A Slick Model: Modeling the Dispersion of Oil Slicks in the Ocean

Table of Contents

Abstract

Oil spills cause millions of dollars in damages and destroy habitats. These repercussions highlight the need to contain oil spills quickly. Oil dispersion models are a set of procedures that model the dispersion of spills. Though oil dispersion models currently exist, these models require a lot of computational power. This project proposed a simple mathematical model to represent the dispersion of oil slicks.

To reduce computation, two techniques were used: the model tracked only the eight cardinal points, and only one set of ocean currents was applied to those points. The initial shape of the spill was approximated to be a circle of a five meter radius. The set of vectors used for the duration of the spill was then estimated. Once the velocity vectors were calculated, they were applied to the cardinal points on the boundary of the initial five meter circle, which extrapolated the points to new locations based on time. Assuming that the eight points were the outline of the spill, a line was fit through the points using three different methods: a piecewise linear fit, a circle fit, and a piecewise cubic fit.

This model simulated three different historical spills and the results were compared to information provided by newspaper data. T-tests showed there was no significant difference between the designed models and the benchmark. Increasing the sample size will help to verify integrity.

A Slick Model: Modeling the Dispersion of Oil Slicks in Oceans

Modeling the dispersion of oil slicks in the ocean is essential as the world extracts greater amounts of oil from the ocean floor each year, and the potential for spills increases with each additional barrel of oil extracted. According to Olascoaga and Haller (2012), the largest accidental oil spill in history was the Deepwater Horizon oil spill on April 20, 2010. This accident resulted in an estimated loss of four million barrels of oil that spilled into the Gulf of Mexico. The spill had several severe financial and environmental implications: a reduction in tourism, the cost of clean-up efforts, and the revenue loss associated with the barrels of oil spilled (Olascoaga & Haller, 2012). In order to concentrate clean-up efforts and efficiently use monetary resources, modeling the dispersion of oil spilled is necessary.

Two main methods to model the dispersion of oil already exist: satellite imaging (Özgökmen et al., 2016), and advanced Lagrangian mechanics (North et al., 2011). Satellite imaging uses neural networks to identify areas that are covered by a thin layer of oil, and oil detection algorithms are used to detect smaller but thicker areas of oil (Özgökmen et al., 2016). Using interpolation, a continuous time series of gridded values can be derived (Özgökmen et al., 2016). The other popular method for modeling oil dispersion is Lagrangian mechanics. Lagrangian mechanics is a field of study similar to Newtonian mechanics, except that Lagrangian mechanics bypasses Cartesian constraints and can be used in any coordinate system. According to North et al. (2011), Lagrangian Coherent Structures (LCSs) have the potential to forecast imminent shape changes in oil slicks. LCSs can also be used to obtain remarkably accurate predictions of instabilities despite approximations (North et al., 2011).

The problem with most current models is the amount of computational power required to model the spill. As North et al. (2011) stated, many thousands of particles are required to compute their predictions. Each particle in North et al.'s model represents an oil particle, and each oil particle can be subject to various forces making accurate predictions with a limited number of particles difficult. This is why many particles are needed to model the spill (North et al., 2011). Additionally, these particles substantially increase the computing power required for an accurate prediction.

This project aimed to design a simple process to model the dispersion of oil spills. To achieve this, it was essential to understand basic vector operations such as addition and subtraction of vectors, and scalar multiplication. As mentioned by Khan (2008), addition or subtraction is computed by adding or subtracting the components of each vector. Also, scalar multiplication is accomplished by multiplying both components by the scalar (Khan, 2008). Visualizing vector operations was important in understanding how the different vectors are added together with different weights to produce an "average" vector. Along with vector operations, an understanding of elementary calculus was essential. Concepts such as derivatives, integrals, and substitution needed to be thoroughly understood. As Dawkins (2018) taught, derivatives represent the slope of a curve at a given point, whereas integrals represent the area under the curve or the summation of little pieces of area. Substitution is a method to help compute derivatives and integrals (Dawkins, 2018). In addition to integrals, the "shoelace formula", $A = \frac{1}{2} abs\left(\det \begin{bmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{bmatrix}\right)$, where x_i and y_i are vertices, is helpful for calculating the area of polygons (Ahmed et al., 2018). Being able to visualize and understand what each mathematical expression, equation, and transformation represented was essential to the solution of the oil spill computational model. For example, Green's theorem-the relationship between line integrals and double integrals-is hard to understand without the proof (Dawkins, 2018).

