Gaddiel Morales
CS 527

Analyzing Parameters of a Particle Swarm Optimization

Introduction

This lab tests various hyper parameters of a particle swarm optimization (PSO) and how they affect the simulation's ability to converge on a solution. One parameter was chosen as an independent variable and the rest were kept at default values. The parameters, default values and test range are as follows: number of particles, 40, 10 to 100 in increments of 10; inertia, 0.5, 0.1 to 1 in increments of 0.1; cognition, 1, 0.1 to 4 in increments of 0.1; social, 1, 0.1 to 4 in increments of 0.1.

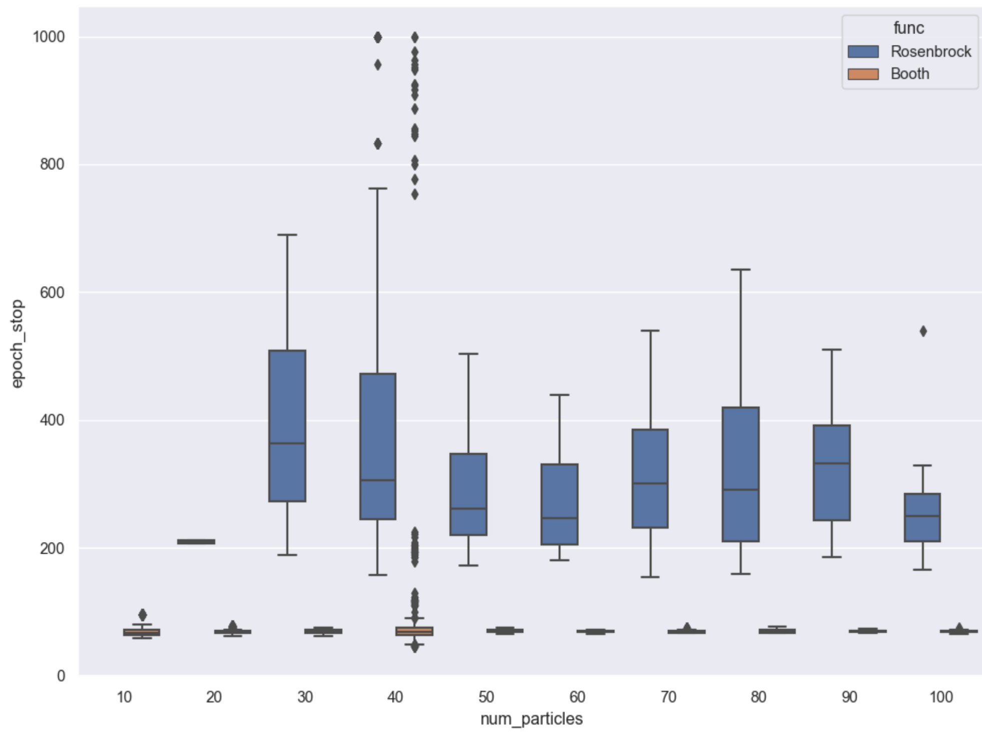
Two sets of simulations were run, one used the Rosenbrock function and the other used the Booth function. Additionally, I simulated 20 tests of each parameter combination since the particle positions are randomly chosen. This allows variation in case one simulation gets stuck in a local optima.

