Project Draft

Title: Assessment of the impacts of the BP oil spill in the Gulf of Mexico on water quality
Notice: Dr. Bryan Runck
Author: Diego Osorio
Date: 11/8/2022

Project Repository: https://github.com/osori050/GIS5571_project
Google Drive Link:
Time Spent: 12 h (including code development)

Abstract

In 2010, an explosion occurred in the Deepwater Horizon semi-submersible mobile offshore drilling unit leading to an oil spill of roughly 210 million gals into the Gulf of Mexico. NOAA measured seawater physicochemical variables then (such as oxygen, salinity, conductivity, SPM, and pH) and projected the daily oil plume's trajectory. That data will be analyzed together to assess the impact of the said catastrophe on water quality over the Louisiana continental shelf. The water variables are expected to have abrupt changes due to the pollutant input.

Problem Statement

In 2010, the Deepwater Horizon semi-submersible mobile offshore drilling unit suffered an explosion that killed 11 workers, injured 17 others, and spilled around 210 million US gal (780000 m^3) of oil into the ocean over 87 days. This catastrophe caused the biggest oil leak in history and severely impacted the quality of water in the Gulf of Mexico (Deepwater Horizon oil spill, 2022; Deepwater Horizon explosion, 2022). Thus, this project aims to assess the impact of the BP oil spill in the Gulf of Mexico on the water by analyzing physicochemical variables such as oxygen, conductivity, and SPM (Suspended particulate matter), among others measurements at that time. Figure 1 shows roughly the area affected by the spill, and Figure 2 shows the locations where the physicochemical variables were measured.