Unlike other methods that use a complex Lagrangian approach (North et al., 2011), this project attempted to model just the border or the circumference of the oil spill—this reduced computational power, as the circumference of the spill is sufficient in determining the area of the spill. A model that requires less computational power without sacrificing the accuracy of the oil spill prediction is important because many countries, including third world or developing countries are gaining access to oil drilling techniques. However, the cost associated with expensive computational equipment such as supercomputers or high-speed microcomputers may be cost prohibitive to the third world or developing countries. Countries like Somalia now have the capability to drill for oil but they do not have access to the supercomputer required to simulate thousands of particles in case of a spill.

Methods

The supervisor for this project was Clive Menezes.

Materials

- Computer for internet access and Excel

Procedures

The general concept for calculating the area of an oil spill was divided into five steps. In the first step, the geographical epicenter of the spill was obtained from news reports or other well log data that was recorded with a governmental agency. Using this geographical position, the ocean current velocity vector was obtained from the Ocean Surface Current Analysis Real-time, OSCAR, database, located on the Physical Oceanography Distributed Active Archive Center, PO.DAAC, website (Earth and Space Research, 2009). Due to the one-third degree latitude resolution of sampling from OSCAR, a position circle was set up to obtain the ocean current velocity vector at eight cardinal positions on the circumference of it i.e. at 0^o, 45^o, 90^o, 135^o, 180º, 225º, 270º, and 315º. The radius of the position circle was estimated to be the amount the oil spill was expected to disperse on the surface of the ocean after some time interval.

More precisely, a computer was used to obtain the ocean current velocity data for the spill's epicenter from the OSCAR, database via the Internet. The current velocity vector in the OSCAR database was for a given position. As a result, the current velocity vector for the epicenter had to be interpolated from the OSCAR database information. In the following formulas, for interpolating, real() is a function that only keeps the real part of complex numbers, [u', v'] is the velocity vector at (x, y), [u_i, v_i] is a velocity vector at (x_i, y_i) given by OSCAR, and d is 3700, the linear distance between two points from OSCAR. $u' = \sum u_i w_i$, $v' = \sum v_i w_i$, $w = \frac{a}{\sum a}$, $a = real\left(\sqrt{\frac{d-\sqrt{(x_i-x\prime)^2+(y_i-y\prime)^2}}{d}}\right)^2$ Next, a position circle, as described in step one, was estimated and the current velocity vector at each of the eight cardinal points was calculated. These eight current velocity vectors were calculated using the process described above. Since vectors from OSCAR are positioned in a grid, the process, above, utilizes a weighted sum to calculate the magnitude of the vector components. Thus, the farther (x', y') was from a point on the grid, (x_i, y_i), the less effect the vector [u_i, v_i] had on the vector [u', v'].

In step two, the ocean current velocity vectors at the eight cardinal points, obtained in the previous step, were assumed to operate on points on the circumference of an oil spill of an arbitrary radius, i.e. the circumference of the initial oil spill. In this project a radius of five meters was chosen.

In step three, the ocean current velocity vectors for the eight cardinal points on the surface of the exemplary circle in step two were extrapolated to obtain the new positions of the cardinal points at the given time.

In other words, the movement of cardinal position particles at the surface of the ocean was found using the function: f(x, y)=(x, y)+.4(Δt)[Ocean Currents], where Δt represents elapsed time from the start of the spill.

Using the new positions of the cardinal points, step four derives a boundary utilizing three different methods: piecewise straight lines, a circle fit, and a piecewise cubic fit. Step four also calculates the area of the spill circumscribed by the given position was calculated using the shoelace formula, the formula for the area of a circle, and Green's theorem.

