Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 6, 1805-1827
2024
Publisher: Learning Gate
DOI: 10.55214/25768484.v8i6.2347

Monthly rainfall prediction for different climatic zones in south Africa for 2024 using a random forest model

Jeremiah Ayodele Ogunniyi[1], Mohamed A. M. Abd Elbasit[2], Ibidun Christiana Obagbuwa[1*]
[1]Department of Computer Science and Information Technology, School of Natural and Applied Sciences, Sol Plaatje University, South Africa; jerryogunniyi@gmail.com (J.A.O.) ibidun.obagbuwa@spu.ac.za (I.C.O.)
[2]Department of Physical and Earth Sciences, School of Natural and Applied Sciences, Sol Plaatje University, South Africa; mohamed.ahmed@spu.ac.za (M.A.M.A.E.).

Abstract: This study predicted 2024 rainfall in different climatic zones in South Africa using a random forest model. Previous studies have shown that random forests performed better than other models for rainfall prediction. South Africa was divided into nine using the Koppen-Geiger climate classification system, and three cities were selected for each climatic zone. Atmospheric datasets from the South African Weather Service and the National Aeronautics and Space Agency were used for this study. The datasets were trained, tested, and validated to assess the model's accuracy. With good forecast ability observed, the random forest was then used for monthly rainfall for 2024. The result of this prediction was then compared with 2022 and 2023 rainfall. The result indicated months where much rain should be expected in various cities and cities that may likely experience droughts. This result is particularly important for agriculture, water resource management, and early drought/flooding warning systems.

Keywords: *Climate zones, Machine learning, Rainfall prediction, Random forest.*

1. Introduction

The use of machine learning models in rainfall prediction has increased worldwide due to its advantages (Crammer et al., 2017). The atmospheric variables affecting rainfall are complex and nonlinear (Gao et al., 2017). Rainfall prediction requires using variables such as dew point, relative humidity, wind speed, temperature, atmospheric pressure, and water vapour. With an ever-growing database using historical datasets, machine learning models can easily reveal patterns and relationships between the variables (Chattopadhyay et al., 2020). These models can handle the complexities in the datasets better than the numerical method of rainfall prediction. In addition, multiple models can be combined to improve the accuracy of the forecast. With evolving weather conditions due to climate change, machine learning models can be designed to continuously learn and adapt to this change (Rolnick et al., 2022). Since the main aim is to improve forecasting accuracy, machine learning models will enhance prediction's precision and reliability, affecting the economy, agriculture, disasters, and water resource management (Sharma et al., 2020). Most research on weather forecasting, especially rainfall prediction, has been carried out outside Africa. Most weather stations in Africa still rely on numerical and statistical methods for their forecast (Meque et al., 2021; Milton et al., 2017). These methods are time-consuming and expensive (Chen et al., 2022). Recent research in Africa has shown the potential for machine learning applications. Bamisile et al. (2020) used machine and deep learning models for solar radiation comparison. These models were applied to Nigeria's four northern states (Borno, Kano, Yobe, Zamfara) using an hourly time step 12 years datasets. Their result revealed a coefficient of determination value of 0.89 when using the support vector regression and 0.95 when using the recurrent neural network.

* Correspondence: ibidun.obagbuwa@spu.ac.za

Bouras et al. (2021) used four machine learning algorithms to forecast crop yield. Their results showed the effectiveness of machine learning models for cereal yield prediction. Their models showed they can accurately predict crop yield with an R square of 0.88. Cedric et al. (2022) used machine models to predict crop yield production in six West African countries. Their result revealed coefficients of determination of 95.3%, 93.15%, and 89.78% for decision tree, K-Nearest Neighbor, and logistic regression, respectively. Sub-Saharan Africa has also used machine learning algorithms to predict malaria occurrence using climate variability (Nkiruka, 2021). They stated that their research results will aid decision-making and adequate preparation for future outbreaks. Despite the advantages of machine learning models in rainfall prediction, the models depended on the availability and reliability of the datasets used. In some climatic zones, getting datasets for all the atmospheric variables needed for rainfall prediction was difficult. Also, three cities were selected in each climatic zone to examine the best model for rainfall prediction. These cities only represent a fraction of the whole climatic zone and can only indicate the best models. Therefore, though this work suggests models to be used in different climatic zones, some models may perform better in cities that were not examined.

Researchers have shown the possibility of using machine learning models for weather prediction worldwide (Bamisile et al., 2021; Diez-Sierra & del Jesus, 2020; Cramer et al., 2017). They have also shown that it is faster, more accurate, does not need high computational power, and, most importantly, is the future of weather prediction (Cedric et al., 2022; Islam et al., 2023; Jose et al., 2022; Shah et al., 2023). Therefore, this study predicts monthly rainfall for twenty-seven cities in South Africa under nine climatic zones.

2. Literature Review

Due to the computational cost and time-consuming nature of numerical weather models, researchers have explored alternative ways of forecasting weather on a large scale using various machine learning algorithms and have found them accurate. Different machine learning models have been used across the world for various forecasts. Ridwan et al. (2020) used four machine learning methods for rainfall forecasting in Terengganu, Malaysia. Using the autocorrelation function and projected error based on historical rainfall data, they used datasets from 10 stations around Terengganu for the study. Their result revealed that Boosted decision tree regression had the best coefficient of determination compared to decision tree forest, neural network regression, and Bayesian linear regression. He et al. (2021) examined the use of machine learning algorithms for sub-seasonal climate forecasting. They focused on predicting temperature and rainfall on a two-week to two-month scale. They used two non-deep learning (DL) models (AutoKNN, MultiLLR) and five machine learning models (Multitask Lasso, Gradient boosted trees (XGBoost), State of the art baselines, Encoder (LSTM)-Decoder (FNN) and CNN-LSTM models) for their research. Their results showed that machine learning models captured the predictability on sub-seasonal time scales with the ability to outperform the baseline set. However, machine learning model results were better with the best-designed models for deep learning. Due to various catastrophic events in South America attributed to weather and climate, Anochi et al. (2021) deployed ML algorithms for weather forecasting modelling in South America. They observed that numerical methods for precipitation prediction could not accurately show precipitation patterns due to the absence of datasets specific to the regions. Their work showed the possibility of using machine learning algorithms for accurate precipitation prediction. They trained their models using 36-year datasets from 1980 to predict rainfall in 2018 and 2019. However, they observed significant errors in the summer months during the rainy season. This was attributed to local processes in the region that the algorithms could not learn and high energy during that period. They then showed that training the networks with Tensorflow for all seasons except spring will make them perform better than those trained with neural networks. Diez-Sierra and del Jesus (2020) investigated how eight statistical and machine learning models performed while predicting daily precipitation in a semi-arid region. They used 36-year rainfall data divided into training and test datasets using the 80-20 principle. The methods used include the random forest, generalized linear model, linear regression, support vector machine, k

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 6: 1805-1827, 2024
DOI: 10.55214/25768484.v8i6.2347

nearest neighbor, and neural network. They stated two advantages of machine learning methods over others: they do not need a known priori, nor do they need to make assumptions on the distribution errors. Their result showed that the hyperparameter chosen affects the machine learning model. They further stated that if wrong parameters are used, it will affect the predictive capability and overfitting of the training models. They suggested that the hyperparameters should take higher values to avoid overfitting. The summary of their results showed that NN performed better than other models while predicting rainfall intensity. SVM, K-NN, and RF closely followed this, while the worst of their models was WT.

As a result of the successes of ML approaches in weather forecasting in South America, Monego et al. (2022) also applied these models to precipitation prediction using gradient boosting (GB). They used the extreme gradient boosting (XGB) and TensorFlow (TF) models to train datasets from January 1980 to February 2020 using the 75-25 principle. They considered meteorological variables such as air temperature at the surface and 850 hPa, surface pressure, specific humidity at 850 hPa, zonal wind component at 500 hPa and 850 hPa, meridional wind component at 850 hPa, as well as rainfall. Their result showed excellent performance concerning pattern recognition for precipitation. They also showed that XGB performed better than TensorFlow deep neural network for all seasons except autumn, while TensorFlow performed better than XGB only in autumn. Baran et al. (2020) compared machine learning models with parametric classification techniques using datasets from 2002 to 2014. Their result showed that multilayer perceptron performed best when average rainfall was used as an additional covariate. They also used these models to predict cloud clover. They stated systematic errors in calibration when several probabilistic classification methods are used. For their work, POLR performed best for two days; however, for long-term forecasts, MLP performed best. They stated that including atmospheric parameters such as pressure, humidity, and temperature can improve the predictive performance of any model. Bamisile et al. (2020) compared the results of global and diffuse solar radiation using machine and deep learning models. For deep learning models, they used the artificial neural network (ANN), convolutional neural network (CNN), and recurrent neural network (RNN).

In contrast, they selected the SVM, polynomial regression (PR), and RF for machine learning algorithms. These models were applied to Nigeria's four northern states (Borno, Kano, Yobe, Zamfara) using an hourly time step 12 years datasets. They stated that the time spent training the machine learning models was reduced compared to the time spent training deep learning models; however, from their results, the deep learning models performed better. Appiah-Badu et al. (2021) predicted precipitation using five ML models (Decision Tree, K-Nearest Neighbour, Multilayer Perceptron, Extreme Gradient Boosting, and Random Forest) in Ghana using 41-year climatic datasets from 2018. They divided the datasets into training and test sets using three different ratios (70:30, 80:20, 90:10) to assess the performance of the models. They analysed four climatic zones in Ghana: Coastal zone, Forest zone, Transitional zone, and Savanah zone. Their result showed that over the coastal zone of Ghana, MLP performed best using the 90:10 principle, while XGB performed best using 80:20. For the three different ratios in the coastal zone of Ghana, KNN performed worst.

Similarly, for the forest zone, at 90:10, MLP performed better than other models, with KNN performing worse. Using 70:30, DT, RF, and XGB, all had a better precision, recall, and f1 score. Over the transitional zones, all three ratios of RF and XGB performed best in the metrics used. The result also showed that MLP performed best using the 90:10 ratio, perhaps the best ratio when considering multilayer perceptron. Finally, over the Savannah zone, RF and XGB performed best in 90:10, 80:20, and 70:30, both with and without rain. The study did not investigate why KNN performed worst in all zones. However, their study showed that machine learning models are suitable for rainfall prediction, especially RF, XGB, and MLP.

Machine learning algorithms have been applied to predict rainfall in various parts of the world. Jose et al. (2022) used different machine learning models to predict daily rain in India. Their result showed that long short-term memory performed best with a coefficient of determination of 0.9. In Australia, noted for the highest extreme temperature in the world, Polishchuk et al. (2021) used a random forest model to predict rainfall to know when there will be wildfire. Their results revealed an accuracy of 85.9% in their training model and 84.7% accuracy in the rainfall prediction.

