



BIG DATA IN CLIMATE CHANGE

Dr. Parul Agarwal

Department of Computer Science and Engineering Jamia Hamdard, New Delhi-62.

Dr. Siddhartha Sankar Biswas

Department of Computer Science and Engineering Jamia Hamdard, New Delhi-62.

ABSTRACT

Big Data often associated with E-Commerce and social media sites amuses us when associated with “Climate Change”. Climate Change has its adverse effects on every part of the world. Rainfall in dry regions, melting of glaciers resulting in rise of sea level, global warming, and deforestation are some of the alarming implications of it. These affects are huge, controversial, challenging yet important. What appears as one of the solutions is Big Data. Big Data comes as a solution in not just providing effective storage of data collected pertaining to disasters or its visualizations but has a potential of predictions and analysis as to how fast and where lie the sources of effect. In addition, big data and its analytical techniques can provide untapped opportunities and solutions to diversify our understanding and accelerate decision making which shall be vital to control the “Climate Change”. The main focus of this paper is to understand the various problems in using big data itself and its solutions by analyzing the methods and challenges that need to be addressed to make Big Data as a promising answer to the universal threat: Climate Change. The need of the hour is to save the Earth and mankind

KEYWORDS : Big Data, predictive analysis, global warming, GHG, 3 V's of Big data, Climate Change,

Introduction:

Climate Change is not just a phenomenon related to Polar Bear, rather deals with the changes of weather patterns over time. These changes might result due to man-made or natural influences. But its basically human influence which has resulted in natural ones. Humans have tremendously misused the environment through deforestation, changes in land use, wastage of natural resources, practices that have led to increase of emissions of CO₂, fossil fuel and Aerosols and Green House gases(GHG) causing ecological imbalance and rise in global temperatures.

The increase in global temperatures has resulted in extreme temperatures, floods, increase in sea levels, melting of glaciers, decline of ice sheets, adverse impact on food production and also resulted in several diseases. Several important information and its sources on climate change are available in successive reports of the Intergovernmental Panel on Climate Change (IPCC) (IPCC, 2001) created by the United Nations (U.N.) and the World Meteorological Organization in 1988. The data generated either as a result of the changes that take place in the environment (i.e. post changes) or as part of analysis which might predict the future, is huge and heterogeneous in nature. The ecologists, environmentalists are placed at various places on the Earth doing this task tenaciously to save the Earth. But, success lies in sharing this data and jointly working on it to synthesize this “Huge or Big Data” arising from scientific results of research projects. This is being adopted in Ecology (Chaudhary et. al., 2010).Its having a great impact on research and literature over the past few years.

Fig 1 from (Faghmous and Kumar, 2014) depicts the components of climate system and the interactions with each other. The original source of image is (IPCC, 2001).

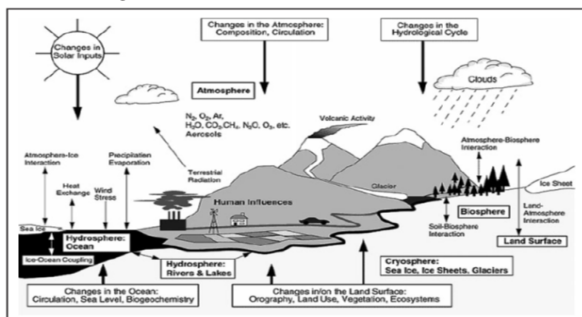


Fig. 1 : Shows the components of the global climate system and their interactions.

Big Data is data which is too huge and has been mainly associated with 3 V's. In (Chen, Mao and Liu, 2014; Grobelnick, 2012), the authors discuss the challenges and opportunities associated with Big Data owing to its characteristics in 3V's Model.

