Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Joint Field Characterisation of Onboard Perception, Mapping and Manipulation on a Sub-13 kg Rocker-Bogie Rover

Shahan Sahay
Vellore Institute of Technology (VIT)
Harsh Yadav
Vellore Institute of Technology (VIT)
Avishkar Jaiswal
Vellore Institute of Technology (VIT)
Ravi Tiwari
ravi.tiwari@vit.ac.in
Vellore Institute of Technology (VIT)

Research Article

Keywords:

Posted Date: November 5th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11256718/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Joint Field Characterisation of Onboard Perception, Mapping and Manipulation on a Sub-13 kg Rocker-Bogie Rover

Shahan Sahay[1], Harsh Yadav[2], Avishkar Jaiswal[2], Ravi Tiwari[2*]

[1]School of Computer Science and Engineering (SCOPE), Vellore Institute of Technology (VIT), Chennai, Tamil Nadu, India.
[2*]School of Electronics Engineering (SENSE), Vellore Institute of Technology (VIT), Chennai, Tamil Nadu, India.

*Corresponding author(s). E-mail(s): ravi.tiwari@vit.ac.in;
Contributing authors: shahan.sahay2023@vitstudent.ac.in;
harsh.yadav2023@vitstudent.ac.in;
avishhkar.jaiswal2023@vitstudent.ac.in;

Abstract

This paper reports the joint outdoor characterisation of detection, mapping, local re-planning and grasping running concurrently on one 4 GB embedded graphics processing unit module aboard a 12.982 kg six-wheel rocker-bogie rover, and the power it draws. Platforms of this class are characterised one subsystem at a time. The high-level stack runs as Robot Operating System 2 (ROS 2) nodes. On a held-out split of 250 annotated outdoor frames spanning overcast and direct sunlight, a half-precision single shot multibox detector on a MobileNetV2 backbone reached 89.4% mean average precision at 0.5 intersection over union while the vehicle sustained 14.5 frames per second with mapping, planning and telemetry resident. Planar light detection and ranging inertial mapping over a bounded 10 m by 10 m grass plot returned 11.2 cm positional root mean square error, with 220 ms local re-planning latency, bounding safe ground speed near $0.34\,\mathrm{m\,s^{-1}}$. Thirty autonomous pick-and-place trials on a 12 cm cylindrical target succeeded 23 times: 76.7%, Wilson 95% interval 0.59–0.88. Mean terminal draw under mixed duty was 38.5 W, exceeding the duty-weighted prediction from three isolated states by 16.3 W and giving about 5.3 h of derived endurance. Perception, mapping, manipulation and terminal power were measured together on one sub-13 kg field platform, and the shortfall shows that a budget built from isolated subsystem states underestimates what this autonomy actually draws.

Keywords: Embedded perception, Energy characterisation, Field robotics, LiDAR SLAM, Mobile manipulation, Rocker-bogie rover

1 Introduction

A small outdoor rover carries one processor, one battery and six drive motors, and its detector, its mapper and its arm compete for all three. Taken apart, the pieces are settled engineering. Depthwise-separable backbones [1] paired with single shot multibox detection heads [2] are the standard detector for this compute envelope, and published throughput for that class of module is collected with the detector as the only significant workload [3]. Two-dimensional simultaneous localisation and mapping (SLAM) from a planar scanner has mature open implementations [4, 5], and Robot Operating System 2 (ROS 2) has made the composition of those pieces routine [6]. Assembled on one vehicle and driven outdoors, they interact. That interaction is what this paper measures.

What the literature does not supply is the cost of running them together on a platform small enough to be built from commercial off-the-shelf (COTS) parts. Field studies of mobile manipulation report grasp outcomes with the whole stack online [7], and studies of robot energy report draw against locomotion regime [8], but the two are rarely measured on the same vehicle in the same sessions. A design team sizing a battery around an onboard perception stack has to add together numbers that were never taken together. This paper measures them together on one build.

The platform is a 12.982 kg six-wheel rocker-bogie rover with a passive differential, built from polyvinyl chloride (PVC) tube and sheet metal around COTS electronics. A 4 GB embedded graphics processing unit (GPU) module hosts perception and planning as ROS 2 nodes. An 8-bit microcontroller unit (MCU) owns the actuator loop and holds sole write access to the drives. A 720p camera, a planar light detection and ranging (LiDAR) scanner and a six-axis inertial measurement unit (IMU) feed the perception stack, and a six degree-of-freedom (6-DoF) arm does the picking. The loop closes onboard. Nothing the vehicle needs in order to stop depends on the radio.

The main contribution of this paper is a joint field characterisation of an integrated low-cost autonomy stack, made on one vehicle on one plot. The contribution has three parts.

1. A stack in which a single shot multibox detector (SSD) head on a MobileNetV2 backbone, quantised to half precision (FP16), runs beside two-dimensional LiDAR-inertial SLAM, a local re-planner and a 6-DoF arm on one embedded GPU module aboard a sub-13 kg COTS platform, characterised outdoors on grass rather than indoors or in simulation.
2. Joint measurement of mean average precision (mAP), throughput in frames per second (FPS), positional root mean square error (RMSE) and autonomous grasp success for the same vehicle, on the same ground, in the same sessions, so that the interaction between them is observed rather than inferred across separate papers.

3. An energy account for that autonomy taken at the battery terminals: three isolated operating states, the duty-weighted prediction they generate, the mean actually measured under a continuous mixed-duty run, and the 16.3 W by which the second falls short of the third. The derived endurance that follows is reported with its sensitivity to the one assumption it rests on.

Three kinds of quantity are kept apart throughout, and the sentence carrying each number says which kind it is. Measured quantities were observed in the field: accuracy, frame rate, mapping error, grasp outcomes, link statistics, deflection, terminal wattage, mass, step height and gradient. Derived quantities follow from measured inputs through a stated model: tractive effort, shoulder torque demand, usable energy, endurance and the speed ceiling. Specified quantities come from vendor datasheets and are never offered as results.

The rest of the paper is organised as follows. Section 2 places the work against articulated mobility on deformable ground, detection on embedded graphics processing units, planar mapping where the geometry is thin, grasping from a mobile base in the field, energy accounting for small platforms and telemetry in the 433 MHz band. Section 3 describes the build: the chassis and drive train, the two-tier computation split, the perception and manipulation chain, and the power and safety paths, with the derived mobility and torque quantities stated as models rather than as results. Section 4 sets out the field protocol, the ground-truth procedure, the trial definitions and, at equal length, what the campaign failed to control. Section 5 reports the measurements across terrain, detection, mapping and re-planning, manipulation, energy and telemetry, and discusses each against the comparison points established in Section 2. Section 6 treats what the data cannot support, one constraint at a time. Section 7 closes.

2 Related Work

2.1 Articulated mobility on deformable ground

The rocker-bogie linkage and its passive differential are long established, and the trade-offs behind the flight implementations are documented in detail [9]. Mobility metrics for platforms of this kind have been formalised largely around geometric obstacle-crossing criteria evaluated on rigid steps and in simulation [10]. On grass the binding constraint is usually different. Traction and load transfer govern the achievable gradient well before actuator torque does, a point made directly in the planetary traction-control literature [11] and developed analytically through wheel-soil interaction models [12]. Recent comparative field work on agricultural soil measured tensile force and obstacle-surmounting capability across four robots of different locomotion classes, and found ground interaction rather than installed torque to be the discriminating variable [13]; instrumented slope trials on a six-wheeled rover reach the same conclusion from proprioceptive data [14]. Section 5.1 reports a case in which the observed gradient ceiling sits far below anything the drive torque would forbid, which is consistent with those accounts.

2.2 Detection on embedded graphics processing units

Surveys of neural-network deployment on this class of module report throughput almost entirely on benchmark datasets with the detector as the only significant workload [3]. A recent empirical study on the same module family reports per-model inference times with nothing else resident [15], and a broader comparison across heterogeneous edge devices measures energy per inference and latency for SSD-class detectors under the same isolated conditions [16]. Aboard a rover the detector shares its processor with the SLAM front-end, the planners and the ROS 2 transport, and its battery with six drive motors. Isolated throughput transfers poorly. The frame rate in Section 5.2 is a system-level rate under concurrent load, and is not comparable with those benchmarks.

