# District and Taluka wise Medium Range Weather Forecast Verification

This project examines the accuracy of medium-range weather forecasts on a Taluka level and District level.


## Submitted By:

**Satyam Kumbhar**
ID: 202319005
**Kaushal Kathiriya**
ID: 202319013
**Drashti Nayakpara**
ID: 202319019
**Piyush Pandya**
ID: 202319023

## Submitted To:

**Dr. S. B. Yadav**
(Asst. Professor )
(Department of Agricultural Meteorology
BACA, AAU, Anand)

# Contents

# Weather Forecasting:

Weather forecasting is **the use of science and technology to predict the weather in a specific location and time period**. It involves making observations of the atmosphere to predict things like cloud cover, rain, snow, wind speed, and temperature before they happen.

Accurate forecasts are particularly vital in sectors like agriculture, transportation, and disaster management, helping farmers optimize their planting schedules, assisting transportation systems in adjusting routes, and aiding in effective disaster response and mitigation strategies.

# Weather Forecast Verification:

Weather forecast verification is an important process that evaluates the **accuracy** and **reliability** of weather predictions. This is done by comparing the forecasted weather conditions with the actual observed weather conditions over a certain time period. Through verification, forecasters can assess the performance of different forecasting models and techniques, identify strengths and weaknesses, and improve the overall accuracy of future forecasts.

# Types Of Weather Forecast:

| Sr No. | Type of Weather Forecast | Validity Period | Application/Uses |
|---|---|---|---|
| 1 | Now-casting | Minutes to a few hours | Provides real-time information for immediate decision-making, such as severe weather warnings, and event planning. Often used for short-term guidance on rapidly changing conditions. |
| 2 | Short-Range Forecast | Up to 1-3 days | Useful for daily weather planning, such as scheduling outdoor activities, and preparing for minor weather-related impacts. Common in media weather reports. |
| 3 | Medium-Range Forecast | 3-7 days | Applied for planning events or travel a week ahead. Helps in preparing for moderate weather changes and is often used in agriculture, retail planning. |
| 4 | Long-Range Forecast | 1-2 weeks | Assists with longer-term planning and decision-making, major event coordination, and seasonal preparations. Provides an outlook on general weather trends. |
| 5 | Extended-Range Forecast | 2 weeks to 1 month | Used for assessing trends and planning in advance, such as for agriculture, energy management, and climate-sensitive industries. Provides a broader view of weather patterns and anomalies. |

# Objectives:

## Accuracy

Assess the accuracy of medium-range forecasts.

## Regional Variation

Analyze how accuracy changes across different Talukas.

## Verification Methods

Apply various techniques to evaluate forecast performance.

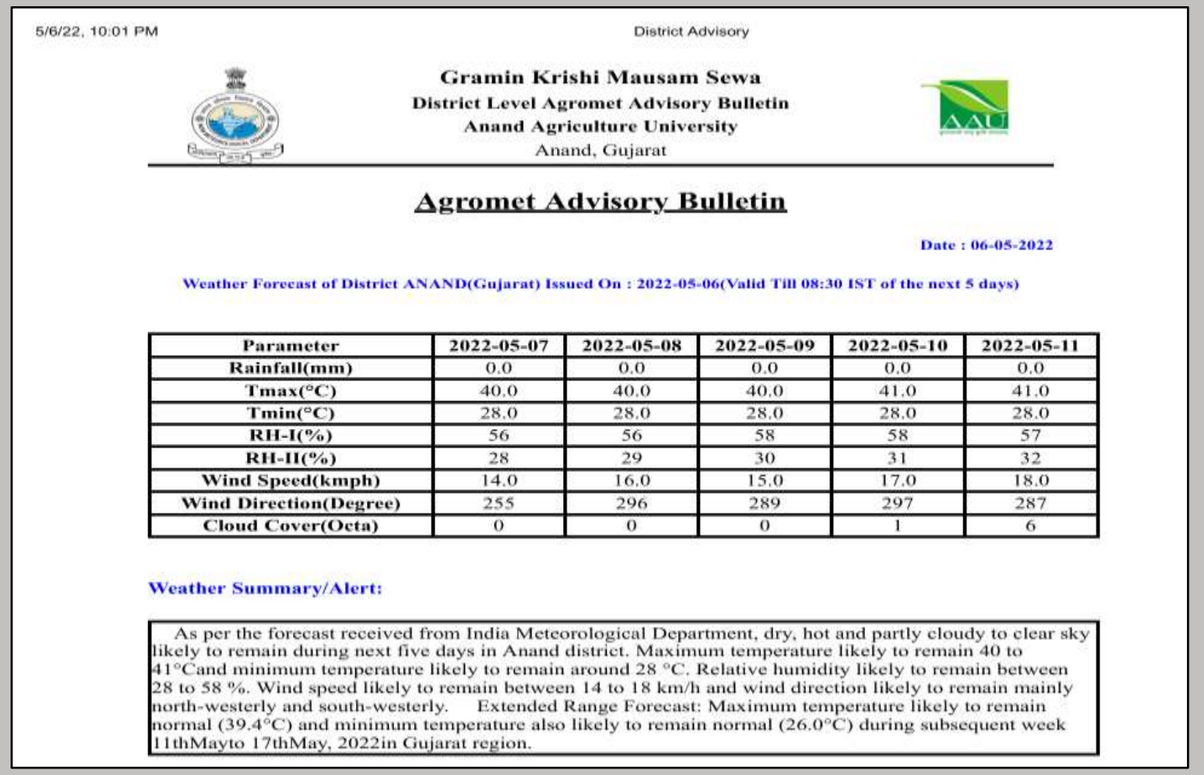
# Data Collection:

## Source:

- Observed Rainfall data (Taluka Level): GSDMA
- Forecasted Rainfall data (Taluka Level): Department of Agricultural Meteorology, AAU, Anand.
- Observed and Forecasted Data (District Level): Department of Agricultural Meteorology, AAU, Anand.

## Format:

- Observed data for Taluka level was in pdf format.



