Data Assimilation and Comparative studies on climate change using Data Mining with statistics approaches

P. Rajesh[1*] and S. Ravishankar[2]

Article History
Received: 07-05-2020
Revised and Accepted : 05-01-2021
Published: 16-03-2021

https://doi.org/10.56343/STET.116.014.003.007
http://stetjournals.com

Abstract

The objective of this paper, is to explore the use of data mining in weather dataset by statistical approach.The dataset used was secondary data of thirty days of weather-related data of three major cities of India *viz.,*. Delhi, Chennai and Bangalore. Data mining is a best tool for retrieving useful and hidden pattern with the help of various statistical techniques. Results showed contrasing trends of weather parameters in the three cities for the same period. The findings will be useful in weather prediction in future.

Key words: data assimilatiion, data mining, normalization, statistical approach, weather data

INTRODUCTION

Data mining is the process of analyzing hidden patterns for using pre-existing data. Data mining is also known as data discovery and knowledge discovery for handing advanced data analysis (Han *et al.* 2011). The major steps involved in a data mining process namely locate the data, data collection, data cleaning, integration, data selection, data transformation and discovering the knowledge (Bocca *et al* 2016). The area of weather forecasting is used to collecting hugs amount of data as possible to find the current weather state of the atmosphere metrics namely temperature, humidity, and wind conditions. Data mining techniques is easy to understand the atmospheric condition and to determine how to find the future atmosphere conditions using regression analysis (Rajesh and karthikeyan, 2019a). In data mining techniques, normalization is one of the most important concepts for prepare a well suitable dataset with unique format. After using the normalization techniques various scales of information are converted into similar scale of information. Various normalization techniques are also used to handling the data analysis. One of the most popular normalization techniques is called maxima and minima normalization (Teixeria and Stephany, 2013; Rathod *et al.,* 2018; Rajesh and Karthikeyan, 2019b).

MATERIALS AND METHODS

Linear Regression Model

Regression analysis is a statistical tool to launch a relationship between two or more variables. One of the variables is named as predictor variable which means value is collected *via* experiments. Another variable is named as response variable which means derived from the predictor. A non-linear relationship where the exponent of any variable is not equal to 1 creates a curve. The general mathematical equation for a linear regression is

$$y = a + bx \qquad \ldots(1)$$

Where y is the response variable, x is the predictor variable and a, b are constants which are called the coefficients.

Normalization

If the data is having different scales then these types of data is not suitable for fit in to the modeling equation. The max-min normalization is used to convert different scales of data into unique format from 0.1 to 0.9. In this case, 0.1 is a minimum value and 0.9 is a maximum (Rajesh and karthikeyan, 2018). After completing the normalizations, the dataset is comfortable one for simply fitting the data into the proposed models.

P. Rajesh

email: *rajeshdatamining@gmail.com*

[1*]Assistant Professor, PG Department of Computer Science, Government Arts College, C.Mutlur, Chidambaram, Tamil Nadu, India.

[2]Assistant Professor, PG Department of Computer Applications, Arignar Anna Government Arts College, Villupuram, Tamil Nadu, India.

P - ISSN 0973 - 9157
E- ISSN 2393 - 9249

Table 1. Time series data for weather metrics of Chennai

Date	High Temp	Low Temp	Humidity (%)	Wind Speed (km/h)	Cloud Cover (%)	Rain (%)	Wind Chill	UV
01-Jan	28	22	79	19	23	1	27	0
02-Jan	31	23	81	17	25	0	30	0
03-Jan	31	22	80	23	22	0	27	4
04-Jan	32	22	78	24	20	0	25	3
05-Jan	30	22	79	24	23	0	24	0
06-Jan	30	25	77	19	19	0	25	0
07-Jan	30	22	81	17	17	7	24	0
08-Jan	30	21	80	24	25	5	21	6
09-Jan	30	22	79	25	18	6	27	5
10-Jan	30	22	77	21	20	4	25	7
11-Jan	29	22	76	20	19	0	25	3
12-Jan	30	22	79	16	23	0	22	0
13-Jan	30	22	80	17	21	0	27	0
14-Jan	31	22	81	15	20	0	25	0
15-Jan	28	25	80	17	18	0	21	7
16-Jan	30	22	79	23	18	3	25	6
17-Jan	30	22	81	24	20	2	24	7
18-Jan	31	22	77	20	19	1	22	7
19-Jan	31	21	78	19	17	7	20	3
20-Jan	32	22	77	18	20	0	25	1
21-Jan	30	22	79	17	16	0	27	0
22-Jan	30	23	89	20	20	0	22	0
23-Jan	31	22	90	15	21	1	26	0
24-Jan	31	21	91	20	17	0	25	0
25-Jan	32	22	87	21	18	5	27	0
26-Jan	32	22	89	17	20	7	23	3
27-Jan	32	22	91	18	18	0	22	2
28-Jan	32	22	92	18	19	4	27	4
29-Jan	32	22	89	20	20	3	28	1
30-Jan	32	22	85	19	21	7	28	1
31-Jan	32	23	84	20	20	0	27	0