Results

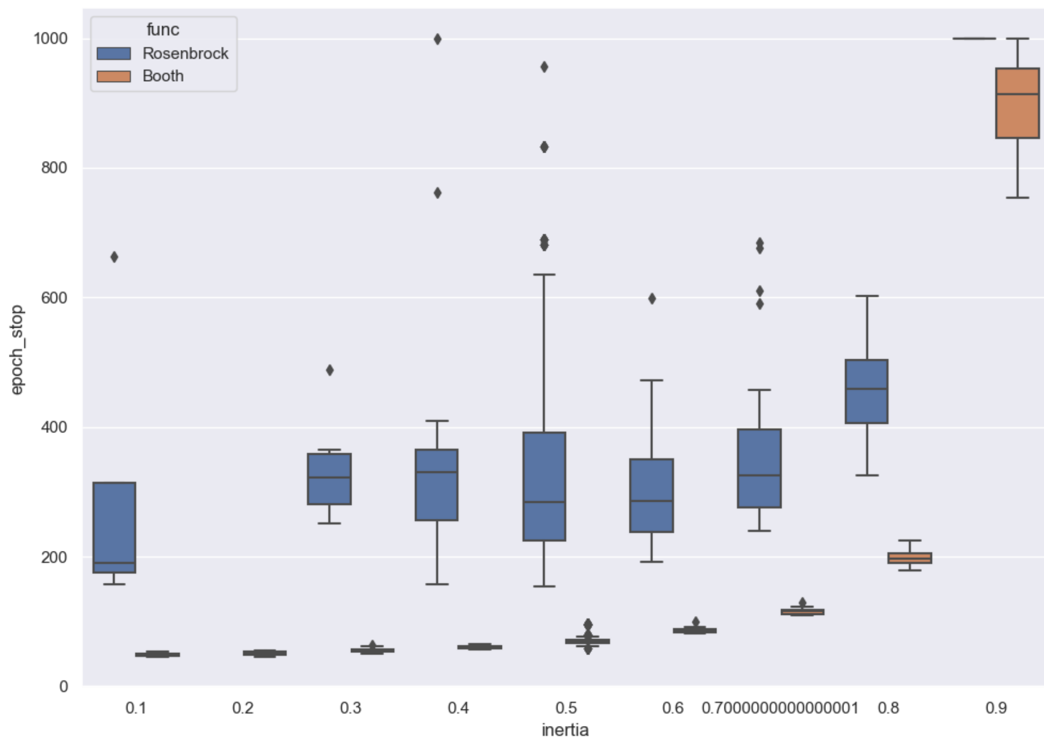
The results of the lab are graphed below. Note that extra simulations were run in order to better study the relationship between social and cognitive parameters. Every combination of values for these were tested. As a result there exists more data, and thus more outliers, for the default values of the parameters of "number of particles" and "inertia". The graphs for the social and cognitive parameters are also much more populated.

Converging Simulations

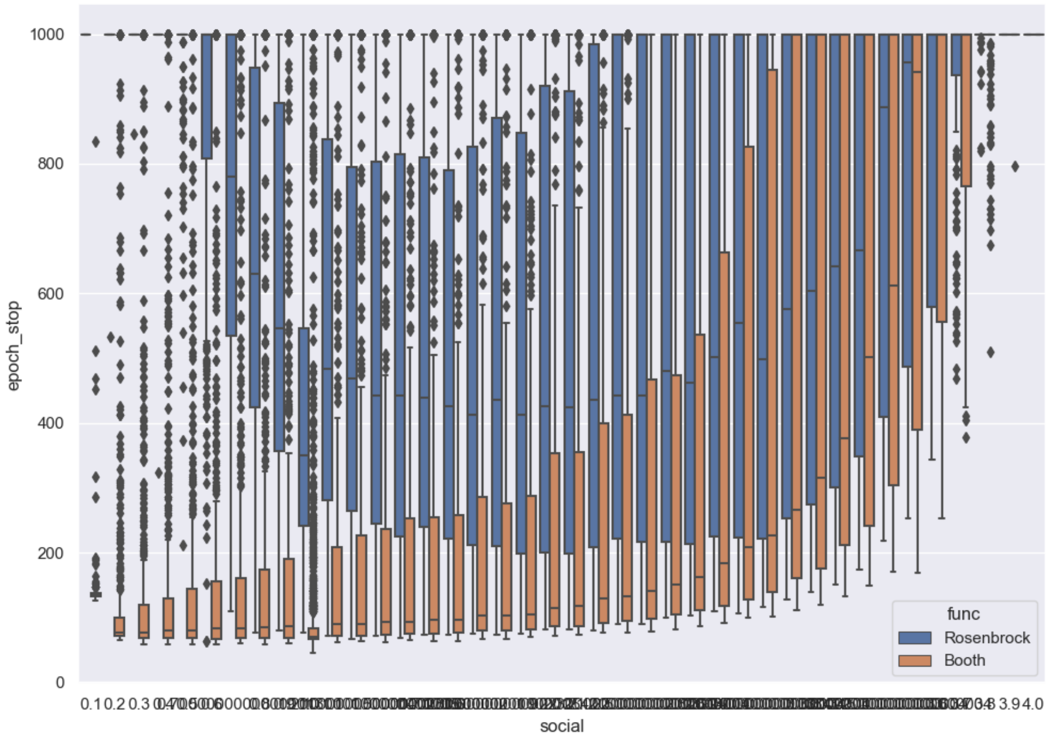
The following graph shows epochs to converge for different numbers of particles. For Rosenbrock, the values closer to 60 seem to increase the variance in epochs needed to converge. Meanwhile, the Booth function did not vary much. There was some extra variance for the booth at around 40.