Since the current velocity vectors were only calculated for selected cardinal points, interpolation was necessary to solve for the entire boundary of the spill after the selected time interval. The piecewise straight-line interpolation scheme required straight lines to be plotted through the each of the eight cardinal points and to have their domains restricted to the two points they pass through. Since this utilized straight lines, the shoelace formula was used to calculate the area of the spill. The encompassing circle scheme attempted to encapsulate half of the cardinal points. To do as such, the eight cardinal points were converted from Cartesian to polar coordinates using the transformation $x = r * \cos\theta$ and $y = r * \sin\theta$ where (x, y) represents the coordinate of the point before the transformation and (θ, r) represents the coordinate of the point after the transformation. After the transformation, the average magnitude r value for the eight polar points were calculated. Using this average as a radial length, a circle was drawn to represent the region of the oil spill. Since this method produced a circle, the formula for the area of a circle was used to calculate the area of the spill. The region incorporating the eight cardinal points using the piecewise cubic fit scheme was determined as follows. In an interval $[x_i, x_{i+1}]$, the interpolant looked like $f(x) = d_i(x - x_i)^3 + c_i(x - x_i)^2 + y_i'(x - x_i) + y_i$ where $c_i = \frac{3s_i - 2y_i' - y_{i+1}'}{x_{i+1} - x_i}$, $d_i = \frac{y_i' + y_{i+1}' - 2s_i}{(x_{i+1} - x_i)^2}$, $s_i = \frac{y_{i+1} - y_i}{x_{i+1} - x}$, and y_i', y_{i+1}' were the derivative values at (x_i, y_i) and (x_{i+1}, y_{i+1}), respectively. These derivative values were calculated using the Steffen method:

$y_i' = (sign(\Delta_{i-1}) + sign(\Delta_i))min\left(|\Delta_{i-1}|, |\Delta_i|, \left|\frac{h_i\Delta_{i-1}+h_{i-1}\Delta_i}{2(h_{i-1}+h_i)}\right|\right)$ where $h_i = x_{i+1} - x_i$, $\Delta_i = \frac{y_{i+1}-y_i}{h_i}$, and sign() is a function that transfers the sign of the number in its parenthesis to the number one. The interpolant required four points to model an interval between two points. Because of this, the interpolant could not model the intervals near the edge of the boundary. To overcome this deficiency, the interpolant was found in terms of y instead of x near the edges. Since this method produced piecewise curves some of which were in relation to y and others x, integrals had to be used. To find the area of this boundary, the double integral could have been taken. However, double integrals would produce very large numbers so it was preferable to use utilize Green's theorem and convert the double integrals to line integrals.

Finally, in step five, the integrity of the model was verified. This was done by comparing predicted spill with historical spills.

In other words, the model was run on historical data to simulate previous spills. After running the model, the model's estimated areas were compared to the historical actual areas. Repeating steps one to four two more times helped to verify if the model was accurate. The aforementioned procedures are described in the flowchart of Figure 1 and explained in more detail below.

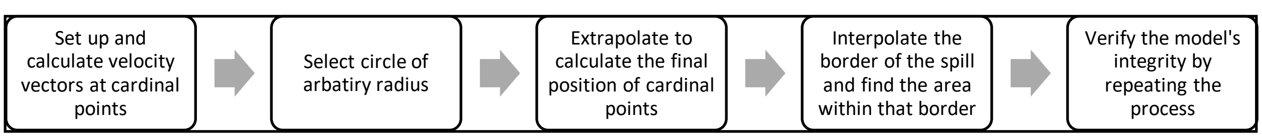


Figure 1. Flowchart of designed process

The process described above was formatted into excel sheets to reduce manual computation. This is explained in Appendix A.

Results

After, news articles about the Sanchi Oil Tanker Collision were found, the information from the articles including time and location of the spill, were inputted into the Excel file where the formulas for the model were located. The area of the spill was then calculated using three different methods. The areas calculated for the Sanchi spill can be found in the “Sanchi” column of Table 1. The newspaper gave the area of the spill as 101 km^2. Using a piecewise linear fit, the area was calculated as 180 km^2. The circle fit method resulted in an area of 241 km^2. Utilizing the piecewise cubic fit the area was outputted as 175 km^2.