2.3 Planar mapping where the geometry is thin

Grid mapping with Rao-Blackwellised particle filters [4], scan matching without odometry [17], sparse pose adjustment [18], submap-based loop closure [5] and their maintained successors [19] are the reference implementations, and their published accuracy is predominantly indoor. Outdoors the failure mode is geometric rather than computational. A horizontal scan plane over open ground returns few persistent vertical features, and the registration problem becomes ill-conditioned along the unconstrained direction. That degeneracy has been formalised [20], is detectable online [21], and more recent registration work feeds the detection back into the optimisation so that weakly constrained directions are not updated [22]; learned estimators now predict non-localisability from raw scans before the optimisation is attempted [23]. Work reporting planar SLAM accuracy on open terrain has to state how the environment was made observable. Section 4.1 states it here.

2.4 Grasping from a mobile base in the field

Field mobile manipulation success rates are shaped by base positioning error, target geometry and disturbance. A full-stack deployment in an unmodified grocery store recorded 61% over 335 pick attempts with the entire perception, planning and navigation stack running online [7], and its failure taxonomy is dominated by item properties and by causes upstream of the gripper; that system assumes its navigation places the base within $\pm 40\,\mathrm{cm}$ of the target. An orchard manipulator using onboard perception reported 70% over 56 outdoor leaf-grasping approaches, against 63% over 75 approaches on an indoor plant model [24]. Results of that kind, with the trial count stated, are the correct comparison point for the grasp figure reported in Section 5.4.

2.5 Energy accounting for small platforms

Power modelling for wheeled robots is well developed on the locomotion side, where draw is strongly regime-dependent and can be predicted from physical parameters [25], and energy has long been treated as a hard constraint on deployment planning [26]. More recent work instruments computation and control together and identifies the lack of coordination between them as a principal source of inefficiency [8]. What is

rarely reported for a small outdoor platform is the delta attributable to the perception stack itself, measured at the battery terminals of the same vehicle. Section 5.5 reports it.

2.6 Telemetry in the 433 MHz band

Field measurements of 433 MHz propagation consistently show achieved range far below datasheet line-of-sight claims once vegetation and ground clutter are present [27–29]. That literature motivates a design decision reported in Section 3.4: the halting path is not carried on the radio.

Each subsystem above has been characterised well in isolation. Detection throughput, mapping error, grasp success and power draw for one low-cost outdoor vehicle, measured together, have not.

3 System Description

Figure 1 shows the sensing, computation, actuation, power and safety paths and where the authority boundaries fall. Table 1 lists the specification with the provenance of each entry. Photographs of the vehicle on the test plot appear in Figure 2.

3.1 Chassis and drive train

The chassis is a six-wheel rocker-bogie with a passive differential coupling the left and right rockers, so that chassis pitch follows the mean of the two rocker angles and the payload deck stays closer to horizontal than either side [9]. Wheel radius is 0.0625 m. Measured system mass, weighed with every subsystem installed and the battery fitted, is 12.982 kg. Six geared planetary motors drive the wheels through independent servo drives specified at 30 V and 5 A, with a continuous torque near 1.5 N m each. Structural members are PVC tube; the deck is sheet metal. Figure 2b shows the linkage and wheel mounting.

Two derived quantities set the mobility envelope, and neither is a measurement. The first estimates the force the vehicle must overcome to roll on level grass:

$$F_r = C_{rr}\, m\, g \approx 0.12 \times 12.982 \times 9.81 \approx 15.3 \text{ N}, \tag{1}$$

where F_r is the resistive force (N), C_{rr} is the dimensionless coefficient of rolling resistance, m is the measured system mass (kg) and g is the acceleration due to gravity, taken as $9.81\ \mathrm{m\,s^{-2}}$. The value of C_{rr} was assumed from the terrain class rather than obtained by drawbar pull, so Equation (1) returns a model output and not an observation, and its error is dominated by that assumption rather than by the mass term. The gravitational component along a slope of inclination θ (degrees) is $F_g = mg\sin\theta$, which is 35.1 N at the 16° ceiling reported in Section 5.1.

The second derived quantity converts available motor torque into a tractive force at the wheel-ground contact:

$$F_a = \frac{n\,\tau_c}{r} = \frac{6 \times 1.5}{0.0625} = 144 \text{ N}, \tag{2}$$

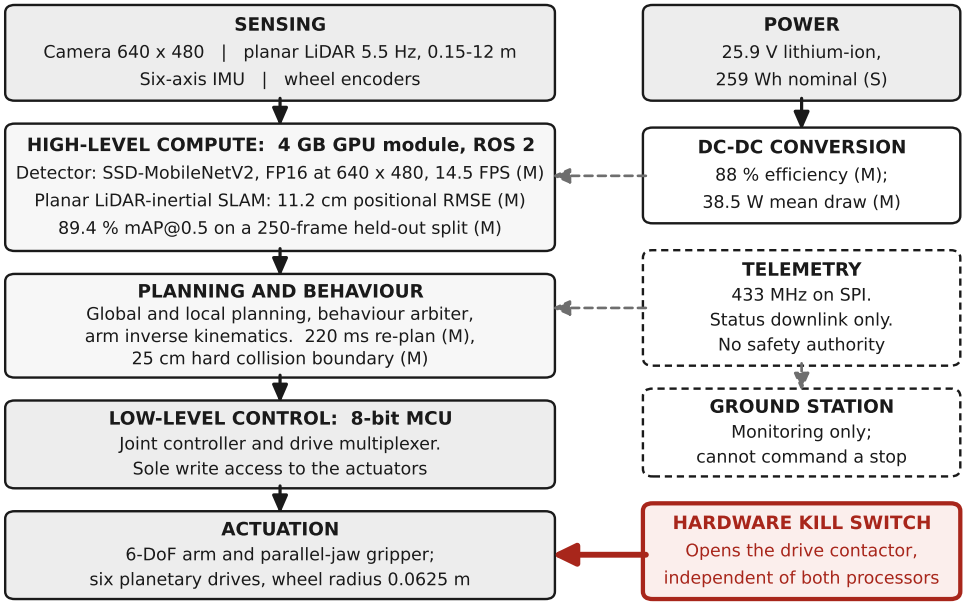


Fig. 1 Sensing, computation, actuation, power and safety paths of the vehicle.
Read the left column downward: every arrow into the actuation block originates at the 8-bit microcontroller unit (MCU), so the 4 GB graphics processing unit (GPU) module never writes to an actuator directly. The right column carries the two paths that matter for safety. Telemetry (dashed) terminates at the ground station and holds no command authority, while the kill path (heavy arrow) reaches the drive contactor without passing through either processor, which is why the 38 ms to 88 ms link latencies of Section 5.6 never sit in a stopping path. Values printed inside the boxes are those reported in Section 5. Acronyms: LiDAR, light detection and ranging; IMU, inertial measurement unit; SLAM, simultaneous localisation and mapping; RMSE, root mean square error; FPS, frames per second; FP16, half precision; mAP@0.5, mean average precision at 0.5 intersection over union; DoF, degree of freedom; SPI, serial peripheral interface.

where F_a is the tractive effort (N), n is the number of driven wheels (dimensionless, here six), τ_c is the continuous torque available at each wheel (N m, specified) and r is the wheel radius (m, measured). Equation (2) is a torque-to-force conversion under a no-slip assumption. It is an upper bound and it was never the binding constraint. The comparison that makes this concrete is with vehicle weight: mg is 127.4 N, so the drive train could in principle develop a tractive coefficient of $F_a/mg \approx 1.13$. A rigid vehicle able to develop a coefficient above unity is not limited by torque at any attainable slope. At the observed 16° ceiling the vehicle was developing

$$\mu_{req} = \frac{F_g + C_{rr}\, m\, g \cos\theta}{m\, g} = \sin\theta + C_{rr}\cos\theta \approx 0.39, \tag{3}$$

where μ_{req} is the dimensionless tractive coefficient demanded at the contact patch. The vehicle stopped while requiring about 0.39 and while its drive train could have supplied 1.13, a factor of about 2.9. What closed that gap was wheel-grass friction and load transfer off the leading wheels, not installed torque, which is the limit that traction control exists to address [11, 12]. Equation (3) is therefore the useful bound and Equation (2) is the one a reader should not quote as a capability.

