Simulation for Transatlantic Navigation of Small Sailing Vessels

CX 4230 Group 33

James Theriault, jtheriault7@gatech.edu

Sandaru Fernando, sfernando8@gatech.edu

Sean Fish, sfish7@gatech.edu

https://github.gatech.edu/sfish7/ocean-sim

Video

Abstract

Autonomous navigation of small-scale sailboats across the Atlantic Ocean is a challenging task, with only one successful transatlantic crossing recorded for vessels less than 2.4 meters in length (The Microtransat Challenge, n.d.). This project provides a simulation to inform optimizations of the path for this task, specifically as a resource for autonomous vehicles of this nature. This is not intended to determine the perfect route for a vessel, but instead to provide a tool for planning.

The utility of simulation is limited by its ability to accurately model factors of interest, and model validation benefits greatly from data. However, since the data is scarce for small-scale transatlantic crossings, this project leverages data and models relating to identified key factors to synthesize a simulation. This simulation will account for conditions in the Atlantic Ocean that will significantly impact the navigation of a small sailboat. Factors that will be examined include wind, surface currents, and maritime traffic.

Project Description

Successful transatlantic crossings by small autonomous sailing vessels are difficult and require planning. When running fully autonomously, these vehicles must be equipped with information to make decisions without human intervention. The goal of this project is to create a simulation that serves as a tool for creating navigational resources for such vessels. One potential example is the generation of cost maps through simulated trials to determine areas of higher risk.

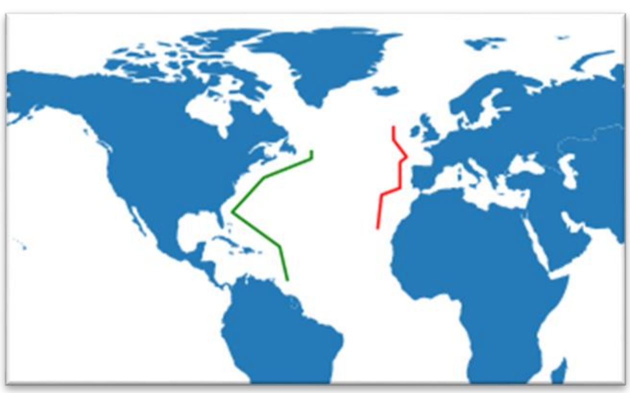


Figure 1 - Map with lines defining successful transatlantic crossings

The major relationship of interest in this project is that between the sailing vessel and the ocean. To investigate this, the modeling of several aspects of the ocean will be required. However, the ocean itself is a vast and complex system, and creating a model which incorporates all its nuances is far beyond the scope of this simulation.

This project is not centered on ocean dynamics and simplifying assumptions will be made to model only the components most relevant to navigation. Since sailing vessels move with the waters they are in and utilize wind for propulsion, surface currents and winds are the two primary factors we are considering. Additionally, our simulation will model some of the hazards that certain regions of the ocean may contain. Specifically, we investigate the density of larger vessels which may interfere with operations of autonomous boats. This was chosen due to the number of failures in transatlantic crossing attempts caused by larger vessels like fishing boats picking up the autonomous vehicles (The Microtransat Challenge, n.d.).

In addition to the model of the environment, a model of the vehicle itself is required to characterize the relationship between them. The dynamics of sailboats are also rather complex and can vary significantly based on hull and sail configuration. To simplify the problem, there will be assumptions made regarding the dynamics of the sailboat as well. The simulation of the sailboat will not involve detailed physics calculations as one may find in common robotics simulators. For this project, the sailboat dynamics will be simplified to that of a lightweight and passive system without active navigation, with the potential of future expansion to include navigation.

Literature Review

Relevant Factors

Gibbons-Neff and Miller (2011) recognized several factors that are important for the task of navigation. These factors include vehicle dynamics, winds, currents, sea states, and sunlight, in addition to other seasonal and human influences.

Vehicle Dynamics

The state of a sailboat can be represented as position, velocity, angle, and wind direction (Melin, Dahl, & Waller, 2015). This state can be modeled using various nonlinear differential equations. The dynamics of sailboats are time-variant and stochastic, and some of the interactions between the sailboat and its environment can be difficult to model (Abrougui & Nejim, 2019). Models can be incorporated into control algorithms and autopilots that modify the rudder angle and sail angle to control the movement of the sailboat, with promising results. Given the utility of modeling vehicle dynamics, these characteristics will be modeled in the simulation. Due to the long-range nature of this simulation, the model of the vehicle dynamics may be highly abstracted.

Environmental Factors

Climatological data from NOAA Pilot Charts were utilized by Gibbons-Neff and Miller (2011) for long-term planning. These charts, which present data for specific months, discretize bodies of water into cells, each with statistical information on winds, prevailing currents, gales, and sea ice. The data was used to calculate ranges of traversal times, which were then used to estimate the desirability of the area. These charts also presented hurricane probabilities, which were also incorporated into planning.

Du et al. (2018) similarly utilize grid discretization for the purpose of path-planning and also use slope fields for modeling wind flow. Regarding path estimations, they find that the total length of a route is positively correlated with minimum voyage time. However, when sailing upwind, they find that there is a critical value under which the minimum sailing time will increase as the route length decreases, due to the need for tacking maneuvers to progress. This can improve time estimates for vehicles and exemplifies how simulations of relevant factors can aid navigation. Wang et al. (2019) also utilize slope fields to model surface currents for optimizing energy consumption.

Techniques for simulating maritime traffic have utilized cellular automata and statistical distributions for environmental forces to estimate velocities for large vessels. (Qi et al., 2017). The acceleration of the ship is described to have two components, a regular and a random one, considering the effects of weather and driver behavior.

Abrougui and Nejim (2019) contribute the use of current in addition to wind for modeling sailboat navigation to yield more accurate simulation results. Using physical simulations of wind and current may assist in characterizing behaviors across larger distances for the purpose of long-range planning.

When modeling the ocean, there are complex interactions to consider. For example, Echevarria et al. (2021) describe non-linear interactions between wind-waves and mean flow in which each factor modifies the properties of the other. Deng et al. (2009) investigate interactions between wind stress and ocean surface currents, finding that the relationship can be used to improve ocean circulation models. Physical behaviors are difficult to characterize, and this is one of the reasons the simulation will not directly model the underlying physical mechanisms of the ocean. However, investigations into physical behaviors reveal trends which may be simplified and modeled in the higher-level long-range simulation.

