Designation: E2991/E2991M − 17

Standard Test Method for Evaluating Response Robot Mobility: Traverse Gravel Terrain[1]

This standard is issued under the fixed designation E2991/E2991M; the number immediately following the designation indicates the year of original adoption or, in the case of revision, the year of last revision. A number in parentheses indicates the year of last reapproval. A superscript epsilon (ε) indicates an editorial change since the last revision or reapproval.

INTRODUCTION

The robotics community needs ways to measure whether a particular robot is capable of performing specific missions in unstructured and often hazardous environments. These missions decompose into elemental robot tasks represented individually as standard test methods and practices. The associated test apparatuses and performance metrics provide a tangible language to communicate varying mission requirements. They also enable repeatable testing to establish the reliability of essential robot capabilities.

The ASTM International Standards Committee on Homeland Security Applications (E54) specifies standard test methods and practices for evaluating such robot capabilities. These standards facilitate comparisons across diverse models or multiple configurations of a single model. They support robot researchers, manufacturers, and user organizations in different ways. Researchers use the standards to understand mission requirements, encourage innovation, and demonstrate break-through capabilities. Manufacturers use the standards to evaluate design decisions, integrate emerging technologies, and harden developed systems. User organizations leverage the resulting robot capabilities data to guide purchasing decisions, align deployment objectives, and focus training with standard measures of operator proficiency. Associated usage guides describe how such standards can be applied to support these various objectives.

The overall suite of standards addresses critical subsystems of remotely operated response robots, including maneuvering, mobility, dexterity, sensing, energy, communications, durability, proficiency, autonomy, logistics, safety, and terminology. This test method addresses the robotic mobility on gravel terrain.

1. Scope

1.1 The purpose of this test method is to specify the apparatuses, procedures, and performance metrics necessary to quantitatively measure a teleoperated ground robot's capability of traversing gravel terrain. The primary performance metric for this test method shall be a robot's possession of such a capability with a specified statistical significance level.

1.2 Average rate of advance over the specified terrain shall be the secondary performance metric for this test method. The measure shall be calculated only when a robot under test has completed a statistically-significant number of repetitions.

1.3 This test method can also be used to measure the operator proficiency in performing the specified task. The corresponding performance metric may be the number of completed task repetitions per minute over an assigned time period ranging from 10 to 30 minutes.

1.4 This test method is a part of the mobility suite of ground response robot test methods, but this test method is stand-alone and complete. This test method applies to ground systems operated remotely from a standoff distance appropriate for the intended mission. The system includes a remote operator in control of all functionality and any assistive features or autonomous behaviors that improve the effectiveness or efficiency of the overall system.

1.5 The apparatus, specified in Section 6, can only test a limited range of a robot's capabilities. When the robot has been tested through the limit or limits of the apparatus, a note shall be associated with the results indicating that the robot's actual capability may be outside of the limit or limits imposed by the test apparatus. For example, the size of the gravel terrain test apparatus could possibly affect the acceleration of the robot under test and, in turn, the resulting average rate of advance.

[1] This test method is under the jurisdiction of ASTM Committee E54 on Homeland Security Applications and is the direct responsibility of Subcommittee E54.09 on Response Robots.

Current edition approved Sept. 1, 2017. Published October 2017. DOI: 10.1520/E2991_E2991M-17.

1.6 *Performing Location*—This test method may be performed anywhere the specified apparatuses and environmental conditions can be implemented.

1.7 *Units*—The values stated in either SI units or inch-pound units are to be regarded separately as standard. The values stated in each system may not be exact equivalents; therefore, each system shall be used independently of the other. Combining values from the two systems may result in nonconformance with the standard. Both units are referenced to facilitate acquisition of materials internationally and minimize fabrication costs.