Similarly, Raval et al. (2021) used different machine learning models to predict rainfall using 10-year datasets. Their result revealed that the logistic regression model had the highest classification with an f1 and precision of 86.9% and 97.1%, respectively. Other researchers, such as He et al. (2022), Islam et al. (2023), and Sachindra et al. (2018), among others, have all used machine learning models to predict rainfall in Australia with an accuracy above 80%. In South America, Ferreira and Reboita (2022) showed that applying machine learning models in rainfall prediction has led to a 75% error reduction in rainfall estimates. Their result also indicated the ability of their models to reproduce both the spatial and temporal variation for dry and wet seasons across South America. Despite high topographic gradients and unstable climatic conditions in South America, researchers have shown the possibility of deploying machine learning models for rainfall prediction with high accuracy (Anochi et al., 2021; Gómez et al., 2023).

Yu et al. (2017) compared random forests and support vector machine models for real-time radar-derived rainfall forecasting and observed that random forests performed better in one-hour-ahead forecasting. Similarly, Zainudin et al. (2016) conducted a comparative analysis of data mining techniques for Malaysia rainfall prediction, and their results showed that random forest performed much better than other models used for their study. For rainfall prediction across various ecological zones of Ghana, random forest, extreme gradient boosting, and multilayer perceptron performed well (Appiah-Badu et al., 2021). For the improvement of rainfall estimation in northern Algeria using Meteosat second-generation datasets, Ouallouche et al. (2018) showed that random forests performed better statistically compared to an artificial neural network and support vector machine. In India, Shah et al. (2023) used a random forest-based model for their rainfall prediction and observed an accuracy rate of 90% and a probability of detection of 68%. Using a complete ensemble empirical mode decomposition hybridized with a random forest and kernel regression model for monthly rainfall prediction, Ali et al. (2020) performed best compared to other models.

The results from the reviewed literature have shown the accuracy of machine learning models and their better performance compared with numerical weather forecasts. They have also indicated the various machine learning models for various climates and the possibility for improvements in their forecast. However, It is important to note that some machine learning models will be more appropriate for different regions. These results have also established the accuracy of random forests for rainfall prediction worldwide. Therefore, This work uses a random forest model to predict 2024 monthly rainfall.

3. Methodology

We built, trained, and tested a random forest model for weather prediction based on the datasets that we have available. Temperature, rainfall, wind speed, evaporation, relative humidity, dew point, and water vapour datasets over twenty-seven cities in nine climatic zones in South Africa were collected and combined to form a single record. Rainfall was assigned as the target variable, though, in the same way, other weather parameters can be determined. For this work, we restrict it to predicting only rainfall. The models were trained using the 80/20 principle for all the models. The National Aeronautics and Space Administration NASA used MERRA/OMI and the South African Weather Service to obtain data for this work. This research's data include evaporation, temperature, rainfall, wind speed, irradiance, and water vapour. After collecting the datasets, they were cleaned by developing a model that identified and excluded missing data, duplicates, and irrelevant datasets. Clean datasets were retrieved and converted to monthly datasets for this study. The random forest model was trained for rainfall prediction in South Africa using the 70:30 principle and evaluated, after which 2024 monthly rainfall prediction was made and compared with 2022 and 2023 rainfall using random forest. This process is represented in the flow chart in Figure 1.

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 6: 1805-1827, 2024
DOI: 10.55214/25768484.v8i6.2347

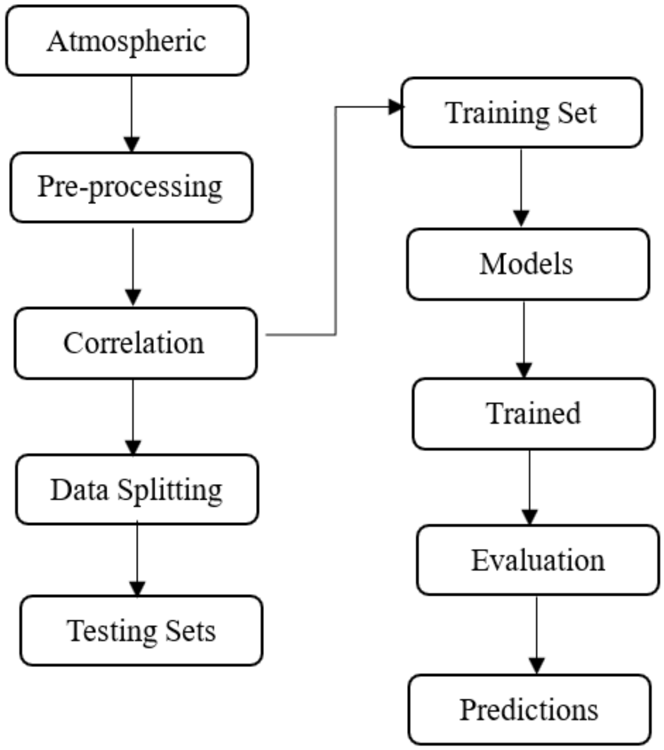


Figure 1.
Schematic view of the data analysis process.

3.1. Study Area

The study location is based on the Koppen-Gieger climate classification of South Africa. South Africa is broadly classified into Arid, subtropical wet (fully humid and dry winter), and subtropical dry (hot summer). This study also considered the sub-group climate classification, such as the cold and semi-arid steppe, cold arid desert, hoot and semi-arid steppe, and the hot arid desert under the arid classification. The subtropical highland with a dry winter and the humid subtropical with a dry winter are considered under the subtropical wet. The sub-divisions are subtropical dry, warm and dry summer, temperate oceanic without dry season, and humid subtropical without dry season. For each sub-division, three locations are randomly selected for rainfall prediction. Figure 1 shows the Koppen-Geiger climate classification for South Africa and the study locations.

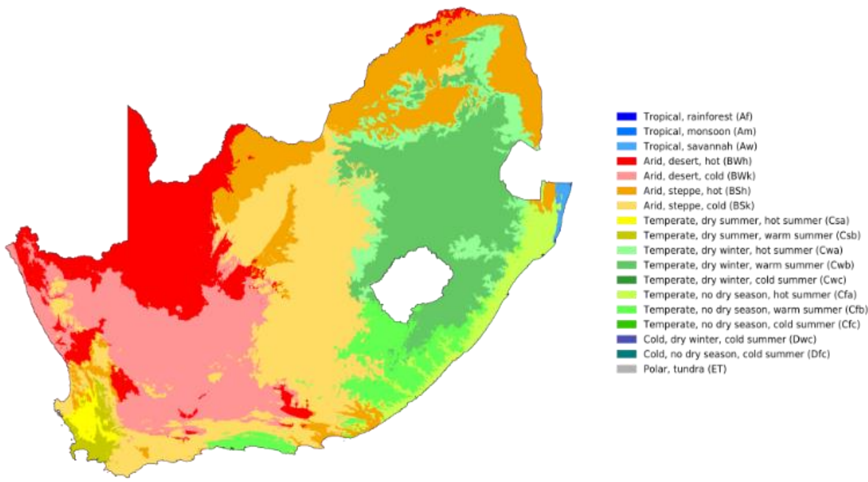


Figure 2.
Koppen-Geiger South Africa weather classification showing the various climates.

4. Result and Discussion

4.1. Cold and Semi-Arid Steppe

This section presents monthly rainfall prediction for different climatic zones in South Africa for 2024 using a random forest model. The prediction was compared to the monthly rainfall of 2022 and 2023, as shown in Figure 3. Under the cold and semi-arid steppe in Bloemfontein, figure 3 shows that more rainfall of over 100mm will be received in the first four months of 2024. A similar amount of rainfall received in January, March, May, and October is expected in Springfontein in 2024. Compared to the previous years, more rainfall is expected in February and winter. While there is an increase in the predicted amount of rainfall in November in 2023, the amount of rainfall in December will be lower than what was experienced in 2023 and about half of the 2022 rainfall. In Springfontein, it is predicted that there will be more rainfall in 2024 than in 2023 throughout the years, and except for March and June, it is also expected that more rainfall will be received than in 2024. This is good news for farmers as they can plan appropriately, as rainfall above 80mm is expected in the last three months of 2024, while rainfall of about 40mm is expected monthly during winter. In Welkom, Figure 3 shows a similar pattern of rainfall to 2023 in 2024, with 2024 expected to receive more rainfall. The amount of rainfall expected from February to April is similar to that Welkom received in 2022. However, more rainfall is expected in the autumn and early spring months (June to September) compared to previous years. Although the rainfall expected in October will be more than that of 2023, the rainfall expected in November and December is similar to that of 2023 and much lower than in 2022 during summer. For the three cities under this climatic zone, more rainfall is expected during autumn in 2024 compared to other years. While Springfontein predicted that there would be more rainfall during the spring and summer of 2024, similar amounts of rainfall are expected in Bloemfontein and Welkom compared to 2023.

