Received: 22 September 2018 | Revised: 8 June 2020 | Accepted: 14 June 2020
DOI: 10.1002/rob.21979

REGULAR ARTICLE

WILEY

Self-reliant rovers for increased mission productivity

Daniel Gaines[1] | Gary Doran[1] | Michael Paton[1] | Brandon Rothrock[1] | Joseph Russino[1] | Ryan Mackey[1] | Robert Anderson[1] | Raymond Francis[1] | Chet Joswig[1] | Heather Justice[1] | Ksenia Kolcio[2] | Gregg Rabideau[1] | Steve Schaffer[1] | Jacek Sawoniewicz[1] | Ashwin Vasavada[1] | Vincent Wong[1] | Kathryn Yu[1] | Ali-akbar Agha-mohammadi[1]

[1]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California, USA

[2]Okean Solutions, Seattle, Washington, USA

Correspondence
Daniel Gaines, Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Dr, Pasadena, California 91109, USA.
Email: Daniel.M.Gaines@jpl.caltech.edu

Abstract

Achieving consistently high levels of productivity has been a challenge for Mars surface missions. While the rovers have made major discoveries and dramatically increased our understanding of Mars, they require a great deal of interaction from the operations teams, and achieving mission objectives can take longer than anticipated when productivity is paced by the ground teams' ability to react. We have conducted a project to explore technologies and techniques for creating self-reliant rovers (SRR): rovers that are able to maintain high levels of productivity with reduced reliance on ground interactions. This paper describes the design of SRR and a prototype implementation that we deployed on a research rover. We evaluated the system by conducting a simulated campaign in which members of the Mars Science Laboratory (Curiosity rover) science team used our rover to explore a geographical region. The evaluation demonstrated the system's ability to maintain high levels of productivity with limited communication with operators.

KEYWORDS
planetary robotics, planning, position estimation, navigation, obstacle avoidance

1 | INTRODUCTION

Maintaining high levels of productivity for Mars rover missions has proven to be challenging. While the operations teams have achieved impressive accomplishments with the rovers, doing so requires significant human effort to develop command products for the rovers and it often takes longer, more days on Mars (aka sols), than anticipated to accomplish objectives. A primary reason for these productivity challenges is the heavy reliance on interaction between the rovers and ground operators to accomplish mission objectives. For example, rovers depend on operators to provide a detailed schedule of activities, select science targets, navigate around slip hazards, and recover from anomalies. When combined with the limited communication opportunities between the rovers and human operators, this reliance on ground interaction results in under-utilization of vehicle resources and increased sol on Mars to accomplish mission objectives.

This issue is anticipated to become increasingly important for future Mars missions as our aging sun-synchronous data relay orbiters are replaced by non-sun-synchronous orbiters. The overflight patterns of non-sun-synchronous orbiters result in reduced opportunities for ground-in-the-loop interactions with the rovers.

We have developed a design for future rover flight software to address these productivity challenges. We refer to the approach as self-reliant rovers (SRR) as the objective of the work is to increase the ability of rovers to accomplish objectives and respond to unexpected conditions with reduced reliance on human intervention. Although our objective is to reduce the reliance on ground support to promote productivity, we are by no means attempting to remove human operator involvement, whether mission engineers or scientists. To the contrary, a major emphasis of our work is to enable operators to provide guidance to rovers without requiring up to date knowledge of the rover and its environment.

In this paper we present our design for developing and operating a SRR. The design is motivated by the results of an extensive case study of Mars Science Laboratory (MSL) operations that we performed during the first year of this project (Gaines et al., 2016). We have developed a prototype implementation of this design and deployed it on a research rover. The prototype includes advancements in goal planning, autonomous science, health assessment, autonomous navigation and global localization.

To evaluate the efficacy of the SRR approach we conducted a simulated exploration campaign in which members of the MSL science team used our rover to explore a geographical region. The evaluation demonstrated that the our approach was able to achieve significant productivity improvements over the current approach to rover design and operation. In particular, the evaluation demonstrated that the SRR approach achieved an 80% reduction in number of sols to complete the campaign.

To provide context for the project, we begin in the next section with background on the current practice for operating Mars rovers with a brief introduction to MSL operations.

2 | BACKGROUND ON MSL MISSION OPERATIONS

One of the challenges of surface missions is the degree to which operations are impacted by a priori unknown and changing environmental conditions. While orbital imagery provides valuable information to guide activity, it does not capture all the conditions that affect the rover. For example, while orbital imagery may indicate that exploring a particular region is promising to achieve a science objective, the specific science targets are not known until additional data is collected from the rover itself. As such, surface operations must be reactive and respond to the results of activity carried out during the previous sol (Martian day). This daily planning activity is referred to as “tactical” operations and is patterned after the tactical operations developed for the Mars Exploration Rovers (Mishkin et al., 2006).

MSL operations augments this tactical process with “strategic” and “supratactical” phases (Chattopadhyay et al., 2014). Strategic planning focuses on developing long-term plans, typically spanning weeks or months, to achieve high-level objectives. Examples of strategic planning include the development of strategies for exploring a large geographical area or a high-level traverse route for reaching a distant objective. The supratactical stage provides a bridge between the long-term strategic plan and the day-to-day, highly reactive tactical process. The process is designed to coordinate the complex science instruments and manage the constraints and resources required to conduct campaigns.

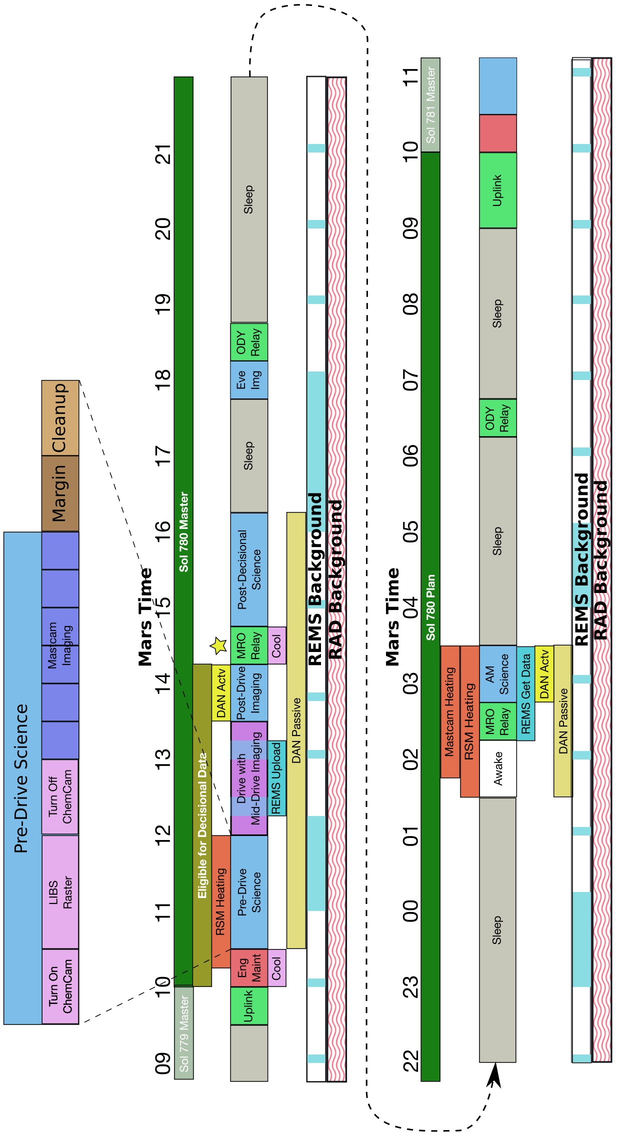


FIGURE 1 Example sol in the life of the rover [Color figure can be viewed at wileyonlinelibrary.com]

2.1 | An example sol in the life of the rover

To provide an idea of how the team operates the rover, Figure 1 illustrates an example sol of rover activity. This is an example of a typical drive sol, derived from an actual sol, Sol 780, command products. Following are some key aspects of the diagram.

The plan for each sol begins with an “Uplink” window in which new commands products may be sent to the vehicle from Earth. There are various downlink windows throughout the sol in which the rover uses orbiter relays to send collected data back to Earth. While there are multiple downlink windows, certain downlinks have increased importance based on the time that data in the relay will reach operators. If data from a relay will reach operators by the start of the next tactical planning shift, then they relay is termed “decisional” because data from the relay can be used to make decisions in for the rover's next plan. Which relays are considered decisional depends on the relative timing between Earth and Mars along with latencies in the orbiter relay process. In Figure 1 the starred "MRO Relay" represents the decisional relay for this sol. It is important to realize that for this plan, only the data collected before this pass could be used to inform the next plan. While the remaining data will eventually be sent to Earth and may be used to inform future plans, it will arrive too late to inform the next plan.

Another important aspect of Figure 1 is how the team structures the sol into "blocks" of activity. For example, the main portion of the rover's day consists of a Pre-Drive Science block, a Drive with Mid-Drive Imaging bock and a Post-Drive Imaging block. The block structure organizes activity into related groups and allows a "Master" sequence to enforce timing between these major types of activity. The latter has to do with uncertainties in predicting the duration of activity in the plan. Due to environmental conditions such as lighting, scene content and terrain, the time to perform imaging and drive activities varies. The team uses the block structure to ensure that if activity in one block runs longer than expected, it can be cut off to avoid interfering with subsequent activity. To protect against loss of data, the team builds "Margin" into each block, to allow activities to run longer than predicted. To deal with cases where durations exceed allocated margin, the team also sequences "Cleanups" after each block, to ensure that any activity is forcefully terminated before the start of the next block.

2.2 | Constrained sols

The vast majority of the surface mission is conducted with the team restricting operations to daytime hours on Earth. The consequence is that the operations team is often out of sync with the activity of the rover on Mars. Figure 2 illustrates the impact this can have on the data available to the team during planning. In the diagram, the end-of-day relay from the rover arrives on the ground late in the Earth day. The team waits until the next Earth day to begin planning. Meanwhile the rover is waking up for its next Mars day without a new set of command products from Earth. By the time the team has completed the tactical process, they must wait for the subsequent Mars morning to uplink the products to the vehicle.

This often limits what the team can have the rover do during the middle sol of Figure 2. If the vehicle were allowed to make significant changes to its state, in particular driving to a new location, this would significantly limit the types of activities the team could command on the subsequent sol. These limited activity sols are referred to as "constrained sols" because the latency of data often restricts the type of activity the team can perform.

A similar situation arises when the team takes days off for weekends and holidays. In these cases, the team will create plans that span multiple sols (aka multisol plans). Again, activities that result in significant changes to vehicle state are limited since they will impact the activity that can be done in later sols of the plan.

Given the current way in which we design and operate rovers, constrained sols are a major detractor from mission productivity. For example, with current surface operations, when the rover drives to a new location it must wait for imagery collected at this location to be sent to Earth and for the science and engineering teams to analyze the data and identify the specific set of activities to perform at the location to meet their current mission objectives. Depending on the phasing of Earth and Mars local time, this can result in an entire sol in which the rover waits for these new activities.

Overall, 41% of sols on the MSL mission are constrained sols simply due to the frequency that the ground operations team is available to respond to the latest downlink with a 9-h shift. More recent orbiters, such as Mars Atmosphere and Volatile EvolutioN (MAVEN), have highly eccentric orbits which do not provide the consistent end-of-sol relay opportunities that MSL has enjoyed with MRO. Therefore, relying on such orbiters would result in an increase in the number of constrained sols.

2.3 | Productivity challenges

In the first year of the SRR project we conducted an extensive case study of MSL operations to identify significant productivity challenges. In the study, we analyzed three different MSL campaigns. Our report provides full details of the study (Gaines et al., 2016). Here we provide a high level summary of the major outcome.

The report identified two primary metrics for mission productivity:

- Percentage of sols making significant contributions toward campaign objectives
- Utilization of vehicle resources

The term "significant contribution" is used to distinguish between certain types of activities that, while arguing providing some support for a campaign objective, was not providing a substantial contribution. For the campaigns we investigated, there were two types of activities that were considered to have low productivity if they were the only activities performed on a sol.

First, there was a specific set of "blind targeting" activities for the ChemCam instrument. These activities were sometimes performed on constrained sols (i.e., sols that followed a drive without ground-in-the-loop. The activity performs a ChemCam observation at a fixed point relative to the rover. Depending on where the rover happens to stop, this could result in an observation of outcrop, a float

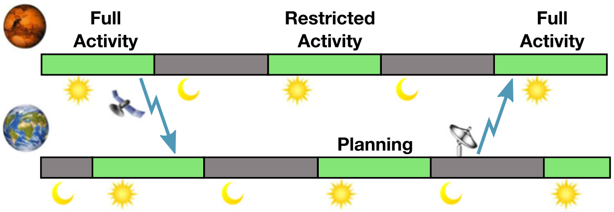


FIGURE 2 Mars activity versus Earth planning [Color figure can be viewed at wileyonlinelibrary.com]

rock, or soil. For sols in which this blind targeting activity was the only activity contributing toward campaign objectives, we considered such sols as having low contribution to campaign objectives.

Second, we considered the productivity of sols in which the rover spent the sol driving to a location of interest. If the rover was able to effectively use resources to make progress toward the operators' goal location, then the sol would be considered to have high productivity. In contrast, cases in which the drive faulted out or the drive distance was limited by operators' visibility of terrain rather than vehicle resources were considered low productivity.

2.3.1 | Sol productivity

For the first metric, we looked at how effectively each sol of a campaign was used toward accomplishing campaign objectives. Figure 3 shows the results.

Sols were binned into the following categories:

- *Campaign*: The sol directly contributed to the campaign objectives with remote sensing and/or drives. These sols are considered "productive" sols.
- *Campaign multisol*: The sol contained significant activity was performed toward the campaign objectives as part of a multisol plan, either due to a weekend or constrained planning. This limited the choices the team had for what to include in this sol and may have been used differently had it not been part of a multisol plan. Despite this qualification, these sols are still considered "productive" sols.
- *Extra drives*: A sol in which unexpected additional drives were required to reach objectives. This includes cases where a planned drive faulted out early and had to be replanned on a subsequent sol.
- *Postdrive multisol*: A sol in which the team was not able to achieve substantial campaign objectives due to lack of data following a drive during a multi-sol plan.
- *Deferred*: A sol in which campaign objectives were unexpectedly deferred due to the need to respond to an issue identified during tactical plan development or in response to an event from received downlink data.
- *Runout*: A sol consisting of very low activity, that is, used in cases the team had to create multisol plans but the tactical timeline capacity did not allow for sufficient time to develop activities for all sols of the plan.

We observed that across the three campaigns, there was a relatively large number of sols with low productivity. Specifically, 48% of the sols were classified as low productivity.

2.3.2 | Resource utilization

The above observation indicates that there is opportunity to increase productivity if these low productivity sols can be used more productively. However, it is not clear if there are sufficient vehicle resources available to be more productive. To answer that question, we performed a detailed analysis of how vehicle resources were used, including energy, time and data volume. The case study report provides the full detail. Here we summarize that we observed that there were, in fact, sufficient resources to be more productive.

As an example, Figure 4 shows an analysis from the Pahrump Hills Walkabout campaign that showed there was an estimated 72 h of unused energy over the 19 sols of the campaign, or an average of 3.8 h per sol of unused energy. The unused energy comes from two sources: shunt energy and rover idle time. When the rover battery has been charged to its capacity, additional generated energy is shunted. This lost energy could have instead been used by the rover to perform additional activity. When activities on the rover complete earlier than expected, the rover remains awake in an idle state until the schedule shutdown time. As with shunt energy, this idle energy would ideally be put to more productive use.

