Emerson Hall

Ramana M. Pidaparti

19 November 2025

Final Testing and Validation

Introduction

The goal for Week 11 was to complete final validation of the UAV simulation system by testing it across all terrains and weather conditions and by comparing the performance of the baseline A* controller and the machine learning enhanced controller. This final evaluation represents the most comprehensive test of the entire multi week project, because it exposes the system to a fully deterministic set of scenarios that vary both the physical environment and the strength of environmental disturbance. The focus this week was on consistency, robustness, and the ability of each mode to maintain stable and safe trajectories in urban, natural, and manmade terrains while winds ranged from calm to storm conditions. The results from these tests give a clear indication of how the system behaves at scale and how the machine learning controller compares to the classical baseline approach.

Method

The Week 11 validation script (Github Link) generated deterministic scenarios for each terrain type and assigned three UAVs to start and goal pairs that were guaranteed feasible using thick occupancy inflation and BFS connectivity. Each terrain presented different challenges. The urban environment contained symmetric blocks and road corridors that created tight navigation gaps. The natural terrain contained a central pond, rocks, and trees that forced wide arcs around obstacles. The manmade harbor featured water channels, piers, and warehouse blocks positioned to produce frequent direction changes.

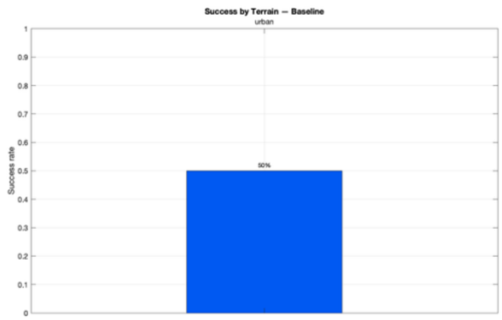
Both the baseline and machine learning modes were tested under calm, breeze, gusty, and storm profiles. The calm condition had no environmental disturbance, while breeze, gusty, and storm introduced progressively stronger background winds and gust magnitudes. Each UAV used repeated A* replanning, path smoothing, shortcutting, and Frenet based tracking with repulsion and boundary sliding. The machine learning mode replaced the core path following behavior with the learned controller from previous weeks.

For every run, the script recorded time to goal, total path length, minimum obstacle clearance, replans, and whether all UAVs reached their destinations. It also produced videos and summary figures for success rate, time and path length distribution, and minimum clearance.

Results

Urban Terrain

The calm condition showed similar performance between the two modes. The baseline completed in 62.9 seconds while the machine learning mode completed in 61.4 seconds. Both handled the map well with no failures. In the breeze condition, the baseline slowed significantly to 95.2 seconds while the machine learning mode finished in 81 seconds, showing that the learned controller handled moderate wind more efficiently. In both gusty and storm conditions, all UAVs failed to reach their goals in both modes, and the simulation ended at the maximum allowed time of 159.8 seconds. The success rate for urban terrain was fifty percent for both modes shown in Figures 1 and 2.



Figures 1 and 2: Success of Urban for Baseline and Machine Learning

The clearance curves for urban terrain showed that both modes maintained similar minimum distances from obstacles shown in Figure 3. Both began near 0.3175 meters at the twenty fifth percentile and rose steadily to approximately 0.4175 meters at the seventy fifth percentile.