ISSN: 2576-8484
Vol. 8, No. 6: 1805-1827, 2024
DOI: 10.55214/25768484.v8i6.2347

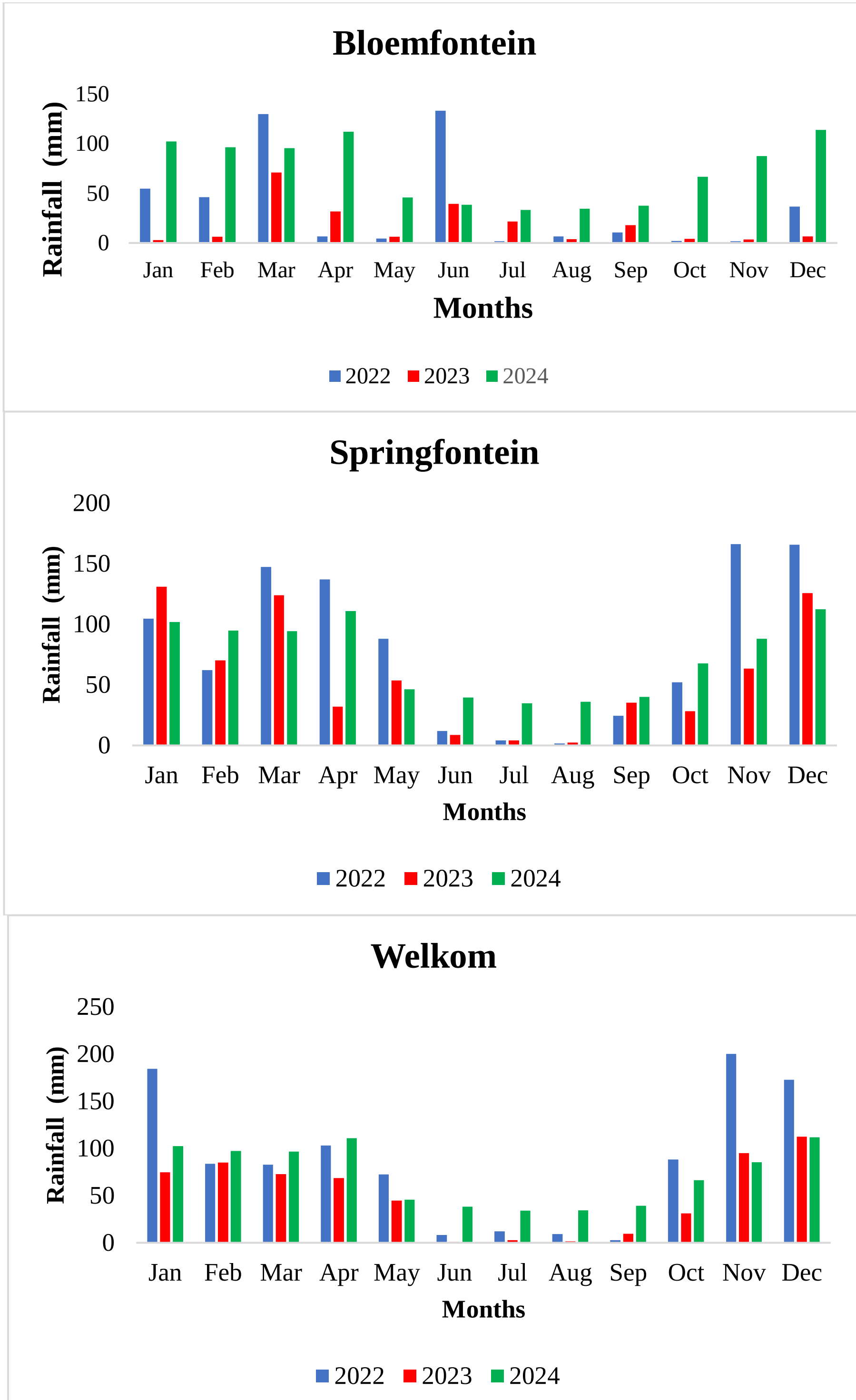
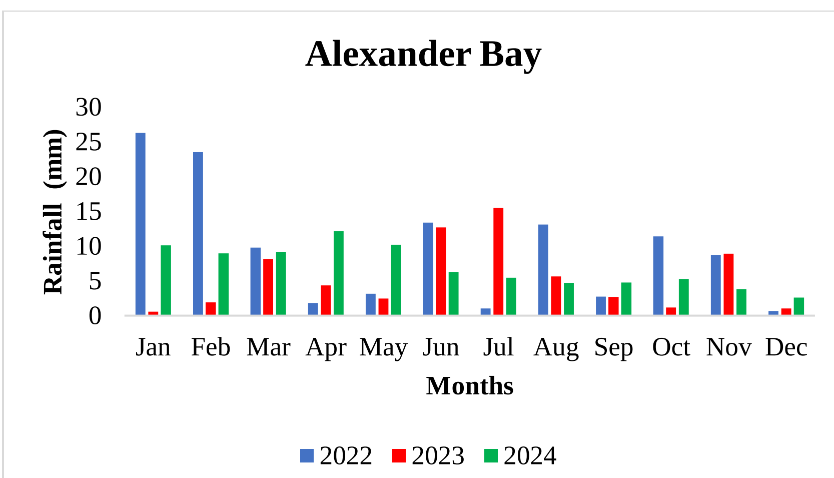


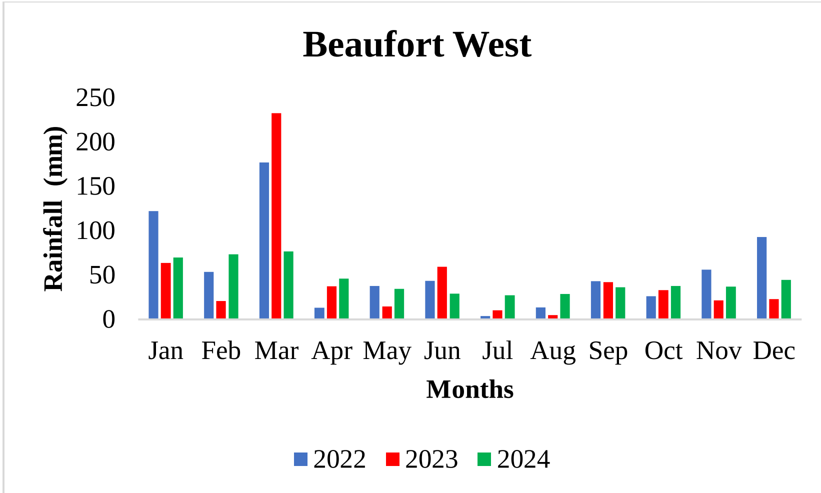
Figure 3.
Rainfall prediction for 2024 compared with 2023 and 2023 rainfall under the cold and semi-arid steppe (Bloemfontein, Springfontein, and Welkom).

ISSN: 2576-8484
Vol. 8, No. 6: 1805-1827, 2024
DOI: 10.55214/25768484.v8i6.2347

4.2. Cold Arid Desert

Figure 4 shows rainfall predictions for Alexander Bay, Beaufort West, and Bristown in the cold, arid desert. The amount of rainfall expected in Alexander Bay is more in 2024 compared to 2023 for the first two months as there was little or no rainfall those months in 2023; however, compared to 2022, rainfall expected in January and February is less than half of what was received in 2022. A similar amount of rainfall is expected in March as in 2022 and 2023. In April and May, more rainfall is expected in 2024, more than double what was received in previous years. While rainfall in Alexander Bay was about 2mm and 4mm for April and May of 2022, in 2023, it was about 4mm and 3mm for April and May, respectively. It is expected that rainfall for those months in 2024 will be about 13mm and 11mm, indicating that more rainfall is expected in the winter of 2024 compared to other years. However, this amount of rainfall is still significantly small as it may all be showers. In autumn, 2023 received more rainfall than expected in autumn 2024, with the amount of rainfall expected to reduce by about half compared to 2023. There is a slight increase in rainfall expected in September, October, and December compared to 2023, though this increase is negligible. Therefore, 2024 is expected to be another dry year in Alexander Bay. In Beaufort West, 2024 is expected to be another dry. Only in the first three months is rainfall above 50mm expected, but all under 80mm. The amount of rainfall received in January 2022 was reduced by almost half compared to January 2023. This amount was maintained in 2023; however, it is predicted that February 2024 rainfall will be more than that of 2022 and 2023. The amount of rainfall expected in February is similar to that expected in March in Beaufort West. This is less than half of what was received in March compared to other years. Although there was a significant decrease from March to April by over 150mm in 2022 and 2023, the reduction in 2024 was about 30mm. The rainfall expected in late winter, early autumn, and spring is similar to 2024. More rainfall is expected in the last three months of the year compared to 2023. In Bristown, a similar rainfall pattern is seen across the months, with higher rainfall in the early parts of the year, reduced during winter and autumn, and increases from late spring to early summer. It is predicted that over 100mm monthly rainfall will be received in 2024 from January to April, an increase from 2023 with no month receiving up to 75mm except in December. Compared with 2022, with high rainfall received, only in January, March, May, September, and December did 2022 receive more rainfall than predicted for 2024. The total amount of rainfall expected in 2024 is similar to the amount received in 2022, almost twice the amount received in 2023. That means that 2024 rainfall in Bristown is expected to double the rainfall received in 2023. Also, in 2024, more rainfall is expected during later winter and early autumn compared to other years. January, April, and December are expected to receive the highest rainfall in 2024 in Bristown.





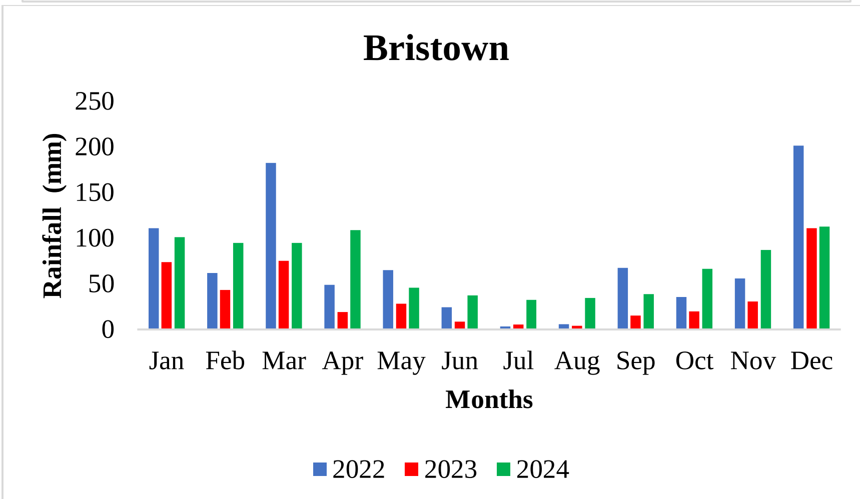


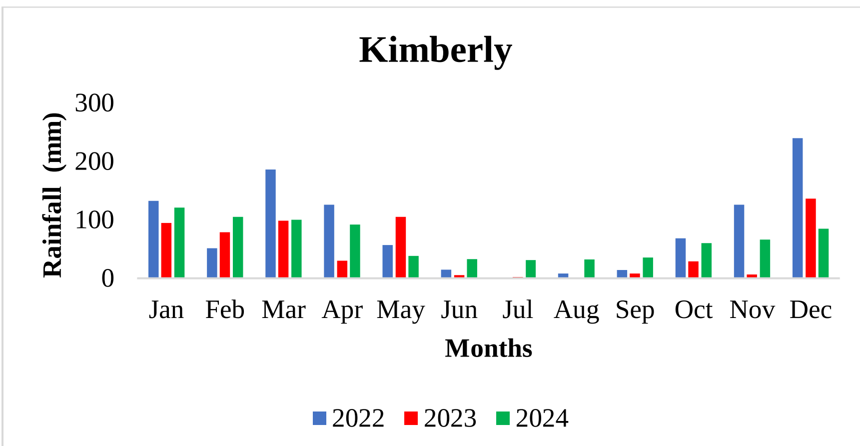
Figure 4.
Rainfall prediction for 2024 compared with 2023 and 2023 rainfall under the cold arid desert (Alexander Bay, Beaufort West, and Bristown).

