# Assignment 3

Maddy Prazeus, 31494978

## Section 1: Report Table

| **Base Model** | **Travelling Sales Person, used to compute the shortest path between n points** |
|---|---|
| Extension Assumptions | Consider Travelling sales person as a ship navigating islands. Introduce a known current which will either speed up / not affect / slow down the ship, depending on the angle between them. Now we seek to minimise total travel time of the ship between the islands. Finally, introduce an unknown stochastic current. |
| Techniques showcased | Simulated Annleaing & Monte Carlo Simulation. |
| Modelling question 1 | How does a constant current impact our ship's ability to navigate the ocean? |
| Modelling question 2 | How does an unknown stochastic current impact our ship's ability to navigate the ocean? |

## Section 2: Introduction - 10%

Travelling Sales Person (TSP) is used to find the shortest path between a number of points. However, I believe in many situations the optimal path is not nessacarily the one with the shortest distance. There are many other external factors that can influence the choice of an optimal path. Such as, travel time, cost, fuel consumption, safety, enjoyment and many others.

With this in mind, I hope to explore TSP beyond the context of a shorter parth to the more general "optimal path". Where the optimal path is determined by total travel time. To facilitate this exploration we consider the following hypothetical:

Our travelling sales person is now a ship who intends to make a stop at a number of islands. The ship would like to find a path between all the islands that minimises its total travel time. To distinguish our example from the base model we now introduce a known constant current. Depending on the angle between the ship and the current, its passage between two islands will either be sped up, slowed down or unaffected. The final extension, will be considering route planning for our ship with an unknown stochastic current. We will use Monte Carlo to simulate a number of possible current magnitudes and directions, and subsequently determine the optimal route to take.

## Key Questions:

1. How does a constant current impact our ship's ability to navigate the ocean?
   1.1 Is there a connection between angle and magnitude of the current and solution path shape?
   1.2 Does the optimal route differ from the shortest?
   1.3 What happens to the total travel time of our ship? Increase / decrease / variable?

2. How does an unknown stochastic current impact our ships ability to navigate the ocean?
   2.1 Is Monte Carlo simulation useful for finding a "good path" (relative to the true optimal on the given day)?
   2.2 Are there any trends in the optimal paths across the simulations?
   2.3 Are there any limitations to Monte Carlo simulation that impact our ability to find a close to optimal path?

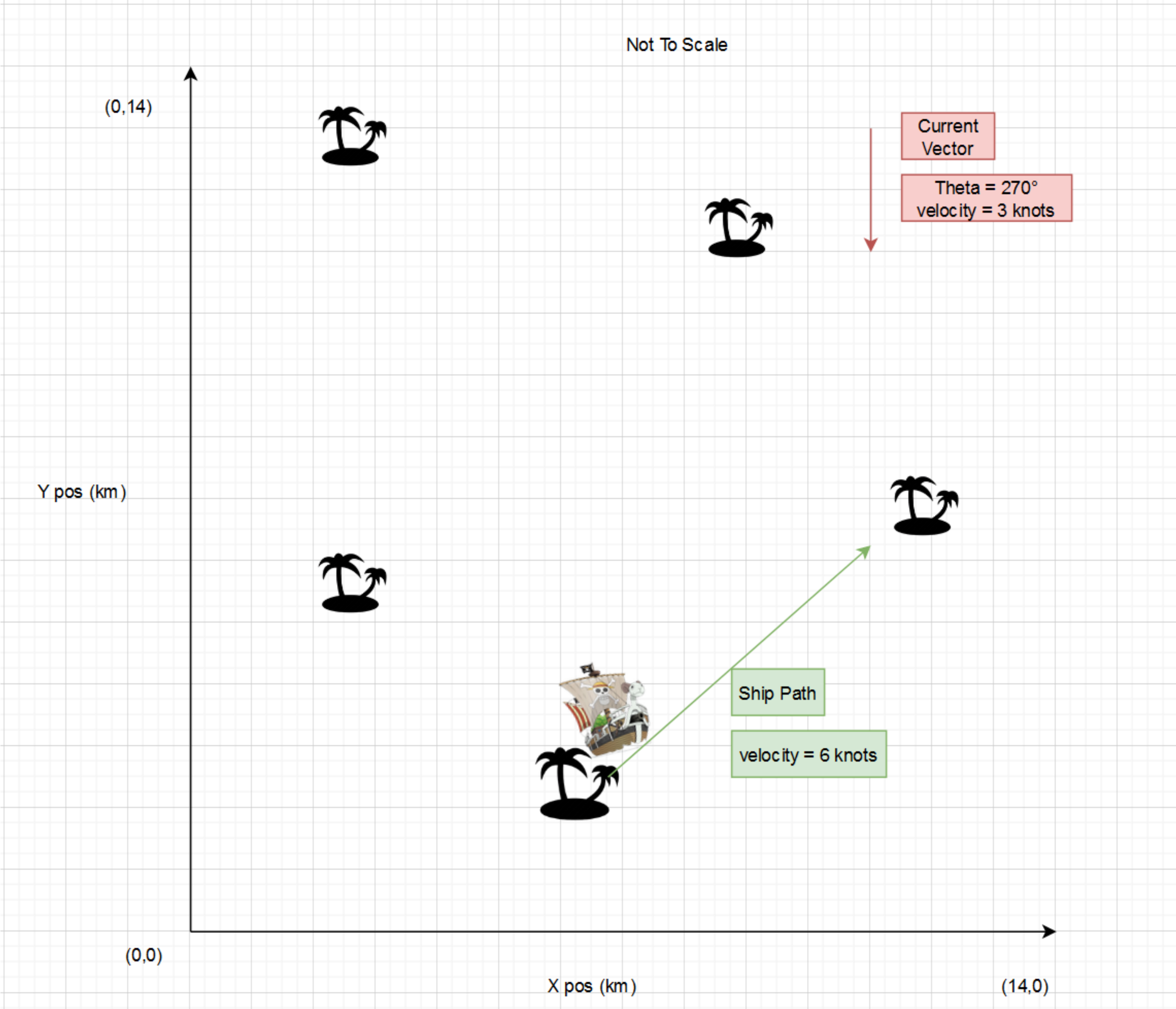
## Technique Justification:

- Simulated Annealing
  - TSP is a NP hard problem. Therefore, to make our model more scalable for a large number of islands we introduce simulated annealing.
  - Simulated annealing is not dependent on the input size of our graph, it scales with the cooling factor and perterbutions / iteration. This allows us to escape from the poor superpolynomial time complexity of TSP.
  - Alternatives such as hill climbing, are not appropriate as they will look past global optimisations in the pursuit of local minimisations.

- Monte Carlo Simulation
  - The current for any given day is not dependent on any other day before or after it. Therefore, it is not possible for us to train a model on previous data or use prediction tools such as Markov chains to plan our route.
  - Given the previous point, we must then use some form of simulation. Since no alternatives are taught in this unit, Monte Carlo is the go to option.

## Section 3: Model Description - 35%

### Model Visualisation




(Figure 1)

### Original Model Assumptions:

- Instead of a Travelling sales person we now consider a ship that must Travel between N islands.
- All islands must be visited once.
- Islands cannot be revisited.
- Each island exists at some point $(x_i, y_i)$.
- The ship starts at an island $(x_0, y_0)$.

### Extended Model Assumptions:

- The ship now moves at a constant velocity v.
- There exists a vector S, which describes the velocity and direction of the ship moving between two islands.
- For any day there exists a constant current C, which is comprised of a magnitude M and a direction $\theta$.
- Given a path for our ship between any two islands. Our ship will either be sped up, slowed down or unaffected by the current C (see derivations for examples).
- We now seek to optimise the total time taken for our ship to complete the voyage between all islands.

See figure 1 for a visualisation of this model.

## Model class and type of analysis:

- Determine the class of model and analysis

For question one our model would fall under Numerical versus Analytical. In question two we extend this to Deterministic vs Stochastic.

