

Using Tweets For Rainfall Monitoring

Luiz Eduardo Guarino de Vasconcelos

Eder Christian Moreira dos Santos

Mário Lemes de Figueiredo Neto

Nelson Jesus Ferreira

Leandro Guarino de Vasconcelos



Introduction

- The number of people connected to the Internet has increased.
- In Brazil, the population is approximately 204 million people and 54% of the Brazilian population access the Internet.
- There are 96 million active social networking accounts in Brazil.
- Content and data posted on social media, it is possible to collect, analyze and extract information and knowledge



Introduction

- ▶ Due to the success of social media in Brazil, especially Twitter, there was the hypothesis to monitor rainfall through tweets.
- ▶ The hypothesis arose from the need to improve rainfall estimation accuracy in the Brazilian current scenario.
- ▶ Measures with rain gauges can be influenced by exposure to wind, which can cause up to 20% underestimation of the measurements



Introduction

- Remote sensing is an alternative. Weather radars are able to represent the spatial structure of rainfall systems, but they have several sources of error inherent in the rainfall intensity
- This work aims to analyze the contents of the Twitter social media, especially the tweets related to rainfall events in order to determine whether this information may contribute to the rainfall monitoring in Brazil



Social Networks

- ▶ A social network is composed by people or organizations connected to a computer network
- ▶ Facebook, Twitter, Youtube
- ▶ Social ties can be exemplified when there is interaction between two most users of any social media; when all the communication generated by these users becomes frequent and generates a connection between them.



Twitter

- ▶ Every social media has a purpose.
- ▶ Twitter is intended to convey information quickly, since each tweet is limited to 140 characters.
- ▶ Twitter allows ease of use of the information through its API.
- ▶ There are more than 316 million monthly active users and over 500 million tweets are sent per day.



Social Network Analysis

- ▶ With the large volume of data and information provided in social media, it is possible to study patterns of interactions between the actors of social networks.
- ▶ The term "Social Network Analysis" (SNA) has existed for a long time and their use has grown in recent years due to the widespread use of social media.
- ▶ A large part of social media has powerful tools that analyze user profiles.



Weather Information

- For the acquisition of meteorological data in Brazil, automatic weather stations (AWS or EMA in portuguese) and meteorological radars are used.
- An automatic weather station collects, minute by minute, weather information representing the area where it is located.
- The data is freely available in real time via internet for the development of weather forecast and the different meteorological products of interest to industry users and the public in general, and for a wide range of applications in research meteorology, hydrology and oceanography.