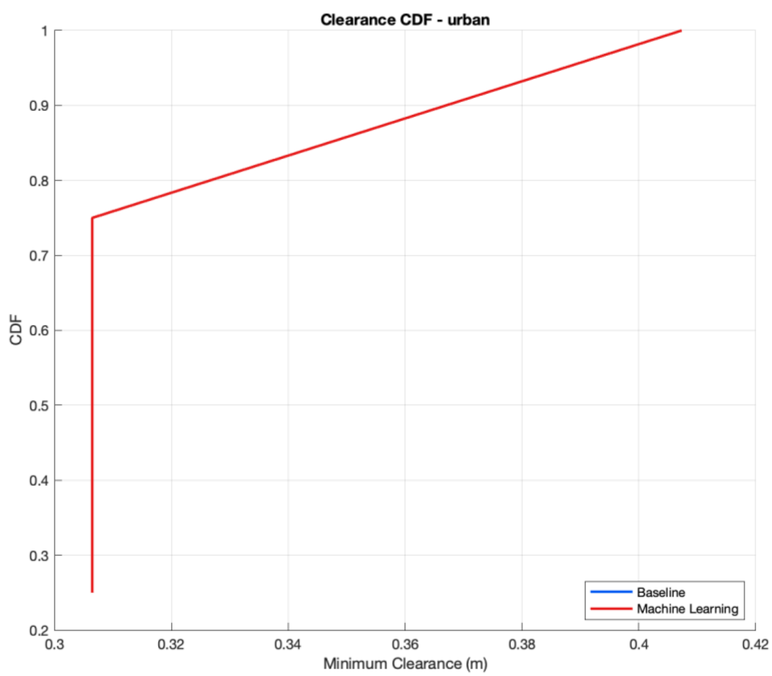


Figure 3: Clearance CDF of Urban

The time distribution shown in Figure 4 revealed that the baseline had a broader spread under wind, with long tails reaching the 160 second limit. The machine learning mode had slightly shorter median times but ultimately reached the same cutoff in the strongest winds. Path lengths for both modes had long upper tails due to repeated replanning, with the longest paths exceeding 575 meters.

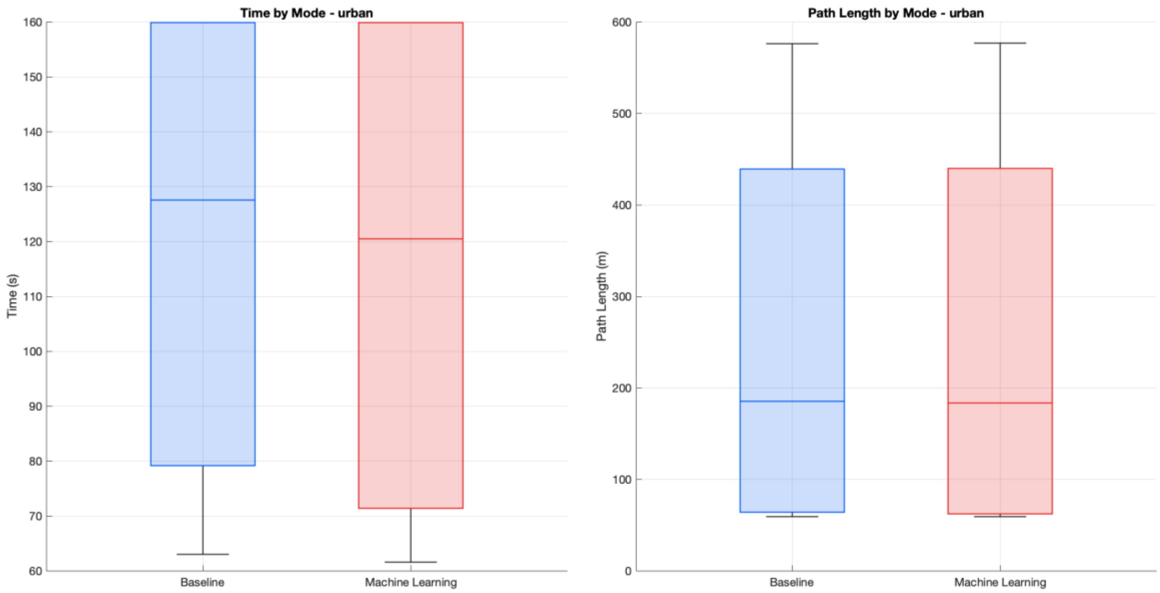
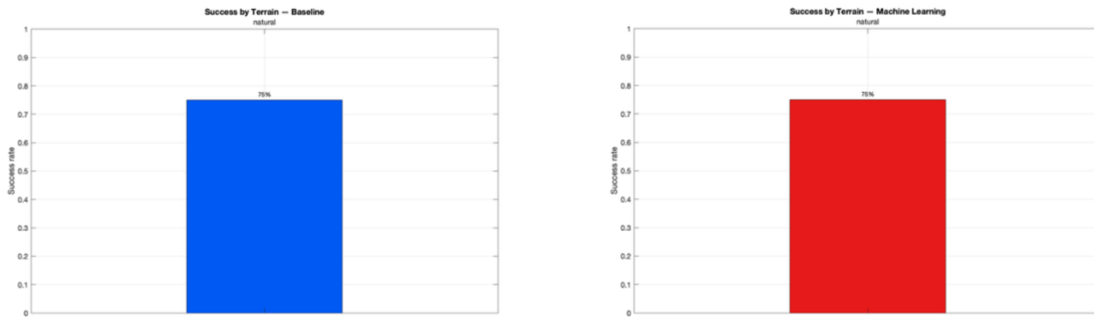


