2026 NDIA MICHIGAN CHAPTER
GROUND VEHICLE SYSTEMS ENGINEERING
AND TECHNOLOGY SYMPOSIUM
AUTONOMY & ROBOTICS TECHNICAL SESSION
AUG. 11-13, 2026 - NOVI, MICHIGAN

STUCK-ASSIST: A LARGE MULTIMODAL MODEL-BASED STUCK RECOVERY ASSIST FOR DRIVING ON OFF-ROAD DEFORMABLE TERRAINS

Mayuresh Bhosale[1], **Jordan A. Whitson, PhD**[3], **Ardalan Vahidi, PhD**[2], **and Yunyi Jia, PhD**[1]

[1]Department of Automotive Engineering, Clemson University, Greenville, SC 29607, USA.
[2]Department of Mechanical Engineering, Clemson University, Clemson, SC 29634, USA.
[3]US Army DEVCOM GVSC

ABSTRACT

Recent advancements in off-road autonomy have shown significant progress in perception, planning, and control frameworks, including end-to-end learning approaches. Comprehensive results have been demonstrated in both simulation and real-world experiments; however, there are significant challenges in critical cases that need further evaluation. One such challenge is the immobilization of autonomous ground vehicles (AGVs) in unstructured off-road environments, which can significantly impact agriculture, space exploration, military operations, and search and rescue missions. Addressing this problem requires recovery strategies that are context-sensitive, adaptable to terrain and vehicle conditions, and effective in integrating multimodal inputs. To this end, this paper investigates the use of a large multimodal model (LMM) providing higher-level planning assistance with human-in-the-loop evaluations for vehicle recovery after immobilization in unstructured off-road terrain. The experimental simulation platform developed was based on the Algoryx (AGX) Dynamics engine for high-fidelity terramechanics interaction and vehicle physics combined with Unreal Engine 5. This platform was further integrated with a driving simulator equipped with steering wheel and pedal interfaces for human-in-the-loop experiments. We evaluated ten representative unstuck scenarios across two deformable terrains (loose sand and compact sand) under two modes: an unskilled baseline, where participants attempted recovery unaided, and a co-intelligence mode, where participants used LMM advisory instructions. The results show that LMM assistance improved stuck recovery rates by 70% compared to unaided and unskilled human driving.

Citation: Mayuresh Bhosale, Jordan A. Whitson, Ardalan Vahidi, Yunyi Jia, "STUCK-ASSIST: A LARGE MULTIMODAL MODEL-BASED STUCK RECOVERY ASSIST FOR DRIVING ON OFF-ROAD DEFORMABLE TERRAINS," In *Proceedings of the Ground Vehicle Systems Engineering and Technology Symposium* (GVSETS), NDIA, Novi, MI, Aug. 11-13, 2026.

1. INTRODUCTION

In critical safety cases such as an AGV stuck or immobilized—defined as the inability to generate sufficient traction on unstructured off-road terrain due to the presence of soft soil, gravel, mud patches, or ditches—conventional autonomy stacks offer limited help [1]. Most existing algorithms prioritize avoidance through perception-driven risk assessment tasks [5], [4] or nominal control with

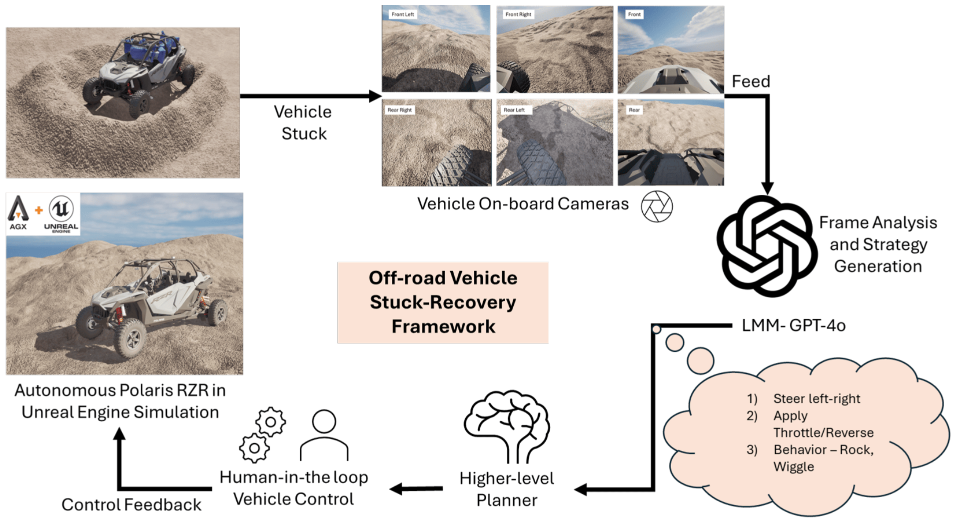

Figure 1: The LMM-assisted stuck recovery framework integrates a high-fidelity simulation platform with vehicle telemetry, GPT-4o advisory instructions, and execution with human-in-the-loop.

slip regulation, torque distribution, and path planning [2]. Once deeply stuck in off-road deformable terrain, these approaches fail because the available sensors provide cues that may not account for the importance of load transfer in AGV wheel-terrain contact, and the required actions are sequential and adaptive to gain traction. This emphasizes the need to develop stuck recovery methodologies for AGVs from a safety-critical point of view.

Off-Road Autonomy: With recent advancements in off-road autonomous driving, significant progress has been made in planning for safe traversal. Self-supervised learning approaches, such as probabilistic and evidential risk-aware mapping [3], [5], quantify risk in the terrain to AGV immobilization. Rather than labeling risky terrain as out-of-distribution data, they treat high-uncertainty regions conservatively. Reinforcement learning-based hierarchical planners have also been developed to minimize the slip [6] and avoid paths that are considered risky [7].