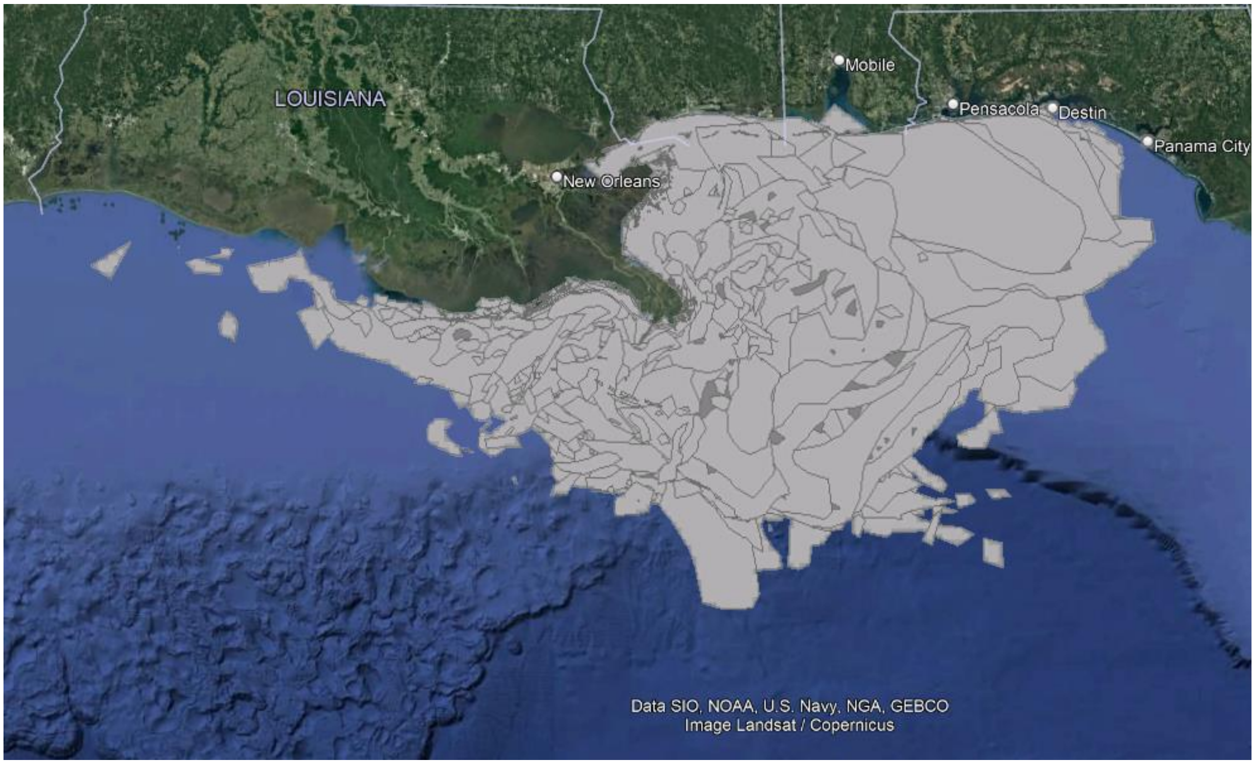

Figure 1. Oil spill plume (ESRI, n.d.)

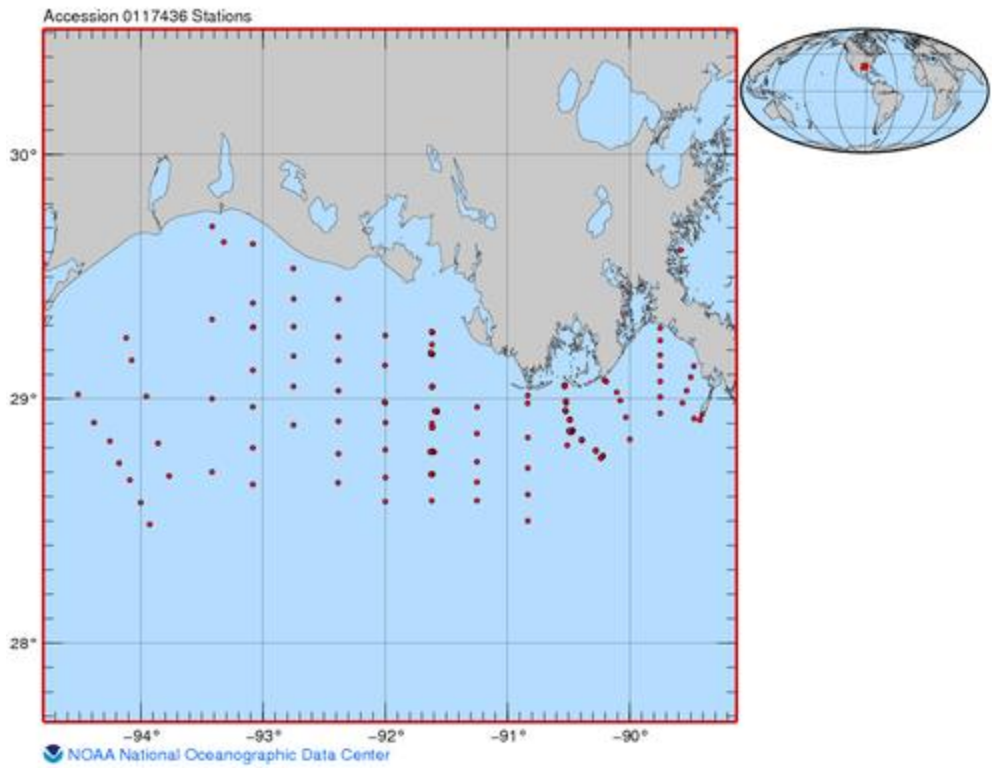

Figure 2. Water variables monitoring station locations (NOAA, 2021)

Table 1. Requirements for the oil impact on water quality

#	Requirement	Defined As	(Spatial) Data	Attribute Data	Dataset	Preparation
1	Area affected by the oil spill	Projection of the trajectory of the spill plume	Polygon geometry	Dates and area	NOAA Data	
2	Water physicochemical variables	Raw input dataset	Tables	Physicochemical variable measurements and dates	NOAA Data	Transformation into GIS layers

Input Data

First, ESRI has created a layer on ArcGIS Online that shows the daily projection of the oil spill plume in the Gulf of Mexico from May 2 to August 5 of 2010 based on NOAA data acquisition. The full extent is from -85.7202 to -93.4310 latitude, and from 27.2551 to 30.4219 longitude (datum WGS 1984). Additionally, there is a collection of physical, chemical, and biological data in the Gulf of Mexico from February 2 to October 28 of 2010 by NOAA recorded in a Microsoft Access database and text files.