We found similar results for the other two campaigns and found that there was an average of 3 h of unused energy across all

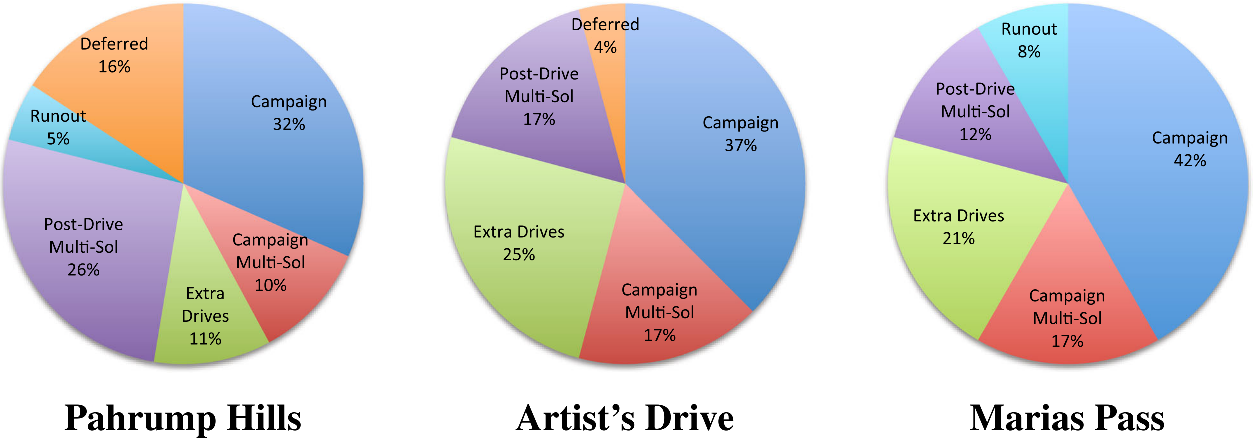


FIGURE 3 Sol productivity for case study campaigns [Color figure can be viewed at wileyonlinelibrary.com]

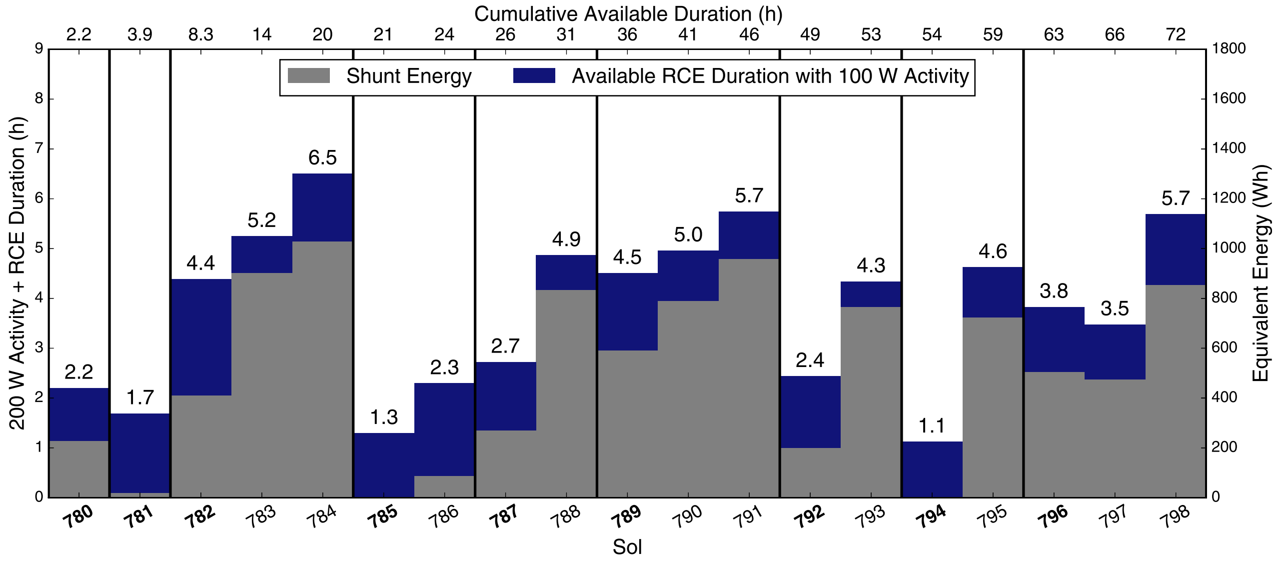


FIGURE 4 Estimate of extra duration availability for Pahrump Hills Walkabout campaign [Color figure can be viewed at wileyonlinelibrary.com]

campaigns. For context, we observed that each sol had an average of 9 h of activity, thus an additional 3 h would represent a 33% increase in overall activity. More importantly, if these hours can be used effectively it would enable the team to increase the productivity of the lower productivity sols from Figure 3.

2.3.3 | Significant productivity challenges

The case study included an analysis of the challenges the team faced in making more effective use of sols and vehicle resources. The report includes a more complete listing and discussion. Here we highlight the three that we believe contributed the most to loss of productivity in the campaigns.

1. *Predicting Vehicle Resource Usage:* It is difficult to predict how long activities will take to complete. Operators tend to overestimate duration to avoid activities being cut off. As a result, activities typically end earlier than expected which contributes to the rover idle time described in Section 2.3.2.
2. *Ground-in-the-loop for target selection and drive planning*: The rover relies on ground operators to pick out specific science targets and to identify paths around slip hazards, such as patches of loose sand. This results in a significant drop in productivity on sols that follow drives during constrained periods of the mission. Even during nonconstrained sols, it constrains the timing of activity that can change the state of the vehicle and activity that acquires decisional data to occur before the decisional pass.
3. *Ground-in-the-loop to respond to outcome of activity*: We observed several instances where the team decided to re-do an activity, or return to a previous location, after observing the data received from the vehicle. This included the need to replan drives that faulted out or require observations that did not have intended results due to lighting conditions or targeting problems.

The above challenges served as a guide for our project. We used these challenges to identify what aspects of the flight software architecture and ground practices to target and what approaches and technologies to develop to address them. The following section describes the resulting system and the changes we identified.

3 | OVERVIEW OF THE SRR DESIGN

We developed the SRR system to address the productivity challenges described in Section 2.3.3. The system is designed within the context of the current rover flight software architecture with changes necessary to address the identified productivity challenges (Weiss, 2013). Figure 5 provides an overview of this architecture and the changes we are introducing with the SRR system.

The architecture consists of components organized into three layers: behaviors, activities, and functions. Each successive layer has a reduced degree of autonomy, fewer interactions with other components, and a narrower scope of system knowledge.

- *Behavior*: Collection of autonomously scheduled activities in service of an over-arching mission goal. Contains broad system knowledge.
- *Activity*: Coordinates function invocations to achieve some high-level spacecraft task. Encompasses knowledge local to the activity being managed.
- *Function*: Primitive action required to achieve a single well-delineated spacecraft objective. Contemplates only highly localized function-specific knowledge.

Following is a summary of the changes we are introducing for the SRR approach. Subsequent sections will provide more details on the most significant changes.

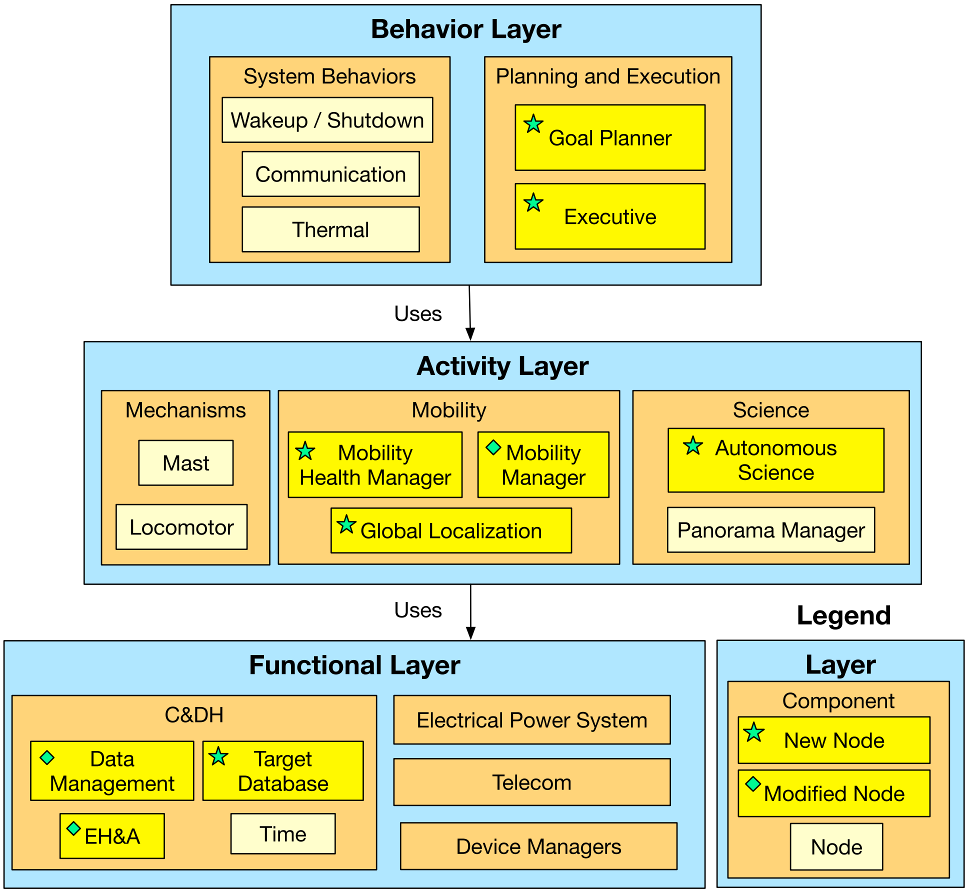


FIGURE 5 Self-Reliant Rover flight software architecture [Color figure can be viewed at wileyonlinelibrary.com]

- *Goal planner*: Generates onboard activity plans to accomplish mission goals. This addresses Productivity Challenge Item 1, from Section 2.3.3, by enabling the rover to respond onboard to actual resource usage. It also addresses Productivity Challenge Item 2 by allowing the rover to respond to new goals identified by onboard autonomous science.
- *Executive*: Executes plans generated by the Goal Planner and provides updates to facilitate replanning.
- *Autonomous science*: Identifies science targets when the rover enters an unexplored area. Increases the scope of guidance that scientists can provide and deepens the integration with onboard planning, as compared with previous autonomous science on MSL (Francis et al., 2017). This addresses Productivity Challenge Item 2 by enabling the rover to identify its own targets, with the help of scientist guidance, without waiting for ground-in-the-loop interaction.
- *Mobility manager*: Improves navigation by reasoning about terrain-dependent slip and by detecting and avoiding sand hazards. This addresses Productivity Challenge Item 2 by enabling the rover to plan drives around sand hazards without relying on ground assistance.
- *Mobility health manager*: Increases the robustness of mobility activities and the scope of faults from which the rover can autonomously recover by leveraging model-based fault detection and isolation. This addresses Productivity Challenge Item 3 by enabling the rover to recover from certain types of mobility faults that would otherwise require the rover to wait for ground intervention.
- *Global localization*: Maintains high quality position knowledge over long traverse distances via onboard global localization (a technique that previously required ground operator support).
- *Target database*: Facilitates communication about targets of interest among scientists, engineers, and onboard autonomous components by leveraging previous ground operations tools onboard.
- *Data management:* Provides queryable onboard data product access to autonomous components such as onboard science analysis.
- *EH&A*: Provides onboard access to engineering, housekeeping, and accountability telemetry for use by autonomous reasoning components.

4 | ONBOARD PLANNING

We address a number of the aforementioned productivity challenges via the inclusion of an onboard *goal planner*. The planner incorporates up-to-date knowledge of onboard resource levels and vehicle state to generate sequences of activities that fulfill high-level mission objectives. This allows the team to use less conservative modeling of activity resource use and duration, which in turn contributes to less idle time due to unused margin.

The inclusion of a planner also enables the system to respond to new objectives identified by onboard autonomous science, described in the next section. This combination of onboard planning and autonomous science allows the rover to make better use of the low productivity sols identified in Section 2.3.1.

Ground-based planning teams retain the ability to command specific actions, but the primary means of guiding rover operations becomes the crafting of these high-level goals. In this way, the onboard goal planner supplements and enhances, rather than replaces, the traditional tactical planning process.

4.1 | Campaign intent

A significant challenge to maintaining high rover productivity under reduced operator interaction is conveying operator guidance and objectives without requiring operators to have up to date knowledge of the rover and its environment. Our approach is motivated by prior operations practice. In traditional operations, each planning cycle begins with a review of the current long term objectives of the mission presented in the context of the latest available rover state data (Chattopadhyay et al., 2014). The operators assimilate all the various objectives, state data, and mission knowledge to synthesize a high quality plan that makes progress toward the goals while respecting limited rover resources such as time, energy, and data volume.

The team will typically have several high-level objectives to pursue. For example, during MSL's Pahrump Hills Walkabout campaign, the primary focus of the mission was to collect observations of exposed outcrop forming the basal layer of Mount Sharp (Gaines et al., 2016). This required driving the rover to several locations and acquiring high quality Mastcam and ChemCam observations selected locally at each stop.

Concurrently, the team also pursued a variety of supplementary objectives. During this campaign, the comet Siding Spring (Comet C/2013 A1) would pass Mars closer than any other known comet flyby of Earth or Mars. The operations team thus incorporated comet observations into the rover plans. In addition, the team planned ongoing periodic observations to study clouds, dust devils, and atmospheric opacity. A wide range of recurring engineering activities also had to be included: instrument calibrations, telemetry collection, and system configuration management.

Importantly, the quality of the plan is not just a function of what activities are scheduled; it depends on how well they relate to the current objectives and to each other. Each individual outcrop observation was valuable, but understanding the geology of the region required accumulation of a variety of observations that were spatially distributed throughout the area. Periodic tasks such as atmospheric measurements and engineering activities had similar preferred temporal patterns that the team must try to match.

We developed the concept of *campaign intent* to convey such information to the rover so that it may generate its own prudent in situ plans when human guidance is prohibitively delayed. Campaign intent specifies a set of goals for the rover and the relationships among those goals. We gleaned three initial types of campaign intent from MSL scenarios, as summarized in Figure 6.

- *Class sampling*: Choose observation targets that best exemplify a particular feature (e.g., layering). Once identified, the targets form a goal set. Value typically accumulates with additional samples from the set, but eventually reaches a point of diminishing returns.
- *Temporally-Periodic sampling*: Schedule goals to match a repeating temporal pattern (e.g.,hourly). The preferred goal cadence typically allows at least some timing flexibility.
- *State-based sampling*: Trigger goals based on the evolution of the rover/terrain state (e.g., at every 50 m traveled). The state criteria is typically expressed as a preferred cadence with some flexibility.

4.2 | Using campaign intent to guide planning

Our approach to plan generation is based on branch-and-bound search. Starting from the empty plan, each iteration of search expands a chosen partial plan into many possible successor plans (the branches). Each potential successor is scored and must exceed a running threshold of plan quality (the bound) To be retained for future expansion; otherwise it is pruned (along with all its descendants). Specifically, the optimistic maximum quality of any plan based on the candidate partial plan must exceed the pessimistic minimum quality prediction of all other candidates already considered. Plan quality is evaluated as the degree of satisfaction of the campaign intents, which may be both priority tiered and utility weighted by the user. The frontier of un-expanded partial plans is periodically sorted by estimated final plan quality, yielding a hybrid of depth-first and best-first expansion order.

Partial plans are always expanded forward in time by appending one of the possible subsequent actions to the growing plan. The possible actions include mandatory goals (such as communication passes), auxiliary actions (such as sleep periods), as well as all the possible goals introduced by campaign intents. For temporal and state-based campaigns, this is just the next instance of the periodic goal, timed within its allowed cadence. For unordered goal set campaigns, each remaining un-attempted goal becomes a possible addition. In the limit, the search will thus evaluate (or justifiably prune) all possible combinations and orderings of campaign goals.

The complete search can be very time intensive, but is guaranteed to return an optimal plan according to the expressed campaign preferences. Even without running to completion, the search can return the best plan encountered so far. This anytime algorithm feature allows the rover to limit its planning time and proceed to be productive with a reasonable (but not provably optimal) plan. Minor plan perturbations during execution are accommodated by time-efficient repair strategies (e.g., to shift actions forward after a small driving delay), while major disruptions (such as an insurmountable obstacle in a drive, or the injection of an entirely new goal) invoke a full replanning cycle so that all goals are reconsidered. More details on the planning algorithm can be found in Russino et al. (2019).