Conceptual Model

The environmental model consists of a map of the northern Atlantic Ocean, with axes defined by latitude and longitude. To enable greater computational efficiency, simple assignment of environmental characteristics, and extensibility, and due to common use in vehicle path-planning, the map is discretized into a grid. The environmental characteristics of each grid cell is determined by aggregating the datapoints from existing ocean datasets. The grid represents several qualities:

- Surface Current component vectors (m/s)
- Wind component vectors (m/s)
- Density of maritime traffic (Number of AIS transponders in a grid cell)

The model of the vessel can be described as:

- Latitude position (degrees)
- Longitude position (degrees)
- Surface Current Drag Coefficient
- Wind Drag Coefficient

The properties of each grid cell will be used to determine the effect of the ocean conditions on the vessel. The environmental model is like a set of flow field diagrams, and our trajectories are calculated in a manner akin to Euler's method. The step size is associated with the time step of each simulation update and the coefficients of drag of the vehicle.

The drifting vessel does not have any control, and its trajectory is solely defined by the environment, position, and drag characteristics.

Termination Conditions

The simulation will be terminated when any of the following occur:

- Beached vessel: The vessel moves off valid ocean grid cells (onto land)
- Interference: The vessel is picked up by a simulated vessel
- Maximum time steps: The simulation has reached maximum specified time steps
- Invalid run: The vessel moves off the map

General Assumptions

Reducing the scope of the simulation makes the creation of the model more tractable. Assumptions were made to eliminate factors which are out of scope for the current model. The following assumptions were made for this simulation:

- Currents, winds, and maritime traffic are consistent throughout the year
- The vessel has no active control
- Component vectors of currents and winds are independent and modeled by Gaussian noise
- Areas beyond the North Atlantic have no effect on the environment
- A degree of latitude and longitude can be approximated as 111000 meters (United States Naval Academy, 2019)
- Simulation runs that exceed the boundaries of the map are invalid
- There is a 10% chance a maritime vessel in range will interfere with the autonomous vessel if it is in range
- Visibility of maritime vessels is 450 meters (Eckhardt, 2012)

Simulation Model

Implementation

This simulation is implemented in the Python programming language in a Jupyter Notebook environment. Development made use of Google Colab, which allows for the use of cloud resources and enables collaborative editing. There were two components of the simulation that made heavy use of libraries:

- Geographical Data – GeoPandas and GeoPy assisted in working with, processing, and visualizing geographical data
- Visualization – Matplotlib and Plotly enabled creating custom visualizations with geographical data

Running the Model

The easiest way to get started with the simulation involves first cloning the Git repository, and then uploading the .ipynb file as a Colab Notebook. Once the notebook is connected to a runtime, data.zip should be uploaded. After this, all cells should be run consecutively. The notebook will take about 10 minutes to run.

Environmental Model

Surface Currents and Winds

The simulation model utilizes data available from resources such as NOAA (NOAA, 2022) and Earth & Space Research (Dohan, 2022) to plot winds and surface currents at various points on the ocean.

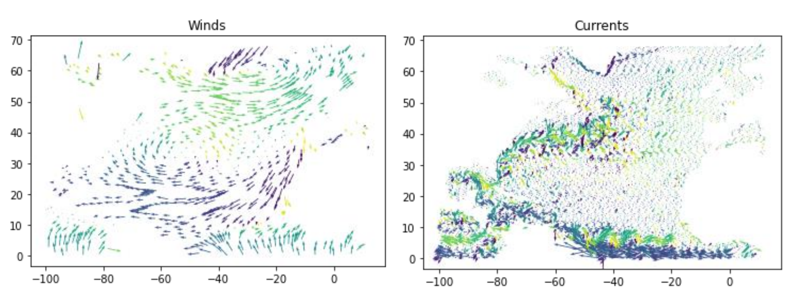


Figure 2 - Surface Currents and Winds Data

Maritime Traffic

The simulation model incorporates AIS data from the NOAA Office for Coastal Management (2022) for modeling maritime traffic.

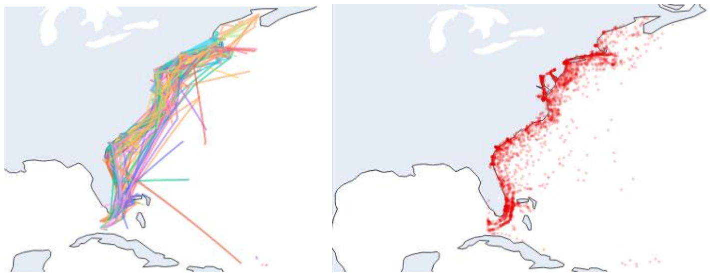


Figure 3 - Data visualization of maritime traffic

Regridding

To simplify the simulation model, the surface winds dataset and surface current dataset are regridded into a discrete grid, through the use of an aggregation function like the mean.