## Referenced Algorithms, Derivations and Mathematical Results:

### Current Impact on Ship

Let S be the velocity vector of a ship and C be the velocity vector of the current. We can define the impact of C on S as it travels between islands as the vector projection P of C onto S.

$$P = |C|cos(\theta)\hat{S}$$

Where $\theta$ is the angle between S and C if they are connected tail to tail.

Given $cos(\theta) = \frac{a.b}{|a||b|}$ for two tail connected vectors a and b. We can simplify the above:

$$P = |C|\frac{S.C}{|S||C|} \times \hat{S}$$

$$P = \frac{S.C}{|S|} \times \hat{S}$$

If $\theta < \pi/2$ then we will add |P| to our ship's velocity. Otherwise if $\theta > \pi/2$ then we will subtract |P| from our ship's velocity.

Note that if $\theta = \pi/2$, i.e. they are perpendicular. Then $S.C = 0$ and $P = 0$. Hence, there will be no change in the ship's velocity.

## Section 4: Results - 35%

### Dependencies

Before running any of the code in `analysis.py` or `main.py`, please ensure the following dependences are installed: `psutil, multiprocessing, tqdm, numpy and matplotlib`

### Question 1

Let tsp_instance = `[(47, 45), (16, 51), (48, 39), (7, 41), (67, 8), (27, 76), (34, 6), (25, 18), (66, 77), (15, 99), (61, 27), (49, 72), (84, 20), (87, 62), (98, 59)]`

And ship_speed = 60

Given the above inputs I generated a number of solution tours for a range of current directions (See Section 1 in `analysis.py`):

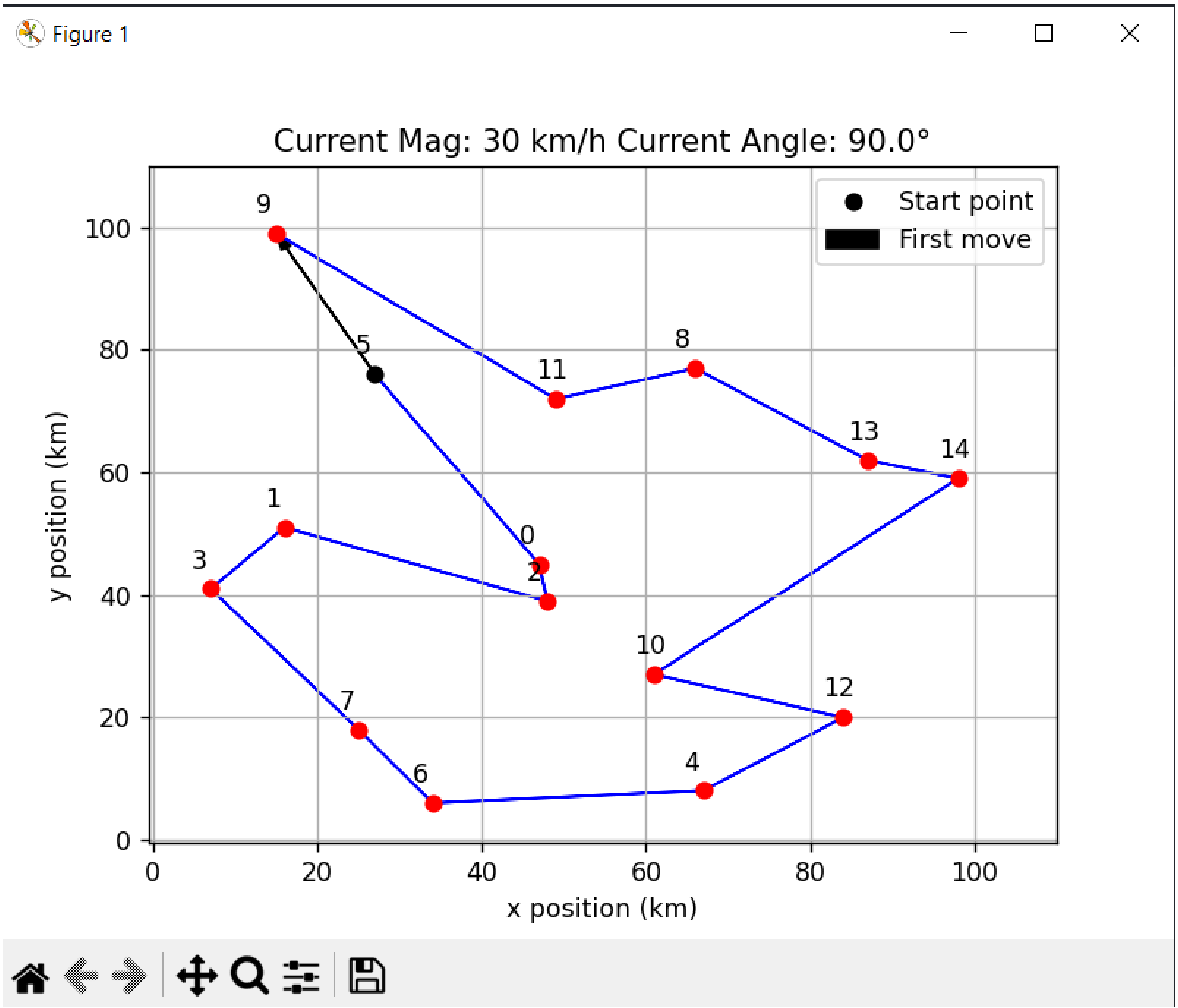


Figure 1
Current Mag: 30 km/h Current Angle: 90.0°
Start point
First move
y position (km)
x position (km)
0
20
40
60
80
100
9
5
11
8
13
14
1
0
2
3
10
12
7
6
4
(Figure 2)