Table 2. Time series data for weather metrics of Delhi

Date	High Temp	Low Temp	Humidity (%)	Wind Speed (km/h)	Cloud Cover (%)	Rain (%)	Wind Chill	UV
01-Jan	22	9	77	17	15	0	27	3
02-Jan	23	9	77	17	17	5	26	5
03-Jan	24	9	79	16	20	4	30	4
04-Jan	24	8	80	15	21	0	27	7
05-Jan	23	8	80	17	20	7	30	3
06-Jan	23	9	80	17	21	0	25	7
07-Jan	23	10	80	21	15	7	25	0
08-Jan	23	10	77	20	20	8	27	0
09-Jan	23	9	74	17	21	7	25	0
10-Jan	21	9	75	17	17	8	25	0
11-Jan	23	9	81	17	19	11	27	0
12-Jan	21	12	75	15	18	6	26	1
13-Jan	18	13	74	16	17	7	26	2
14-Jan	18	10	80	20	19	6	27	1
15-Jan	18	9	78	21	22	6	25	4
16-Jan	12	10	79	23	19	0	25	3
17-Jan	18	9	83	16	21	0	26	5
18-Jan	18	10	85	20	18	0	26	7
19-Jan	12	9	81	17	21	6	25	9
20-Jan	13	8	81	19	22	5	27	1
21-Jan	16	5	80	27	19	7	24	2
22-Jan	15	8	75	16	18	1	24	1
23-Jan	15	9	78	18	19	7	25	4
24-Jan	17	8	73	17	18	6	19	3
25-Jan	14	5	74	27	21	7	18	5
26-Jan	12	8	77	23	25	8	24	7
27-Jan	13	9	79	21	18	4	25	9
28-Jan	11	9	77	24	18	0	27	7
29-Jan	14	7	79	17	19	0	28	6
30-Jan	10	6	77	20	21	0	28	3
31-Jan	14	5	78	18	25	2	28	2

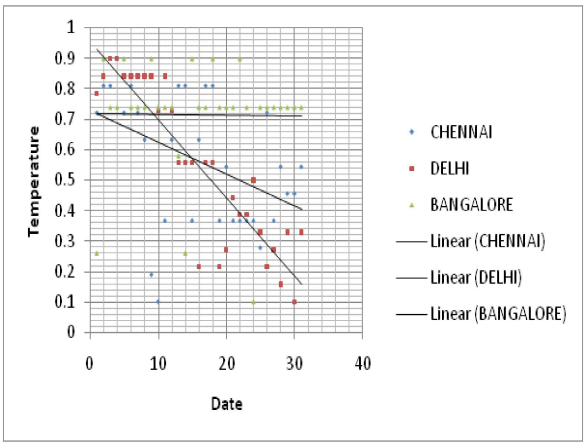

Fig. 1. Comparison of Temperature in the three cities during the study period. Regression lines are fit for the normalized data

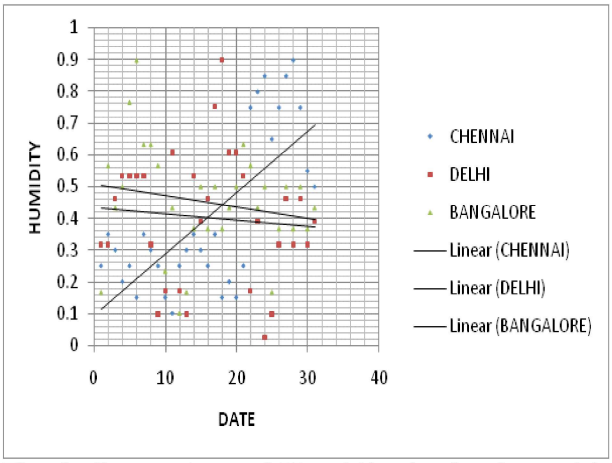

Fig. 2. Comparison of Humidity in the three cities during the study period. Regression lines are fit for the normalized data