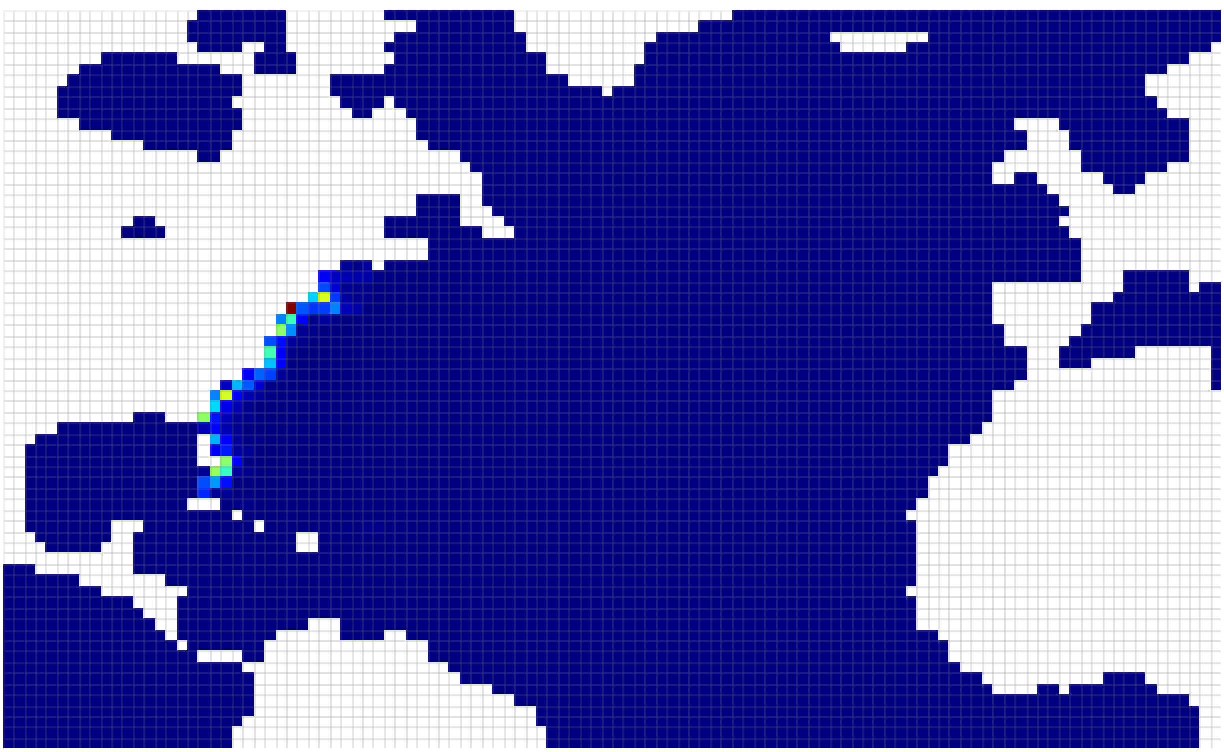


Figure 4 - Regridded vessel density

The regridding function is shown on the next page.

```python
def grid_data(gdf: geopandas.GeoDataFrame, scale: tuple, dim: tuple, aggfunc='mean'):
  """
  Args:
    gdf: GeoPandas Dataframe to be gridded
    scale: 2-tuple with the size of cells in x,y directions (degrees)
    dim: Total dimensions being considered

  Returns:
    cell: New gdf with average gridding applied
  """
  # Create cells

  # Properties of the grid:
  if dim is not None:
    xmin, ymin, xmax, ymax = dim
  else:
    xmin, ymin, xmax, ymax = gdf.total_bounds

  x_cell_size = scale[0]
  y_cell_size = scale[1]

  print(x_cell_size, "x size ", y_cell_size, "y size")

  # Iterate over all cells and create bounding box
  grid_cells = []
  for x0 in np.arange(xmin, xmax, x_cell_size):
    for y0 in np.arange(ymin, ymax, y_cell_size):
      x1 = x0+x_cell_size
      y1 = y0+y_cell_size
      grid_cells.append(shapely.geometry.box(x0,y0,x1,y1))

  print(len(grid_cells), "cells created")

  # Collect all boxes into geodataframe
  cell=geopandas.GeoDataFrame(grid_cells, columns=['geometry'])

  # Dissolve data into cells by averaging
  merged = geopandas.sjoin(gdf, cell, how='left', op='within')
  dissolve = merged.dissolve(by='index_right', aggfunc=aggfunc)

  # Transfer data to cells
  for column, data in dissolve.iteritems():
    if column != "geometry":
      cell.loc[dissolve.index, column] = data.values
  return cell
```

Figure 5 - Regridding Code

Vehicle Model

For the scope of this project, the vehicle model is simplified as coordinates in terms of latitude and longitude with an assumed drag coefficient of 1 for surface currents (NOAA's Atlantic Oceanographic and Meteorological Laboratory. (n.d.)) and winds. In the context of the simulation, this indicates that the surface current and wind speed will directly transmit to the vessel speed without scaling.

Simulation

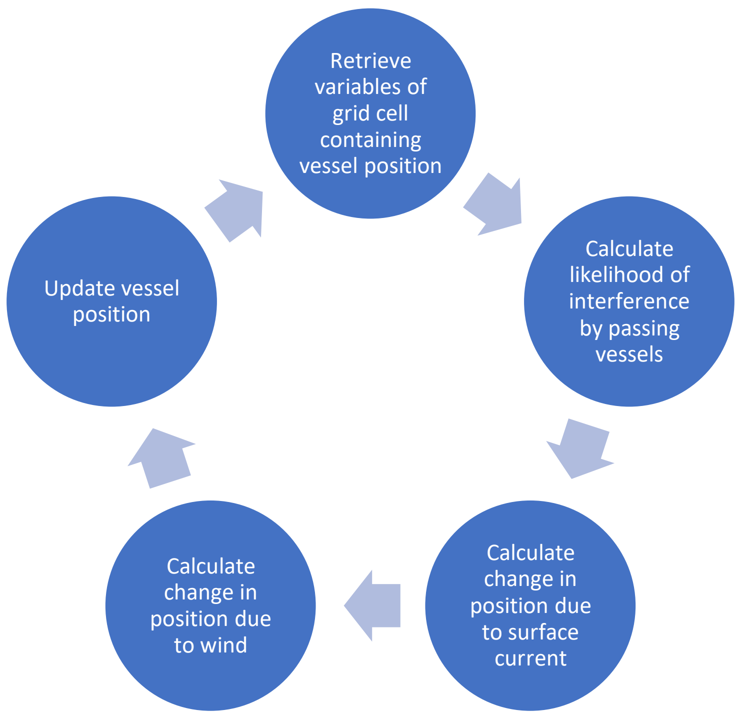


Figure 6 - Simulation Loop

The simulation can be described as a simple loop, where termination conditions are evaluated and new position updates are calculated based on environmental conditions and the size of the timesteps.

