



Oceanids C2: An Integrated Command, Control, and Data Infrastructure for the Over-the-Horizon Operation of Marine Autonomous Systems

Catherine A. Harris^{1*}, Alvaro Lorenzo-Lopez¹, Owain Jones¹, Justin J. H. Buck², Alexandra Kokkinaki², Stephen Loch², Thomas Gardner² and Alexander B. Phillips¹

¹ Marine Autonomous and Robotic Systems, National Oceanography Centre, Southampton, United Kingdom, ² British Oceanographic Data Centre, National Oceanography Centre, Liverpool, United Kingdom

OPEN ACCESS

Edited by:

Leonard Pace,
Schmidt Ocean Institute,
United States

Reviewed by:

Adrian Tramallino,
Schmidt Ocean Institute,
United States
Philip Andrew McGillivray,
United States Coast Guard Pacific
Area, United States

*Correspondence:

Catherine A. Harris
cathh@noc.ac.uk

Specialty section:

This article was submitted to
Ocean Observation,
a section of the journal
Frontiers in Marine Science

Received: 15 November 2018

Accepted: 07 May 2020

Published: 08 June 2020

Citation:

Harris CA, Lorenzo-Lopez A, Jones O,
Buck JJH, Kokkinaki A, Loch S,
Gardner T and Phillips AB (2020)
Oceanids C2: An Integrated
Command, Control, and Data
Infrastructure for the Over-the-Horizon
Operation of Marine Autonomous
Systems. *Front. Mar. Sci.* 7:397.
doi: 10.3389/fmars.2020.00397

Long-range Marine Autonomous Systems (MAS), operating beyond the visual line-of-sight of a human pilot or research ship, are creating unprecedented opportunities for oceanographic data collection. Able to operate for up to months at a time, periodically communicating with a remote pilot via satellite, long-range MAS vehicles significantly reduce the need for an expensive research ship presence within the operating area. Heterogeneous fleets of MAS vehicles, operating simultaneously in an area for an extended period of time, are becoming increasingly popular due to their ability to provide an improved composite picture of the marine environment. However, at present, the expansion of the size and complexity of these multi-vehicle operations is limited by a number of factors: (1) custom control-interfaces require pilots to be trained in the use of each individual vehicle, with limited cross-platform standardization; (2) the data produced by each vehicle are typically in a custom vehicle-specific format, making the automated ingestion of observational data for near-real-time analysis and assimilation into operational ocean models very difficult; (3) the majority of MAS vehicles do not provide machine-to-machine interfaces, limiting the development and usage of common piloting tools, multi-vehicle operating strategies, autonomous control algorithms and automated data delivery. In this paper, we describe a novel piloting and data management system (C2) which provides a unified web-based infrastructure for the operation of long-range MAS vehicles within the UK's National Marine Equipment Pool. The system automates the archiving, standardization and delivery of near-real-time science data and associated metadata from the vehicles to end-users and Global Data Assembly Centers mid-mission. Through the use and promotion of standard data formats and machine interfaces throughout the C2 system, we seek to enable future opportunities to collaborate with both the marine science and robotics communities to maximize the delivery of high-quality oceanographic data for world-leading science.

Keywords: marine autonomous systems, over-the-horizon operations, autonomous gliders, data curation, near-real-time data, human-robot interface, microservices architecture, software development

1. INTRODUCTION

Due to the scale and complexity of the extreme environment of the world's oceans, acquiring marine measurements at an optimal spatial and temporal resolution is very challenging. Traditional ship-based methods are often prohibitively expensive and only enable observation in the locality of the research ship, resulting in very limited snapshots of marine ecosystems. The Argo array of 4,000+ floats addresses this by providing coarse global coverage of temperature, salinity and velocity measurements (Roemmich et al., 2009), but are unable to provide spatially targeted measurements, measurements from the deep-ocean, sampling of seasonal ice zones, marginal seas, and boundary currents, due to the drifting nature of float technology (Roemmich and the Argo Steering Team, 2009).

Advances in long-endurance Marine Autonomous Systems (MAS) (Eriksen et al., 2001; Manley and Willcox, 2010; Furlong et al., 2012; Roper et al., 2017), piloted over-the-horizon, i.e., without an operator or research ship nearby, offer an opportunity to bridge the gap between research vessels and float technology, significantly reducing reliance on research ships, enabling multiple robotic vehicles to operate simultaneously in an area for an extended period of time and providing an improved composite picture of the marine environment. Underwater gliders in particular have been viewed as a critical component of future observing systems (Testor et al., 2010, 2019; Liblik et al., 2016), while Long Range Autonomous Underwater Vehicles (AUVs) are being considered for monitoring disperse decommissioned oil and gas infrastructure (Jones et al., 2019).

