Rainfall prediction plays a vital role in the management of water resources, directly impacting agricultural productivity, urban water supply, and hydroelectric power generation. Accurate rainfall prediction is crucial for effective water management and the mitigation of natural disasters like floods and droughts [1]. To predict rainfall, different approaches have been employed, including statistical models, machine learning algorithms, and artificial intelligence techniques [2]. Statistical models like the Markov chain model and the Auto-regressive Integrated Moving Average (ARIMA) model have been extensively used for rainfall prediction [3]. Furthermore, machine learning techniques, such as artificial neural networks (ANNs), decision trees, and support vector machines (SVMs), have contributed to the improvement of rainfall prediction [4]. In recent years, researchers have also explored the potential of advanced deep learning techniques like long short-term memory (LSTM) and convolutional neural networks (CNNs) for accurate rainfall prediction [5].

This study addresses two key aspects related to rainfall prediction. Firstly, it focuses on the utilization of deep learning techniques, particularly the Long Short-Term Memory (LSTM) algorithm, to construct a predictive model for daily rainfall. While previous research extensively explored the application of Artificial Neural Networks (ANN), Multi-Layer Perceptron (MLP), and linear regression models in rainfall prediction, there is a significant research gap regarding the implementation of deep learning methods specific to the Jordan area. This gap is crucial to address as weather patterns and parameters vary across different locations, making models developed for one area unsuitable for another.

To support this research endeavor, a novel dataset has been assembled, encompassing climate data collected from 1988 to 2017 and rainfall records spanning from 1987 to 2018. The dataset includes a wide range of meteorological variables such as temperature, precipitation, wet bulb temperature, dry bulb temperature, and vapour pressure. It is structured to facilitate temporal analysis, with columns indicating the day, month, and year of each data point. Moreover, the dataset has undergone preprocessing, allowing for the calculation of monthly and yearly means for various meteorological variables. This preprocessing step enhances the dataset's utility for research purposes, enabling a comprehensive understanding of climate conditions throughout the dataset's timeframe and aiding the identification of significant patterns and trends that may not be apparent at the daily level. Additionally, rigorous data cleaning and preprocessing procedures have been implemented to ensure the dataset's accuracy and reliability for subsequent analyses.

A thorough examination of the existing literature highlights the potential for enhancing the accuracy of rainfall prediction through the utilization of deep-learning models. Overall, this study presents a deep learning-based
approach, utilizing the LSTM algorithm, for daily rainfall prediction in the Jordan area. It is supported by a comprehensive dataset that offers valuable insights into temperature variability, precipitation patterns, and the thermodynamic properties of the atmosphere. By addressing the research gap and considering the location-specific factors, this study contributes to the advancement of rainfall prediction techniques and provides a foundation for improved water resource management and natural disaster mitigation.

The subsequent sections of this paper are organized as follows: Section 2 provides the background of the research and presents a comprehensive literature review. Section 3 outlines the methodology employed in this study. The experimental results are presented and discussed in Section 4. Finally, Section 5 concludes the paper by summarizing the key findings and presenting recommendations for future research.

II. BACKGROUND AND LITERATURE REVIEW

Water plays a crucial role in households, agriculture, and industries, making it an indispensable element. However, Jordan, a country characterized as semi-arid, faces severe water scarcity issues, negatively impacting various aspects of life, particularly critical sectors that directly affect citizens' well-being [6]. The limited water resources in Jordan pose a significant challenge, exacerbating its vulnerability to the effects of climate change. Studies have indicated a declining trend in both annual and seasonal precipitation in the country, which is expected to persist [7].

A comprehensive study examining the impact of climate change on water resources in Jordan utilized observational data and hydrological models. The findings highlighted a significant decrease in annual and seasonal precipitation, particularly during the winter and spring seasons. The study also projected that this decreasing trend in precipitation is likely to continue based on climate change scenarios [7]. Additionally, another study conducted by [8] investigated the variability and trends of precipitation in Jordan using observational data from weather stations. The results demonstrated a negative trend in precipitation, with a notable reduction in the number of rainy days and total rainfall. The study further identified the changes in circulation patterns of the Mediterranean Sea and the Red Sea as the primary factors contributing to the precipitation changes, particularly in the northern and central regions of Jordan [8].