Weather Information

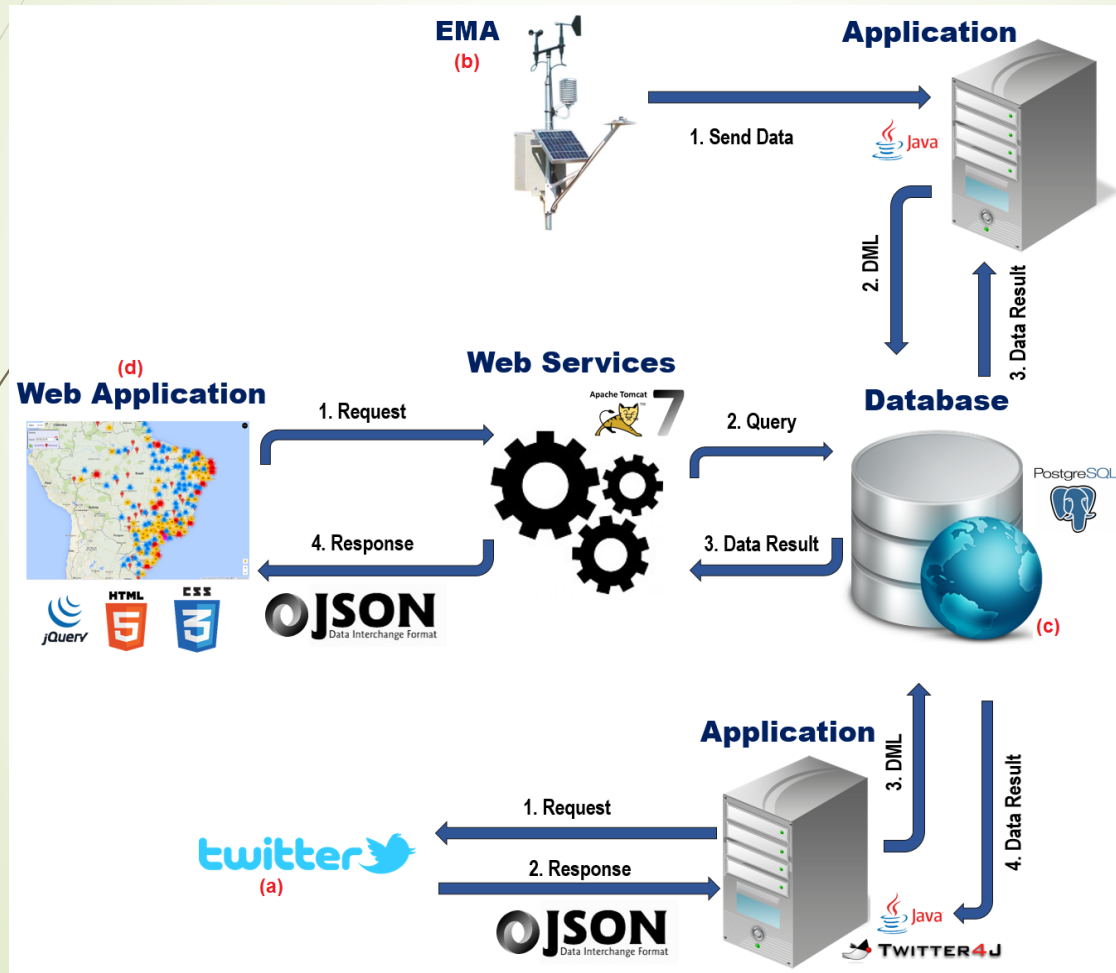
- The radars are used to perform the meteorological monitoring on vulnerable cities the occurrence of floods, flash floods and geological events such as landslides.
- In Brazil there is the monitoring of rainfall using satellite images. The satellite-based measurements establish relationship with meteorological variables estimated by existing devices on satellites.



Case Study

- To collect the tweets is necessary to consume the data from the Twitter API
- The Twitter API allows extract various information of the social network

Case Study – Application Architecture





Case Study

- The Java programming language, along with the Twitter4J, which is an abstraction of the Twitter API.
- Database PostgreSQL to the storage of all data collected by the application;
- Web service in Java, hosted on Apache Tomcat 7 Server;
- Web application that consumes and uses the web service to view the data on a map.

Experimental Results

Word (in Portuguese)	Word (in English)	Amount of tweets	%
chuva	rain	793,132	71,15
chuveisco	drizzle	786	0,07
trovoada	thunderstorm	10,544	0,95
nevoeiro	fog	5,310	0,05
tempestade	storm	76,231	6,84
raio	lightning	60,087	5,39
garoa	drizzle	4,934	0,44
trovão	Thunder	19,012	1,71
temporal	temporal	112,505	10,09
relâmpago	lightning	26,158	2,35
dilúvio	flood	23,790	0,21
chuarada	rain	4,219	0,38
precipitação	<u>rainfall</u>	1,456	0,38

Words used to rainfall Monitoring



Experimental Results

- #chuva (rain) hashtag was defined
- We collected 1,328,221 tweets (posted by anybody) between days 08 September 2015-08 October 2015.
- The collection was done through the application.
- From the collected data, it was found that only 0.1% of tweets had the latitude and longitude in the tweets.
- To increase the amount of tweets with geolocation, the city defined in the user profile was used which is available in each tweet.
- This allowed 23.78% of tweets were geolocation information.