The next graph shows how inertia affects the PSO. For Rosenbrock, this parameter did not have huge impact, but it did slightly improve effectiveness at around 0.3. For Booth there seems to be some kind of exponential trend with high values causing more difficulty in convergence.

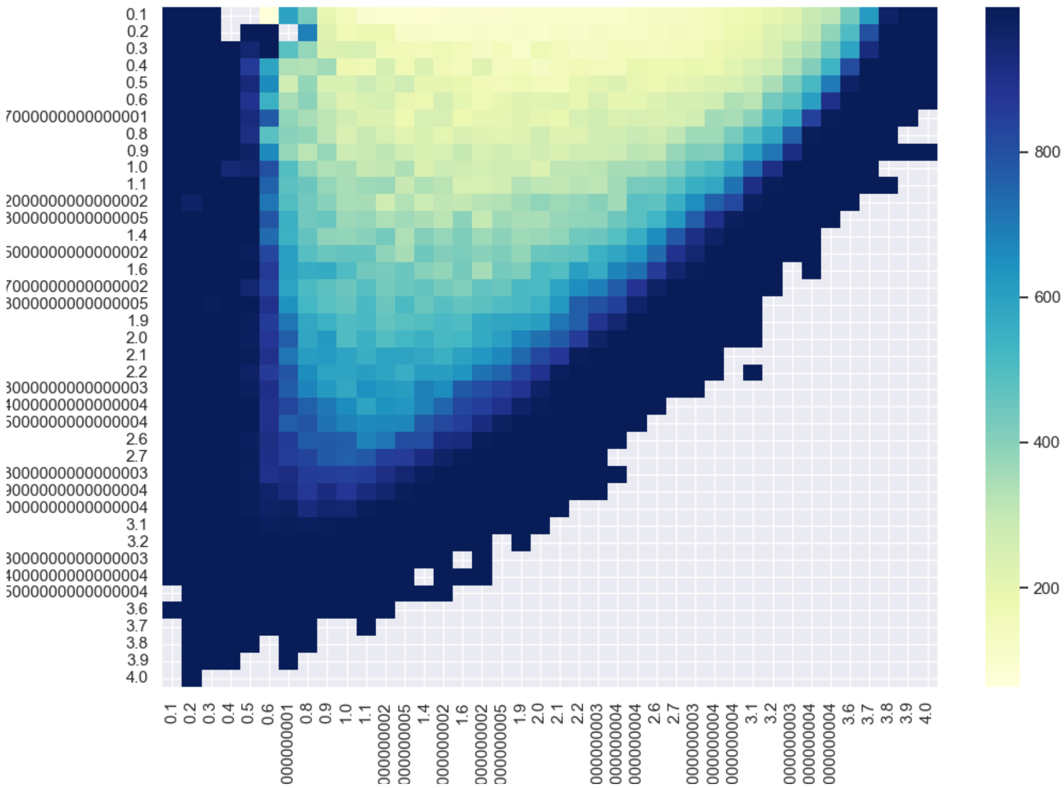


The next graph shows the results of varying the social parameter. In this case, Rosenbrock had a clear performance increase at around 2. The high and low ends caused the simulation to have difficulty in convergence. The booth function experienced a very different effect. Low values saw less variance and more convergence, while high values saw an extreme increase in variance and decreased convergence performance.