Several factors contribute to the decline in Jordan's annual per capita water share, which currently stands at less than ninety cubic meters. Notably, the increasing population numbers and expanding economic activities demanding substantial water quantities impose significant pressure on the water resources [9]. Moreover, since 1948, Jordan has experienced substantial influxes of refugees, including the influx of Palestinian refugees over the past decades and the current wave of Syrian refugees, with over one million Syrian refugees currently residing in the country [10].

Jordan's climate exhibits diverse conditions, ranging from a Mediterranean climate to a desert climate, with an overall arid nature. Winter temperatures typically range from 9-13°C in the southern and northern highlands, while desert regions experience temperatures of 19-22°C. During summer, temperatures in the Jordan Valley range between 38-39°C, and in desert regions, they fall within 26-29°C. The majority of precipitation, approximately 75%, occurs during the winter months. Jordan's climate is also influenced by meteorological phenomena like the Dry Sirocco (Khamsin) winds, leading to significant temperature fluctuations, sometimes reaching an increase of up to 15°C [11].

According to data from the Jordan Meteorological Department, the total annual rainfall in Jordan displayed significant variation between 1987 and 2018, ranging from a low of around 100 millimeters (mm) in 1991 and 2002 to a high of over 500 mm in 1992 and 1995. The average annual rainfall during this period was approximately 250 mm [12]. Rainfall patterns in Jordan are highly erratic and can be influenced by various factors, including the El Nino-Southern Oscillation (ENSO), the North Atlantic Oscillation (NAO), and the Mediterranean Oscillation (MO). These climate patterns have an impact on the strength and trajectory of weather systems that bring rainfall to the region. Rainfall is typically heaviest in the northwest and the highlands, while the eastern desert region receives the least amount of precipitation [12].
The rainy season in Jordan typically spans from November to April, with most of the rainfall occurring in December and January [12]. In recent decades, there has been an overall decrease in rainfall in Jordan, and climate change is believed to be a contributing factor, as indicated by several studies [13]. This decline in rainfall has had significant implications for the country's water resources, leading experts to express concerns about potential future water shortages [13].

The regions of Jordan that experience the highest levels of precipitation are the elevated highlands, where long-term annual averages range from 300 to 600mm. As one moves eastward from the highlands, there is a noticeable decrease in precipitation, with a more pronounced decrease towards the west (Figure 1). For instance, within a distance of 10 km and an altitude difference of 1200 m, the average annual precipitation decreases from 600 mm in the northern highlands to 250 mm in the Jordan Valley to the west. The decrease in precipitation towards the east is less pronounced than towards the west; for example, from 300 mm/year in Shoubak to 50 mm/year approximately 30 km eastward in the Jafr area [10].

A more comprehensive understanding of Jordan’s water situation reveals that only around 3% of the country’s total area receives an average annual precipitation exceeding 300 mm. This minimum amount is necessary for rain-fed agriculture under the prevailing climate conditions in the country. Consequently, it can be inferred that approximately 83% of the total precipitation falls over areas unsuitable for rain-fed agriculture, leaving only 17% of precipitation potentially beneficial for such purposes. The remaining 83% of precipitation requires costly technical interventions to make it partially usable. Some of the precipitation water flows through wadis and is collected in reservoirs, while the rest infiltrates the ground, replenishing the groundwater resources. Being a semi-arid country, Jordan experiences increased salt contents in precipitation water due to atmospheric dust and the low levels of rainfall [10], [14]. The Jordanian government attributes the issue of water scarcity to factors such as heavy reliance on groundwater, declining rainfall, and limited surface water availability.