1.8 *This standard does not purport to address all of the safety concerns, if any, associated with its use. It is the responsibility of the user of this standard to establish appropriate safety, health, and environmental practices and determine the applicability of regulatory limitations prior to use.*

1.9 *This international standard was developed in accordance with internationally recognized principles on standardization established in the Decision on Principles for the Development of International Standards, Guides and Recommendations issued by the World Trade Organization Technical Barriers to Trade (TBT) Committee.*

2. Referenced Documents

2.1 *ASTM Standards:*[2]

C33/C33M Specification for Concrete Aggregates

D5821 Test Method for Determining the Percentage of Fractured Particles in Coarse Aggregate

E2521 Terminology for Evaluating Response Robot Capabilities

E2592 Practice for Evaluating Response Robot Capabilities: Logistics: Packaging for Urban Search and Rescue Task Force Equipment Caches

E2803 Test Method for Evaluating Emergency Response Robot Capabilities: Mobility: Confined Area Obstacles: Inclined Planes

3. Terminology

3.1 *Definitions:*

3.1.1 The following terms are used in this test method and are defined in Terminology E2521: abstain, administrator or test administrator, emergency response robot or response robot, fault condition, operator, operator station, remote control, repetition, robot, teleoperation, test event or event, test form, test sponsor, test suite, testing target or target, testing task or task, and trial or test trial.

3.1.2 The following terms are used in this test method and are defined in *ALFUS Framework Volume I*:[3] autonomous, autonomy, level of autonomy, operator control unit (OCU), and semi-autonomous.

4. Summary of Test Method

4.1 A robot under test traverses from one end of the gravel terrain to the other and back while alternating left and right turns around the pylons to complete a figure-8 path. The robot is credited with a fixed forward distance for each completed figure-8 path repetition.

5. Significance and Use

5.1 Traversing on terrains with small aggregate (such as Number 8 or smaller gravel, per Specification C33/C33M) could pose problems for ground robots because the aggregate may become incrementally packed into the locomotion subsystems (such as driving sprockets, belts, chains, tire treads, or track pads) leading to jamming, slippage, or other failures, and thus adversely affecting a robot's mobility. This test method addresses aforementioned issues of mobility.

NOTE 1—Larger-sized gravel might not be as easily packed into robotic locomotion subsystems but might present different types of mobility challenges such as angular, rough, sharp, or broken aggregate pieces interfering with wheels, tracks, or other types of locomotion mechanisms. These issues are out of the scope of this test method.

5.1.1 Small gravel based terrains are non-rigid in nature and could cause a robot to turn-in-place or dig-in when the robot is negotiating a tight turn. Certain robotic locomotion mechanisms might be designed for other mobility purposes and might not create sufficient traction against the specified gravel terrain. As such, extensive testing within this type of terrain may expose robot design or reliability issues and lead to field maintenance or repair.

5.1.2 The gravel traverse capabilities could be affected by additional factors such as the weight and its distribution, ground contact areas, and control schemes for the robot. As such, extensive testing within this type of terrain may also lead to innovations in robot design.

5.2 Key features of response robots are that they are remotely operated from safe standoff distances, deployable at operational tempos, capable of operating in complex environments, sufficiently hardened against harsh environments, reliable and field serviceable, durable or cost-effectively disposable, and equipped with operational safeguards. As such, a major advantage of using robots in response operations is to enhance the safety and effectiveness of responders or soldiers.

5.3 This test method aligns user expectations with actual capabilities to understand the inherent capability trade-offs in deployable systems at any given cost. For example, a design issue of the number of batteries to be packed on a robot could affect desired weight, endurance, or cost. Appropriate levels of understanding can help ensure that requirement specifications are articulated within the limit of current capabilities.

5.4 This test method provides a tangible representation of essential robot capabilities with quantifiable measures of performance. When considered with other related test methods in the suite, it facilitates communication among communities of robot users and manufacturers. As such, this test method can be used to:

[2] For referenced ASTM standards, visit the ASTM website, www.astm.org, or contact ASTM Customer Service at service@astm.org. For *Annual Book of ASTM Standards* volume information, refer to the standard's Document Summary page on the ASTM website.

[3] NIST Special Publication 1011-I-2.0 *Autonomy Levels for Unmanned Systems (ALFUS) Framework Volume I: Terminology*, Version 2.0. Available from National Institute of Standards and Technology (NIST), 100 Bureau Dr., Stop 1070, Gaithersburg, MD 20899-1070, http://www.nist.gov.