Table 3. Time series data for weather metrics of Bangalore

Date	High Temp	Low Temp	Humidity (%)	Wind Speed (km/h)	Cloud Cover (%)	Rain (%)	Wind Chill	UV
01-Jan	27	22	74	16	18	7	21	1
02-Jan	31	22	80	18	19	6	23	5
03-Jan	30	21	78	17	19	7	24	0
04-Jan	30	24	79	27	18	8	27	9
05-Jan	31	23	83	23	17	4	24	3
06-Jan	30	21	85	21	19	0	24	5
07-Jan	30	21	81	24	22	0	25	0
08-Jan	30	22	81	17	19	0	19	9
09-Jan	31	22	80	20	21	2	18	3
10-Jan	30	22	75	18	18	1	24	5
11-Jan	30	23	78	19	21	0	25	4
12-Jan	30	22	73	17	22	1	27	7
13-Jan	29	23	74	23	19	2	28	3
14-Jan	27	24	77	24	18	1	28	7
15-Jan	31	24	79	24	19	6	28	0
16-Jan	30	25	77	19	18	0	27	0
17-Jan	30	22	79	17	21	0	30	0
18-Jan	31	20	77	24	25	0	23	0
19-Jan	30	20	78	25	18	0	24	0
20-Jan	30	21	79	21	18	0	27	1
21-Jan	30	22	81	20	19	1	24	2
22-Jan	31	21	80	16	18	1	24	1
23-Jan	30	21	78	17	19	7	25	4
24-Jan	26	24	79	15	18	6	19	3
25-Jan	30	23	74	17	21	7	18	5
26-Jan	30	21	77	23	25	8	24	7
27-Jan	30	21	79	21	18	4	25	9
28-Jan	30	22	77	24	18	0	27	7
29-Jan	30	22	79	17	19	0	28	6
30-Jan	30	22	77	20	21	0	28	3
31-Jan	30	23	78	18	25	2	28	2

Table 4. Normalized data using equation 2 for weather metrics of Chennai

Date	High Temp	Low Temp	Humidity (%)	Wind Speed (km/h)	Cloud Cover (%)	Rain (%)	Wind Chill	UV
01-Jan	0.722	0.3	0.25	0.42	0.722	0.214	0.633	0.1
02-Jan	0.811	0.5	0.35	0.26	0.9	0.1	0.9	0.1
03-Jan	0.811	0.3	0.3	0.74	0.633	0.1	0.633	0.557
04-Jan	0.9	0.3	0.2	0.82	0.456	0.1	0.456	0.443
05-Jan	0.722	0.3	0.25	0.82	0.722	0.1	0.367	0.1
06-Jan	0.811	0.9	0.15	0.42	0.367	0.1	0.456	0.1
07-Jan	0.722	0.3	0.35	0.26	0.189	0.9	0.367	0.1
08-Jan	0.633	0.1	0.3	0.82	0.9	0.671	0.1	0.786
09-Jan	0.189	0.3	0.25	0.9	0.278	0.786	0.633	0.671
10-Jan	0.1	0.3	0.15	0.58	0.456	0.557	0.456	0.9
11-Jan	0.367	0.3	0.1	0.5	0.367	0.1	0.456	0.443
12-Jan	0.633	0.3	0.25	0.18	0.722	0.1	0.189	0.1
13-Jan	0.811	0.3	0.3	0.26	0.544	0.1	0.633	0.1
14-Jan	0.811	0.3	0.35	0.1	0.456	0.1	0.456	0.1
15-Jan	0.367	0.9	0.3	0.26	0.278	0.1	0.1	0.9
16-Jan	0.633	0.3	0.25	0.74	0.278	0.443	0.456	0.786
17-Jan	0.811	0.3	0.35	0.82	0.456	0.329	0.367	0.9
18-Jan	0.811	0.3	0.15	0.5	0.367	0.214	0.189	0.9
19-Jan	0.367	0.1	0.2	0.42	0.189	0.9	0.011	0.443
20-Jan	0.544	0.3	0.15	0.34	0.456	0.1	0.456	0.214
21-Jan	0.367	0.3	0.25	0.26	0.1	0.1	0.633	0.1
22-Jan	0.367	0.5	0.75	0.5	0.456	0.1	0.189	0.1
23-Jan	0.367	0.3	0.8	0.1	0.544	0.214	0.544	0.1
24-Jan	0.367	0.1	0.85	0.5	0.189	0.1	0.456	0.1
25-Jan	0.278	0.3	0.65	0.58	0.278	0.671	0.633	0.1
26-Jan	0.722	0.3	0.75	0.26	0.456	0.9	0.278	0.443
27-Jan	0.367	0.3	0.85	0.34	0.278	0.1	0.189	0.329
28-Jan	0.544	0.3	0.9	0.34	0.367	0.557	0.633	0.557
29-Jan	0.456	0.3	0.75	0.5	0.456	0.443	0.722	0.214
30-Jan	0.456	0.3	0.55	0.42	0.544	0.9	0.722	0.214
31-Jan	0.544	0.5	0.5	0.5	0.456	0.1	0.633	0.1