5/6/22, 10:01 PM District Advisory

**Gramin Krishi Mausam Sewa**
**District Level Agromet Advisory Bulletin**
**Anand Agriculture University**
Anand, Gujarat

**Agromet Advisory Bulletin**

**Date : 06-05-2022**

**Weather Forecast of District ANAND(Gujarat) Issued On : 2022-05-06(Valid Till 08:30 IST of the next 5 days)**

| Parameter | 2022-05-07 | 2022-05-08 | 2022-05-09 | 2022-05-10 | 2022-05-11 |
|---|---|---|---|---|---|
| Rainfall(mm) | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| Tmax(°C) | 40.0 | 40.0 | 40.0 | 41.0 | 41.0 |
| Tmin(°C) | 28.0 | 28.0 | 28.0 | 28.0 | 28.0 |
| RH-I(%) | 56 | 56 | 58 | 58 | 57 |
| RH-II(%) | 28 | 29 | 30 | 31 | 32 |
| Wind Speed(kmph) | 14.0 | 16.0 | 15.0 | 17.0 | 18.0 |
| Wind Direction(Degree) | 255 | 296 | 289 | 297 | 287 |
| Cloud Cover(Octa) | 0 | 0 | 0 | 1 | 6 |

**Weather Summary/Alert:**

As per the forecast received from India Meteorological Department, dry, hot and partly cloudy to clear sky likely to remain during next five days in Anand district. Maximum temperature likely to remain 40 to 41°Cand minimum temperature likely to remain around 28 °C. Relative humidity likely to remain between 28 to 58 %. Wind speed likely to remain between 14 to 18 km/h and wind direction likely to remain mainly north-westerly and south-westerly. Extended Range Forecast: Maximum temperature likely to remain normal (39.4°C) and minimum temperature also likely to remain normal (26.0°C) during subsequent week 11thMayto 17thMay, 2022in Gujarat region.

# Data Extraction and Manipulation:

## Data Extraction

- For data extraction from pdf format, Python libraries such as pdfplumber, OS, Pandas, Tabula, datetime, PyMuPDF were used.

## Data Manipulation

- Missing data was added manually from Anand Agromet Advisory data.

```
[ ]  final_df.to_csv('/content/2023_district_60_1.csv', index=False)
```

```
pdf_path = '/content/Anand_District Advisory E (4).pdf'
dfs = tabula.read_pdf(pdf_path, pages=1)
df = dfs[0]
df_transpose = df.T
df_transpose
```

```
WARNING:tabula.backend:Got stderr: Aug 19, 2024 7:40:11 PM org.apache.pdfbox.pdmodel.font.PDType0Font toUnicode
WARNING: No Unicode mapping for CID+0 (0) in font CAAAAA+Times New Roman
```

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|---|
| **Parameter** | Rainfall(mm) | Tmax(°C) | Tmin(°C) | RH-I(%) | RH-II(%) | Wind Speed(kmph) | Wind Direction(Degree) | Cloud Cover(Octa) |
| **2023-01-14** | 0.0 | 28.0 | 12.0 | 42.0 | 25.0 | 11.1 | 296.0 | 0.0 |
| **2023-01-15** | 0.0 | 29.0 | 12.0 | 38.0 | 23.0 | 16.1 | 30.0 | 0.0 |
| **2023-01-16** | 0.0 | 29.0 | 13.0 | 38.0 | 25.0 | 15.5 | 25.0 | 0.0 |
| **2023-01-17** | 0.0 | 29.0 | 14.0 | 35.0 | 24.0 | 15.6 | 34.0 | 3.0 |
| **2023-01-18** | 0.0 | 29.0 | 15.0 | 31.0 | 19.0 | 17.3 | 32.0 | 3.0 |

# Methodology:

1

## Data Analysis

Statistical methods are used to compare observed and forecasted data.

2

## Verification Metrics

Quantitative measures assess forecast skill, such as Ratio score, HK score, Threat score and RMSE.

3

## Results and Conclusion

Examine accuracy differences across districts and talukas and give proper conclusion.

# Data Analysis and Forecast verification metrics:

The results verified using ratio score, HK score and RMSE for rainfall and RMSE for other parameters which are as below.

- **Ratio Score**
- **HK Score(Hanssen and Kuipers index, peirce’s skill score)**
- **Threat Score(critical success index)**
- **RMSE(Root mean squared error)**
- **Correct, usable or unusable forecast**

# Ratio score:

- Ratio score measures the accuracy of forecast out of total forecasts issued.
- The ratio score close to zero indicates imperfect forecast and close to hundred indicates perfect forecast.
- The ratio score varies from 0 to 100 percent.
- It is calculated by using the equation:

$$\text{Ratio score} = \frac{\text{correct forecast}}{\text{total no.of forecast}} \times 100 = \frac{YY+NN}{n} \times 100$$

Where,

YY = no. of case forecasted yes and observed yes
NN = no. of case forecasted no and observed no
n = total no. of observations

# HK score:

- HK score is the ratio of economic saving over climatology due to the forecaster to that of the perfect forecaster.
- The score varies between -1 and +1.
- It is calculated as:

$$H.K.\,Score = \frac{YYNN - YNNY}{(YY + YN)(NY + NN)}$$

- HK Score = 1: Perfect forecast, all predictions are correct.
- HK Score = 0: No skill, the forecast is no better than random guessing.
- HK Score < 0: Negative skill, the forecast is worse than random guessing.

# RMSE (Root Mean Squared Error):

- The values indicate the degree of error in the forecast.
- The root mean square error (RMSE) of all weather parameters was worked out for the absolute error between observed and forecasted values.
- The lower values of RMSE indicate less difference between observed and forecasted values.

$$\text{RMSE} = \frac{\sqrt{\Sigma(F-O)2}}{n}$$

Where,

F = forecasted value, O = observed value, n = no. of observations

# Threat Score:

- Measures the fraction of observed and/or forecast events that were correctly predicted.
- It can be thought of as the accuracy when correct negatives have been removed from consideration, that is, TS is only concerned with forecasts that count. Sensitive to hits, penalizes both misses and false alarms. Does not distinguish source of forecast error.
- Its range is 0 to 1, with a value of 1 indicating a perfect forecast.