Experimental Results

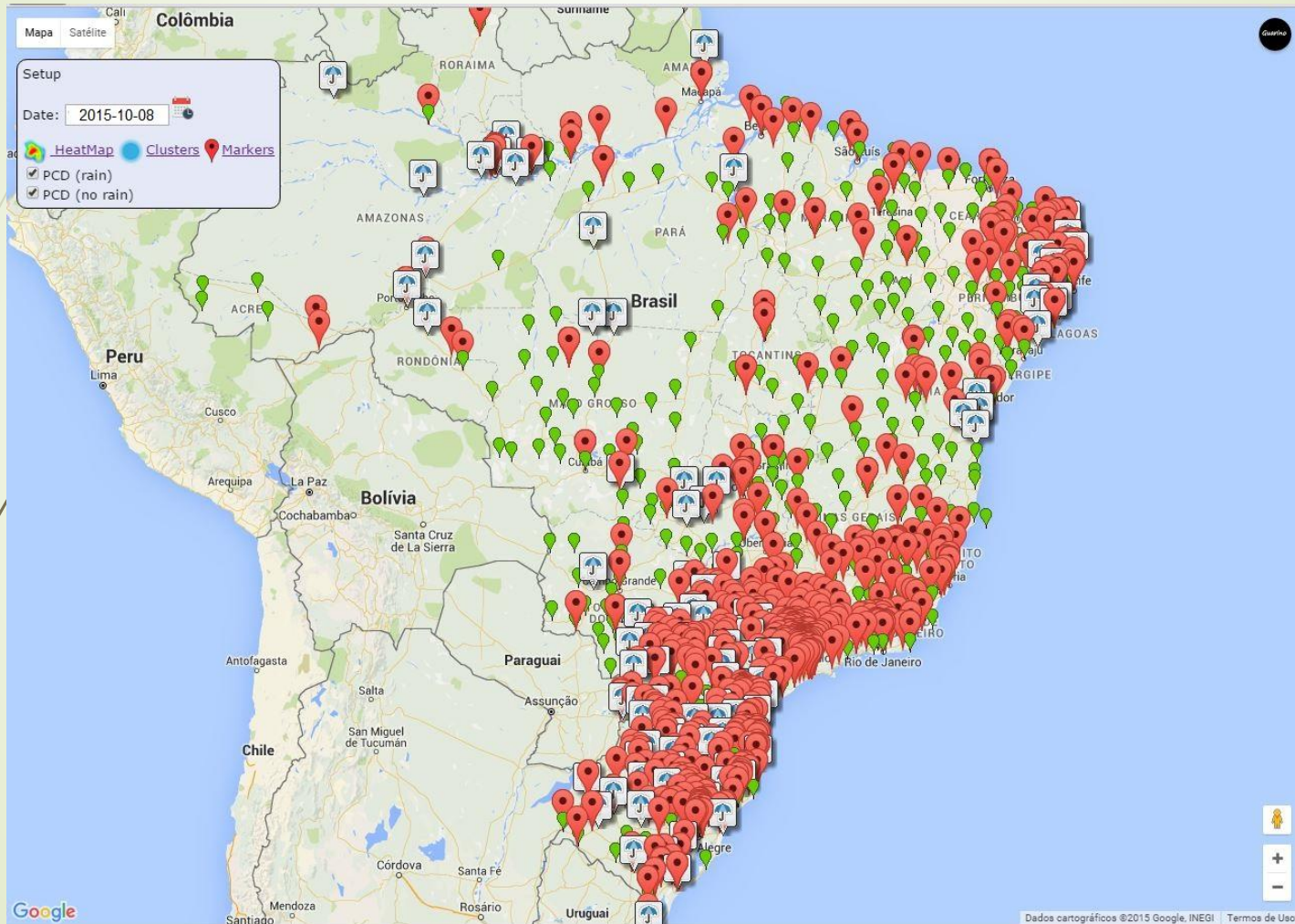