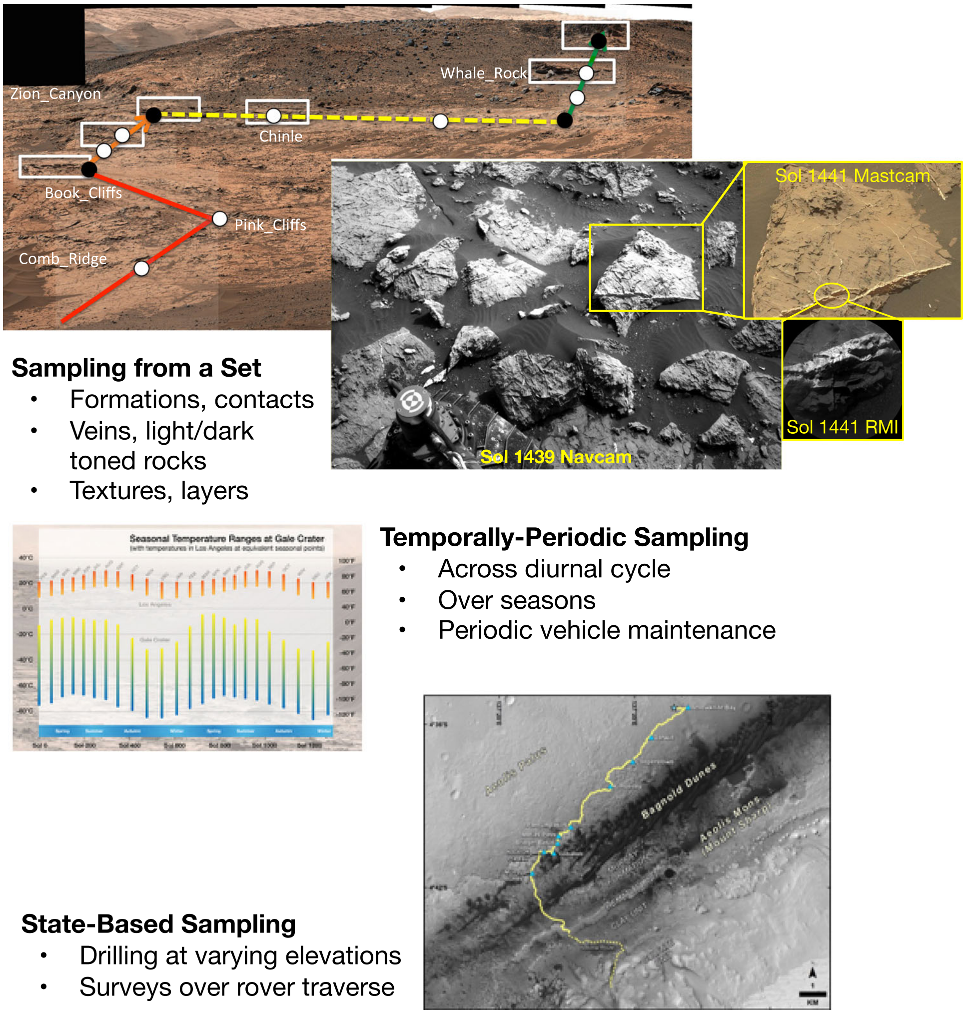


FIGURE 6 Summary of campaign intent types [Color figure can be viewed at wileyonlinelibrary.com]

Figure 7 shows an example plan generated by the search algorithm. The planning model derives from the operational MSL activity model and features important mission aspects such as science campaign activities, communication windows, regenerative sleeping, and device heating.

The campaign objectives provided to the rover in this example include: a goal set campaign with a distant Mastcam target (entailing a long-range traverse), a temporal campaign with recurring atmospheric opacity (tau) measurements every 3 h, and state-based campaign with mid-drive survey actions after every 75 m traveled. The resultant plan demonstrates how the planner synthesizes the campaign relationships to coordinate rover activity, including pausing the ongoing drive action to interleave other objectives.

5 | SCIENTIST-GUIDED AUTONOMOUS SCIENCE

As identified in our case study, one factor affecting the productivity of rover missions is the requirement of ground-in-the-loop for selecting science targets using the highest-resolution imagery acquired after arrival at the new site. If scientists would like to take observations and measurements at several locations, they must use at least one ground-in-the-loop cycle at each location. With limited communication opportunities, this can stretch out the exploration of a series of locations over many sols.

5.1 | Using campaign intent to guide autonomous science

To alleviate the requirement of ground-in-the-loop for science target selection, we have developed a collection of capabilities that enable "scientist-guided" autonomous exploration of previously unseen locations. Because no algorithmic approach can currently match the nuanced decision-making of human geologists in selecting targets, we adopt a strategy to mitigate the limitations of autonomous science systems. The strategy recognizes that there are several common geological observations that can be aided and executed by an algorithmic approach, such as recognizing familiar rock types, identifying changes between rocks of different types, and noting rock textures such as layering.

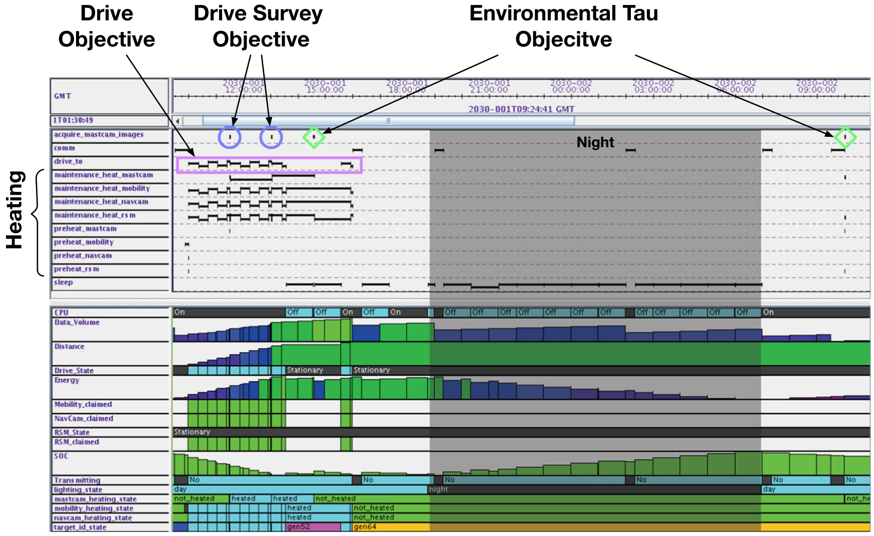


FIGURE 7 Example generated plan illustrating a long-range drive objective that was split up to support two different types of campaign objectives [Color figure can be viewed at wileyonlinelibrary.com]

We call our strategy "scientist-guided" to emphasize that we attempt to incorporate as much prior geologic/contextual knowledge as possible into the behavior of the algorithm. At the most abstract level, we parameterize autonomous science by (1) what type of targets to identify in a scene, and (2) how to measure the targets that have been identified. Using auxiliary information available during strategic planning, such as orbital imagery, scientists can decide what types of geologic features are likely to be at any given location. Then, scientists can decide the best strategy for measuring each of the features they expect to find (e.g., which instrument to use to take images or spectral observations). Each of these high-level directives correspond to a "class sampling" campaign intent (described in Section 4.1) that provides guidance, such as "take Mastcam images of layering in the light gray outcrop" or "take Laser-Induced Breakdown Spectroscopy (LIBS) measurements of the tan-colored outcrop." Finally, scientists can provide information to the planner to specify the relative priority or utility of performing each campaign.

5.2 | Novel tools for scientist-guided autonomy

We build upon existing work to develop new tools that scientists can use to specify (1) what they would like to measure and (2) how they would like to measure it. Starting with the first category, the types of targets scientist-guided autonomy can identify are individual geologic units, "contacts" between units (regions where two units are visibly adjacent), and layering within units. To identify individual units, we use the existing pixel-wise classification software TextureCam (Thompson et al., 2012). The generality of TextureCam allows it to be used for identifying everything from specific geologic units for surface missions to atmospheric features such as clouds seen from space during orbital missions (Chien et al., 2016). Many of the new technologies we propose use the output of TextureCam, a probability that each pixel belongs to a class of interest, as an input for processing and target selection. Training a TextureCam classifier requires that examples of a particular feature of interest have been seen previously.

A geologic contact between two units can reveal key information about the history of formation of the units as well as what occurred in the intervening time between when each of the units was deposited. Accordingly, we have developed a new algorithm called Finding Oriented Regions of Contact (FORC) that uses the output of TextureCam to determine where two particular geologic units are in contact with each other. For geologic units A and B, FORC works by first running TextureCam on board to produce probability maps $\mathbf{P}_A$ and $\mathbf{P}_B$ containing the probabilities that each pixel in an image $\mathbf{I}$ belongs to units A or B, respectively. These probabilities are combined to produce a "contact score" for each pixel in the image, which corresponds to the probability that the two units are adjacent at that pixel. After the contact score is derived, any number of algorithms (some of which are described below) can be used to select targets for observation based on this score. An example using the FORC algorithm is shown in Figure 8. At an area known as Marias Pass, the Curiosity rover encountered a contact between the lighter-toned Murray unit and darker-toned Stimson unit (Newsom et al., 2016). Using previously acquired Navcam images, we trained TextureCam to label each of these units. Then, we applied the trained models on an image of the contact from Sol 992. Using the output of TextureCam, the FORC algorithm was then used to produce a contact score, shown on the far right of Figure 8. The pixels with contact score exceeding 0.5 are highlighted. Although there are some small false-positive regions, the approach successfully identifies much of the visible contact region between the two units.

Layering or stratification is another essential feature used to geologically characterize a new area. Layered rocks can be formed by the deposition of materials via water, air, or other geological processes such as lava flows. Understanding the depositional environment of a rock can inform the conditions present at the time of its formation, which can give clues about past habitability. In addition to identifying regions with layering, it is also desirable to infer the *orientation* of layering within each region to inform follow-up measurement strategies. For example, it could be desirable to acquire LIBS measurements *across* the layers within a rock to determine how the depositional environment changed over time. Accordingly, the novel Fast Oriented-Layer Detector (FOLD) algorithm is designed to both detect and determine the orientation of layers within a scene. Figure 9 shows an example of FOLD operating in "sliding-window" mode to detect layering across a large field-of-view image. The image was taken by the MSL navigation camera, and shows a layered butte within the "Murray Buttes" region of Gale Crater. The lower image in the figure shows a heat map (white is the "hottest") of the signal-to-noise ratio (SNR) at each pixel within the image. Small black bars show the orientation of layering at each pixel. FOLD not only detects the layering within the butte, but also the layering of blocks that have fallen from their original locations.

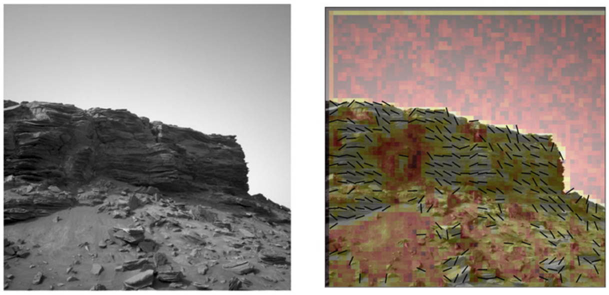


FIGURE 9 Left: MSL Navcam image of a butte in the Murray Buttes region of Gale Crater. Right: FOLD results, with hotter colors indicating higher SNR for layer detection and black bars showing the inferred orientation for the regions of the image with the highest SNR. SNR, signal-to-noise ratio [Color figure can be viewed at wileyonlinelibrary.com]

The second category of algorithms must take the output of TextureCam, FORC, or FOLD and determine how to measure the identified features of interest. The diverse onboard target selection (DOTS) algorithm addresses the scenario in which scientists are interested in acquiring "diverse" measurements of a geologic region. That is, if TextureCam detects several similarly relevant but disjoint geologic regions within a scene, scientists might prefer to take a measurement of each separate region rather than taking several measurements within one region. Such a measurement strategy allows scientists to assess the geochemical diversity of a scene rather than focusing measurements on one isolated area. DOTS is designed to take pixel-wise scores such as TextureCam probabilities or FORC contact scores and produce a set of specific point targets to be measured. DOTS works by greedily selecting a target that has the highest probability of belonging to a region of interest, then removing from consideration all points that can be reached from the target by a path that stays (with high probability) inside a contiguous patch of the same region. This process continues until the desired number of targets have been found, or no relevant geological regions remain unmeasured. An example of DOTS is shown in Figure 10.

Finally, it is common to take a series of remote sensing measurements in a line or grid to cover a region of interest either with a point-measurement tool like ChemCam LIBS or a camera like Mastcam. In both cases, enabling a rover to emulate ground-targeted behavior requires the capability to plan rasters and mosaics on board. We have developed the Onboard Raster And Mosaic Planner (OnRAMP) algorithm to provide this capability. As with DOTS, OnRAMP works by using the output of other software such as TextureCam to identify a feature or region of interest. Then, using a dynamic programming algorithm, OnRAMP plans a series of point measurements or image frames to construct a raster or mosaic that optimally measures the identified feature or region. In our field experiments, we use OnRAMP to plan image mosaics of features that are too large to fit in a single frame. These include mosaics of contact regions or layering within outcrop.

The tools described above are designed to enable scientists to mix and match capabilities to provide the appropriate guidance to autonomous science at each location of interest. For the purposes of our field experiments, to reduce planning complexity for the scientists, we create predefined "scripts" that the scientists can use more easily by only specifying a few key parameters. As an example, scientists might request "2–4 LIBS observations of light gray unit at location *A*" or "1–2 images of the contact between the dark gray and

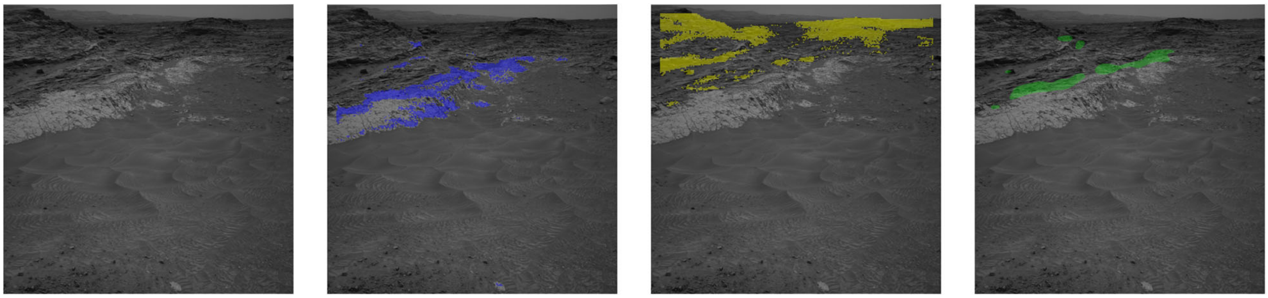


FIGURE 8 Left: The Murray–Stimson contact at Marias Pass. Center: The Murray (blue) and Stimson (yellow) units are identified using TextureCam. Right: FORC is used to derive a contact score, with the highest-valued regions highlighted (green) [Color figure can be viewed at wileyonlinelibrary.com]

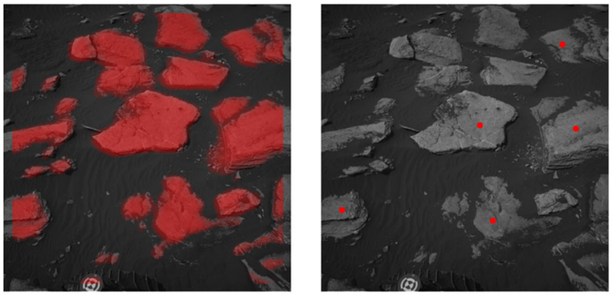


FIGURE 10 Left: MSL Navcam image with outcrop classifications (exceeding 50% confidence) from TextureCam in red. Right: diverse onboard target selection results, showing a set of point-measurement locations proposed to measure the identified outcrop [Color figure can be viewed at wileyonlinelibrary.com]

tan units at location *B*." Thus, scientists have a menu of options to choose from rather than needing to construct each combination of capabilities from scratch.