**Contingency Table**

| | | Observed | | |
|---|---|---|---|---|
| | | yes | no | Total |
| Forecast | yes | *hits* | *false alarms* | *forecast yes* |
| | no | *misses* | *correct negatives* | *forecast no* |
| Total | | *observed yes* | *observed no* | *total* |

*Threat score (critical success index)* - $TS = \frac{hits}{hits + misses + false\ alarms}$ (also denoted CSI)

# Correlation Coefficient:

- The correlation coefficient (often denoted as *r*r) quantifies the strength and direction of a linear relationship between two variables.
- In the context of forecasting, it measures how closely the forecast values align with the observed values on a scatter plot.
- Range:
  - The value of *r*r ranges from -1 to 1
  - r=1: Perfect positive linear relationship.
  - r=−1: Perfect negative linear relationship.
  - r=0: No linear relationship.

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}}$$

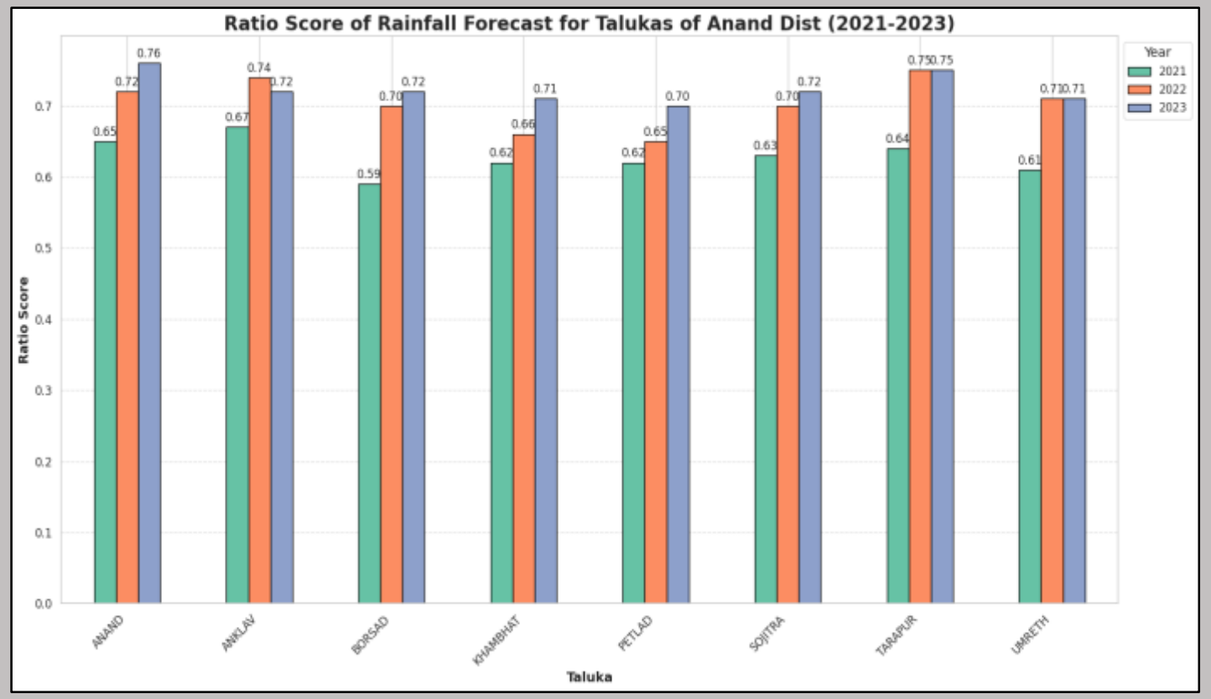
$r$ = correlation coefficient
$x_i$ = values of the x-variable in a sample
$\bar{x}$ = mean of the values of the x-variable
$y_i$ = values of the y-variable in a sample
$\bar{y}$ = mean of the values of the y-variable

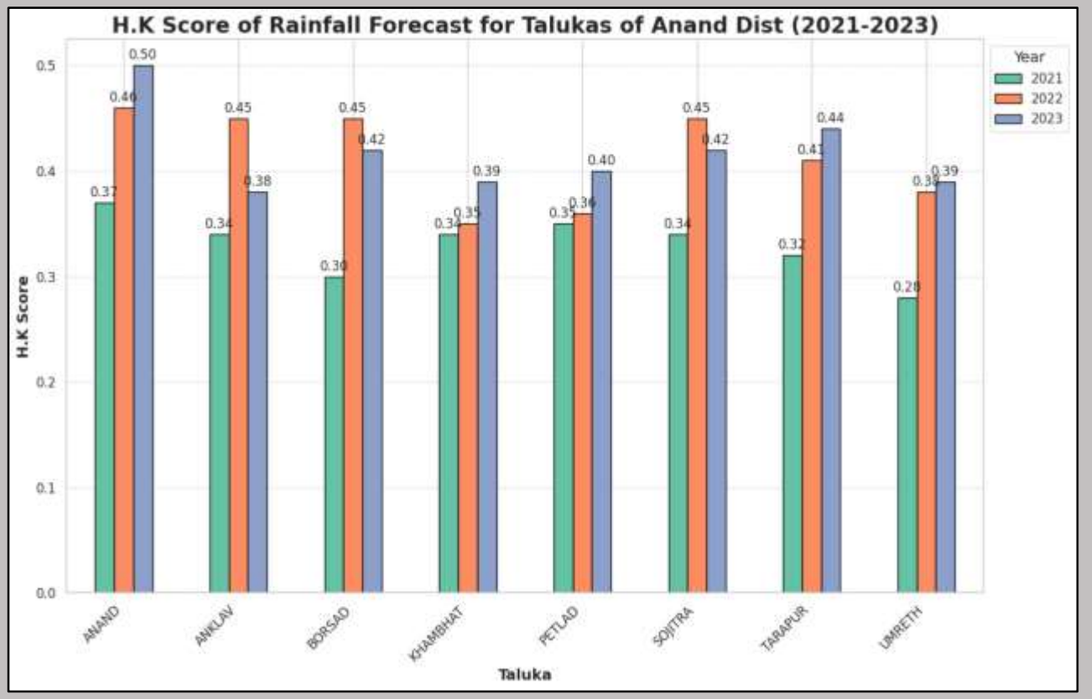
# Correct, Usable and Unusable:

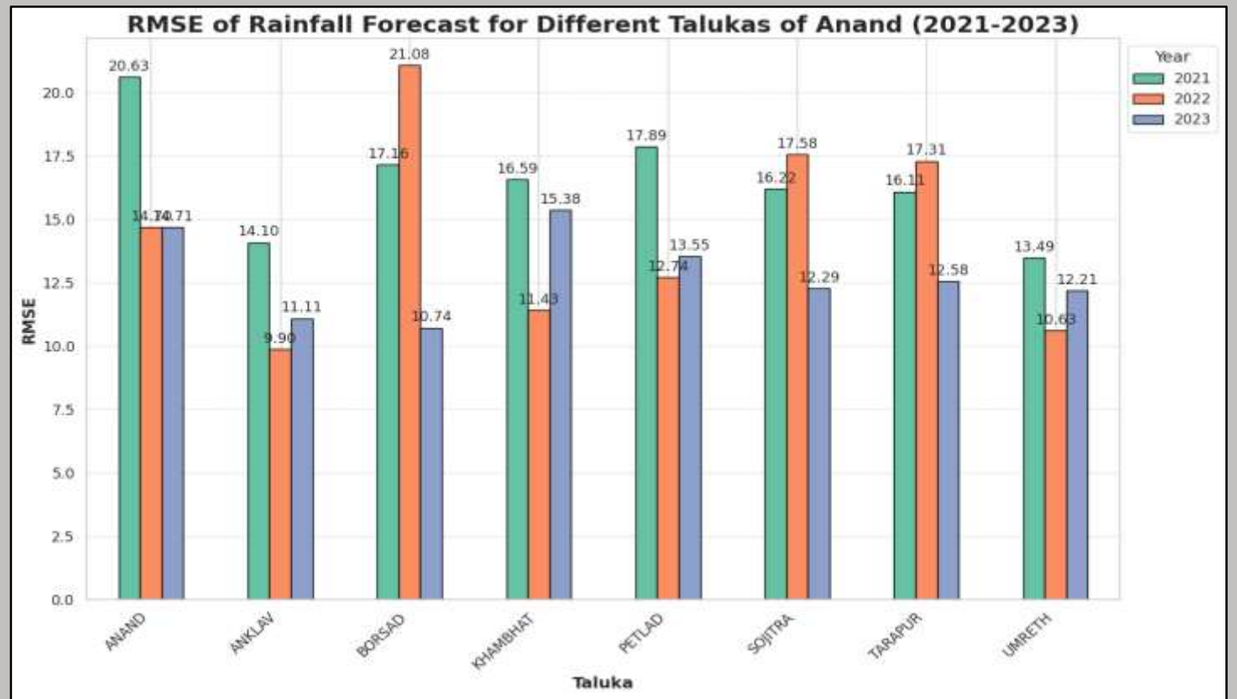
| Parameters | Correct | Usable | Unusable |
|---|---|---|---|
| | diff = Observed - Forecasted | | |
| **Rainfall** | diff < 1 mm | 1 mm < diff < 10mm | diff > 10 mm |
| **Temperature** | diff < 1 °C | 1 °C < diff < 2 °C | diff > 2 °C |
| **Humidity** | diff < 1 % | 1 % < diff < 10 % | diff > 10 % |

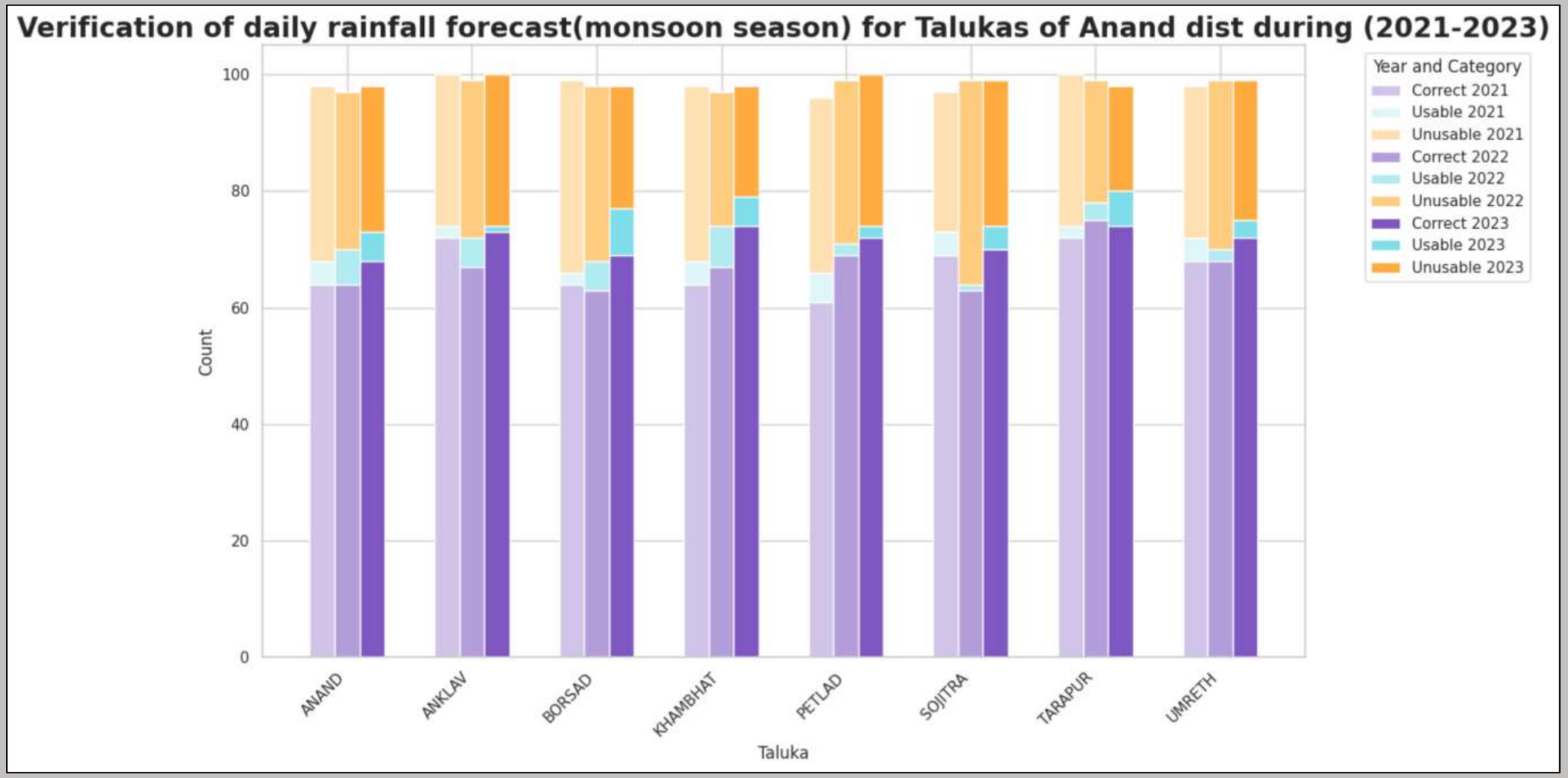
# Results:

| Accuracy of forecasted rainfall at taluka level (Anand) (2021-2023) | | | | | | | |
|---|---|---|---|---|---|---|---|
| Taluka | Year | Ratio Score | HK Score | RMSE | Correct | Usable | Unusable |
| ANAND | 2021 | 0.65 | 0.37 | 20.63 | 117 (64) | 7 (4) | 55 (30) |
| | 2022 | 0.72 | **0.46** | 14.7 | 117 (64) | 11 (6) | 50 (27) |
| | 2023 | **0.76** | **0.5** | 14.71 | 125 (68) | 10 (5) | 45 (25) |
| ANKLAV | 2021 | 0.67 | 0.34 | 14.1 | 131 (72) | 3 (2) | 48 (26) |
| | 2022 | **0.74** | **0.45** | **9.9** | 122 (67) | 9 (5) | 49 (27) |
| | 2023 | 0.72 | 0.38 | **11.11** | 133 (73) | 2 (1) | 47 (26) |
| BORSAD | 2021 | 0.59 | 0.3 | 17.16 | 117 (64) | 3 (2) | 61 (33) |
| | 2022 | 0.7 | **0.45** | 21.08 | 116 (63) | 9 (5) | 54 (30) |
| | 2023 | 0.72 | 0.42 | **10.74** | 126 (69) | 14 (8) | 38 (21) |
| KHAMBHAT | 2021 | 0.62 | 0.34 | 16.59 | 117 (64) | 7 (4) | 55 (30) |
| | 2022 | 0.66 | 0.35 | **11.43** | 123 (67) | 12 (7) | 42 (23) |
| | 2023 | 0.71 | 0.39 | 15.38 | 135 (74) | 9 (5) | 35 (19) |
| PETLAD | 2021 | 0.62 | 0.35 | 17.89 | 112 (61) | 10 (5) | 55 (30) |
| | 2022 | 0.65 | 0.36 | **12.74** | 126 (69) | 4 (2) | 51 (28) |
| | 2023 | 0.7 | 0.4 | 13.55 | 131 (72) | 3 (2) | 48 (26) |
| SOJITRA | 2021 | 0.63 | 0.34 | 16.22 | 126 (69) | 8 (4) | 44 (24) |
| | 2022 | 0.7 | **0.45** | 17.58 | 116 (63) | 2 (1) | 64 (35) |
| | 2023 | 0.72 | 0.42 | **12.29** | 129 (70) | 7 (4) | 45 (25) |
| TARAPUR | 2021 | 0.64 | 0.32 | 16.11 | 132 (72) | 3 (2) | 47 (26) |
| | 2022 | **0.75** | 0.41 | 17.31 | 138 (75) | 5 (3) | 39 (21) |
| | 2023 | **0.75** | **0.44** | **12.58** | 135 (74) | 11 (6) | 33 (18) |
| UMRETH | 2021 | 0.61 | 0.28 | 13.49 | 124 (68) | 7 (4) | 47 (26) |
| | 2022 | 0.71 | 0.38 | **10.63** | 125 (68) | 3 (2) | 53 (29) |
| | 2023 | 0.71 | 0.39 | **12.21** | 132 (72) | 5 (3) | 44 (24) |


Ratio Score of Rainfall Forecast for Talukas of Anand Dist (2021-2023)
Ratio Score
Taluka
Year
2021
2022
2023
ANAND
ANKLAV
BORSAD
KHAMBHAT
PETLAD
SOJITRA
TARAPUR
UMRETH


H.K Score of Rainfall Forecast for Talukas of Anand Dist (2021-2023)
H.K Score
Taluka
Year
2021
2022
2023
ANAND
ANKLAV
BORSAD
KHAMBHAT
PETLAD
SOJITRA
TARAPUR
UMRETH