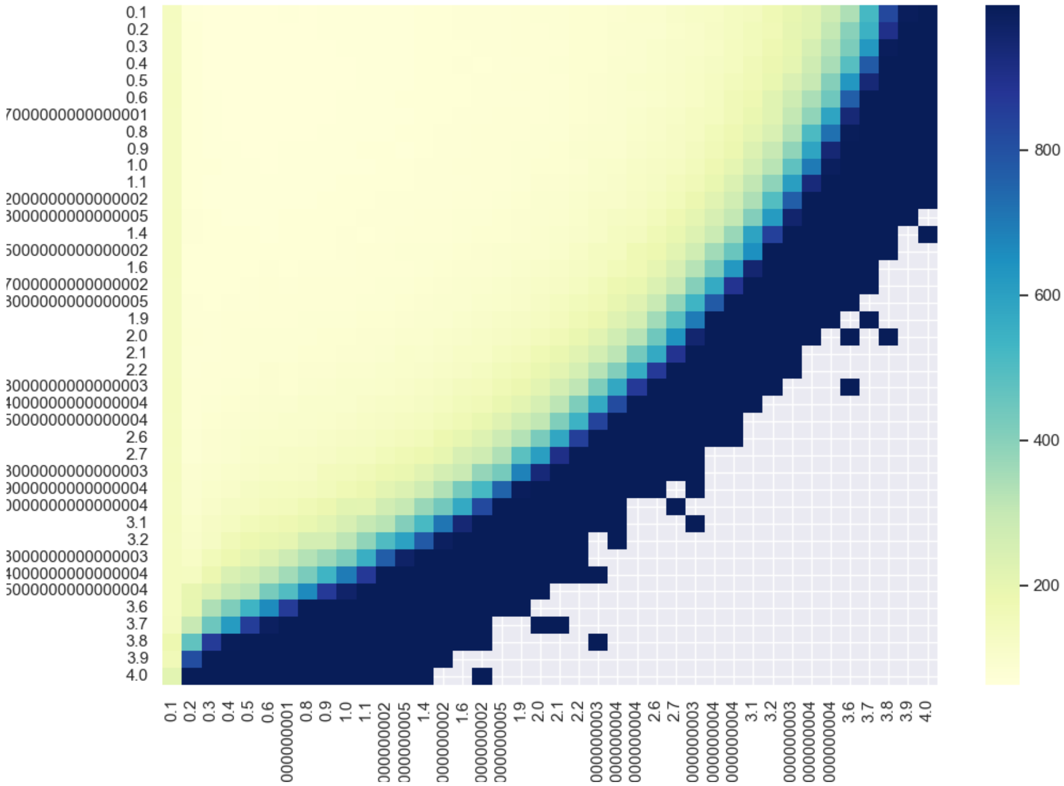


The final parameter tested was cognition. For Rosenbrock, this parameter caused worse performance which trended in a linear pattern. Meanwhile the Booth function experienced an exponential behavior where higher values quickly worsened the simulation's performance.

In order to better understand how social and cognitive parameters relate to each other, two heat maps were created. Social represents the rows while cognition represents the columns.



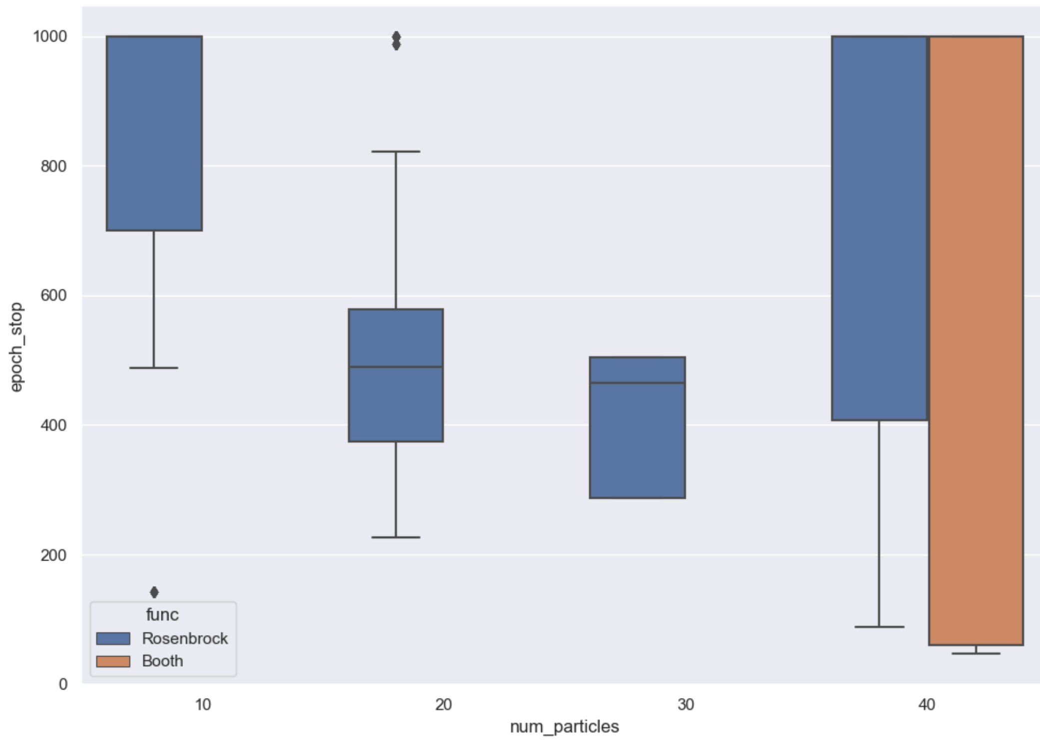
The graph above represents the results of social and cognitive tests involving the Rosenbrock function. It shows the PSO converges faster at a combination of low cognition values, and a social value of around 2.