6 | SLIP-AWARE NAVIGATION

This section provides an overview of the slip-aware navigation system. The navigation systems equipped on the Mars rover missions, Mars Exploration Rover (MER) and MSL, rely on the grid-based estimation of surface traversability applied to local terrain (GESTALT) algorithm (Goldberg et al., 2002) to detect and avoid local geometric hazards and the **D*** algorithm (Stentz & Mellon, 1993) to plan global paths to goals. These methods have enabled operators to provide high-level autonomy goals to the rovers, increasing mission efficiency.

However, geometry alone is not sufficient to guarantee success in every Martian environment. Both MER and MSL operators have experienced hazardous conditions due to otherwise geometrically benign terrain such as sand dunes. These hazards can create adverse conditions such as wheel slip and sinkage. Excessive slip in terrain with little traction can cause rovers to become stuck. An example is the MER rover Spirit, which became stuck in a sand pit and eventually lost power due to its inability to tilt and orient its solar panels. As a result, when current Mars rovers pass through environments with hazards undetectable by the on-board nav systems, operators must resort to manual control with slow and deliberate commands, resulting in a loss in efficiency. In response, this paper proposes a navigation system that can reason about geometry *and* terrain type to plan safe reliable paths to science targets and enable a larger role in autonomy for future Mars Rovers.

The slip-aware navigation system, highlighted in Figure 11, is built upon the GESTALT system (Goldberg et al., 2002) and contains the following components: (i) stereo vision, (ii) visual odometry (VO), (iii) terrain classification, (iv) traversability assessment, (v) path planning, and (vi) execution. The input to system is a synchronized pair of stereo images from the rover's navigation cameras. Image data is processed by the OpenCV (Bradski, 2000) block matching algorithm to obtain a dense, 3D point cloud. Concurrently, the left stereo image is processed by the TextureCam (Thompson et al., 2012) terrain classifier to detect sand hazards. Further details on this classifier can be found in Section 5.2. Both texture and depth information are then processed by the Jet Propulsion Laboratory (JPL) VO method detailed in Howard (2008) to compute the relative motion between images. The dense stereo point cloud, sand hazard classification, and rover pose is incorporated into an occupancy grid map, and assessed for both geometric and terrain hazards in the traversability-assessment module. This map is furthermore labeled with a-priori terrain information to inform the slip-aware planner. Geometric Hazards are assessed and mapped using a modified version of the GESTALT (Goldberg et al., 2002) mapping method to account for geometry and terrain type. This map information is used by the RRT# sample-based planner (Arslan & Tsiotras, 2016) to plan safe paths around geometry- and terrain-based hazards that minimizes expected slip.

Our navigation system plans paths on a map that builds upon the data structure detailed in (Goldberg et al., 2002)—an occupancy-grid map fitted to a local ground plane with point-cloud statistics. The slip-aware navigation system improves on this map structure by adding terrain information for each point in the stereo point cloud. An example of this map structure is seen in Figure 12. Point clouds are accumulated to compute geometry and terrain statistics at each cell in the map. To assess the traversability of the map at each cell, a plane the size of the rover is centered and fitted to the containing points. Each cell in the map contains the following information: (i) maximum step-size, (ii) roughness, (iii) slope, and (iv) terrain information. Terrain information comes in the form of a discrete

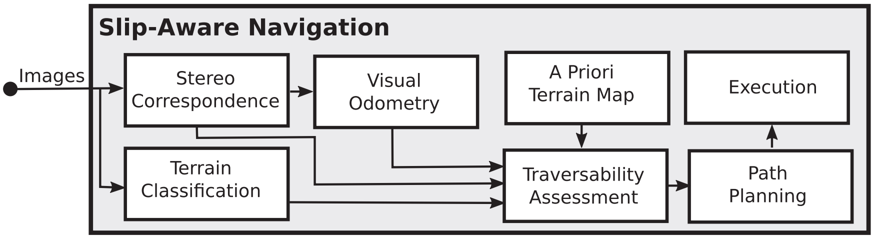


FIGURE 11 Illustration of the slip-aware navigation pipeline. This navigation system uses both geometry and texture from stereo images and a priori orbital maps, assess hazards to the rover, and plan safe paths in challenging environments with high-slip risks. This will allow rover operators to plan longer autonomous traverses in difficult terrain

probability distribution for the three terrain types of interest: soil, sand, and flagstone. This information can either come from an a priori map, with the use case being an orbital map labeled by human operators, or from the on-board terrain classifier. This information is used to compute a goodness metric, $\mathbf{g}(\cdot)$, for a given map cell, $\mathbf{c}$:

$$\mathbf{g}(\mathbf{c}_i) = \min(\mathbf{f}(\text{step}), \mathbf{f}(\text{roughness}), \mathbf{f}(\text{slope}), \mathbf{f}(\text{sand})), \tag{1}$$

where $\mathbf{f}(\cdot)$ computes the goodness of each parameter and returns a value between 0 and 1, where 0 is lethal and 1 is benign:

$$\mathbf{f}(x, \alpha, \beta) = \begin{cases} 1, & x \le \alpha, \\ \dfrac{\beta - x}{\beta - \alpha}, & \alpha \le x \le \beta, \\ 0, & x \ge \beta. \end{cases}$$

The parameters $\{\alpha, \beta\}$ refer to the minimum and maximum safe values for each goodness parameter. This goodness map, where each cell contains a goodness metric is thresholded to obtain a binary obstacle map with C-space expansion. Lethal obstacles can be seen in Figure 12 as pink areas with light blue borders.

The slip-aware navigation system plans safe paths that avoids geometric- and terrain-based hazards by employing the sample-based planner, RRT# (Arslan & Tsiotras, 2016) and the terrain-aware traversability map to make informed decisions on expected wheel slippage. The sample-based planner constructs a random graph where vertices contain robot poses and edges link poses by vehicle-constrained motion primitives (Pivtoraiko et al., 2009) in the form of pairs of fixed distance arcs. During planning, new vertices are considered as viable if they do not intersect with any geometric or terrain obstacles (sand) in the map. The cost of edges in the graph, and distances between vertices, is a function of the edge's motion primitive distance weighted by an expected slip profile for each terrain type. Terrain slip profiles map slope to expected rover slip for a given terrain type. This planner furthermore takes into account direction of travel when adding a new sample.

We experimentally validate the slip-aware navigation system with an isolated test in the northeast corner of the JPL Mars Yard. This Mars analogue environment consists of a 20 × 12 m area seen in Figure 13 and the top corner of Figure 20. This section of the Mars Yard consists of loose soil and bedrock, colored in Figure 13 as light and dark brown, respectively. Half of the slope consists of slippery, loose sand, and half of the slope consists of grippy bedrock. This information is feed to the system in the a priori map, with slip profiles for soil and bedrock at 17 degrees characterized as 50% and 0%, respectively. The test consisted of starting the rover at the bottom of a 17 deg slope (green circle), commanding the rover to drive to a waypoint at the base of the slope (black circle), and then to the final goal at the top of the slope (gold star). This was done with the slip-aware navigator, and the slip-unaware navigator, which only is identical to the previous but only minimizes arc distance when planning paths.

Results of the test are seen in Figure 13. The path planned for the slip-unaware planner is depicted by a red line, and the planned path for the slip-aware planner is depicted by a blue line. During the first traverse, from the rover start to the waypoint both planners plan very similar paths. The geometric obstacle, a large boulder is avoided to arrive at the waypoint. On the second leg of the traverse which takes the rover directly up the path, the slip-aware planner diverts the rover to the less slippery bedrock to reach the goal. This shows the planners ability to plan longer paths to avoid excessive slip.

7 | GLOBAL LOCALIZATION

Position knowledge of the rover is traditionally estimated using visual odometry and inertial measurements. Although visual odometry is reliable, position estimates accumulate drift errors on the order of 2% of the distance traveled. This poses a significant challenge for executing long autonomous drive plans due to the growth in uncertainty of the target waypoint, keep-out zones, and so forth.

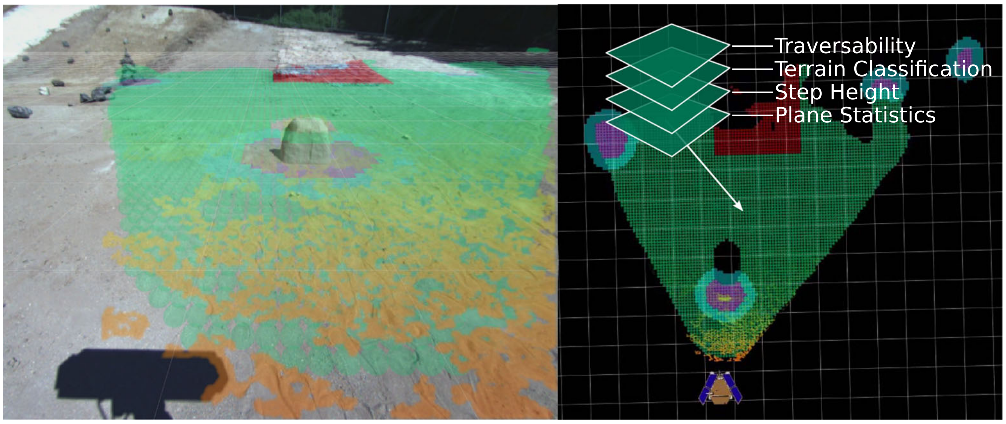


FIGURE 12 The traversability map used by the slip-aware planner. This 2D Grid map is based on the GESTALT traversability map (Goldberg et al., 2002) with the addition of labeled terrain information from either an a priori map or an onboard terrain classifier. Traversability is a function of underfoot plane statistics, maximum step tolerances, and terrain type. The figure in the map is colored by a priori terrain type, where green is loose soil, and red is embedded bedrock. Obstacles are colored as pink with light blue expansion zones [Color figure can be viewed at wileyonlinelibrary.com]

For MSL, this drift is corrected manually by rover operators using visual alignment of navcam imagery to orbital HiRISE imagery. To estimate this alignment, a mosaic of navcam stereo images are taken to cover a full panoramic around the rover. These images are then orthographically projected and salient surface features are manually tie-pointed to compute a correction offset (Figure 14).

The SRR employs both longer drives as well as multi-sol operations without the involvement of ground operators to perform localization corrections. To achieve this, a similar method of aligning navcam imagery to orbital maps is used in an automated manner onboard the rover. Instead of keypoint tie-pointing, the images are aligned using a match criteria on both the image intensities and the elevation map. Although position estimates accumulate drift, the attitude of the rover is known to high accuracy using the IMU and solar alignment. As a result, navcam stereo imagery can be accurately projected into an orthographic map and aligned in azimuth, leaving only the translation to be estimated.

The navcam images are projected with a spatial resolution of 25 cm/px to match orbital imagery from HiRISE. Similarly the elevation maps are projected to match the HiRISE DEMs at 1 m/px. Global position is estimated using a template match of both the surface imagery and elevation maps in the vicinity of the last VO estimate, using a search range bounded by the worst-case drift.

$$MI(X, Y) = \sum_{x \in \mathscr{X}} \sum_{y \in \mathscr{Y}} p_{XY}(x, y) \log \left(\frac{p_{XY}(x, y)}{p_X(x) p_Y(y)} \right). \tag{2}$$

The matching criteria for the imagery uses mutual-information, or relative entropy, between the images (Ansar & Matthies, 2009). This measures the statistical dependence between corresponding pixels of a candidate alignment. Mutual information is used instead of more conventional correlators such as sum of absolute differences or normalized cross correlation for robustness to differences in lighting or surface conditions when the orbital images were acquired. Mutual information is defined in Equation (2), where X and Y are the orthoprojected navcams and HiRISE images respectively, with pixel intensities x and y. The joint probability p_{XY} and marginals p_X and p_Y are estimated with histograms. The elevation map alignment uses a conventional sum of squared differences correlator. The overall matching score is a weighted sum between the image and elevation scores.

Conducting global localization experiments in the Mars Yard is problematic due to the large number of artificial structures, as well as the volatility of object movement within the yard. Instead, we experiment with navcam data taken from the MSL traverse between sols 1 and 2003, consisting of 6391 navcam images. The ground-truth trajectory is obtained from the manually aligned rover position at the end of each drive. The correction offset from this alignment is then interpolated backwards to the beginning of the drive to form a smooth path. The localization errors are shown in Figure 15. Experiments were conducted by localizing individual navcam images, and mosaics of all images captured from the same location, resulting in a mean error of 0.76 and 0.91 m respectively.

8 | MOBILITY HEALTH ASSESSMENT

Autonomous rover science is only practical if successful activity completion can be assured. The rover must respond appropriately to hazards and system failures, sacrificing science activities where necessary to preserve system health, but also reliably recognizing and recovering from routine interruptions, many of which currently require ground-in-the-loop resolution. Existing rover fault protection can be elaborate (MSL, for instance, has over 1000 distinct fault monitors) but is defensive in nature, whereas future rovers must also ensure operational efficiency. Simply expanding the scope of fault protection is unlikely to provide this additional capability.

Model-based reasoning offers one approach to compensate for failures, whether due to uncertainty in the environment, plan, or rover performance itself. The SRR prototype incorporates Model-based Off-Nominal State Identification and Detection (MONSID) (Kolcio & Fesq, 2016), which analyzes command and sensor data in

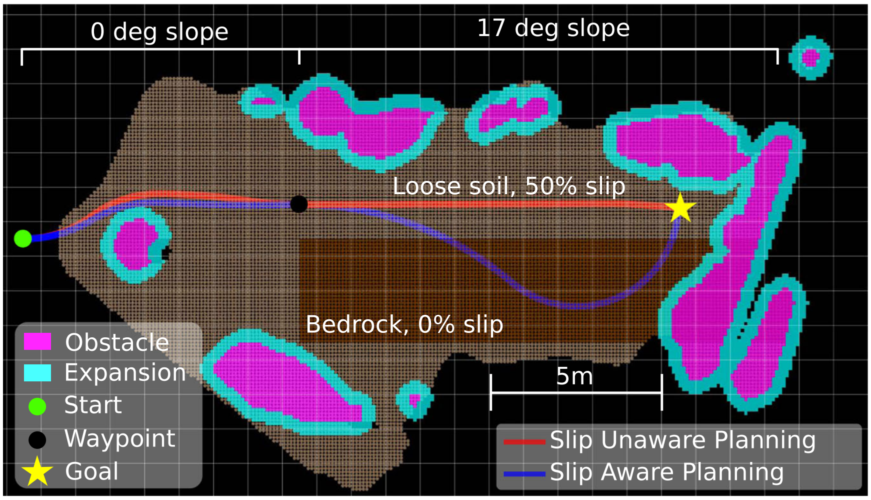


FIGURE 13 Slip-aware navigation results from a small test in the JPL Mars Yard [Color figure can be viewed at wileyonlinelibrary.com]

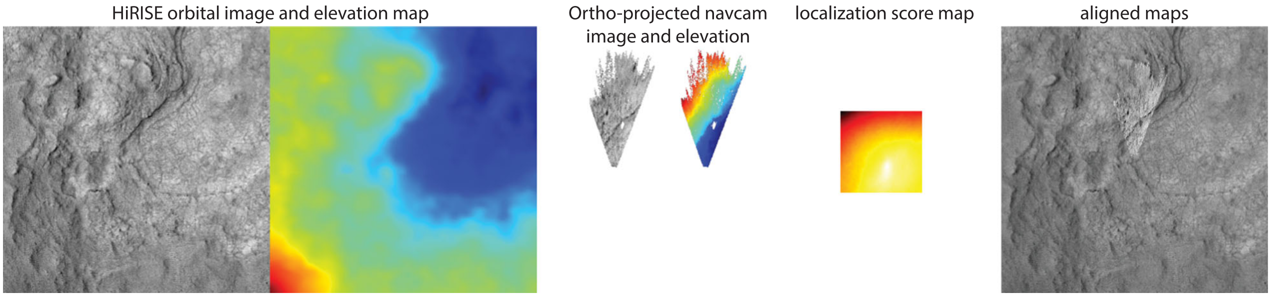


FIGURE 14 Global localization utilizes automated alignment of navcam image and elevation maps to onboard orbital maps [Color figure can be viewed at wileyonlinelibrary.com]

real-time, constructing an estimate of system health that is available to other autonomy components on-board. MONSID utilizes the constraint suspension technique (Fesq, 1993) to perform fault detection and diagnose likely causes of failure. MONSID consists of a diagnostic engine and a system model, the latter containing simplified physics models comprising a network of numerical constraints between sensed and internally computed parameters. The model also relates these constraints to physical or logical components of the host system, allowing inconsistencies to be linked to responsible components.