The first reservoir, Sultani, was built in 1962. The number of reservoirs currently in Jordan is 23 reservoirs. Table 1 shows the list of reservoirs in Jordan. The 23 reservoirs have provided Jordan with a total of 355 million cubic meters of water. The size of the reservoirs ranged between large and small in an effort by governments to collect the most significant possible amount of rain and torrential water in Jordan.
Table 1. List of reservoirs in Jordan

<table>
<thead>
<tr>
<th>Reservoir Name</th>
<th>Period operation</th>
<th>Capacity (MCM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Sultani</td>
<td>1962 – today</td>
<td>1.2</td>
</tr>
<tr>
<td>2. Ziqlab Dam</td>
<td>1964 -today</td>
<td>4</td>
</tr>
<tr>
<td>3. Qatana</td>
<td>1964 – today</td>
<td>4</td>
</tr>
<tr>
<td>4. Shueib dam</td>
<td>1969 – today</td>
<td>1.4</td>
</tr>
<tr>
<td>5. Kafrein</td>
<td>1968 – today</td>
<td>8.5</td>
</tr>
<tr>
<td>6. Wadi Al-Arab</td>
<td>1986-today</td>
<td>16.8</td>
</tr>
<tr>
<td>7. Talal</td>
<td>1987-today</td>
<td>75</td>
</tr>
<tr>
<td>8. Rajil</td>
<td>1992 -today</td>
<td>3.5</td>
</tr>
<tr>
<td>9. Siwaqa</td>
<td>1993-today</td>
<td>2.5</td>
</tr>
<tr>
<td>10. Karama</td>
<td>1997-today</td>
<td>55</td>
</tr>
<tr>
<td>11. Mujib</td>
<td>1999-today</td>
<td>29.8</td>
</tr>
<tr>
<td>12. Wala</td>
<td>2002- today</td>
<td>9.3</td>
</tr>
<tr>
<td>13. Madoneh</td>
<td>2003- today</td>
<td>0.09</td>
</tr>
<tr>
<td>15. Wihda</td>
<td>2006-today</td>
<td>110</td>
</tr>
<tr>
<td>16. Wadi Butum</td>
<td>2011-today</td>
<td>0.47</td>
</tr>
<tr>
<td>17. Ibn Hamad</td>
<td>2015-today</td>
<td>4</td>
</tr>
<tr>
<td>18. Kufranjah</td>
<td>2017-today</td>
<td>7.8</td>
</tr>
<tr>
<td>19. Karak</td>
<td>2017-today</td>
<td>2</td>
</tr>
<tr>
<td>20. Wadi Fidan</td>
<td>2017-today</td>
<td>3.4</td>
</tr>
<tr>
<td>21. Zarqa Ma’en</td>
<td>2017-today</td>
<td>2</td>
</tr>
<tr>
<td>22. Ajloun</td>
<td>2018-today</td>
<td>1</td>
</tr>
<tr>
<td>23. Rahma</td>
<td>2019-today</td>
<td>0.4</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>—</td>
<td><strong>354.86</strong></td>
</tr>
</tbody>
</table>

Al Wihda reservoir, located on the Yarmouk River, is the largest reservoir in Jordan and is shared between Jordan and Syria. It spans 110 meters in length and has a storage capacity of 115,000,000 cubic meters. Its primary purpose is to supply water to Jordan and provide hydroelectric power to Syria. The King Talal reservoir, established in 1977, is the second-largest reservoir in Jordan. It serves irrigation and electricity generation purposes, with a storage capacity of 75 million cubic meters. According to [9], studies indicate a projected decrease in Jordan's per capita water share to 60 cubic meters annually for all uses by 2040, compared to the current share of approximately 90 cubic meters [9].

Rainfall estimation plays a vital role in various applications, including reducing traffic accidents and congestion, improving water management, and mitigating floods. However, traditional theory-driven numerical weather prediction (NWP) approaches face challenges, such as limited understanding of physical processes, extracting useful knowledge from large volumes of observational data, and the need for substantial computational resources [15]. Deep learning methods, which have demonstrated success in domains like computer vision, speech recognition, and time series prediction, show promise in extracting temporal and spatial features from meteorological data, which is characterized by its vast geospatial nature. Deep learning-based weather prediction (DLWP) is expected to complement conventional methods [16].

