Conducting robotics field trials: experiences, alternatives and best practices

Travis Manderson, Juan Camillo Gamboa Higuera, Sandeep Manjanna and Gregory Dudek

***Abstract*—We describe some of the challenges and key considerations in doing marine robotics field deployments. These ideas were developed based on over a decade of conducting annual multi-university field deployments in the Caribbean Sea using a range of different vehicles. We describe the logistics associated with running a field trial in a remote location and provide an insight into the planning that goes into these deployments. Deploying robots in real world environment comes with challenges and conducting these deployments away from the home base adds to these challenges. We conclude by providing suggestions regarding the conduct of high-value field experiments based on our experience.**

I. INTRODUCTION

In this paper we discuss and describe our experience conducting field experiments in marine robotics over more than a decade. Our work has spanned algorithm and systems development, but in this paper we focus on the challenges and consideration of field deployment irrespective of the specific data being collected. Doing large scale experiments far from home entails a range of pragmatic challenges different from most other domains, and we outline some of our own hard-won rules of conduct in this context.

Notably, field experiments "in the wild" can make exceptional demands not only to the equipment being used, but also on the people who are participating. It is common knowledge that researchers participating in field experiments or outdoor robotics competitions occasionally break down physically or psychologically. Taking account of fatigue, psychological factors and physical limitations is often a significant consideration that has received some attention in the context of long-term work under high stress levels [1]. To our knowledge, has not been explicitly acknowledged in context of experimental design or in the robotics literature.

Our work has addressed on several types of field robotics including driving [2], flying [3], walking [4], swimming [5, 6, 7, 8] and surface [9] vehicles, often working in tandem [10]. In this paper we focus primarily on the experiments in the marine domain, since they provide the richest confluence of logistic challenges. In particular, we have developed technologies for navigation, mapping and human interaction on fringing coral reefs (thus depths under 40 m/120 ft) and regularly conduct experiments on reefs around the island nation of Barbados. This island is hardly the most adversarial context for experimental work, but it is remote from our laboratories and replete with many experimental challenges. The fact that it is also a comfortable place to recline between experiments does not detract from the very real environmental challenges faced there, nor from the immense logistic obstacles posed by working there.

It is common knowledge that most of our planet is covered by water, and that most of that water is in the world's oceans. As a consequence, environmental conditions in the world's oceans are critical to the habitability and prosperity of the entire planet. That makes it singularly ironic that the ocean's are poorly understood even in shallow water; we have better maps of the moon than that of the Atlantic ocean!

The relative lack of data, and particularly robot-collected data, from the oceans is a direct result of the physical, observational and logistic challenges to making observations. Notably the fact that robotic systems that work undersea are generally unable to communicate over radio, need to travel large distances due to the scale of the domain, and do not have access to GPS signals [11].

II. CLASSES OF FIELD TESTING

The gold standard of for scientific and engineering progress is the acquisition of repeatable quantifiable performance metrics that explain how well an approach works, and be be used for further analysis. Ideally, a single evaluation scenario should be repeated sufficiently to obtain confidence bounds, and starting conditions should be varied to identify the impact of the initialization. This is true irrespective of the amount of advance simulation, since no amount of simulation is a replacement for actual experiments.

In outdoor field testing, however, environmental conditions or even robot characteristics may change to the extent that such repeatability is hard to achieve, especially in the early stages of development of an approach. For example, the landing of the Hyabusa2 spacecraft on an asteroid cannot be executed more than once [12]. This challenge with respect to repeatability is doubly true when there are humans in the loop who can become fatigued to change their mental state, making HRI in the field a singularly demanding domain. As a result, alternative experimental standards are frequently required.

To make the most out of the limited opportunities for field testing, it is useful to plan for a variety of experiments with different levels of risk and effort. The idea is to be prepared for failures, and to minimize the chances of coming back home empty-handed. An illustrative case is the unexpected landing of the Philae lander during the Rosetta mission [13]: even if the execution didn't go according to the plan, the mission could hardly be considered a failure. Field experiments can have a variety of successful outcomes, including

- The deployment of the system leads to a new scientific discovery
- The experiments validate the designed system under a variety of test scenarios

(a) (b) (c) (d)

Fig. 1: Overview of experiment deployment - a) Aerial view of the ocean deployment in good weather used for site survey and planning, b) beach deployment of Aqua, c) initiating underwater swimming experiment, d) executing experiment with safety diver. During deployment and experiments using Aqua we use one lead operator, one observer free to assist, and one or more videographers.