5.4.1 Inspire technical innovation and guide manufacturers toward implementing combinations of capabilities necessary to perform essential mission tasks.

5.4.2 Measure and compare essential robot capabilities. This test method can establish the reliability of the system to perform specified tasks, highlight break-through capabilities, and encourage hardening of developmental systems.

5.4.3 Inform purchasing decisions, conduct acceptance testing, and align deployment objectives with statistically significant robot capabilities data captured through repeated testing and comparison of quantitative results.

5.4.4 Focus operator training and measure proficiency as a repeatable practice task that exercises actuators, sensors, and operator interfaces. The test method can be embedded into training scenarios to capture and compare quantitative scores even within uncontrolled environmental variables. This can help develop, maintain, measure, and track very perishable skills over time and enable comparisons across squads, regions, or national averages.

5.5 Although this test method was developed for response robots, it may be applicable to other domains. Different user communities can set their own thresholds of acceptable performance within the test method for various mission requirements.

5.6 It is recommended that users of this test method consider their particular robot requirements when interpreting the test results. The capability evaluated in this test method alone shall be interpreted according to the scope of this test method and shall not be considered as an overall indication of the capability of the robot's mobility system nor of the entire robotic system. A single test method only captures the specified single aspect of a robot's capabilities. A more complete characterization of a robot's capabilities requires test results from a wider set of test methods.

6. Apparatus

6.1 The test apparatus is a fixed-size gravel terrain through which the robot has to traverse (see Figs. 1 and 2 for illustrations).

6.1.1 The terrain size may be scaled to provide for various levels of mobility constraints for robots under test depending on testing requirements. Three typical terrain sizes are nominally 3.6 m [12 ft] long by 1.2 m [4 ft] wide, 7.2 m [24 ft] long by 2.4 m [8 ft] wide, and 14.4 m [48 ft] long by 4.8 m [16 ft] wide, respectively. A test sponsor is authorized to specify other test apparatus sizes to suit particular deployment requirements.

NOTE 2—Test Method E2803 and this test apparatus can be combined to form a gravel test terrain on an incline.

6.1.2 Each of the terrains is fully covered with packed Number 8 aggregate (with nominal size of 2.36 to 9.5 mm, or 0.09 to 0.37 in.), as specified in Specification C33/C33M, for a minimal depth of 15 cm [6 in.], and evenly spread throughout the apparatus floor. The gravel shall be dry, both visibly and to the touch within the entire apparatus, including from the surface through the bottom.

6.1.3 Each terrain is recommended to be fully surrounded by walls that are nominally 1.2 m [4 ft] tall and are typically built out of plywood or oriented strand boards (OSB). Lumber with a nominal cross-section of 5 by 25 cm [2 by 10 in.] could also be used to surround the terrain and help further contain the gravel.

6.1.4 The ends of each of the terrains have end zones that are half as long as the width of the terrain and feature walls completely painted with alternating black-and-white, vertical, and nominally 30 cm [12 in.] wide stripes.

6.1.5 Four pylons define a figure-8 traverse path (see Fig. 2). The two timer pylons are placed at a distance from each other equal to the width of the terrain and centered between the end