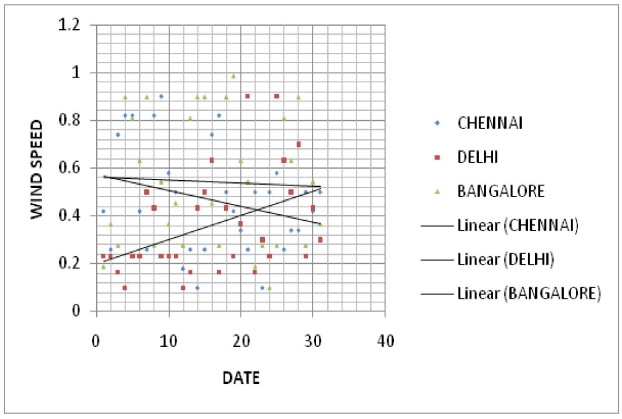

Fig. 3. Comparison of wind speed in the three cities during the study period. Regression lines are fit for the normalized data

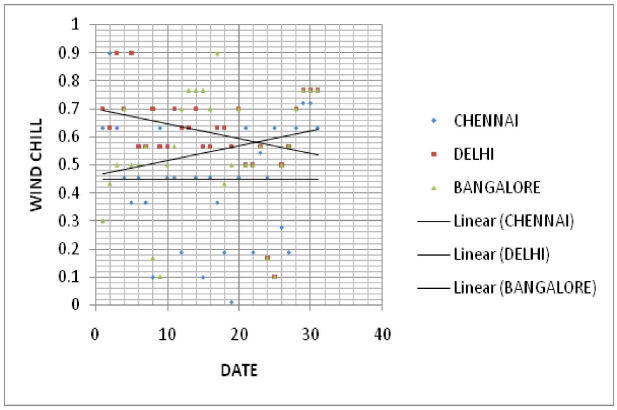

Fig. 4. Comparison of Wind chill in the three cities during the study period. Regression lines are fit for the normalized data

Table 5. Normalized data using equation 2 for weather metrics of Delhi

Date	High Temp	Low Temp	Humidity (%)	Wind Speed (km/h)	Cloud Cover (%)	Rain (%)	Wind Chill	UV
01-Jan	0.786	0.5	0.3182	0.2333	0.1	0.1	0.7	0.367
02-Jan	0.843	0.5	0.3182	0.2333	0.26	0.464	0.633	0.544
03-Jan	0.9	0.5	0.4636	0.1667	0.5	0.391	0.9	0.456
04-Jan	0.9	0.4	0.5364	0.1	0.58	0.1	0.7	0.722
05-Jan	0.843	0.4	0.5364	0.2333	0.5	0.609	0.9	0.367
06-Jan	0.843	0.5	0.5364	0.2333	0.58	0.1	0.567	0.722
07-Jan	0.843	0.6	0.5364	0.5	0.1	0.609	0.567	0.1
08-Jan	0.843	0.6	0.3182	0.4333	0.5	0.682	0.7	0.1
09-Jan	0.843	0.5	0.1	0.2333	0.58	0.609	0.567	0.1
10-Jan	0.729	0.5	0.1727	0.2333	0.26	0.682	0.567	0.1
11-Jan	0.843	0.5	0.6091	0.2333	0.42	0.9	0.7	0.1
12-Jan	0.729	0.8	0.1727	0.1	0.34	0.536	0.633	0.189
13-Jan	0.557	0.9	0.1	0.1667	0.26	0.609	0.633	0.278
14-Jan	0.557	0.6	0.5364	0.4333	0.42	0.536	0.7	0.189
15-Jan	0.557	0.5	0.3909	0.5	0.66	0.536	0.567	0.456
16-Jan	0.214	0.6	0.4636	0.6333	0.42	0.1	0.567	0.367
17-Jan	0.557	0.5	0.7545	0.1667	0.58	0.1	0.633	0.544
18-Jan	0.557	0.6	0.9	0.4333	0.34	0.1	0.633	0.722
19-Jan	0.214	0.5	0.6091	0.2333	0.58	0.536	0.567	0.9
20-Jan	0.271	0.4	0.6091	0.3667	0.66	0.464	0.7	0.189
21-Jan	0.443	0.1	0.5364	0.9	0.42	0.609	0.5	0.278
22-Jan	0.386	0.4	0.1727	0.1667	0.34	0.173	0.5	0.189
23-Jan	0.386	0.5	0.3909	0.3	0.42	0.609	0.567	0.456
24-Jan	0.5	0.4	0.0273	0.2333	0.34	0.536	0.167	0.367
25-Jan	0.329	0.1	0.1	0.9	0.58	0.609	0.1	0.544
26-Jan	0.214	0.4	0.3182	0.6333	0.9	0.682	0.5	0.722
27-Jan	0.271	0.5	0.4636	0.5	0.34	0.391	0.567	0.9
28-Jan	0.157	0.5	0.3182	0.7	0.34	0.1	0.7	0.722
29-Jan	0.329	0.3	0.4636	0.2333	0.42	0.1	0.767	0.633
30-Jan	0.1	0.2	0.3182	0.4333	0.58	0.1	0.767	0.367
31-Jan	0.329	0.1	0.3909	0.3	0.9	0.246	0.767	0.278