- The failures of the system lead to insight on improved designs
- The deployment of the system results in valuable data and footage that help in telling the story that motivated the research in the first place.
- The members of the team conducting the field experiments gain valuable experience for future experiments.

Given the effort required to run field trials, it is advisable prepare for tests targeting multiple objectives similar to the list above; starting with low risk tests, for which success is guaranteed, building up to more ambitious but higher risk tests.

A. *Classes of testing*

Field experiments, especially those conducted far from home, can provide an ensemble of opportunities and risks. When experiments take place in exotic locations, or have high overhead, there may be limited opportunities for the kind iteration and repeatability required for perfect quantitative evaluation. Nevertheless, there are different classes of field tests that enable scientific progress. Here we list a few.

- quantification of performance: the gold standard for scientific progress is quantitative performance evaluation with certainty bounds, but these can be very hard to achieve especially in the face of environmental variability.
- feasibility/competence verification: in very challenging environments, or with sufficiently novel designs, simply proving an approach works can be sufficient.
- Enumeration of failure modes: a parallel of basic validation is an enumeration of how failures occur and performance factors, which can be used to spur further research.
- User satisfaction/capability assessment: in addition to the measurement and quantification of performance on task metrics, the extent to which a system is satisfactory to a user population, or can be configured, provides an alternative and important aspect of evaluation.
- Discovery of failed/poor assumptions: associate with basic validation is the identification of the underlying causes of any under-performance, and the physical factors responsible.
- Discovery of new phenomena/properties: one key aspect of conducting real experiments in the real world is the discovery of new phenomena. This is the touchstone of experimental science going back to Galileo, but is rarely explicitly recognized in our discipline. It is sometimes referred to as serendipity.

B. *Pre-experiment planning*

In our experience, doing experiments away from the lab entails four major kinds of activity: travelling with equipment and people to the target location, setting up a remote experimental facility, conducting experiments, and returning home. For experiments close to home, several of these steps are trivial, but for many kinds of field robotics, field studies take place far from home. In our own cases, experimental trials are typically thousands of miles from home away from easily accessible equipment and involve setting up a very substantial remote laboratory. This means packing not just robotics systems, but lithium batteries, spare parts and tools.

Since the changes of a failure away from home and significant, we need to plan for make repairs on site. Over the 12 years we (Dudek) have been coordinating experiments in Barbados, I have seen groups from most of the participating universities experience hardware failures. These have spanned leaking “waterproof” housings, overheated CPUs, blown power supplies, and failed flash RAM. On packing list now includes all these kinds of spare parts and many more. As a consequence, a key role for remote experimental sessions is a head of packing who manages the shipping lists, inventories each case of equipment and coordinates the planning of all equipment.

The actual experiment planning for a field trial is done in two phases. First is to plan the experiments and prepare the schedule before the field trip. Second is to plan or revise each day’s experiments while on the trip, overview the schedule for the day, and update the experiment schedule according to the progress from previous days.

Aside from the particular research interests of each team member, we assign roles during planning to ensure that the workload is evenly distributed. These roles include the *principal investigators* who lead the team, an *experiments manager* who decides on scheduling of experiments and helps in resolving resource conflicts, a *packing lead* who helps preparing packing lists and manages the inventories before departure and after arrival, a *documents lead* who reminds

Fig. 2: Experiment team and transported experiment facilities, (a) circa 2005 and (b) 2018

people about any paperwork that's required for the experiments (permits, immigration documents, insurance) and *experiment leads* who deal with the specific details for each experiment. Note there may be other roles depending on the type of experiments; e.g. a *diving equipment lead* that ensures that diving equipment is available for the experiments, and *videographers* who document the progress of experiments.