Table 2. Input data for oil spill analysis

#	Title	Purpose in Analysis	Link to Source
1	NOAA_Gulf_Oil_Spill	Analysis of the extension of the oil spill plum in the Gulf of Mexico over time	http://maps1.arcgisonline.com/arcgis/rest/services/NOAA_Gulf_Oil_Spill/MapServer
2	Physical, chemical, and biological data collected in the Gulf of Mexico from 02 Feb 2010 to 28 Oct 2010 (NCEI Accession 0117436)	Assessment of physicochemical parameters	https://www.ncei.noaa.gov/archive/archive-management-system/OAS/bin/prd/jquery/accession/download/117436

Methods

As the NOAA Gulf Oil Spill layer lies on ArcGIS Rest API, a request is needed by querying the data. The following link shows in i) orange the main URL of ArcGIS Rest API (resource URL), ii) red the specific service and function (operation), iii) blue the function inputs for search (parameters), and iv) green the response type:

http://maps1.arcgisonline.com/arcgis/rest/services/NOAA_Gulf_Oil_Spill/MapServer/0/query?where=SHAPE_Area%3E0&text=&objectIds=&time=&geometry=&geometryType=esriGeometryEnvelope&inSR=&spatialRel=esriSpatialRelIntersects&relationParam=&outFields=*&returnGeometry=true&returnTrueCurves=false&maxAllowableOffset=&geometryPrecision=&outSR=&returnIdsOnly=false&returnCountOnly=false&orderByFields=&groupByFieldsForStatistics=&outStatistics=&returnZ=false&returnM=false&gdbVersion=&returnDistinctValues=false&resultOffset=&resultRecordCount=&queryByDistance=&returnExtentsOnly=false&datumTransformation=&parameterValues=&rangeValues=&f=pjson

The data is retrieved with a GET request and the content is written in a JSON file which is then converted to a shapefile by using the tool JSON To Features.

Similarly, the water variables are downloaded through a GET request as a tar file which is later decompressed. Since the data is stored in a Microsoft Access database, the pyodbc package in Python can be used to access the data. Nonetheless, as shown in Figure 3, the code generates an error message as the Microsoft Access Driver is not present.

```
try:
    con_string = r'Driver={Microsoft Access Driver (*.mdb, *.accdb)};DBQ=E:\ArcGIS_1\Project\BP\0117436\2.2\data\0-data\NGOMEX_2010_Data_NODC_v2.accdb;'
    conn = pyodbc.connect(con_string)
except pyodbc.Error as e:
    print('Error in Connection', e)

Error in Connection ('IM002', '[IM002] [Microsoft][ODBC Driver Manager] Data source name not found and no default driver specified (0) (SQLDriverConnect)')
```

Figure 3. Block of code to access MS Access databases

Hence, the data needs to be pulled from the text files. Since the data in the files are separated by strings (- and |) which altogether recreate cells as shown in Figure 4, a cleaning process is performed by opening the files and writing to a new file the lines without the '-' strings. Later, the latter file is opened as a pandas DataFrame where the NaN and whitespace-populated columns are dropped.

ID	Station ID	Station	Date	depth	Inorganic SPM	Organic SPM	Total SPM
1	42	C5	05/17/10	0.00	2.20	4.40	6.60
2	40	C6C	05/17/10	0.00	5.00	3.80	8.80
3	40	C6C	05/17/10	3.20	4.00	2.00	6.00

Figure 4. Raw data in text files

In future steps, a point shapefile will be created with the coordinates of the stations in the 2010_StationList file. Then, through joins, data about salinity (PSU), conductivity (S/m), oxygen (mg/L and %), SPM (inorganic, organic, and total in mg/L), and pH will be integrated into this shapefile with the dates of the measurements. As the crude oil has a density lower than that of seawater, the oil spill plume is expected to have floated in the ocean; thus, only measurements taken at 0 m depth will be considered. The area of interest (AOI) will be defined as a rectangle including the extent of the bounding box of the stations plus a buffer of around 11 km (6.9 miles).