3.2 Computation and software organisation

Computation splits across two tiers. The GPU module hosts the detector, the SLAM front-end, the global and local planners, the behaviour arbiter and the arm inverse kinematics as ROS 2 nodes communicating over loopback; the middleware distribution is Eloquent Elusor. The MCU owns the real-time path. It multiplexes velocity commands onto the six servo drives and steps the arm joints, and it is the only device on the vehicle with write access to an actuator. The split keeps the actuator loop deterministic while the perception stack runs at whatever rate it achieves, which on grass was not constant. One consequence matters for the timing budget in Section 5.3: a command issued by the GPU module is a request to the MCU, and the MCU applies it at its own rate.

3.3 Perception and the manipulation chain

The detector is an SSD head [2] over a MobileNetV2 backbone [1], trained offboard and exported to FP16 for onboard inference at 640×480. It has a single class, a 12 cm cylindrical sample container. The SLAM front-end fuses planar scans at a specified 5.5 Hz over a specified 0.15–12 m range with IMU attitude and wheel odometry, producing an occupancy grid and a map-frame pose. A detection bounding box is projected onto the ground plane through the fixed camera mounting geometry, which gives the target a position in that same frame. The global planner treats that position as a goal, and the arm treats it as a grasp pose once the base has stopped. The chain is open-loop from the moment the base halts: no visual servoing runs during the final approach, and Section 5.4 shows where that costs trials.

The shoulder joint is the binding actuator. A NEMA 17 stepper rated at 0.44 N m nominal drives a 5:1 planetary reduction, giving an estimated 2.2 N m holding torque. The static demand at full extension follows from two moments about the shoulder axis:

$$\tau_{req} = m_p\, g\, L + m_a\, g\, L_c \approx 1.18 + 0.88 \approx 2.06 \ \text{N m}, \tag{4}$$

where τ_{req} is the static holding torque demanded at the shoulder (N m), m_p is the payload mass (kg, here 0.40), L is the horizontal reach from the shoulder axis to the payload (m, here 0.30), m_a is the distributed mass of the arm structure outboard of the shoulder (kg, here about 0.60) and L_c is the horizontal distance from the shoulder axis to the centre of mass of that structure (m, here about 0.15). Equation (4) is a static estimate under a rigid-link assumption and ignores dynamic loading during the lift. Estimated capacity exceeds estimated demand by about 7%, which is a thin margin. The 0.40 kg figure reported in Table 2 is not this estimate: it is the heaviest

Table 1 Platform specification. Rows are grouped by subsystem in the order the subsystems are introduced in Section 3; within a group, measured entries precede specified entries, and ties are broken by order of mention in the text. Provenance tags: (M) measured on the build, (S) specified by the vendor datasheet, (A) architectural choice carrying no numeric value. Acronyms: PVC, polyvinyl chloride; GPU, graphics processing unit; MCU, microcontroller unit; ROS 2, Robot Operating System 2; LiDAR, light detection and ranging; IMU, inertial measurement unit; DoF, degree of freedom; SPI, serial peripheral interface.

Item	Value	Provenance
System mass, as flown	12.982 kg	(M)
Wheel radius	0.0625 m	(M)
Suspension	six-wheel rocker-bogie, passive differential	(A)
Structure	PVC tube frame, sheet-metal deck	(A)
Drive units	six geared planetary motors on servo drives	(S)
Drive supply limits	30 V, 5 A per drive	(S)
Continuous torque per drive	about 1.5 N m	(S)
High-level compute	4 GB embedded GPU module	(S)
Low-level control	8-bit MCU	(S)
Middleware	ROS 2, Eloquent Elusor	(A)
Scanner	planar LiDAR, 5.5 Hz, 0.15–12 m	(S)
Inertial sensing	six-axis IMU	(S)
Wheel odometry	quadrature encoders on all six drives	(S)
Camera	720p universal serial bus camera, run at 640×480	(S)
Manipulator	6-DoF arm, parallel-jaw gripper	(A)
Shoulder actuator	NEMA 17, 0.44 N m nominal, 5:1 reduction	(S)
Battery pack	25.9 V lithium-ion, 259 Wh nominal	(S)
DC-DC conversion efficiency	88%	(M)
Telemetry	433 MHz transceiver on SPI	(S)
Halting path	hardware kill switch on the drive contactor	(A)

Every entry is a property of the build. No entry in this table is a performance result. Performance results appear in Table 2 and Table 3.

payload carried through repeated handling without joint slip. The two agreeing is a consistency check between a model and an observation, not a second measurement.

3.4 Power and the safety path

A 25.9 V lithium-ion pack of 259 Wh nominal capacity supplies the drives directly and the logic rails through DC-DC conversion measured at 88% efficiency. Telemetry uses a 433 MHz transceiver on a serial peripheral interface (SPI) bus and carries a status downlink only.

The radio holds no safety authority. Emergency halting is a hardware kill switch that opens the drive contactor independently of both processors and of the link; it is the red mushroom head visible on the deck in Figures 2a and 2c. That decision was taken before the link was characterised, and the propagation literature supports it [27, 28]. A stop command that arrives 88 ms late, or that fails to arrive in more than one transmission in nine, is not a stop command. The cost of the decision is an operational range limit, and Section 5.6 states what that limit is.

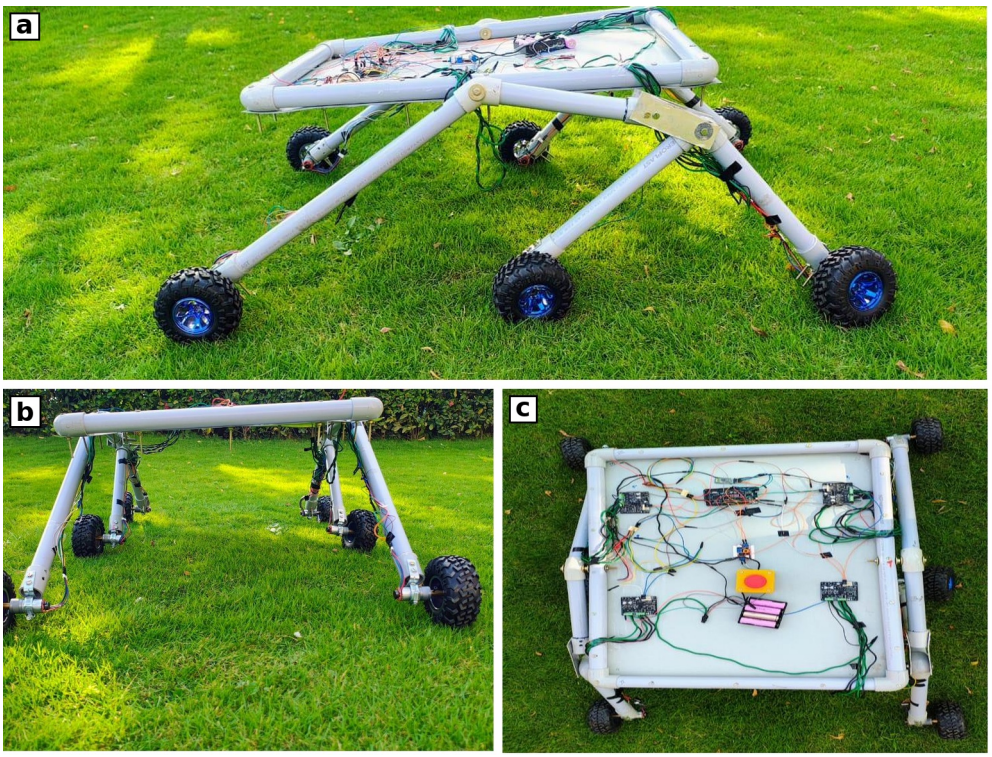


Fig. 2 The vehicle on the grass test plot, photographed under the conditions in which every number in this paper was recorded.
(a) General view. All six wheels carry independent drives, and the rocker and bogie links that keep the payload deck near horizontal are visible along each side. **(b)** The suspension from the side. The ground clearance seen here is what sets the 7.8 cm step-crossing result of Section 5.1. **(c)** The deck from above: servo drives at the corners, the microcontroller near the centre, the battery pack alongside it, and the red mushroom head of the hardware kill switch described in Section 3.4. The wiring visible here is the physical form of the two-tier split drawn in Figure 1.

4 Field Methodology

All trials were conducted outdoors on a level-to-undulating grass plot under daylight, across several sessions. Figure 3 sets out the five stages and the order in which they ran. It also records the central limitation of the campaign, which belongs in the running text and not only in the artwork: one vehicle was built, so every metric carries $N = 1$ at the level of the hardware, and apart from the 30 grasp trials and the 250-frame detector split each figure comes from a single measurement run.