MONSID was used previously to model and diagnose rover electrical power subsystems (Kolcio et al., 2017), illustrating its suitability to rover-specific challenges. A second MONSID model was developed for integrated testing with the Self-Reliant Rover, concentrating on Athena mobility systems and associated sensors, and exercised with seeded faults before full-scale demonstration in simulated autonomous science scenarios (Kolcio et al., 2019).

The MONSID mobility model is summarized in Figure 16. Orange boxes represent rover components or pseudocomponents. Blue ovals indicate command or sensor values provided from the system. Data provided to components, and connections between components, flow through ports, indicated by green boxes. These connections represent transfers of state variable estimates into components or from one component to another. State variable estimates are tested within each component against mathematical constraints, which are reevaluated whenever new data arrives.

In this specific model, each of Athena's six wheels is associated with separate steering and drive motors, and each motor corresponds to a distinct MONSID component. Position encoders provide feedback to each controller. At the system level, Athena control software sends steer angle and drive distance commands to each wheel individually, coordinating the six wheels to follow a commanded path described by a series of arcs. The arc commands are the only external inputs to the MONSID chassis component, which represents the system-level controller.

As command and sensor data arrive, MONSID propagates new data through the model. In each component, evaluation of constraints produces one or more predictions about other state

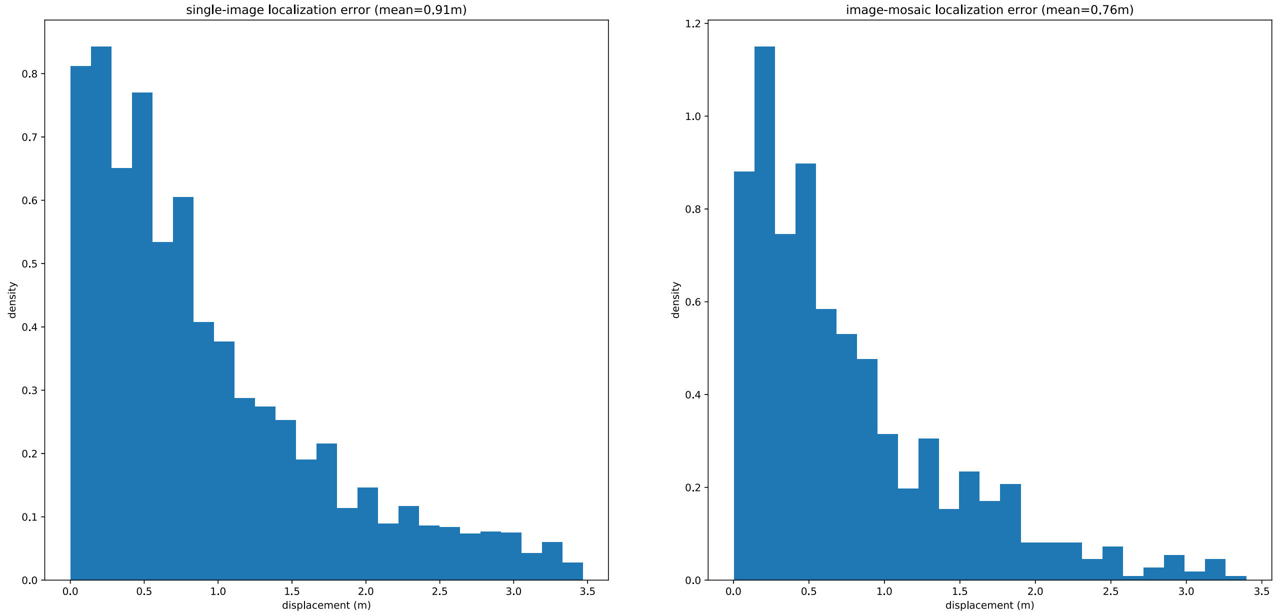


FIGURE 15 Error histograms for global localization on Mars Science Laboratory traverse using single navcam images (left) and navcam mosaics (right) [Color figure can be viewed at wileyonlinelibrary.com]

variables, which are then compared against sensor data or predictions provided from other components. Any significant discrepancy in these estimates implies that either the sensor data are wrong, or there has been a violation of our behavioral assumptions. This allows MONSID to detect failures both in rover software and hardware, whether due to command errors, mechanical or electrical faults, or terrain.

Model-based diagnosis thus addresses the following needs of the Self-Reliant Rover:

- Detect and classify recoverable mobility faults: Determine whether interruptions are terrain-induced or caused by mechanical failure, and whether the rover should autonomously retreat and avoid the problematic terrain.
- Sense errors in terrain knowledge: Determine whether trajectory deviations are caused by mechanical failure, sensor failure, or incorrect assessment of terrain. In the latter case, terrain knowledge may be recovered by comparison to alternate models of terrain characteristics.
- Recognize emergent,unknown,or surprise behavior: A significant autonomy challenge is the risk of unexpected behavior in response to novel or untested conditions. However, due to its reliance upon physical principles instead of purpose-built monitors, model-based health assessment can often detect and correctly classify novel system behavior, including problems arising from its own autonomy software.

MONSID model development included testing against Athena mobility data captured over a variety of terrain, in nominal operating condition and tests with injected system faults. Fault testing included hardware faults, such as simulated drive and steer motor failures, and faults more consistent with software or command errors such as steer motor polarity errors. Improper commanding is a fault case of particular interest to autonomous system developers because commands sourced on-board cannot be verified in the traditional manner, and these faults may be ignored by a traditional fault management system. However, where incorrect commands conflict with the expected system behavior, they can be detected and isolated in a systematic manner.

A typical result from a simulated command error is shown in Figure 17. Command errors affecting only a single wheel did not impede overall rover motion (and thus might go undetected for some time), but did result in unwanted drag with the potential to cause irreversible damage.

The MONSID system reacted to this test case by finding deviations in several mobility system components——not only was the sensed angle inconsistent with other steer angles, but the steering error propagated through several other constraints and tripped additional discrepancy alerts. Detection occurred before starting the forward drive, thus providing an opportunity to react to the fault before the rover suffered damage.

After detecting the fault, MONSID attempted to isolate a likely cause. In this case the features initially detected indicated one or more problems with either of the right corner wheels or with the

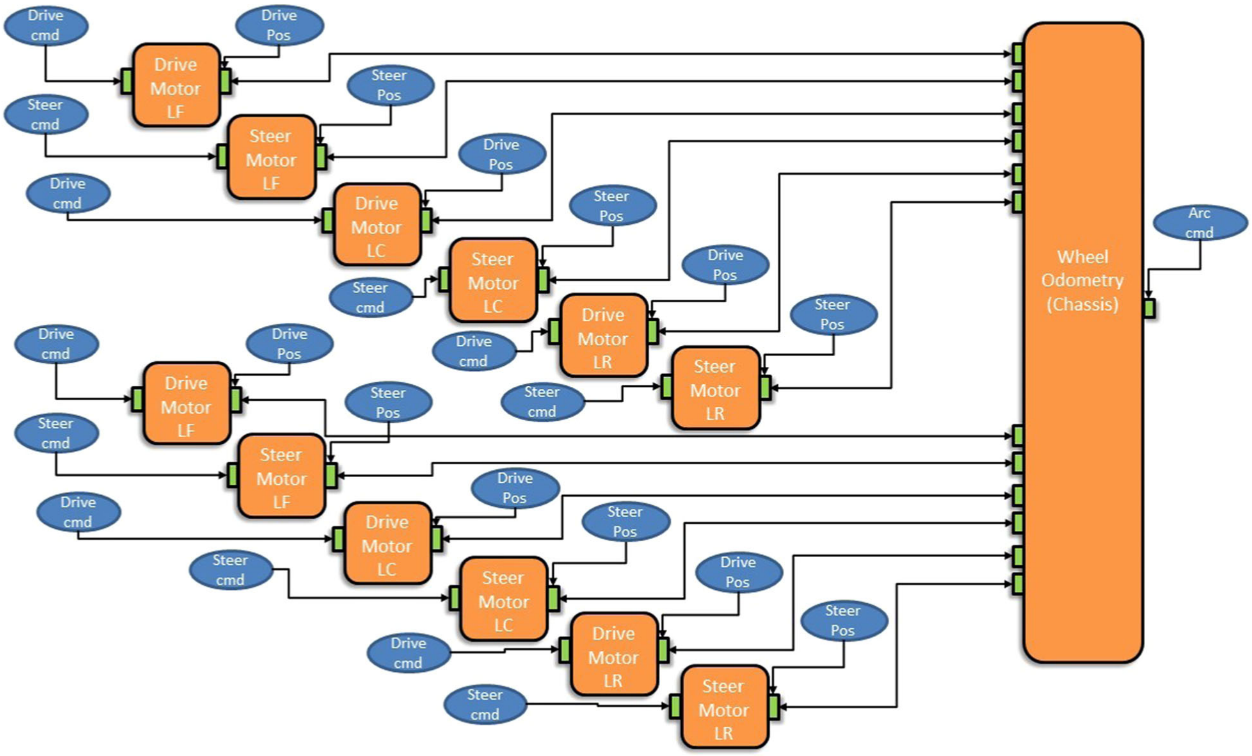


FIGURE 16 MONSID model of Athena mobility components [Color figure can be viewed at wileyonlinelibrary.com]

control software. However, as multiple constraints were violated, successive suspension of violated constraints quickly confirmed that failure of the right rear wheel is the correct solution. Algorithmically this is somewhat equivalent to traditional troubleshooting, but with the advantage of proceeding without predefined tests or checklists, relying instead only on the system model.

The most difficult task for health assessment is the ability to sense and react appropriately to previously unknown behavior. MONSID was given such an opportunity in early 2018: The Athena team was surprised when one of the front wheels rose from the surface of a steep slope while the other five wheels remained in contact, behavior thought impossible due to Athena's suspension geometry. The behavior was replicated and found to be caused by unusually high traction in a center wheel, coupled with slippage of both rear wheels. This led the center wheel to drive forward relative to the rover as a whole, rotating the bogie in the process.

This behavior is interesting because it only manifests at the system level——all individual rover components operate in familiar and acceptable ways. It is therefore unlikely that this behavior would be halted, and perhaps not even noticed, by a traditional fault protection system. The actual root cause is a violation of more fundamental assumptions about the rover, specifically a change in wheel geometry while on flat terrain. These assumptions are implicit in the MONSID constraints, and thus the novelty of this situation is detected without difficulty, despite having never been observed in previous years of testing and experience with Athena.

Mission operations would probably desire the Self-Reliant Rover to suspend further motion if this occurred, allowing controllers to thoroughly analyze the new behavior and check for previously unknown vulnerabilities. MONSID's responsibility in this case is simply to detect the event and classify it as a ground-recoverable fault. However, MONSID also provides diagnostic information to assist in event analysis, in this case isolating the fault to the center and rear wheels instead of any control fault——indeed, MONSID exonerates the wheel that actually rises from the surface. MONSID's conclusion is correct, despite initial appearances to the contrary.

9 | SYSTEM EVALUATION: MARS YARD WALKABOUT CAMPAIGN

We have developed a prototype implementation of the SRR approach using the Athena research rover. The Athena rover, shown in Figure 18, is roughly the size of the Spirit and Opportunity rovers from the Mars Exploration Rovers mission. Similar to those rovers, Athena has six wheels with rocker-bogie suspension with navigation cameras mounted on a pan/tilt mast. It should be noted that our system is still in a prototype stage in which not all bugs have been worked out. However, we wanted to conduct a fairly ambitious evaluation of the system in this relatively early stage to help guide future development. To facilitate this evaluation, we developed a "checkpointing" capability that allowed us to save a snapshot of the state of execution and resume execution from a previous checkpoint. This enabled us to restart execution from saved checkpoint if a problem was encountered without having to restart from the very beginning. As such, the results that follow represent a composite of executions.

FIGURE 17 A simulated software error leading to an incorrect steering angle [Color figure can be viewed at wileyonlinelibrary.com]

To evaluate the ability of the SRR approach to increase productivity we conducted a simulated campaign in which actual planetary scientists used our system to explore a geographical region. We selected a "walkabout" campaign for our evaluation. A walkabout is a reconnaissance campaign in which the rover makes an initial pass over the region of interest performing remote sensing observations. The operators use information collected from the walkabout to make an informed decision about which locations to revisit for more in-depth study including potential sampling operations. We selected a walkabout campaign for the evaluations because it has been found to be an effective means of exploring a region of interest (Yingst et al., 2017) and is anticipated to be used in the Mars 2020 mission to help identify sampling locations.

As mentioned in Section 7, evaluating global localization in the Mars Yard is problematic due to artificial structures. Further, including health assessment in the evaluations would have resulted in the artificial injection of faults which could perturb the results. As such, the global localization and health assessment components were not included in the walkabout evaluations. Instead, more targeted evaluations were performed on these components in as described in Sections 7 and 8.

9.1 | Objectives

Figure 19 summarizes the objectives of the evaluation. For the primary objective, we wanted to evaluate how well the SRR approach enabled the operators to maintain high levels of productivity with limited ground-in-the-loop interactions with the rover. As discussed in Section 2.3, the strong reliance ground-in-the-loop interactions was a major factor in limiting productivity in the MSL campaigns that we studied. As noted in Section 1, this reliance on ground-in-the-loop interactions is expected to become an increased liability with non-sun-synchronous relay orbiters.

The productivity metrics we used are based on the MSL case study we conducted in the first year of the project (Gaines et al.,

FIGURE 18 The Athena research rover [Color figure can be viewed at wileyonlinelibrary.com]

2016). These metrics relate to how long it takes to accomplish campaign objectives. The percentage of sols that contribute toward campaign objectives provides a measurement of how effectively each sol is used during the campaign. For our simulated campaigns, our participating scientists were the judge of whether the activities in a sol made “significant contributions” toward campaign objectives. The objectives of the campaign were to collect observations that characterize geographical areas. The scientists made an assessment as to whether the observations performed by the rover provided a satisfactory characterization of the area. To be considered a productivity sol, the sol must include activities to satisfactorily characterize at least one location or drive an efficient route to a target of interest to the scientists.

In the MSL campaigns we studied, we observed that almost half of the sols were not making significant contributions toward campaign objectives. We wanted to determine if the SRR approach could improve this measurement. The related metric, number of sols to complete objectives, is simply the number of sols required for the team to complete the simulated campaign. The number of locations surveyed is a rough measure of the quality of the campaign. Generally speaking, more locations surveyed will increase the scientists' understanding of the region enabling them to make more informed decisions about which locations to revisit.

9.2 | Methodology

For our region of interest, we constructed geological scenes in the JPL Mars Yard. Figure 20 shows the area we created. We simulated a larger area by applying an 8x scaling factor to the actual Mars Yard dimensions. This allowed us to simulate a longer-duration mission than would otherwise be feasible in the Mars Yard. We similarly scaled time between the simulated mission and actual rover activity to match realistic activity durations from MSL operations.