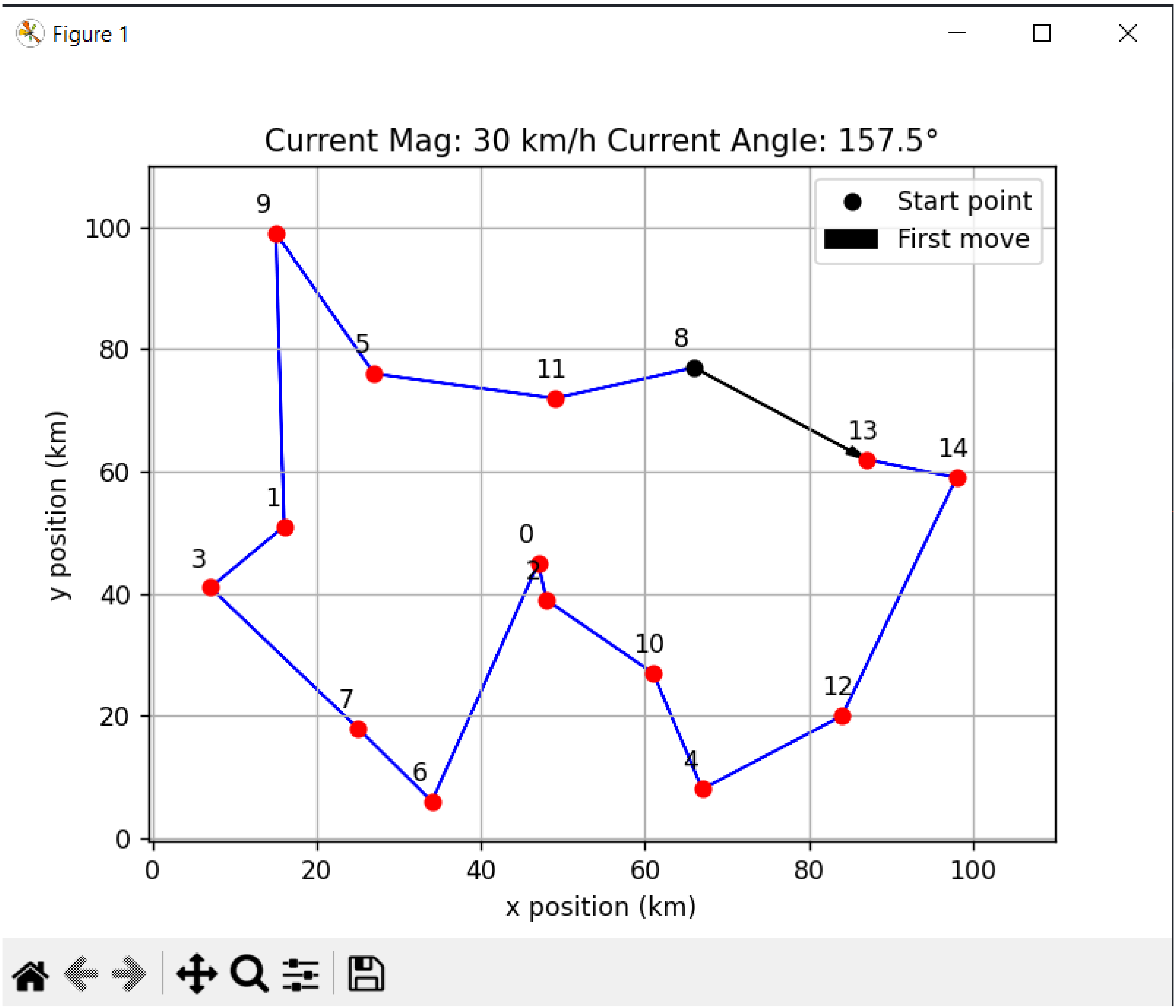


Figure 1
Current Mag: 30 km/h Current Angle: 157.5°
Start point
First move
y position (km)
x position (km)
0
20
40
60
80
100
9
5
11
8
13
14
1
3
0
2
10
12
7
6
4

(Figure 3)

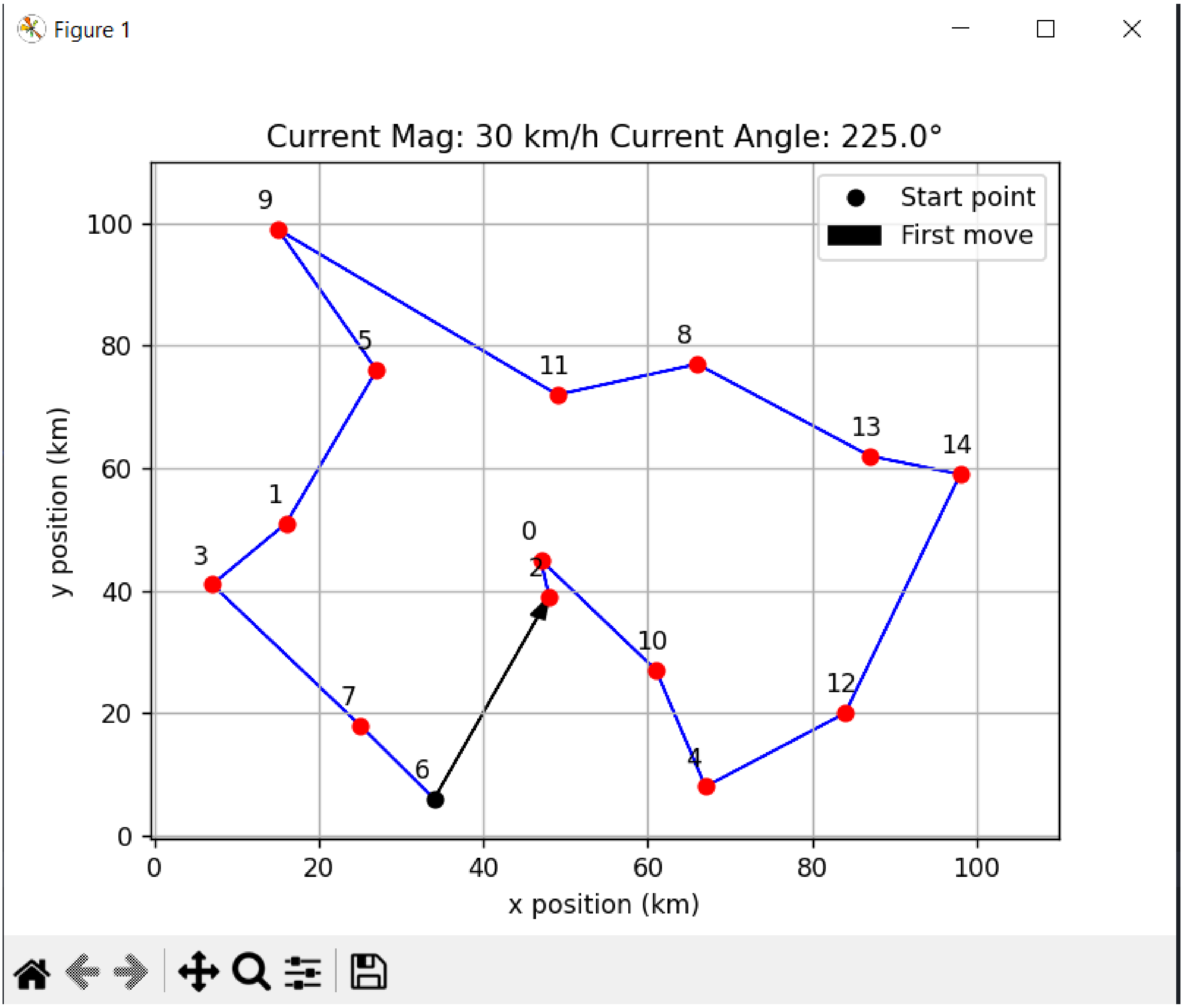
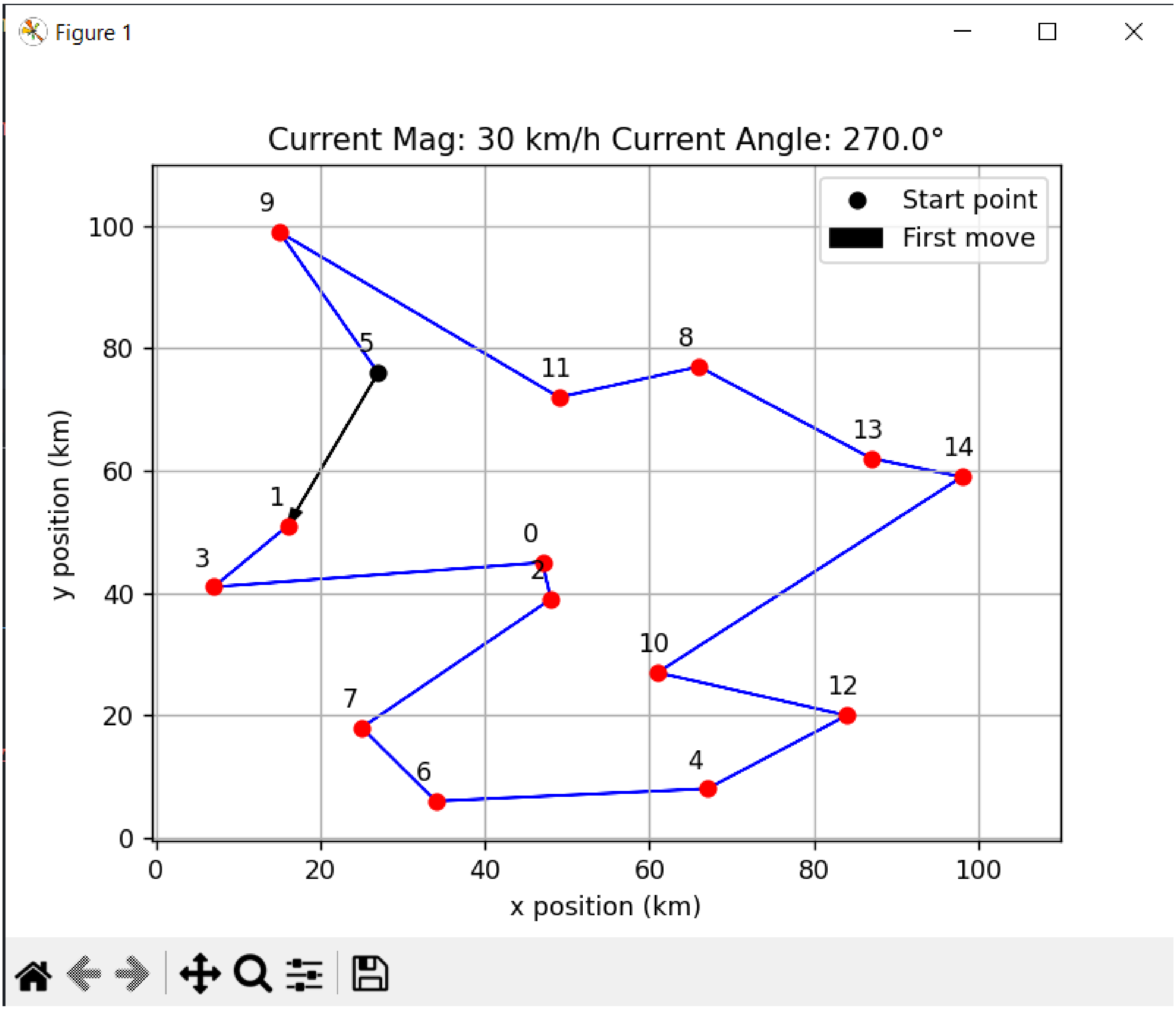