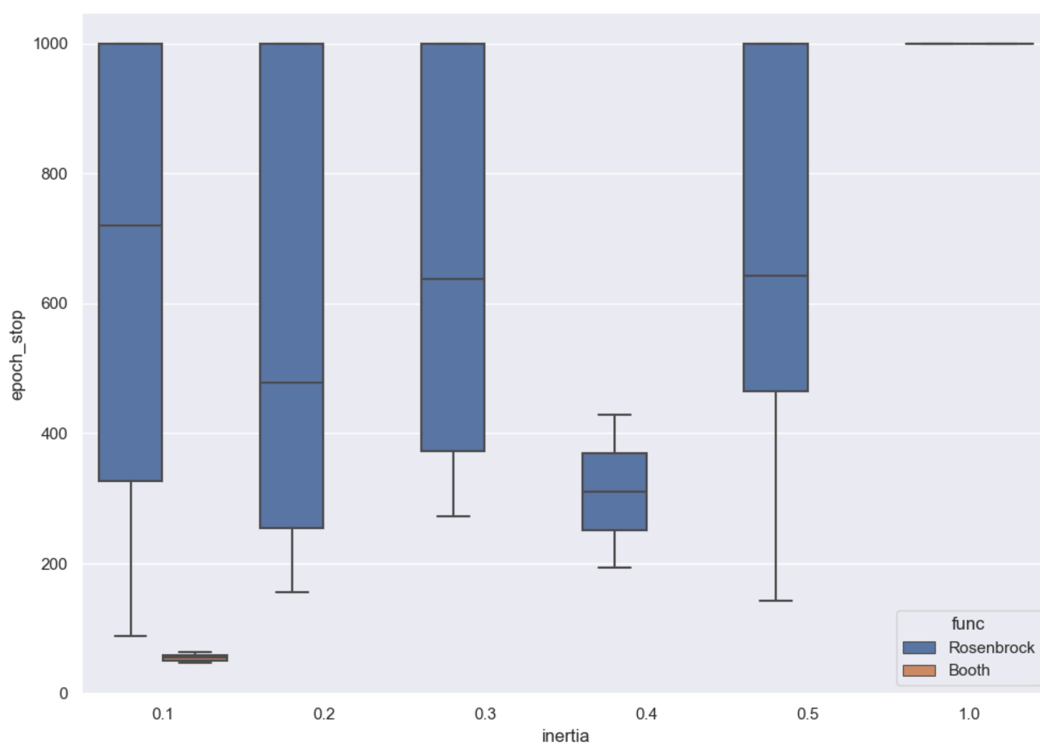
The graph above shows social and cognitive parameter tests using the Booth function. In this case, keeping either value low allowed for quick convergence.

Diverging Simulations

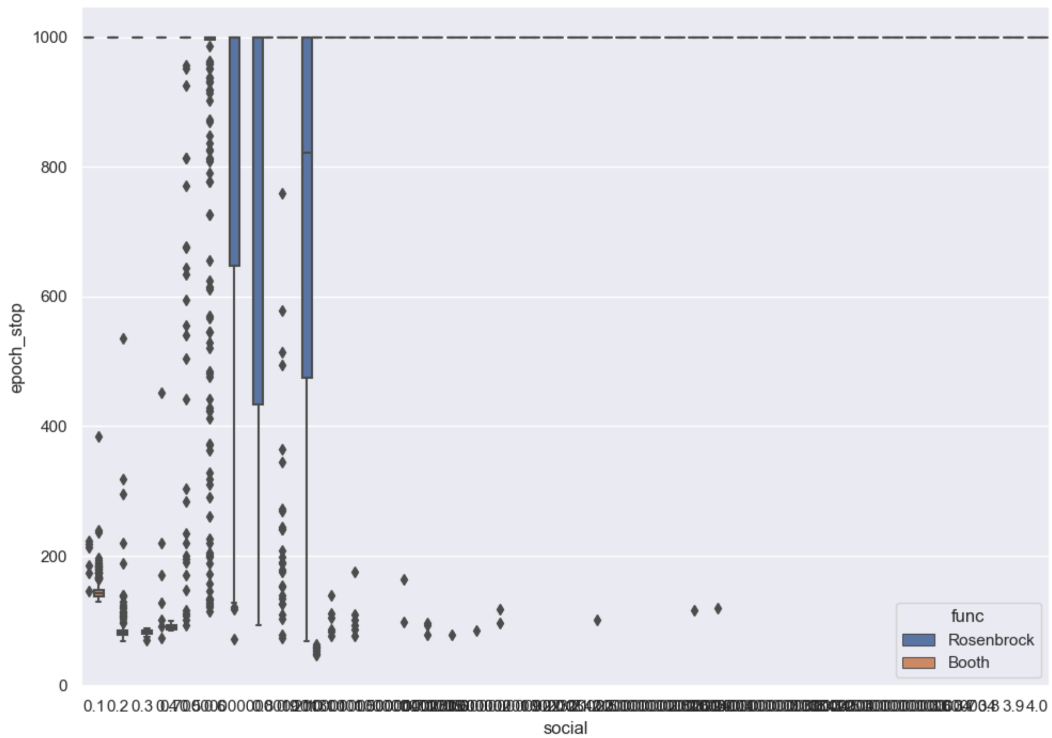
While simulations that did not converge are not as common, there are some occurrences. The following graphs allow for better analysis of how parameters may have impacted the PSOs inability to converge.