4.1 Site and preparation

The test area was $10\,\mathrm{m} \times 10\,\mathrm{m}$ and its perimeter was deliberately marked with traffic cones and temporary fencing. This choice limits the external validity of the mapping result and a closing caveat is the wrong place for it. A horizontal scan plane over open grass returns almost nothing persistent, and without vertical structure the scan

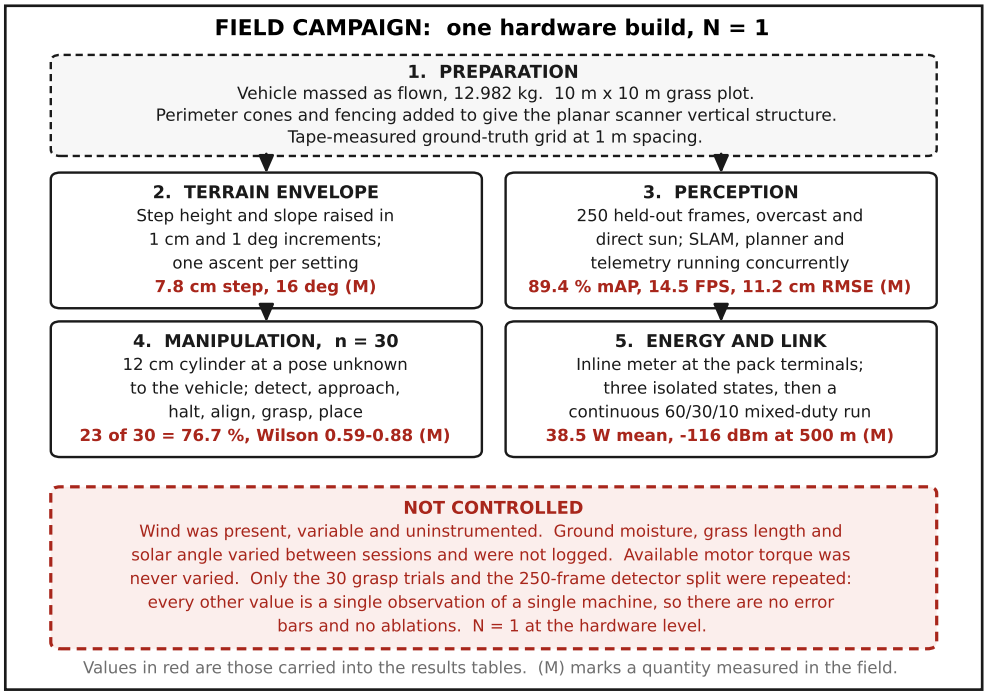


Fig. 3 Trial protocol.
The dashed preparation panel records the artificial vertical structure that makes the 11.2 cm mapping result a bounded-plot result. Stages 2 to 5 ran in the order shown, and the values printed in red at the foot of each stage panel are those carried into Table 2. The dashed closing panel records what was left uncontrolled across all four measurement stages, including the single-build condition $N = 1$ stated in the text of Section 4. (M) marks a quantity measured in the field.

matcher has no geometry to register against, which is the degenerate case described by Zhang et al. [20]. The cones and fencing were introduced to make the environment observable to the sensor. The mapping accuracy reported in Section 5.3 therefore belongs to a bounded plot and does not transfer unmodified to open terrain.

4.2 Ground truth procedure

Ground truth for localisation was a tape-measured grid referenced to the plot corners at 1 m spacing, sketched in Figure 4. The vehicle was commanded to a sequence of grid intersections, halted at each, and its map-frame pose compared against the surveyed coordinate. Positional error was summarised as the root mean square error of the planar offsets over those stops. Tape measurement over a 10 m span carries an error of a centimetre or two. Against an 11.2 cm result that is not dominant, but neither is it negligible, and a total station or real-time kinematic (RTK) reference would be needed before quoting the figure more precisely. Because the comparison was made at halted stops, the result understates error in motion by an amount this campaign did not measure.

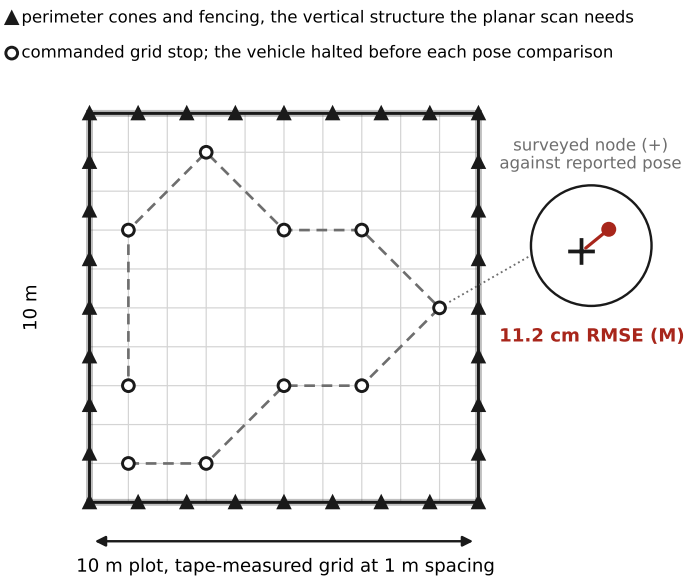


Fig. 4 Geometry of the mapping trial.
The plot is 10 m square with a tape-measured grid at 1 m spacing; triangles mark the perimeter cones and fencing that supply the vertical structure the planar scanner needs, and open circles mark commanded grid stops at which the vehicle halted before its map-frame pose was compared against the surveyed coordinate. The inset at the upper right expands one such comparison: the cross is the surveyed node, the filled marker the reported pose, and the offsets over the whole sequence summarise to the 11.2 cm root mean square error of Section 5.3. Marker positions and the stop sequence drawn here are illustrative of the procedure and are not the surveyed coordinates.

4.3 Trial protocols

Step height and slope were raised in 1 cm and 1° increments, with a single ascent attempted at each setting, until the vehicle stalled or lost six-wheel contact. Neither ceiling was re-measured. Each is a single observation.

The detector was evaluated on a held-out split of 250 annotated frames, disjoint from training, captured on the same terrain across overcast and direct-sun lighting. Frame rate was logged on the vehicle with the SLAM stack, the planner and the telemetry node running, so the throughput reported below is a system rate and not a detector rate.

Manipulation was assessed over 30 autonomous pick-and-place trials. The 12 cm cylinder was placed at a pose unknown to the vehicle within the plot, and Figure 5 shows the arrangement. A trial scored as a success only if the vehicle detected the target, approached, grasped and retained the object through placement without operator intervention. Failures were logged with an attributed cause at the time of the trial. The success rate is reported with its Wilson 95% interval, which is the interval recommended for small samples of this size [30].



Fig. 5 The autonomous manipulation trial arrangement, identical for all 30 trials. **(a)** The 12 cm cylindrical target placed on the plot at a pose unknown to the vehicle, with the six degree-of-freedom (6-DoF) arm on the deck. **(b)** The same arrangement from the approach side. The arm is mounted at a fixed station, so its workspace is fixed relative to the wheels and any error in where the base comes to rest has to be absorbed by that workspace; failures clustered where that error exceeded ±2.5 cm, as reported in Section 5.4. **(c)** The parallel-jaw gripper and wrist at closer range. The jaws are driven through a single geared pair from one servo, which is why grip force is not independently commanded and why compliance in the jaws, measured as 5.1 mm of deflection at 400 g, is a property of the printed linkage rather than of a controller.

Power was logged with an inline meter at the battery terminals, first in three isolated states and then over a continuous mixed-duty run. Link quality was swept at 25, 100, 300 and 500 m.

4.4 What was not controlled

Wind was present, variable and uninstrumented. Ground moisture, grass length and solar angle varied between sessions and were not logged. Available motor torque was never varied, so the reading of the gradient ceiling in Section 5.1 is an inference from an observed failure mode rather than the result of an experiment that manipulated torque. Only the grasp trials and the detector split were repeated. There are no error bars and no ablations, and where a trial count permits an interval the interval is given; elsewhere the figures are one observation of one machine.

No configuration of the vehicle other than the one described in Section 3 was tested. The perception nodes were never run with any part of the stack disabled, so nothing in this campaign isolates the contribution of an individual subsystem to any measured quantity.