Figure 1
Current Mag: 30 km/h Current Angle: 225.0°
Start point
First move
y position (km)
x position (km)
0
20
40
60
80
100
9
5
8
11
13
14
1
0
3
2
10
7
12
6
4
(Figure 4)



(Figure 5)

Across figures 2-5 there appears to be a relationship between current angles close to the vertical poles $\pi/2$ and $3\pi/2$ and the horizontal poles 0 and $\pi$. More specifically for angles close to the vertical pole, our solution tours seems to favour vertical movement. Similarily, for angles closer to the horizontal pole, we see a solution path that favours horizontal movement. Thus, answering question 1.1, that there is a positive relationship between angle direction and favoured direction of solution path.

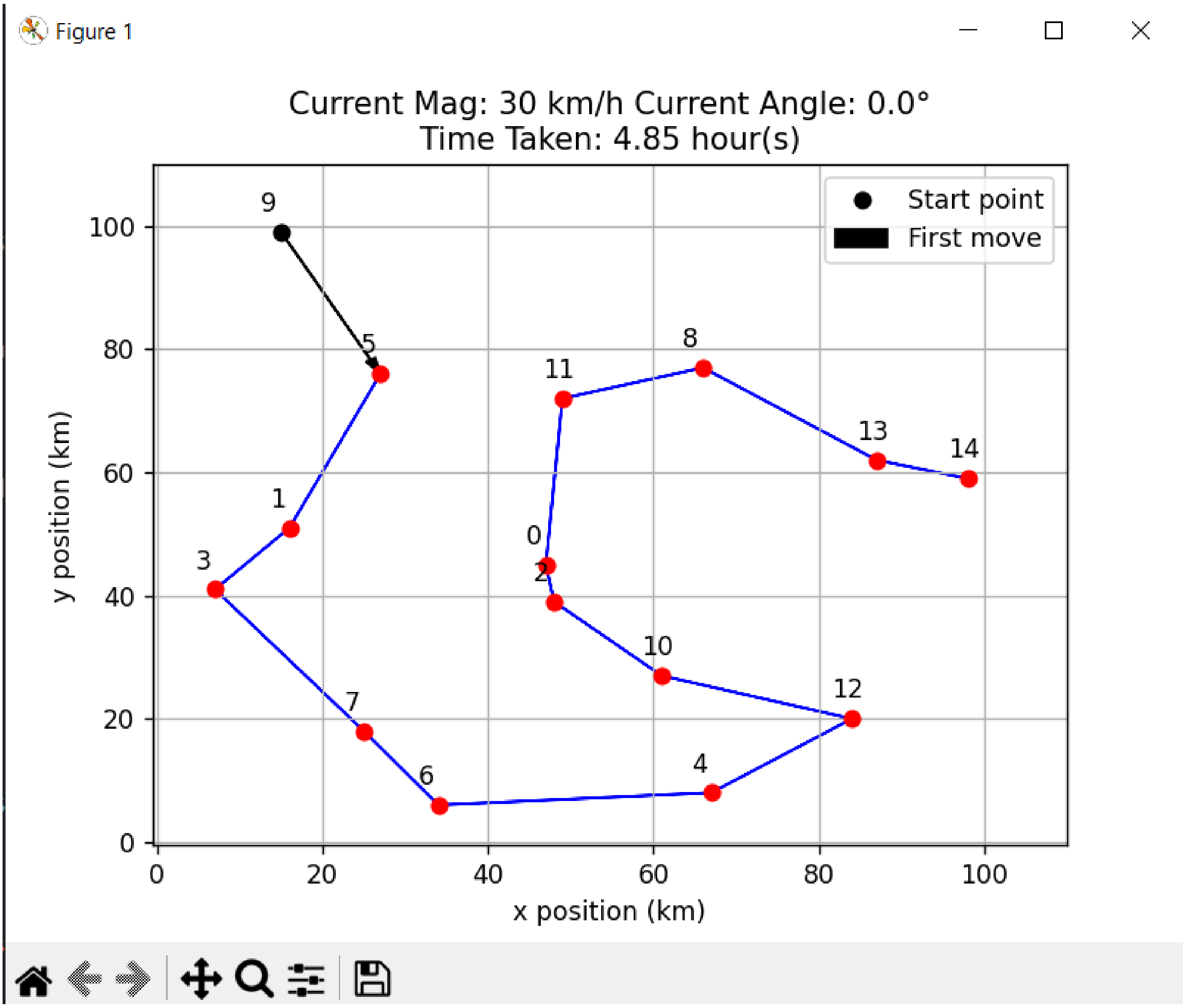
My next step in analysis was to compare tours that are impact by current with ones are not, i.e. optimal path vs shortest path. However, when doing this I found that I was getting very similar or identical tours to the ones displayed in figures 2-5. Inferring that there is no impact from introducing current, and identifying a fundamental flaw in my model. After further inspection, I realised that since the sum of all the turns that the boat makes will always result in 360° since the tour forms a closed polygon.

Given this flaw, I propose the following modification to the extended model:

- The ship is not required to return to its starting position after visiting the final island.

This extension will be used henceforth in the analysis section.

Using the same initial inputs as figures 2-5 and varyign the angle of the current I created the following figures. (See Section 2 in `analysis.py`)




(Figure 6)