Simulation code on the next page.

```python
def simulation(gdf, start_coords, vehicle_params, max_iterations=50,
               rng=None, timestep=10000):
  # Start condition
  coords = [start_coords]
  grid_cell_area = 111000.0**2
  curr_coeff = 1
  factor = timestep

  # Simulation loops until max_iterations or termination condition
  for i in range(max_iterations):
    co = coords[i]
    try:
      # Current vector components
      curr_x = gdf.cx[co[0], co[1]].head(1).u_x.values[0]
      curr_y = gdf.cx[co[0], co[1]].head(1).v_x.values[0]
      # Wind vector components
      wind_x = gdf.cx[co[0], co[1]].head(1).u_y.values[0]
      wind_y = gdf.cx[co[0], co[1]].head(1).v_y.values[0]
      # Ship density
      num_ships = gdf.cx[co[0], co[1]].head(1).n.values[0]
    except:
      if (within_area(gdf, co[0], co[1])):
        return ('Termination: Beached', coords)
      return ('Termination: Invalid', coords)

    # Calculate interference occurence
    if num_ships > 0 and rng is not None:
      pr_int = prob_interference(num_ships, grid_cell_area)
      r = rng.random()
      if r < pr_int:
        return ('Termination: Interference', coords)

    # Get current deltas
    cd_x, cd_y = current_delta(curr_x, curr_y, coeff=1)
    # Get wind deltas
    wd_x, wd_y = wind_delta(wind_x, wind_y, coeff=1)
    # Calculate new position
    d_x = factor * (cd_x + wd_x)
    d_y = factor * (cd_y + wd_y)
    new_x = co[0] + d_x
    new_y = co[1] + d_y
    if math.isnan(new_x) or math.isnan(new_y):
      return ('Termination: Beached', coords)

    # Add new position to list
    coords.append([new_x, new_y])

  return ('Termination: Max Time Steps', coords)
```

Figure 7: Simulation code

Termination Conditions

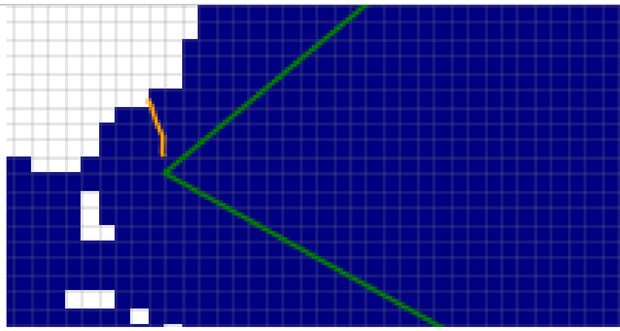
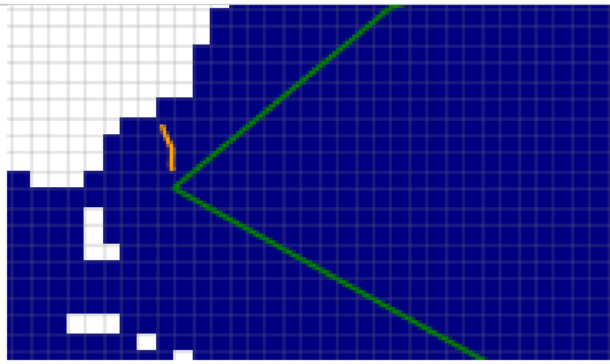
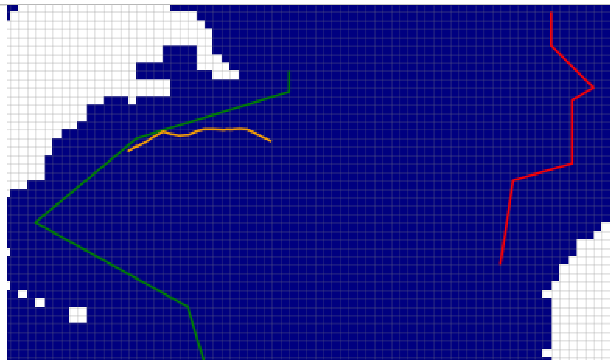
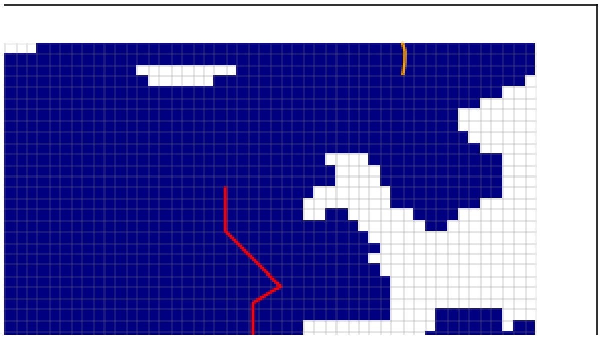
Beached	
Interference	
Maximum Timesteps	
Invalid Run	

Table 1- Termination Condition Examples

Verification

The simulation was verified as a faithful implementation of the conceptual model through iterative development. Variables and values were checked between data frames to ensure validity of data being imported. The conceptual model was simple enough that the code implementation did not require significant verification.

Experimental Results and Validation

This simulation was largely developed as a potential resource for generating cost maps and testing navigation strategies for the task of crossing the North Atlantic Ocean. For this project to be useful, we first need to validate the trajectories generated and make comparisons to real examples.

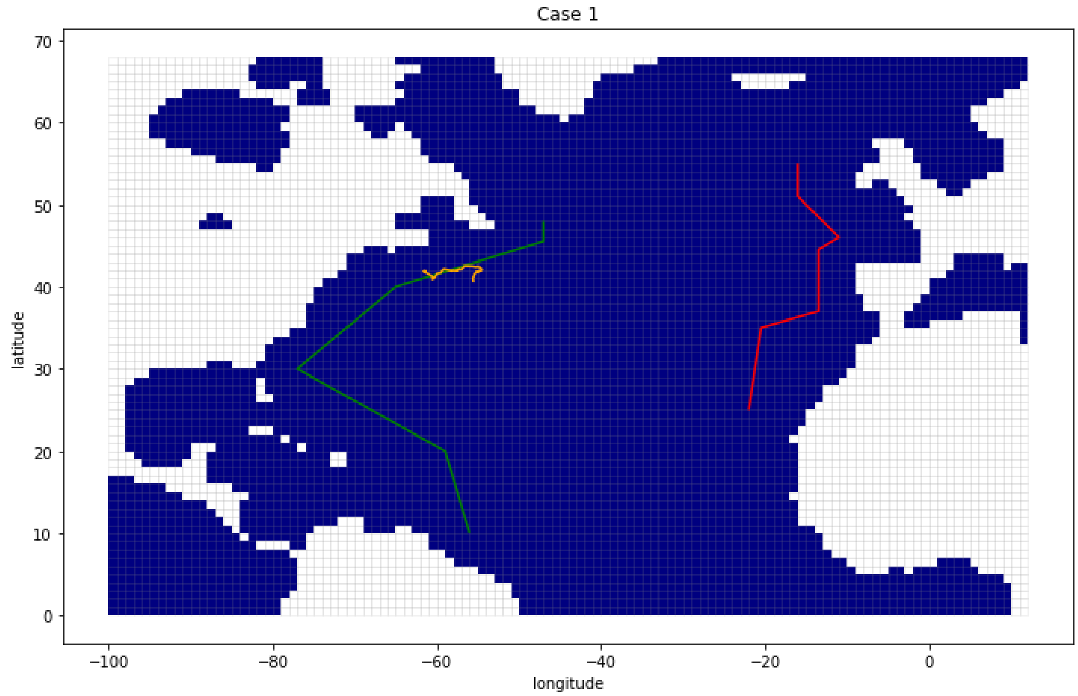
For trajectory validation, trajectories from NOAA Fisheries (n.d.) were used as base cases for comparison. Three cases are compared, with 50 timesteps used for the validation. The timesteps are adjusted to match the time between position reports in the base cases. Interference is disabled for this validation. Two metrics are used:

1. Average distance error by timestep
    - The simulated trajectory is calculated starting at a reported position in the base case
    - The distance error is found between the next reported position and the next simulated position
    - The process repeats from the next reported position in the base case
    - The average distance error is reported
2. Total distance error
    - The distance error is found between the base case at the final timestep and the simulated run at the final timestep

Metric 1 is calculated once. Metric 2 is calculated 5 times and averaged, with the environment varied between runs. During the development of these metrics, it was discovered that beaching is extremely common when starting near the shore. The starting timestep is adjusted for Metric 2 to avoid the issue of immediate beaching skewing results. The timestep for all 3 cases is 43200 seconds, or 12 hours.

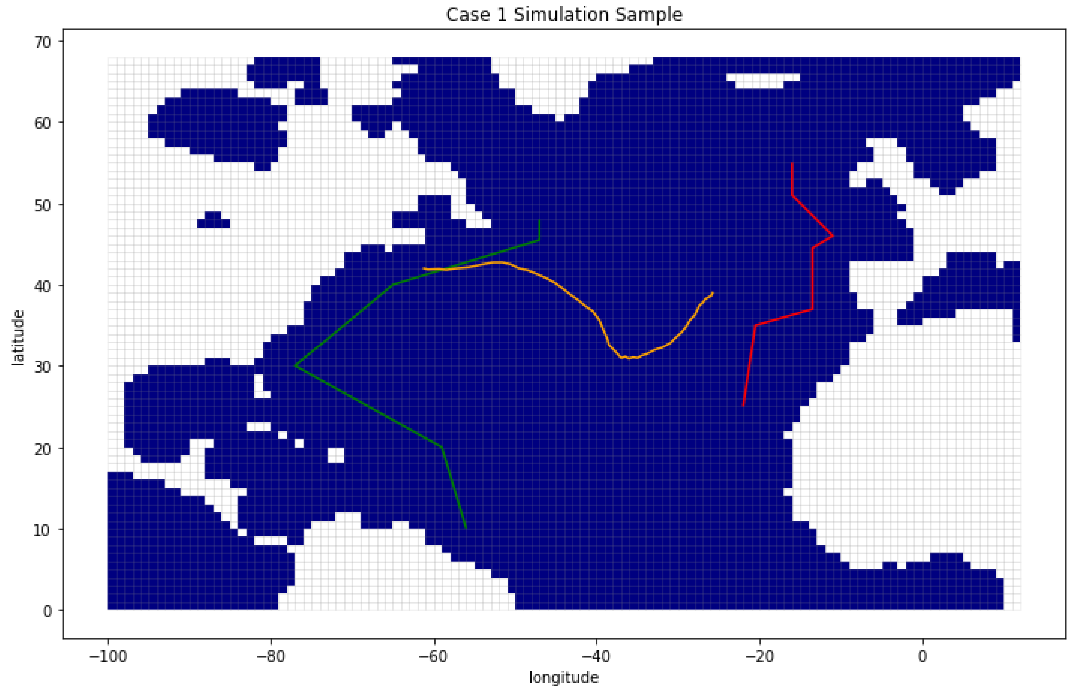
We expect the total distance error to increase over time due to the propagation of error. However, it is possible that the trends in ocean currents would result in less error overall than expected.

Case 1



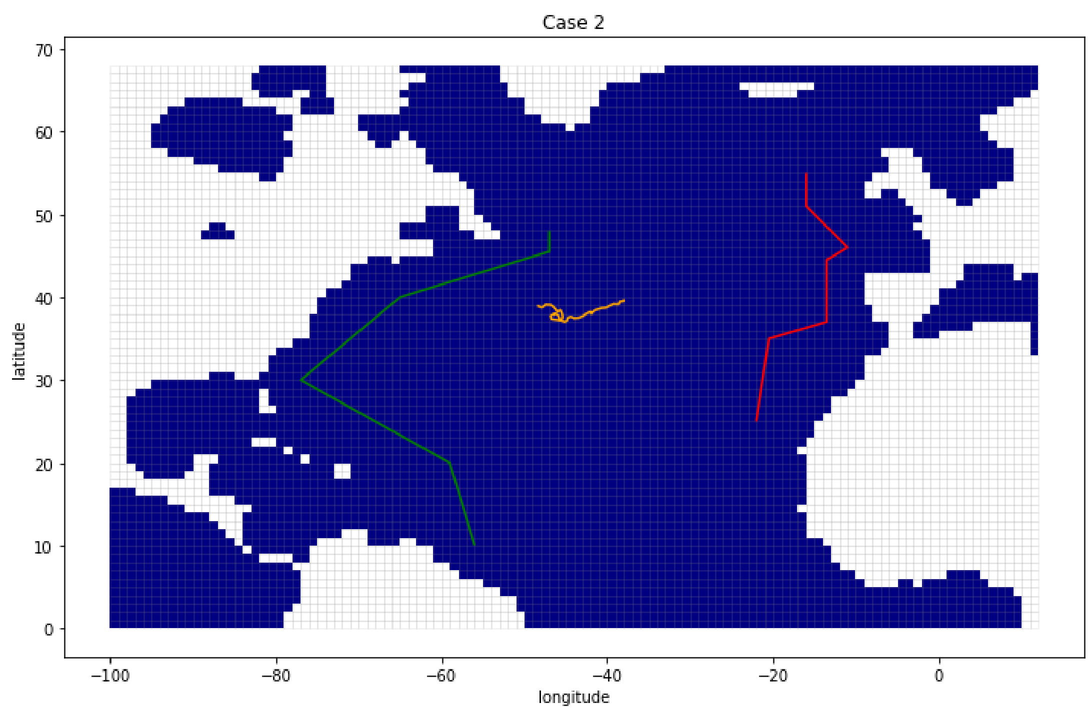
	Average Error	Standard Deviation
Metric 1	.740	.250
Metric 2	31.393	0.977

We find that the average error for each timestep is approximately .737 of a degree, the length of one side of a grid cell. The standard deviation is rather high, indicating high variability in this metric. The average error is much higher for the full path, which is expected. However, this distance is a large proportion of the total path distance.