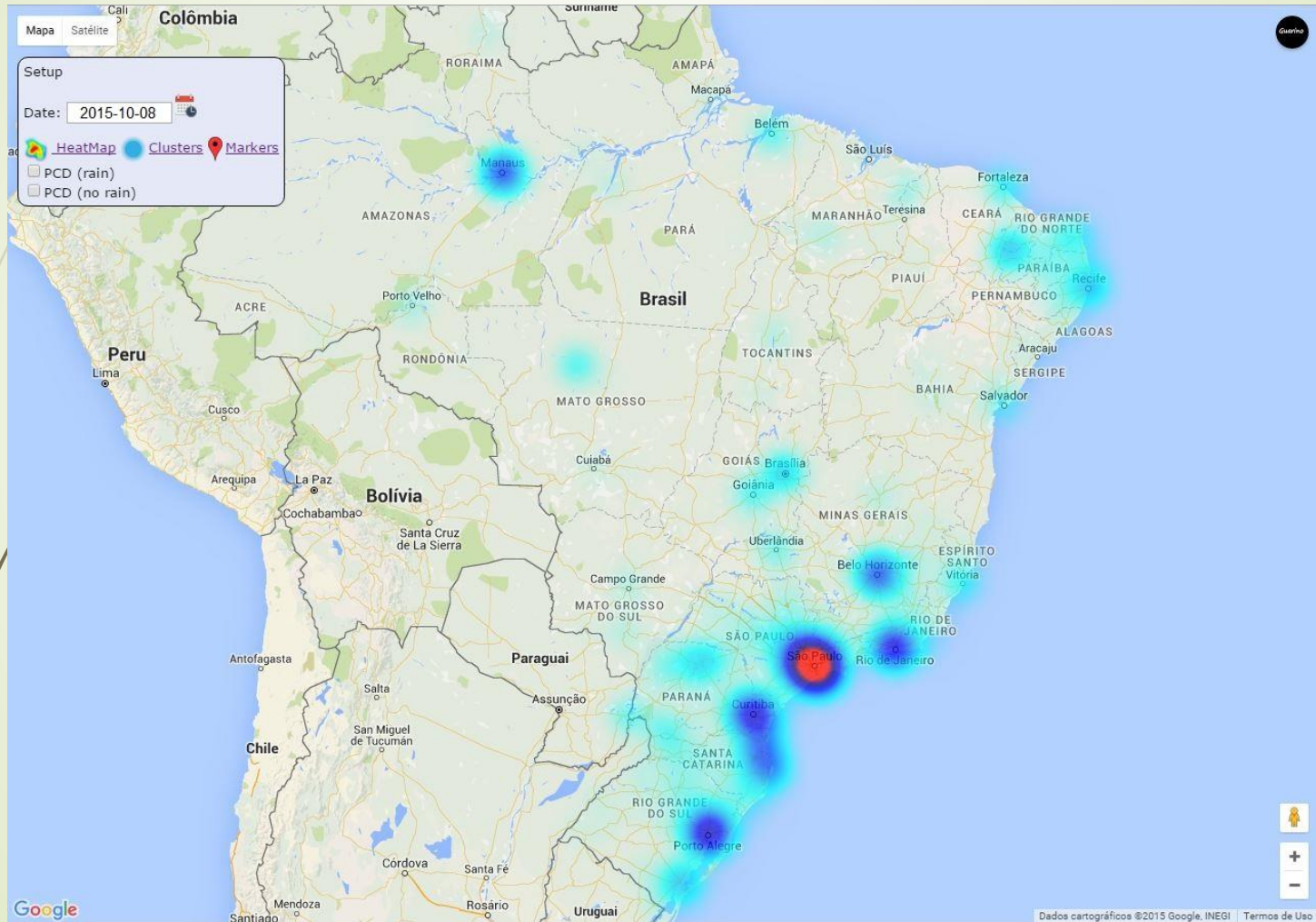