Previous studies have utilized machine learning algorithms, including Multilayer Perceptron (MLP), for rainfall prediction based on personal experience and observation of rainfall parameters. However, the potential of deep learning in rainfall prediction is limited, particularly when utilizing sensor-based datasets. Recent surveys indicate that MLP is the most commonly used neural network model for rainfall forecasting [17], [18], [19]. This research
J. Electrical Systems 20-7s (2024): 1696-1704

aims to employ machine learning algorithms and artificial intelligence to conduct an analytical study on rainfall levels in Jordan and their impact on the water supply to Jordanian reservoirs over the next twenty years. Machine learning (ML) enables machines to perform tasks without explicit programming by training them on past data to predict future information. Researchers have suggested using ML systems to forecast water levels, utilizing various datasets collected worldwide [20]-[23]. The methods range from traditional machine learning approaches to those that rely on deep learning (DL) algorithms. Support Vector Machines (SVM) have been widely used by researchers [24]-[26].

III. METHODOLOGY

The objective of this study is to develop a predictive model for daily rainfall using deep learning techniques, specifically the Long Short-Term Memory (LSTM) algorithm. The research aims to address the research gap in rainfall prediction specific to the Jordan area and overcome the limitations of previous models that may not be suitable for this region due to variations in weather patterns and parameters. The methodology involved assembling a dataset incorporating climate data and rainfall records, followed by preprocessing procedures to ensure its suitability for research purposes. Subsequently, the Long Short-Term Memory (LSTM) algorithm was employed to develop the predictive model for daily rainfall. Figure 2 summarizes the steps followed in this research. The upcoming subsections will provide a detailed explanation of research steps.

A. Dataset

The dataset used in this study has been assembled from data acquired from the Jordanian Ministry of Water and Irrigation Jordan Valley Authority. The dataset encompasses a comprehensive collection of daily rainfall records ranging from 1987 to 2018. In addition to the rainfall data, consistent records of water levels in Jordanian reservoirs have been meticulously recorded since their establishment. Furthermore, climatic conditions in Jordan have been systematically documented based on records provided by the Ministry of Water and Irrigation.

The compiled dataset comprises climate data collected between 1988 and 2017, encompassing a wide range of meteorological variables, such as temperature, precipitation, wet bulb temperature, dry bulb temperature, and vapor pressure. Organized in a manner that facilitates temporal analysis, the dataset includes specific columns for the day, month, and year of each data point. Additionally, it provides information on daily maximum and minimum temperatures, mean temperature, and the temperature difference between the maximum and minimum values. This comprehensive arrangement allows for a detailed examination of temperature variability within a given day. The inclusion of wet bulb temperature, dry bulb temperature, and vapor pressure in the dataset provides valuable insights into the thermodynamic state of the air and the amount of water vapor present in the atmosphere. These variables contribute to a better understanding of overall climate conditions during the dataset’s covered period.

Spanning several decades, from 1988 to 2017 for climate records and 1987 to 2018 for rainfall records, the dataset enables the exploration of long-term climate trends and variability. This temporal coverage is essential for
investigating changes in temperature and precipitation patterns over time and their potential connections to larger-
scale climate phenomena or human activities. The dataset demonstrates suitability for various climate research
applications, including the study of temperature variability, precipitation patterns, and the thermodynamic properties
of the atmosphere.

