Reliability of high cloud data between GLOBE Observer App and GEO satellite during the 2019 Fall Cloud Challenge

Mia L., Kimberly T., Victoria T. and Trina T.

MacArthur Fundamental Intermediate School

Santa Ana, CA

Abstract

The presence of low-level and mid-level clouds can lead to inaccuracies in high cloud data collection. This can lead to satellite or ground observation errors. Satellite errors can be reduced with adjustments to instrument sensitivity, while observer errors can be improved with training. Data from the Fall Cloud Challenge was collected to observe when satellite data and student data agreed. If there was agreement, low and mid clouds did not interfere with the accuracy of high cloud data collection. The inverse was considered as well. Over 1000 observations were made, only 395 observations were "reliable" data. Of this reliable data, it was concluded that 328 (83%) agreements were found. Upon further analysis, only 82 observations were taken under the condition in which there was the presence of both low/mid clouds as well as high clouds, the inverse was found. Only 14 (17%) observations consisted of an agreement of high cloud data between student and satellite. **This did support our hypothesis that low/mid cloud presence can interfere with high cloud data recording. Further high cloud data collecting in the presence of both high clouds and low/mid clouds needs to be collected to gather greater support for our hypothesis.**

Investigation

Problem: Low and mid clouds interfere with the satellite high clouds data.

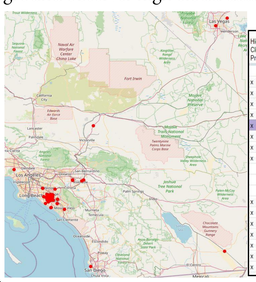
Hypothesis: If the low and mid clouds interfere with the satellite high cloud observations, then the data may not be recorded accurately. This can be determined by the amount of disagreements between satellite data and ground observer data on Globe Observer App.

Key Terms

- **high cloud agreement**: GEO satellite and the students reported similar data.
- **reliable data**: analyzed data found to be dependable in conducting investigation
- **GEO satellite match**: Observer data taken when satellite was overhead

Introduction

From October 15 to November 15 of the year 2019, students at Douglas MacArthur Fundamental Intermediate School participated in the Fall Cloud Challenge. The data was gathered at the convenience of each student and MacArthur was recognized as a Top Observer for the GLOBE North American Region. Over 1000 observations were made within that month, not all of them were used in this study. In analyzing this data, many challenges arose that lead to a need to separate data into two categories: reliable and unreliable. The purpose was to ensure the data used in further analysis was of high quality. The reliable data was further analyzed to determine whether the presence of low-level and mid-level clouds caused inaccuracies that led to disagreement between satellite high cloud data and ground observer high cloud data as stated in our hypothesis.

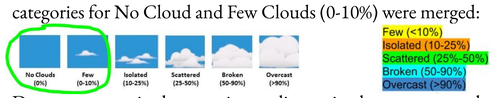


High Clouds Present	High Cloud Data Agrees	Student High Cloud Cover	GEO High Cloud Cover	Low or Mid Clouds Present	Mid Cloud Cover	GEO Mid Cloud Cover	Low Cloud Cover	GEO Low Cloud Cover	GEO/Terra/Aqua Satellite Match Table	Student Total Cloud Cover %	GEO Total Cloud Cover
	x	isolated	4.69	x	isolated	0	-99	15.31	https://scool.l	20-25%	19.39
x		isolated	5.21	x	few	9.38	isolated	11.46	https://scool.l	20-25%	26.05
x		isolated	8.6	x	-99	5.38	-99	23.66	https://scool.l	25-50%	37.64
x		isolated	21.43	x	-99	4.08	isolated	6.12	https://scool.l	20-25%	31.63
x		isolated	21.43	x	-99	4.08	-99	6.12	https://scool.l	25-50%	31.63
x	x	isolated	22.55	x	-99	1.56	-99	31.37	https://scool.l	20-25%	55.88
x		isolated	50.0	x	scattered	21.09	broken	2.41	https://scool.l	50-90%	74.7
x		isolated	90.72	x	-99	9.78	-99	0	https://scool.l	20-25%	100
x		overcast	0	x	overcast	0	overcast	16.33	https://scool.l	90-100%	16.33
x		overcast	0	x	overcast	11.34	overcast	42.27	https://scool.l	90-100%	53.61
x		overcast	0	x	-99	22.11	-99	43.16	https://scool.l	90-100%	65.27
x		overcast	0	x	overcast	42.86	overcast	54.29	https://scool.l	90-100%	96.81
x		overcast	0	x	overcast	0	overcast	5.18	https://scool.l	90-100%	5.18
x		overcast	0	x	overcast	0	overcast	22.92	https://scool.l	90-100%	22.92
x		overcast	0	x	overcast	0	overcast	79.59	https://scool.l	90-100%	79.59
x		overcast	4.17	x	-99	1.04	-99	62.5	https://scool.l	90-100%	67.71
x		overcast	5.43	x	overcast	13.64	overcast	11.96	https://scool.l	90-100%	38.43
x		overcast	9.28	x	overcast	84.54	overcast	6.19	https://scool.l	90-100%	100.01
x	x	overcast	73.91	x	overcast	7.61	overcast	1.09	https://scool.l	90-100%	82.61