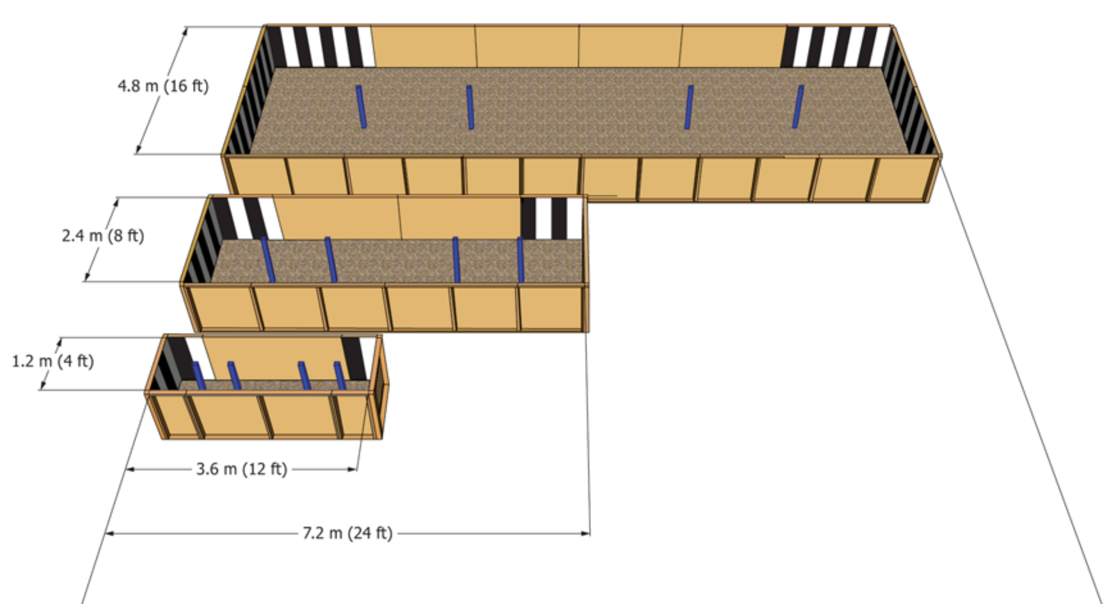

FIG. 1 Confined Area Terrains: Gravel; Three Sizes

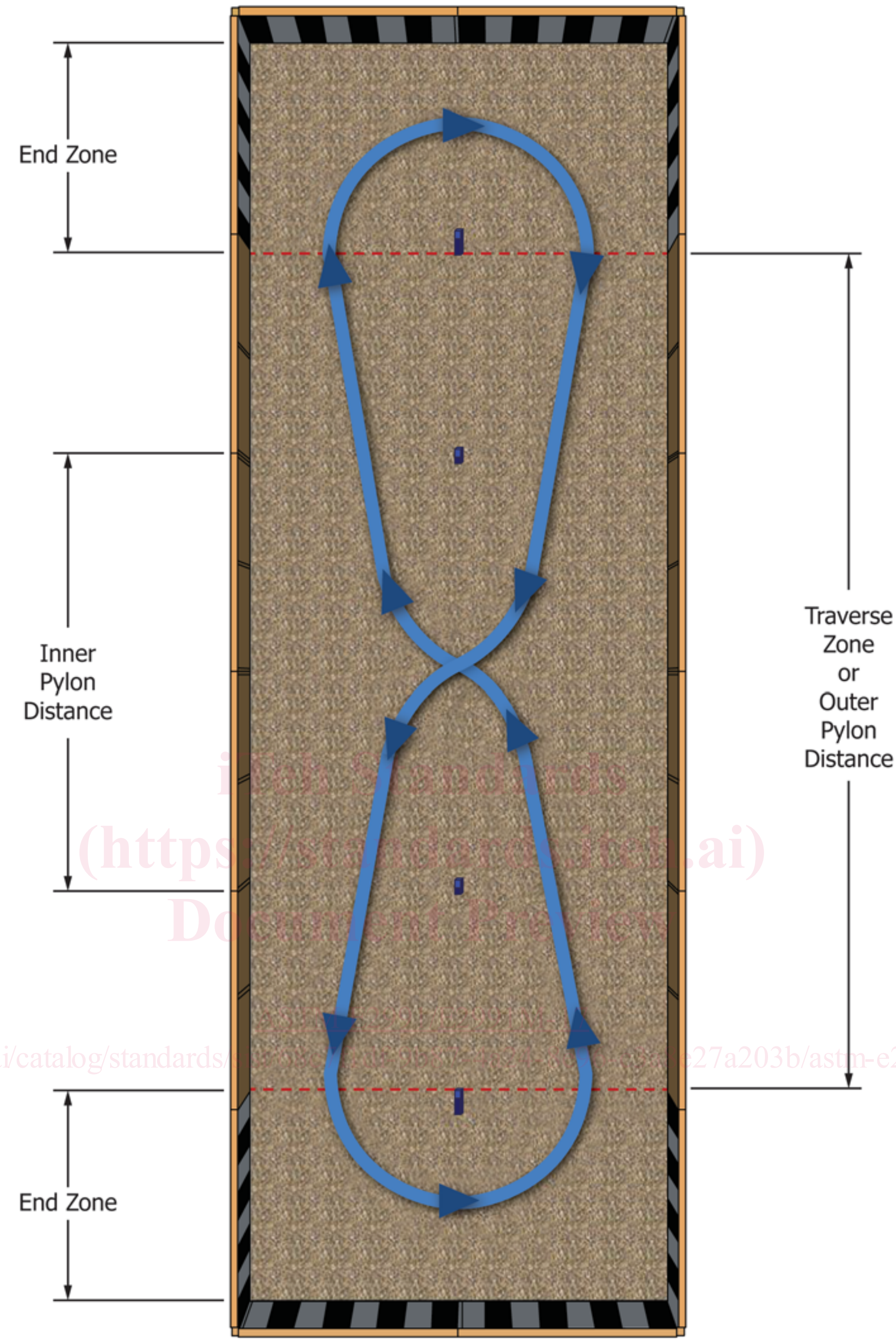

FIG. 2 Confined Area Terrains: Gravel; Figure-8 Path

zones and between the sidewalls. The two outer pylons are placed at the edge of each end zone and centered between the sidewalls. Pylons that are typically used for general traffic guidance and have nominal heights of at least 30 cm [1 ft] and nominal cross-section diameters of at most 30 cm [1 ft] shall be acceptable for use in this test method.

6.2 In terms of lighting, the lit test condition is specified as indoors (with furnished lighting), typically measured at nominally 150 to 300 lux or outdoors (in the daylight), typically measured at up to 1000 lux, nominally. The dark test condition is specified as nominally 0.1 lux. The dark test condition is not specified to be lower than 0.1 lux because of the implementation cost concerns. 0.1 lux is dark enough to require the robot to have onboard illumination—a test robot capability—in order to perform the task. However, it is recognized that actual operational environments may be darker.

NOTE 3—When this test apparatus is implemented in an International Organization for Standardization (ISO) freight container, the 0.1 lux darkness could be achieved by turning off all the inside lighting sources and covering the entrance of the container entirely with light-blocking drapes.

6.3 Appropriate measuring devices shall be used to measure the robot's performance within an apparatus.

