

**WEATHER FLOW PATTERNS VISUALIZATION USING DATA MINING TECHNIQUE WITH EMPHASIS ON RAIN FALL.**

**OLAGUNJU MUKAILA (MCPN, MIEEE)**  
Department Of Computer Science,  
Kwara State Polytechnic, Ilorin,  
Nigeria.  
olamukaila@yahoo.com

**TAIWO A. O.**  
Department of Mathematics,  
University Of Ilorin,  
Ilorin, Nigeria.

**MABAYOJE M.A. (MCPN, MBAN)**  
Department of Computer Science,  
University of Ilorin,  
Ilorin-Nigeria.  
mabayoje01@gmail.com  
mabayoje.mak@unilorin.edu.ng

**AMJEN A.O.**  
Department of Computer Science,  
University of Ilorin,  
Ilorin-Nigeria.

**ADEBAYO OLAWALE S.**  
Department of Cyber Security Science,  
School of ICT, FCT Minn,  
olawaldan@yahoo.com

**ABSTRACT**

The essence of this paper is to study the phenomenon which may be too slow or too fast to be predicted.

Data mining technique with emphasis on visualization is used to forecast the rainfall pattern of Atlanta in USA with MATLAB program, the rainfall pattern can easily be monitor and access for over a period of time. This enable the regular planning to be carry out especially where rainfall is the major sources of water.

**Keywords:** Weather, Visualization, Data-mining, Rainfall, Forecast.

**1.0 INTRODUCTION**

Data mining according to (David, L.O. and Dutsay, D. 2001) refers to the analysis of the huge quantity of data that are stored in computers.

Also, H. Havenstein 2006, defined data mining as the process of analyzing data from different perspectives and summarizing what into useful information.

According to the Gartner Group, data mining is the process of discovering meaningful new

conclusion, patterns and trends arising through large amount of data mined in organisations, using pattern recognition techniques as well as statistical and mathematical techniques.

- 1.1 Data mining has a lot of applications according to (H. Tayssir (2006), which include:
- I. In politics, it can be used for identification of potential voters in an election.
  - II. In the banking system, it can also be used to detect fraud in the system.
  - III. In the field of medicine, data mining can be used for diagnosis purposes.
  - IV. In telecommunication, it can be used to determine user's behaviors.
  - V. In a manufacturing firm, it is used in quality control.
  - VI. In terrorist detection and related offences.

**2.0 MATERIALS AND METHODS.**

The material used in this paper is the rainfall data for Atlanta Georgia in United States of America. The used data is given in figure 1 below. The method used is one of the data mining methods, which is CRISP-DM (Cross Industry Standard and Process for Data Mining). CRISP-DM consists of six phases.

According to Daniel T. (2004) which include:

- I. Business / research understanding
- II. Data understanding phase
- III. Data preparation phase
- IV. Modeling phase
- V. Evaluation phase

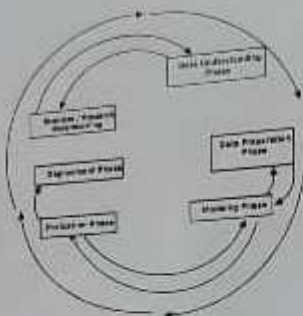


Fig 1. CRISP-DM CYCLE

The CRISP-DM is used to solve the current problems. The CRISP-DM is to provide a solution to the above problem. As follows:

1. **Business understanding:** This is the first stage in solving data mining problem. The essence of this stage is to approximate project objectives and to predict the rainfall pattern and trends. This stage is very important in the sense that it is the stage at which the objectives of the project are being defined.
2. **Data collection:** This is the second stage in CRISP-DM, and it involves the collection of data. The data collections for this paper are given in table 1.
3. **Data preparation:** the data used for this project are arranged in such a way that it was approximated in order to give the accurate result.
4. **Modeling Development:** The table 1 given below.