4.3. Hot and Semi-Arid Steppe

Figure 5 shows the rainfall prediction for 2024 compared to 2022 and 2023 rainfall for Kimberly, Mahikeng, and Port Elizabeth under the hot and semi-arid steppe climate classification. In Kimberly, the figure shows that it is expected to receive about 20mm more rainfall in January than in January 2023 and 10mm less in January 2022. However, more rainfall is expected in February. There was an increase in rainfall by about 25mm in February 2022 and 2023; this estimated increase is predicted for February 2023. In March, a similar amount of rainfall compared to 2023 is expected in Kimberly in 2024, with a slight reduction in April. However, the estimated rainfall for April 2024 will be about three times what was experienced in April 2023 and about 30mm lower than in April 2022. Compared to May 2022 and May 2023, lower rainfall is expected in Kimberly in 2024. From autumn (June) to early spring (September), more rainfall is expected in 2024, although monthly estimates are below 35mm. It is predicted that there will be a gradual increase in rainfall expected till the end of the year. The predicted values are higher than Kimberly's rainfall in 2023 except in December but lower than 2022 rainfall. It is also predicted that there will be an annual increase of about 200mm in 2024 rainfall compared to 2023; however, this is still about 250mm lower than 2022 rainfall. In 2022, Mahikeng received rainfall of over 150mm in January, April, November, and December; in 2023, in February, the amount of rainfall

Mahikeng received is estimated to be around 270mm. However, in 2024, only January, April, and December are the predicted monthly estimates above 100mm. There is an increase of about 60mm in January rainfall between 2023 and the predicted 2024 values, but there is almost a 200mm difference in February. The amount of rainfall predicted for March, though lower in 2024, is quite similar to that of 2023. In April, Mahikeng received about 200mm of rainfall in 2022, about 70mm in 2023, and a predicted estimate of over 110mm in 2024. Similar values in May are seen in both 2023 and 2024 estimates. Like Kimberly, there is expected to be more rainfall in autumn and early spring compared to other moths and then a gradual increase in 2024 rainfall till the end of the year. It is also predicted that the rainfall in November and December will be similar to 2023 rainfall in those months. The predicted annual rainfall in 2024 is within the same range as that of 2023 but much lower than the rainfall received in 2022 in Mahikeng.

In Port Elizabeth, more rainfall is predicted for 2024 for the middle to late spring and summer months compared to 2023. Among all locations, this is the first city that has predicted more rainfall in 2024 than what was received both in 2022 and 2023. In January 2022, Port Elizabeth received about 20mm of rainfall, which increased to 70mm in January 2023; January 2024 estimates are predicted to be over 100mm. Similarly, more rainfall is expected for February compared to 2022 and 2023. However, lesser rainfall is predicted for March compared with the previous two years. The rainfall expected in April is more than twice what was experienced in 2022 and 2023. From May to September, except for August, it is predicted that lesser rainfall will be experienced compared to 2023. For October and November, rainfall estimates for 2024 will be at least thrice what was received in 2022 and 2023, while that of December will be similar to the amount of rainfall received in December 2022.



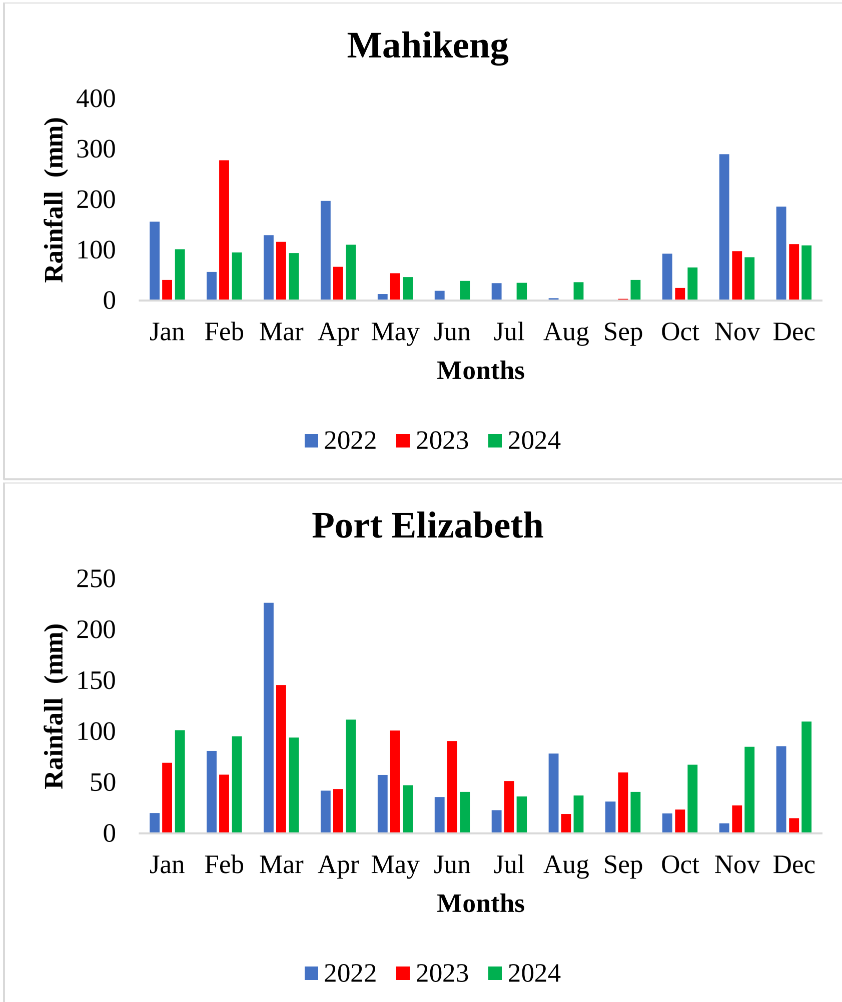


Figure 5.
Rainfall prediction for 2024 compared with 2023 and 2023 rainfall under the hot and semi-arid desert (Kimberly, Mahikeng, and Port Elizabeth).

4.4. Hot, Arid Desert

In Lauville, a predicted decrease of about 450mm in rainfall is expected in 2024 compared to 2022 and 2023. Therefore, with monthly rainfall from January to April 2024 around 100mm, it is suggested that proper planning be made, especially for those involved in farming, to mitigate the water shortage. Although the graph shows a similar amount of rainfall as shown in Figure 6 for 2023 and 2024, the most significant change can be seen in the last three months of the year, where expected rainfall in 2024 is less than half the amount received in 2023.

Contrary to Lauville, the amount of rainfall estimated for 2024 in Musina is more than the amount received in 2023, while the predicted rainfall for 2024 in Upington is more than twice the amount received in 2023. In 2023, Musina received significant rain in January, February, and December, with February and December rainfall of over 200mm. While Musina has only a few months of rainfall above 100mm as predicted, the amount of rainfall expected is well spread compared to other months. While there was almost a drought between May and September of 2022 and 2023, monthly rainfall of about 40mm is expected in 2024, which will gradually increase to over 110mm by the end of the year. In Upington, as shown in Figure 6, only in January, February, March, and December did it receive rainfall above 40mm in 2023. 2023 rainfall in other months was insignificant. Though there was more rainfall in 2022 compared to 2023, only in March did the 2022 rainfall exceed the predicted rainfall for 2024. A similar amount of rainfall in January, February, May, June, and December 2022 is predicted in 2024.

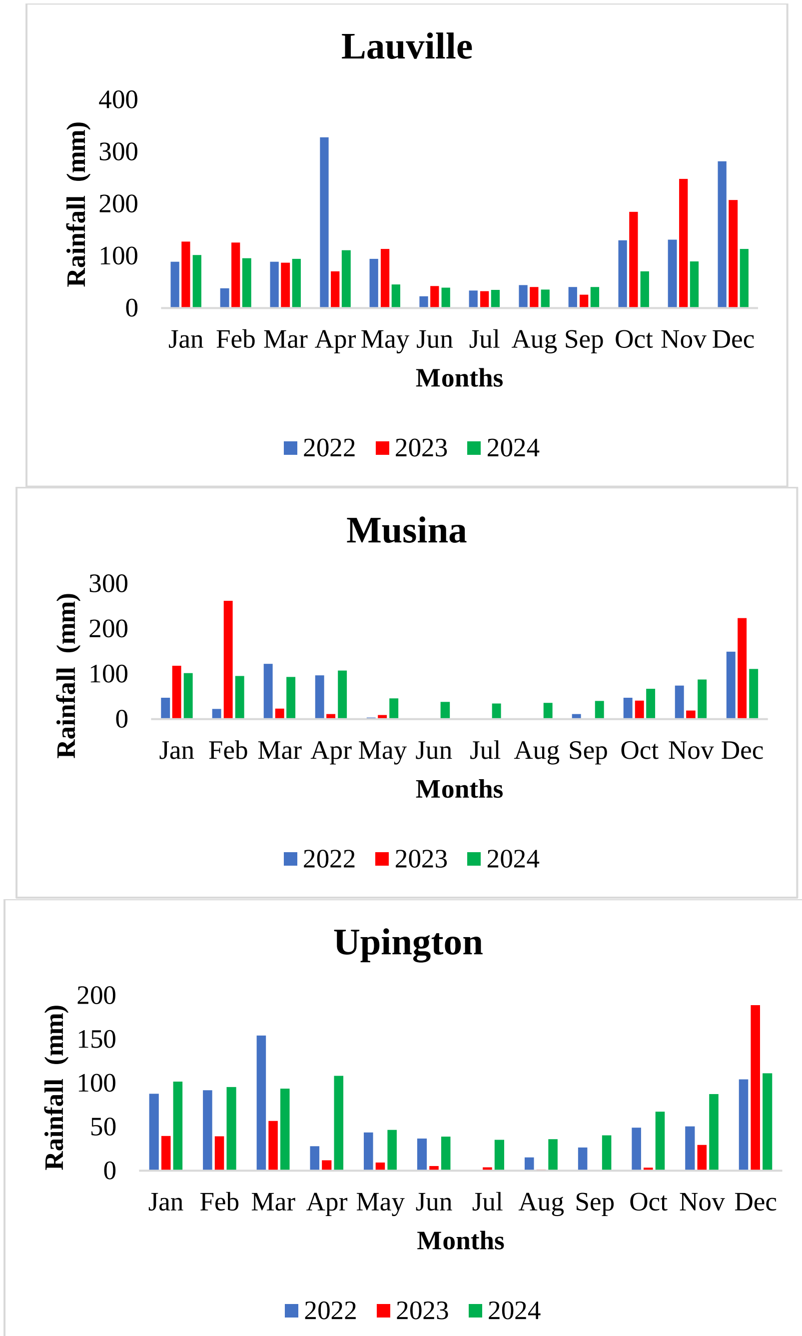


Figure 6.
Rainfall prediction for 2024 compared with 2023 and 2023 rainfall under the hot arid desert (Lauville, Musina, and Upington).