Three planetary scientists from the MSL mission participated in the evaluations. We prepared strategic guidance for the scientists similar to the guidance they would be provided for an actual campaign on Mars. This included the labeled imagery in Figure 20 showing units and features identified from “orbital” data. The team was also provided with the contextual information and strategic guidance for the campaign shown in Figure 21.

This included contextual background for the campaign, specific objectives and a high-level “sol path.” The sol path specifies how the week of sols will be broken down into planning sessions. This is primarily driven by the pattern of relay orbiter overflights that determine when ground-in-the-loop cycles are available. Because we are interested in the impact of non-sun-synchronous orbiters, we used a projected overflight pattern based on the MAVEN orbiter. This resulted in the sol path shown in Figure 21c in which the seven sols are broken up into three planning sessions consisting of two sols, three sols and two sols, respectively.

9.3 | Results

9.3.1 | Plan 1: Sols 33, 34

We met with the scientists to collect their objectives for the first phase of the walkabout campaign, the execution of Sols 33 and 34. This first session raised an interesting choice the scientists have for

Primary Objective: Evaluate ability of SRR approach to enable productivity for rover operations with reduced ground-in-the-loop interactions

- Productivity metrics:
    - Percentage of sols making significant contributions toward campaign
    - Number of sols to complete objectives
    - Number of locations surveyed during campaign

Secondary Objective: Collect feedback on use of Self-Reliant Rover technologies to support rover operations with reduced ground-in-the-loop interactions

FIGURE 19 Objectives of evaluation

interacting with the rover. They would be sending objectives to the rover which would then have two sols to accomplish these objectives. The scientists would then review the collected data and send objectives for the next three sols.

With current rover operations, the team requires ground-in-the-loop interaction To specify objectives. As a result, if the team were operating an MSL-style rover, they would provide only the objectives the rover could accomplish in the upcoming two sols.

In contrast, with the SRR approach, the scientists have the option of providing more objectives to the rover than the rover is expected to accomplish in two sols. There are a few advantages of over-subscribing the rover in this way. First, by giving the rover a longer-term view of the team's objectives, it can generate a higher-quality plan, though this comes at the cost of computational complexity. Second, if the rover is able to accomplish objectives more quickly, or with lower resource consumption, than expected, it will be able to work on these additional objectives. Third, though rare, there are occasionally problems with the uplink process (e.g., hardware problems at the DSN station) that result in loss of uplink. This is a major loss of productivity for traditional operations as the rover is left largely idle until operators are able to send command products at the next uplink opportunity, which is at least one more sol in the future. In contrast, if the team gave the SRR additional objectives, it can continue to pursue those, even if the uplink is missed.

In our first planning session, the scientists chose to come up with the complete set of objectives for the full walkabout campaign and provide these up front to the rover. They would still have the opportunity to revise objectives the rover has not yet accomplished in subsequent planning sessions, based on the data obtained so far. Figure 22 shows the initial set of objectives that the scientists provided. It is notable that it took the team only 1 h to provide the full set of objectives. Similar to actual operations, the scientists chose the names for target and location names by selecting collection of names centered around a theme. We opted for a light-hearted theme with a collection of Muppet names. The team made effective use of the names with some of their assignments such as assigning Muppets with similar coloring to the locations, such as Fozzie for a tan unit and Pepe for an area with reddish rocks.

The scientists objectives included manually-targeted observations (Figure 22b) in the area for which they had imagery from the rover along with guidance for autonomous science (Figure 22c) in the areas they have not yet seen. Manually targeting observations represent the predominant means that operators interact with an MSL-style rover. It has the advantage that scientists can specify the exact targets for which they wish to collect data but has the disadvantage that it requires a ground-in-the-loop interaction. The team must receive imagery from the rover's location so that the team can select targets and the rover must remain in that location until it receives the manually targeted observations from the team. By providing a means for scientists to guide autonomous science, we are attempting to allow operators to direct the rover's autonomous selection of targets with a reduced reliance on ground-in-the-loop interactions. However, the system still supports operator-targeted observations to take full advantage of the scientists' input when the team is able to exploit a ground-in-the-loop interaction.

Another design decision we needed to make was how to solicit priority specifications from the operators. While our planner supports a continuous range of priorities, providing such granularity can be problematic for operations. A large amount of time can be lost by the team trying to assign such precise priority values. Instead, we opted for a simple interface in which the team would select from just two priority levels: Normal and High.

We provided the initial set of objectives to the Athena rover for execution of the first two sols of the walkabout. The execution of Sols 33 and 34 took less time that modeled allowing the rover to arrive at Fozzie with enough time to complete the survey at this location before the end of sol 34. This was expected behavior as we

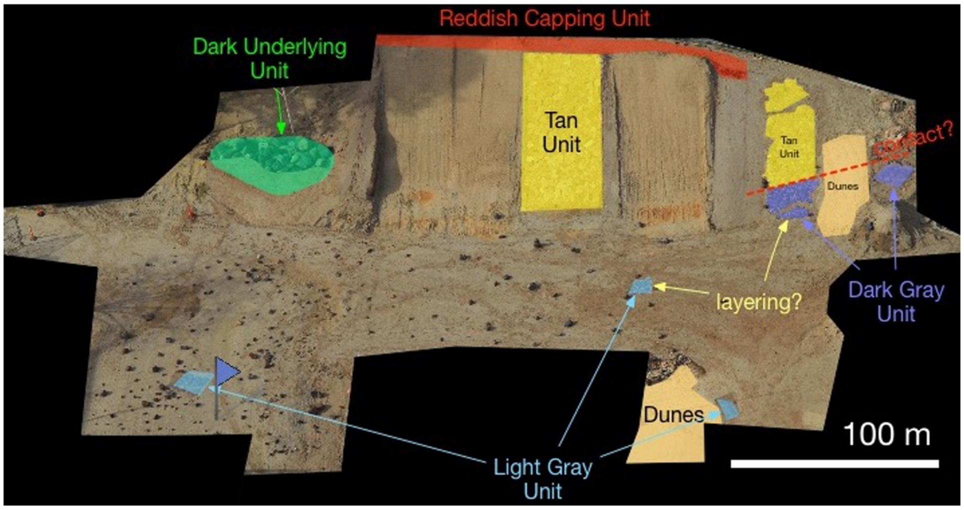


FIGURE 20 Aerial view with labels from the simulated Strategic Science Working Group [Color figure can be viewed at wileyonlinelibrary.com]

(a)

The rover has completed a long drive to this new region of interest, where several distinct units have been identified from orbit. In past sols we've seen examples of similar-looking materials in isolation, but we are now approaching an area where several of these units are co-located, possibly in contact with each other. Our strategic goals for this area are to characterize the diversity of units in this locality and to understand their relationships. The team agreed on a campaign that uses multiple passes through the area. The first pass will acquire remote science to accomplish the above goals and to select targets to return to in a second pass for contact science and sampling.

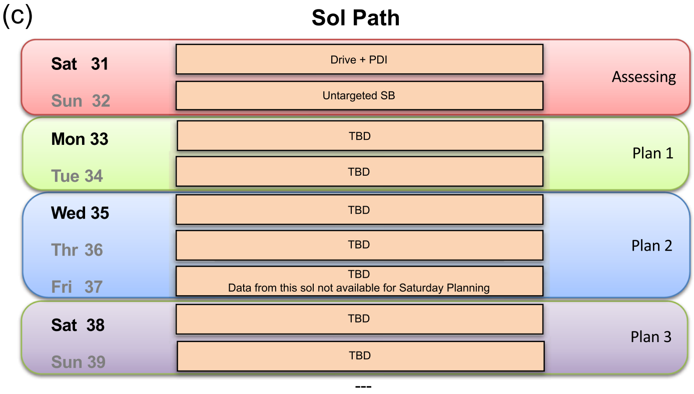
Context for campaign.

(b)

Conduct a walkabout of the Mars Yard

- We are making an initial pass through this area. We will return to perform sampling.
- Use the upcoming week of sols to efficiently collect diagnostic data you will need to select sampling locations
  - Perform triage of the field area to identify interesting locations for further investigation
  - Acquire Mastcam observations of potentially interesting locations

Objectives of walkabout



Sol Path

Sat 31	Drive + PDI	Assessing
Sun 32	Untargeted SB	
Mon 33	TBD	Plan 1
Tue 34	TBD	
Wed 35	TBD	Plan 2
Thr 36	TBD	
Fri 37	TBD Data from this sol not available for Saturday Planning	
Sat 38	TBD	Plan 3
Sun 39	TBD	

High-level sol path for the campaign.

FIGURE 21 Strategic objectives for the Mars Yard walkabout campaign. (a) Context for campaign. (b) Objectives of walkabout. (c) High-level sol path for the campaign [Color figure can be viewed at wileyonlinelibrary.com]

intended to slightly over-estimate the durations of rover activity. This was done as it is generally less disruptive to replan and take advantage of surplus time than to replan to resolve conflicts that arise when activities take longer than expected.

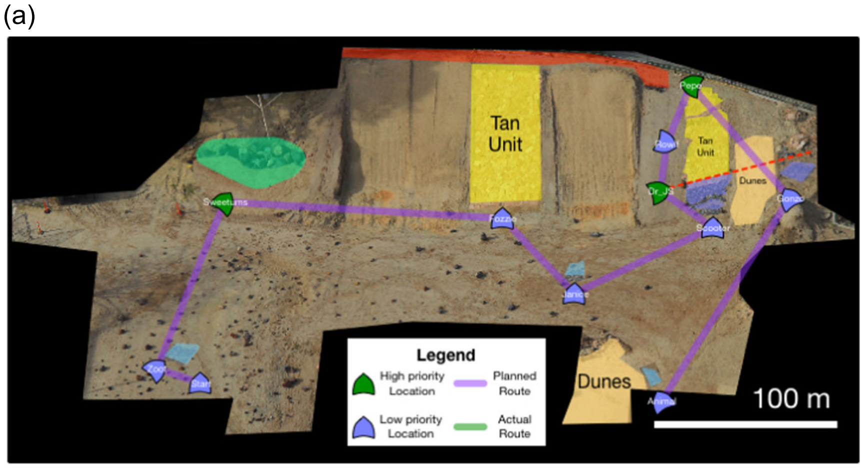
Figure 23 shows the results of autonomous science for the first two sols. The figures show the targets selected for each detector selected by the scientists along with the confidence of the classification. In general, the system did a good job of identifying targets. In the cases that there were examples of a given class, the rover picked out good targets such as dark outcrop at Zoot, dark rock at Sweetums and tan outcrop and reddish rock at Fozzie. Note that the detectors are not perfect and included some miss-classifications. For example, The edge of a light gray outcrop was falsely classified as dark outcrop at Zoot, soil was incorrectly classified as tan outcrop at Fozzie (but with relatively low confidence) and the shadow of a rock was classified as a reddish rock at Fozzie. Also at Fozzie, the detectors missed some examples of light red rock. This was due to the training set used for reddish rocks consisting of dark red rocks. While additional training would likely improve these results, we do not expect to achieve perfect classification with these autonomous science detectors and were pleased to have some false examples to present to the scientists to illustrate the range of behavior the system can exhibit.

The Zoot results show another example classification error in which the rover's shadow was mistaken for dark outcrop. In practice it is not difficult to filter out the rover's shadow. The flight software on the actual mars rovers tracks the sun position and could be updated to know where in a given image its own shadow falls. This would allow filtering out targets that are covered by the rover's shadow.

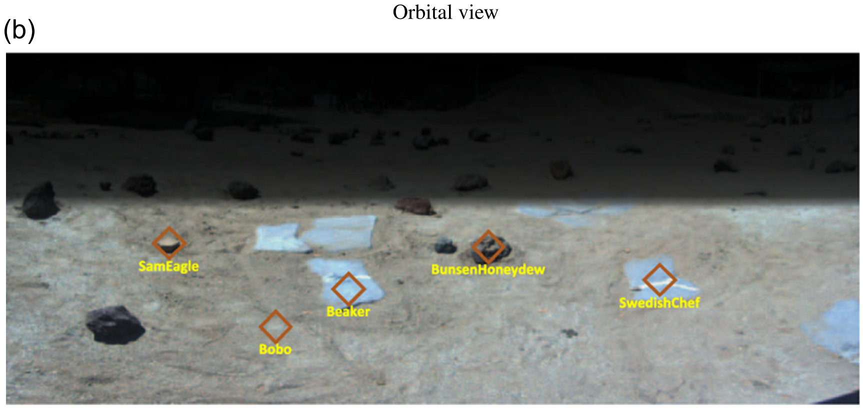
9.3.2 | Plan 2: Sols 35, 36, 37

We met with the scientists and presented the results of the first plan's execution. We asked if they wanted to provide new objectives or change any of the objectives already provided to the rover but not yet accomplished. The team considered whether they wanted to return to any previously visited areas to perform additional observations but decided that the rover had collected a sufficient amount of data at those locations to meet the walkabout objectives and preferred that the rover continue on to new areas.

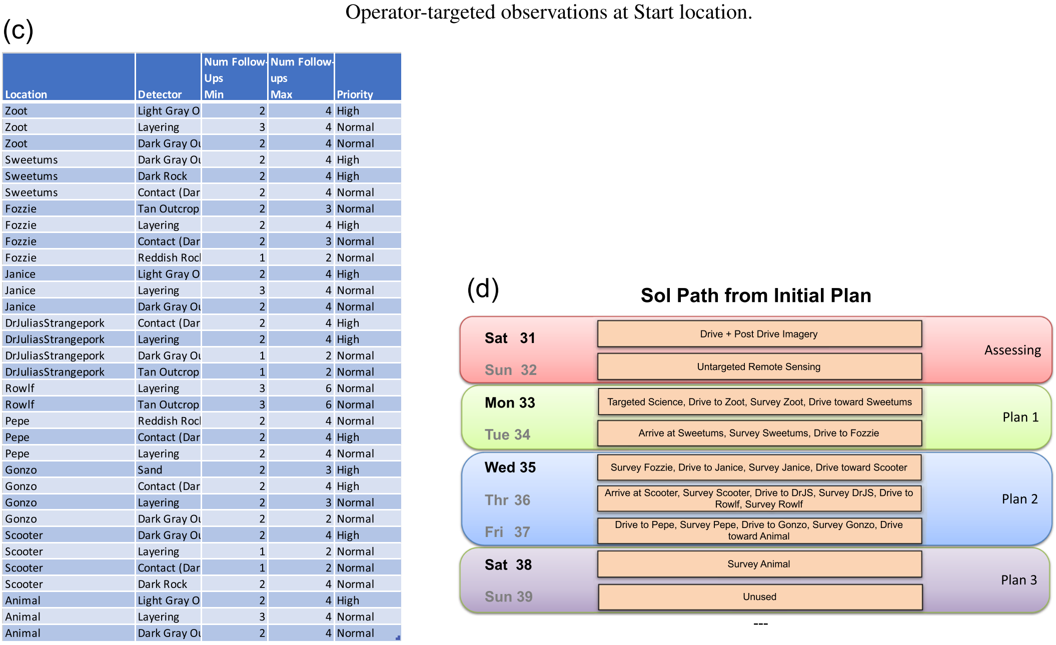
They did, however, make some adjustments to the autonomous science guidance they had provided in the previous plan. Most notably was removal of contact detection from some locations that did not have a noticeable contact from the orbital imagery. This decision was the result of the scientists learning about the capability of the contact detector from the example results at Sweetums and Fozzie. Their intent for the contact detector was to identify visual variations within a scene, such as a transition from lighter to darker gray in outcrop. However, the contact detector we have developed at this point sometimes struggles to detect contacts where the color variation is subtle. As such, the scientists opted to reserve its use for locations expected to have a stronger contact. This resulted in removing contact detection at Gonzo and Scooter. They considered removing contact detection from Pepe but decided to retain it to see if there was a contact between reddish rock and tan outcrop at this location. This was a useful interaction as it points to the value of developing a detector that can identify such variations within a scene and propose follow-up observations to document the changes.



Orbital view



Operator-targeted observations at Start location.