News articles about the Bonga Oil Spill were found. The information from the articles, time and location of the spill, were inputted into the Excel file where the formulas for the model were located. The file, then, calculated the area of the spill using three different methods. The areas calculated for the Bonga spill can be found in the “Bonga” column of Table 1. The newspaper gave the area of the spill as 906 km^2. Using a piecewise linear fit, the area was calculated as 944 km^2. The circle fit method resulted in an area of 800 km^2. Utilizing the piecewise cubic fit the area was outputted as 668 km^2.

News articles about the Ennore Oil Spill were found. The information from the articles, time and location of the spill, were inputted into the Excel file where the formulas for the model were located. The file, then, calculated the area of the spill using three different methods. The areas calculated for the Ennore spill can be found in the “Ennore” column of Table 1. The newspaper gave the area of the spill as 0.034 km^2. Using a piecewise linear fit, the area was calculated as 0.002 km^2. The circle fit method resulted in an area of 0.315 km^2. Utilizing the piecewise cubic fit the area was outputted as 0.013 km^2.

Boundary obtainment method	**Sanchi**	**Bonga**	**Ennore**
Newspaper	101 km^2	906 km^2	0.034 km^2
Piecewise linear fit	180 km^2	944 km^2	0.002 km^2
Circle fit	241 km^2	800 km^2	0.315 km^2
Piecewise cubic fit	256 km^2	668 km^2	0.013 km^2

Table 1. Area affected by oil spills according to different boundary obtainment methods

To analyze the raw data, three t-tests were performed between the newspaper and the three regression methods. The tests determined if there was a significant difference between the newspaper reported area and the model calculated areas. The formula for a t-test is $t = \frac{\overline{x_1} - \overline{x_2}}{\sqrt{\frac{S_1^2 + S_2^2}{n}}}$. To compute a t-test, it is necessary to find the mean and standard deviation of each group. The newspaper had a mean of 335.678 and a standard deviation of 496.487. The piecewise linear fit had a mean of 374.667 with a standard deviation of 501.203. The circle fit had a mean of 347.105 with a standard deviation of 410.265. The piecewise cubic fit had a mean of 281.004 with a standard deviation of 346.380. These figures can be found in Table 2.

Boundary obtainment method	**Mean**	**Standard deviation**
Newspaper	335.678 km^2	496.487 km^2
Piecewise linear fit	374.667 km^2	501.203 km^2
Circle fit	347.105 km^2	410.265 km^2
Piecewise cubic fit	281.004 km^2	346.380 km^2

Table 2. Averages and standard deviation of the different boundary obtainment methods

Plugging in these values into the formula yields a t-value of 0.0957 between the newspaper and the piecewise linear fit, 0.0307 between the newspaper and the circle fit, and 0.1564 between the newspaper and the piecewise cubic fit. These t-values were converted to p-values using an Excel function. The first test converted to a p-value of 0.466. The second test converted to a p-value of 0.489. The third test converted to a p-value of 0.445. These t and p values can be found in Table 3. Using a p significance level of 0.05, these p values mean that there is no significant difference between the designed model and the benchmark-newspaper reports.

Boundary obtainment method	**t-value**	**p-value**
Piecewise linear fit	0.0957	0.932
Circle fit	0.0307	0.978
Piecewise cubic fit	0.1564	0.944

Table 3. t and p values for the different boundary obtainment methods

Discussion

A model for modeling an oil spill was developed and converted into three Excel files. The model takes information from newspapers and outputs a predicted oil spill region. Since newspapers do not report the exact spill region, the reported area affected by the spill was used to verify the integrity of the model.

To verify the integrity of the model, the area outputted from the model was compared to the newspaper's report using a t-test. The results show that there is not a significant difference between the area calculated by the designed model and the benchmark-newspaper reports-using a significance level of 0.05. The p-value between the newspaper and the piecewise linear fit was 0.932. The p-value between the newspaper and the circle fit was 0.978. The p-value between the newspaper and the piecewise cubic fit was 0.944. These p-values are recorded in Table 3. Since all of the methods were shown not to have a significant difference, the goal of this project was achieved. However, more testing needs to be done to verify this because it hard to determine significance when the sample size is small.