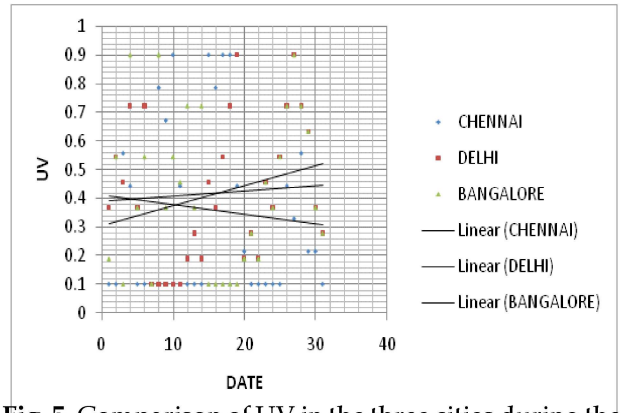

Fig. 5. Comparison of UV in the three cities during the study period. Regression lines are fit for the normalized data

Table 6. Normalized data using equation 2 for weather metrics of Bangalore

Date	High Temp	Low Temp	Humidity (%)	Wind Speed (km/h)	Cloud Cover (%)	Rain (%)	Wind Chill	UV
01-Jan	0.26	0.42	0.1667	0.1889	0.2	0.8	0.3	0.189
02-Jan	0.9	0.42	0.5667	0.3667	0.3	0.7	0.433	0.544
03-Jan	0.74	0.26	0.4333	0.2778	0.3	0.8	0.5	0.1
04-Jan	0.74	0.74	0.5	1.1667	0.2	0.9	0.7	0.9
05-Jan	0.9	0.58	0.7667	0.8111	0.1	0.5	0.5	0.367
06-Jan	0.74	0.26	0.9	0.6333	0.3	0.1	0.5	0.544
07-Jan	0.74	0.26	0.6333	0.9	0.6	0.1	0.567	0.1
08-Jan	0.74	0.42	0.6333	0.2778	0.3	0.1	0.167	0.9
09-Jan	0.9	0.42	0.5667	0.5444	0.5	0.3	0.1	0.367
10-Jan	0.74	0.42	0.2333	0.3667	0.2	0.2	0.5	0.544
11-Jan	0.74	0.58	0.4333	0.4556	0.5	0.1	0.567	0.456
12-Jan	0.74	0.42	0.1	0.2778	0.6	0.2	0.7	0.722
13-Jan	0.58	0.58	0.1667	0.8111	0.3	0.3	0.767	0.367
14-Jan	0.26	0.74	0.3667	0.9	0.2	0.2	0.767	0.722
15-Jan	0.9	0.74	0.5	0.9	0.3	0.7	0.767	0.1
16-Jan	0.74	0.9	0.3667	0.4556	0.2	0.1	0.7	0.1
17-Jan	0.74	0.42	0.5	0.2778	0.5	0.1	0.9	0.1
18-Jan	0.9	0.1	0.3667	0.9	0.9	0.1	0.433	0.1
19-Jan	0.74	0.1	0.4333	0.9889	0.2	0.1	0.5	0.1
20-Jan	0.74	0.26	0.5	0.6333	0.2	0.1	0.7	0.189
21-Jan	0.74	0.42	0.6333	0.5444	0.3	0.2	0.5	0.278
22-Jan	0.9	0.26	0.5667	0.1889	0.2	0.2	0.5	0.189
23-Jan	0.74	0.26	0.4333	0.2778	0.3	0.8	0.567	0.456
24-Jan	0.1	0.74	0.5	0.1	0.2	0.7	0.167	0.367
25-Jan	0.74	0.58	0.1667	0.2778	0.5	0.8	0.1	0.544
26-Jan	0.74	0.26	0.3667	0.8111	0.9	0.9	0.5	0.722
27-Jan	0.74	0.26	0.5	0.6333	0.2	0.5	0.567	0.9
28-Jan	0.74	0.42	0.3667	0.9	0.2	0.1	0.7	0.722
29-Jan	0.74	0.42	0.5	0.2778	0.3	0.1	0.767	0.633
30-Jan	0.74	0.42	0.3667	0.5444	0.5	0.1	0.767	0.367
31-Jan	0.74	0.58	0.4333	0.3667	0.9	0.3	0.767	0.278