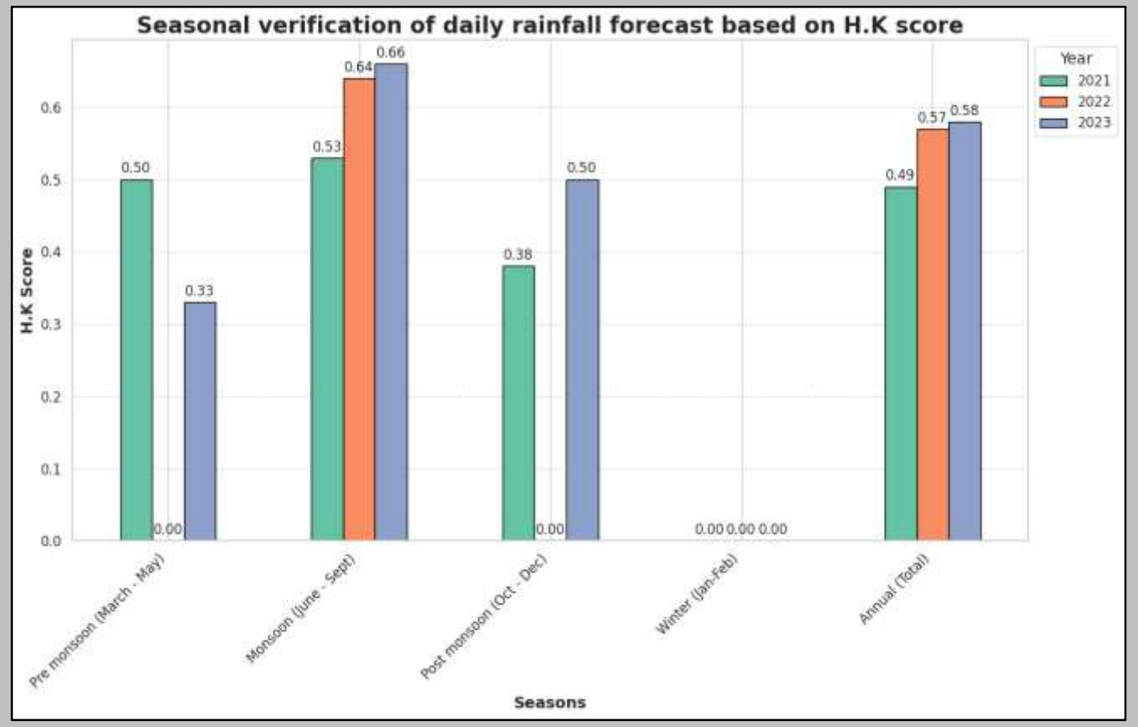

RMSE of Rainfall Forecast for Different Talukas of Anand (2021-2023)
RMSE
Taluka
Year
2021
2022
2023
ANAND
ANKLAV
BORSAD
KHAMBHAT
PETLAD
SOJITRA
TARAPUR
UMRETH

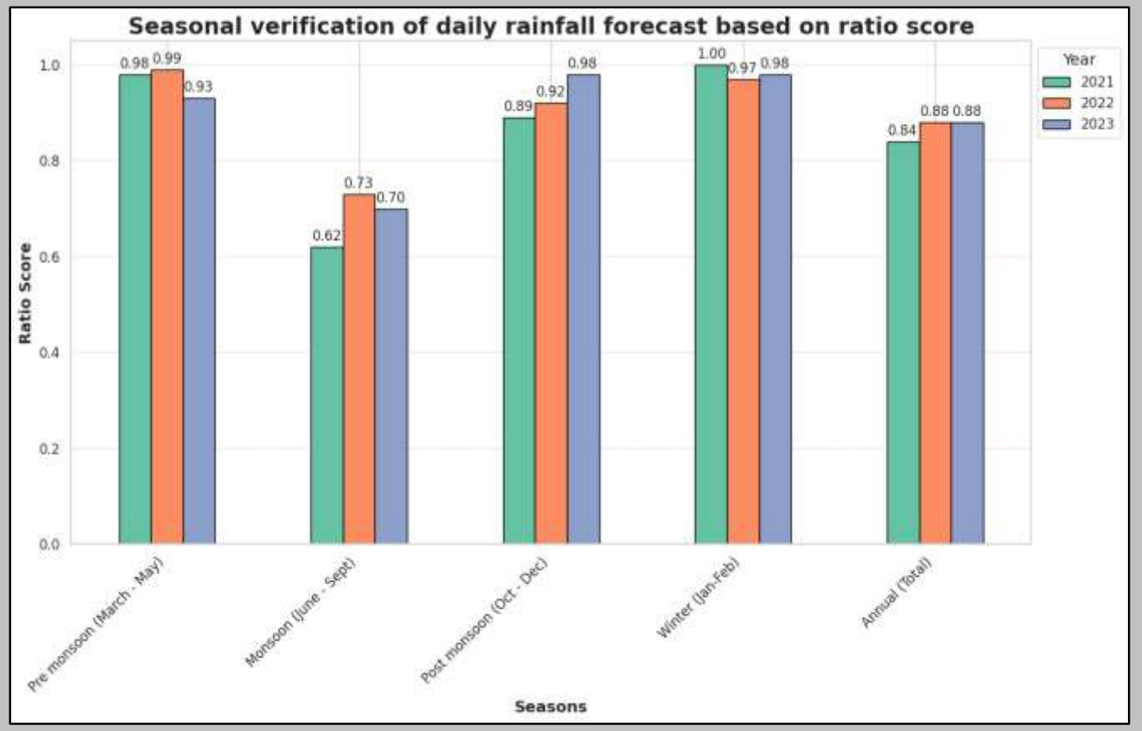

Verification of daily rainfall forecast(monsoon season) for Talukas of Anand dist during (2021-2023)
Year and Category
Correct 2021
Usable 2021
Unusable 2021
Correct 2022
Usable 2022
Unusable 2022
Correct 2023
Usable 2023
Unusable 2023
Count
0
20
40
60
80
100
ANAND
ANKLAV
BORSAD
KHAMBHAT
PETLAD
SOJITRA
TARAPUR
UMRETH
Taluka