5 Results and Discussion

5.1 Terrain envelope

The vehicle cleared 7.8 cm steps and held ascent on grass to 16° under the increment procedure of Section 4.3. Each ceiling is one observation of one machine: the step was raised until the leading wheels no longer lifted onto it, and the slope until the vehicle stalled, with a single ascent attempted at each setting and neither ceiling re-measured. The ground clearance that sets the first of those numbers is visible in Figure 2b.

Ascent ended in wheel slip and in loss of load on the leading wheels. Geometry was not the ceiling, and by the argument of Section 3.1 neither was motor torque:

Table 2 Measured and derived performance of the vehicle. Rows run in subsystem order, mobility, perception and planning, manipulation, then energy and link, and within a subsystem in the order the quantity is introduced in the text; where a derived row is computed from a measured row, the measured row precedes it. Provenance tags: (M) measured in the field, (D) derived through a stated model, (A) an architectural choice adopted at design time. No cell is blank. Acronyms: RMSE, root mean square error; mAP@0.5, mean average precision at 0.5 intersection over union; FPS, frames per second; RSSI, received signal strength indication.

Quantity	Value	Provenance
Maximum obstacle clearance	7.8 cm	(M)
Maximum traverse gradient on grass	16°	(M)
Tractive coefficient demanded at that gradient	0.39	(D)
Tractive coefficient available from drive torque	1.13	(D)
SLAM positional RMSE, bounded plot	11.2 cm	(M)
Local re-planning latency	220 ms	(M)
Hard collision boundary	25 cm	(A)
Speed ceiling holding that boundary	0.34 m s^{-1}	(D)
Detection mAP@0.5, 250-frame held-out split	89.4%	(M)
Vision frame rate under concurrent load	14.5 FPS	(M)
Autonomous grasp success, $n = 30$	23 of 30, 76.7%	(M)
Wilson 95% interval on that rate	0.59–0.88	(D)
Manipulator stable load	0.40 kg	(M)
End-effector deflection at 400 g	5.1 mm	(M)
Mean terminal power, mixed duty	38.5 W	(M)
Continuous endurance	about 5.3 h	(D)
RSSI at 500 m	−116 dBm	(M)
Packet loss at 500 m	11.5%	(M)

Every value comes from the single hardware build described in Section 3. Apart from the grasp trials and the detector split, each is a single observation, so no row carries a dispersion estimate.

the drive train can develop a tractive coefficient near 1.13 and the vehicle stopped while demanding about 0.39. What bound the climb was friction and weight distribution. Because available torque was never varied, this reading is an inference from the observed failure mode rather than a controlled result, and it is consistent with field comparisons in which ground interaction rather than installed torque discriminates between platforms [13, 14]. The wheels were still turning when the vehicle stopped.

5.2 Onboard detection

The detector scored 89.4% mAP at 0.5 intersection over union (IoU) on the 250-frame held-out split while the vehicle sustained 14.5 FPS with the full stack running. Both numbers come from the same sessions on the same ground, which is the point of reporting them together, and both describe one detector on one target class.

Read the accuracy narrowly. The split is disjoint from training but was captured on the same plot, on the same object, with the same camera, so 89.4% is an in-domain figure and carries no claim about a second object or a second site. It is also a single split rather than a cross-validated mean, so it has no dispersion estimate attached to it. What the number does support is the claim the paper needs: on this

compute a quantised MobileNetV2-SSD held its accuracy with the mapping, planning and telemetry nodes resident alongside it, across lighting that ranged from overcast to direct sun.

Published throughput for this class of module is obtained without that concurrent load [3, 15, 16], so 14.5 FPS is not comparable with those figures either. Treat it instead as a design input. It fixes a 69 ms detection interval, which sits inside the 182 ms scan period, so the detector keeps up with the scanner and does not itself lengthen the obstacle-to-command chain. That containment is the assumption the speed ceiling of Section 5.3 rests on.

5.3 Mapping and the local planner

Positional RMSE against the tape-measured grid was 11.2 cm, and the local planner produced a revised trajectory 220 ms after an unmapped obstacle entered the scan. The hard collision boundary was 25 cm. Those three numbers interlock, and Figure 6 shows how. Worst-case delay from obstacle appearance to command change is the 182 ms scan period plus 220 ms of planning, about 402 ms. Subtracting the localisation error from the boundary leaves 13.8 cm of travel over that interval, which gives a derived speed ceiling

$$v_{max} = \frac{d_b - e_{RMSE}}{t_{scan} + t_{plan}} = \frac{0.250 - 0.112}{0.182 + 0.220} \approx 0.34 \ \mathrm{m\,s^{-1}}, \tag{5}$$

where v_{max} is the ceiling on ground speed ($\mathrm{m\,s^{-1}}$), d_b is the hard collision boundary (m, configured), e_{RMSE} is the positional RMSE (m, measured), t_{scan} is the scanner period (s, specified as the reciprocal of 5.5 Hz) and t_{plan} is the measured local re-planning latency (s). Equation (5) is a derived bound, not a measured maximum speed, and it is optimistic: the 11.2 cm term was obtained at halted grid stops and understates error in motion, and the detection interval of 69 ms sits inside the scan period only while the stack keeps up. The budget also excludes the interval between the planner issuing a command and the MCU applying it, which Section 3.2 describes and this campaign did not instrument, so 402 ms is a floor on the delay rather than a ceiling.

The second qualification is the one raised in Section 4.1. Removing the cones and fencing removes most of the scan-matching constraint, and no measurement was taken in that condition. The 11.2 cm figure is bounded-plot accuracy, obtained where artificial vertical structure made the environment observable, and it says nothing about open grass [20, 21].

5.4 Autonomous manipulation

Twenty-three of 30 autonomous trials succeeded. The rate is 76.7% with a Wilson 95% interval of 0.59 to 0.88 [30]. That interval spans nearly thirty points, so the point estimate should never be quoted alone, and 30 trials cannot separate the failure modes statistically. It does not support a claim of reliable autonomous grasping; it supports a claim that a stack of this kind grasps a known single-class target outdoors more often than not.

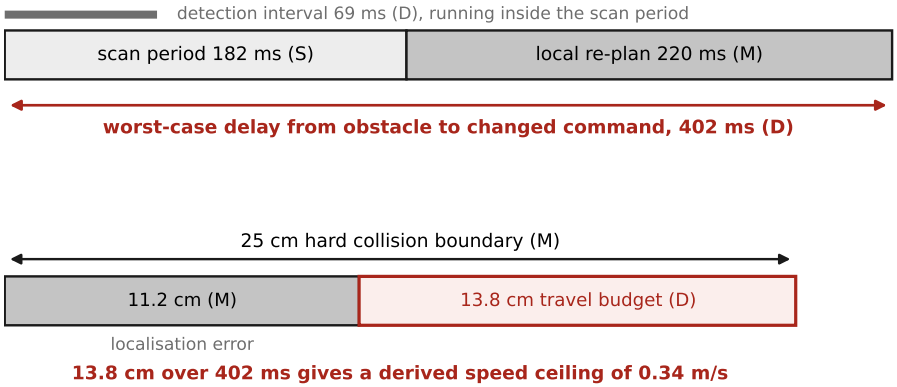


Fig. 6 How the three interlocking measurements of Section 5.3 produce a speed ceiling. Above: the worst-case delay from an obstacle entering the scan to a changed command is one scan period plus one re-planning latency, 402 ms in total; the 69 ms detection interval fits inside the scan period and does not add to the chain. Below: of the 25 cm hard collision boundary, 11.2 cm is consumed by localisation error, leaving 13.8 cm of travel budget. The quotient of the shaded budget and the total delay is the 0.34 $\mathrm{m\,s^{-1}}$ bound of Equation (5), which is a ceiling on commanded speed and not a measured maximum. (M) measured, (D) derived, (S) specified.

The failure log does show a pattern. Unsuccessful attempts clustered where base alignment error exceeded ±2.5 cm at the moment the vehicle halted, and during gusts. Both causes sit upstream of the gripper, which is where the arrangement in Figure 5 would predict them: the arm is mounted at a fixed station, so its workspace is fixed relative to the wheels, and whatever the base gets wrong the arm must absorb without any closed-loop correction during approach.