In the graph above shows the cases of non-convergence for different values of particles, Rosenbrock had more occurrences of non-convergence. Meanwhile Booth only had issues at 40 particles. In general it seems using less particles caused more difficulty in the PSO.

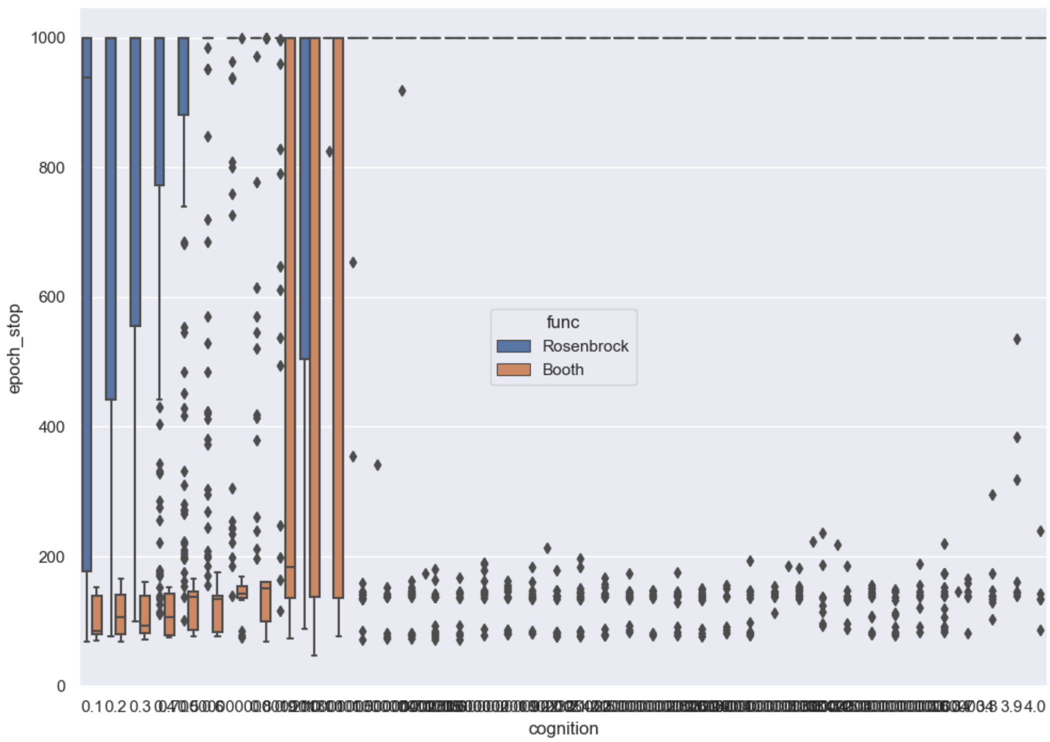


The graph above shows values of inertia. Rosenbrock again had more PSO tests that didn't converge, while the Booth function had little variance in the epochs.