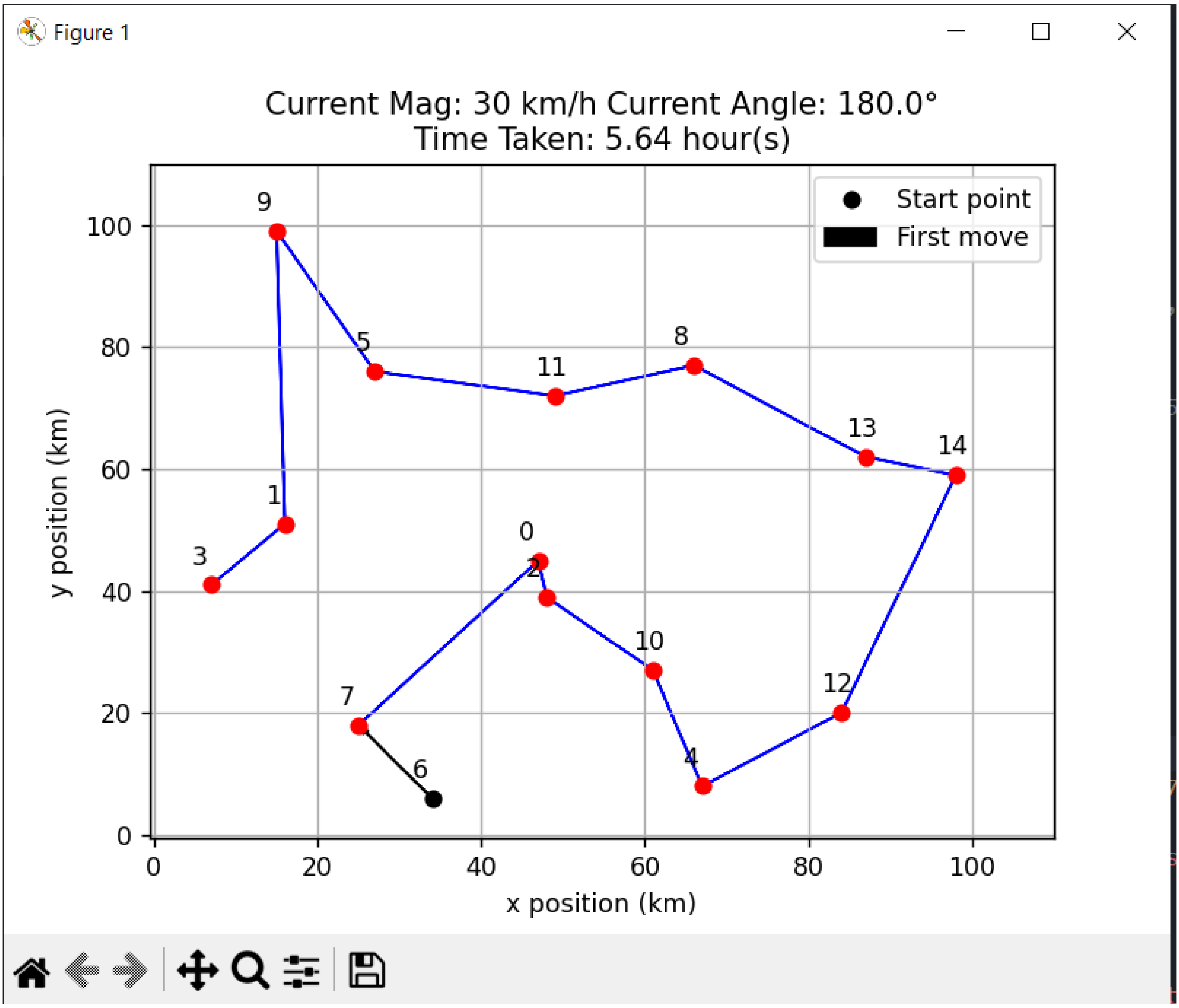


Figure 1
Current Mag: 30 km/h Current Angle: 180.0°
Time Taken: 5.64 hour(s)
Start point
First move
y position (km)
x position (km)
0
20
40
60
80
100
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
(Figure 7)

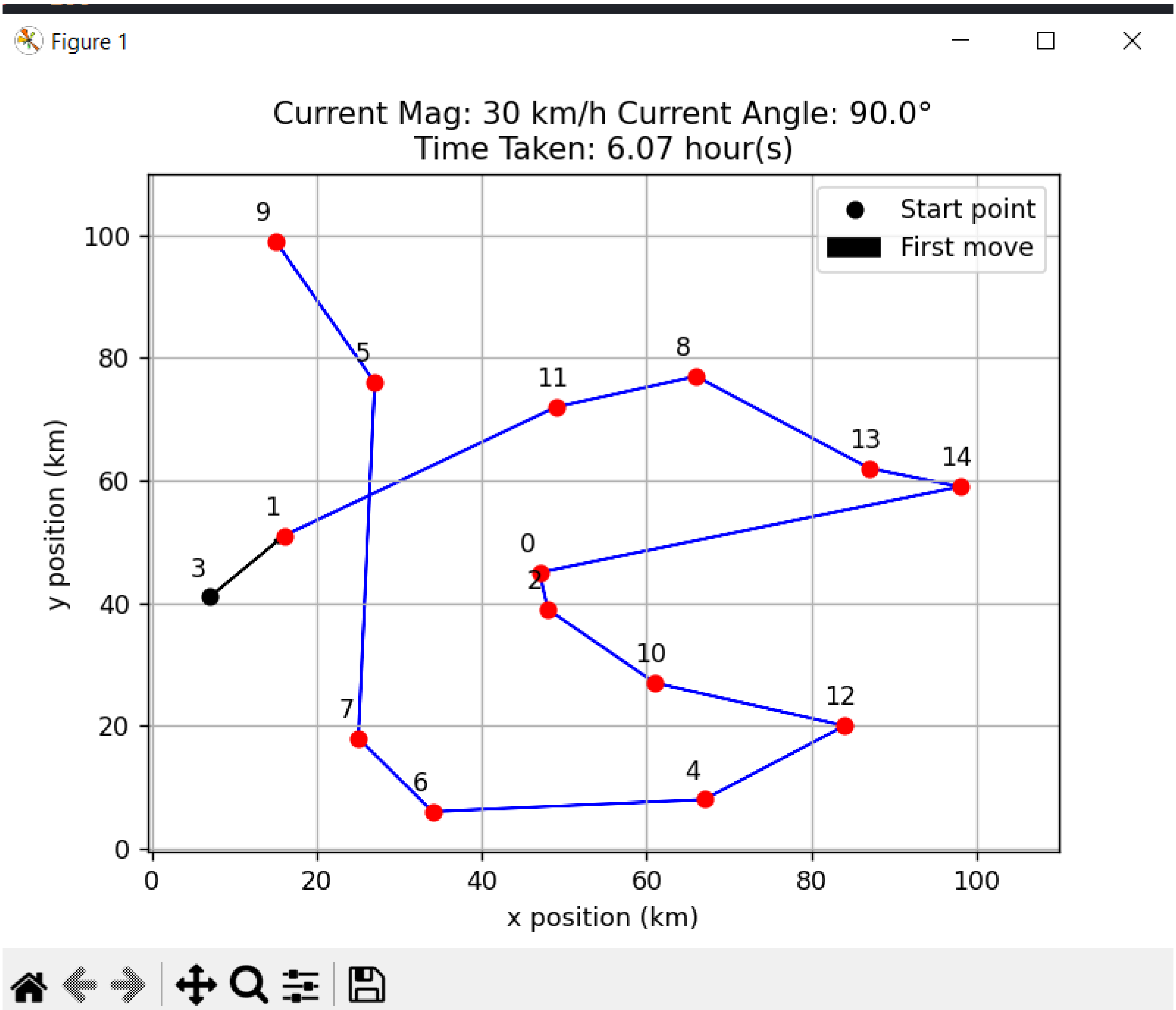
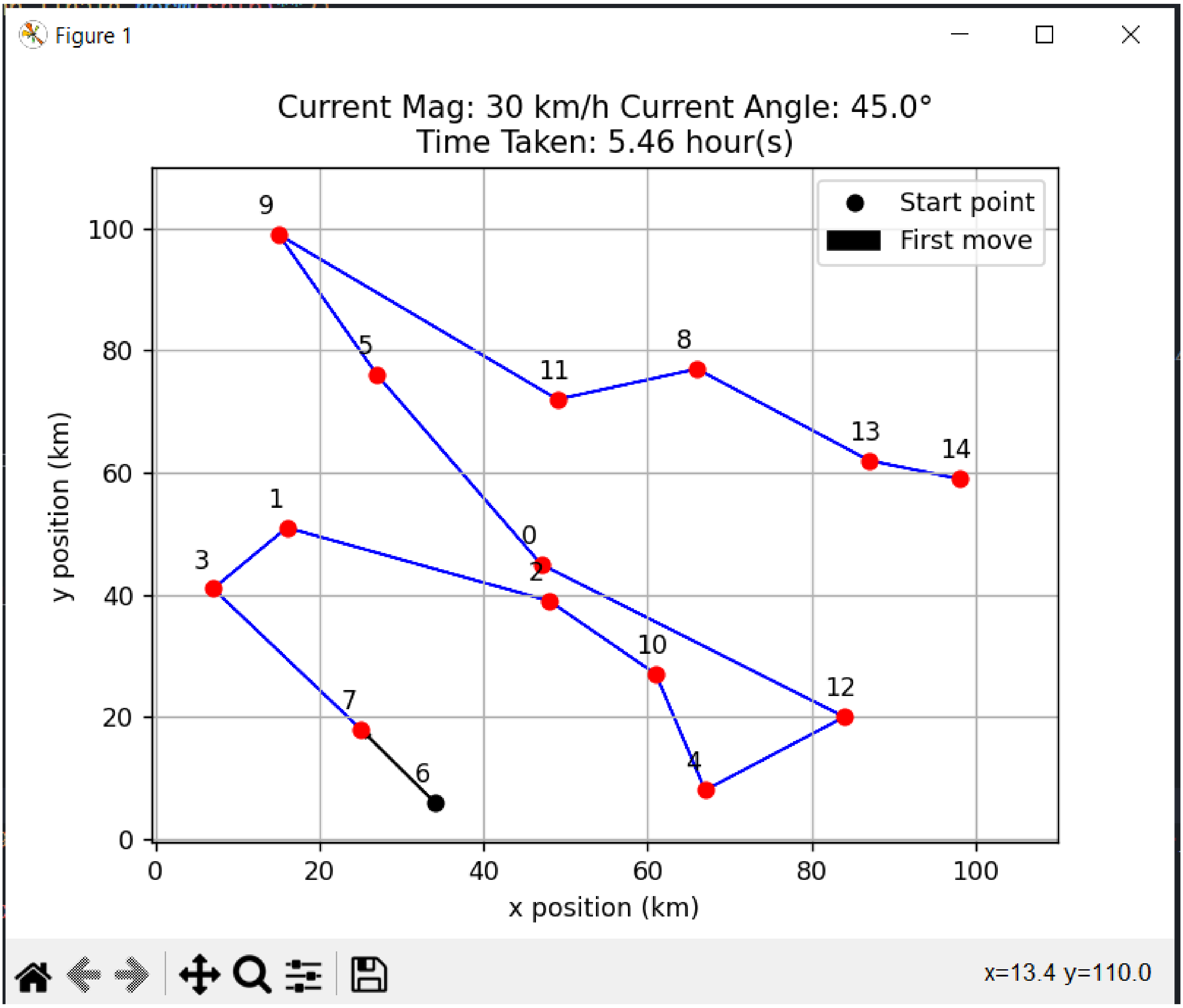