4.5. Subtropical Wet

4.5.1. Humid Subtropical with Dry Winter

Figure 7 shows the predicted rainfall for 2024 compared with rainfall received in 2022 and 2023 in Dundee, Louis Trichardt, and Nelspruit. In Dundee, there is an observed decrease in rainfall received in January 2022 and January 2023 by about 20mm and a further decrease by 20mm in the estimated values of January 2024. However, in February, the rainfall predicted for 2024 is similar to the rainfall received in 2022 and about 80mm less than the February rainfall of 2023. More rainfall is predicted in March 2024 compared to previous years. It is also predicted that the rainfall expected in April will be thrice the amount received in April 2023 but less than that of 2022. However, in May, the amount of rainfall received in the past two years and predicted rainfall for 2024 are similar. Between June and September, it is predicted that there will be more rainfall than the same period combined in 2022 and 2023. From mid-spring to December, there is predicted to be less rainfall in 2024 compared to previous years. The amount of rain expected in November is less than half what was received in the past two years, and the amount expected in December will be about 60mm lower than in 2023 and 100mm lower than the amount of rainfall expected in December 2022. Summarily, it is predicted that there will be lower rainfall in 2024 compared to 2022 and 2023 despite having more rainfall in the autumn of 2024.

In Louis Trichardt, it is predicted that the annual rainfall of 2024 will be more than those of 2022 and 2023. In 2023, significant months of rainfall were January, February, and December, all receiving rainfall above 100mm, while March, October, and November received about 40mm of rainfall. In 2022, a significant amount of rainfall was received in March, April, November, and December, while in 2023, it is predicted that the first four and last three months will experience significant rainfall. It is also predicted that there will not be any dry months in 2024 compared to other years. Predicted rainfall for winter months would be about 40mm. However, despite 2024 having more rainfall than the previous two years, the predicted monthly rainfall would still be lower than the months of heavy rainfall in 2022 and 2023. For instance, in February 2023, about 250mm of rainfall was received, while the predicted rainfall for 2024 would be about 100mm. However, the amount of rainfall received in January 2023 and predicted January 2024 are similar. The rainfall expected In March, April, October, and November are also similar to the amount received in 2022 during these months.

The amount of rainfall expected in 2024 is similar to the annual rainfall in 2022 in Nelspruit and lower than the 2023 rainfall. In January 2024, it is predicted that there will be more rain compared to January 2022 and 2023 by about 30mm. 2023 had months of heavy rain, like February, November, and December, where rainfall exceeded 200mm and even 250mm in November. Also, in 2022, rainfall in November and December was above 150mm; however, in 2024, months with heavy rain would receive only about 100mm of rainfall. These months are January, April, and November, although a significant amount of rainfall is predicted for February, March, and November. It is also predicted that there will be more rainfall in 2024 winter and early spring compared to 2022 and 2023.

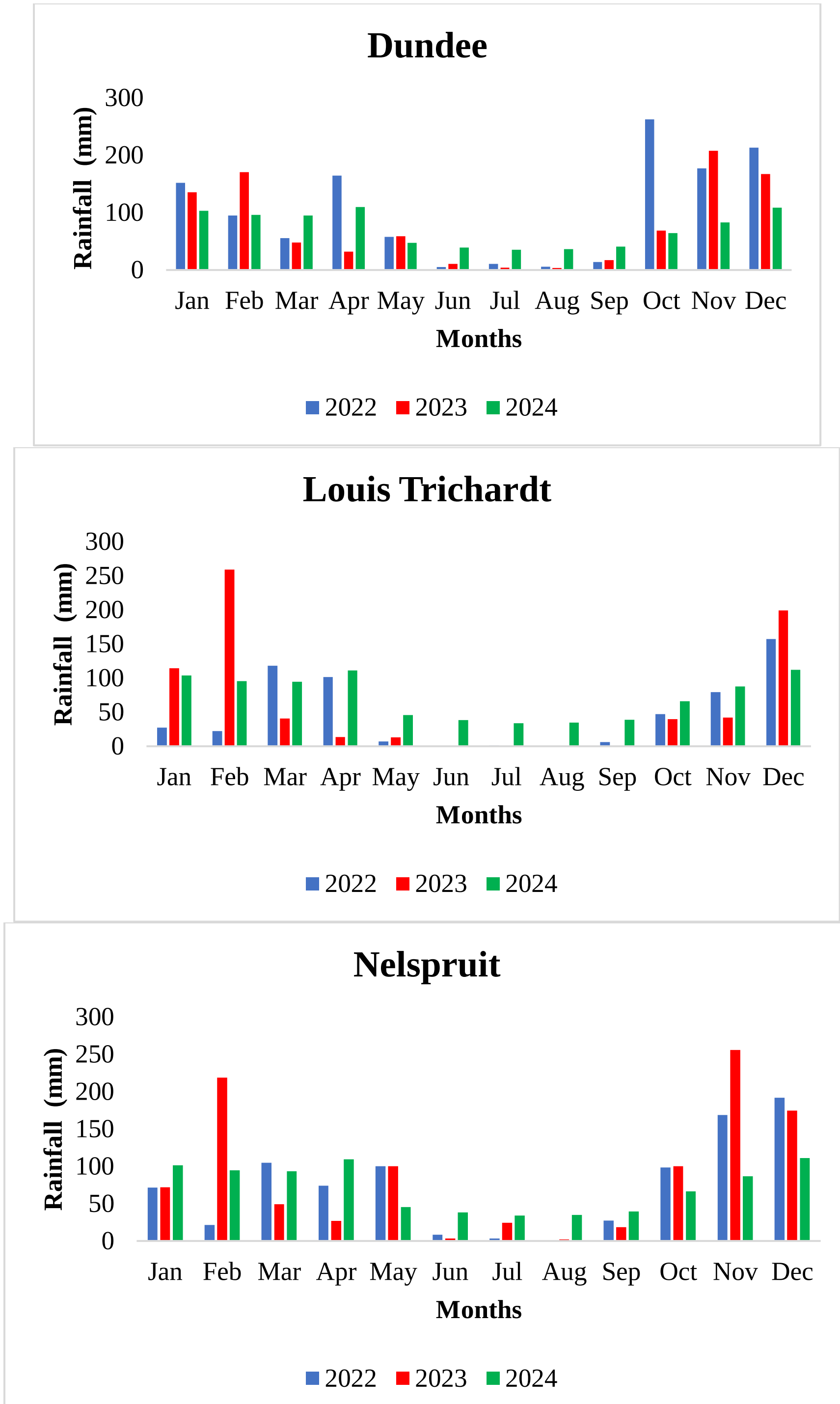


Figure 7.
Rainfall prediction for 2024 compared with 2023 and 2023 rainfall under the humid subtropical with dry winter (Dundee, Louis Trichardt, and Nelspruit).

4.5.2. Subtropical Highland with Dry Winter

Figure 8 shows the predicted rainfall for 2024 for the subtropical highland with a dry winter and compares it with 2022 and 2023 results for Harrismith, Johannesburg, and Newcastle. The result shows that there was more rainfall in 2022 than in 2023, which is what is predicted for 2024, although 2024 estimates are higher than 2023 for Harrismith. In 2022, there were five months with monthly rainfall above 150mm (January, April, October, November, and December), with February and March also experiencing high rainfall. Compared to 2023, rainfall above 100mm was only received in January, November, and December. While only January, April, and December of 2024 have rainfall above 100mm, other months are predicted to have an equally high amount of rain. Between January 2022 and January 2023, there is a decrease of about 30mm in rain received; a further reduction of about 20mm is predicted for January 2024. The predicted amount of rain for February and March is within 20mm of what was received in 2022 and 2023, with 2024 predicted to have more rain. It is also predicted that there will be more rain in April of 2024 by about 50mm compared to what was received in April 2023, but it will be lower than the amount of rainfall received in 2022 in Harrismith. In May, a decrease of about 10mm is predicted compared to 2023 rainfall. However, more rainfall is predicted for 2024 from June to September. Though this climatic zone has a dry winter, there is predicted to be more rain during winter than in previous years. It is also predicted that there will be a gradual increase in 2024 rainfall from September till the end of the year; however, it would still be lower than the 2022 and 2023 rainfall.

In the commercial city of Johannesburg, the total amount of rainfall predicted for 2024 would be a little lower than the annual rainfall of 2023 and much lower than the annual rainfall of 2022. Like Harrismith, Johannesburg had five months of rain above 150mm in 2022 (January, February, April, November, and December), while there were four months above 100mm in 2023 (February, May, November, and December). This reduced the predicted rainfall for 2024 (January, April, and December) to three months. This reveals that more rain should be expected with varying intensity in the early and late months of the year. Compared to 2024, in the first five months of the year, only in January and April will the predicted rainfall be more than that of 2023, while compared with 2022, 2024 predicted rainfall will only be higher in March and May. A monthly average of 100mm of rainfall is predicted for the first four months of 2024. It is also predicted that there will be more rain in winter and early spring compared to other years. The predicted rainfall for the last two months of 2024 is much lower than that of 2023 and 2022.

Newcastle also had four months (April, October, November, and December) of rainfall above 150mm, with rainfall in two months exceeding 200mm (October and December) in 2022. In 2023, rainfall over 150mm was recorded only in February and November, with December having a high amount of rainfall. In 2024, none of the months is predicted to have rainfall of over 150mm, with only three months (January, April, and December) having rainfall above 100mm. Predicted rain for January 2024 is similar to what was received in 2022 and 2023, while that of February is similar to 2022 February rainfall and much lower than 2023 February rainfall. For March, the predicted rainfall was about twice what was received in the previous two years.

Similarly, for April, the predicted rainfall is more than twice what was received in 2023, though lower than the 2022 April rainfall, while the predicted rainfall for May is within the same range as recorded in 2022 and 2023. As with other cities under the dry winter climatic zone, more rainfall is expected in 2024 compared to other years. Rainfall predicted for October is similar to 2023 October rainfall but much lower than that of 2023. However, less rainfall is predicted for the last two months of the year compared to these months in 2022 and 2023, with November receiving about half the amount of rainfall recorded in 2022 and 2023.