$$X' = a + \frac{(X - X_{\min})(b - a)}{X_{\max} - X_{\min}} \qquad \ldots(2)$$

In equation (2), for 'X', Corresponding i^{th} value, X_{min} means minimum value in corresponding column, X_{max} means maximum value in corresponding column, 'a' means minimum value and 'b' means maximum value. The same technique is used in various researches for converting different scales of data into similar format (Serrano, 2019; Rajesh *et al*. 2019; Rajesh 2020) .

RESULTS AND DISCUSSION

The weather data in 2020 of Chennai, Delhi and Bangalore including various metrics namely, data on the weather details, highest temperature for the

Table 7. Comparison of high temperature for Chennai, Delhi and Bangalore

Chennai High Temp.	Delhi High Temp.	Bangalore High Temp.
0.7222	0.7857	0.26
0.8111	0.8429	0.9
0.8111	0.9	0.74
0.9	0.9	0.74
0.7222	0.8429	0.9
0.8111	0.8429	0.74
0.7222	0.8429	0.74
0.6333	0.8429	0.74
0.1889	0.8429	0.9
0.1	0.7286	0.74
0.3667	0.8429	0.74
0.6333	0.7286	0.74
0.8111	0.5571	0.58
0.8111	0.5571	0.26
0.3667	0.5571	0.9
0.6333	0.2143	0.74
0.8111	0.5571	0.74
0.8111	0.5571	0.9
0.3667	0.2143	0.74
0.5444	0.2714	0.74
0.3667	0.4429	0.74
0.3667	0.3857	0.9
0.3667	0.3857	0.74
0.3667	0.5	0.1
0.2778	0.3286	0.74
0.7222	0.2143	0.74
0.3667	0.2714	0.74
0.5444	0.1571	0.74
0.4556	0.3286	0.74
0.4556	0.1	0.74
0.5444	0.3286	0.74

Table 8. Comparison of humidity of Chennai, Delhi, and Bangalore

Chennai Humidity	Delhi Humidity	Bangalore Humidity
0.25	0.3182	0.1667
0.35	0.3182	0.5667
0.3	0.4636	0.4333
0.2	0.5364	0.5
0.25	0.5364	0.7667
0.15	0.5364	0.9
0.35	0.5364	0.6333
0.3	0.3182	0.6333
0.25	0.1	0.5667
0.15	0.1727	0.2333
0.1	0.6091	0.4333
0.25	0.1727	0.1
0.3	0.1	0.1667
0.35	0.5364	0.3667
0.3	0.3909	0.5
0.25	0.4636	0.3667
0.35	0.7545	0.5
0.15	0.9	0.3667
0.2	0.6091	0.4333
0.15	0.6091	0.5
0.25	0.5364	0.6333
0.75	0.1727	0.5667
0.8	0.3909	0.4333
0.85	0.0273	0.5
0.65	0.1	0.1667
0.75	0.3182	0.3667
0.85	0.4636	0.5
0.9	0.3182	0.3667
0.75	0.4636	0.5
0.55	0.3182	0.3667
0.5	0.3909	0.4333