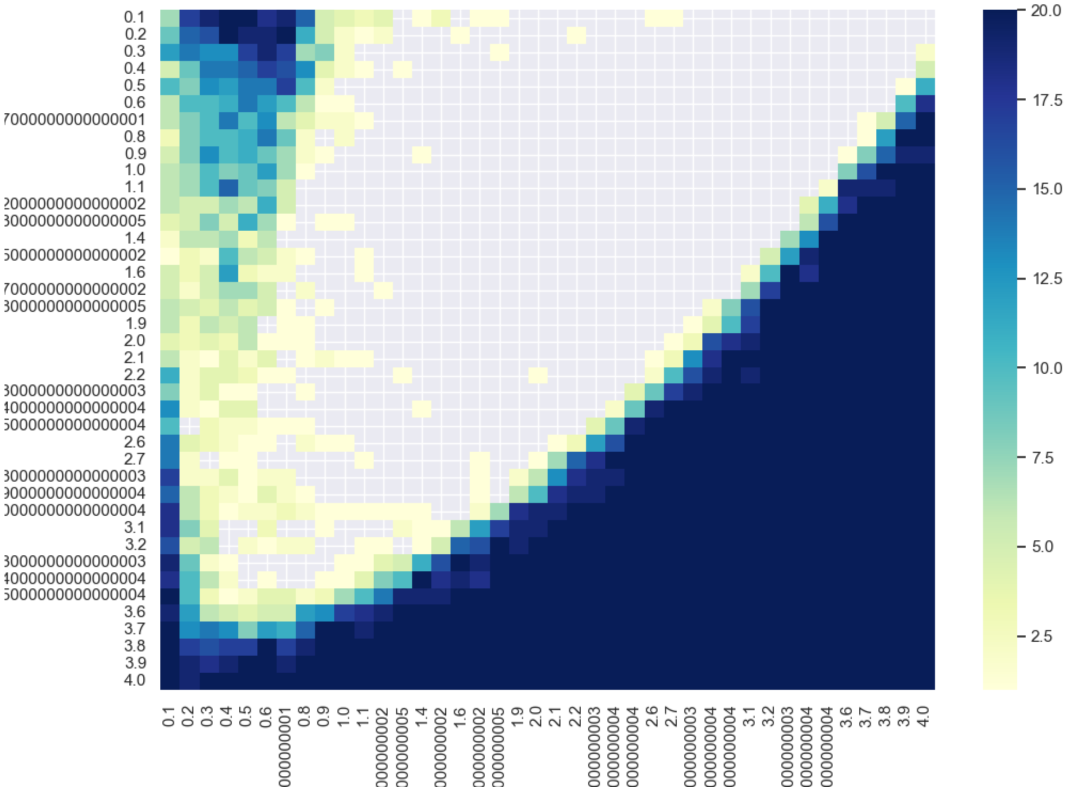


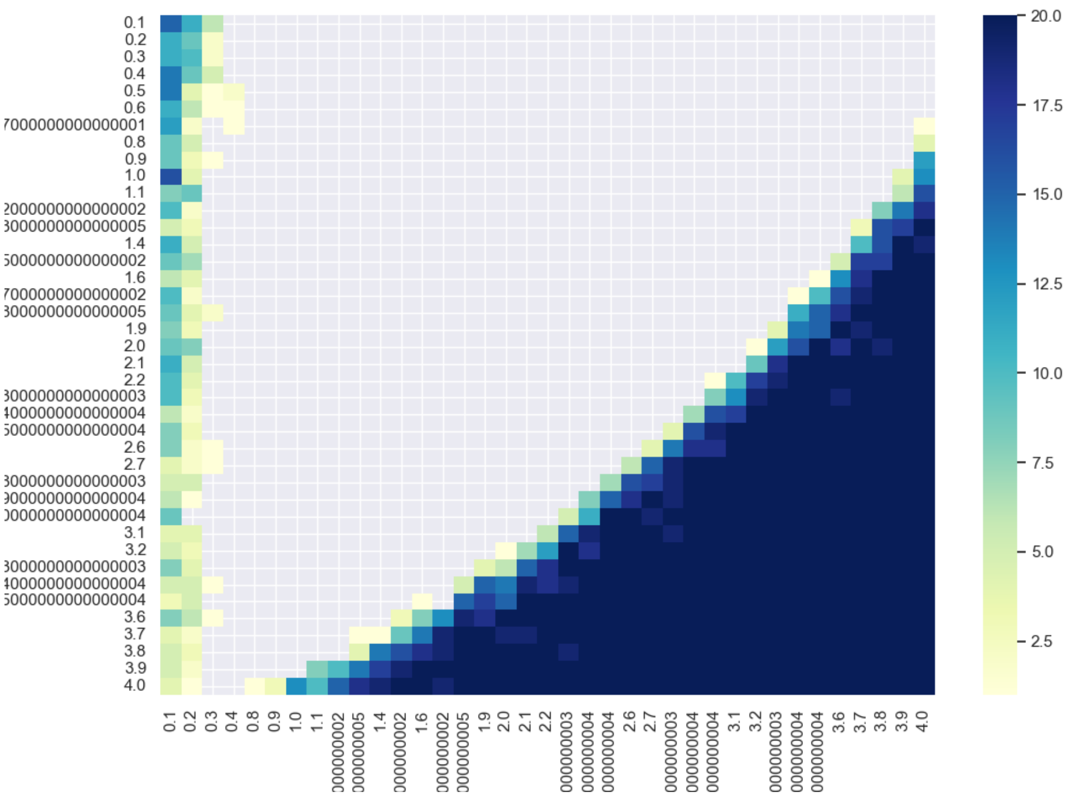
The graph above shows the lower social values caused the simulation to not converge early in the simulation, and more often. Rosenbrock had a higher variance in the values. Booth appeared to end the earlier.