Two published field results locate the figure. A full-stack deployment picking diverse items in an unmodified grocery store recorded 61% over 335 attempts [7], with failures likewise concentrated upstream of the gripper, and an orchard manipulator reported 70% over 56 outdoor approaches [24]. Neither comparison is controlled. A single-class cylinder on open grass is a far easier perception problem than a grocery shelf or a leaf in a canopy, so the higher rate here reflects task difficulty at least as much as system quality. The contrast that is informative runs the other way: the grocery-store system tolerates a base positioning error of ±40 cm because its arm reach and its grasp planner absorb that error, while this vehicle fails beyond ±2.5 cm. The difference is in the arm, not in the navigation.

Structural compliance was not the limiting factor. End-effector deflection at 400 g was 5.1 mm, roughly 4% of the 12 cm target diameter and small against the alignment tolerance that failures actually violated. The chain fails where detection projects a target position, the base stops with finite positioning error, and the fixed workspace absorbs the remainder. More of the seven failures were logged against base alignment than against gusts. That is a raw count and not a ranking: with wind uninstrumented and $n = 30$, these data cannot separate the two causes, and Section 6.5 states why.

5.5 Energy cost of autonomy

Three isolated states were logged at the terminals: sensing-only idle at 12.2 W, nominal locomotion at 28.5 W and peak manipulation at 14.8 W, the arm working only while the base is halted. A continuous mixed-duty run apportioning 60% locomotion, 30% idle and 10% manipulation returned a mean of 38.5 W over the run. Weighting the three isolated readings by that same duty split predicts 22.2 W. The measured mean exceeds the prediction by 16.3 W, which is more than the entire idle draw, and Figure 7 shows the size of the gap against the states that were supposed to explain it.

That gap belongs to the model rather than to the instrument. Isolated readings were taken with the other subsystems quiescent, whereas under mixed duty the perception stack, the drive controllers and the arm hold current concurrently, and motor start-stop transients enter the average. Additive duty weighting assumes the states are separable and on this vehicle they are not, which is the coordination effect that instrumented studies of computation and control also report [8], and is consistent with power models in which locomotion draw is strongly regime-dependent [25]. The mixed-duty figure is used for endurance because it is the only one measured under the load the vehicle actually sees.

Usable energy is derived from the nominal pack capacity, the permitted depth of discharge (DoD) and the measured conversion efficiency as

$$E_u = E_n \cdot \mathrm{DoD} \cdot \eta = 259 \times 0.90 \times 0.88 \approx 205 \ \mathrm{Wh}, \tag{6}$$

where E_u is the usable energy at the load (Wh), E_n is the nominal pack energy (Wh, specified), DoD is the permitted depth of discharge expressed as a fraction of nominal capacity (dimensionless, assumed 0.90 to protect cycle life) and η is the measured DC-DC conversion efficiency (dimensionless, 0.88). Continuous endurance follows as

$$T_e = \frac{E_u}{\bar{P}} = \frac{205}{38.5} \approx 5.3 \ \mathrm{h}, \tag{7}$$

where T_e is the derived continuous endurance (h) and $\bar{P}$ is the measured mean terminal power under mixed duty (W).

That figure was not verified by running a pack to cutoff. It is a quotient of Equation (6) and a measured mean draw, and it inherits whatever error sits in the depth-of-discharge assumption, which is the only input to Equation (6) that was adopted rather than measured or specified. The right panel of Figure 7 shows how far that matters. Sweeping DoD from 0.70 to 1.00 moves the derived endurance from 4.1 h to 5.9 h, so the assumption is worth about 1.8 h of quoted runtime and a reader who disagrees with 0.90 can read off their own figure. The measured input, the 38.5 W, is the part that is not in dispute.

Two consequences follow for a design team. The first is that a power budget assembled from isolated subsystem measurements will underestimate, on this vehicle by 16.3 W, and the error is larger than the entire idle state. The second is that energy, not accuracy, is what constrains how long this autonomy can run [26]: at 38.5 W the vehicle has about five hours, and every subsystem that shares the pack spends part of it.

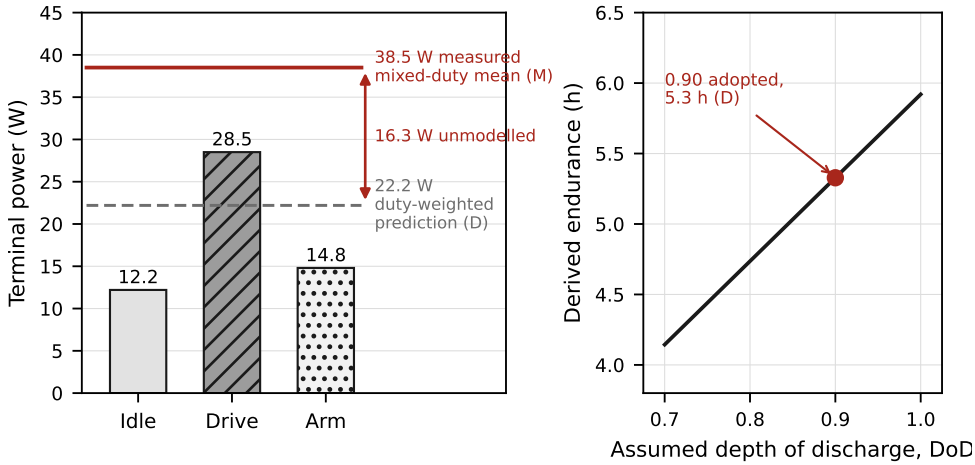


Fig. 7 Measured power states at the battery terminals and the sensitivity of the derived endurance. Left: the three isolated power states measured at the battery terminals, plotted against the duty-weighted prediction they generate (dashed) and the mean actually measured under a continuous mixed-duty run (solid). Idle is the sensing-only state, Drive is nominal locomotion, Arm is peak manipulation with the base halted. The 16.3 W between the two lines is the part of the draw that additive duty weighting does not capture, and it exceeds the whole idle state. Right: how the derived endurance of Equation (7) depends on the one input to Equation (6) that was assumed rather than measured. The marked point is the 0.90 depth of discharge (DoD) adopted here; across the plotted range the quoted endurance moves by about 1.8 h. No pack was run to cutoff. (M) measured, (D) derived.

5.6 Telemetry and the halting path

Link quality degraded monotonically with range over a single sweep, and Figure 8 plots the three quantities together. Received signal strength indication (RSSI) fell from −78 dBm at 25 m to −116 dBm at 500 m. Packet loss rose from no measurable loss at 25 m, through 1.8% at 100 m and 5.4% at 300 m, to 11.5% at 500 m. Round-trip latency grew from 38 ms to 88 ms across the same sweep, which is consistent with retransmission under an automatic repeat request (ARQ) scheme rather than with propagation delay: 50 ms of extra delay over 475 m of extra path cannot be time of flight.

Characterisation stopped at 500 m against a vendor line-of-sight claim of a different order, and terrain profile, antenna height and orientation were not recorded. These figures describe one deployment geometry and one sweep, which is the same qualification the published propagation measurements attach to theirs [27–29].

At 500 m the link drops more than one packet in nine. That is adequate for an advisory status stream and unacceptable for anything else, which is why the halting path in Section 3.4 bypasses the radio and both processors. The cost of that architecture is an operational range limit set by line of sight to the operator rather than by the link: the vehicle cannot safely be commanded beyond the distance at which a person can reach the switch, which is far shorter than 500 m.

Table 3 Energy accounting for the vehicle. Rows follow the order of the derivation in Section 5.5: the three isolated states first, then the duty-weighted prediction they generate, then the measured mixed-duty mean and the endurance derived from it. Where a stage carries both a measured and a derived row, the measured row precedes. Provenance tags: (M) measured at the battery terminals, (D) derived through a stated model, (S) specified by the vendor datasheet, (P) a parameter adopted rather than measured or specified. The duty weights are the proportions of run time apportioned to each state during the continuous mixed-duty run and are dimensionless; the dash in the weight column marks rows to which no duty weight applies.