NOTE 4—A stopwatch may be used to measure the time and an electronic beam-breaking sensor could be used to count the number of repetitions. Video captured with four simultaneous feeds is highly recommended for recording the tests. The four simultaneous video feeds are: an overall view of the apparatus, a detailed view of the robot performing the task, a view of the OCU's display showing the view from the robot's onboard camera, and a view of the operator's hands controlling the robot. This type of record helps verify test results and enhance operator training.

7. Hazards

7.1 In addition to 1.8, users of this test method shall also address equipment preservation and human-robot coexistence concerns. Safety setups such as belays and containment walls shall be used when there are such concerns. Environmental conditions, such as high or low temperatures, excessive moisture, and rough terrains can be stressful, exceed the respective ranges within which the robot is built to properly operate, or damage robotic components. These conditions can also cause unexpected robot motions that, in turn, can have negative effects on the humans that are nearby or on the robot itself.

7.2 Identify all the emergency stop (E-stop) button(s) on the robot chassis and the OCU before operating or interacting with the robot.

7.3 While the robot is active and the E-stop button is disengaged, avoid:

7.3.1 The areas directly in front of and behind the robot,

7.3.2 The reachable radius of the robot's manipulator, as equipped, and

7.3.3 Touching the robot other than to engage the E-stop button.

8. Procedure

8.1 Ensure that the apparatus and environmental conditions are set up properly according to the Apparatus Section 6, including the size of the test terrain, the moisture level of the gravel, and the lighting within the test apparatus.