Figure 1
Current Mag: 30 km/h Current Angle: 90.0°
Time Taken: 6.07 hour(s)
Start point
First move
y position (km)
x position (km)
0
20
40
60
80
100
9
5
11
8
13
14
1
3
0
2
10
12
7
6
4

(Figure 8)




(Figure 9)

Depending on the configuration of the islands the TSP algorithm will now favour the following two strategies: Move approximately with the current as much as possible, and if this is not possible attempt to move close to perpendicular with the current (as atleast it then wont have a negative impact). Figures 6-8 clearly utilise a mix of the two above strategies, moving with the current where possible and moving close to perpendicular when not. In figure 9 we see that since there are minimal opportunities to move with the current (45°). It instead almost entirely moves perpendicular to the current through all islands. Which once again answers question 1.1, that there is a relationship between current angle and ship tour (except now we have valid data).

As seen in Figure 9, the aforementioned strategies can produces optimal tours that do not minimise total distance travelled. To then answer question 1.2, focusing on min/maxing interactions with the current versus focusing on minimising distance evidently produces wildly different output tours.

After running the TSP simulated annealing solver with no current and the same island configuration 10 times I got a sample mean travel time of 5.454 hrs and a sample standard deviation 0.159. The small standard deviation reveals that on average a tour without a current will take less time than a tour with a current. However, there are still edge cases as seen in Figure 6. Where for specific current directions in relation to island configurations there can be a much faster result. So to answer question 1.3 on average introducing a current increases the time taken for our tour. However, specific combinations of inputs can result in outliers that reduce it.

### Question 2

We now consider the situation where we have an unknown stochastic current. Our goal is to generate a number of possible scenarios (currents) and determine from this an optimal path for our ship to take the following day.

To do this we will use Monte Carlo to simulate a number of possible currents (unique directions and magnitude), and the corresponding solution paths for each of them. We then loop over all the solutions and find which performs on average better with the generated currents. This solution is presented as the estimated optimal path for a given island configuration and will be used by the ship on the following day

Utilising the same island configuration as in the above sections I created the following figures (see Section 4 of `analysis.py`, note this section will take a little while to run as I increased the cooling ratio from 0.9 -> 0.99 for the true optimal path, see line 97)

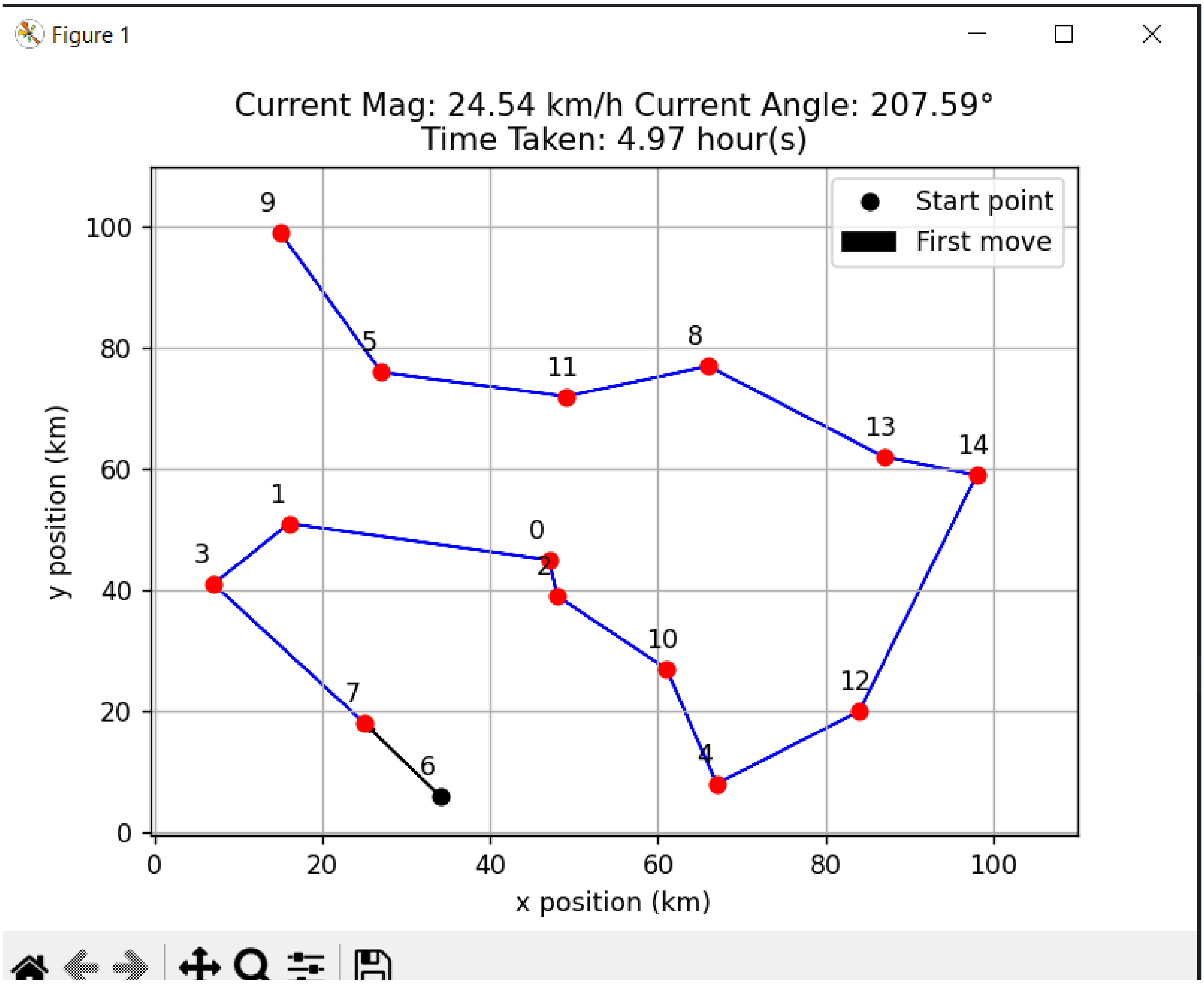


Figure 1
Current Mag: 24.54 km/h Current Angle: 207.59°
Time Taken: 4.97 hour(s)
Start point
First move
y position (km)
x position (km)
0
20
40
60
80
100
9
5
11
8
13
14
1
0
3
2
10
12
7
4
6