Quantity	Value	Duty weight	Provenance
Sensing-only idle draw	12.2 W	0.30	(M)
Nominal locomotion draw	28.5 W	0.60	(M)
Peak manipulation draw	14.8 W	0.10	(M)
Duty-weighted prediction	22.2 W	1.00	(D)
Mixed-duty mean, measured	38.5 W	1.00	(M)
Excess over prediction	16.3 W	–	(D)
Nominal pack energy	259 Wh	–	(S)
Permitted depth of discharge	0.90	–	(P)
DC-DC conversion efficiency	0.88	–	(M)
Usable energy	about 205 Wh	–	(D)
Continuous endurance	about 5.3 h	–	(D)

The depth of discharge carries the (P) tag because 0.90 was adopted to protect cycle life rather than measured or specified; it is the only such input in this table. The endurance row is a quotient of the usable energy and the measured mixed-duty mean; no pack was discharged to cutoff.

6 Limitations

The limitations below are constraints on what the reported data can support. Each is stated as a boundary on inference rather than as a shortfall to be corrected later.

6.1 Single build and single run

One vehicle was built. Every number in Table 2 has $N = 1$ at the level of the hardware, and only the 30 grasp trials and the 250-frame detector split were repeated at the level of the trial. No metric has an error bar because no metric has a variance to report. The practical consequence is specific: build-to-build variation in manufacture, in wiring loss, in wheel-tyre condition and in gearbox break-in is entirely absent from these figures, and any of those could move a quantity such as the 38.5 W mean by an amount this campaign cannot bound. Nothing here separates a property of the design from a property of this particular assembly.

6.2 No ablation of the autonomy stack

Every measurement was taken with the whole stack resident. The detector, the SLAM front-end, the planners and the telemetry node were never switched off individually, and the vehicle was never run in a perception-free configuration. Three quantities inherit that directly. The 14.5 FPS is a system rate whose shortfall against an isolated

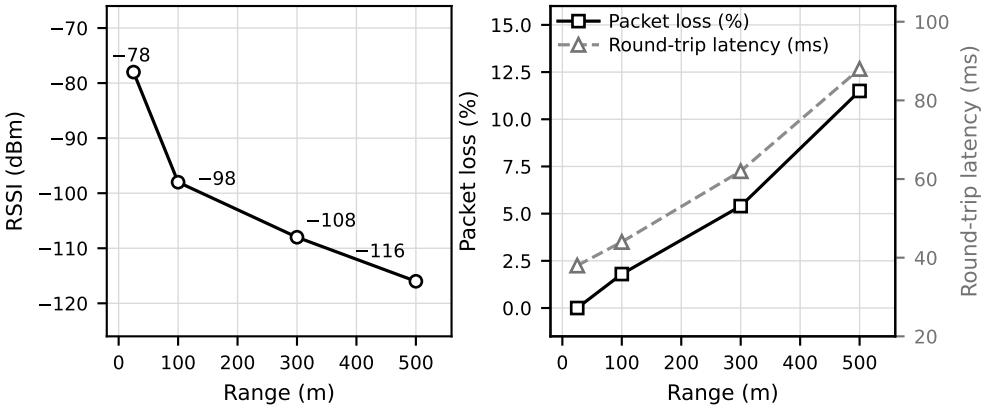


Fig. 8 Measured link quality against range over a single sweep at 25, 100, 300 and 500 m. Left: received signal strength indication (RSSI), which falls by 38 dB across the sweep. Right: packet loss on the left axis and round-trip latency on the right. The two rise together, and the latency growth is attributed to automatic repeat request (ARQ) retransmission rather than to propagation delay, because the extra path length accounts for a negligible fraction of the extra 50 ms. All four points are single measurements taken in one deployment geometry with antenna height and orientation unrecorded.

benchmark cannot be apportioned between the SLAM front-end, the planner and the middleware. The 220 ms re-planning latency is measured with the detector competing for the same processor. The 38.5 W mean contains locomotion, sensing, inference and arm current in proportions this campaign did not separate, which is precisely why the duty-weighted decomposition in Section 5.5 fails by 16.3 W. Nothing here supports a statement of the form “perception costs x watts” or “mapping costs y milliseconds” on this vehicle. Recovering those would need the same build measured with named nodes disabled, under the same duty profile, and that experiment was not performed.

6.3 Scope of the detection result

The 89.4% describes one architecture, one training set, one target class and one site. The held-out split is disjoint from training but not independent of it in any stronger sense: the frames were captured on the same plot, under the same mounting geometry, of the same 12 cm cylinder. A second object, a second background or a second camera height would each require its own measurement, and none was taken. The figure is also a single split rather than a cross-validated mean, so it carries no dispersion estimate and the paper reports none. Against the published benchmarks for this class of module [3, 15, 16] the comparison runs only one way: those figures are obtained with the detector as the sole resident workload, so they bound what this vehicle could achieve without concurrency, and 14.5 FPS says nothing about how much of the shortfall belongs to which competing node.

6.4 The mapping result belongs to a bounded plot

The perimeter cones and fencing were introduced to give the planar scanner vertical structure it would not otherwise have found on open grass. The 11.2 cm RMSE is an accuracy obtained under artificial observability, and the honest reading is that a horizontal planar scan is the wrong sensor for open terrain rather than that this stack maps open terrain to a decimetre [20]. No measurement was taken with the markers removed, so the size of the degradation is unknown and cannot be estimated from these data. A related gap sits in the ground truth itself: tape measurement over 10 m carries a centimetre-scale error that is small against 11.2 cm but not negligible against it, and the comparison was made at halted stops, so error in motion is unmeasured and the speed ceiling of Equation (5) inherits that optimism.

6.5 Sample size in the manipulation trials

Thirty trials give a Wilson 95% interval of 0.59 to 0.88 around the 76.7% point estimate [30]. Any comparison against a published rate that falls inside that band is uninformative, which includes both field results cited in Section 5.4. The failure taxonomy is subject to the same limit. Two causes were logged, base alignment error beyond ± 2.5 cm and gust disturbance, and with seven failures in total no ranking between them survives contact with the interval. Separating wind-induced from alignment-induced failure would need on the order of a hundred trials with wind instrumented, and the present data cannot substitute for that.

6.6 Derived quantities inherit unmeasured assumptions

Four equations in this paper return model outputs whose inputs were not all measured. The rolling resistance coefficient in Equation (1) was assumed from terrain class rather than obtained by drawbar pull, so the 15.3 N resistive force and the 0.39 tractive coefficient of Equation (3) both carry that assumption. The tractive effort of Equation (2) rests on a datasheet torque and a no-slip contact, neither verified on the vehicle. The shoulder demand of Equation (4) is static and ignores dynamic loading, which is the loading a 7% margin is least able to absorb. The usable energy of Equation (6) rests on a depth of discharge that was adopted rather than measured. None of these five appears in this paper as an experimental result, and none should be quoted as one.

6.7 Scope of the link characterisation

The link was swept once, to 500 m, in one geometry, with terrain profile, antenna height and antenna orientation unrecorded. Each of those is known from the propagation literature to move achieved range substantially [27, 29], so the four points describe a deployment rather than a device. The safety argument built on them is robust to that weakness in one direction only: the measurements are sufficient to show that the link is unfit to carry a stop command, and insufficient to specify a usable command range for anything else.

7 Conclusion

A 12.982 kg COTS rocker-bogie rover ran its full autonomy loop onboard a single low-power GPU module and was characterised outdoors on grass. Detection reached 89.4% mAP@0.5 on a 250-frame held-out split at 14.5 FPS with the rest of the stack running concurrently. Planar LiDAR-inertial SLAM returned 11.2 cm positional RMSE in a plot marked with artificial vertical structure, and the local planner replied 220 ms after an obstacle entered the scan, which together bound safe ground speed near 0.34 m s^{-1}. Twenty-three of 30 autonomous grasps succeeded, 76.7% with a Wilson 95% interval of 0.59 to 0.88, with failures clustered where base alignment exceeded ±2.5 cm. Mean terminal draw under mixed duty was 38.5 W, giving about 5.3 h of derived endurance. The measurement that travels furthest beyond this vehicle is the one that failed. Weighting three isolated power states by the duty split they were meant to describe predicted 22.2 W; the vehicle drew 38.5 W under that split, and the 16.3 W difference is larger than the whole sensing-only idle state. A budget assembled the way most small-platform budgets are assembled, subsystem by subsystem, would have been wrong by 42% of the measured mean on this build. The mobility argument points the same way: a drive train able to develop a tractive coefficient of 1.13 stopped while demanding 0.39, so the datasheet torque never described the limit either. Two of the numbers a designer would most likely take from a specification sheet, power and traction, were the two the field measurement contradicted. All results reported here are single-build, largely single-run, and are reported as such.