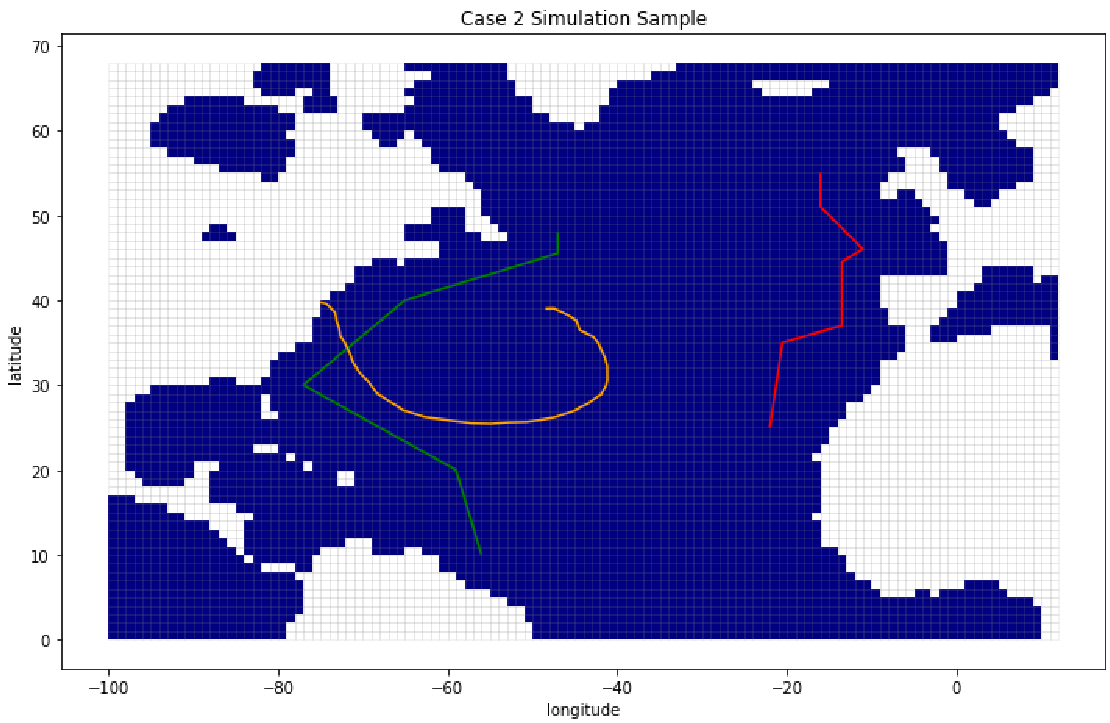
By sampling a simulation run from Metric 2, we see that the simulated vessel drifts far farther than that of the base case. However, the initial part of the trajectory does move towards the same general direction.

Case 2



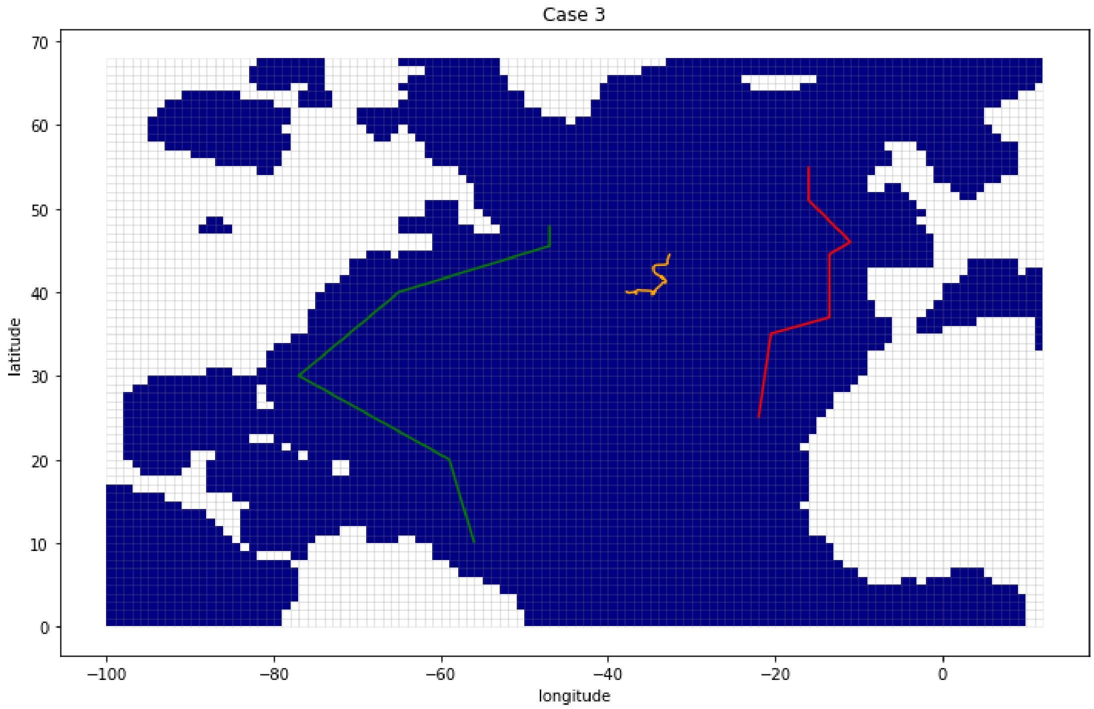
	Average Error	Standard Deviation
Metric 1	.812	.276
Metric 2	34.227	2.343

The results for case 2 are comparable to those of case 1. However, one of the simulation runs for Metric 2 was terminated due to the vessel being beached, while all the others were due to the maximum time steps being reached as expected. This is unexpected as the location is in the middle of the ocean.