To facilitate access to marine measurement technologies, the UK's Natural Environment Research Council (NERC) funds the National Marine Equipment Pool (NMEP) hosted at the National Oceanography Centre (NOC). The NMEP is the largest centralized marine scientific equipment pool in Europe, providing scientific instruments and equipment capable of sampling from the sea surface to the deep ocean. It includes the Marine Autonomous and Robotic Systems (MARS) fleet of 40+ robotic platforms, comprising Autonomous Underwater Vehicles, Remotely Operated Vehicles (ROVs), Unmanned Surface Vehicles (USVs), and underwater gliders (see **Figure 1A**). To fulfill our obligations to the NERC Data Policy and to "ensure the continuing availability of environmental data of long-term value..."¹, scientific data collected from the MARS platforms is delivered and archived by the British Oceanographic Data Centre (BODC) (a data assembly center, DAC), where it is made available to the wider scientific community and the general public.

MARS vehicles are now routinely used in simultaneous single or multi-vehicle campaigns over geographically disparate locations (see **Figure 2**) and are involved in many large-scale national and international research programmes. Recent examples include: the use of underwater gliders to collect a continuous record of the full water-column within the Overturning in the Sub-polar North Atlantic Programme

(OSNAP) (Houpert et al., 2018); the acoustic harvesting of data from the RAPID array of trans-Atlantic moorings (Cunningham et al., 2007); deep ocean exploration in the Southern Ocean (Salavasidis et al., 2018; Garabato et al., 2019) and under the Filchner-Ronne ice shelf in Antarctica (McPhail et al., 2019) (see **Figure 1B**). The annual MASSMO (Marine Autonomous Systems in Support of Maritime Operations) deployments have seen heterogeneous fleets of USVs and underwater gliders operating in shelf and shelf-break locations to demonstrate the capabilities of MAS each year since 2014. In May 2017, MASSMO 4 saw eight vehicles collect oceanographic and passive acoustic data in the Faroe-Shetland Channel (see **Figure 1C**).

Whilst the popularity of MAS vehicles continues to grow, the further expansion of heterogeneous multi-vehicle operations is currently limited by a number of factors: firstly, MAS are typically highly specialized vehicles, each model of which operates using its own command interface, without cross-platform standardization; secondly, the data produced by each vehicle is typically in a custom format, making the automated ingestion of observational data for near-real-time analysis and subsequent data discovery very difficult, resulting in a time-consuming manual process; thirdly, the majority of MAS vehicles do not provide machine-to-machine interfaces, limiting the development and usage of multi-vehicle operating strategies and the automated delivery of data to end-users.

To ensure the MARS fleet is used to its full potential, the Oceanids Command and Control (C2) project is developing the required piloting tools and data services to streamline operation of MAS platforms operating over-the-horizon. In this paper, section 2 highlights the challenges for large scale over-the-horizon operation and section 3 reviews existing work addressing these issues. Section 4 describes the approach adopted and section 5 provides an overview of the C2 system. Section 6 addresses the C2 piloting infrastructure and section 7 describes the C2 data system. Section 8 describes ongoing developments and applications to real-world science deployments. Finally, section 9 presents conclusions from the work to date.

2. PROBLEM DEFINITION

Traditionally, deployments of marine autonomous systems (MAS) have consisted of a vehicle being launched from a support ship, performing a single pre-defined mission (lasting in the order of one day) and returning to the ship for recovery, data analysis, and redeployment on the next mission (German et al., 2008; McPhail, 2009). As a result of increases in battery and power technology, long-endurance vehicles such as underwater gliders are being deployed in increasingly large fleets for months at a time, without the need for a support ship. This paradigm shift has many implications for the piloting and operation of fleets of such long-endurance vehicles. In traditional deployments, a plan for a single vehicle would be constructed by highly-skilled pilots using prior knowledge of the environment ascertained from support ship observations of the deployment site. However, long-endurance vehicles perform multiple missions in a single deployment, periodically establishing communications with a

¹Natural Environment Research Council (NERC) data policy. Available online at: <https://nerc.ukri.org/research/sites/data/policy/data-policy/> (accessed November 12, 2018).