1999	2.5
2000	4.0
2001	2.5
2002	1.0
2003	1.0
2004	2.0
2005	2.2
2006	1.1
2007	3.0
2008	2.8
2009	2.0

3.0 DISCUS  
 Table 1 is the data used in this paper. The data

Fig 2. Fall

*Weather Flow Patterns Visualization Using Data Mining Technique with Emphasis on Rain Fall*

	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
1999	5.24	2.57	1.77	1.19	4.42	7.42	5.57	1.26	4.13	2.41	1.34	2.27
2000	4.89	3.76	2.67	1.75	1.01	3.31	2.7	4.02	4.57	0.40	5.02	2.01
2001	3.77	2.61	9.03	4.7	3.31	6.78	3.54	1.03	7.13	0.79	0.47	1.32
2002	4.25	2.94	3.40	1.80	3.52	2.41	2.53	5.77	6.23	5.94	5.30	5.23
2003	3	3.81	7.08	3.44	3.86	7.34	5.32	3.42	2.41	3.45	4.17	1.29
2004	2.84	4.6	1.04	2.8	1.58	5.80	2.1	3.03	13.83	1.13	7.24	4.82
2005	1.77	1.18	1.49	4.24	1.98	2.85	14.83	6.21	0.07	1.98	2.92	3.02
2006	5.1	1.5	2.52	1.01	1.60	5.8	1.21	6.81	2.11	1.04	4.25	1.03
2007	3.08	2.43	1.31	1.73	1.15	3.88	1.65	3.01	2.12	2.47	0.06	4.78
2008	2.42	4.88	1.17	1.22	1.1	0.55	7.17	1.77	0.73	3.10	2.04	4.30
2009	1.80	1.7	0.13	1.10	4.59	1.14	5.02	5.24	0.74	0.11	0	0

**3.0. DISCUSSION OF THE RESULTS**

Table 1 is the source of data for the analysis in this paper. The data is summarized as follows.



Fig 2: Falling Pattern and Trend of Rainfall (in Whanta) for January (1999-2009)

continuous process and trends using through large amount of data stored in repositories, using pattern recognition techniques as well as statistical and mathematical techniques.

3.1 Data mining has a lot of applications according to (H. Bayens *et al.*, 2006), which include:

- I. In politics, it can be used for identification of potential voters in an election.
- II. In the banking system, it can also be used to detect fraud in the system.
- III. In the field of medicine, data mining can be used for diagnosis purposes.
- IV. In Telecommunication, it can be used to determine user's behavior.
- V. In a manufacturing firm, it is used in quality control.
- VI. In training decisions and related activities.

### 3.0 MATERIALS AND METHODS

The material used in this paper is the rainfall data for Atlanta, Georgia in United States of America. The used data is shown in figure 1 below. The method used is one of the data mining methods, which is CRISP-DM (Cross-industry standard process for data mining). CRISP-DM consists of six phases. According to (Daniel T. L. 2000) which include:

- I. Business / research understanding
- II. Data understanding phase
- III. Deployment phase
- IV. Data preparation phase
- V. Modeling phase
- VI. Evaluation phase

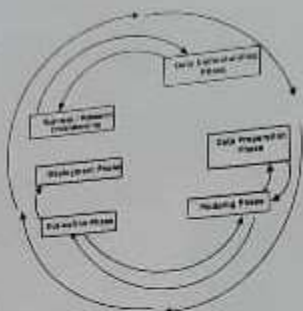


Fig. 1: CRISP-DM CYCLE

The CRISP-DM is used to solve the chosen problems. The CRISP-DM is to provide a solution to the above problem. As follow:

1. **Business understanding:** This is the first stage in solving data mining problem. The essence of this stage is to determine project objectives and to predict the rainfall pattern and trends. This stage is very important in the sense that it is the stage at which the objectives of the project are being defined.

2. **Data collection:** This is the second stage in CRISP DM, and it involves the collection of data. The data collection for this paper are given in table 1.

3. **Data preparation:** the data used for this project are arranged in such a way that it was approximated in order to give the accurate result.

4. **Modeling Development:** The table 1 given below:

Year  
Month

1999	1.4
2000	1.6
2001	2.1
2002	2.1
2003	1
2004	1.4
2005	2.0
2006	3.1
2007	1.2
2008	1.4
2009	2.0

### 3.0 DISCUS

Table 1 is the data  
for paper. The d

Fig 2: Fall

complexity, patterns, and trends using through large amount of data stored in repositories, using pattern recognition technologies as well as statistical and mathematical techniques.