Due to the large volume of data, one day per week or every other week will be chosen to carry out the spatial interpolation to see the temporal trend of the impacts of the crude oil on Louisiana shelf seawater. For this, 4 interpolation methods will be used: global and local polynomial interpolations, which are inexact interpolators, and inverse distance weighted and radial basis functions, which are exact interpolators. These 4 models will be compared head to head to see which fits better.

Regarding the oil spill shapefile, it will be clipped to the AOI and only 1 plume trajectory projection per week or every other week will be considered (same day of the week as in the water variables dataset). Finally, animations will be generated for each variable and interpolation method. That is, it is expected to have around 20 time-series animations (5 variables by 4 interpolators). I would like to have for each animation the oil spill trajectory time series overlaying with a degree the transparency on top to see if the changes in the water variables are related to the oil spread. Figure 5 and Figure 6 illustrate the workflow to follow.

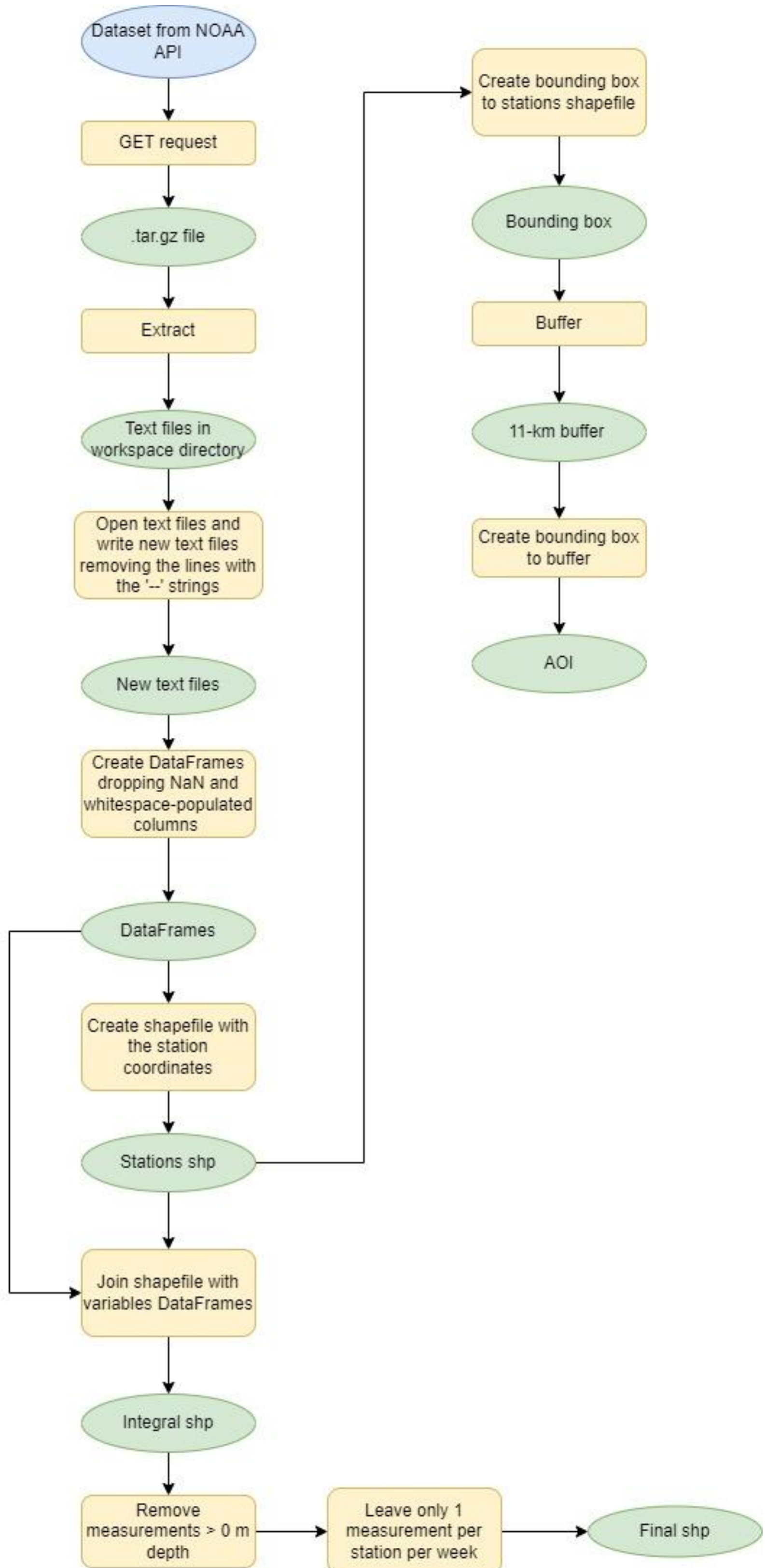

Figure 5. Flow diagram of the transformation of NOAA's data to create the AOI and the final shapefile with the water variable values