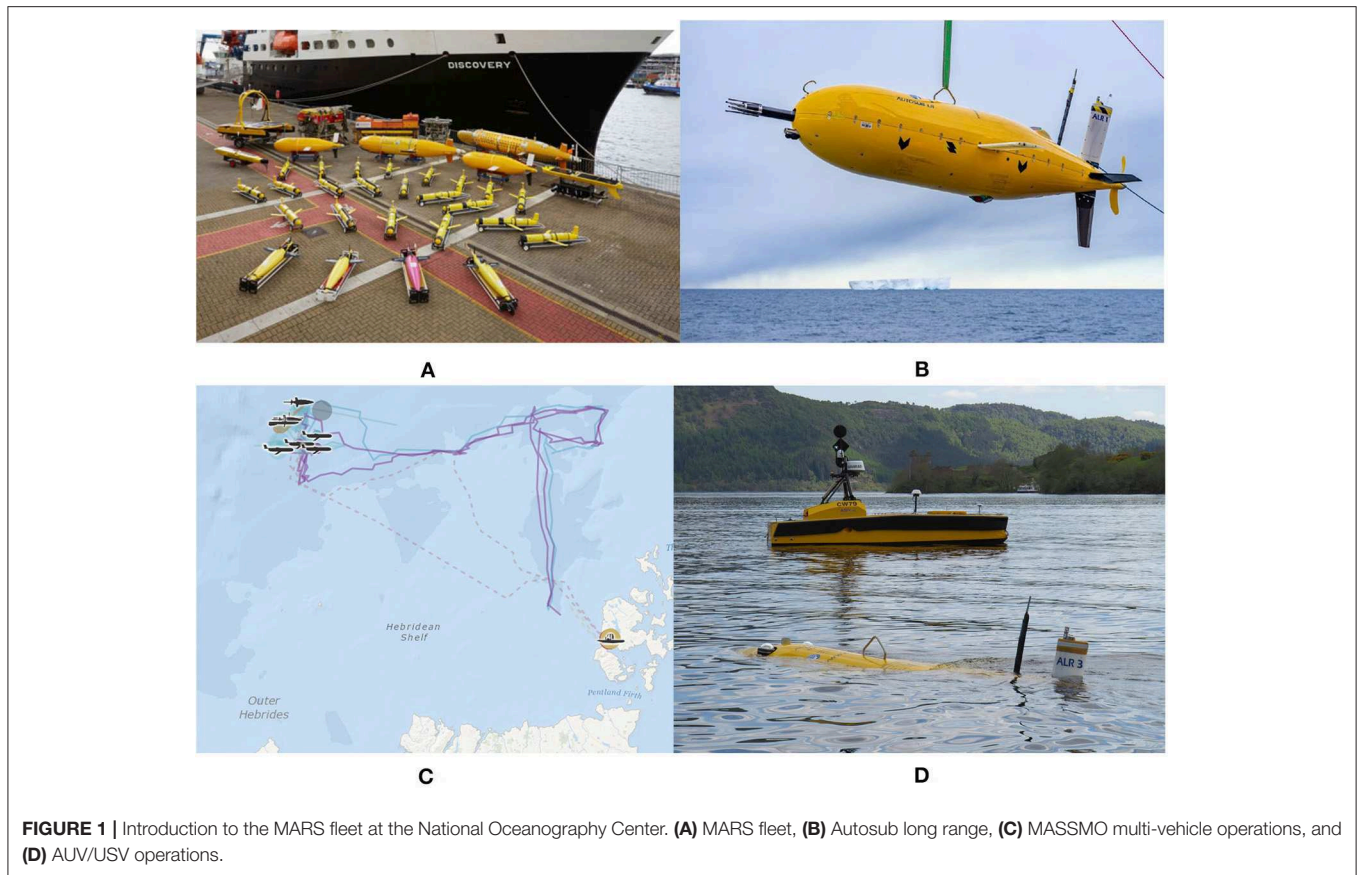


FIGURE 1 | Introduction to the MARS fleet at the National Oceanography Center. **(A)** MARS fleet, **(B)** Autosub long range, **(C)** MASSMO multi-vehicle operations, and **(D)** AUV/USV operations.