Based off of the results, it was concluded that the circle fit boundary obtainment method was most accurate of the three methods presented. This was known because the circle fit had the highest p-value, indicating that it was the least different from the benchmark. It was also determined that the piecewise linear fit was second most accurate. This was identified because the piecewise linear fit had the second highest p-value. This meant it was the second least different from the benchmark. Finally, it was resolved that the piecewise cubic fit was the least accurate. This was recognized because the piecewise cubic fit had the lowest p-value. This meant it was the most different when compared to the other two methods.

The model could possibly obtain even higher p-values if the sources of error in the model were reduced. These errors arose from three main issues: the curvature of the Earth, assumptions made while creating the model, and accuracy/lack of input data. The first issue caused the area calculated by the model to be farther from the benchmark near the poles. This was because when converting from longitude to distance from the epicenter of the spill, a constant ratio was used instead of a variable ratio. A constant ratio did not take into account the fact that the distance between degrees of longitude shrinks to 0 at the poles. The second issue specifically refers to assuming oil spills will spread in a radial motion. When calculating the position circle, mentioned in step one, it was assumed that the oil would spread radially. However, this is not always the case. If a spill were to spread linearly, in one direction, the area calculated by the model would be farther from the benchmark. The third issue comes from a lack of input data. This means that there was not enough information about the initial conditions of the spill to provide an accurate model. For example, the distance from the surface of the ocean and the initial velocity of the oil were not provided in the news articles so these had to be approximated- the fiver meter arbitrary circle mentioned in step two. These approximations probably caused the areas calculated to be off from the benchmark slightly. The third issue of error also mentions the problem with the accuracy of input data. A case of this is the fact that the times provided were only accurate to days. If specific times were given, the area calculated by the model would be closer to the benchmark. Another situation of this was, the resolution of OSCAR was one-third of a degree of latitude. If the resolution of OSCAR was higher, the velocity vectors would not have been interpolated. In the end, the curvature of the Earth, assumptions made while creating the model, and the accuracy/lack of input data caused the majority of error in this project.

This project is relevant to the oil field because revenue loss associated with oil spills is quite high. To minimize this loss, it is necessary to quickly contain the spill. This project can help accomplish this by modeling the position of the spill. In order for this project to be applicable to real situations it would be important to reduce the error in the model. This can be done by taking into account the curvature of the Earth. To do this, the whole process can be done in relation to latitude and longitude instead of converting it to distance, in meters. Or, another model can be made with the assumption that the spill will move linearly. Then, a test can be used to determine which model would be most effective for that situation.

There are many factors and anomalies that occur in the physical world. These are what make an accurate model almost impossible.

References

Ahmed, D., Anh, V., Arashiro, T., Azarnoush, A., Goh, Z., Han, J., ... & Vreugdenhil, A. (2018, June 6). Area of a Triangle. Retrieved December 9, 2018, from https://brilliant.org/wiki/triangles-calculating-area/

Dawkins, P. (2018, June 6). Calculus I - Pauls Online Notes. Retrieved October 7, 2018, from http://tutorial.math.lamar.edu/Classes/CalcI/CalcI.aspx

Earth and Space Research. (2009, May 1). OSCAR third degree resolution ocean surface currents. Retrieved from https://podaac.jpl.nasa.gov/dataset/OSCAR_L4_OC_third-deg

Khan, S. (2008, September 4). *Vector operations review*. Retrieved October 7, 2018, from https://www.khanacademy.org/math/precalculus/vectors-precalc/combined-vector-operations/a/vector-operations-review