Further methods addressing traversal over deformable terrain typically focus on learning terrain parameters [8], [9] or compensating for unknown dynamics [10] for model-based control strategies but have not explicitly addressed stuck recovery situations [20]. Complementary efforts such as Stochastic Traversability Evaluation and Planning for Risk-Aware Off-road Navigation [4] have demonstrated avoidance of hazardous traversals due to water risk or negative obstacles and developed recovery behaviors including Wiggle, Escape Lethal, and Tilt Recovery. However, these efforts did not explicitly target stuck recovery as a distinct research problem.

LMMs for Autonomous Driving: In contrast to traditional off-road recovery and path re-planning, human operators typically employ multi-step procedures: they assess visual and kinematic signals, adjust strategies across multiple attempts, and execute concise, interpretable control inputs (e.g., “steer left while slightly pulsing the throttle” or “reverse to firmer ground, then approach diagonally”). Such interpretation and adaptation is critical in stuck scenarios, as highlighted in off-road self-recovery guidebooks [17], [18].

Adding on these human-like decision-making qualities, recent frameworks based on LMM highlight their capability for adaptive autonomous driving with reasoning. Within these approaches, GPT-DRIVER [11] and OPENEMMA [12] align the motion planning problem with language modeling, which generates accurate trajectories with enhanced chain-of-thought reasoning. In addition, LMMs have been applied to behavior planning, where DRIVEMLM [13] utilizes high-level decisions generated by LMMs to control the vehicle in the CARLA simulator. Consequently, DRIVEGPT4 [14] and LANGPROP [15] demonstrated potential for end-to-end learning capability, where LANGPROP [15] iteratively refines the policy through optimization and code generation. Together, these studies demonstrate how LMMs can improve vehicle autonomy through reasoning-based trajectory generation, control system integration, and adaptive policy refinement.

For off-road scenarios, GOONDAE [16] introduced a denoising auto-encoder-based driver assistance platform for teleoperation, explicitly treating unskilled driver inputs as noisy expert demonstrations, leading to improved control performance under simulated off-road conditions.

Although extensive research in the past few years has advanced LMMs for autonomous driving, most remain focused on on-road tasks such as trajectory planning and traffic management. Off-road efforts, such as GOONDAE [16], concentrate primarily on teleoperation stability rather than stuck recovery. To address these gaps, we propose an LMM-assisted human-in-the-loop architecture for AGV recovery in off-road scenarios. In contrast to previous research that concentrated primarily on perception and navigation, our methodology prioritizes stuck recovery as a distinct challenge and leverages an LMM to produce interpretable, context-sensitive advisory strategies that individuals can decide to follow.

Our main contributions are as follows:

- **Establishing stuck recovery as a safety-critical off-road autonomy requirement.** We demonstrate and highlight the safety-critical importance of stuck-recovery scenarios by presenting the first comprehensive study of these situations on deformable terrain.
- **Investigating and establishing LMM's capabilities as advanced stuck-recovery advisors.** We demonstrate that LMM-generated actions (including rocking, wiggle, reverse escape, and low-throttle escape) can efficiently assist human operators in demanding recovery operations by combining a fine-tuned GPT-4o with knowledge of vehicle recovery and exteroceptive and proprioceptive data.
- **Enhancing the accuracy of human-AI collaboration for off-road.** When compared to unaided unskilled drivers, our human-AI trials in standardized off-road immobilization scenarios show quantifiable increases in recovery success rates.

Collectively, these contributions go beyond existing research on traversability estimation and off-road navigation by specifically addressing stuck recovery as a unique and important safety concern and by demonstrating how future autonomous recovery systems can be supported by human-AI co-intelligence under the guidance of LMMs.

2. METHODS

2.1. System Framework

To evaluate LMM-assisted recovery strategies in unstructured off-road scenarios, an experimental methodology was designed to combine a simulation platform with human-in-the-loop evaluation. The overall framework, as illustrated in Fig. 1 integrates high-fidelity simulation software developed in Unreal Engine 5, as highlighted in Section 2.2; followed by an LMM learning strategy for advisory instructions with designated prompts, as highlighted in Section 2.3; the overall control architecture explained in Section 2.5; and a human-machine user interface (UI) for advisory instructions highlighted in Section 2.6.

2.2. High-Fidelity Simulation Environment Setup

To investigate vehicle immobilization in unstructured off-road environments, a high-fidelity experimental framework was developed using Unreal Engine 5.4 [22] in conjunction with AGX Dynamics [21]. This framework integrates a physics engine capable of capturing rigid-body vehicle dynamics, drivetrain and deformable terrain dynamics. The simulated platform was designed

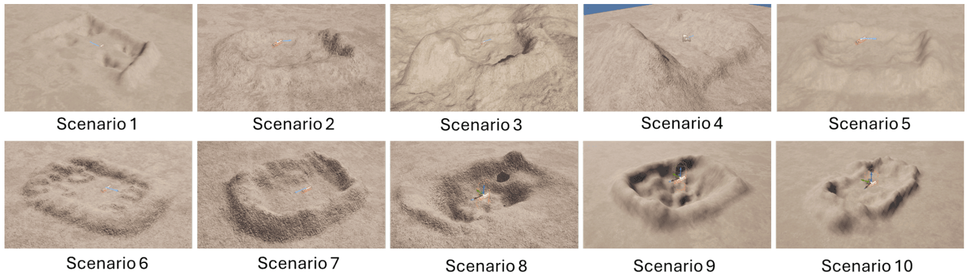

Figure 2: Ten scenarios created in Unreal Engine 5 platform for stuck escape strategy, encompassing two different off-road terrain conditions, including soft and compact sand with uphill, slopes and wedges to evaluate stuck recovery strategies

to resemble a Polaris RZR Pro R 4, incorporating rigid-body representations and detailed drivetrain architecture. The baseline drivetrain model was derived from the AGX Dynamics Wheel Loader template, comprising a combustion-engine torque source, a torque converter, a multi-speed gearbox, and all-wheel-drive capability with differentials.