C. Autonomous deployment of an amphibious vehicle: From land to water

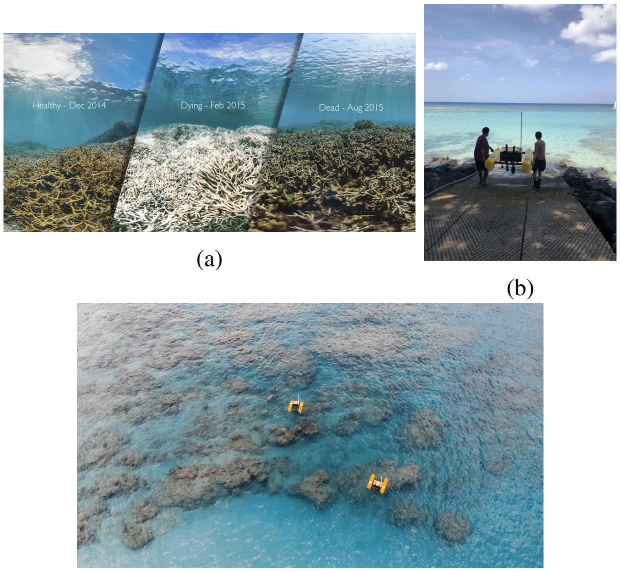
We recently designed legs that allow the Aqua platform walk and swim successfully without swapping the legs. The design was based on intuitions built from our experience in walking with compliant semi-circular legs and swimming with flexible paddles. The new design, affectionately named the "bone-saw legs", combines both types of legs and flipper into one cohesive part, using synthetic material.

Amphibious legs may enable practical and low-cost deployment of the robot for underwater environmental monitoring. A potential user (a scientist in the field) would only need to deploy the robot on land and indicate to the robot which way to go with either waypoints, and example image of interest, or other scientific measuring objective. Using its legs in walking mode, the robot can autonomously direct itself to the water and switch its gait to swimming when it detects it is more efficient to swim than walk. Afterwards, the robot may carry out its mission: for example, using the vision-based navigation and tracking algorithms we've developed, the robot may collect data for a fixed period of time before returning to land and walking home.

This is an example of a high risk experiment: design and testing iterations take extended periods of time, scheduling is highly dependent on weather conditions, and requires coordination by team members on land and in water. In this case, we focused on providing qualitative evaluation on the performance of the system, demonstrating the intended autonomous amphibious capabilities by a proof-of-concept.

D. Deployment and reef survey with autonomous surface vehicles

An important part of the planning for field experiments is choosing an appropriate location to conduct experiments. Considerations often include the appropriateness as related to the specific requirements, travel, local support, accommodations as well as whether our research can benefit the local community. One such example is our work on coral reef survey with autonomous surface vehicles. Barbados has a McGill facility which is situated on the beach with access to ocean. This makes the location attractive in terms of accommodation and the access to deployment site (Fig. 3b). Also the North and South Bellairs reefs right behind the McGill facility provide right environment to run our reef survey experiments. This has also excited local marine conservation authorities and has been appreciated by the local communities.



(a)

(b)

(c)

Fig. 3: Coral reef monitoring with robotic boats. a) Effect of global warming on Corals [14], b) Two people deploying the boat into the ocean, c) Aerial view of two robotic boats surveying the region

Along with running robotic-related experiments, we investigate techniques to survey and monitor the health of the coral reefs which can be tracked over the years. This is beneficial to both robotic community and marine biologists who are interested in assessing how the health of coral reefs changes over time. We built and deployed two autonomous surface vehicles for monitoring the health of reefs around the island of Barbados (Fig. 3c). Increase in the ocean temperatures has resulted in widespread coral bleaching at an ever-increasing rate (Fig. 3a). Improved monitoring will enhance the currently poor understanding of the spatial and temporal dynamics of coral bleaching and thus the effects of climate change on these fragile ecosystems could be measured.

Surface vehicles in ocean need to fight constant waves and wind. Changes in tides and weather also result in changes in

Fig. 4: Daylight experiment that extended into a challenging and potentially dangerous evening scenario.

Fig. 5: Managing and occasionally untangling large amounts of fibre-optic cable was, at one time, a recurring annoyance

the conditions of the sea surface, such as ocean roughness, visibility underwater, and exposed reef rocks. These cause an increased load on the boats and in turn reduce the time budget to finish the survey. This motivated us to design a non-uniform adaptive sampling technique [9] that generates paths to efficiently sample data and then mathematically model a scalar field in a given region of interest. Tides play an important role in experiment schedules for the day as they affect the boat deployments.