(Figure 10)

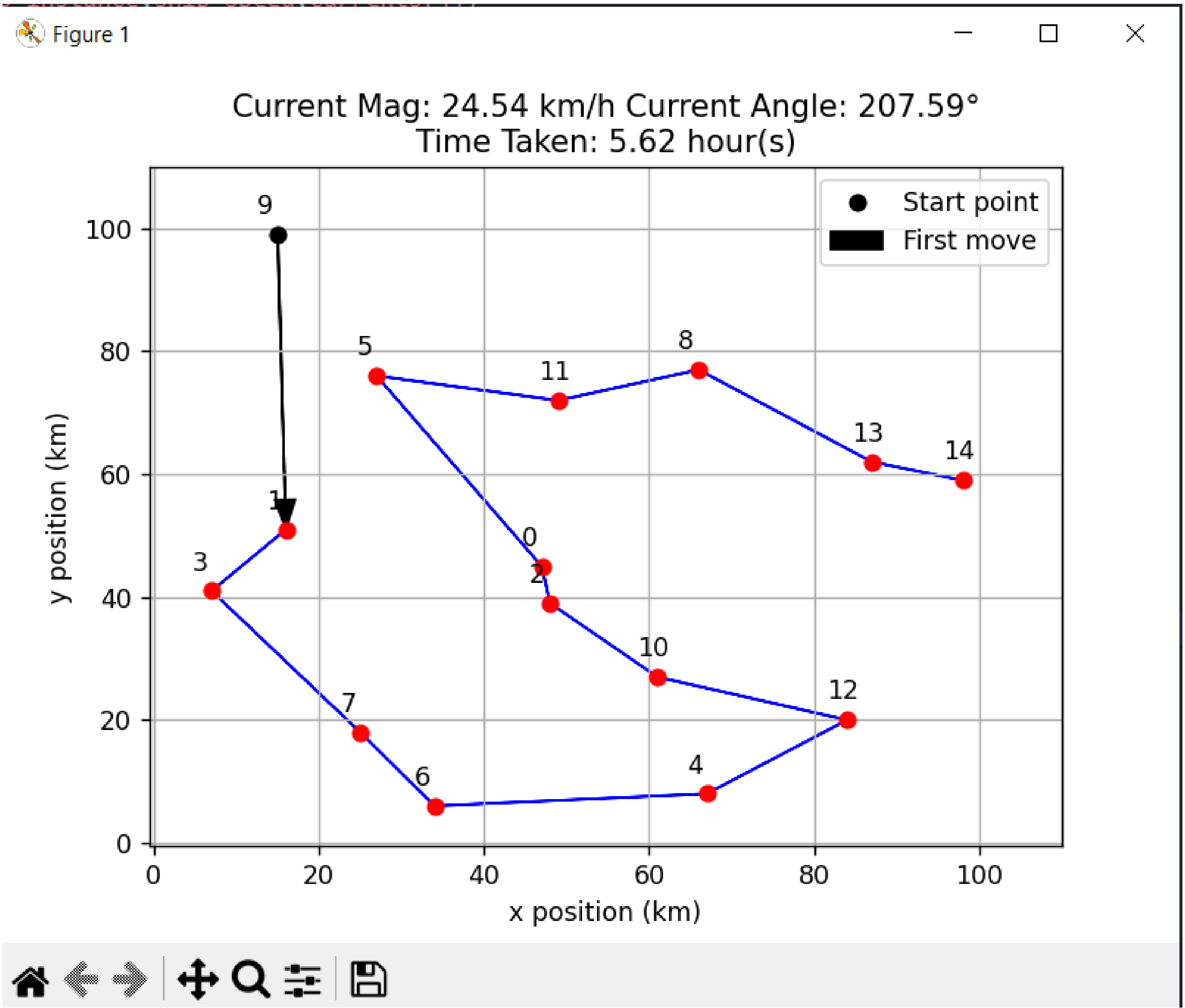


Figure 1
Current Mag: 24.54 km/h Current Angle: 207.59°
Time Taken: 5.62 hour(s)
Start point
First move
y position (km)
x position (km)
0
20
40
60
80
100
9
5
11
8
13
14
1
3
0
2
10
12
7
6
4

(Figure 11)