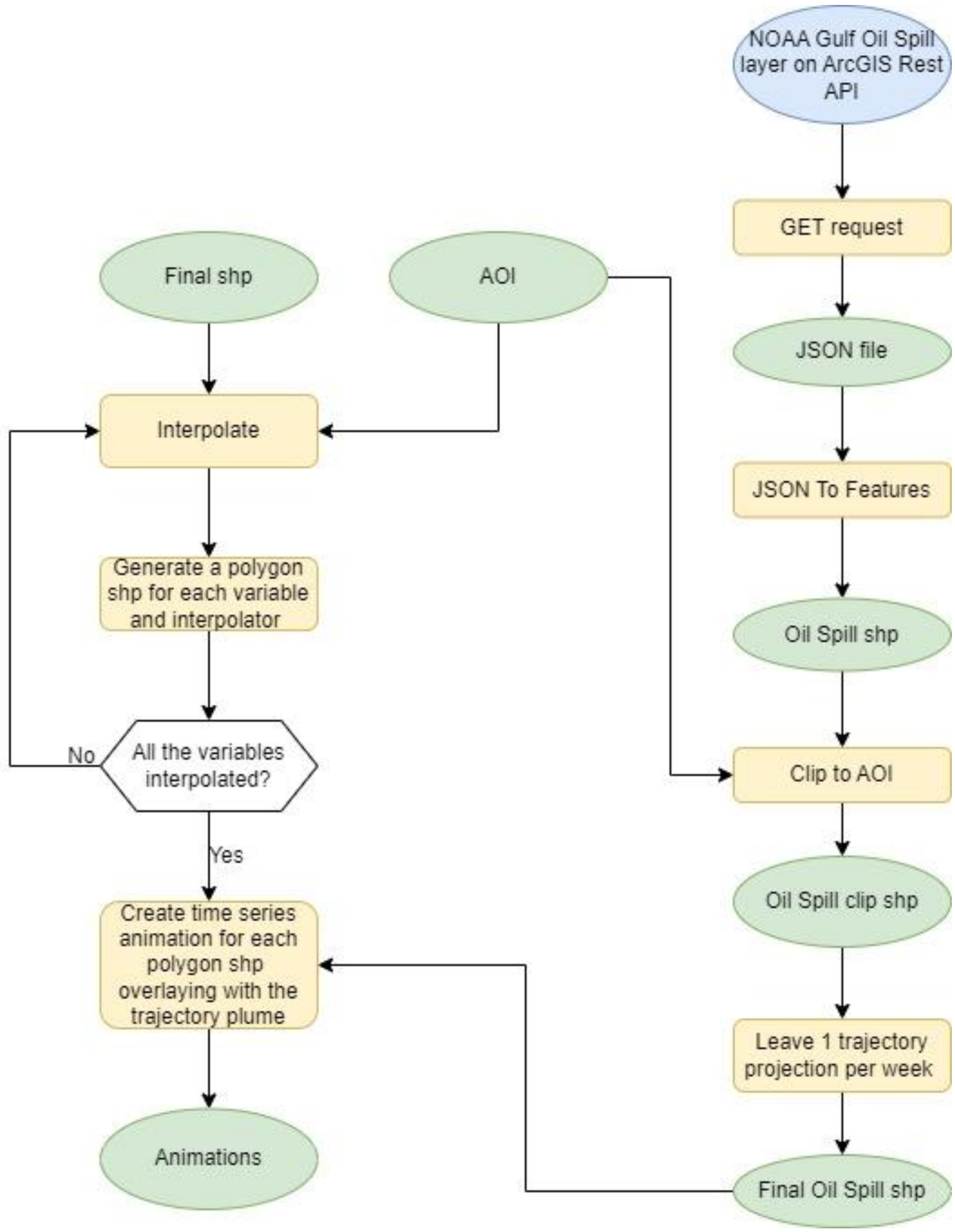

Figure 6. Diagram of the time series animation workflow

Results

It is expected to see some variables changing abruptly over time because of the oil spill plume expansion. Likewise, the furthest stations from the spill are expected to have recorded the spill impacts days later than the closest ones. In this section, there will be some screenshots of the animations to illustrate the most significant impacts on the physicochemical variables in the seawater of the Louisiana continental shelf according to each interpolator.

Results Verification

There will be 3 ways to verify the results. First, some water variables started being measured and recorded from February 2, 2010, and the oil spill occurred on April 20, 2010, which allows comparing water quality before and after the accident. Second, the plume trajectory was projected up to August 5, 2010, and as shown in Figure 1 and Figure 2, the plume does not overlap 100% with the stations (AOI), so the stations further to the west will still be a reference of the water quality status. Third, since two exact interpolators will be used, the results from these two will be compared to the data values at the stations, which should be the same.

Discussion and Conclusion

As the physicochemical variables are expected to change considerably, the magnitude of the change will be compared to reference information on the seawater status prior to the accident to draw conclusions about the impact of the oil spill. The interpolators will also be compared to decide which models are better for these kinds of problems. Moreover, in this section, I will talk about the methods (such as packages, functions, etc.) that helped me get the results as well as those that did not and why. I will explain the new tools and knowledge acquired and how that can be applied to other problems.

References

