ISBN: 978-1-948012-17-1

Ethics and Information Technology (ETIT)

DOI: <http://doi.org/10.26480/et.02.2020.53.56>

COMPARATIVE ANALYSIS OF TWO MODIFIED REGRESSION ALGORITHMS DURING THE PREDICTION OF EVOLUTION OF ICE SHEET THICKNESS

Arnav Ghosh^a, Sayan Majumder^{b*}, Dipanwita Ghosh^c, Subhas Halder^d^a*Dream Institute of Technology, Thakurpukur, Kolkata 700104, India*^b*Gargi Memorial Institute of Technology, Baruipur, Kolkata 700144, India*^c*Maulana Abul Kalam Azad University of Technology, Kolkata 700064, India*^d*Gargi Memorial Institute of Technology, Baruipur, Kolkata 700144, India**Corresponding Author Email: sayanmajumder90@gmail.com

This is an open access article distributed under the Creative Commons Attribution License CC BY 4.0, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ARTICLE DETAILS

Article History:

Received 25 October 2020

Accepted 26 November 2020

Available online 03 December 2020

ABSTRACT

Though research of polar ice sheet has advanced recently, it is still unknown the recent evolution in this field. Nowadays, loss of ice mass from ice sheet of polar region is very much common issue due to global warming. Approximately, 269 billion tons of ice is melting every year since 2002, which rises the sea level by 20 feet, every year. In this paper, we have predicted the thickness of ice sheet in the year 2017, using two regression analysis techniques. We have applied regression techniques with method of least square (MLS) and least absolute deviation (LAD) to predict the evolution of ice sheet thickness, from which we have concluded a comparison between the results obtained from these two algorithms. The dataset used is from NASA. Python programming language is used here to find the result after which we can conclude the better regression technique.

KEYWORDS

MLS, LAD, Ice sheet, Regression.

1. INTRODUCTION

Huge uncertainties in future prediction of rise of sea level are due to melting of Ice sheet in the polar region, mainly Greenland. Not only global warming, but also dynamic discharge from marine causes the rises of temperature with irreversible loss of mass of ice sheet.

In Davos, a one day workshop was organized (Polar 2018) by the experts of ISMASS in Switzerland on 15th of June, 2018. The subject of that workshop was to discuss about rise of temperature and prediction of ice sheet (Hanna et al., 2013). A study summarized the details of the workshop held in Portland, USA in the year 2013 (Hanna et al., 2014). Here, we know about advancement of modelling of ice sheet prediction from last six years. Surface mass is also changed in SMB (Surface Mass Balance) due to glaciers terminated from marine (Hanna et al. 2016; Hanna et al., 2018). Others study also described the mapping of velocity of ice sheet (Rignot et al., 2014). We have to find the SMB (surface mass balance) of Greenland ice sheet (Rignot et al., 2019), where we can find the floating ice, which is melting day by day into water. Satellite is a good option to track these changes. Using these satellites, we can easily detect and locate the place. We can find mainly 3 satellites to measure the change of ice sheet-

- CryoSat
- EnvySat
- ERA & ICESat.

Correction algorithm was used to track the EnvySat data (Filament and Rémy, 2012). Coastal areas of Greenland are getting thinner day by day (Hanna et al., 2016; Majumder et al., 2019). The rate of mass change is also increasing. GrIS (Greenland ice sheet) mass loss nearly quadrupled during

the year 2002–2012 (Bevis et al., 2019). The sea level of Greenland ice sheet was nearly one and half time than Antarctic (Box et al., 2013; Box et al., 2018).

The difference between mass gain and mass loss of glaciers, every year is more than 15%, which causes 10cm sea level change. If we change the ice sheet, then process associate of GIA (Glacial Isostatic Adjustment) must be changed. We can find mainly two types of polar ice sheet data-

- Antarctica Ice Sheet
- Greenland Ice Sheet

The thickness of ice mass of Greenland changes very little per year, but we can found a lot of differences. The climate of Antarctica is much colder than Greenland. -30 degrees is the lowest temperature in northern region, whereas in the south section, this is -20 degrees. The mass of Greenland ice sheet is huge but Antarctica is still the biggest one. Single large ice sheet is responsible here to cover entire Greenland, while periphery is arranged with small glaciers.