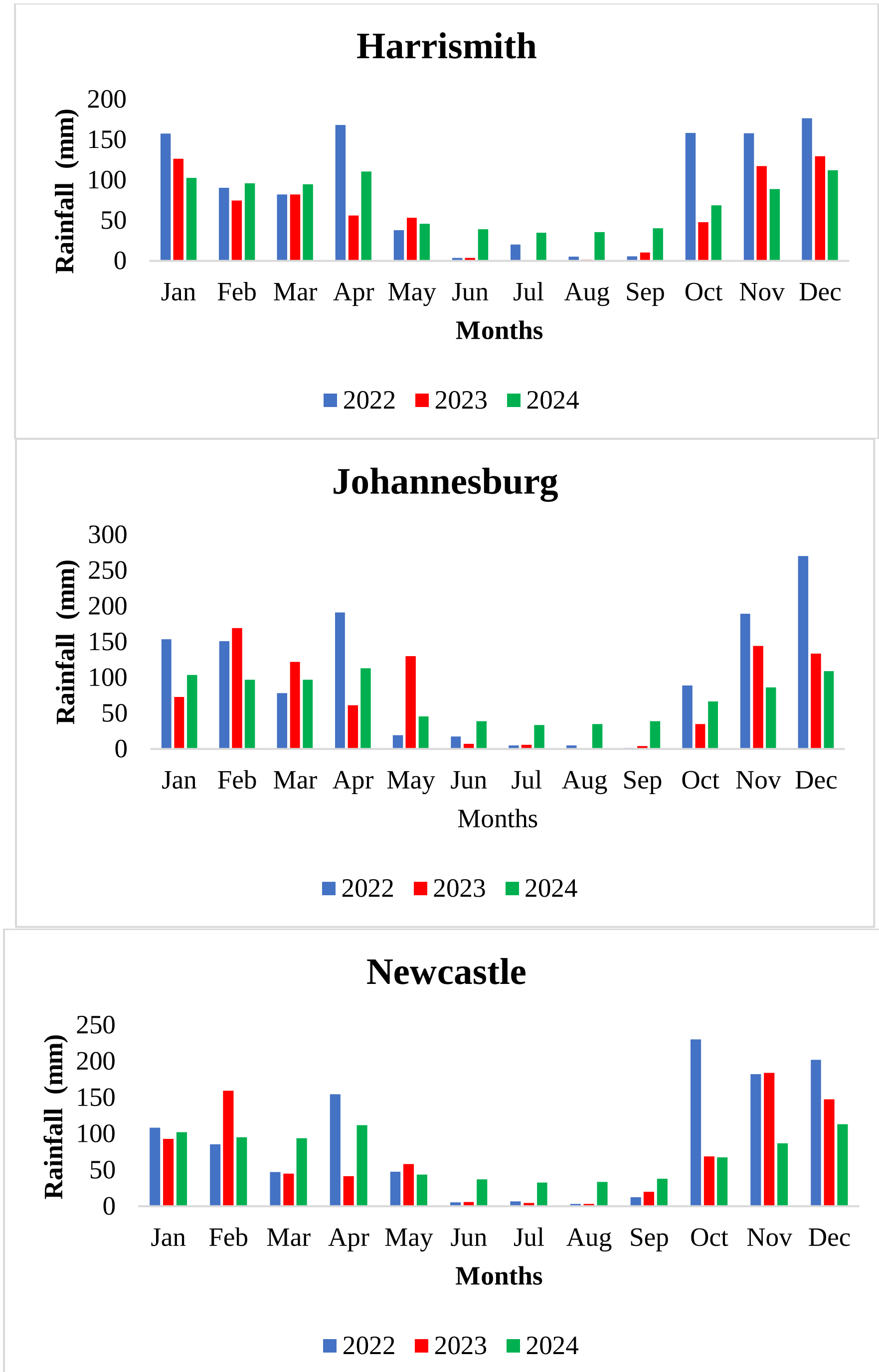


Figure 8.
Rainfall prediction for 2024 compared with 2023 and 2023 rainfall under the subtropical highland with dry winter (Harrismith, Johannesburg, and Newcastle).

4.6. Subtropical Dry

4.6.1. Humid Subtropical Without Dry Season

Figure 9 shows the 2024 rainfall prediction with 2022 and 2023 rainfall for Durban, Port Edward, and Richards Bay under the humid subtropical without a dry season. These locations were the primary motivation for this study due to the flooding in April 2022. Parts of KwaZulu Natal province, including Durban, Richards Bay, and Port Edward in the Eastern Cape, received rainfall above the expected average, leading to the loss of lives and destruction of properties and infrastructures. As seen in the three figures in Figure 9, there is no dry season throughout the year as rain is expected to fall with varying degrees of intensity. In Durban, predicted rainfall in January and February is about 20mm lower than 2023 rainfall but higher than 2022 in those months. However, in March, both the predicted rainfall and previously recorded rainfalls of 2022 and 2023 fall within the same range. As mentioned earlier in the study, April 2022 received an unusual amount of rainfall. Durban received almost 350mm of rain, most happening on the 11th and 12th of April. This is responsible for the high rainfall profile in April as seen in the graph. December 2022 also received high rainfall, although it was more spread across the days than in April. In April 2023, rainfall received in Durban drastically reduced to about 70mm; it is predicted that this will increase in 2024 to over 110mm but will experience a decrease to about 50mm in May. The predicted rainfall for winter and early spring compared with those of 2022 and 2023 are within the same range. Therefore, the same amount of rainfall experienced and recorded in Durban in 2022 and 2023, as well as winter and early spring, should be expected in 2024. However, from October to December 2024, rainfall was predicted to be less than half what was recorded in 2023. This means less rainfall should be expected towards the end of the year.

In Port Edward, 2024, rainfall predicted is within the same range as rainfall recorded in 2022 January and February. This was lower than the 2023 January rainfall and more than the 2023 February rainfall. The predicted rainfall is about 40mm lower in March than in 2022 and 2023. As with other cities under this climatic condition, 2022 April rainfall was about 300mm compared to 60mm recorded in 2023. It is predicted that there will be an increase of about 50mm in April rainfall from 2023. May predicted rainfall is about 20mm lower than in 2022 and 2023. Like Durban, predicted 2024 rainfall and recorded rainfall for 2022 and 2023 during winter months are similar. Although rain would still be expected to fall, the intensity would be low. High rainfall was still recorded in the spring of 2022 and December, which is also predicted for 2024, but at a lower amount. The amount of rain expected in September and October is less than a third of what was recorded during these months in 2022, and less than half of what was recorded in 2023 November is predicted for April 2024. However, a similar amount of rainfall should be expected in December 2024 compared to December 2023.

In Richards Bay, much rainfall was received in January, April, October, November, and December of 2022, with four months (April, October, November, and December) receiving over 200mm of rain in 2023, five months (February, May, October, November, and December) received rainfall exceeding 150mm. However, in 2024, it is predicted that only in January, April, and December will monthly rainfall exceed 100mm. This is reflected in the expected annual rainfall, which is over 300mm lower than the annual rainfall of 2023 and over 500mm lower than the average rainfall of 2022. Despite 2024 being predicted to have lower amounts of rainfall than the previous years, a significant amount of rain is expected in the first four months of 2024, with March rainfall exceeding those of March 2022 and March 2023. It is also predicted that there will be more rain in April compared to April 2023. The figure also shows that there would be rain during winter months, which will gradually increase till the end of the year. Rainfall for October, November, and December is predicted to be at least 80mm lower than in 2023 and 130mm lower than in 2022.

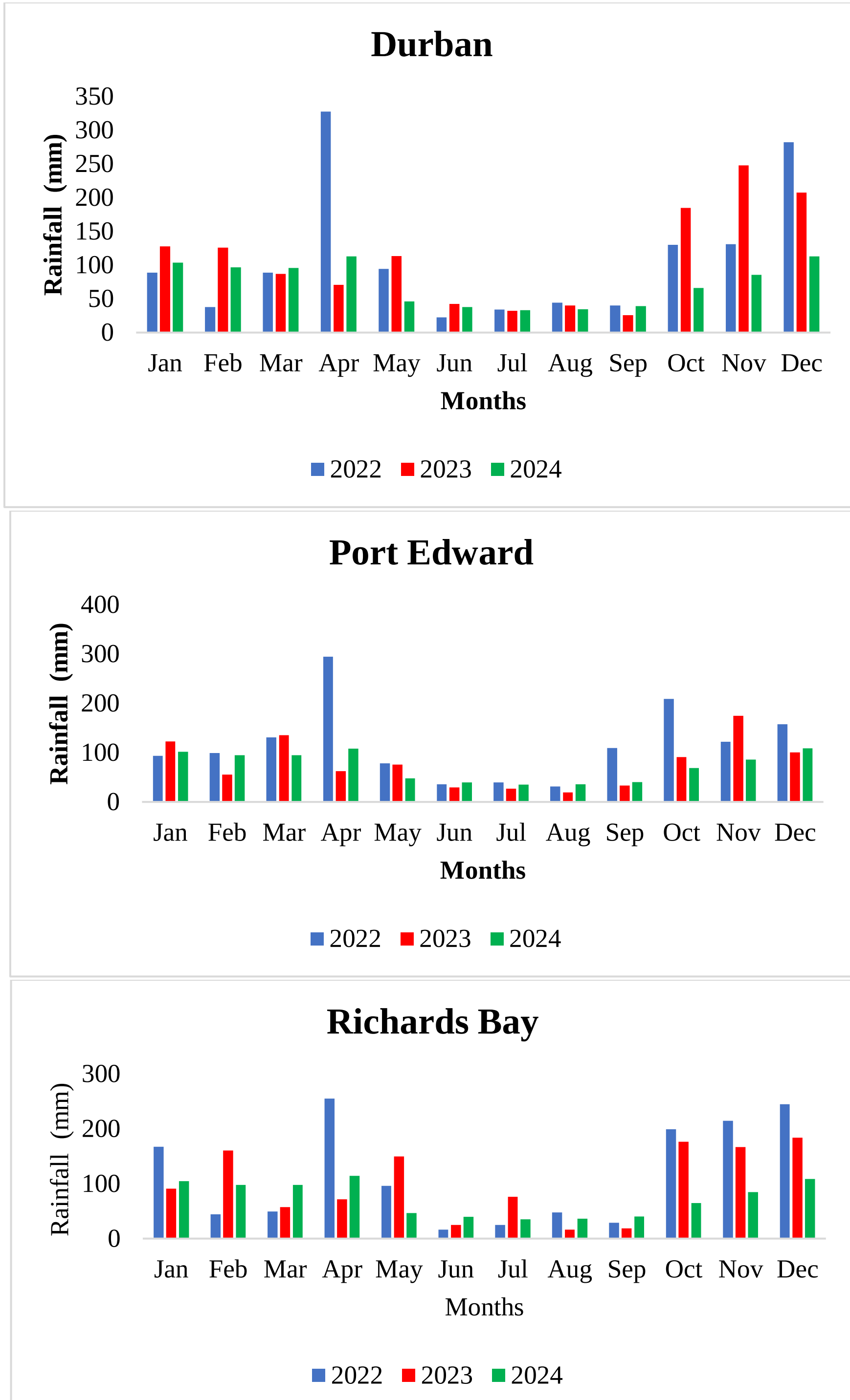


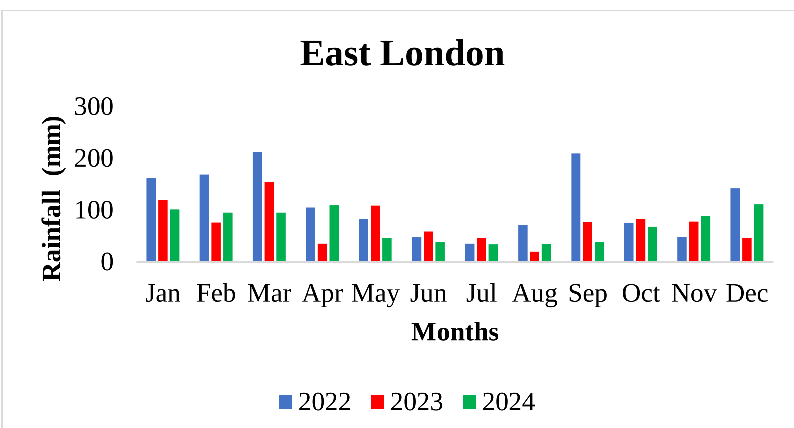
Figure 9.
Rainfall prediction for 2024 compared with 2023 and 2023 rainfall under the humid subtropical without dry season (Durban, Port Edward, and Richards Bay).