Figure 4: Time by Mode and Path Length by Mode in Urban

Natural Terrain

Natural terrain showed better overall reliability than urban terrain. The calm runs were nearly identical, with the baseline at 53.8 seconds and the machine learning mode at 54.8 seconds. In breeze conditions both modes again produced almost the same performance, with times near 71 seconds. In gusty conditions, the baseline completed in 70.8 seconds while the machine learning mode required 71.5 seconds. Storm conditions resulted in failures for both modes, with all UAVs timing out at 159.8 seconds. The overall success rate for natural terrain was seventy five percent shown in Figures 5 and 6.



Figures 5 and 6: Success by Natural of Baseline and Machine Learning

The clearance plots indicated that both controllers maintained safe distances while maneuvering around the pond and terrain obstacles. The baseline started around 0.375 meters at the lower quartile and reached 0.4675 meters at the upper quartile. The machine learning mode began slightly higher at 0.4375 meters and reached roughly 0.4725 meters at the upper quartile.

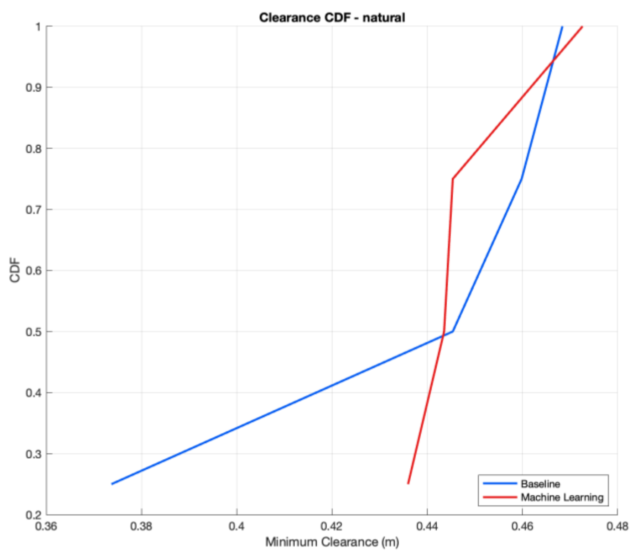


Figure 7: Clearance CDF of Natural

Time distributions for both modes shared nearly identical ranges, while path length distributions remained compact under successful trials and extended to around 187.5 meters for the longest cases. The distributions are shown in Figure 8.