Terrain model: To better capture the stuck situations, it was necessary to utilize the simulated interaction between vehicle tires and deformable terrain, where wheel sinkage, the bulldozing effect, and shear failure define vehicle stuck recovery. We therefore utilized the AGX Dynamics terrain model, employing a multiscale terramechanics formulation based on the Bekker–Wong pressure–sinkage theory and Mohr–Coulomb plasticity [21] as equation 1. The terrain is discretized into a voxel-based height field, with each cell retaining state variables including density, compaction, and shear strength. Following are the main governing equations for terramechanics from AGX dynamics.

$$\sigma_z^v(z) = \sigma_s \left[1 - \left(\frac{z}{\sqrt{A_s + z^2}} \right)^3 \right] \tag{1}$$

where $\sigma_{vz}(z)$ is the local stress at depth z from the surface, A_s is the projected contact patch area between tire and terrain, and σ_s is the uniform surface stress.

Upon the establishment of stresses, shear failure is assessed for each voxel utilizing the Mohr–Coulomb yield criterion as in equation 2.

$$\tau = c + \sigma \tan(\phi + \psi_{\text{dil}}) \tag{2}$$

where τ represents shear stress, σ denotes normal stress, c signifies cohesion, ϕ indicates the internal friction angle, and ψ_{dil} refers to the dilatancy angle. This regulates the initiation of terrain failure, slope instability, and wheel slippage.

Additionally, AGX Dynamics allows us to incorporate dilatancy and soil compaction, which defines the voxel density changes under repeated loading and how this affects the effective friction angle. Since this increases the soil density, these effects are important because they help provide more realistic terrain consequences.

The voxel-based stress propagation law, the Mohr–Coulomb shear strength criterion, and the compaction–dilatancy connection collectively allow us to simulate essential terrain phenomena, including wheel sinkage, rutting, slope instability, and immobilization. This fidelity is crucial for examining vehicle recovery strategies in unstructured off-road scenarios.

In addition to the high-fidelity physics, we designed 10 different scenarios, as shown in Fig. 2, that were evenly divided into five "easy" and five "hard" cases. Difficulty was defined based on qualitative analysis of terrain features; softer, more deformable soils with steep inclines and higher wedges that required more escape times were categorized as "hard". Whereas scenarios with simpler terrain geometry were categorized as "easy". Combined, these scenarios include soft sand and compact sand terrain, uphill entrapment, downhill ruts, side slopes, obstacle wedges, and hill-surrounded pits. The scenarios were designed to systematically test the vehicle against various off-road conditions that could lead to induced immobilization, as shown in the top-left of Fig. 1.

2.3. Large Multimodal Model Learning Strategy

Expanding upon this simulation environment, we present a new LMM human-in-the-loop strategy that was trained using both expert knowledge and physics-based vehicle data. The GPT-4o model [19], an LMM with strong text and visual processing capabilities, was refined to create the advisory instruction module. We developed a specific fine-tuning dataset from two complementary sources: domain knowledge adaptation and vehicle data derived from simulation, as described below, to customize the model for off-road mobility and recovery tasks.

2.3.1. Domain Knowledge Adaptation

The books "Total Approach to Getting Unstuck OFF-Road" [18] and the "How to Off-Road Guide" [17] were utilized as the basis of the knowledge-based reasoning module encompassing the vast experience they offer in off-road recovery. The texts described systematic ideas of tire-soil interaction, vehicle dynamics, and recovery methods for a variety of terrains, such as rocky slopes, mud, snow, and sand. By combining various sources, we intend to ensure that the model had professional recovery methods and fundamental off-road driving principles that went beyond what could be learned from only unstructured simulation data. Question-answer tokens designed for the GPT-4o fine-tuning format were created from relevant sections, such as traction recovery, hill ascent/descent control, and vehicle rocking procedures, as indicated below:

Guide Book Training Pair Sample

User: What action should an autonomous off-road vehicle take when it experiences high slip on a sandy incline?

Assistant:
Strategy: Stabilize traction and replan ascent.
Rationale: On sandy slopes, sustained high

Guide Book Training Pair Sample (continued)

slip rapidly degrades bearing capacity; a brief retreat and path adjustment preserves momentum.
Plan:

1. Abort current climb and hold position to assess slope aspect and compaction.
2. Reverse 1 vehicle length to firmer ground while keeping wheels straight to avoid bulldozing.
3. Re-approach on a shallower traverse that reduces required drawbar pull. Use a short momentum run-up from firm ground.

2.3.2. Simulation-Derived Vehicle Data

Structured vehicle data pairs were formatted to include vehicle telemetry, on-board camera feeds as images, and labeled successful or unsuccessful recovery attempts. These data included steering angle, throttle or brake response, GPS locations, wheel speed and terrain condition through images targeting the multiple-vehicle views in the following format:

Vehicle Data Sample

Timestamp: xxx sec
Image URL: Frame xxx
Steering Angle (rad): xxx
Steering Command (rad): xxx
Throttle Command (0-1): xxx
Brake Command (0-1): xxx
Wheel Speeds (rad/s): [FL: xx, FR: xx, RL: xx, RR: xx]
World Location and Rotation: [xxxx, xxxx, xxxx, xxxx, xxxx, xxxx]

2.4. Structured Prompt–Response Design

To ensure consistency, reduce ambiguity, and generate explicit control commands, a structured prompt for test evaluation further guided the learned and fine-tuned LMM. This method removes unnecessary responses and ensures compatibility with the human–machine interface (HMI), in contrast to general-purpose conversation responses. The steering and throttle settings were adjusted with bounds $[-1.0, 1.0]$ in order to be consistent with the simulation interface, aligning from left-to-right steering lock. Throughout all simulation iterations, this initial prompt was used consistently:

Prompt to LMM

System Instruction:
You are an off-road driving assistant helping recover a stuck AGV. You must respond ONLY in this format:
Behavior: <one of: Rocking, Directional escape, Wiggle, Low torque escape, Reverse escape>
Throttle: <float between -1.0 and 1.0>
Steering: <float between -1.0 Left and 1.0 Right>

User Inputs: Same format as per vehicle data sample

LMM Response

Assistant:
Behavior: xx, Throttle: xx, Steering: xx

This integrated example shows the complete pipeline; simulation-derived data are wrapped into a constrained prompt, and the LMM returns an interpretable command. The behavior label provides the high-level strategy, throttle, and steering input with a wiggle maneuver as a behavior.

2.5. Control Architecture

The vehicle actuation loop was executed via a ROS 2 control stack that integrated the simulator, driver inputs, and GPT advisory commands. Participants engaged with the vehicle using a steering wheel and pedal assembly, which produced continuous impulses for steering, throttle, and braking. These were published on specific ROS topics for steering, throttle and brake command and accepted by the Unreal–AGX simulation interface.

Vehicle telemetry, comprising steering angle, wheel speeds, wheel slip ratios, longitudinal and lateral velocities, and global position, was simultaneously published on ROS topics and later accepted by the GPT4o-based LMM. Control operations were executed within a closed-loop architecture: the driver applied commands, the simulated vehicle reacted under AGX physics, and the resultant states were transmitted to both the LMM and the driver user interface explained in subsection 2.6.

In co-intelligence mode, GPT-4o advising signals were analyzed into organized control recommendations (maneuver type, steering goal, and throttle/brake modulation). These were not implemented as independent actions but were presented to participants, maintaining shared autonomy in which human operators executed the control inputs as advice.

Figure 3: Driver-facing HMI portraying a simulated off-road vehicle cockpit. The display features lateral, forward, and backward views, and the UI supports LMM-assisted instructions with arrows for direction, forward/reverse movement, and maneuver behavior.

2.6. Human-Simulator User Interface for Instruction Delivery

The HMI was established as a driver-centered dashboard, similar to a layout in a real Polaris RZR cockpit, and was fully incorporated within the Unreal Engine simulation environment. As illustrated in Fig. 3, the HMI provided both visual situational awareness and multimodal advisory assistance.

2.6.1. Camera perspectives

Several scene capture feeds from cameras were displayed to mimic a natural perspective within the driver's field of vision, encompassing forward, rear, and lateral views. This offered participants spatial context typically provided by mirrors or head motions in a physical vehicle.

2.6.2. Advisory overlay

The LMM's guidance was given in a two-layered instruction layout with textual indicators and a directional arrow. **Textual indicators** showing recovery behavior strategy, together with the following movements for advice: **Wiggle:** Instructs for lateral steering movements while adjusting to the suggested throttle or brake instructions to take advantage of slight variations in the contact patch, allowing to gain traction in terrains such as loose mud or sand. **Rocking:** Rapidly switching between the throttle or reverse or punching the throttle to break traction from loose sand and build up momentum, utilizing the compaction of the terrain to gain traction. **Reverse Escape:** directing the car to a less steep surface or allowing it to back up enough to gain momentum for forward movement. **Directional Escape:** Instructions to select an alternate path by steering in one direction (left/right) rather than making an aggressive escape. **Low Throttle Escape:** Instructions to maintain low throttle input, limit tire slip, reduce soil shear, and gradually increase or maintain traction—especially on highly deformable terrain. **Forward or Reverse** Text with the forward command is shown in green, and the reverse command is shown in red.

Directional arrow represents changes in direction based on the LMM instructions for steering values as shown in Fig. 3. The steering values are derived as direct input from LMM intending to choose a goal point.

This design allowed participants to rapidly associate advisory instructions with control actions without reading for extended periods of time by lowering the cognitive translation effort.

3. EXPERIMENTAL DESIGN

Figure 4: A participant is interacting with the stuck-recovery simulation setup with co-intelligence mode, demonstrating the LMM-assisted recovery interface.

3.1. Participants

Ten volunteers with no previous off-road driving experience were recruited to eliminate expert bias. All participants were given a demonstration of the simulation experiment setup and a trial test on a random scenario to become familiar with the scenarios and the two testing modes as shown in Fig. 4.

3.2. Testing Parameters (Scenarios and Modes)

Every participant undertook trials across ten immobilization situations specified in the Method Section 2.2: that includes soft sand pit, uphill entrapment, downhill rut, side-slope embedment, diagonal obstacle wedge, loose gravel ascent, mud patch, rocky step, mixed deformable-rigid transition, and compound obstacle.