4.6.2. Temperate Oceanic Without a Dry Season

Rainfall prediction for the temperate oceanic without dry season (East London, George, and Mthatha) for 2024 compared with rainfall values for 2022 and 2023 is shown in Figure 10. The first figure in Figure 10 shows that while a steady increase in the amount of rainfall received in 2022 for the first three months is observed, in 2024, there is a slight decrease of less than 5mm between January and March. In 2023, there was a decrease of about 40mm between January and February and an increase of about 80mm between February and March. In April, the predicted values are within the same range with 2022 rainfall and thrice 2023 rainfall. Though there is no dry season in this climatic region, the rainfall predicted and recorded for previous years is reduced during winter compared to other months. It is also seen in the graph that there would be a lesser amount of rainfall between May and July in 2024 compared to 2023, as well as in September and October. The result also reveals that the annual average rainfall for 2024 will be lower than that of 2023 and much lower than the 2022 yearly rainfall.

However, in Geroge, there is an increase of about 200mm between the 2022 annual rainfall and the 2023 annual rainfall. A further 200mm increase in rainfall is predicted between 2023 and 2024. This is exciting news for a region that receives little rainfall. Predicted January rainfall is about four times what was recorded in 2022 and 2023. Also, the predicted rainfall for February is about four times that of 2023 and about 20mm more than the 2022 February rainfall in George. However, there would be lower rainfall in March compared to other years, and it would be about 100mm lower than the 2023 March rainfall. This would be compensated in April as it is predicted that the 2024 April rainfall would be about 90mm more than what was recorded in 2023 and 100mm more than the 2022 April rainfall. Though there is no dry season for this climatic zone, the result shows that in five months (April, July, August, October, and November) 2022, rainfall was below 20mm, with July receiving about 5mm the entire month. August 2023 also received rainfall below 10mm. This value is expected to rise to about 40mm in 2024. Finally, the results in George indicate that there would be more rainfall between October and December by a wide margin than what was recorded in 2022 and 2023.

While in George, there was an increase of about 200mm from 2022 to 2023 and from 2023 to 2024, in Mthatha, the reverse is the case. There was an annual decrease of about 250mm from 2022 rainfall to 2023 rainfall. It is predicted that there will be a further decrease of about 300mm between 2023 rainfall and 2024 rainfall. As with this climatic zone, there was no dry season, as rainfall was experienced and predicted during winter. It is predicted that January 2024 rainfall will be about half of what was recorded in 2023 and March and May. February rainfall would be about 10mm more than 2023 February rainfall but 10mm lower than 2022 February rainfall. It is also predicted that there will be more rainfall in April 2024 by about 20mm compared to April 2023. Between June and August, rainfall predictions for 2024 and recorded values for 2022 and 2023 are within the same range. Compared to 2023, lower rainfall is predicted in September and November; in October and December, the amount of rainfall expected would be similar to that of 2023.



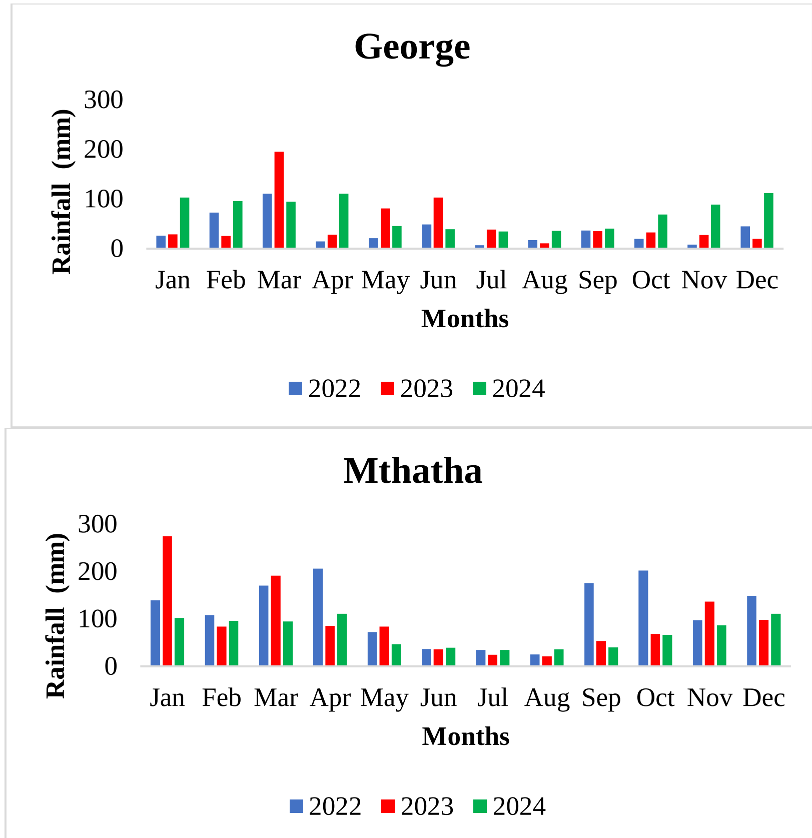


Figure 10.
Rainfall prediction for 2024 compared with 2023 and 2023 rainfall under the temperate oceanic without dry season (East London, George, and Mthatha).

4.6.3. Warm and Dry Summer

For the warm and dry summer, rainfall prediction for 2024 compared with recorded values of 2022 and 2023 is presented in Figure 11. In Bredasdorp, it is shown that there was little rainfall in the summer months of December and January of 2023, while low rainfall was recorded in April and November of 2022. The predicted rainfall values estimate that there will be much more rainfall from October to December than in previous years. Similarly, more rainfall is predicted 2024 for January, February, and April than in previous years. It is predicted that the rain in March would be considerably lower than that of 2023 by almost 200mm, while the predicted rainfall for June 2024 would be about 100mm lower than the amount of rain received in 2023. A similar amount of rainfall in 2023, May, and July is predicted for 2024.

In Cape Town, dry summer was only experienced in 2023 among the three years, with 2022 December receiving considerable rainfall. Most rainfall recorded in Cape Town in 2023 was during autumn, early winter, and early spring. In January 2023, less than 5mm of rainfall was received compared to a predicted value of about 100mm in January 2024. Similarly, the predicted value for February is about five times what it was in February 2023. In March, from year to year, there is a steady increase in the amount of rainfall received from 2022 to 2024, although 2022 rainfall in March was much lower than in 2023 and 2024. This pattern is also seen in April, with 2023 and 2024 values closer to what was obtained in 2024. While there was an increase in rain from April to June for both 2023 and 2024, the predicted values in 2024 are expected to reduce during this period. Predicted values in July for 2024 and recorded rainfall values for 2022 and 2023 are similar. The observed increase in

ISSN: 2576-8484
Vol. 8, No. 6: 1805-1827, 2024
DOI: 10.55214/25768484.v8i6.2347

2022 rainfall from October to December can also be seen in the predicted values for 2024, while in 2023, there was a decrease in these months.

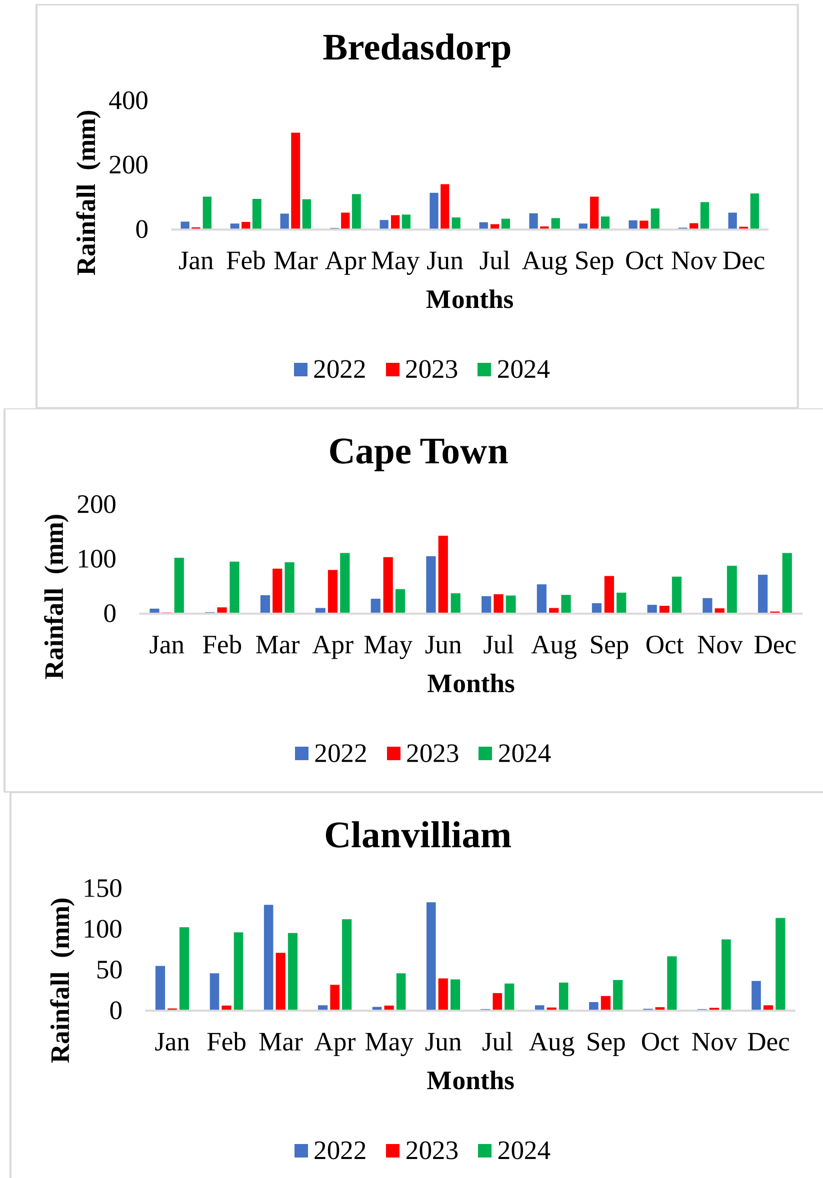


Figure 11.
Rainfall prediction for 2024 compared with 2023 and 2023 rainfall under the warm and dry summer (Bredasdorp, Cape Town, and Clanvilliam).