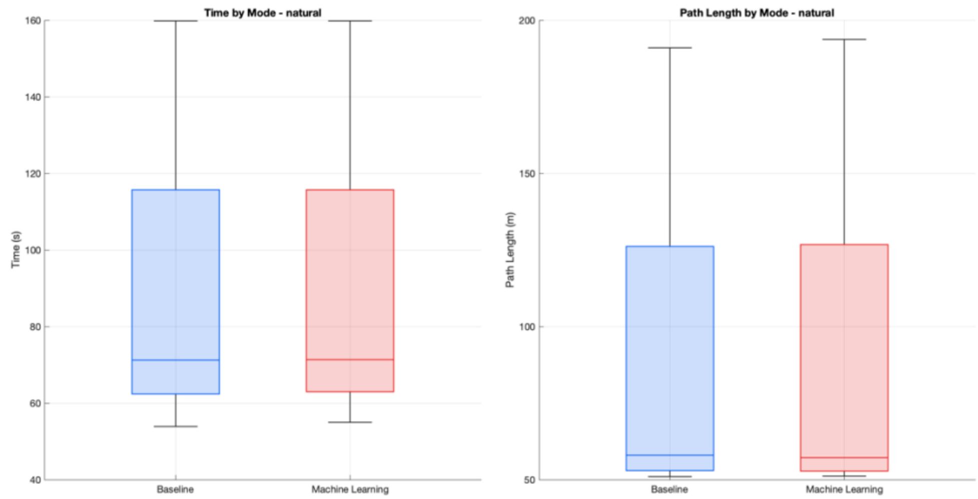


Figure 8: Time by Mode and Path Length by Mode for Natural

Manmade Terrain

Manmade terrain produced the clearest differences between the two modes. In calm conditions, both modes completed successfully at approximately 65 seconds, although both briefly planned trajectories that cut through buildings before correcting themselves. In breeze conditions, the machine learning mode completed more quickly at 54.2 seconds compared to 61.2 seconds for the baseline. In gusty conditions the baseline still completed successfully at 56.3 seconds, but the machine learning mode failed when the red UAV did not reach its goal. In storm conditions both controllers experienced failures for two UAVs and timed out at 159.8 seconds. The final success rate for manmade terrain was seventy five percent for the baseline but only fifty percent for the machine learning mode shown in Figure 9 and 10.

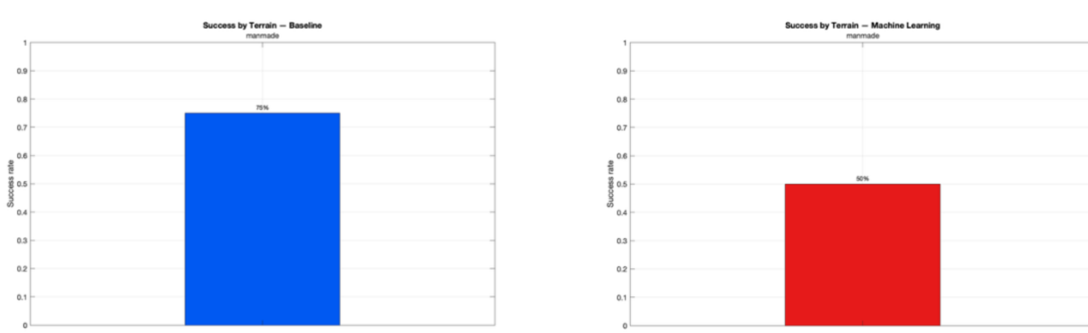


Figure 9 and 10: Success by Manmade for Baseline and Machine Learning

The clearance plot shown in Figure 11 showed that the baseline maintained a minimum clearance beginning around 0.25 meters and rising to approximately 0.5145 meters, while the machine learning mode began slightly higher near 0.3065 meters and ended near 0.527 meters.

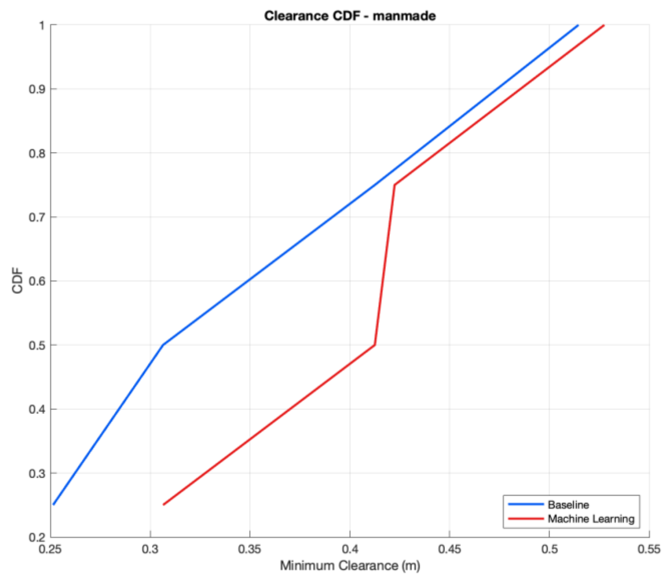


Figure 11: Clearance CDF for Manmade

The time distribution shown in Figure 12 indicated that the machine learning mode had faster completion in mild wind but showed large variance and occasional failure in high wind. Path lengths displayed a wide range for both modes, with upper values reaching more than 300 meters due to rerouting around the harbor structures.