8.2 Ensure that the robot system configuration has been identified and documented.

8.2.1 The particular system configuration to be tested shall be comprehensively identified and uniquely named by using the make, model, and applicable configuration name as provided by the manufacturer. This identification process includes measuring the time required to bring the system to the operationally ready state, called setup time. The process, then, involves measuring and documenting the dimensions and weights of all the subsystem, components, and as-shipped packaging. These include the robot, OCU, and other sustainment and maintenance items such as power sources and spare parts. This identification process also lists subsystems, payloads, and items in the field-maintenance kit. These include tools and consumable items such as duct tape, cable ties, and other items. Documentation shall also include detailed photographs of all of the above as well as videos of routine maintenance tasks (for example, battery change). The system configuration shall remain the same for all relevant tests to enable direct comparison of performance and to identify capability trade-offs between different configurations. Any number of identified system configurations can be subjected to testing.

8.3 Determine the number of required repetitions if measuring system capabilities.

8.3.1 Test trials shall produce enough successful repetitions to demonstrate the reliability of the system or operator necessary for the envisioned missions. The higher the ratio of successful repetitions to faults, the more reliable the system or operator. The more repetitions completed with that ratio, the more confidence may be placed in that reliability. The calculated reliability and confidence levels can be determined from statistical tables. Some missions may require higher reliability. Others may be more resilient to failure and can accommodate lower reliability. A test trial of 30 repetitions or more is recommended to establish a system's capability. Operator proficiency trials are typically time limited as specified in 8.4.

8.3.2 A reasonable starting threshold may be at least 80 % reliability with 80 % confidence. This can be achieved by performing 30 repetitions with 27 or more successes. When 30 repetitions is not feasible, this reliability and confidence may still be achieved if the first test trial includes 20 repetitions with 19 or more successes or 10 successful repetitions.

8.3.3 Multiple trials are allowed to improve the performance of the same tested robot configuration. If this is the case, the latest 30 consecutive repetitions from across the multiple trials shall be considered together when determining the ratio of successful repetitions to faults.

8.4 Select the desired timer increment duration in minutes.

8.4.1 Timers shall be used to capture the elapsed time of the trial. They provide deterministic indications of start and end times to minimize the uncertainty of the elapsed time of the trial.

8.4.2 Sequential timer increments, in minutes, shall be used to measure system capabilities with operators designated by the manufacturer. A reasonable timer increment may be the estimated time to complete five or more successful repetitions. Use enough timer increments to complete the required number of repetitions necessary to calculate the desired reliability and confidence in the system to perform the task. If sequential timer increments start adding up to more than the overall test plan can accommodate, truncate the trial at the end of a timer increment and calculate the resulting reliability and confidence separately. This indicates that the system may be reliable but less efficient than expected, although a low number of repetitions will limit confidence in that assertion.

8.4.3 Time limited trials, in minutes, shall be used to measure operator proficiency. A reasonable timer increment may be the estimated time to complete five or more successful repetitions. When measuring and comparing operator proficiency, it is important to equalize the elapsed time of operation for "expert" and "novice" operators to normalize for fatigue. However, shorter elapsed times resulting in fewer successful repetitions reduces confidence in the measured reliability of the operator to perform the task.

8.5 Select an operator to perform the test.

8.5.1 To measure system capabilities, the operator or operators shall be designated by the manufacturer in order to align interests and to ensure the best possible robot performance. The best results obtained with the manufacturer-designated operator shall be used as the 100th percentile of operator proficiency for the given system. Any other operator can measure her or his proficiency as a percentage. Example levels of operator proficiency may be "Novice" (0 % to 39 %), "Proficient" (40 % to 79 %), and "Expert" (80 % to 100 %). Operator proficiencies may be compared across regional or national averages.

8.5.2 Practice is optional before testing. The operator shall be familiar with the test procedure, the apparatus settings, and the environmental conditions enough to be tested.

8.5.3 During a test trial, the operator shall be located remotely at a station that is out of direct sight and sound of the robot in the test apparatus while still maintaining communications with the robot.