The dataset underwent several processing steps to enhance its utility for various research applications. One of
the key processes involved calculating monthly and yearly means for different meteorological variables. By
aggregating the data at these temporal scales, it becomes possible to examine larger-scale climate patterns and
trends. The calculation of monthly means for temperature and precipitation variables, including maximum and
minimum temperature, mean temperature, wet bulb temperature, dry bulb temperature, and vapor pressure, provides
valuable insights into the overall climate conditions for each month. This analysis allows for the identification of
monthly variability in temperature and precipitation patterns, as well as the detection of any potential anomalies or
outliers within the data. Moreover, calculating yearly means for the same set of variables enables the investigation
of long-term climate trends and variability. This approach helps to understand changes in temperature and
precipitation patterns over time and their potential relationships with larger-scale climate phenomena or human-
induced activities. By examining annual averages, researchers can gain a comprehensive understanding of the
dataset's temporal characteristics and draw meaningful conclusions about climate change and its impacts.

Overall, the processing steps applied to the dataset, including the calculation of monthly and yearly means,
contribute to its enhanced utility and facilitate a more comprehensive analysis of temperature and precipitation
patterns at different temporal scales.

Furthermore, the calculation of monthly and yearly means provides a more comprehensive picture of the climate
conditions over the period covered by the dataset (1988-2017 for climate records and 1987-2018 for rainfall
records), which can aid in identifying patterns and trends that may not be apparent when analyzing the data at the
daily level. Overall, the processing step of calculating monthly and yearly means for various meteorological
variables has dramatically enhanced the dataset’s utility for climate research. It allows for a more comprehensive
understanding of the climate conditions over the period covered by the dataset. It facilitates the identification of
patterns and trends that may need to be apparent when analyzing the data at the daily level.

The cleaning and processing of the climate data dataset is a crucial step in ensuring the accuracy and reliability
of the data for subsequent analyses. In this case, the dataset had many missing and unrealistic values, which needed
to be addressed to ensure the research findings’ validity. One approach to handling missing values is to remove the
rows that contain missing data, which may result in a loss of information. However, a more appropriate approach
was taken in this case: to apply reasonable restrictions to ensure the data is correct. This approach helps minimize
the loss of information while ensuring that the data is accurate and reliable.

B. Deep Learning Model for Rainfall Prediction

To predict rainfall, we employed deep-learning techniques to predict rainfall. Specifically, we developed a deep-
learning-based model for rainfall prediction in Jordan, the selected location. The model architecture consists of
multiple layers, each serving a specific function, as outlined below:

Input Layer: The input layer comprises artificial input neurons responsible for receiving preprocessed weather
data, which is then passed on to subsequent layers for further processing [1].

LSTM: The LSTM (Long Short-Term Memory) is a specialized type of Recurrent Neural Network (RNN) that
effectively captures long-term dependencies (Miao et al., 2020). LSTMs were specifically designed to overcome
the challenge of learning long-term dependencies. This characteristic allows them to handle such dependencies more
effectively compared to other models.

By utilizing the input layer and LSTM architecture, our proposed deep-learning model demonstrates the
capability to capture and leverage temporal dependencies in the rainfall prediction process. This approach ensures
that the model can effectively analyze and predict rainfall patterns in the Jordan region.

IV. RESULTS

Data analysis and visualization is a crucial step in understanding and interpreting the information contained
within a dataset. By analyzing and visualizing data, patterns and trends can be identified and used to make informed
predictions. Additionally, data visualization allows for the accessible communication of complex information to a
wide audience, making it an essential tool in fields such as business, science, and research. In short, data analysis
and visualization enable extracting insights and knowledge from data, which can be used in further deep analysis.

Figure 3 illustrates the rainfall between 1987 and 2018. Each bar represents a year, and the bar is divided into
sections, with each section representing a month. The color of each section represents the percentage of rainfall for
that month in that year. From the figure, 2017, 1995, and 1999 had the lowest annual rainfall, while 1993 had the highest annual rainfall, followed by 1994, 2002, 1991, 1988, and 2013. This information can provide insight into the rainfall patterns in the region over the past several decades. For example, the low rainfall in specific years (such as 2017, 1995, and 1999) may indicate a period of drought, while the high rainfall in other years (such as 1993) may indicate a period of heavy rainfall or a higher likelihood of flooding.