Location	Detector	Num Follow-Ups Min	Num Follow-ups Max	Priority
Zoot	Light Gray O	2	4	High
Zoot	Layering	3	4	Normal
Zoot	Dark Gray Ou	2	4	Normal
Sweetums	Dark Gray Ou	2	4	High
Sweetums	Dark Rock	2	4	High
Sweetums	Contact (Dar	2	4	Normal
Fozzie	Tan Outcrop	2	3	Normal
Fozzie	Layering	2	4	High
Fozzie	Contact (Dar	2	3	Normal
Fozzie	Reddish Rocl	1	2	Normal
Janice	Light Gray O	2	4	High
Janice	Layering	3	4	Normal
Janice	Dark Gray Ou	2	4	Normal
DrJuliasStrangepork	Contact (Dar	2	4	High
DrJuliasStrangepork	Layering	2	4	High
DrJuliasStrangepork	Dark Gray Ou	1	2	Normal
DrJuliasStrangepork	Tan Outcrop	1	2	Normal
Rowlf	Layering	3	6	Normal
Rowlf	Tan Outcrop	3	6	Normal
Pepe	Reddish Rocl	2	4	Normal
Pepe	Contact (Dar	2	4	High
Pepe	Layering	2	4	Normal
Gonzo	Sand	2	3	High
Gonzo	Contact (Dar	2	4	High
Gonzo	Layering	2	3	Normal
Gonzo	Dark Gray Ou	2	2	Normal
Scooter	Dark Gray Ou	2	4	High
Scooter	Layering	1	2	Normal
Scooter	Contact (Dar	1	2	Normal
Scooter	Dark Rock	2	4	Normal
Animal	Light Gray O	2	4	High
Animal	Layering	3	4	Normal
Animal	Dark Gray Ou	2	4	Normal

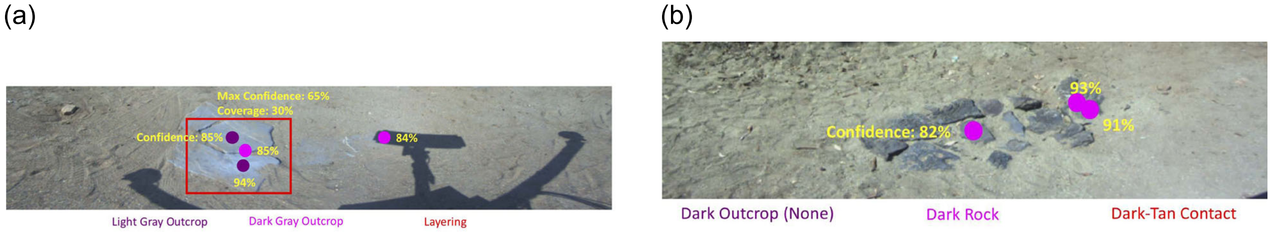
Autonomous science guidance

Sol Path from Initial Plan

Sat 31	Drive + Post Drive Imagery	Assessing
Sun 32	Untargeted Remote Sensing	
Mon 33	Targeted Science, Drive to Zoot, Survey Zoot, Drive toward Sweetums	Plan 1
Tue 34	Arrive at Sweetums, Survey Sweetums, Drive to Fozzie	
Wed 35	Survey Fozzie, Drive to Janice, Survey Janice, Drive toward Scooter	Plan 2
Thr 36	Arrive at Scooter, Survey Scooter, Drive to DrJS, Survey DrJS, Drive to Rowlf, Survey Rowlf	
Fri 37	Drive to Pepe, Survey Pepe, Drive to Gonzo, Survey Gonzo, Drive toward Animal	
Sat 38	Survey Animal	Plan 3
Sun 39	Unused	

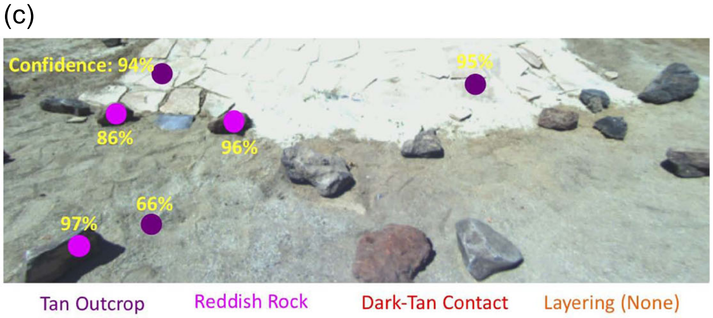
High level sol path

FIGURE 22 Objectives and expected plan for Sols 33, 34. (a) Orbital view. (b) Operator-targeted observations at Start location. (c) Autonomous science guidance, (d) High level sol path [Color figure can be viewed at wileyonlinelibrary.com]



Autonomous science at Zoot

Autonomous science at Sweetums



Autonomous science at Fozzie

FIGURE 23 Autonomous science results from Sols 33, 34. (a) Autonomous science at Zoot. (b) Autonomous science at Sweetums. (c) Autonomous science at Fozzie [Color figure can be viewed at wileyonlinelibrary.com]

This second session with the scientists provided another interesting decision point for the scientists. The rover had already performed autonomous science at Fozzie, with the results available to the team, and the rover would still be at this location when it receives the next uplink. This provided the scientists to select their own targeted observations at Fozzie before proceeding on with the rest of the walkabout. Figure 24 shows the targets the scientists chose, in diamonds, along with the targets already acquired by the rover. The result demonstrates the value of combining autonomously targeted and scientist targeted observations. The scientists benefited from the rover already performing some observations as they could spend less time at this location to complete the survey. The scientists chose two additional observations of tan outcrop with leveraged the two observations already performed by the rover. The scientists chose to space their observations a uniform distance from the rover's selection enabling them to measure the lateral heterogeneity of the outcrop. They also selected one of the light reddish rocks that was missed by the rover.

The execution of the second plan highlighted the use of the slip-aware navigator. While traveling to and from the Gonzo target, the rover was able to detect and maneuver around sand hazards. These drives would have required additional ground-in-the-loop cycles with current operations because current rovers are unable to identify sand hazards on their own and orbital imagery often has insufficient resolution to identify sand hazards ahead of time.

Figure 25 shows some of the autonomous science results from the second plan. As with the first plan, autonomous science generally did a good job of identifying targets when there were examples of the intended class in the scene. The scientists were particularly happy with the results at Dr. Julias Strangepork, Figure 25a. This was one of the most important locations as this is where orbital imagery showed that the tan and gray units would be in contact. The rover was able to identify examples of tan and gray outcrop and successfully identified and documented the contact between the units. In addition, the rover identified layering in the scene. This was an important observation as it enabled the scientists to confirm the orbital impression from Figure 20.

This location also raised an important lesson for our future work. While the rover correctly identified examples of the tan unit, the locations it selected for follow-up were tightly clustered. The scientists indicated that they would prefer these targets be better distributed across the outcrop. This was not a problem with the underlying tan outcrop detector. As shown in Figure 26 it did a good job of identifying almost all of the tan outcrop. The problem is that we choose follow-ups by picking the contiguous regions of the image with highest average confidence. This can result in clustering results. Instead, we learned that we should also take into account spatial distribution when selecting targets.

9.3.3 | Plan 3: Sols 38, 39

The rover was traveling toward the final location, Animal, at the end of Sol 37 but had not yet reached the target when it had to stop for the night. The expected plan for Sol 38 was for the rover to complete the drive to Animal and perform the requested survey of the location.

The scientists decided to let the rover complete the objectives without modification. And even though there was another sol available for them to use for execution, they decided that the rover

FIGURE 24 Operator-targeted observations at Fozzie location [Color figure can be viewed at wileyonlinelibrary.com]

had sufficiently surveyed the region to meet the objectives of the walkabout. Rather than surveying additional locations, they decided to conclude the first pass of the walkabout.

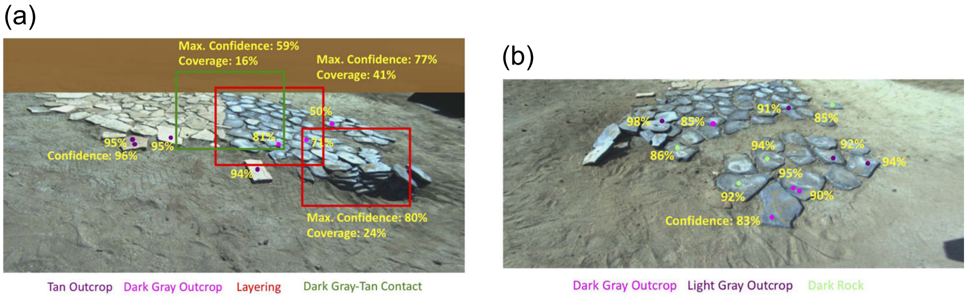
9.3.4 | Discussion

Following the field test, we met with the scientists for a final session to evaluate the results of the walkabout campaign. To determine how well the rover performed in the walkabout we asked the scientists to review the results of each visited location and assess if the location had been sufficiently surveyed to meet the campaign objectives. While there were cases where the scientists would have selected different observations from those selected by the rover, they concluded that each location had been sufficiently surveyed. Further, they concluded, given the locations that were visited, that the walkabout had successfully achieved the strategic objective of surveying the Mars Yard region.

As a means of evaluating the productivity improvement of the SRR approach, we performed an analysis to determine the best-case execution of the campaign using an MSL-style rover. The analysis was based on the ground-in-the-loop decisions that would be required by an MSL-rover, limiting the amount and type of activity that can be performed each sol. For each ground-in-the-loop cycle, an MSL-style rover would be limited to surveying at most one location, due the reliance on operators selecting observations, and driving to the next location. For the two drive segments that required the rover to avoid sand, an MSL-style rover would require additional ground-in-the-loop cycles as it would rely on human operators to identify the sand hazards and plan paths around them.

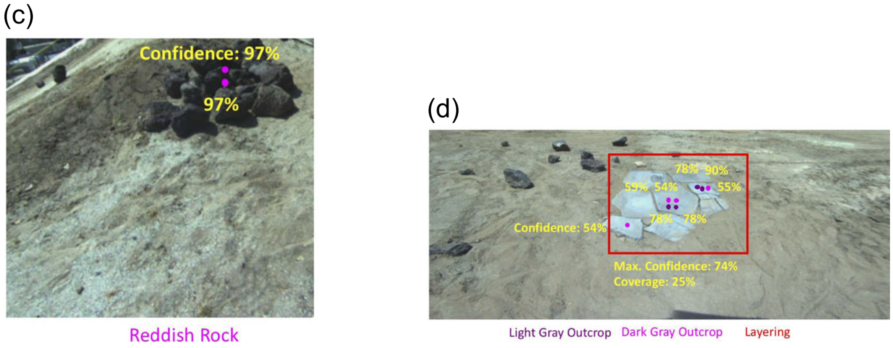
Note that this analysis represents a best-case scenario as it did not incorporate factors such as drive faults, operations procedure times or mission risks and policies. Such factors would result in reducing the productivity measurements of this comparison mission. As such, by excluding these factors, we are making a more conservative assessment of the SRR approach versus this best-case MSL campaign.

Figure 27 provides quantitative measurements of the productivity improvements achieved by the SRR approach. In Figure 27a we provide a breakdown of sol productivity similar to the ones we performed for the MSL case study in Section 2.3.1. The MSL approach was projected to have 32% low productivity sols due to the need to wait for ground-in-the-loop for performing location surveys and to guide the rover around areas with sand hazards. In contrast, the SRR approach was able to make significant progress toward campaign objectives on each sol. This results in SRR achieving a 47% increase in productive sols.



Autonomous science at Dr. Julias Strangepork

Autonomous science at Scooter



Autonomous science at Pepe

Autonomous science at Janice

FIGURE 25 Autonomous science results from Sols 35, 36, 37. (a) Autonomous science at Dr. Julias Strangepork. (b) Autonomous science at Scooter. (c) Autonomous science at Pepe. (d) Autonomous science at Janice [Color figure can be viewed at wileyonlinelibrary.com]

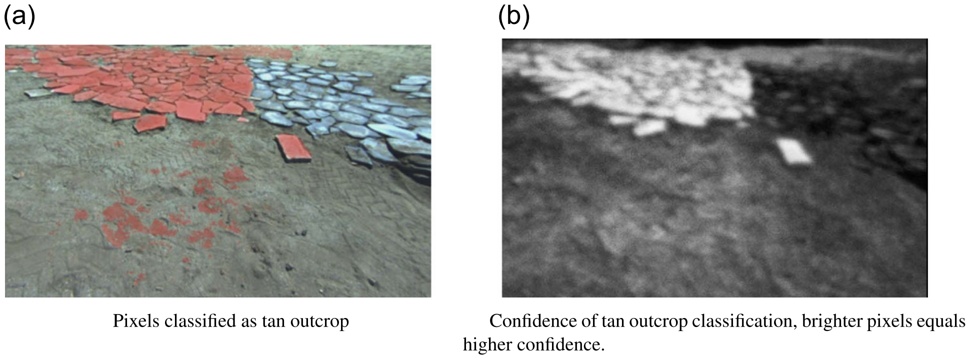


FIGURE 26 Detection of tan outcrop at Dr. Julias Strangepork. (a) Pixels classified as tan outcrop. (b) Confidence of tan outcrop classification, brighter pixels equals higher confidence [Color figure can be viewed at wileyonlinelibrary.com]

Figure 27b compares the number of sols required to survey all of the locations selected by the scientists. An MSL-style rover would require 28 sols to perform the survey while the SRR completed the campaign in six sols. This represents an 80% reduction in sols to complete the campaign.

Finally, in Figure 27c we compare the number of locations the rovers would be able to survey if the project chose to restrict the walkabout to a single week. An MSL-style rover would survey only three locations while the SRR surveyed 11 locations. This is a 267% increase in number of locations surveyed.

Overall, the walkabout campaign demonstrated that the SRR approach is able to maintain high levels of productivity with limited ground-in-the-loop cycles. The approach provided mechanisms that allow the scientists to effectively guide the rover's behavior despite the limited communication opportunities.

The key to the increase in productivity was the ability of the SRR approach to conduct effective characterization activities with reduced ground-in-the-loop dependency. This was achieved with a combination of several components in our system. The autonomous science component allows the rover to analyze geographical scenes and select targets for follow-on observations guided by scientist input. The Goal Planner allows the activity plan to be dynamically updated to safely perform the observations selected by the Autonomous Science component and reserve resources for subsequent scientist objectives. The Navigation system was needed to safely maneuver around areas of loose sand that would require ground intervention with currently Mars rover capabilities.

10 | RELATED WORK

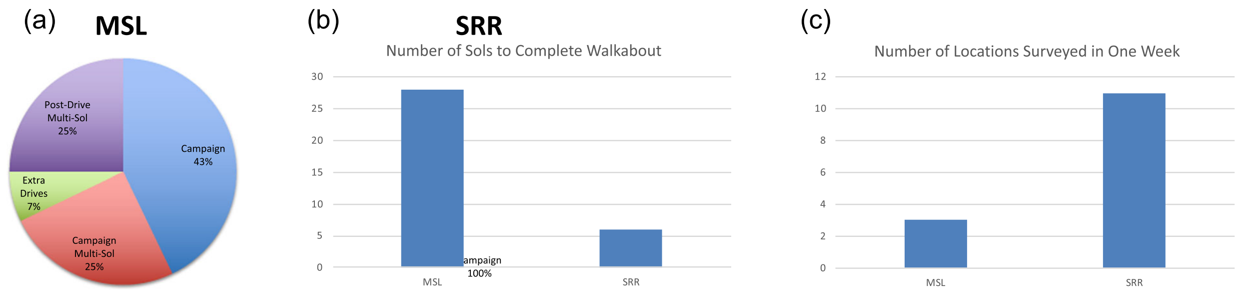
Shalin, Wales, and Bass conducted a study of Mars Exploration Rovers operations to design a framework for expressing the intent for observations requested by the science teams (Shalin et al., 2005). Their focus was the use of intent to coordinate planning among human operators and the resulting intent was not captured in a manner that would be conducive for machine interpretation. Our approach codifies some of the fields in their framework in a way suitable for the rover. In particular, the authors defined a "related observations" field as a way for scientists to identify relationships among different observations, which need not be in the same plan. Our work on campaign intent can be seen as a way of defining a specific semantics to these types of relationships to facilitate reasoning about these relationships by the rover.