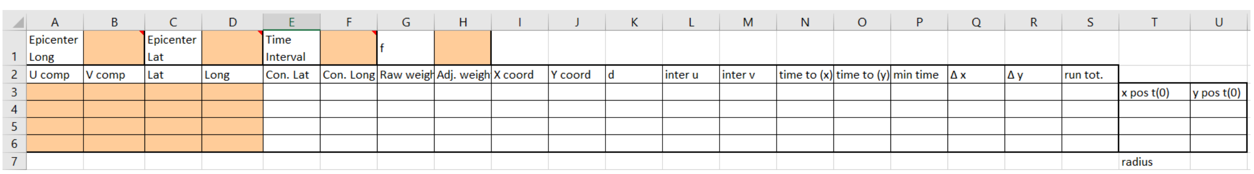
North, E. W., Adams, E. E., Schlag, Z., Sherwood, C. R., He, R., Hyun, K. H., & Socolofsky, S. A. (2011). Simulating oil droplet dispersal from the Deepwater Horizon spill with a Lagrangian approach. *Geophys. Monogr. Ser*, *195*, 217-226.

Olascoaga, M. J., & Haller, G. (2012). Forecasting sudden changes in environmental pollution patterns. *Proceedings of the National Academy of Sciences*.

Özgökmen, T. M., Chassignet, E. P., Dawson, C. N., Dukhovskoy, D., Jacobs, G., Ledwell, J., ... & Poje, A. C. (2016). Over what area did the oil and gas spread during the 2010 Deepwater Horizon oil spill?. *Oceanography*, *29*(3), 96-107.

Appendix A

To reduce manual computation, the steps contained in the procedures were converted to formulas in Excel. In Figure A1, the position circle radius is estimated. Cells B1, D1, and F1 were used as inputs for the epicenter longitude, the epicenter latitude, and the time interval respectively. Regions A2:D6 and A7:D10 were used as input for the u components, the v components, the latitude, and the longitude of ocean current velocity vectors, respectively. Region E2:M6 and E7:M10 interpolated vectors based on the inputs. Columns M and N calculated the time it would take for the particle move into another set of vectors. Column P took the minimum value of column M and N. Using the minimum time to leave the set of vectors, cell P, columns Q and R calculated the x and y position of the particle. Column S kept a running total of the time. When the total goes to zero, about two iterations. Columns T and U calculate the approximated median radius-U7.