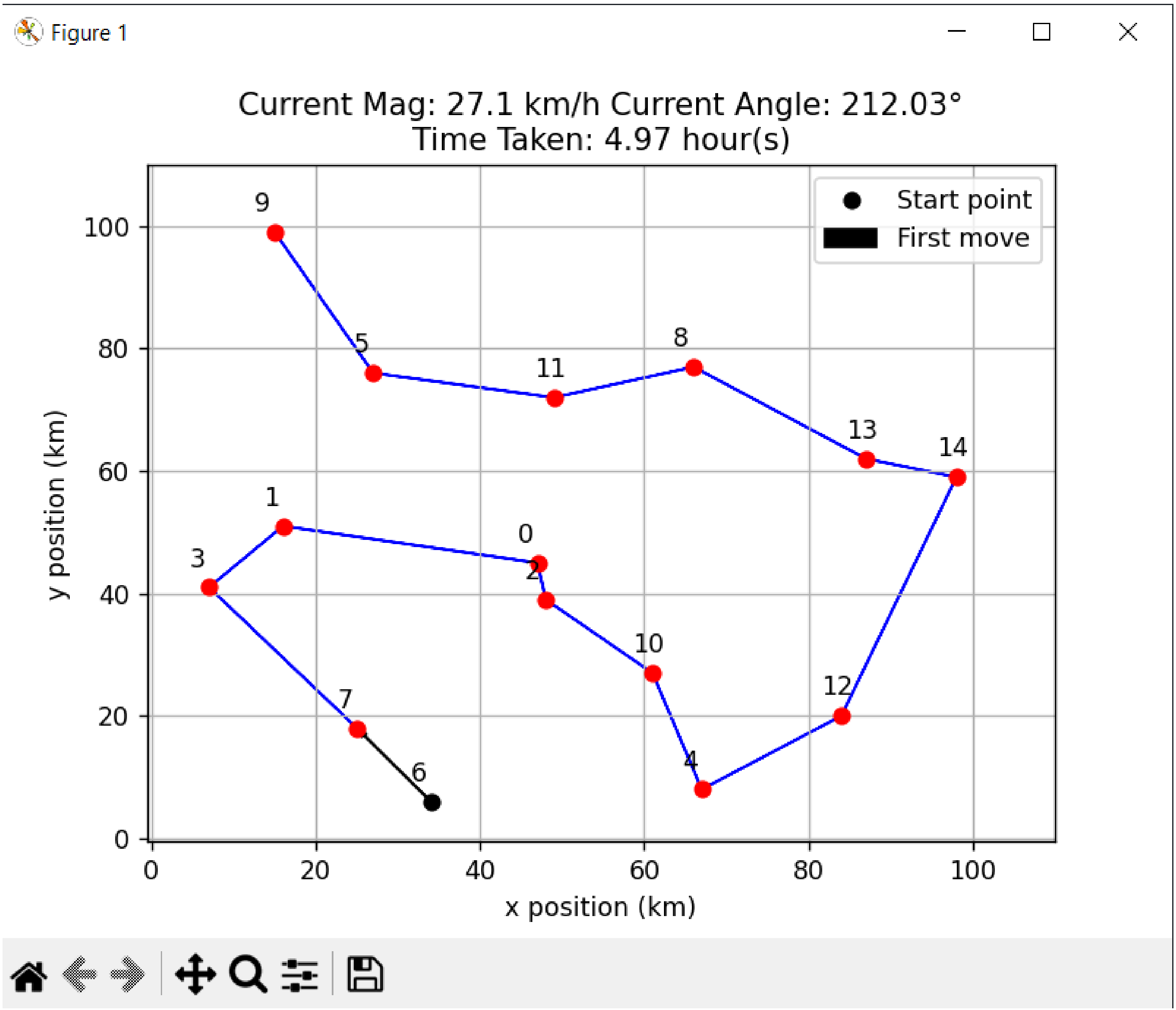
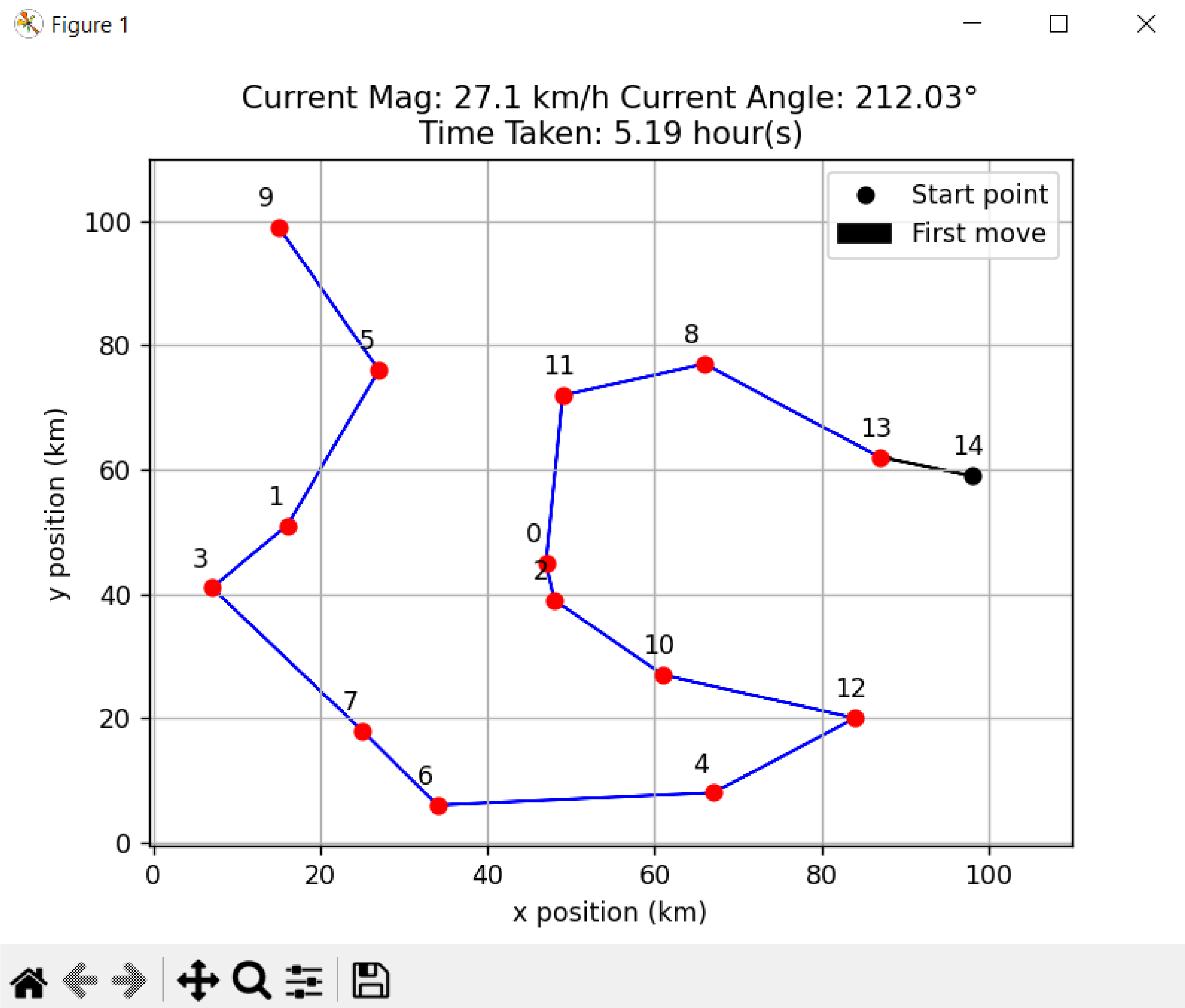


Figure 1
Current Mag: 27.1 km/h Current Angle: 212.03°
Time Taken: 4.97 hour(s)
Start point
First move
y position (km)
x position (km)
0
20
40
60
80
100
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
(Figure 12)




(Figure 13)

Figures 10 and 12 showcase our estimated optimal path found by montecarlo. Figures 11 and 13, identify the true optimal path after sampling the true current. After running this process a number of times I observed that the % difference between the true and total travel time was never more than 10%. In some cases Monte Carlo found a solution which was better! This can be attributed to the performance limitations of Simulated annealing in comparison to the multi-processor utilisation of MonteCarlo. MonteCarlo can cover a larger combination space in a similar amount of time as Simulated Annealing.

I would argue that a % error difference of < 10% definitely meets Quesiton 2.1's criteria of "good". As such, I would recommend MonteCarlo as a useful technique in this scenario.

With respect to question 2.2 there is no trend we can find between generated solution paths. As prevdu previsously identified when answering question 1.1, angle and shape of a tour are entirely dependent on the angle and magnitude of the current. Since each solution utilises a randomly generated current, there is not recogniseable trend between them.

In reflection to the code I have written and results I have been able to obtain, MonteCarlo simulation is ultimately bound by our Simulated Annealing TSP. More specifically, we must limit the perturbations / iteration or cooling rate for it to be able to run in a reasnoble amount of time. This will likely introduce innaccuracy issues for configurations with large numbers of islands, as we will unable to effectively explore for each generated solution. As such to answer question 2.3 the key limitation is island number. For smaller amounts this solution approach will run fine. However, for larger it will be unable to generate good potential solutions for each given current.

### Summary

- **Question 1:**
  - **A constant current will on average increase the travel time of the ship.**
  - **A constant current will introduce new tour paths which differ significantly from the ones generated using the base model.**
  - **Said new tour paths are directly related to current angle, key preferences of the algorithm are to maximise paths that travel with the current and for all others try to move perpendicular .**
- **Question 2:**
  - **Monte Carlo is effective at generating an on average best tour to take for an unknown current for a given island configuration.**
  - **There are no trends between potential solutions, as each solution is dependent on its randomly generated current.**
  - **There is a key limitation to this method which is the underlying Simulated annealing implementation. For larger island amounts this will likely result in point 1 no longer being true. I.e. it will degrade to become a poor tool for generating an on average good path.**

## Section 5: List of Algorithms and Concepts - 5%

- **Travelling Sales Person used for base model.**
- **Simulated annealing used to find a solution for a given TSP arrangement and super polynomial time complexity.**
- **MonteCarlo simulation used to generate many possible potential solution tours for a unknown stochastic current.**
- **Vector Projections are used to determine the impact of a current on the ships passage.**
- **Multiprocessing is used to speedup our MonteCarlo simulations by fully utilising the systems available logical cores.**