Chart with markers - October 8, 2015 –
48,500 tweets about rainfall. 102 AWS received mm of rainfall and 439 AWS had no rainfall that day

Experimental Results



Heat Map chart - October 8, 2015

Experimental Results

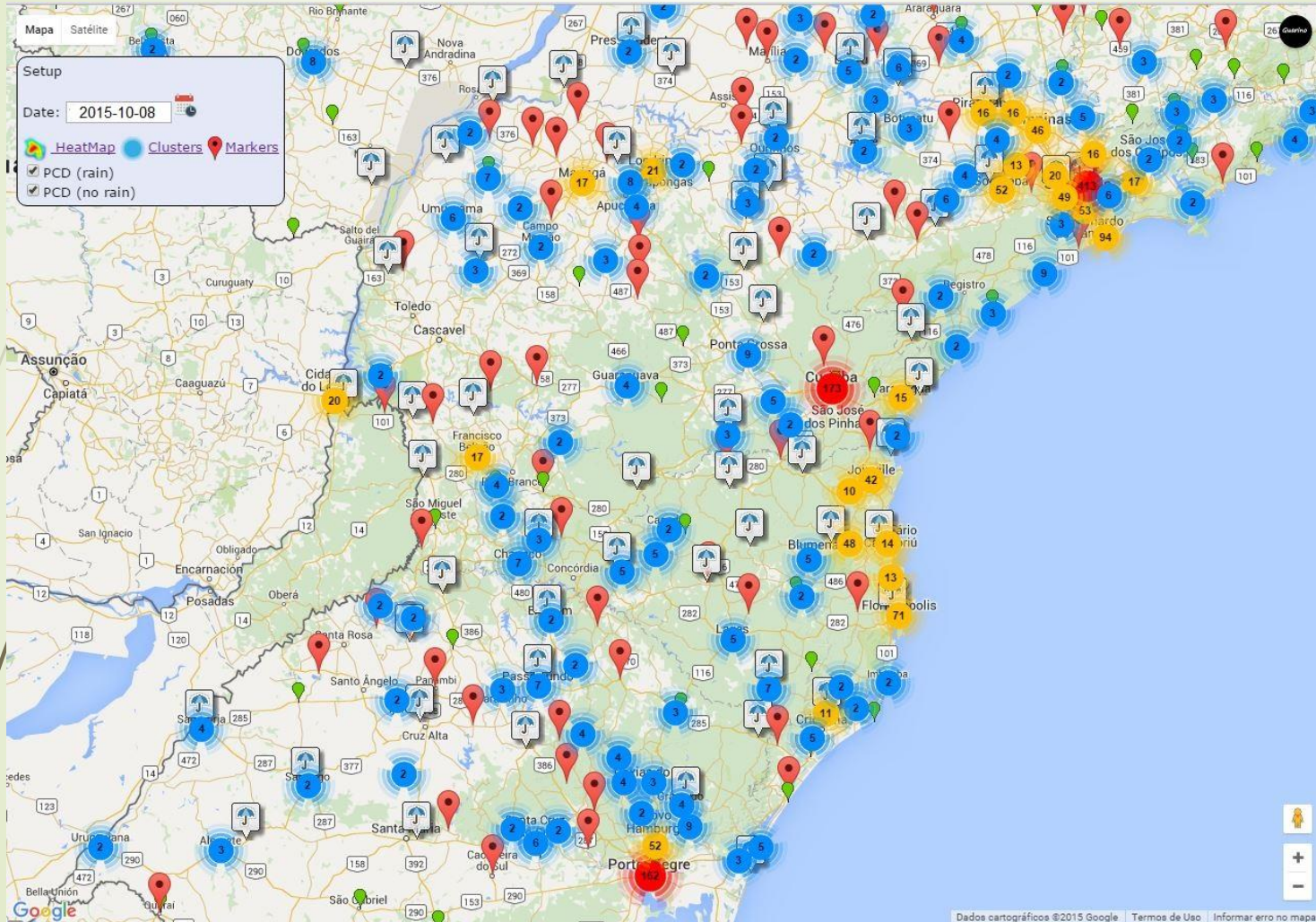
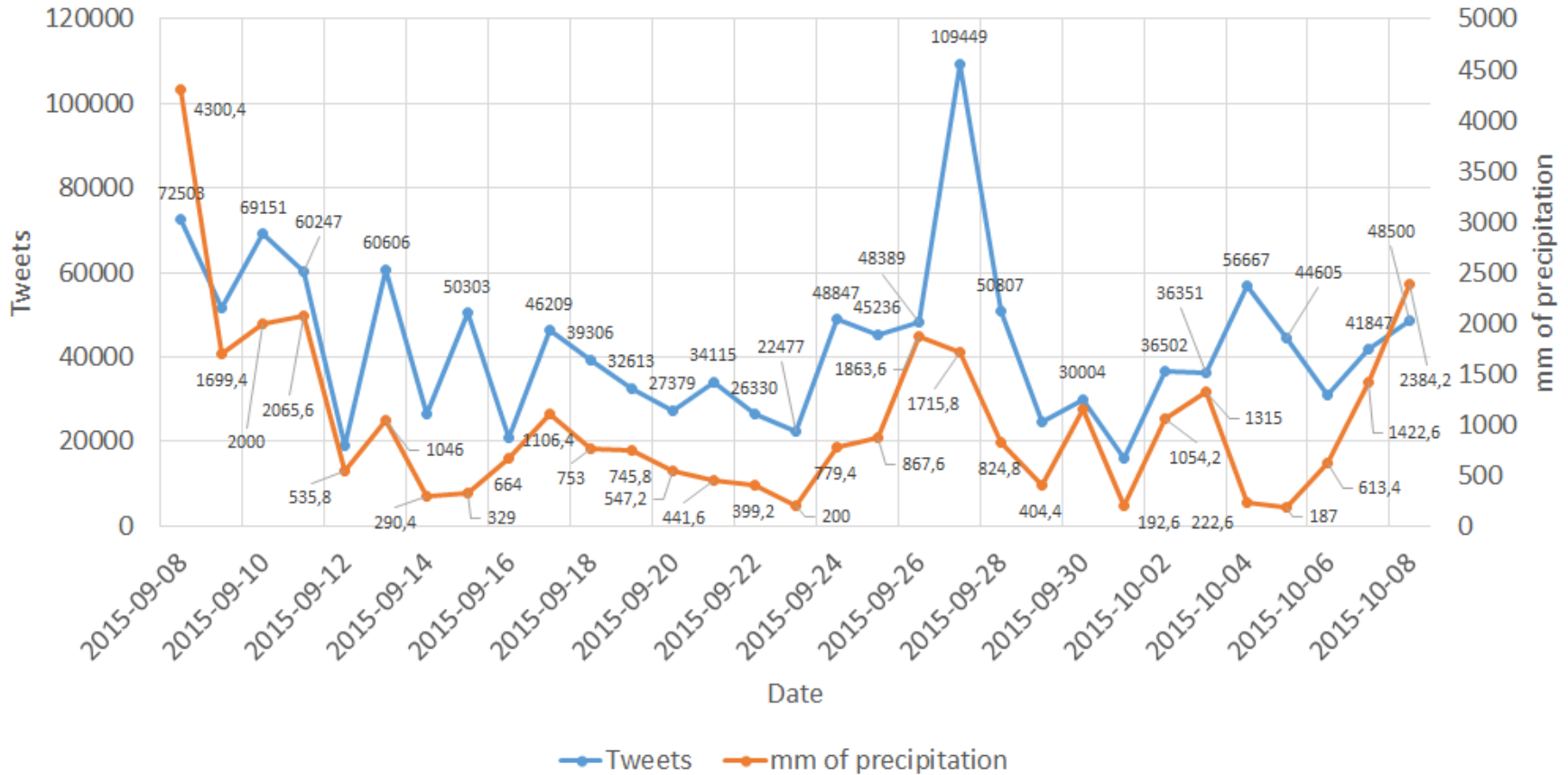


Chart with tweets (clusters) and AWS's of the Southern Brazil region – October 8, 2015

Experimental Results

Tweets and mm of precipitation per day



Amount of tweets and mm of the rainfall per day



Conclusion

- We performed social network analysis of Twitter, in particular, the tweets related to rainfall events in order to determine whether this information could contribute to the monitoring of precipitation events in Brazil.
- There is a close relationship exists between the amount of tweets and the amount of mm precipitation detected by AWS.
- The solution used in this study is a tool that can assist in monitoring rainfall events. There is the possibility, for example, to provide warnings about certain weather conditions and especially supplement the information provided by other monitoring means



Conclusion

- Future Works
- Examine minimum and maximum distance of tweets regarding the AWS with and without precipitation statement;
- Monitor hourly tweets relating them to rain gauges and radars;
- Visually compare the tweets map with satellite images processed by DSA/CPTEC/INPE.