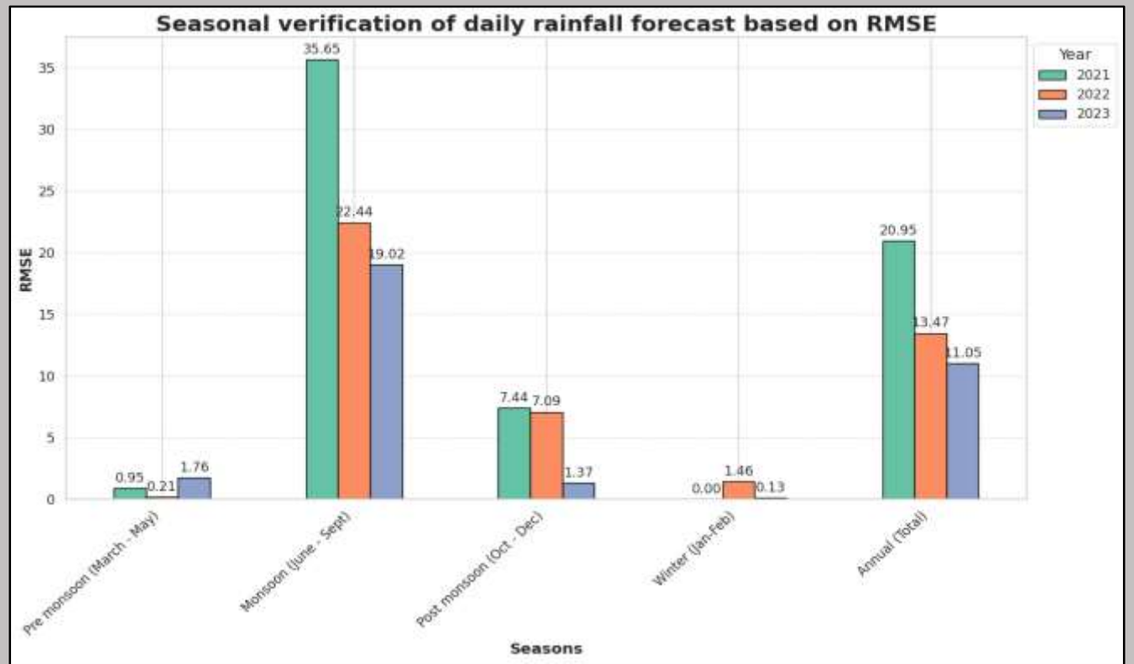
| Accuracy of forecasted rainfall Anand District (2021-2023) | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Period** | **Year** | **Ratio Score** | **HK Score** | **RMSE** | **Correct** | **Usable** | **Unusable** |
| **Pre monsoon (March - May)** | 2021 | **0.98** | 0.5 | **0.95** | 98.89 | 0 | 1.11 |
| | 2022 | **0.99** | 0 | **0.21** | 97.8 | 0 | 0 |
| | 2023 | 0.93 | 0.33 | **1.76** | 98.84 | 0 | 1.16 |
| **Monsoon (June - Sept)** | 2021 | 0.62 | **0.53** | 35.65 | 35.53 | 6.58 | 57.89 |
| | 2022 | 0.73 | **0.64** | 22.44 | 42.7 | 6.74 | 50.56 |
| | 2023 | 0.7 | **0.66** | 19.02 | 47.67 | 5.81 | 46.51 |
| **Post monsoon (Oct - Dec)** | 2021 | 0.89 | 0.38 | 7.44 | 95.12 | 1.22 | 3.66 |
| | 2022 | 0.92 | 0 | 7.09 | 100 | 0 | 0 |
| | 2023 | **0.98** | 0.5 | **1.37** | 98.89 | 1.11 | 0 |
| **Winter (Jan-Feb)** | 2021 | **1** | **_** | **0** | 100 | 0 | 0 |
| | 2022 | **0.97** | 0 | **1.46** | 100 | 0 | 0 |
| | 2023 | **0.98** | 0 | **0.13** | 100 | 0 | 0 |
| **Annual (Total)** | 2021 | 0.84 | 0.49 | 20.95 | 82.41 | 1.95 | 15.64 |
| | 2022 | 0.88 | **0.57** | 13.47 | 84.16 | 1.86 | 13.98 |
| | 2023 | 0.88 | **0.58** | 11.05 | 85.31 | 1.88 | 12.81 |

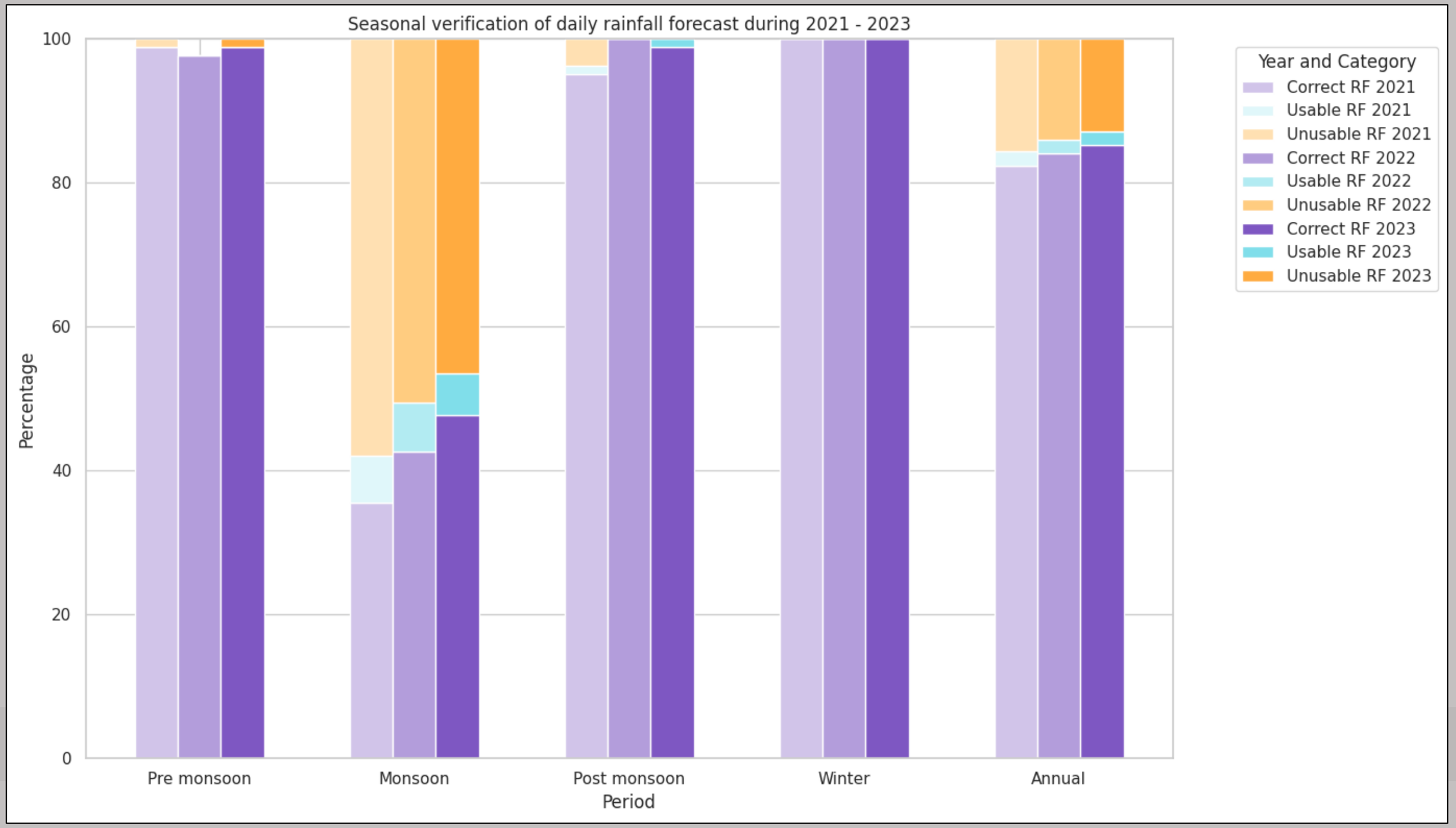
## Correlation coefficient between observed and predicted weather parameters