Material and Methods

- **NASA Globe Observer App**
- **phone or digital device with a camera**
- **Cloud Chart**

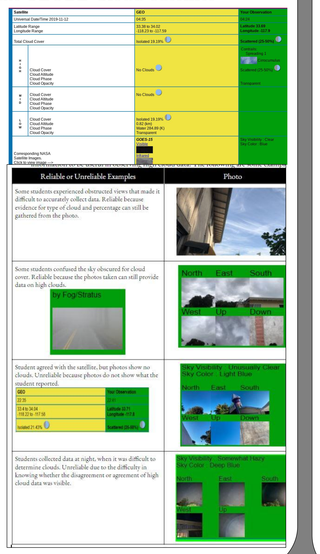
1. Cloud data could be recorded on paper or directly inputted into the NASA Globe Observer App on a phone or other digital device with a camera.
2. All Globe Observer App data was compiled in a Google Sheet for analysis.
3. Data was categorized as reliable and unreliable comparison reports.
4. Percentages for high cloud cover were taken from NASA Globe website (The categories for No Cloud and Few Clouds (0-10%) were merged:



5. Data was categorized as agreeing or disagreeing between ground observer high cloud data and GEO satellite high cloud data to investigate hypotheses.

Reliable and unreliable data

1. Initially, over 1000 ground observations were recorded on either paper or through the Globe Observer App.
2. To increase data quality, only data taken with Globe Observer App was further analyzed since the reliability of this data increases with the App guiding cloud accuracy. The Globe Observer App totaled 581 observations.
3. Of the 581, only 511 were GEO satellite matches that provided a **GLOBE Cloud Satellite comparison**; the rest were considered unreliable.
4. Of the 511, it was determined that a disagreement in total cloud cover between the GEO satellite and ground observation would represent too great a difference to consider reliable. Reliable data was determined to be that which had no greater difference in total cloud cover than 20% between satellite and observer. That left 395 observations.
5. Questionable data was individually analyzed and determined to be reliable or unreliable. The following table shows examples. In order to sort the reliable and unreliable data, we had to analyze the **GLOBE Cloud Satellite comparisons**.
6. Data that might be unreliable was analyzed individually to see if it provided enough information to be useful in observing high cloud data. The following are some examples:
7. Of the 395 reliable observations, data was further separated into the following categories:
   a. Total high cloud agreement between satellite and observer
   b. Total high cloud agreement in the presence of high clouds
   c. Total high cloud agreement in the absence of high clouds
   d. Total high cloud agreement in the presence of low/mid clouds
   e. Total high cloud agreement in the presence of high clouds and low/mid clouds



Results and discussions

Of the over 1000 observations, 581 ground observations were recorded using Globe Observer App, which was considered to control for higher quality data collecting and was considered more reliable. Of the 581 ground observations, 511 were GEO satellite matches. This meant that the GEO satellite and ground observation was recorded within 15 minutes of each other and therefore could reveal agreement and disagreement in high cloud data. Of the 511 GEO satellite matches, only 395 were considered reliable enough to use for further analysis of the agreement and disagreement of high cloud data (**Table 1**).

Table 1: Chart of reliable data

Total Globe Observer App Data	Total Geo Satellite Matches	Total Reliable Data
581	511	395

Table 2: Reliable data categorises

Analysis of Reliable Data = within 395 Observations			
High Cloud Data Agreement between satellite and student	**High Cloud Agreement in High Cloud presence**	**High Cloud Agreement collected in High Cloud absence**	**Data collected in Presence of low/mid clouds**
328	82	246	178
83%	20.8%	62.3%	45.1%

To determine ground observations were reliable, the total high cloud cover needed to be within 20% variation of the satellite data. High cloud cover that was greater than 20% different, was considered unreliable. In addition, reliability was determined by analyzing photos and **GLOBE Cloud Satellite comparison** charts.

However, it was also noted that some of the above categories show a lack of essential components that were expected to draw conclusions in our hypothesis. For example, some of the observations were recorded without the presence of high clouds. Other data was collected without the presence of low/mid clouds. Having both high clouds and low/mid clouds present would lead to an ideal condition to investigate the interference of low/mid clouds on data collection of high clouds.