Their framework also includes information that we agree is essential for effective communication among operators but that we do not currently express to the rover. For example, the "scientific hypotheses" field is used to indicate what high-level campaign objective is being accomplished by the requested observation. We are not yet providing these higher-level campaign objectives to the rover, though it is an interesting area of future research.

Mali views intent as a means for a user to place constraints on the types of plans a planner is allowed to produce such as only generating plans that have at most one instance of a class of actions or that plans must limit the use of a particular action (Mali, 2016). The primary role of our use of intent is to allow the planner to assess the value of achieving a given set of goals. However, some of our campaign intent does imply constraints and preferences on how, or more specifically, when goals are accomplished. For example, the periodic campaign intent specifies a timing relationship among goals and a preference on how close to comply with the desired timing.

There are some similarities between our campaign definitions and those used for Rosetta science planning (Chien et al., 2015). Both use campaigns to express requests for variable-sized groups of observations with relationships and priorities. Rosetta plans covered much longer time periods (e.g., weeks) and required more complex temporal patterns, such as repeating groups of observations. But observation patterns were primarily driven by the predictable trajectory of the spacecraft, allowing relationships to be expressed as temporal constraints. This is not sufficient for rovers, where many observations are dictated by the rover location and surrounding terrain, and the duration of many activities cannot be accurately predicted. State-based and goal set relationships more accurately represent some of the science intent found on surface missions.

There have been a variety of autonomous science systems deployed or proposed for rovers including the AEGIS system running on



Sol productivity. Number of sols to complete campaign. Locations surveyed in campaign.

FIGURE 27 Quantitative evaluation of SRR productivity. (a) For brevity, SRR pie chart not shown as it is 100% productivity, resulting in SRR yielding 32% increase in sol productivity. (b) SRR achieved 80% reduction in sols to complete campaign. (c) SRR achieved 267% increase in locations surveyed in one week. (a) Sol productivity. (b) Number of sols to complete campaign. (c) Locations surveyed in campaign. SRR, Self-Reliant Rover [Color figure can be viewed at wileyonlinelibrary.com]

the opportunity and curiosity rovers (Francis et al., 2017), and the SARA component proposed for an ExoMars rover (Woods et al., 2009). These systems allow the rover to identify targets in its surroundings that match scientist-provided criteria. The introduction of campaign relationships broadens the scope of the type of guidance that scientists can provide these systems, allowing scientists to express the amount of observations they would like for their different objectives along with the relative priorities of the high-level objectives.

There have been several integrated rover systems with similar objectives to our work including PRoViScout (Paar et al., 2012), Zoe (Wettergreen et al., 2014), and OASIS (Castano et al., 2007). The PRoViScout project has similar objectives to our work (Paar et al., 2012). These systems include autonomous science capabilities to enable onboard identification of science targets. Similar to our approach, they select follow-up observations for identified targets and submits these requests to an onboard planner to determine if there are sufficient resources to accomplish these new objectives. The campaign intent concepts we have developed would also be applicable to PRoViScout as a way to increase the expressivity for providing scientist intent to the rover.

There is an active area of research in intent recognition (Sukthankar et al., 2014). The general goal of this area is to identify the objectives of other agents (human or otherwise) from observations of the agents' actions. In contrast, in our work, it is acceptable for users to explicitly identify their intent, rather than require the system to attempt to infer intent. Indeed, there is interest in the operations team to clearly document their intent for the purpose of communication among teams and as a record of what activity was planned for the rover and why. As such, rather than try to infer user intent, our objective is to increase the expressivity of the rover's interface to more closely reflect mission intent.

The Mars 2020 mission is planning to incorporate onboard scheduling to improve resource utilization of the rover (Rabideau & Benowitz, 2017). Similar to the SRR approach, the use of onboard scheduling is intended to allow the Mars 2020 rover to use current vehicle knowledge when generating schedules to accomplish mission objectives. This will reduce the loss of productivity that results from the difficulty in predicting how much resources (e.g., time and energy) activities will consume. The SRR approach is addressing additional productivity challenges by improving the ability of rovers to identify their own objectives, to incorporate a richer set of guidance from operators and to reason about slip hazards as it navigates. The SRR and Mars 2020 planner face similar planning challenges including integrating scheduling and execution (Chi et al., 2018), the need for ground tuning of parameters to address onboard search limitations (Chi et al., 2019), and the difficulty of managing complex rover resources (Chi et al., 2020).

The navigation system presented in this paper is most similar to the system presented in (Helmick et al., 2009). They propose a system with the same high-level machinery: (i) a GESTALT-based vision pipeline, (ii) a terrain classifier, and (iii) a slip-aware planner. However, their system is not capable of making decisions based on direction of travel. When direction of travel is not considered, then the system is forced to make more conservative plans. An example is if the rover is planning a path on a steep slope containing soil, it might be too dangerous to drive up the slope due to expected slippage, but driving downhill would be safe.

Model-based diagnostics are being actively studied in other space contexts where traditional fault protection and spacecraft safing is insufficient, notably crewed spaceflight. There are many efforts to mature similar technologies in NASA Exploration, such as in Aaseng et al. (2015). Additional work is ongoing to mature model-based diagnostics with particular focus on verification and validation of model-based systems, including an effort related to this development as described in Nikora et al. (2018).

11 | CONCLUSIONS

We have presented an approach for increasing the authority of autonomous rovers to increase mission productivity. Our approach includes the ability for ground operators to provide guidance to the system without requiring up to date knowledge of the rover's state and its surroundings.

We have implemented a prototype of this approach on the Athena test rover. The prototype includes advances in the areas of goal planning, autonomous science, health assessment, autonomous navigation, and global localization. We evaluated the prototype in a simulated walkabout campaign with Mars Science Laboratory scientists. The evaluation demonstrated that the SRR approach can provide significant increase in surface mission productivity with limited communication with operators. The results showed the SRR approach reduced the duration of the walkabout by 80% compared to the duration that would have been required with current rovers.

To date, we have focused on rover activities that involved remote sensing and driving. While this makes up a large portion of rover activity, there are other rover activities such as robotic arm science and sampling that we have not addressed. These activities have their own type of productivity challenges that would benefit from a study similar to what we have conducted for driving and remote sensing. In addition, it would be helpful to conduct more simulated campaigns with different groups of scientists and in a larger area. This would enable the inclusion of the global localization and mobility health assessment components into the simulated campaign evaluations.

In addition, we will be pursuing the inclusion of some of these autonomy components into the Mars 2020 extended mission. In particular, global localization and the new autonomous science techniques we developed would be good candidates for incorporation during Mars 2020 extended mission as these technologies could be integrated without changes to the existing flight software, similar to the deployment of AEGIS on the Mars Science Laboratory rover (Francis et al., 2017).

ACKNOWLEDGMENTS

This study was conducted at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration. This study was funded by the Jet Propulsion Laboratory Research and Technology Development program. We are extremely grateful to the MSL team for their support. Special thanks to Diana Blaney, Abigail Fraeman and Vivian Sun for their help with the project.

ORCID

Daniel Gaines http://orcid.org/0000-0002-8853-9136
Gary Doran https://orcid.org/0000-0003-1233-2224
Brandon Rothrock https://orcid.org/0000-0003-2237-6589
Raymond Francis https://orcid.org/0000-0002-9031-5238
Ashwin Vasavada https://orcid.org/0000-0003-2665-286X

REFERENCES

Aaseng, G., Barszcz, E., Valdez, H., & Moses, H. (2015). Scaling up model-based diagnostic and fault effects reasoning for spacecraft. In *Proceedings of the AIAA SPACE 2015 Conference and Exposition*. AIAA.

Ansar, A. & Matthies, L. (2009). Multi-modal image registration for localization in titan's atmosphere. In *IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 2009*, IEEE (pp. 3349–3354).

Arslan, O., & Tsiotras, P. (2016). Incremental sampling-based motion planners using policy iteration methods. *CoRR*, abs/1609.05960.

Bradski, G. (2000). The OpenCV Library. Dr. Dobb's Journal of Software Tools.

Castano, R., Estlin, T., Anderson, R. C., Gaines, D. M., Castano, A., Bornstein, B., … Judd, M. (2007). OASIS: Onboard Autonomous Science Investigation System for opportunistic rover science. *Journal of Field Robotics*, *24*(5), 379–397.

Chattopadhyay, D., Mishkin, A., Allbaugh, A., Cox, Z. N., Lee, S. W., Tan-Wang, G., & Pyrzak, G. (2014). The Mars Science Laboratory supratactical process. In *Proceedings of the SpaceOps 2014 Conference*, Pasadena, CA.

Chi, W., Agrawal, J., Chien, S., Fosse, E., & Guduri, U. (2019). Optimizing parameters for uncertain execution and rescheduling robustness. In *Proceedings of International Conference on Automated Planning and Scheduling (ICAPS)*, Berkeley, CA.

Chi, W., Chien, S., & Agrawal, J. (2020). Scheduling with complex consumptive resources for a planetary rover. In *Proceedings of International Conference on Automated Planning and Scheduling (ICAPS)*, Nancy, France.

Chi, W., Chien, S., Agrawal, J., Rabideau, G., Benowitz, E., Gaines, D., … Biehl, J. (2018). Embedding a scheduler in execution for a planetary rover. in international conference on automated planning and scheduling. In *Proceedings of International Conference on Automated Planning and Scheduling (ICAPS)*, Delft, Netherlands.

Chien, S., Doubleday, J., Thompson, D. R., Wagstaff, K., Bellardo, J., Francis, C., … Piug-Suar, J. (2016). Onboard autonomy on the Intelligent Payload EXperiment CubeSat mission. *Journal of Aerospace Information Systems*, 1–9.

Chien, S., Rabideau, G., Tran, D., Doubleday, J., Nespoli, F., Ayucar, M., … Kueppers, M. (2015). Activity-based scheduling of science campaigns for the rosetta orbiter. In *Proceedings of IJCAI 2015*, Buenos Aires, Argentina.

Fesq, L. (1993). *Marple: An Autonomous Diagnostician for Isolating System Hardware Failures*. PhD thesis, University of California at Los Angeles.

Francis, R., Estlin, T., Doran, G., Johnstone, S., Gaines, D., Verma, V., … Bornstein, B. (2017). AEGIS autonomous targeting for ChemCam on Mars Science Laboratory: Deployment and results of initial science team use. *Science Robotics*, *2*(7), eaan4582. 10.1126/scirobotics.aan4582

Gaines, D., Doran, G., Justice, H., Rabideau, G., Schaffer, S., Verma, V., … Estlin, T. (2016). *Productivity challenges for Mars rover operations: A case study of Mars Science Laboratory operations*. Technical Report D-97908, Jet Propulsion Laboratory.

Goldberg, S. B., Maimone, M. W., and Matthies, L. (2002). Stereo vision and rover navigation software for planetary exploration. In *Proceedings of IEEE Aerospace Conference* (Vol. 5, pp. 2025–2036).

Helmick, D., Angelova, A., & Matthies, L. (2009). Terrain adaptive navigation for planetary rovers. *Journal of Field Robotics*, *26*(4), 391–410.

Howard, A. (2008). Real-time stereo visual odometry for autonomous ground vehicles. In *Proc. of the Int. Conf. on Intelligent Robots and Systems (IROS)* (pp. 3946–3952).

Kolcio, K. & Fesq, L. (2016). Model-based off-nominal state identification and detection for autonomous fault management. In *Proceedings of the 2016 IEEE Aerospace Conference*. IEEE.

Kolcio, K., Fesq, L., & Mackey, R. (2017). Model-based approach to rover health assessment for increased productivity. In *Proceedings of the 2017 IEEE Aerospace Conference*. IEEE.

Kolcio, K., Mackey, R., & Fesq, L. (2019). Model-based approach to rover health assessment —mars yard discoveries. In *Proceedings of the 2019 IEEE Aerospace Conference*. IEEE.

Mali, A. D. (2016). Expressing user intent in planning by instance rewriting. In *Proceedings of the 2016 IEEE 28th International Conference on Tools with Artificial Intelligence (ICTAI 2016)*, San Jose.

Mishkin, A. H., Limonadi, D., Laubach, S. L., & Bass, D. S. (2006). Working the Martian night shift: The MER surface operations process. *IEEE Robotics and Automation Magazine*, *13*(2), 46–53.

Newsom, H., Belgacem, I., Jackson, R., Ha, B., Vaci, Z., Wiens, R., … Watkins, J. (2016). The materials at an unconformity between the Murray and Stimson formations at Marias Pass, Gale Crater, Mars. In *Lunar and Planetary Science Conference* (vol. 47).

Nikora, A., Srivastava, P., Fesq, L., Chung, S., & Kolcio, K. (2018). Assurance of model-based fault diagnosis. In *Proceedings of the 2018 IEEE Aerospace Conference*. IEEE.

Paar, G., Woods, M., Gimkiewicz, C., Labrosse, F., Medina, A., Tyler, L., … Kapellos, K. (2012). PRoViScout: A planetary scouting rover demonstrator. In *Proceedings of SPIE Vol. 8301 ntelligent Robots and Computer Vision XXIX: Algorithms and Techniques*.

Pivtoraiko, M., Nesnas, I. A. D., & Kelly, A. (2009). Autonomous robot navigation using advanced motion primitives. In *2009 IEEE Aerospace Conference* (pp. 1–7).

Rabideau, G. & Benowitz, E. (2017). Prototyping an onboard scheduler for the Mars 2020 rover. In *Proceedings of the International Workshop on Planning and Scheduling for Space*, Pittsburgh, PA.

Russino, J., Gaines, D., Schaffer, S., & Wong, V. (2019). Pathogen: Using campaign intent to guide onboard planning for a self-reliant rover. In *Proceedings of the International Workshop on Planning and Scheduling for Space*, Berkeley, CA.

Shalin, V. L., Wales, R. C., & Bass, D. S. (2005). Communicating intent for planning and scheduling tasks. In *Proceedings of HCI International*, New Jersey.

Stentz, A., & Mellon, I. C. (1993). Optimal and efficient path planning for unknown and dynamic environments. *Int. Journal of Robotics and Automation*, *10*, 89–100.

Sukthankar, G., Goldman, R. P., Geib, C., Pynadath, D. V., & Bui, H. H. (2014). An introduction to plan, activity, and intent recognition. In G. Sukthankar, R. Goldman, C. Geib, D. Pynadath, & H. H. Bui (Eds.), *Plan, activity, and intent recognition*. Waltham, MA: Elsevier.

Thompson, D. R., Abbey, W., Allwood, A., Bekker, D., Bornstein, B., Cabrol, N. A., … Wagstaff, K. L. (2012). Smart cameras for remote science survey. In *Proceedings of the International Symposium on Artificial Intelligence, Robotics, and Automation in Space*.

Weiss, K. (2013). An introduction to the jpl flight software product line. In *Proceedings of the 2013 Workshop on Spacecraft Flight Software (FSW-13)*, Pasadena.

Wettergreen, D., Foil, G., Furlong, M., & Thompson, D. (2014). Science autonomy for rover subsurface exploration of the Atacama desert. *AI Magazine*, *35*(4), 47–60.

Woods, M., Shaw, A., Barnes, D., Price, D., Long, D., & Pullan, D. (2009). Autonomous science for an ExoMars rover-like mission. *Journal of Field Robotics*, *26*(4), 358–390.

Yingst, R., Berger, J., Cohen, B., Hynek, B., & Schmidt, M. E. (2017). Determining best practices in reconnoitering sites for habitability potential on Mars using a semi-autonomous rover: A geoheuristic operational strategies test. *Acta Astronautica*, *132*, 268–281.

How to cite this article: Gaines D, Doran G, Paton M, et al. Self-reliant rovers for increased mission productivity. *J Field Robotics*. 2020;1–26. https://doi.org/10.1002/rob.21979