In this paper, we have used the dataset of Greenland ice sheet from NASA, from where we have extracted 3 features namely, temperature, thickness and elevation and applied the techniques namely MLS and LAD to predict the evolution of thickness of ice sheet. At last, we have made a comparison between these two algorithms to find the best one.

In the regression model the goal is to get a linear relationship between data points. Here we have achieved linearity in the data using MLS and LAD regression model. The method of least squares (MLS) is mostly used regression tools for prediction models. Theoretical properties of least

Quick Response Code	This paper was presented at	Access this article online	
	International Conference on Contemporary Issues in Computing (ICCIC-2020) - Virtual IETE Sector V, Salt Lake, Kolkata From 25th-26th July 2020	Website: www.intelcomp-design.com	DOI: 10.26480/et.02.2020.53.56

square method have been extensively studied and are fully understood (Majumder et al., 2018). Despite its many properties, the LS estimate can be sensitive to outliers. LAD regression predicts the result smoothly even if outliers are present.

In Modified least-square regression model, the regression parameters are found by minimizing the sum of squared difference of the predicted value and observed values of a dependent variable.

In Modified Least absolute deviation regression model, instead of minimizing sum of square of residuals, we minimize the absolute value of the residuals. This provides more robust solution when outliers are present. Gradient descent is also used with both regression techniques so that we can find the loss function and predict the thickness of ice sheet. We can abbreviate the proposed algorithms as-MLS-Descent Algorithm, LAD-Descent Algorithm

Afterwards, we will get the proposed methodology section, where we will discuss about our two algorithms in brief step by step so that first, loss function must be calculated using regression and then gradient descent technique is also used. This part is followed by results and discussion section. In that section, our results will be discussed with proper diagrams and tables. For our dataset representation, we have included a graph where thickness of ice sheet is shown.

Then we must include a comparison table between two algorithms. This section will be followed by conclusion, where we can conclude our thought between method of least square and least absolute deviation regression (Majumder et al., 2020). The graph also represents the comparison or error estimation between these two algorithms. This error should help to predict the accurate mass of ice.

At last, references of the paper are enlisted with proper citation.

2. PROPOSED METHODOLOGY

2.1 MLS regression algorithm

Algorithm 1

Step 1: Plotting the given data of thickness and elevation.

Step 2: Mean normalization is used to get every feature into approximately $-1 \leq x_i \leq 1$ range,

$$x_i = (x_i - \mu_i) / s_i$$

Where, μ_i is the average value of x in training set & s_i is the standard deviation

Step 3: Find the cost function J ,

$$J = \left(\sum_{i=1}^n ((x_i \cdot \theta) - y_i)^2 \right) / (2 * m)$$

Where, m is the number of training examples and θ is the cost function coefficient.

Step 4: Gradient descent algorithm is used to minimize θ that will be used in predicting thickness of ice sheet.

$$\theta := \theta - \alpha \cdot \frac{\partial J}{\partial \theta}$$

α is the learning rate.

Step 5: Predicting the thickness of ice sheet

$$h = \theta \cdot x$$

Where x is the test set data.

Step 6: Standard error of estimation

$$err = \sqrt{\left(\frac{J}{m} \right)}$$

Finish: Plot the learning curve using cost J over number of iterations performed.

2.2 LAD regression algorithm

Algorithm 2

Step 1: Plotting the given data of thickness and elevation.

Step 2: Mean normalization is used to get every feature into approximately $-1 \leq x_i \leq 1$ range

$$x_i = (x_i - \mu_i) / s_i$$

Where, μ_i is the average value of x in training set & s_i is the standard deviation

Step 3: Find the cost function J

$$J = \left(\sum_{i=1}^n |(x_i \cdot \theta) - y_i| \right) / (2 * m)$$

Where m is the number of training examples and θ is the cost function coefficient.