![Fig. 3. Rainfall between 1987 and 2018](image)

Additionally, the differences in rainfall between years and between months can indicate the variability and unpredictability of rainfall in the region. Figure 4 gives an idea about the seasonal patterns of rainfall; by looking at the percentage of rainfall each month, we can conclude that some months are rainier than others; this could be used in agriculture and other related fields. Overall, this figure provides valuable information about the rainfall patterns in the region over the past several decades, which can be used to understand the impact of rainfall on various aspects of life in the region and to make informed decisions about water management and other related issues.

![Fig. 4. Rainfall prediction](image)

A. Evaluation and assessment of forecasting models

Researchers have utilized many evaluation methods such as mean absolute error (MAE), mean square error (MSE), correlation coefficient (R), coefficient of determination (R2) and root mean square error (RMSE) [1], [27], [13]. It can be noted that most researchers used RMSE in their studies followed by R2 [27].

1) Root Mean Squared Error (RMSE): Root Mean Squared Error (RMSE) measures how well a model can predict the actual values of a dataset compared to the predicted values. It is commonly used in regression problems. The RMSE is calculated as:

\[
RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2}
\]

Where \(y_i\) are the actual values, \(\hat{y}_i\) are the predicted values, and \(n\) is the number of observations in the dataset. The RMSE is the square root of the average squared differences between the predicted and actual values and measures the average magnitude of the error. The smaller the value of RMSE, the better the model is at predicting the actual values. It is a commonly used metric in machine learning to evaluate the performance of a model. Please note that RMSE is sensitive to outliers and is affected by the scale of the target variable, so it is essential to consider these facts when interpreting the results. Figure 5 shows the predicted vs true values.
B. R-squared:

The coefficient of determination, also known as R-squared, measures how well a model can explain the variation in the target variable. It ranges from 0 to 1, where 1 means the model explains all the variations in the target variable, and 0 means the model does not explain any variations in the target variable [27].

The R-squared is calculated as:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$$

Where $y_i$ are the actual values, $\hat{y}_i$ are the predicted values, $\bar{y}$ is the mean of the actual values, and n is the number of observations in the dataset.

R-squared is a commonly used metric in regression problems [15]; it provides a quick way to understand the goodness of fit of a model. It is easy to interpret, and a value of 1 means that the model fits the data perfectly, while a value of 0 means that it does not explain any of the variances in the target variable. R-squared value does not guarantee a good model; it can be misleading if there are many predictor variables or if the model needs to be more balanced. Figure 6 shows the values of model training.

V. CONCLUSION

Rainfall prediction is crucial, especially for countries that depend on rainfall as a primary source of water. As climate change and other factors continue to impact rainfall patterns globally, the ability to accurately predict rainfall can inform decision-making by government bodies and water resource managers. This study focusses on Jordan that has limited water resources and a heavy reliance on rainfall for drinking water, irrigation, and other uses. The use of artificial intelligence techniques is critical in enhancing rainfall prediction. Recurrent Neural Network (RNN) and Support Vector Machine (SVM) are used to analyze collected data set of rainfall between 1987 and 2017. The data set pre-processed carefully before the analysis to eliminate and replace the missing values. Data visualized in daily, monthly, and yearly means for various meteorological variables such as: the maximum, minimum temperature recorded each day, the mean temperature, and the difference between the maximum and minimum temperature. These variables help in testing the wet bulb temperature, dry bulb temperature, and vapour pressure in the dataset. Where it provides insight into the air’s thermodynamic state and the atmosphere’s water vapour content, and it will assist in examining the temperature variation within a given day during a specific month in a year and give a complete vision on the overall climate conditions during the year. This research is significant as it underscores the importance of integrating advanced technologies into water resource management strategies. The implications of accurate rainfall prediction are far reaching, from better management of water resources to more sustainable agricultural practices. Ultimately, the research discussed in this article can contribute to more effective resource
management, a key consideration in a world grappling with growing populations and increasing environmental pressures.

REFERENCES