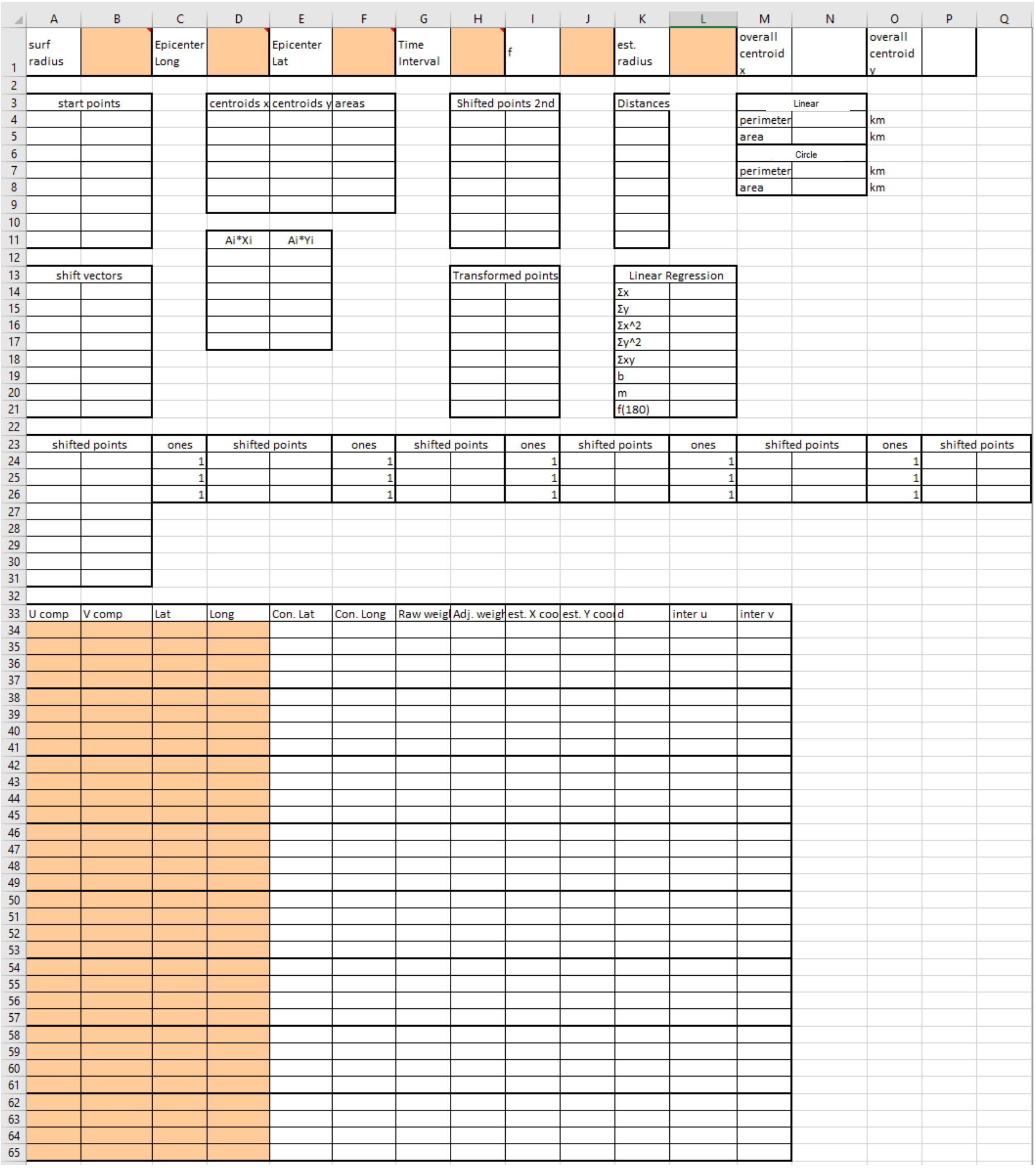


	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
1	Epicenter Long		Epicenter Lat		Time Interval		f														
2	U comp	V comp	Lat	Long	Con. Lat	Con. Long	Raw weigh	Adj. weigh	X coord	Y coord	d	inter u	inter v	time to (x)	time to (y)	min time	Δ x	Δ y	run tot.		
3																				x pos t(0)	y pos t(0)
4																					
5																					
6																					
7																				radius	

Note. Shaded boxes signify areas for input.

Figure A1. Excel file for estimating the position circle radius

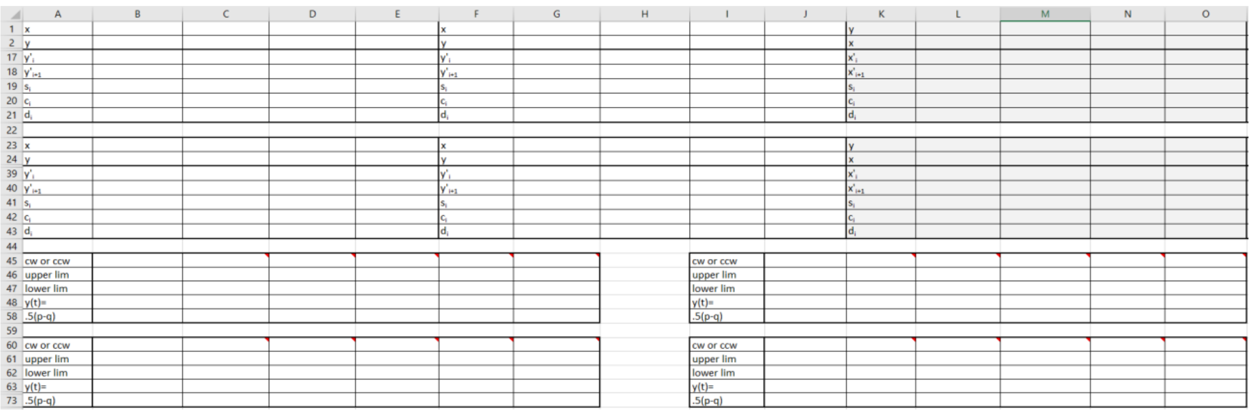
In Figure A2, regions A33:A65, B33:B65, C33:C65, D33:D65 were used as input for the u components, the v components, the latitude, and the longitude of vectors, respectively. Cells B1, D1, F1, H1, L1 were used as inputs for the radius of the circle at the surface, the epicenter longitude, the epicenter latitude, the period of time the spill takes place, and the median radius, respectively. Region A3:B11 calculated the boundary points using the inputted surface radius. Region E33:M65 interpolated vectors for the boundary points. Region A13:B21 multiplied the results from E33:M65 by the time interval and .4 to convert the vectors from meter per second to displacement. In region A23:B31, the shift vectors were added to the start points. Region D23:R26 formatted region A23:B31 so that the shoelace formula could be performed. In cells D3:E9, the spill is divided into sub triangles and the centroids of these triangles are calculated. In F3:F9 the area of these triangles are calculated using the shoelace formula. Region D11:E17 multiplies the areas by the centroid as an intermediate step for finding the overall centroid of the spill. Cells M1:P1 calculated the overall centroid of the spill. Region H3:I11 shifts the points again so that they are centered around the origin. H12:I21 performs a polar transformation on the points. Cell N5 and N8 calculate the area of the spill using a piecewise linear fit and a circle fit, respectively.