In Clanvilliam, significant rainfall was recorded in 2022 in January, February, March, June, and December. Despite it being a dry summer, in 2023, it was mostly a dry year, with monthly rainfall below 20mm recorded in nine months (January, February, May, July, August, September, October, November, and December). In 2022, seven months of rainfall were recorded below 20mm (April, May, July, August, September, October, and November). In the predicted rainfall estimates for 2024, only four months (May, June, July, and August) are expected to have rainfall values below 40mm. In 2023, only

April had monthly rainfall of over 40mm recorded, but in 2022, there was rainfall of about 130mm recorded in March and June. The predicted values of 2024 show that rainfall of over 90mm is expected in five months (January, February, March, April, and December).

5. Conclusion

This research provides insight into the amount of rainfall expected in South Africa in 2024. Random forest was used for this study to predict 2024 monthly rainfall for different cities in nine climatic zones. Based on the Koppen-Geiger climate classification system, South Africa was divided into three groups and further subdivided. The broad three climatic zones are the arid, subtropical wet, and subtropical dry climate classification. The arid climate classification was further divided into four: the cold and semi-arid steppe, cold arid desert, hot and semi-arid steppe, and the hot arid desert. The subtropical wet climate classification was divided into humid subtropical with dry winter and subtropical highland with dry winter, while the subtropical dry climate classification was divided into three: humid subtropical without dry season, temperate oceanic without dry season, and warm and dry summer. Three locations were selected in each category, and random forest is used for 2024 rainfall prediction. The results were then compared with 2022 and 2023 rainfall in those locations. In some locations, more rainfall is expected in 2024 compared to previous years, while in some locations, less rainfall was predicted. Where there would be more rainfall, it is good news for farmers. In comparison, in regions where lower rainfall is expected, farmers need to look for alternative ways to combat the foreseeable challenges with low rainfall.

Funding:
This work is funded by the DSCI-NICIS National e-science Postgraduate Teaching and Training Platform (NEPTTP)

Availability of Data and Materials:
The datasets used for this work are available on request from the South African Weather Service and Giovanni site accessible via https://giovanni.gsfc.nasa.gov/giovanni/

Acknowledgments:
The authors gladly recognize the infrastructure support offered by Sol Plaatje University for this study.

References

[1] Ali, M., Prasad, R., Xiang, Y. and Yaseen, Z.M., 2020. Complete ensemble empirical mode decomposition hybridized with random forest and kernel ridge regression model for monthly rainfall forecasts. Journal of Hydrology, 584, p.124647.
[2] Anochi, J.A., de Almeida, V.A. and de Campos Velho, H.F., 2021. Machine learning for climate precipitation prediction modeling over South America. Remote Sensing, 13(13), p.2468.
[3] Appiah-Badu, N.K.A., Missah, Y.M., Amekudzi, L.K., Ussiph, N., Frimpong, T. and Ahene, E., 2021. Rainfall prediction using machine learning algorithms for the various ecological zones of Ghana. IEEE Access, 10, pp.5069-5082.
[4] Bamisile, O., Oluwasanmi, A., Ejiyi, C., Yimen, N., Obiora, S. and Huang, Q., 2021. Comparison of machine learning and deep learning algorithms for hourly global/diffuse solar radiation predictions. International Journal of Energy Research.
[5] Baran, Á., Lerch, S., El Ayari, M. and Baran, S., 2021. Machine learning for total cloud cover prediction. Neural Computing and Applications, 33(7), pp.2605-2620.
[6] Bouras, E.H., Jarlan, L., Er-Raki, S., Balaghi, R., Amazirh, A., Richard, B. and Khabba, S., 2021. Cereal yield forecasting with satellite drought-based indices, weather data and regional climate indices using machine learning in Morocco. Remote Sensing, 13(16), p.3101.

[7] Cedric, L.S., Adoni, W.Y.H., Aworka, R., Zoueu, J.T., Mutombo, F.K., Krichen, M. and Kimpolo, C.L.M., 2022. Crops yield prediction based on machine learning models: case of west african countries. Smart Agricultural Technology, p.100049.

[8] Chattopadhyay, A., Nabizadeh, E. and Hassanzadeh, P., 2020. Analog forecasting of extreme-causing weather patterns using deep learning. Journal of Advances in Modeling Earth Systems, 12(2), p.e2019MS001958.

[9] Chen, C.Y., Yeh, N.C., Chuang, Y.C. and Lin, C.Y., 2022. Development of a Low-Cost Portable Cluster for Numerical Weather Prediction. Electronics, 11(17), p.2769.

[10] Cramer, S., Kampouridis, M., Freitas, A.A. and Alexandridis, A.K., 2017. An extensive evaluation of seven machine learning methods for rainfall prediction in weather derivatives. Expert Systems with Applications, 85, pp.169-181.

[11] Diez-Sierra, J. and del Jesus, M., 2020. Long-term rainfall prediction using atmospheric synoptic patterns in semi-arid climates with statistical and machine learning methods. Journal of Hydrology, 586, p.124789.

[12] Ferreira, G.W. and Reboita, M.S., 2022. A new look into the South America precipitation regimes: Observation and Forecast. Atmosphere, 13(6), p.873.

[13] Gao, T., Wang, H.J. and Zhou, T., 2017. Changes of extreme precipitation and nonlinear influence of climate variables over monsoon region in China. Atmospheric Research, 197, pp.379-389.

[14] Gómez, D., Aristizábal, E., García, E.F., Marín, D., Valencia, S. and Vásquez, M., 2023. Landslides forecasting using satellite rainfall estimations and machine learning in the Colombian Andean region. Journal of South American Earth Sciences, 125, p.104293.

[15] He, S., Li, X., DelSole, T., Ravikumar, P. and Banerjee, A., 2021, May. Sub-seasonal climate forecasting via machine learning: Challenges, analysis, and advances. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 35, No. 1, pp. 169-177).

[16] He, R.R., Chen, Y., Huang, Q. and Kang, Y., 2019. LASSO as a tool for downscaling summer rainfall over the Yangtze River valley. Hydrological Sciences Journal, 64(1), pp.92-104.

[17] He, R., Zhang, L. and Chew, A.W.Z., 2022. Modelling and predicting rainfall time series using seasonal-trend decomposition and machine learning. Knowledge-Based Systems, 251, p.109125.

[18] Islam, M.S., Hossain, A., Khatun, A. and Kor, A.L., 2023, February. Evaluation of the Performance of Machine Learning and Deep Learning Techniques for Predicting Rainfall: An Illustrative Case Study from Australia. In 2023 International Conference on Electrical, Computer and Communication Engineering (ECCE) (pp. 1-5). IEEE.

[19] Jose, D.M., Vincent, A.M. and Dwarakish, G.S., 2022. Improving multiple model ensemble predictions of daily precipitation and temperature through machine learning techniques. Scientific Reports, 12(1), pp.1-25.

[20] Meque, A., Gamedze, S., Moitlhobogi, T., Booneeady, P., Samuel, S. and Mpalang, L., 2021. Numerical weather prediction and climate modelling: Challenges and opportunities for improving climate services delivery in Southern Africa. Climate Services, 23, p.100243.

[21] Nkiruka, O., Prasad, R. and Clement, O., 2021. Prediction of malaria incidence using climate variability and machine learning. Informatics in Medicine Unlocked, 22, p.100508.

[22] Ouallouche, F., Lazri, M. and Ameur, S., 2018. Improvement of rainfall estimation from MSG data using Random Forests classification and regression. Atmospheric Research, 211, pp.62-72.

[23] Polishchuk, B., Berko, A., Chyrun, L., Bublyk, M. and Schuchmann, V., 2021, September. The rain prediction in Australia based Big Data analysis and machine learning technology. In 2021 IEEE 16th International Conference on Computer Sciences and Information Technologies (CSIT) (Vol. 1, pp. 97-100). IEEE.

[24] Raval, M., Sivashanmugam, P., Pham, V., Gohel, H., Kaushik, A. and Wan, Y., 2021. Automated predictive analytics tool for rainfall forecasting. Scientific Reports, 11(1), p.17704.

[25] Ridwan, W.M., Sapitang, M., Aziz, A., Kushiar, K.F., Ahmed, A.N. and El-Shafie, A., 2021. Rainfall forecasting model using machine learning methods: Case study Terengganu, Malaysia. Ain Shams Engineering Journal, 12(2), pp.1651-1663.

[26] Rolnick, D., Donti, P.L., Kaack, L.H., Kochanski, K., Lacoste, A., Sankaran, K., Ross, A.S., Milojevic-Dupont, N., Jaques, N., Waldman-Brown, A. and Luccioni, A.S., 2022. Tackling climate change with machine learning. ACM Computing Surveys (CSUR), 55(2), pp.1-96.

[27] Sachindra, D.A., Ahmed, K., Rashid, M.M., Shahid, S. and Perera, B.J.C., 2018. Statistical downscaling of precipitation using machine learning techniques. Atmospheric research, 212, pp.240-258.

[28] Shah, N.H., Priamvada, A. and Shukla, B.P., 2023. Random forest-based nowcast model for rainfall. Earth Science Informatics, 16(3), pp.2391-2403.

[29] Sharma, R., Kamble, S.S., Gunasekaran, A., Kumar, V. and Kumar, A., 2020. A systematic literature review on machine learning applications for sustainable agriculture supply chain performance. *Computers & Operations Research*, *119*, p.104926.

[30] Yu, P.S., Yang, T.C., Chen, S.Y., Kuo, C.M. and Tseng, H.W., 2017. Comparison of random forests and support vector machine for real-time radar-derived rainfall forecasting. Journal of hydrology, 552, pp.92-104.

[31] Zainudin, S., Jasim, D.S. and Bakar, A.A., 2016. Comparative analysis of data mining techniques for Malaysian rainfall prediction. Int. J. Adv. Sci. Eng. Inf. Technol, 6(6), pp.1148-1153.