The next graph shows the cognition values for tests that did not converge. For all values, there were occurrences where the simulation did not converge. Lower values appear to cause more variance in epoch times. Booth started with low variance, then quickly increased.



The final two graphs show heat maps comparing Social in the x axis and cognition y axis. In both it seems low values caused an increase of non-converging tests, especially for the social parameter. The first graph is the Rosenbrock tests, and there seems to be a hotspot of non-convergence near the top left. Additionally, there is a large area centered on a social of 2.2 and a cognition of 1.4 where values would converge. The booth function also showed that some social values increase the odds of convergence, but increasing cognition worsened performance. When both values are high, It is extremely likely that the system will not converge. Additionally, if the social value is high, setting cognition low will make the system more likely to converge.


0.1
0.2
0.3
0.4
0.5
0.6
7000000000000001
0.8
0.9
1.0
1.1
2000000000000002
3000000000000005
1.4
5000000000000002
1.6
7000000000000002
3000000000000005
1.9
2.0
2.1
2.2
3000000000000003
4000000000000004
5000000000000004
2.6
2.7
3000000000000003
9000000000000004
0000000000000004
3.1
3.2
3000000000000003
4000000000000004
5000000000000004
3.6
3.7
3.8
3.9
4.0
20.0
17.5
15.0
12.5
10.0
7.5
5.0
2.5


0.1
0.2
0.3
0.4
0.5
0.6
7000000000000001
0.8
0.9
1.0
1.1
2000000000000002
3000000000000005
1.4
5000000000000002
1.6
7000000000000002
3000000000000005
1.9
2.0
2.1
2.2
3000000000000003
4000000000000004
5000000000000004
2.6
2.7
3000000000000003
9000000000000004
0000000000000004
3.1
3.2
3000000000000003
4000000000000004
5000000000000004
3.6
3.7
3.8
3.9
4.0
20.0
17.5
15.0
12.5
10.0
7.5
5.0
2.5

Analysis

In conclusion, this lab tested PSOs with varying parameters of the number of particles, inertia, social, and cognition. Cognition and social parameters appeared to have the largest impact on the model's performance.

In general, simulations performed better with a higher amount of particles. With Rosenbrock being a more strict test, it seemed that a value of 55 was an optimal number of particles. Inertia higher values tended to cause more divergence, especially with Booth. This is likely due to particles continuously moving towards and then passing optimal values. The best values appear to be around 0.3 and 0.4. Social appeared to give better performance towards a value of 2 and cognitive appeared to be better at low values, though the lowest values caused some divergence when social was low. In general, it seemed that social was best set to a value of around 2 and cognition would then be good at around 1.4.

It seems that cognition and social parameters should be scaled inversely in order to achieve good performance. If both are high, the simulation is sure to diverge. If the social value is high, cognition should be low and while cognition is high, a lower social value may help. However if social is low, both very high and very low values of cognition may still cause the simulation to diverge.

A possible reason for cognition and social parameters conflicting with each other is that cognition and social parameters pull particles towards the personal or global optimal value. With high values, the particle is strongly pulled straight to that optima, with the pull getting stronger as they approach. This effect is even worse for cognition because for many particles, that personal best is not the optimal value, so the particles will all be scattered, lowering the fitness score. With a high social, the particles go to the globally known optima without exploring for the best possible solution. For the number of particles, this experiment found no instance above 40 where the simulation did not converge. More particles means the more area the collective is able to explore, and the higher a chance for at least one of the particles to find a very good optima. For inertia, while a little amount may help a particle pass through the boundary region of an optima to get into the center, high values will cause the particles to "slide" around more and so they take longer to converge. At very high values, it may even be the case that the particles will continuously orbit the optima. A similar behavior may be why there is a random hot spot at low social and high cognition. The particles go to their personal best optima, then on a random chance get pulled toward a much better global optima, then on the next step return to their personal optima.