| Period | Year | Rainfall | Tmax | Tmin | RHmax | RHmin | Cloud Cover |
|---|---|---|---|---|---|---|---|
| Pre monsoon (March - May) | 2021 | **1** | 0.67 | **0.88** | 0.38 | 0.65 | 0.53 |
| | 2022 | 0 | 0.72 | **0.85** | 0.31 | 0.45 | 0.40 |
| | 2023 | 0.62 | **0.82** | **0.82** | 0.47 | 0.59 | 0.42 |
| Monsoon (June - Sept) | 2021 | 0.27 | 0.63 | 0.22 | 0.60 | 0.64 | 0.54 |
| | 2022 | 0.33 | **0.79** | 0.51 | 0.63 | 0.50 | 0.68 |
| | 2023 | 0.49 | 0.73 | 0.18 | 0.62 | 0.61 | 0.64 |
| Post monsoon (Oct - Dec) | 2021 | **0.79** | **0.79** | **0.80** | 0.49 | 0.66 | 0.50 |
| | 2022 | -0.02 | **0.83** | **0.79** | 0.27 | 0.58 | 0.52 |
| | 2023 | 0.63 | **0.82** | **0.84** | 0.32 | 0.70 | 0.15 |
| Winter (Jan-Feb) | 2021 | 0 | 0.64 | 0.38 | 0.14 | **0.76** | 0.62 |
| | 2022 | 0 | 0.57 | 0.47 | 0.35 | 0.54 | 0.33 |
| | 2023 | 0 | 0.58 | 0.26 | 0.16 | 0.61 | 0.28 |
| Annual (Total) | 2021 | 0.41 | **0.83** | **0.87** | 0.55 | **0.84** | 0.73 |
| | 2022 | 0.41 | **0.87** | **0.92** | 0.49 | **0.80** | 0.73 |
| | 2023 | 0.54 | **0.82** | **0.92** | 0.51 | **0.84** | **0.76** |


Seasonal verification of daily rainfall forecast based on H.K score
H.K Score
Seasons
Year
2021
2022
2023
Pre monsoon (March - May)
Monsoon (June - Sept)
Post monsoon (Oct - Dec)
Winter (Jan-Feb)
Annual (Total)
0.50
0.00
0.33
0.53
0.64
0.66
0.38
0.00
0.50
0.00 0.00 0.00
0.49
0.57
0.58


Seasonal verification of daily rainfall forecast based on ratio score
Ratio Score
Seasons
Year
2021
2022
2023
Pre monsoon (March - May)
Monsoon (June - Sept)
Post monsoon (Oct - Dec)
Winter (Jan-Feb)
Annual (Total)
0.98
0.99
0.93
0.62
0.73
0.70
0.89
0.92
0.98
1.00
0.97
0.98
0.84
0.88
0.88


Seasonal verification of daily rainfall forecast based on RMSE
RMSE
Seasons
Year
2021
2022
2023
Pre monsoon (March - May)
Monsoon (June - Sept)
Post monsoon (Oct - Dec)
Winter (Jan-Feb)
Annual (Total)
0.95
0.21
1.76
35.65
22.44
19.02
7.44
7.09
1.37
0.00
1.46
0.13
20.95
13.47
11.05


Seasonal verification of daily rainfall forecast during 2021 - 2023
Percentage
100
80
60
40
20
0
Pre monsoon
Monsoon
Post monsoon
Winter
Annual
Period
Year and Category
Correct RF 2021
Usable RF 2021
Unusable RF 2021
Correct RF 2022
Usable RF 2022
Unusable RF 2022
Correct RF 2023
Usable RF 2023
Unusable RF 2023

# Conclusion:

## Taluka Level -

### Improvement in Forecast Accuracy:

- ❑ General increase in Ratio Scores and HK Scores across most talukas from 2021 to 2023
- ❑ Decrease in RMSE values over the years, indicating reduced forecast errors.

### Enhanced Forecast Quality:

- ❑ Increase in the number of correct forecasts over time.
- ❑ Higher proportion of usable forecasts, making them more practical for decision-making.
- ❑ Decrease in unusable forecasts, suggesting better performance and reliability.

# District Level -

- **High Accuracy in Pre-Monsoon and Winter:**
  Achieved nearly 100% accuracy in both the pre-monsoon and winter seasons

- **Lower Accuracy and Higher RMSE in Monsoon:**
  The monsoon season presented significant challenges, with lower accuracy (average Ratio Score of 0.68, HK Score of 0.61) and significantly higher RMSE.

- **Steady Improvement in Monsoon Rainfall Forecasts from 2021 to 2023**

- **Strong Correlation in Annual Temperature Forecasts:**
  Annual temperature forecasts for Tmin showed consistently high Correct Counts with strong correlation values

# Thank You