E. Anecdotes about things "going sideways"

In the interest of grounding some of the issues above in concrete examples, we undertake in this section to describe some of the notable of myriad challenges that have arisen in our own experiments of those of our collaborators.

a) Time extension: In marine experiments involving scuba divers, our vision-based experiments generally need to end before dark. This is due to both the required operating conditions for the robot as well as the safety of the diver to can have difficulty diving after dark. Despite that, it has occurred several times that experimental sessions needed more time that the full day offered, and an experimental session extended into the early evening. To accommodate this, our dive teams have trained in the particular skills needed for night diving, and carry multiple flashlights to permit this extended operation (as illustrated in Fig. 4).

b) Why we dislike remote control: In the early years of our work on marine robotics, we used a tethered vehicle that could be controlled from ship-board. While typically used a light-weight tether the had limited impact on maneuverability, it had other disadvantages. These included the hypothetical risk of snagging the tether (rarely a serious issue in itself) as well as the more inconvenient issue of the tether becoming a tangled mess that could require substantial unexpected time and effort to untangle, potentially delaying other activities (as illustrated in Fig. 5).

III. PRESCRIPTIVE ADVICE

Based on our experience, we make the following suggestions regarding the conduct of high-value field experiments. As indicated above, our prescriptive advice is targeted especially to scenarios where the overhead of visiting an experimental site, or conducting an experiment is high.

1) Experiment planning: When planning field experiment that have high cost, it is often critical to assure that the undertaking comes back with at least some data, even of poor quality, both to assure the psychological needs of the participants as well as to provide assurances to participants or funding sources. The fact that experiments are being conducted by human beings cannot be neglected: people benefit from positive feedback and encouragement to achieve maximum productivity. In an experimental setting, a loss of confidence by the team members can even lead to a complete collapse of team focus and motivation.

This suggests a graded approach to the conduct of experiments:

- Place easy experiments first. They build confidence, test the equipment and underlying assumptions, and validate any preconceptions.
- Do some experiments with assured success, even when the experimental value is low. This is the corollary of the item above.
- Do the highest risk experiments last: if the equipment is damaged, you have already locked in some value.
- Research activities can sometimes be leveraged by superficial considerations like off illustrative photographs. These are especially important for teaching. Due to their relevance to the pursuit of funding, such illustrative pictures are informally known as "Money shots". The acquisition of such images should have a privileged place in an agenda.
- Keep some very high risk or improbable experiments on the agenda, in case time and energy allows for it. This avoids the risk of a wasted opportunity.

2) Contingency planning: A reality of experimental robotics is that not every experiment is a success, sometimes due to problems with the experiment itself, and sometimes due to exogenous factors like weather or hardware resources. For experimental sessions with high logistic overhead, for example on the open ocean, it's important not to waste experiment time. This leads to several simple rules for thumb for contingency planning.

- Have backup tests and alternative plans
- Be prepared in advance to allow different experiments to swap time slots if conditions warrant it
- Prepare for weather contingencies
- For each time slot, have both preferred and backup plans
- This can imply having a roster of "bonus" activities, but it is important keep expectations bounded.

Note that having swappable experiments makes the overall logistics much more difficult, especially when human resources differ between the experiments.

3) Experiment scheduling: In scheduling outdoor or underwater experiments, several rules of thumb can be used to maximize value and reduce the impact of failures.

- Have a clear sense of how to value experiments, so that plans can be adapted to maximize utility.
- Occasional failures are inevitable, so plans should be structured to avoid having them cascade. For example, if a series of experiments all depend on a single fragile piece of equipment, and alternative experimental pathway should be prepared as back, if possible.
- "Pull the plug when necessary". That is, exercise a degree of brutal decisiveness if an experiment if failing repeatedly. Failure to give up on a single experiment can otherwise lead to a particular form of cascading failure.