Deepwater Horizon explosion. (2022, August 26). Retrieved from Wikipedia: https://en.wikipedia.org/wiki/Deepwater_Horizon_explosion

Deepwater Horizon oil spill. (2022, September 20). Retrieved from Wikipedia: https://en.wikipedia.org/wiki/Deepwater_Horizon_oil_spill

ESRI. (n.d.). *NOAA_Gulf_Oil_Spill (MapServer)*. Retrieved from ArcGIS REST Services Directory: http://maps1.arcgisonline.com/arcgis/rest/services/NOAA_Gulf_Oil_Spill/MapServer

NOAA. (2021). *Physical, chemical, and biological data collected in the Gulf of Mexico from 02 Feb 2010 to 28 Oct 2010 (NCEI Accession 0117436).* Retrieved from DATA.GOV: https://catalog.data.gov/dataset/physical-chemical-and-biological-data-collected-in-the-gulf-of-mexico-from-02-feb-2010-to-28-oc

Self-score

Category	**Description**	**Points Possible**	**Score**
Structural Elements	All elements of a lab report are included **(2 points each)**: Title, Notice: Dr. Bryan Runck, Author, Project Repository, Date, Abstract, Problem Statement, Input Data w/ tables, Methods w/ Data, Flow Diagrams, Results, Results Verification, Discussion and Conclusion, References in common format, Self-score	28	28
Clarity of Content	Each element above is executed at a professional level so that someone can understand the goal, data, methods, results, and their validity and implications in a 5 minute reading at a cursory-level, and	24	24

	in a 30 minute meeting at a deep level **(12 points)**. There is a clear connection from data to results to discussion and conclusion **(12 points)**.		
Reproducibility	Results are completely reproducible by someone with basic GIS training. There is no ambiguity in data flow or rationale for data operations. Every step is documented and justified.	28	28
Verification	Results are correct in that they have been verified in comparison to some standard. The standard is clearly stated **(10 points)**, the method of comparison is clearly stated **(5 points)**, and the result of verification is clearly stated **(5 points)**.	20	20
		100	100