1.3 Data mining has a lot of applications according to (H. Havens et al, 2006), which include:

- I. In politics, it can be used for identification of potential Voters in an election.
- II. In the banking system, it can also be used to detect fraud in the system.
- III. In the field of medicine, data mining can be used for diagnosis purposes.
- IV. In Telecommunication, it can be used to determine user's behavior.
- V. In a manufacturing firm, it is used to quality control.
- VI. In Internet Services and related activities.

2.0 MATERIALS AND METHODS.

The material used in this paper is the rainfall data for Adama, Georgia in United States of America. The used data is shown in figure 1 below. The method used is one of the data mining methods, which is CRISP-DM (Cross Industry Standard Process for Data Mining). CRISP-DM consists of six stages:

- According to (Chen, 2004) which include:
- I. Business / research understanding
  - II. Data understanding phase
  - III. Data preparation phase
  - IV. Modeling phase
  - V. Evaluation phase

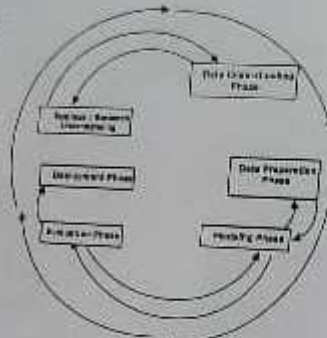


Fig. 1. CRISP-DM CYCLE

The CRISP-DM is used to solve the chosen problem. The CRISP-DM is to provide a solution to the above problem, as follows:

1. **Business understanding:** This is the first stage in solving data mining problem. The essence of this stage is to document project objectives and to predict the initial problem and trends. This stage is very important in the sense that it is the stage at which the objectives of the project are being defined.
2. **Data collection:** This is the second stage in CRISP-DM, and it involves the collection of data. The data collections for this paper are given in table 1.
3. **Data preparation:** the data used for this project are arranged in such a way that it was approximated in order to give the accurate result.
4. **Modeling Development:** The table 1 given below.

14

200  
2001  
2002  
2003  
2004  
2005  
2006  
2007  
2008  
2009

3.0 D  
Table 1 is  
this paper

Fig. 2: Fu

The fig 2 above was generated with MATLAB program when all years i.e. 1999 to 2009 was plotted against all January between 1999 to 2009 it clearly shown January 2006 has the highest rainfall and the rainfall is irregular.



Fig 3: Falling Pattern and Trend of Rainfall for Atlanta for February (1999-2009)

The fig 3 above was generated by plotted all the Februarys against all the years from 1999 to 2009. The fig 3 generated clearly shown that the highest rainfall was recorded between 2005 and 2006.

The Fig 3 clearly shown that the rainfall pattern is irregular.



Fig 4: Falling Pattern and Trend of Rainfall for Atlanta for March (1999-2009)

4.1.

*Weather Flow Pattern Visualization Using Data Mining Technique with Emphasis on Rain Fall.*

was  
highest

The Fig. 4 above, was generated also, by plotted all the months against all the years from 1999 to 2009. March 2001 has the highest rainfall. With the Fig. 4 generated above the rainfall pattern is irregular

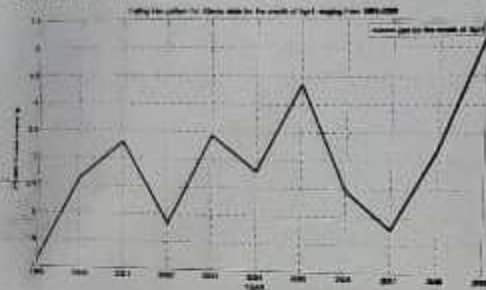


Fig 5: Falling Pattern and Trend of Rainfall for Atlanta for Aprils (1999-2009)

was  
and

The Fig. 5 above was generated by plotted all the years between 1999 to 2009 against all the April between 1999 to 2009. Year 2009 has the highest amount of rainfall and the rainfall is irregular.

