table VI, with 70% of the dataset used for training and the remaining 30% designated for testing. Similar to the baseline LSTM model [9], the baseline RNN model, according to reference [38], has a window size of 3 and a hidden size of 64. With a learning rate of 0.001%, the Adam optimiser is used to train

TABLE VI. COMPARATIVE ANALYSIS WITH THE STATE-OF-THE-ART MODELS

| Model	RMSE Training Loss	R^2 Training Score	Training Correlation Coefficient	RMSE Testing Loss	R^2 Testing Score	Testing Correlation Coefficient
Proposed Model	7.6	84.8	94.1	10.8	68.75	89.1
Proposed Model without Soil Data	8.1	82.8	91	11	65.5	83.2
LSTM [9]	95.1	74.81	93.63	11.21	64.53	88.87
RNN [38]	17	45.1	77.1	18	55.1	55
CNN-RNN [16]	19	48	68.6	21	59	50
Gradient Boost Algorithm [39]	15	45	90.9	8.9	55.78	87.9
GNN [40]	17.4	51.8	86.02	14	58.5	87.8
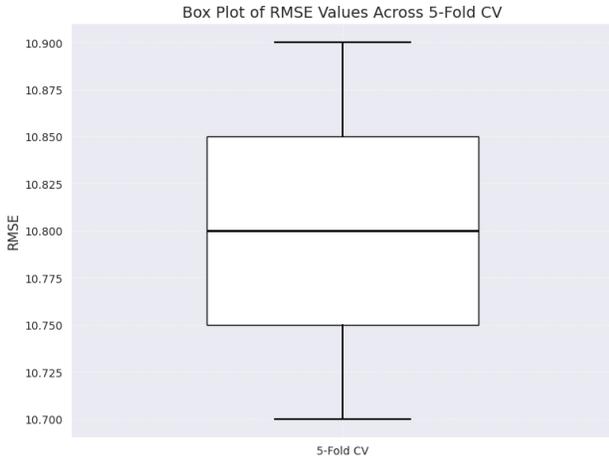
Fig. 8. Statistical analysis of RMSE for proposed model for 5-folds.

both LSTM and RNN models. Performance is then optimised by iteration over 60 epochs.

The proposed model demonstrates a 3.29% improvement in correlation coefficient during training and a 6.62% improvement during testing compared to the proposed model without soil data. This improvement suggests that incorporating soil data enhances the model's ability to encode each year's past accumulation and future-aware agro-environmental context by updating the candidate cell, resulting in stronger correlations and high performance. The proposed model is compared with Gradient boosting [39], which focuses on improving the performance in a sequential manner. However, it shows high errors compared to proposed model. The proposed model shows low RMSE and high correlation coefficient that is 1.45% high when tested.

IV. CONCLUSION

In this study for Bajra crop yield prediction, an deep learning model is proposed. The study proposes a region-based spatiotemporal model with an attention-guided Bidirectional LSTM (Long-Short Term Memory) framework for CYP, termed as G-BiLSTM. The proposed model reproduces the spatial relationships between districts via GCN (Graph Convolution Network) -based immediate neighbour extraction. Further, a Bidirectional LSTM is used to model multi-year CYP temporal features, allowing each annual observation to be encoded using both past and future temporal context. A

variance-reduced and comprehensible representation is produced by integrating an attention mechanism to adaptively highlight the most informative years within a temporal window. Using 15 agro-environmental characteristics, including understudied elements like saline and alkaline soil composition, the framework is assessed on a dataset that includes 32 districts in Rajasthan over 13 years (2007–2019). The proposed model achieves 7.6% if RMSE training loss with 10.8 RMSE testing loss. The proposed model demonstrates a 3.29% improvement in correlation coefficient during training and a 6.62% improvement during testing compared to the proposed model without soil data.

The proposed model outperforms the baselines by leveraging the past, future temporal context and highlighted information of important temporal points. The existing study focus on a single and scope area of cultivation for a single crop. However, in future, the study will focus on the prediction of vast area and various crops like sugarcane. The current model captures the regional graph based on fixed adjacency matrix, in future, the author plan to capture the dynamically changing information in real-time. The existing model uses a single type of data from open-source, the future work will focus on using satellite imagery data as well. This will help improve the model predictions and scope.

V. LIMITATIONS AND FUTURE SCOPE

Though the proposed model shows good performance, there are some limitations.

The research is based on thirteen years of curated district-level data in Rajasthan, and this restricts the time modeling horizon. Although some agro-environmental processes (e.g. salt degradation or climatic oscillations) might have cumulative behavior (that is, over longer times), the three-year sliding window adopted is a proportionate and statistically balanced design choice, within the limited time frame. Further investigation into the temporal dataset will be done in the future, and studies of the multi-scale or hierarchical temporal modeling techniques will be done to capture more long-term oscillatory patterns.

The proposed study is modeled using historical variability and tested on a held-out three-year test set to predict, as real-world forecasting does not know future values of those variables. Nevertheless, regimes may become too extreme so as to keep predictive power (e.g. long periods of multi-year droughts not seen in the observed pattern), as is the case with most data-driven models. In the future, the study will

be enhanced by stress-testing in extreme climatic conditions and explicit failure mode analysis to enhance robustness and transparency.