The 395 observations that were considered reliable, were further analyzed and categorized (**Table 2**). Of the 328 there were agreement in 83% of the reliable data. This appeared to not support our hypothesis.

Further analysis was needed in order to determine the results in the ideal condition. By ideal condition, it was data collected in the presence of both high clouds and low/mid clouds. This condition would best show the accuracy of agreement between the satellite and ground observation on high cloud data.

Table 3: Reliable data in ideal condition

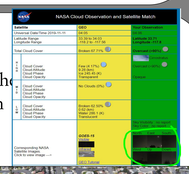
Analysis of data in presence of high and low/mid clouds = 82		
All data collected in ideal condition	**Total High Cloud Agreement Presence of low/mid clouds**	**Total High Cloud Disagreement in presence of low/mid clouds**
82	14	68
100%	17%	83%

Under these conditions, only 82 observations were left (**Table 3**). Within these 82, only 14 or 17% if the total observations contained agreeing high cloud data.

While this supports our initial hypothesis that low/mid clouds interfere, and therefore cause more disagreement, between ground observations and satellites, the available data was determined not to be enough to fully support our hypothesis. More data, with ideal conditions, would need to be collected.

Difficulties

Looking at the data there were some inconsistencies with the what the student recorded and the pictures that were taken by that student. Reliability was determined by referencing the **GLOBE Cloud Satellite comparison** charts.



Much of our data was recorded around our school. Obstacted made it difficult to record accurate cloud data most of the time,

Suggestions and recommendations

1. More data collections can fulfill the ideal conditions of high cloud data when there are low and mid level clouds present.
2. Further analysis is needed to interpret the 2 distinct results from our high cloud observations. We found an 83% agreement between ground observers when observing all high cloud data agreement. However, when both high and low/mid clouds were present, 14% agreement was found.
3. Continue studies on categorizing data reliability and creating standards for reliability could help guide the research.
4. Continue studdies can determine why the disagreement occured (sensitivity of satellite or ground observers)
5. Further training for ground observers could improve data reliability. These strategies could be tested to find there impact of gathering reliable data.

Acknowledgments

2019 Fall Challenge
GLOBE North American Region
TOP OBSERVER
MacArthur Fundamental School

Our team would like to thank Mr. Peter Falcon for providing essential guidance that helped initiate our data collecting and investigation. Also, for providing us with resources and information that supported us throughout our cloud data research, including training our teacher, Mrs. Flores, so that she could help us. We would also like to thank the many individuals that make Globe Student Research Symposium possible, since life-long skills and experiences are acquired through this opportunity. Special thanks to Jennifer Bourgeault for informing and updating our team through the stay-at-home orders. We would also like to thank Mr. John Power, Ms. Walter, and Ms. Patter from UCI for asking us important questions about our data, and helping us look at our data through different perspectives. Our team would also like to acknowledge Ms. Valerie Armstrong, MacArthur's 6th grade science teacher, for her input on our data analysis, and for inviting UCI graduate students to discuss our data. Finally, we cannot imagine the completion of this research without the time and guidance of Mrs. Marilé Colón Robles. We are grateful for her support and encouragement throughout this entire time, without which we may have given up or may have become lost in cloud data.

References

- Atkinson, Joe. "GLOBE Fall Cloud Challenge Rakes in the Observations." *NASA*, NASA, 22 Nov. 2019, www.nasa.gov/feature/langley/globe-fall-cloud-challenge-rakes-in-the-observations
- Dodson, J. Brant, et al. "Eclipse across America: Citizen Science Observations of the 21 August 2017 Total Solar Eclipse." *Journal of Applied Meteorology and Climatology*, vol. 58, no. 11, 2019, pp. 2363–2385., doi:10.1175/jamc-d-18-0297.1.
- Dunbar, Brian. "Why Does NASA Have Its Head in the Clouds?" *NASA*, NASA, www.nasa.gov/audience/foreducators/5-8/features/F_Head_in_the_Clouds.html.
- "GLOBE Observer." *GLOBE.gov*, observer.globe.gov/do-globe-observer/clouds.
- "NASA GLOBE Cloud Data." *NASA*, NASA, scool.larc.nasa.gov/cgi-bin/NASA-GLOBESatMatch.cgi?observation_id=116-173847-32263218-20191111040 6.
- Robles, Marile Colon, et al. "Integrating Tech: Making Science Come Alive With Clouds." *Science Scope*, vol. 043, no. 04, 2019, doi:10.2505/4/ss19_043_04_8.
- Robles, Marilé Colón, et al. "Clouds Around the World: How a Simple Citizen Science Data Challenge Became a Worldwide Success." *Bulletin of the American Meteorological Society*, 2020, doi:10.1175/bams-d-19-0295.1.