8.5.4 During a test trial, any human communication to the operator about the robot's status within the test apparatus shall be considered a fault. However, this should not deter communications regarding the safety of the robot or personnel.

8.6 Place or drive the robot to the start point within the apparatus as specified in Section 6.

8.7 Start the timer increment.

8.8 Perform the task as specified in 8.8.1 and 8.8.2 and repeat until the timer increment expires—see 8.4 on how the expiration dictates the numbers of task repetitions.

8.8.1 Traverse the figure-8 path on the terrain in any driving orientation, from one end zone to the other and back, while alternating left and right turns around the pylons.

8.8.2 For scoring purposes, a successful repetition shall be awarded when the robot returns to the start point and is ready to perform the task again. However, if the timer increment expires, half repetitions can be awarded if the robot has successfully entered the far end zone (see Fig. 2).

8.8.3 Once a robot begins a test trial, the task shall be performed until the specified number of timer increments has expired.

8.9 *Handle Exceptions:*

8.9.1 *Fault Conditions*—When a fault occurs, maintain the trial timer while the robot situation is assessed. To continue the test trial, return the robot to the start point, where minor maintenance is allowed, although may be time limited as specified by the test sponsor. No spare parts shall be used although consumables may be allowed as defined in 8.2 (for example, duct tape and cable ties). The end of the test trial shall be declared if the robot cannot continue. A maintenance report shall be generated, including the information of the occurrence time, repair time, cause remedy, and tools used. Video recording of all of the maintenance activity may be beneficial for later verification, referencing, or training purposes. The following occurrences shall be considered faults resulting in an unsuccessful repetition:

(1) Any robot contact with the pylons or any other non-contact areas as defined in the Apparatus Section.

(2) Any robot situation requiring physical interaction by the operator or others during the trial (for example, a stuck or disabled robot).

(3) For systems with short endurance, battery changes may be allowed without faults in order to achieve statistically significant repetitions.

8.9.2 *Inconsequential Exceptions*—Abnormal, momentary system behaviors that exhibit but do not negatively impact the task execution shall be noted as a part of the test results but the corresponding repetition or repetitions shall not be declared a failure or failures. Examples include momentary software reset or frozen screen within the OCU, momentary loss of communication, etc. Faults shall be declared once such behaviors incapacitate the robot.

8.9.3 *Administrative Pauses*—Anytime during the test or training when the test apparatus needs adjustments or repairs for reasons not the fault of the operator or the robot, the timer shall be paused, the robot shall be stopped, and the apparatus shall be repaired before continuing the trial.

9. Calculation of Results

9.1 Count the total numbers of successful repetitions and faults.

9.2 The calculation of the average rate of advance is performed using the following equation:

$$(\text{number of successful repetitions}) \times (\text{distance per repetition in meters})/\text{timer increments in minutes}$$

9.3 The average rate of advance is calculated only when a test trial is successful. The calculation is recommended to be done to the tenths place.

9.4 A representative set of robots with different sizes, weights, and mobility mechanisms were selected for testing. The 7.2 m [24 ft] long by 2.4 m [8 ft] wide size apparatus was used. Their performance was calculated using this process and the results are shown in Table 1.

10. Report

10.1 The following information on test conditions and test administration shall be documented (except where specifically noted as "Optional") on a test form:

10.1.1 *Standard Suite and Task Name*—For this test method, "mobility" and "gravel terrain traverse" shall appear on the test form.

10.1.2 *Standard Identifier*—ASTM standard designation.

10.1.3 *Date*—The date on which the test was conducted.

10.1.4 *Facility Name*—Laboratory or field name.

10.1.5 *Location*—City and state; and country when not in the United States.

TABLE 1 Test Results for Mobility: Confined Area Terrains: Gravel

Robot	Weight, kg	Length, cm	Mobility Mechanism	Average Rate of Advance, m/min
A	<5	<25	Wheel Based	20
B	5–10	<50	Track Based	25
C	5–10	<50	Wheel Based	50
D	25–30	50–75	Track Based	N/A
E	75–100	75–100	Track Based	10