The agro-environmental characteristics (soil characteristics, salinity, rainfall, climate) of agro-environmental modeling are done at a macro district level using the current graph formulation. Other socio-economic variables, management practices, level of irrigation, and dynamics of pests are not encompassed because of the lack of uniform longitudinal data of the 32 districts on a district-wide basis. The work in the future will develop the graph to incorporate the socio-economic factors.

The proposed model is scalable because the adjacency matrix and feature representation can be applied to other areas, but it is not covered by current empirical cross-regional validation. The model will be tested in various agro-climatic regions in future to enhance the generalizability claims.

The BiLSTM part is being trained on past sequences in which all the years used in the training set are known. Although this means that prediction on unknown test years is realistic, the framework is not at present formulated to carry out in-season forecasting in real-time, where future information of the seasons is unknown. The problem of extending the model to operational forecasting situations will be dealt with in future studies.

REFERENCES

[1] L. Visaria and P. Visaria, "Prospective population growth and policy options for India, 1991–2101," *Population Research and Policy Review*, vol. 15, pp. 57–75, 1996.

[2] D. Elavarasan and P. D. Vincent, "Crop yield prediction using deep reinforcement learning model for sustainable agrarian applications," *IEEE Access*, vol. 8, pp. 86886–86901, 2020.

[3] T. Saranya, C. Deisy, S. Sridevi, and K. S. M. Anbananthen, "A comparative study of deep learning and Internet of Things for precision agriculture," *Engineering Applications of Artificial Intelligence*, vol. 122, p. 106034, 2023.

[4] A. Sharma, M. Georgi, M. Tregubenko, and A. Tselykh, "Enabling smart agriculture by implementing artificial intelligence and embedded sensing," *Computers & Industrial Engineering*, vol. 165, p. 107936, 2022.

[5] P. M. Gopal and R. Bhargavi, "A novel approach for efficient crop yield prediction," *Computers and Electronics in Agriculture*, vol. 165, p. 104968, 2019.

[6] S. Sharma, S. Rai, and N. C. Krishnan, "Wheat crop yield prediction using deep LSTM model," *arXiv preprint arXiv:2011.01498*, 2020.

[7] K. Jhajharia and P. Mathur, "Machine learning based crop yield prediction model in Rajasthan region of India," *Iraqi Journal of Science*, pp. 390–400, 2024.

[8] S. K. Sharma, D. P. Sharma, and K. Gaur, "Machine learning techniques for crop yield forecasting in semi-arid (3A) zone, Rajasthan (India)," *Current Agriculture Research Journal*, vol. 11, no. 3, pp. 1–10, 2023.

[9] K. Jhajharia, P. Mathur, S. Jain, and S. Nijhawan, "Crop yield prediction using machine learning and deep learning techniques," *Procedia Computer Science*, vol. 218, pp. 406–417, 2023.

[10] A. Vashisth and A. Goyal, "Prediction of mustard yield using different machine learning techniques: A case study of Rajasthan, India," *International Journal of Biometeorology*, vol. 67, no. 3, pp. 539–551, 2023.

[11] P. S. and M. G., "Performance evaluation of best feature subsets for crop yield prediction using machine learning algorithms," *Applied Artificial Intelligence*, vol. 33, no. 7, pp. 621–642, 2019.

[12] R. A. Medar, V. S. Rajpurohit, and A. M. Ambekar, "Sugarcane crop yield forecasting model using supervised machine learning," *International Journal of Intelligent Systems and Applications*, vol. 11, no. 8, pp. 11–19, 2019.

[13] E. Khosla, R. Dharavath, and R. Priya, "Crop yield prediction using aggregated rainfall-based modular artificial neural networks and support vector regression," *Environment, Development and Sustainability*, vol. 22, pp. 5687–5708, 2020.

[14] S. Dharmaraja, V. Jain, P. Anjoy, and H. Chandra, "Empirical analysis for crop yield forecasting in India," *Agricultural Research*, vol. 9, pp. 132–138, 2020.

[15] M. Shahhosseini, G. Hu, S. Khaki, and S. V. Archontoulis, "Corn yield prediction with ensemble CNN–DNN," *Frontiers in Plant Science*, vol. 12, p. 709008, 2021.

[16] S. Khaki, L. Wang, and S. V. Archontoulis, "A CNN–RNN framework for crop yield prediction," *Frontiers in Plant Science*, vol. 10, p. 1750, 2020.

[17] S. K. Sharma, D. P. Sharma, and K. Gaur, "Crop yield predictions and recommendations using random forest regression in 3A agroclimatic zone, Rajasthan," *Journal of Data Acquisition and Processing*, vol. 38, no. 2, pp. 1635–1645, 2023.

[18] K. O'Shea and R. Nash, "An introduction to convolutional neural networks," *arXiv preprint arXiv:1511.08458*, 2015.