Note. Shaded boxes signify areas for input.

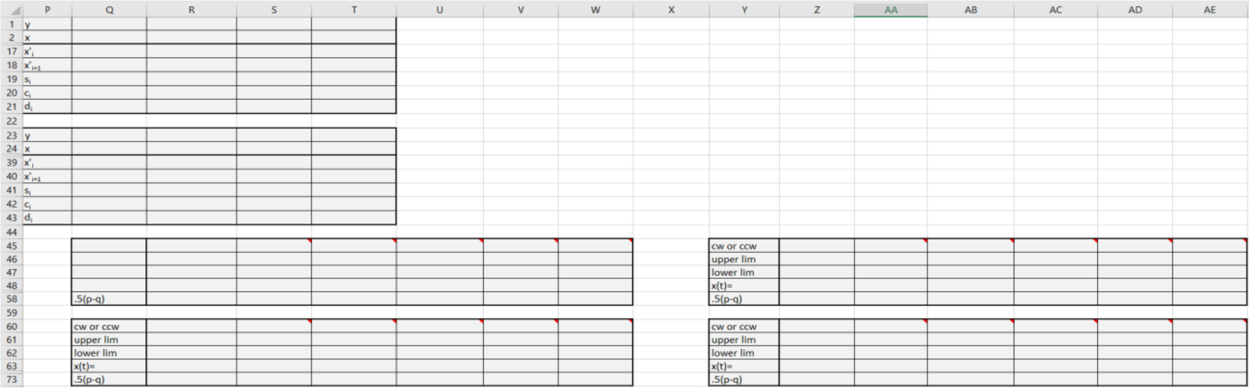
Figure A1. Excel file for extrapolating cardinal points and calculating simple area

In Figure A3, the twice shifted points are used. Only one combination of four points will be discussed because the process is just repeated for the other combination of points in different regions. In region A1:E21, monotonous cubic interpolation is carried out. Cell C21, C20, C17, and C2 are the coefficients of a 3rd order polynomial. These coefficients are moved to region D48:G48. In cell B45 the direction of integration is specified, clockwise or counter-clockwise. In B46 the upper limit of the line integral is inputted. B47 is the lower limit of the line integral. Region A48:G58 simulate integration for line integrals by moving the coefficients to different columns, representing a change to their degree, and diving them by that columns' degree. The output is given in cell B58. The total area is a sum of all of the individual line integrals, calculated in cell AE74.



Note. Shaded cells signify the formulas are in terms of y

Figure A3. Excel file for calculating area of the cubic interpolated boundary



Note. Shaded cells signify the formulas are in terms of y

Figure A3 Continued. Excel file for calculating area of the cubic interpolated boundary