I general, maximizing utility can depend on making potentially harsh decisions regarding the termination of fruitless experimental pathways. When the total ensemble of experiments is the product of a team effort, such decisions can entail substantiate impact on the team dynamics since a lot of personal ego can be attached to any single experiment. As a result, the decision to terminate an experimental pathway should be delegated to the most senior person possible.

IV. CONCLUSION

In this paper we attempted to distill over a decade of experience conducting field trails into a set of guiding rules, principles and policies. Robotics systems are often build to be robust in the face of failures. Or experience suggests that field experiments themselves should be planned with a similar degree of robustness.

While robotics experiments in unstructured terrain, and particular marine environments, can have exceptional value and importance, they are also exceptionally challenging in terms of logistics and execution.

REFERENCES

[1] R. Flin, P. O'Connor, and K. Mearns, "Crew resource management: improving team work in high reliability industries," *Team Performance Management: An International Journal*, vol. 8, no. 3/4, pp. 68–78, 2002. [Online]. Available: https://doi.org/10.1108/13527590210433366

[2] T. Manderson, R. Cheng, D. Meger, and G. Dudek, "Navigation in the Service of Enhanced Pose Estimation," in *International Symposium on Experimental Robotics (ISER)*, 2018.

[3] T. Manderson, A. Holliday, and G. Dudek, "Gaze Selection for Enhanced Visual Odometry During Navigation," in *Proceedings of the Conference on Computer and Robot Vision (CRV)*, 2018.

[4] S. Manjanna and G. Dudek, "Autonomous gait selection for energy efficient walking," in *2015 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2015, pp. 5155–5162.

[5] J. Sattar, G. Dudek, O. Chiu, I. Rekleitis, P. Giguere, A. Mills, N. Plamondon, C. Prahacs, Y. Girdhar, M. Nahon *et al.*, "Enabling autonomous capabilities in underwater robotics," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2008, pp. 3628–3634.

[6] D. Meger, J. C. Gamboa Higuera, A. Xu, P. Giguere, and G. Dudek, "Learning legged swimming gaits from experience," in *Robotics and Automation (ICRA), 2015 IEEE International Conference on*. IEEE, 2015, pp. 2332–2338.

[7] F. Shkurti, W. D. Chang, P. Henderson, M. J. Islam, J. C. Gamboa Higuera, J. Li, T. Manderson, A. Xu, G. Dudek, and J. Sattar, "Underwater Multi-Robot Convoying Using Visual Tracking by Detection," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, Vancouver, Canada, sep 2017, pp. 4189–4196.

[8] J. C. Gamboa Higuera, D. Meger, and G. Dudek, "Synthesizing neural network controllers with probabilistic model-based reinforcement learning," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2018, pp. 2538–2544.

[9] S. Manjanna, H. van Hoof, and G. Dudek, "Policy search on aggregated state space for active sampling," in *2018 International Symposium on Experimental Robotics (ISER)*, 2018, pp. 1–7.

[10] F. Shkurti, A. Xu, M. Meghjani, J. C. Gamboa Higuera, Y. Girdhar, P. Giguere, B. B. Dey, J. Li, A. Kalmbach, C. Prahacs, K. Turgeon, I. Rekleitis, and G. Dudek, "Multi-Domain Monitoring of Marine Environments using a Heterogeneous Robot Team," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Portugal, 2012, pp. 1447–1753.

[11] G. Dudek and M. Jenkin, "Inertial sensors, gps, and odometry," *Springer Handbook of Robotics*, pp. 477–490, 2008.

[12] R. Yamada, H. Araki, K. Yamamoto, H. Senshu, N. Namiki, H. Noda, K. Matsumoto, F. Yoshida, S. Abe, and N. Hirata, "Albedo observation of c-type asteroid ryugu using the hayabusa2 lidar," in *Lunar and Planetary Science Conference*, vol. 50, 2019.

[13] J. Biele, S. Ulamec, M. Maibaum, R. Roll, L. Witte, E. Jurado, P. Muñoz, W. Arnold, H.-U. Auster, C. Casas *et al.*, "The landing (s) of philae and inferences about comet surface mechanical properties," *Science*, vol. 349, no. 6247, p. aaa9816, 2015.

[14] R. Vevers, "Xl catlin seaview survey," The Ocean Agency.