The 3 V's are Volume, Variety and Velocity. Volume refers to the amount of data generated which could be in terabytes(10¹²) or even petabytes(10¹⁵) bytes of data. Variety refers to the heterogeneous nature of data, owing to data being collected from multiple sources and its representation may vary. These could be in form of text, Images and so on. The Velocity refers to the fast changing data. Owing to its features, supporting technology like Cloud Computing can play a major role. This has been discussed in (Agarwal & Biswas, 2017).We can now very well connect how Big Data- a concept primarily associated with Social Networking Sites and E-Commerce also has its association with Climate Change. Thus the need is to channel and harness the benefits of big data and its analytical techniques to fight climate change.

Problems and Solutions

The major sources of climate data are: in situ, remote sense, paleo climatic and climate model output. These have been discussed in (Faghmous & Kumar, 2013). Each of these data sources is unique and faces its own challenges which can be accessed the paper. Data related to Climate science, in particular climate change has several characteristics. Several issues that need to be addressed are many. Proper understanding of the data is essential. We need to know how the data is being generated and its purpose as well as understanding the inherent biases in the data will enable the researchers to modify the results. Also, the data is fast changing and extremely dynamic in nature. So a need to club the static and dynamic data generated in streams is needed. Various instruments, satellites capture the climate data and the abundance of data is huge.

We know that the data would be heterogeneous data emerging from multiple sources and in multiple representation formats. Every data cannot be represented in its binary form, for example hurricane, rather could be geo spatial in nature too. Maybe the data representation could be fuzzy (Zadeh, 1965) too. Correct perspective of the analyst in case of a disaster like hurricane, volcano, drought is essential in drawing a correct conclusion.

In 2009, a team of Google reported in (Lazer, Kennedy, King & Vespignani, 2014) that flu related queries can be analyzed to detect

flu outbreaks much faster than the Centers for Disease Control and prevention. But this did not last long, Google made predictions which were bad.

Call detail records, Cell Phones, Social media, Twitter posts could be used to examine areas affected by earthquakes, generate storm warnings, predict droughts and many more.

Conclusion

Several work has been initiated but still a long way to go. Both in understanding the big data and its potentials and also reaping its benefits for combating climate change. But with respect to the environment the time is less as the effects of climate change are fast emerging. So, a need to generate actionable insights and depth analysis of big data is required to prevent the change. The researchers of both the domains need to collaboratively work in collecting, storing and analyzing the fast flowing data to predict the future.

REFERENCES

1. Agarwal, P., & Biswas, S.S.(2017). "Big Data on Cloud: A Review". International Journal of Advanced Research in Computer Science, Vol. 8(2), ISSN No:0976-5697.
2. Chaudhary, V.B., Walters, L., Bever, J.D., Hoeksema, J.D. & Wilson, G.W.T. (2010). Advancing synthetic ecology: a database system to facilitate complex ecological meta-analyses. Bull. Ecol. Soc. Am., in press.
3. Chen, M., Mao, S., Liu, Y. "Big Data: A Survey" published online in Springer Science + Business Media, New York, 2014.
4. Faghmous, J. H., & Kumar, V. Spatio – Temporal data mining for climate: Advances, Challenges and opportunities. In W. Chu, Ed., Data Mining and Knowledge Discovery for Big data, Springer, Heidelberg, Germany, 2013, pp 83-116.
5. Faghmous, J. H., & Kumar, V. (2014). A Big Data Guide to Understanding Climate Change: The Case for Theory-Guided Data Science. Big Data, 2(3), 155–163. <http://doi.org/10.1089/big.2014.0026>.
6. Grobelnik, M. 2014. Big data Tutorial. <http://Videlectures.net/eswc2012grobelnikbigdata/>. accessed [14.06.2017].
7. Intergovernmental Panel on Climate Change. Climate Change 2001: The Scientific Basis. Cambridge, UK: Cambridge University Press, 2001.
8. Lazer, D., Kennedy, R., King, G., Vespignani, A. The parable of Google Flu: Traps in Big Data Analysis. Science, 2014; 343: 1203-1205
9. Zadeh, L. 1965. Fuzzy Sets. Information and Control. Vol. 8, pp 338-353.