Table 9. Comparison of wind speed of Chennai, Delhi and Bangalore

Chennai Wind Speed	Delhi Wind Speed	Bangalore Wind Speed
0.42	0.2333	0.1889
0.26	0.2333	0.3667
0.74	0.1667	0.2778
0.82	0.1	0.9
0.82	0.2333	0.8111
0.42	0.2333	0.6333
0.26	0.5	0.9
0.82	0.4333	0.2778
0.9	0.2333	0.5444
0.58	0.2333	0.3667
0.5	0.2333	0.4556
0.18	0.1	0.2778
0.26	0.1667	0.8111
0.1	0.4333	0.9
0.26	0.5	0.9
0.74	0.6333	0.4556
0.82	0.1667	0.2778
0.5	0.4333	0.9
0.42	0.2333	0.9889
0.34	0.3667	0.6333
0.26	0.9	0.5444
0.5	0.1667	0.1889
0.1	0.3	0.2778
0.5	0.2333	0.1
0.58	0.9	0.2778
0.26	0.6333	0.8111
0.34	0.5	0.6333
0.34	0.7	0.9
0.5	0.2333	0.2778
0.42	0.4333	0.5444
0.5	0.3	0.3667

Table 10. Comparison of wind chill of Chennai, Delhi, and Bangalore

Chennai Wind Chill	Delhi Wind Chill	Bangalore Wind Chill
0.6333	0.7	0.3
0.9	0.6333	0.4333
0.6333	0.9	0.5
0.4556	0.7	0.7
0.3667	0.9	0.5
0.4556	0.5667	0.5
0.3667	0.5667	0.5667
0.1	0.7	0.1667
0.6333	0.5667	0.1
0.4556	0.5667	0.5
0.4556	0.7	0.5667
0.1889	0.6333	0.7
0.6333	0.6333	0.7667
0.4556	0.7	0.7667
0.1	0.5667	0.7667
0.4556	0.5667	0.7
0.3667	0.6333	0.9
0.1889	0.6333	0.4333
0.0111	0.5667	0.5
0.4556	0.7	0.7
0.6333	0.5	0.5
0.1889	0.5	0.5
0.5444	0.5667	0.5667
0.4556	0.1667	0.1667
0.6333	0.1	0.1
0.2778	0.5	0.5
0.1889	0.5667	0.5667
0.6333	0.7	0.7
0.7222	0.7667	0.7667
0.7222	0.7667	0.7667
0.6333	0.7667	0.7667

Table 11. Comparison of UV in Chennai, Delhi, and Bangalore

Chennai UV	Delhi UV	Bangalore UV
0.1	0.3667	0.1889
0.1	0.5444	0.5444
0.5571	0.4556	0.1
0.4429	0.7222	0.9
0.1	0.3667	0.3667
0.1	0.7222	0.5444
0.1	0.1	0.1
0.7857	0.1	0.9
0.6714	0.1	0.3667
0.9	0.1	0.5444
0.4429	0.1	0.4556
0.1	0.1889	0.7222
0.1	0.2778	0.3667
0.1	0.1889	0.7222
0.9	0.4556	0.1
0.7857	0.3667	0.1
0.9	0.5444	0.1
0.9	0.7222	0.1
0.4429	0.9	0.1
0.2143	0.1889	0.1889
0.1	0.2778	0.2778
0.1	0.1889	0.1889
0.1	0.4556	0.4556
0.1	0.3667	0.3667
0.1	0.5444	0.5444
0.4429	0.7222	0.7222
0.3286	0.9	0.9
0.5571	0.7222	0.7222
0.2143	0.6333	0.6333
0.2143	0.3667	0.3667
0.1	0.2778	0.2778

Tables 12.a-e Comparison of Temperature, Humidity, Wind Speed, Wind Chill and UV using Descriptive Statistics for Chennai, Delhi, and Bangalore

Table 12(a). Temperature

Descriptive Statistics	Chennai	Delhi	Bangalore
Mean	0.5616	0.5442	0.7141
Median	0.5444	0.5571	0.74
Mode	0.3667	0.8429	0.74
SD	0.2112	0.2518	0.1793

Table 12(b). Humidity

Descriptive Statistics	Chennai	Delhi	Bangalore
Mean	0.4048	0.4026	0.4505
Median	0.3	0.3909	0.4333
Mode	0.25	0.3182	0.5
SD	0.2425	0.1967	0.1691

Table 12(c). Wind Speed

Descriptive Statistics	Chennai	Delhi	Bangalore
Mean	0.4665	0.3602	0.5416
Median	0.42	0.2333	0.5444
Mode	0.26	0.2333	0.2778
SD	0.2189	0.2077	0.2652

Table 12(d). Wind Chill

Descriptive Statistics	Chennai	Delhi	Bangalore
Mean	0.4498	0.614	0.5473
Median	0.4556	0.6333	0.5667
Mode	0.6333	0.5667	0.5
SD	0.2084	0.1592	0.2036

Table 12(e). Ultraviolet (UV)