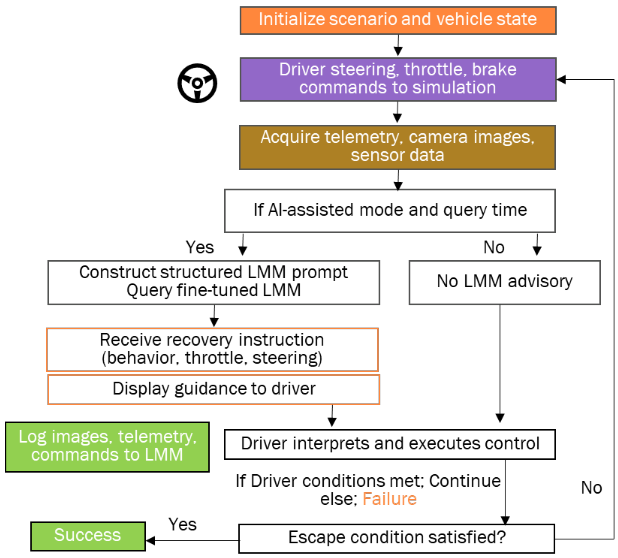

Figure 5: AI-agent-based stuck escape evaluation algorithm

Two modes were assessed; **Unaided unskilled baseline:** Where participants lacking prior experience in off-road driving attempted recovery after getting stuck through their intuition only. These trials recorded the natural response of human drivers without expertise to recover. **Co-intelligence mode:** Where participants were provided with LMM-generated advisory instructions through the UI, specifying behavior maneuver, steering direction, and throttle/brake/reverse commands. The intention of the instruction was conveyed as advisory and not explicit to follow, where participants have complete autonomy to adhere to or reject the guidance instruction. This allowed assessment of whether LMM-generated advisory instruction could enhance recovery outcomes.

This dual-mode protocol was implemented for three reasons from an experimental design standpoint; **Benchmark against inexperienced intuition:** By choosing human participants with no prior off-road experience, the unaided unskilled mode creates a lower baseline. **Evaluation of human-AI collaboration:** The relationship between human adaptability to perception and structured advising recommendations from an LMM is demonstrated by the co-intelligence mode. In contrast to a fully autonomous controller, this setup emphasizes shared autonomy, with the human controlling and regulating at the control-actuation level while the LMM provides higher-level planning strategies. **Diversified terrain robustness:** The observed effects are not limited to a particular terrain type; rather, they generalize across a wide range of immobilization situations when both approaches are evaluated across diverse scenarios. This resulted in a within-subject design in which each participant encountered all scenarios both with and without LMM assistance. The LMM-guided test evaluation algorithm is shown in Fig. 5.

This approach directly addresses the primary research question: *How much can LMM-based planning help inexperienced drivers recover a vehicle from immobilization in challenging, deformable terrain?*

3.3. Experimental Equipment

We utilize Unreal Engine 5.4 with AGX Dynamics as our simulation platform operating on a single Nvidia RTX A6000 GPU. We use a force-feedback steering wheel and pedal setup based on the Thrustmaster FGT Rumble driving simulator.

3.4. Test Protocol

Each trial started with the vehicle in a uniform stationary state with scenario-specific obstacle conditions, a preset position, and soil characteristics. The availability of instructions was dependent upon the chosen mode. For the **unskilled and unaided baseline** mode, the advisory panel is hidden, and participants had to rely only on their own intuition. In the **co-intelligence condition** mode, the advisory UI was activated, allowing participants to utilize the LMM instructions for recovery assistance.

The termination criteria for the immobilization scenario are described as follows. **Success** attempt was recorded when the vehicle surpassed the preset location from the immobilization scenario, exceeding more than one vehicle length. **Failure** was recorded if the participant chose to end the attempt based on their subjective evaluation that further recovery was not possible.

3.5. Randomization and counterbalancing

To prevent the ordering effects, the scenario sequence was randomized for each participant. To balance the exposure to two modes, unskilled versus co-intelligence, the trials were randomly interspersed. This guarantees that the participants switched from modes and conditions instead of running all the trials at once per mode or sequentially. This block randomization approach removes any bias from the learning experience, adaptation to a single mode, or fatigue, ultimately ensuring comprehensive within-subject evaluation with modes and scenarios.

4. EXPERIMENTAL RESULTS

The assessment was performed by examining the multimodal logs collected from all participant trials. Performance was evaluated using the following quantitative metrics: escape success rate (ESR) and escape time (ET) for successful trials. Comparisons were conducted within participants between the unskilled baseline and co-intelligence modes throughout the ten immobilization scenarios.

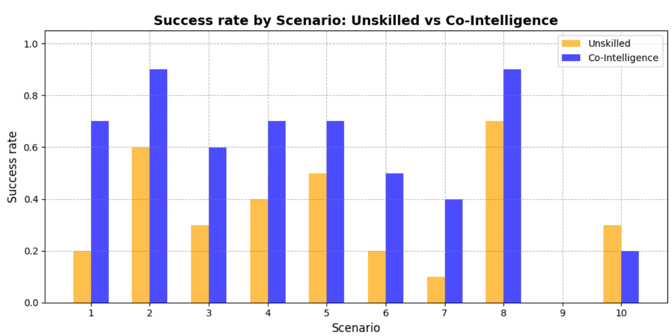

Figure 6: Success rates of escape across 10 off-road scenarios, comparing unaided unskilled driving (baseline) with co-intelligence mode with LMM guidance. The co-intelligence mode consistently outperforms unaided driving.

4.1. Stuck Escape Success Rate

Fig. 6 illustrates the success rates per scenario. Except for scenarios 9 and 10, the success rate in the co-intelligence mode exceeded the unskilled mode. Out of 100 attempts from 10 participants, an unskilled driver, when unaided, resulted in 33 successful recovery attempts. On the other hand, the co-intelligence mode improved the outcome to 56 successful recovery attempts out of 100. Overall, the co-intelligence mode resulted in almost a 70% increase in the ESR, with the most significant outcomes improved in sand pits (S1, S2), uphill entrapments (S3, S8), and wedge barriers (S5). In contrast, in these scenarios, unaided unskilled drivers ceased their efforts.

When categorizing by difficulty from easy to hard scenarios, as in Fig. 7, success rates with the co-intelligence mode were approximately 75% for easy scenarios and 38% for hard scenarios; in contrast, success rates with the unaided unskilled mode were approximately 48% for easy scenarios and 18% for hard scenarios. Specifically, in hard scenarios, the co-intelligence mode showed a 111% improvement in success rates compared to the baseline, whereas in easy scenarios the improvement was 56%. This suggests that the advisory instruction, being systematic and sequential, effectively assists in hard scenarios where an unaided, unskilled driver usually fails the recovery.