Supplementary information. Not applicable.

Acknowledgements. Not applicable.

Declarations

- **Funding.** No funding was received to assist with the preparation of this manuscript.
- **Conflict of interest.** The authors declare that they have no competing interests.
- **Ethics approval and consent to participate.** Not applicable. This study involved no human participants and no animals.
- **Consent for publication.** All authors have read the manuscript and consent to its publication. The photographs reproduced in Figure 2 and Figure 5 were taken by the authors, show equipment built by the authors, and contain no identifiable individuals.
- **Data availability.** The data supporting the findings of this study are available from the corresponding author on reasonable request. They comprise the 250-frame annotated detector evaluation split, the grid-stop pose log underlying the 11.2 cm root mean square error, the 30-trial grasp log with attributed failure causes, the terminal power logs for the three isolated states and the mixed-duty run, and the four-point link sweep.
- **Materials availability.** Not applicable. The vehicle is a one-off build assembled from commercially available parts, listed in Table 1.

- **Code availability.** The launch configuration, the detector export settings and the analysis scripts used to produce the reported figures are available from the corresponding author on reasonable request.
- **Author contribution.** Shahan Sahay and Ravi Tiwari conceived the overall research idea. Shahan Sahay and Harsh Yadav conducted the field experiments and generated the data. Avishkar Jaiswal contributed to data management, validation and analysis, while Ravi Tiwari provided guidance on the research concepts and methodology. Shahan Sahay drafted the manuscript, while Harsh Yadav and Avishkar Jaiswal contributed to the development and refinement of the methodology. Ravi Tiwari critically reviewed and revised the manuscript. All authors reviewed and approved the final version of the manuscript.

References

[1] Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., Chen, L.-C.: MobileNetV2: inverted residuals and linear bottlenecks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 4510–4520 (2018)

[2] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y., Berg, A.C.: SSD: single shot multibox detector. In: Proceedings of the European Conference on Computer Vision (ECCV), pp. 21–37 (2016)

[3] Mittal, S.: A survey on optimized implementation of deep learning models on the NVIDIA Jetson platform. J. Syst. Archit. **97**, 428–442 (2019)

[4] Grisetti, G., Stachniss, C., Burgard, W.: Improved techniques for grid mapping with Rao-Blackwellized particle filters. IEEE Trans. Robot. **23**(1), 34–46 (2007)

[5] Hess, W., Kohler, D., Rapp, H., Andor, D.: Real-time loop closure in 2D LIDAR SLAM. In: Proceedings of the IEEE International Conference on Robotics and Automation (ICRA), pp. 1271–1278 (2016)

[6] Macenski, S., Foote, T., Gerkey, B., Lalancette, C., Woodall, W.: Robot Operating System 2: design, architecture, and uses in the wild. Sci. Robot. **7**(66), eabm6074 (2022)

[7] Huang, I., Cheng, R., Kim, S., Kruse, D., Matl, C., Kaul, L., Hancock, J.C., Harikumar, S., Tjersland, M., Borders, J., Helmick, D.: Practical insights on grasp strategies for mobile manipulation in the wild. arXiv:2504.12512 (2025)

[8] Liu, L., Zhong, R., Willcock, A., Fisher, N., Shi, W.: An open approach to energy-efficient autonomous mobile robots. In: Proceedings of the IEEE International Conference on Robotics and Automation (ICRA), pp. 11569–11575 (2023)

[9] Harrington, B.D., Voorhees, C.: The challenges of designing the rocker-bogie suspension for the Mars Exploration Rover. In: Proceedings of the 37th Aerospace Mechanisms Symposium (AMS) (2004)

[10] Thüer, T.: Mobility evaluation of wheeled all-terrain robots: metrics and application. PhD thesis, ETH Zürich (2009)

[11] Iagnemma, K., Dubowsky, S.: Traction control of wheeled robotic vehicles in rough terrain with application to planetary rovers. Int. J. Robot. Res. **23**(10–11), 1029–1040 (2004)

[12] Ishigami, G., Miwa, A., Nagatani, K., Yoshida, K.: Terramechanics-based model for steering maneuver of planetary exploration rovers on loose soil. J. Field Robot. **24**(3), 233–250 (2007)

[13] Wirkus, M., Hinck, S., Backe, C., Babel, J., Riedel, V., Reichert, N., Kolesnikov, A., Stark, T., Hilljegerdes, J., Küçüker, H.D., Ruckelshausen, A., Kirchner, F.: Comparative study of soil interaction and driving characteristics of different agricultural and space robots in an agricultural environment. J. Field Robot. **41**(6), 2009–2042 (2024)

[14] Gerdes, L., Pérez del Pulgar, C., Castilla Arquillo, R., Azkarate, M.: Field assessment of force torque sensors for planetary rover navigation. J. Intell. Robot. Syst. **111**, 122 (2025)

[15] Swaminathan, T.P., Silver, C., Akilan, T.: Benchmarking deep learning models on NVIDIA Jetson Nano for real-time systems: an empirical investigation. arXiv:2406.17749 (2024)

[16] Alqahtani, D.K., Cheema, M.A., Rodriguez, M.A., Toosi, A.N.: A comprehensive evaluation of deep learning object detection models on heterogeneous edge devices. arXiv:2409.16808 (2026)

[17] Kohlbrecher, S., von Stryk, O., Meyer, J., Klingauf, U.: A flexible and scalable SLAM system with full 3D motion estimation. In: Proceedings of the IEEE International Symposium on Safety, Security and Rescue Robotics (SSRR), pp. 155–160 (2011)

[18] Konolige, K., Grisetti, G., Kümmerle, R., Burgard, W., Limketkai, B., Vincent, R.: Efficient sparse pose adjustment for 2D mapping. In: Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 22–29 (2010)

[19] Macenski, S., Jambrecic, I.: SLAM Toolbox: SLAM for the dynamic world. J. Open Source Softw. **6**(61), 2783 (2021)

[20] Zhang, J., Kaess, M., Singh, S.: On degeneracy of optimization-based state estimation problems. In: Proceedings of the IEEE International Conference on Robotics and Automation (ICRA), pp. 809–816 (2016)

[21] Zhen, W., Zeng, S., Scherer, S.: Robust localization and localizability estimation with a rotating laser scanner. In: Proceedings of the IEEE International Conference on Robotics and Automation (ICRA), pp. 6240–6245 (2017)

[22] Tuna, T., Nubert, J., Nava, Y., Khattak, S., Hutter, M.: X-ICP: localizability-aware LiDAR registration for robust localization in extreme environments. IEEE Trans. Robot. **40**, 452–471 (2024)

[23] Nubert, J., Walther, E., Khattak, S., Hutter, M.: Learning-based localizability estimation for robust LiDAR localization. In: Proceedings of the IEEE/RSJ

International Conference on Intelligent Robots and Systems (IROS), pp. 17–24 (2022)

[24] Mortazavi, M., Cappelleri, D.J., Ehsani, R.: RoMu4o: a robotic manipulation unit for orchard operations automating proximal hyperspectral leaf sensing. arXiv:2501.10621 (2025)

[25] Dogru, S., Marques, L.: A physics-based power model for skid-steered wheeled mobile robots. IEEE Trans. Robot. **34**(2), 421–433 (2018)

[26] Mei, Y., Lu, Y.-H., Hu, Y.C., Lee, C.S.G.: Deployment of mobile robots with energy and timing constraints. IEEE Trans. Robot. **22**(3), 507–522 (2006)

[27] Ahmad, K.A., Salleh, M.S., Segaran, J.D., Hashim, F.R.: Impact of foliage on LoRa 433MHz propagation in tropical environment. AIP Conf. Proc. **1930**, 020009 (2018)

[28] Petäjäjärvi, J., Mikhaylov, K., Roivainen, A., Hänninen, T., Pettissalo, M.: On the coverage of LPWANs: range evaluation and channel attenuation model for LoRa technology. In: Proceedings of the 14th International Conference on ITS Telecommunications (ITST), pp. 55–59 (2015)

[29] Xu, T., Ma, D., Fang, W., Huang, Y.: Experimental study on the propagation characteristics of LoRa signals in maize fields. Electronics **14**(11), 2156 (2025)

[30] Brown, L.D., Cai, T.T., DasGupta, A.: Interval estimation for a binomial proportion. Stat. Sci. **16**(2), 101–133 (2001)