Descriptive Statistics	Chennai	Delhi	Bangalore
Mean	0.3581	0.4183	0.4183
Median	0.2143	0.3667	0.3667
Mode	0.1	0.3667	0.1
SD	0.2927	0.2381	0.2561

particular day, lowest temperature for the day, humidity values in percentage, wind speed metrics in km/h, cloud covering status in percentage, rainfal details in percentage, sunrise and sunset metrics also represented in the form of time scale and UV means ultraviolet range in numbers are given in tables 1-3, respectively. The 11 metrics in tables 1-3 are in different scales and so not suitable for interpretation. so, by using equation 2 different scale of data were converted into similar format ranging from 0.1 to 0.9 and presented in tables 4-6.

Comparison of normalized weather parameters data of Chennai, Delhi and Bangalore are given in tables 7-12. Descriptive statistics of weather parameters of Chennai, Delhi and Bangalore are given in tables 12a-e. Comparison of weather parameters by regression lines fit to the normalized data are shown in figures 1-5. The temperature almost same in Bangalore during the study period, while it showed opposing trends with regard to Delhi and Chennai (Fig.!). While the trend in humidity was similar in Chennai and Bangalore, it showed an increasing trend in Delhi (Fig.2). The wind speed and wind chill remained almost similar in Delhi, but showing opposite trends in Chennai and Bangalore (Figs. 3 and 4). The UV showed increasing trends in Bangalore and

Delhi, while a decreasing trend was found in Chennai (Fig.5).

CONCLUSIONS

In this paper we explained how to use normalized data and descriptive statistics to compare the weather conditions of three major cities. Development of such regression models of this study will be useful to predict future weather conditions.

REFERENCES

Bocca, F. F., Henrique, L. and Rodrigues, A., 2016. The effect of tuning, feature engineering and feature selection in data mining applied to rainfed sugarcane yield modelling. *Comput. Electron. Agric.*, 128: 67–76. https://doi.org/10.1016/j.compag.2016.08.015

Han Jiawei , Micheline Kamber and Jian Pei, 2011. Data Mining: Concepts and Techniques. The Morgan Kaufmann Series in Data Management Systems, Morgan Kaufmann Publishers.Waltham, USA.

Rajesh, P. and Karthikeyan, M., 2018. Predication of Labour Demand in Agriculture Based on Comparative Study of Different Data Using Data Mining and Stochastic Approach. *Int. J. Eng. Sci. Invention*, 2: 86-97.

Rajesh, P. and Karthikeyan, M. and R. Arulpavai, R., 2019. Data Mining Algorithm to predict the Factors for Agricultural Development using Stochastic Model. *IJRTE*, 8(3): 2713-2717. https://doi.org/10.35940/ijrte.C4963.098319

Rajesh, P. and Karthikeyan, M., 2019. Data Mining Approaches to Predict the Factors that Affect the Groundwater Level using Stochastic Model. *AIP Conf. Proc.* 2177: 020079-1–020079-11. https://doi.org/10.1063/1.5135254

Rajesh, P. and Karthikeyan, M., 2019a. Data Mining Approaches to Predict the Factors that Affect the Agriculture Growth using Stochastic Model. *Int. J. Comput. Sci. Eng*. 7(4): 18-23. https://doi.org/10.1063/1.5135254

Rajesh, P. and M. Karthikeyan, 2019b, "Data Assimilation of Gross Domestic Product (GDP) in India using Stochastic Data Mining Approach", *J. Comput. Theor. Nanos.*, Vol. 16(4), pp. 1478–1484. https://doi.org/10.1166/jctn.2019.8061

Rajesh, P., 2020. Developing Data Mining Based Stochastic Model for Investigate the Parameters for Increasing Wheat and Groundnut Productions. *Adalya.*, 9(2): 319-327.

Rathod, S. S., Singh, K. N., Patil, S. G., Naik, R. H., Ray, M. and Meena, V. S. 2018. Modeling and forecasting of oilseed production of India through artificial intelligence techniques. *Indian J. Agric. Sci.* 88(1): 22–27.

Serrano, J. M., Delgado, G., Aranda, V., Calero, J. and Vila, M. A., 2009. Using fuzzy data mining to evaluate survey data from olive grove cultivation. *Comput. Electron. Agric.*, 65(1): 99–113. https://doi.org/10.1016/j.compag.2008.08.001

Teixeira de Lima, G. R. and Stephany, S. 2013. A new classification approach for detecting severe weather patterns. *Comput. Geosci.*, 57: 158–165. https://doi.org/10.1016/j.cageo.2013.04.016