[19] Z. C. Lipton, J. Berkowitz, and C. Elkan, "A critical review of recurrent neural networks for sequence learning," *arXiv preprint arXiv:1506.00019*, 2015.

[20] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," *arXiv preprint arXiv:1609.02907*, 2016.

[21] H. Jiang, C. Zhang, Y. Qiao, Z. Zhang, W. Zhang, and C. Song, "CNN feature based graph convolutional network for weed and crop recognition in smart farming," *Computers and Electronics in Agriculture*, vol. 174, p. 105450, 2020.

[22] Q. Zhang, B. Li, Y. Zhang, and S. Wang, "Suitability evaluation of crop variety via graph neural network," *Computational Intelligence and Neuroscience*, 2022.

[23] X. Song, P. Wang, J. Yu, X. Liu, J. Liu, and R. Yuan, "Relationships between precipitation, soil water and groundwater at Chongling catchment with the typical vegetation cover in the Taihang mountainous region, China," *Environmental Earth Sciences*, vol. 62, pp. 787–796, 2011.

[24] T. Sudha, B. Ramesh, D. P. Biradar, V. C. Patil, N. S. Hebsur, S. S. Adiver, and S. Geeta, "Documentation of cultivation practices of cotton in different soil types," *Karnataka Journal of Agricultural Sciences*, vol. 24, no. 5, pp. 688–691, 2011.

[25] W. Shi, F. Tao, and Z. Zhang, "A review on statistical models for identifying climate contributions to crop yields," *Journal of Geographical Sciences*, vol. 23, pp. 567–576, 2013.

[26] S. Vashisht, P. Kumar, and M. C. Trivedi, "Enhanced GRU-BiLSTM technique for crop yield prediction," *Multimedia Tools and Applications*, vol. 83, no. 41, pp. 89003–89028, 2024.

[27] Z. Jia, K. Wu, H. Wang, W. Zeng, Y. Guo, and D. Liang, "An improved crop yield prediction using CNN–BiLSTM model with attention mechanism," *Journal of the ASABE*, vol. 67, no. 6, pp. 1459–1467, 2024.

[28] L. Zhang, C. Li, X. Wu, H. Xiang, Y. Jiao, and H. Chai, "BO-CNN-BiLSTM deep learning model integrating multisource remote sensing data for improving winter wheat yield estimation," *Frontiers in Plant Science*, vol. 15, p. 1500499, 2024.

[29] U. Nirosha and G. Vennila, "Enhancing crop yield prediction for agriculture productivity using federated learning integrating with graph and recurrent neural networks model," *Expert Systems with Applications*, p. 128312, 2025.

[30] Z. Hao, V. P. Singh, and Y. Xia, "Seasonal drought prediction: Advances, challenges, and future prospects," *Reviews of Geophysics*, vol. 56, no. 1, pp. 108–141, 2018.

[31] A. Gunawat, S. K. Dubey, and D. Sharma, "Development of indices for aridity and temperature changes pattern through GIS mapping for Rajasthan, India," *Climate Change and Environmental Sustainability*, vol. 4, no. 2, pp. 178–189, 2016.

[32] Department of Water Resources, Government of Rajasthan, "Annual rainfall data," [Online]. Available: https://water.rajasthan.gov.in/

[33] Ministry of Agriculture and Farmers Welfare, Government of India, "Area and production statistics of Bajra," [Online]. Available: https://aps.dac.gov.in/

[34] Department of Agriculture, Government of Rajasthan, "District-wise soil characteristics," [Online]. Available: https://agriculture.rajasthan.gov.in/

[35] Department of Agriculture, Government of Rajasthan, "Agro-climatic zones and soil type data," [Online]. Available: https://agriculture.rajasthan.gov.in/

[36] I. G. Tende, K. Aburada, H. Yamaba, T. Katayama, and N. Okazaki, "Development and evaluation of a deep learning based system to predict district-level maize yields in Tanzania," *Agriculture*, vol. 13, no. 3, p. 627, 2023.

[37] T. Sudha, B. Ramesh, D. P. Biradar, V. C. Patil, N. S. Hebsur, S. S. Adiver, and S. Geeta, "Documentation of cultivation practices of cotton in different soil types," *Karnataka Journal of Agricultural Sciences*, vol.

24, no. 5, pp. 688–691, 2011.

[38] N. Bali and A. Singla, "Deep learning based wheat crop yield prediction model in Punjab region of North India," *Applied Artificial Intelligence*, vol. 35, no. 15, pp. 1304–1328, 2021.

[39] O. E. Oluwole, O. O. Edgar, and B. D. Fredrick, "Integrating soil nutrients and location weather variables for crop yield prediction," *International Journal of Innovative Science and Research Technology*, vol. 8, no. 3, pp. 2317–2323, 2023.

[40] E. S. M. El-Kenawy, A. A. Alhussan, N. Khodadadi, S. Mirjalili, and M. M. Eid, "Predicting potato crop yield with machine learning and deep learning for sustainable agriculture," *Potato Research*, pp. 1–34, 2024.