Difficulty	Both	Neither	Only Co-intelligence	Only Unskilled
Easy	23	12	14	1
Hard	7	29	12	2

Table 1: Success outcomes for easy versus hard difficulty across modes.

Table 1 illustrates trial-wise results concerning the difficulty. In easy scenarios, 23 trials resulted in success in both modes, 14 succeeded solely in co-intelligence mode, and only 1 trial succeeded in unaided unskilled driving. In hard scenarios, 7 trials succeeded in both the modes, 12 trials succeeded in only the co-intelligence mode, and only 2 trials succeeded in only the unaided unskilled mode. Thus, there are more successful attempts for co-intelligence than in the unaided unskilled mode. However, 29 hard and 12 easy scenarios failed to recover in both modes, emphasizing the severity of the immobilization conditions.

This investigation demonstrates that co-intelligence greatly increased the number of recoverable trials, especially in challenging situations where unskilled participants rarely succeeded when unaided.

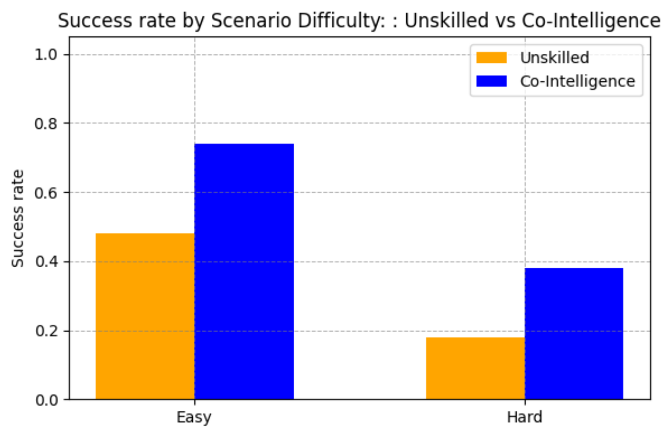

Figure 7: Success rate comparison of unskilled mode versus co-intelligence mode for easy and difficult scenarios. Co-intelligence mode shows significant improvement in a hard difficulty environment.

4.2. Stuck Escape Time to Completion

Fig. 8 depicts the mean escape times (ET) that resulted in successful recoveries across all the trials—compared with unaided unskilled successful recoveries, which resulted in a lower mean ET of 169 seconds. In contrast, co-intelligence's successful attempts resulted in 216 seconds of mean ET. Although the mean ET was higher for the co-intelligence mode, overall success rates were substantially higher. This suggests that participants often required more time and had more confidence in guidance by structured advisory instructions from the LMM model.

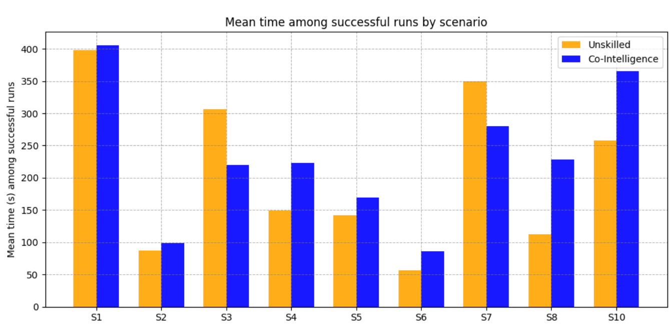

Figure 8: Average escape times (ET) for successful recoveries in both modes. Although co-intelligence mode results in increased recoveries, it often requires more time due to the systematic use of structured advisory procedures.

This pattern highlights a compromise between reliability and time efficiency. Unaided unskilled drivers occasionally depend on harsh acceleration or repeated steering inputs, often resulting in quick escape times and more failures. In contrast, drivers utilizing co-intelligence assistance preferred systematic, low-risk control techniques (eg, slight wiggle, rocking, and low-throttle maneuvers), not because they were conservative, but because the LMM guidance encouraged this behavior. This resulted in consuming additional time and enhanced escape success rates.

4.3. LMM Instruction Compliance with Human Drivers

Across the trials, human drivers maintained mean compliance to follow LMM instructions at a rate of 53.9% suggesting effectively relying on the model's guidance for just over half of the decisions. The relationship between the instruction compliance and success rates highlights a sophisticated collaborative filtering. Rather than following the instructions of the LMM blindly, human drivers acted as an essential layer in the experiment, filtering the hallucinations. When compared to behavior-type compliance, rocking was the most adhered-to instruction, with a 64.72% compliance rate. Whereas reverse escape, directional escape, and low torque escape showed moderate compliance rates of 50.9%, 51.98%, and 54.94% respectively. Wiggle behavior showed the least compliance rate of 45.21% suggesting that this behavior was most difficult for human drivers to interpret.

Table 2 suggests that the compliance rates across all the scenarios are remarkably stable, suggesting interpretable guidance by LMM is consistent across the obstacles and terrain factors. The consistent LMM compliance suggests that LMM is effective at identifying potential recovery paths and behaviors, while human driver provides a filter to ensure the viability of those physical paths in real time.