Step 4: Gradient descent algorithm is used to minimize θ that will be used in predicting thickness of ice sheet.

$$\theta := \theta - \alpha \cdot \frac{\partial J}{\partial \theta}$$

Where, α is the learning rate.

Step 5: Predicting the thickness of ice sheet

$$h = \theta \cdot x$$

Where x is the test set data.

Step 6: Standard error of estimation

$$err = \sqrt{\left(\frac{J}{m} \right)}$$

Finish: Plot the learning curve using cost J over number of iterations performed.

Above 2 algorithms are executed for the dataset of Greenland ice sheet to predict the thickness of the ice sheet, in the year 2017 based upon the temperature and elevation.

From the given dataset, 1st we have normalized the dataset. To extract the features from the dataset. We have taken 2 features namely, temperature and elevation based on which thickness is predicted. Then we have calculated the cost function or loss function. Using Gradient Descent algorithm, the value of theta is minimized. Here, the value of theta is defined as the cost function coefficient. We, have taken 80% of our data as training set and rest is for test set. At last standard error has been estimated, from where we have concluded the result. Thickness of ice sheet is defined as h . X_i and Y_i are two variables as elevation and temperature. Depending upon these two, we have predicted the thickness of mass of Greenland ice.

MLS algorithm use least square method to find the cost function or loss function, where LAD algorithm uses the absolute deviation value to find this loss function. We have downloaded the dataset from <https://nsidc.org/greenland-today/>. The dataset was a record of 2017. We have found a graph of dataset as well as 2 graphs as our result with two regression algorithms. At last a detailed comparison is given to elaborate the picture of these two techniques.

3. RESULTS AND DISCUSSIONS

Figure 1: Dataset Representation- Elevation of ice sheet for 2017 vs Thickness of mass of ice

In the above figure, we can observe, the thickness of ice sheet along x axis, whereas elevation along y axis. As, the elevation is increasing, thickness is decreasing. Our dataset is represented by this picture.

Figure 2: Ice Sheet thickness prediction using MLS regression algorithm

From the above figure, we can predict, the thickness of ice sheet, that is loss function along y axis, whereas number of iterations along x axis. We can observe that, loss of mass is increasing with number of iterations are increasing.

Figure 3: Ice Sheet thickness prediction using LAD regression algorithm

From the above picture, we can predict, the thickness of ice sheet, that is cost function along y axis, whereas number of iterations along x axis. We can observe that, loss of mass is increasing with number of iterations are increasing.

Here, we have used least absolute deviation technique to find the loss function, whereas in previous case, we have used method of least square technique.

At last, we have concluded a detailed comparison between these two algorithms, which is enlisted in a table.

Table 1: Comparison between MLS and LAD Algorithms

Method of Least Square algorithm	Least Absolute Deviation Algorithm
1. Time taken – 350 sec.	1. Time taken- 270 sec.
2. Error estimated is higher	2. Error estimated is lower
3. Not robust due to presence of outlier	3. Outlier does not effect on this. So, this is more robust.
4. It is tougher to minimize regression coefficient in case of MLS.	4. It is easier to minimize regression coefficient in case of LAD.

Standard errors calculated by two algorithms are compared and we can observe that least absolute deviation algorithm gives better result as error is minimized in this case.

On the other hand, least square method can't minimize the error as good as other one. Prediction of thickness of ice is same in both of the cases but time taken by least square method is much higher than absolute deviation method. Absolute deviation method also can detect if outlier is present in case of large dataset, which is rarely possible for MLS.