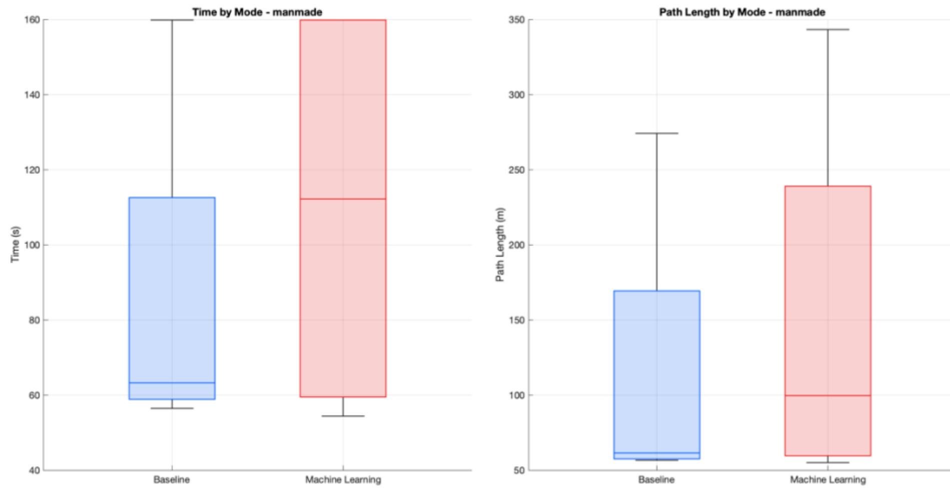


Figure 11: Time by Mode and Path Length by Mode for Manmade

Discussion

The validation results demonstrated that the overall system handled calm and mild wind conditions well in all terrains, but strong wind severely limited performance. In calm scenarios the two controllers behaved almost identically, with small timing differences that reflected natural variation in replanning and trajectory smoothing. In breeze conditions the machine learning mode often completed runs faster, particularly in urban and manmade terrains, which suggests that the learned controller provided smoother or more decisive path following when disturbances were moderate.

However, the machine learning mode struggled in high wind conditions, especially in manmade terrain, where one UAV failed to reach its goal in gusty conditions even though the baseline succeeded. This indicates that the ML controller may be more sensitive to crosswind induced drift, especially in environments with tight corridors or sudden direction changes. In gusty and storm scenarios none of the UAVs in either mode succeeded in urban and natural terrains, which shows that wind magnitude eventually exceeds both controllers' ability to maintain progress toward the goal.

The clearance distributions showed consistent safety margins across modes, which means that the path planners and obstacle inflation strategies remained effective in all terrains. The path length and time distributions provided a clearer distinction between modes, as the machine learning controller tended to take slightly shorter or faster paths in moderate wind but suffered from more variability in high wind. These observations highlight that the machine learning controller improved performance under some conditions but did not fully replace the reliability of classical methods.

Conclusion

Week 11 successfully validated the full UAV simulation system across urban, natural, and manmade terrains under four levels of wind disturbance. The results showed that both the baseline and machine learning modes performed reliably in calm and mild conditions, but both struggled as wind strength increased. The machine learning controller demonstrated improved efficiency in several environments but showed reduced reliability under strong wind or in more complex terrains. The baseline remained more consistent but sometimes slower.

This final validation provided a complete comparison between the two methods and highlighted where each approach performed best. The results mark the conclusion of the multi week development process and establish a foundation for future work, which could improve the machine learning controller's robustness or combine it with traditional planning to create a hybrid system that adapts to environmental difficulty.