4.4. Results at the Scenario Level

At the scenario level, the results displayed this trade-off:

- **Uphill entrapments (S3, S7):** Co-intelligence resulted in reduced ET (see Fig. 8) by guiding

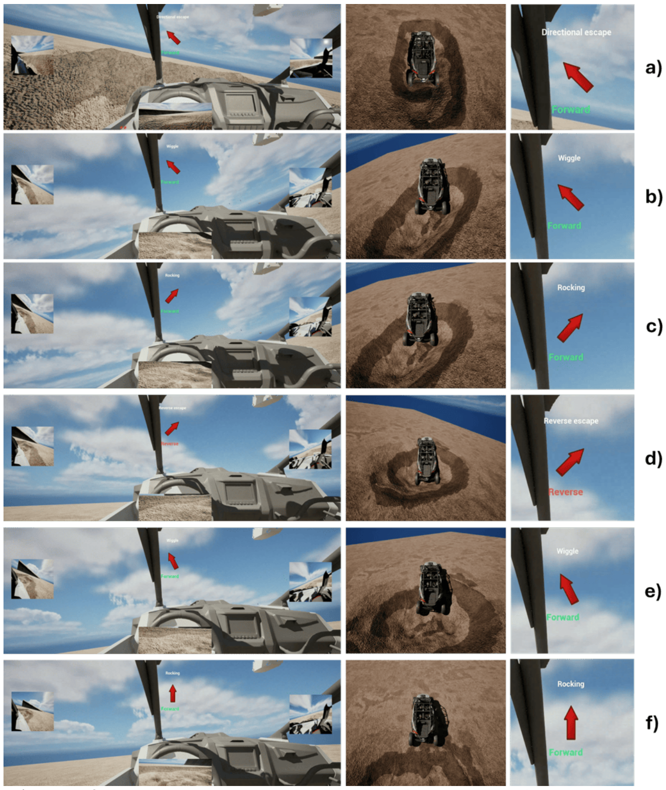

Figure 9: Step-by-step evaluation of LMM-assisted recovery from being stuck in scenario 8. The figure shows a top-down view of the flow of advisory instructions followed. The left column entails the driver's view, the middle column view, unavailable to the driver, is a bird's-eye view for demonstration purposes only, and the right column is an enlarged representation of LMM instructions. a) The vehicle's initial position is inside the uphill entrapment ditch. LMM guides the driver for leftward steering with directional escape behavior and forward throttle. Through these assisted instructions, the driver attempts to climb up the steep hill. b) In addition to the previous step, to gain more traction through the contact patch, LMM instructs wiggle movement in the forward direction for lateral wheel displacement. c) Despite the previous inputs, the vehicle cannot gain traction for further left-forward direction movement. As a result, LMM instructs the driver to rock the throttle in a right-forward direction to gain traction through a jerking movement. d) With no significant movement achieved, LMM adapts to the situation and instructs the driver to reverse while steering right, allowing the vehicle to attempt to align in a different, more favorable direction. e) After sufficient vehicle reverse movement, LMM instructs in forward wiggle movement while steering left, allowing the vehicle's rear wheels to approach the entrapment edge. f) Finally, while maintaining momentum, the LMM assists the driver in applying the throttle in a rocking pattern while steering straight. This results in a sudden gain in traction, ultimately allowing the vehicle to recover from the stuck situation.

Table 2: Mean Compliance and Standard Deviation Across Experimental Scenarios

Scenario	Mean Compliance (%)	Std. Deviation (%)
Scenario 1	52.56	4.94
Scenario 2	54.76	3.56
Scenario 3	53.50	4.23
Scenario 4	51.06	5.11
Scenario 5	52.50	2.07
Scenario 6	55.80	4.00
Scenario 7	53.74	4.06
Scenario 8	55.01	7.09
Scenario 9	53.95	4.94
Scenario 10	55.95	4.49

participants through rocking and directional escape strategies more, allowing them to carry more momentum efficiently. Simultaneously, it allows for higher success rates than the unaided unskilled mode, as shown in Fig. 6.

- **Sand pits (S1, S3, and S10):** ETs were marginally higher in both modes (see Fig. 8), suggesting higher difficulty due to soft terrain. Although challenging, the co-intelligence mode succeeded more often than the unskilled (see Fig. 6). Additionally, both modes consumed similar ETs, with co-intelligence having slightly increased ETs.

- **Scenarios S8 and S10:** Conservative low-throttle escape assistance cues are often followed by the participant in co-intelligence mode, which often results in increased ETs compared to unskilled mode. Fig. 9 highlights the LMM-guided recovery attempt for Scenario 8. The sequence of inputs demonstrates the applicability and adaptability of LMM guidance, wherein dynamic variations in direction, behavior, and pedal inputs result in successful recovery in challenging situations.

- **Scenario S6:** Due to the easy environment, participants in both modes quickly escaped the scenario compared to all the other scenarios, where co-intelligence resulted in significantly higher ESR (see Fig. 6).

Overall results indicate that co-intelligence assistance enhanced robustness in recovery more than time efficiency. The extended mean ET signifies that participants employed systematic and consistent control inputs rather than the risky trial-and-error methods.

5. CONCLUSION

This paper introduced the first systematic investigation into off-road stuck recovery on deformable terrain through a human-in-the-loop simulation framework. Utilizing the AGX dynamics plugin with Unreal provided a higher-fidelity, online simulation platform to be augmented with a fine-tuned GPT-4o model. This simulation technique demonstrated that LMM guidance improved recovery success rates by almost 70% when compared to unaided and unskilled drivers. Although LMM advisory recovery times were often longer, systematic and conservative maneuvers such as wiggling, rocking, reverse escape, and low-throttle escape resulted in higher success rates. These findings identify stuck recovery as a distinct research challenge in off-road autonomy and illustrate the potential of human–AI collaboration with LMMs functioning as high-level advisors to enhance robustness in safety-critical off-road scenarios. In the future, we intend to improve stuck recovery outcomes by integrating LMM with model-based control approaches that can utilize terrain-tire interactions for traction.

References

[1] H. Shon, S. Choi and K. Huh, "Real-Time Terrain Condition Detection for Off-Road Driving Based on Transformer," in IEEE Transactions on Intelligent Transportation Systems, vol. 25, no. 9, pp. 11726-11738, Sept. 2024, doi: 10.1109/TITS.2024.3368476.