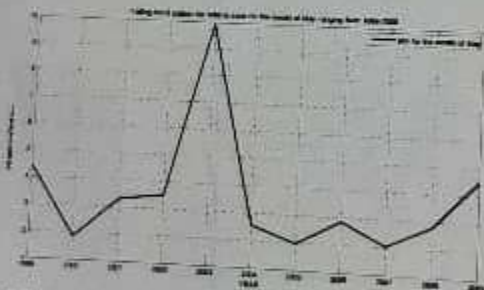


Fig 6: Falling Pattern and Trend of Rainfall for Atlanta for Mays (1999-2009)

Fig. 8 above which were generated with MATLAB clearly shows that the month of July in 2005 has the highest rainfall and it is irregular figure type of rainfall.



Fig 9: Falling Pattern and Trend of Rainfall for Atlanta for Augusts (1999-2009)

Fig. 9 above clearly shows that the rainfall pattern is not regular and month of July in year 2006 has the highest rainfall.



Fig 10: Falling Pattern and Trend of Rainfall for Atlanta for Septembers (1999-2009)



Fig. 6 above was generated by Jentical all on May between 1999 to 2009 against all the years i.e. 1999 to 2009. Year 2003 has the highest amount of rainfall and the rainfall pattern is irregular.



Fig. 7: Falling Pattern and Trend of Rainfall for Atlanta for June (1999-2009)  
 Fig. 7 was generated all the month of June against all the year i.e. 1999 to 2009. 2003 has the highest amount of rainfall.  
 The Fig. 7 clearly shows that the rainfall is irregular type of rainfall.



Fig. 8 - Falling Pattern and Trend of Rainfall for Atlanta for July (1999-2009)

In fig. 10 above, month of September 2004 has the highest rainfall. The rainfall is irregular



Fig. 11: Falling Pattern and Trend of Rainfall for Atlanta for October (1999-2009)

Fig. 11 above shows that the rainfall pattern is irregular.



Fig. 12: Falling Pattern and Trend of Rainfall for Atlanta for November (1999-2009)

Fig. 12 above also takes the same pattern and the rainfall is irregular.

*Weather Flow Patterns Visualization Using Data Mining Technique With Emphasis On Rain Fall*

**Evaluation:** This is the final stage of Crisp-DM method, based on the results generated above, it is easy to predict the rainfall pattern and trends for each month as irregular and this will assist a lot in weather forecast for Atlanta.

#### 4.0 RECOMMENDATIONS AND CONCLUSION

Data mining with visualization techniques greatly helped to discover that the rainfall pattern in the Atlanta is irregular and the following recommendations were made:

1. In areas where water is highly needed, natural sources should not be depended on.
2. Irrigation practices is highly needed.
3. Planting of crops which do not need much water.
4. Construction of dams.
5. Use of fertilizer for highly grow of crops and plants.
6. Avoid carbon emission in order to prevent pollution and weather climates.

#### REFERENCE

- Amari, Y.A (2006): aviation research for science and management students, publishing A: IKK, International, Kano, Nigeria.
- Brian, P.H and Daniel, T.V (2007): Essential midday for engineers and science, 3<sup>rd</sup> edition, published by Elsevier new - York.
- Daniel, T.V (2004): Discovering Knowledge in Data Mining, Published by McGraw-Hill, New York.
- Davis, L.O and Dimpson, D (2004): Advance Data Mining Technique, published by Springer-Verlag, Berlin.
- Hassanain, H (2006): Effort to help Determine electron activities Tailored device display data tools: COP expands reorganizing use computer world, published by McGraw-Hill, New York

- J. Dancos, W.J, Thomas, A.S, Young, A.O, (1974): *Mechanics of Fluid*, published by Edward Arnold LTD, Great Britain.
- Jim, X, Cheng (2003): *Graphical Output Software Tools*, Springer-Verlag New York.
- Onaleye, J.O, Anju, L.B, Ogunju, M, Ameen, A.O, and Rauf, K (2009): Embedded Confusion Matrix in Fuzzy Logic for System Appraisal, *International Journal of Physical Science*, PP 18 - 25.
- Patrick H.O and David N. (1990): "Introduction to convective Heat Transfer Analysis" Mc Graw-Hill Publishers.
- Simon, H.A (1969): *The Sciences of Artificial*, Published MIT Press, London.
- Turkle, L. R (2001): *Visualization of Quantitative Information*, (2nd ed), Graphics Press, Chesham.
- Webster's Ninth, New Collegiate Dictionary: Merriam inc., United States Of America, Access 1997.