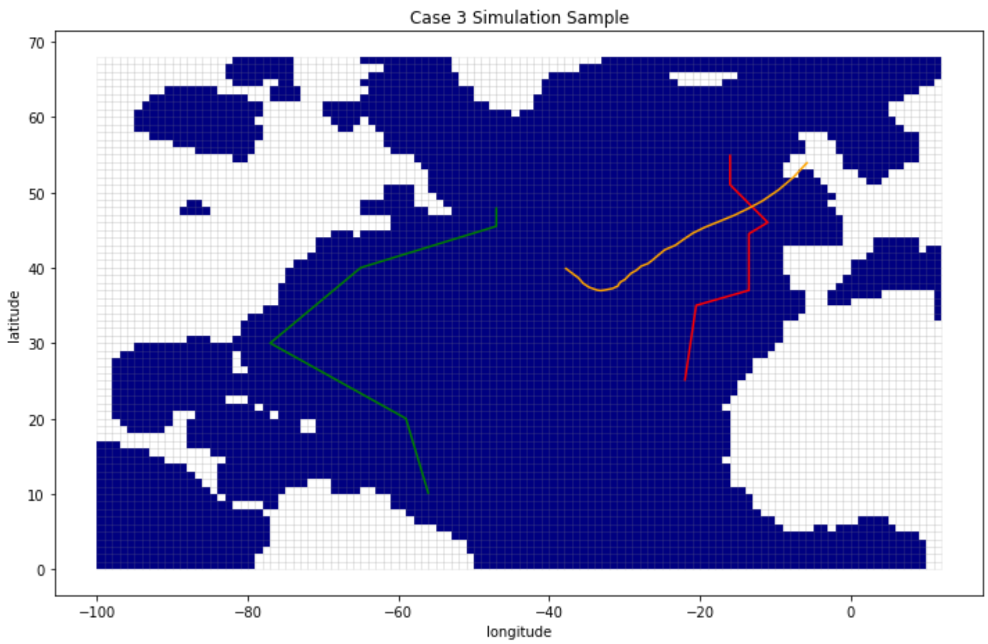
From the sample from the Metric 2 runs, we can see that the projected trajectory goes much farther than the base case.

Case 3



	Average Error	Standard Deviation
Metric 1	1.073	.407
Metric 2	28.053	1.749

In case 3, the error and standard deviation for Metric 1 both increase, while that of Metric 2 is slightly less so than the previous cases. However, the termination case for all 5 runs is beaching, which suggests that the smaller error may be the result of the trajectory being cut short by reaching land.



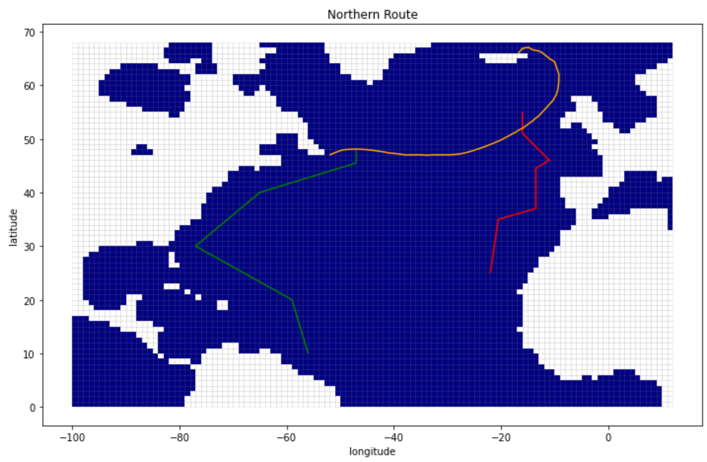
The sample from the Metric 2 runs shows once again that the traveled distance is far greater than in the base case.

Vessel Density

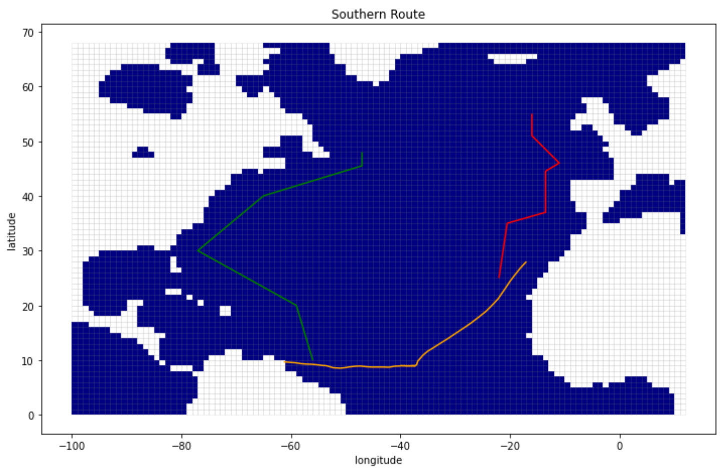
For validation of the vessel density heat map, we plotted the final locations of three Microtransat attempts which were reported as having been caught by fishing boats. Two of the three were caught off the coast of North America, which appears to coincide with the higher densities closer to shore. The third case is further out near Europe.

Experiment

As an initial study of the utility of this simulation, simulated runs were made for 2 corridors of navigation suggested by Gibbons-Neff and Miller (2011). The starting locations of these paths are Cape Spear to the north and Tenerife to the south. The timestep for these simulations is 10000 seconds.



The northern route is from west to east, and crosses both lines which define the Microtransat challenge.



The southern route, which involves travel from east to west, just misses the lines which define the Microtransat challenge.

From these results, we can see that these corridors of navigation do have a high potential for leading to successful crossings.

Summary

In summary, this project provides a simulation environment for transatlantic crossings by small autonomous vessels. In its current state, it provides results that are passable and potentially useful for the purpose of assisting in path planning. While the results in the validation step suggest that our trajectories cover far too much distance in too little time, further tuning of the vessel representation is likely to alleviate this problem to some extent. However, due to the complex nature of the ocean, input from subject matter experts would be helpful in further verification of results and future developments.

Conclusions

From this project, we can conclude that it is feasible to model trajectories of drifting vessels in the North Atlantic using computer simulation. In terms of overall trends in directions, the simulations behaved as expected. On the other hand, the distances covered in timesteps is much higher than ground truth data suggests it should be. This suggests that the data selected for use in the model is valid, and that there is tuning required in how that data affects the trajectories.

Further, we can conclude that the paths suggested by Gibbons-Neff and Miller (2011) are promising and should be investigated further.

Discussion

Limitations

Regarding accurate representation of the system, the simulation does not simulate the true variations in ocean conditions throughout various times of day and times of year. Additionally, the amount of data utilized by the simulation limits its accuracy. The model of the vessel is also simplistic and may be missing key features necessary for more accurate trajectories.