[2] J. Yin et al., "Reliable Global Path Planning of Off-Road Autonomous Ground Vehicles Under Uncertain Terrain Conditions," in IEEE Transactions on Intelligent Vehicles, vol. 9, no. 1, pp. 1161-1174, Jan. 2024, doi: 10.1109/TIV.2023.3317833.

[3] X. Cai et al., "EVORA: Deep Evidential Traversability Learning for Risk-Aware Off-Road Autonomy," in IEEE Transactions on Robotics, vol. 40, pp. 3756-3777, 2024, doi: 10.1109/TRO.2024.3431828.

[4] A. Dixit, D. D. Fan, K. Otsu, S. Dey, A.-A. Agha-Mohammadi, and J. Burdick, “STEP: Stochastic Traversability Evaluation and Planning for Risk-Aware Off-road Navigation; Results from the DARPA Subterranean Challenge,” arXiv preprint arXiv:2303.01614, Mar. 2023.

[5] X. Cai, M. Everett, L. Sharma, P. R. Osteen and J. P. How, "Probabilistic Traversability Model for Risk-Aware Motion Planning in Off-Road Environments," 2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Detroit, MI, USA, 2023, pp. 11297-11304, doi: 10.1109/IROS55552.2023.10341350.

[6] V. Wiberg, E. Wallin, T. Nordfjell and M. Servin, "Control of Rough Terrain Vehicles Using Deep Reinforcement Learning," in IEEE Robotics and Automation Letters, vol. 7, no. 1, pp. 390-397, Jan. 2022, doi: 10.1109/LRA.2021.3126904.

[7] Taghavifar, H., Xu, B., Taghavifar, L., & Qin, Y. (2019). Optimal path-planning of nonholonomic terrain robots for dynamic obstacle avoidance using single-time velocity estimator and reinforcement learning approach. IEEE Access, 7, 159347-159356.

[8] J. Dallas, M. P. Cole, P. Jayakumar and T. Ersal, "Terrain Adaptive Trajectory Planning and Tracking on Deformable Terrains," in IEEE Transactions on Vehicular Technology, vol. 70, no. 11, pp. 11255-11268, Nov. 2021, doi: 10.1109/TVT.2021.3114088.

[9] S. Yu, C. Shen, J. Dallas, B. I. Epureanu, P. Jayakumar and T. Ersal, "A Real-Time Terrain-Adaptive Local Trajectory Planner for

High-Speed Autonomous Off-Road Navigation on Deformable Terrains," in IEEE Transactions on Intelligent Transportation Systems, vol. 26, no. 3, pp. 3324-3340, March 2025, doi: 10.1109/TITS.2024.3520520.

[10] P. Gupta, J. M. Smereka, and Y. Jia, "Actor-Critic Cooperative Compensation to Model Predictive Control for Off-Road Autonomous Vehicles Under Unknown Dynamics," in IEEE International Conference on Robotics and Automation (ICRA), 2025.

[11] Mao, J., Qian, Y., Ye, J., Zhao, H., & Wang, Y. (2023). GPT-Driver: Learning to Drive with GPT. arXiv:2310.01415.

[12] S. Xing, C. Qian, Y. Wang, H. Hua, K. Tian, Y. Zhou, and Z. Tu, “OpenEMMA: Open-Source Multimodal Model for End-to-End Autonomous Driving,” arXiv preprint arXiv:2412.15208, Feb. 2025. [Online]. Available: https://github.com/taco-group/OpenEMMA

[13] Z. Zhang, H. Yang, Y. Ma, et al., “DriveMLM: Aligning Multi-Modal Large Language Models with Behavioral Planning States for Autonomous Driving,” arXiv preprint arXiv:2312.09245, 2023.

[14] J. Xu, C. Chen, T. Wang, et al., “DriveGPT4: Interpretable End-to-End Autonomous Driving via Large Language Model,” arXiv preprint arXiv:2310.01412, 2023.

[15] M. Ishida, Y. Huang, A. Zhan, Y. Li, et al., “LangProp: A Code Optimization Framework using Large Language Models Applied to Driving,” in Proc. Eur. Conf. on Computer Vision (ECCV), 2024.

[16] Y. Cho, H. Yun, J. Lee, A. Ha and J. Yun, "GoonDAE: Denoising-Based Driver Assistance for Off-Road Teleoperation," in IEEE Robotics and Automation Letters, vol. 8, no. 4, pp. 2405-2412, April 2023, doi: 10.1109/LRA.2023.3250008.

[17] "The How To Off-Road Guide", https://www.of4wd.com/education/

[18] B. Wohlers, "The Total Approach to Getting Unstuck Off Road: 4WD Self-recovery & Vehicle-assisted Recovery"

[19] A. Hurst, A. Lerer, A. P. Goucher, A. Perelman, A. Ramesh, A. Clark, and I. Kivlichan, “GPT-4o system card,” *arXiv preprint arXiv:2410.21276*, 2024.

[20] Bhosale, M., Whitson, J. A., and Jia, Y. (February 19, 2026). "A Review of Learning Off-Road Terrain Traversability for Autonomous Ground Vehicles." ASME. J. Auton. Veh. Sys. April 2026; 6(2): 020801. https://doi.org/10.1115/1.4070847

[21] https://www.algoryx.se/agx-dynamics/

[22] https://www.unrealengine.com/en-US

6. CONTACT INFORMATION

Yunyi Jia
McQueen Quattlebaum Professor
Clemson University
Department of Automotive Engineering
4 Research Drive
Greenville, SC
yunyij@clemson.edu

ACKNOWLEDGMENT

This work was supported by Clemson University’s Virtual Prototyping of Autonomy Enabled Ground Systems (VIPR-GS), under Cooperative Agreement W56HZV-21-2-0001 with the US Army DEVCOM Ground Vehicle Systems Center (GVSC).