Figure 4: Estimated Error for MLS Algorithm vs LAD Algorithm

standard error value:

2.6691762601240137

Theta computed from gradient descent:

[753.91617299 -75.18098339 -58.42795502]

Predicted thickness of ice for elevation of 1650ft, (using gradient descent):

1862.3660790442786

Figure 5: Predicted Error by MLS-Descent Algorithm

standard error value:

0.2858762083264863

Theta computed from gradient descent:

[753.91617299 -75.18098339 -58.42795502]

Predicted thickness of ice for elevation of 1650ft, (using gradient descent):

1862.3660790442786

Figure 6: Predicted Error by LAD-Descent Algorithm

4. CONCLUSION

We have predicted the loss of mass of ice sheet of Greenland in the year 2017 using two modified regression analysis algorithms- MLS-Descent Algorithm LAD-Descent Algorithm

The error value predicted by LAD regression algorithm is 2.669, whereas the error predicted by MLS algorithm was 0.286. Cost function or loss function is calculated using there two algorithms. After that gradient descent technique is used to predict the thickness of the ice sheet.

At last, we can say that our modified LAD-Descent algorithm has following advantages over MLS-Descent-

1. This algorithm gives more than 85% better result than MLS one.
2. LAD-Descent algorithm takes much less time to predict the thickness, than MLS Descent algorithm.
3. This algorithm has less computational overhead, so it is easier to compute regression coefficient in this algorithm.
4. LAD-Descent algorithm works better than MLS-Descent, even if outliers are present in the large dataset.

REFERENCES

- Box, J.E., 2013. Greenland ice sheet mass balance reconstruction. Part II: surface mass balance: 1840-2010. *J. Clim.* 26 (18), 6974-6989. <https://doi.org/10.1175/JCLI-D12-00518.1>.
- Box, J.E., Colgan, W.T., Wouters, B., Burgess, D.O., O'Neel, S., Thomson, L.I., Mernild, S.H., 2018. Global sea-level contribution from Arctic land ice: 1971-2017. *Environ. Res. Lett.* 13, 125012.
- Hanna, E., Cropper, T.R., Hall, R.J., Cappelen, J., 2016. Greenland blocking Index 18512015: a regional climate change signal. *Int. J. Climatol.* 36, 4847-4861.
- Hanna, E., Fettweis, X., Hall, R.J., 2018. Brief communication: recent changes in summer Greenland blocking captured by none of the CMIP5 models. *Cryosphere* 12 (10), 3287-3292.

- Hanna, E., Fettweis, X., Mernild, S.H., Cappelen, J., Ribergaard, M.H., Shuman, C.A., Steffen, K., Wood, L., Mote, T.L., 2014. Atmospheric and oceanic climate forcing of the exceptional Greenland ice sheet surface melt in summer 2012. *Int. J. Climatol.* 34, 1022–1037.
- Hanna, E., Navarro, F.J., Pattyn, F., Domingues, C.M., Fettweis, X., Ivins, E.R., Nicholls, R.J., Ritz, C., Smith, B., Tulaczyk, S., Whitehouse, P.L., Zwally, H.J., 2013. Ice-sheet mass balance and climate change. *Nature* 498(7452), 51–59.
- Majumder S. and Bhattacharyya D., 2018. Mitigating wormhole attack in MANET using absolute deviation statistical approach, 2018 IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC), Las Vegas, NV, 317-320
- Majumder S., Bhattacharyya D. 2019. Adopting Machine Learning Technique to Mitigate Various Attacks in MANET- A Survey Report, *International Journal of Scientific Research and Review*, 8(6), 288-295.
- Majumder S., Bhattacharyya D. 2020. Relation Estimation of Packets Dropped by Wormhole Attack to Packets Sent Using Regression Analysis. In: Mandal J., Bhattacharyya D. (eds) *Emerging Technology in Modelling and Graphics. Advances in Intelligent Systems and Computing*, 937. Springer, Singapore
- Rignot, E., Mouginot, J., Morlighem, M., Seroussi, H., Scheuchl, B., 2014. Widespread, rapid grounding line retreat of Pine Island, Thwaites, Smith, and Kohler glaciers, West Antarctica, from 1992 to 2011. *Geophys. Res. Lett.* 41 (10), 3502–3509.
- Rignot, E., Mouginot, J., Scheuchl, B., van den Broeke, M., van Wessem, M.J., Morlighem, M., 2019. Four decades of Antarctic Ice Sheet mass balance from 1979-2017. *PNAS* 116 (4), 1095–1103.