Regarding usefulness, the tendency for the vessel to get beached when near the shore limits the simulation’s utility when attempting to model launches of vessels from land.

The resolution of the grid limits the use of the simulation for path planning to regions far from land.

Future Work

There are many areas of potential improvement throughout the simulation. The most urgent improvement to the model is tuning of parameters to make the distance covered per timestep more consistent with real data. Additionally, a new metric accounting for heading in addition to absolute error would likely be useful for assessment of simulation accuracy, as we found that qualitatively, our generated paths appear to have the correct overall directional trend.

Currently, distance updates are based on approximate translations of meters to degrees of latitude and longitude. This is a rough estimate which we deemed acceptable for this proof-of-concept simulation. Going forward, this could be improved using functions for translating between reference frames. Additionally, future iterations would benefit from additional consideration in map projections to better represent distances, as the effect of the curvature of the Earth is significant at this scale.

The variance in the vectors for each environmental condition are considered as independent. The model would likely have better accuracy using multivariate statistical methods. Sampling of the environmental conditions could possibly be further improved through characterizing the relationships of grid cells with their neighbors. This could potentially incorporate techniques such as cellular automata.

The simulation has limited data on maritime vessels that could cause potential interference. More information from beyond the areas near the east coast of North America would help further validate our model. However, since there is little data from autonomous sailing vessels in general, the probability of interference will still be difficult to estimate.

The vehicle model is very simplistic, and future models would incorporate path planning to better model the decisions of a truly autonomous boat. Further, the effects of the environment on the vessel's movements would be improved with better modeling of the hull and sail designs.

Bibliography

Abrougui, H., & Nejim, S. (2019). Modeling and path following control of an autonomous catamaran sailboat. Naval Engineers Journal, 131(3), 91-105. https://www.ingentaconnect.com/content/asne/nej/2019/00000131/00000003/art00023

Deng, Z., Xie, L., Liu, B., Wu, K., Zhao, D., & Yu, T. (2009). Coupling winds to ocean surface currents over the global ocean. Ocean Modelling, 29(4), 261-268. https://doi.org/10.1016/j.ocemod.2009.05.003

Dohan, Kathleen. (2022). *Ocean Surface Current Analyses Real-time (OSCAR) Surface Currents - Interim 0.25 Degree* (Version 2.0). [Dataset]. ESR. https://doi.org/10.5067/OSCAR-25I20

Du, M., Kang, M., Hou, C., & Xu, J. (2018). Study of Long-term Route Planning for Autonomous Sailboat. In *Proceedings of the International Robotic Sailing Conference*.

Echevarria, E. R., Hemer, M. A., & Holbrook, N. J. (2021). Global implications of surface current modulation of the wind-wave field. *Ocean Modelling*, *161*, 101792. https://doi.org/10.1016/j.ocemod.2021.101792

Eckhardt, B. J. (2012, April 3). *Visibility Factors in Small Boat Collisions - Expert Research*. Robson Forensic. Retrieved April 30, 2022, from https://www.robsonforensic.com/articles/visibility-factors-in-small-boat-collisions-expert-research

Gibbons-Neff, P., & Miller, P. (2011). Route planning for a micro-transat voyage. In *Robotic Sailing* (pp. 183-194). Springer, Berlin, Heidelberg.

Melin, J., Dahl, K., & Waller, M. (2015, August). Modeling and control for an autonomous sailboat: a case study. In World Robotic Sailing championship and International Robotic Sailing Conference (pp. 137-149). Springer, Cham. https://link.springer.com/chapter/10.1007/978-3-319-23335-2_11

NOAA. (2020). *NOAA/NCDC Blended Monthly Global 0.25° Sea Surface Wind Stress, 1987-2011, Lon+/-180*. [Dataset]. NOAA. https://data.noaa.gov/dataset/dataset/noaa-ncdc-blended-monthly-global-0-25-sea-surface-wind-stress-1987-2011-lon-1801

NOAA Fisheries. (n.d.). *2018 Drifter Tracks*. NOAA Fisheries: NEFSC 2018 Drifter Tracks. Retrieved April 30, 2022, from https://apps-nefsc.fisheries.noaa.gov/drifter/index_2018.html

NOAA Office for Coastal Management. (2022, March 24). *AIS data for 2021*. AIS Data for 2021. Retrieved April 30, 2022, from https://coast.noaa.gov/htdata/CMSP/AISDataHandler/2021/index.html

NOAA's Atlantic Oceanographic and Meteorological Laboratory. (n.d.). *Drifter Frequently Asked Questions (FAQs)*. PhOD - Global Drifter program. Retrieved April 30, 2022, from https://www.aoml.noaa.gov/phod/gdp/faq.php#sst

Qi, L., Zheng, Z., & Gang, L. (2017). Marine traffic model based on cellular automaton: Considering the change of the ship's velocity under the influence of the weather and sea. *Physica A: Statistical Mechanics and its Applications*, *483*, 480-494.

The Microtransat Challenge. (n.d.). *The Microtransat Challenge*. Retrieved March 2, 2022, from https://www.microtransat.org/

United States Naval Academy. (2019, May 8). *Approximate Metric Equivalents for Degrees, Minutes, and Seconds*. Approximate metric equivalents for degrees. Retrieved April 30, 2022, from https://www.usna.edu/Users/oceano/pguth/md_help/html/approx_equivalents.htm

VLIZ. (n.d.). *Marine Gazetteer Placedetails*. Marine regions · North Atlantic Ocean (IHO Sea area). Retrieved April 30, 2022, from https://marineregions.org/gazetteer.php?p=details&id=1912

Wang, Y., Yao, P., & Dou, Y. (2019). Monitoring trajectory optimization for unmanned surface vessel in sailboat race. *Optik*, *176*, 394-400.

Appendix: Division of Labor

James Theriault, jtheriault7@gatech.edu	- Background research - Loading Maritime Traffic Datasets - Maritime Traffic Datasets visualization development
Sandaru Fernando, sfernando8@gatech.edu	- Background research - Final paper writing - Loading Surface Current and Wind Datasets - Regridding function development - Noise function development
Sean Fish, sfish7@gatech.edu	- Background research - Vehicle dynamics and path visualization development - Final paper writing and presentation - Simulation function development - Analysis and experiment development