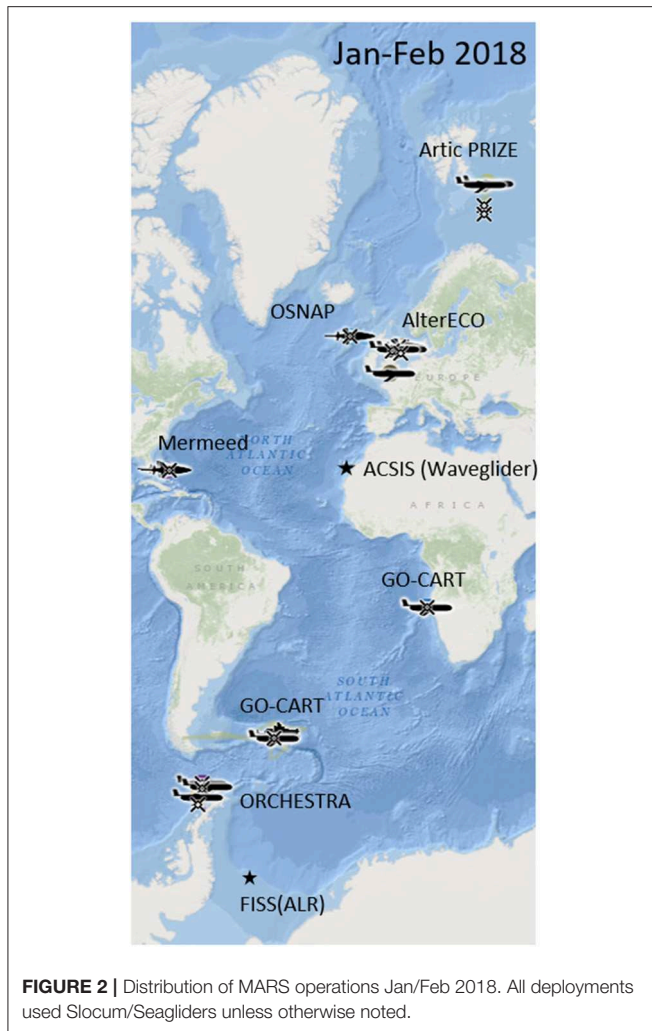
remote pilot via a satellite link. This is known as “over-the-horizon” piloting and requires highly-skilled pilots to be on-call throughout the day and night, to construct new plans based on ocean models and limited observations from the vehicle itself, and to respond to unexpected vehicle states, such as aborts and drifts in calibration. With the increasing frequency, duration, and complexity of MAS deployments, human piloting is becoming more intensive. New data streams from models, forecasts and live observations inform piloting decisions but also increase the time and piloting skills required to take them. The number of highly-skilled pilots cannot feasibly grow at the same rate as the number and duration of MAS operations, highlighting a future scalability problem. If unaddressed, this will limit the number of MAS vehicles which can be deployed simultaneously and may impact mission success rates.

In this section, we discuss the key factors which currently limit the operation of larger, heterogeneous fleets of MAS vehicles and consequently form the major design drivers for our C2 system. Key user interactions with the C2 system and operating modes of the long-range fleet are illustrated in **Figure 3**.

2.1. Diverse Platform Types

The majority of the MARS fleet are long-range platforms which can be piloted over-the-horizon. These can be classified into three distinct vehicle types:

1. **Underwater gliders**—vehicles which use a pump mechanism to make small changes to their buoyancy to move up and down through the water-column (Rudnick et al., 2004). Instead of using a power-hungry propeller, gliders have fixed wings which convert this vertical motion into forwards speed. The resulting sawtooth profile samples the ocean both vertically and horizontally, and requires very little power in comparison to a traditional propeller and motor. As a result, the range of a single glider deployment has been demonstrated as several 1,000 km. Gliders surface and communicate at irregular periods, determined by the dive depth. In shallow water this can be of the order of tens of minutes, whilst for deep-rated gliders diving to 6,000 m this can exceed 24 h.
2. **Long-range autonomous underwater vehicles (LRAUVs)**—are propeller-driven underwater robots which, unlike gliders, are able to move independently in the horizontal and vertical plane (Furlong et al., 2012; Hobson et al., 2012; Roper et al., 2017). This enables them to conduct more complex behaviors while submerged, such as benthic and geological surveys. As LRAUVs can spend days to months underwater without surfacing, opportunities for satellite communication with a remote pilot are highly dependent on the mission plan. Underwater communication channels, such as acoustic modems are very limited in range, with limited application to over-the-horizon operations. The use of Unmanned Surface Vehicles, acting as acoustic communications gateways to



LRAUVs is a promising area of ongoing research (German et al., 2012; Phillips et al., 2018) (see **Figure 1D**). However, the need for a second vehicle type increases piloting complexity and mission risk.

3. **Long-range unmanned surface vehicles (LRUSVs)**—such as Waveglider (Manley and Willcox, 2010), C-Enduro, and Autonaut, are able to maintain continuous satellite communication. As LRUSVs are on the surface constantly, they pose a collision risk to shipping traffic. As a result, LRUSVs are typically supervised constantly by a team of pilots in order to meet existing COLREGS.

In contrast to traditional single-vehicle deployments, MAS deployments are increasingly involving multiple vehicle-types, deployed simultaneously. These heterogeneous fleets have many benefits for data quality and coverage, providing a comprehensive view of both spatial and temporal variation within the operating environment. However, the different capabilities of these vehicle types adds significant piloting overheads, requiring pilots to be trained in the control and operation of multiple vehicles.

In addition to the complexity of multiple vehicle types, vehicles are manufactured by a range of commercial and academic developers. For example, underwater glider models include Seagliders (Eriksen et al., 2001) and Slocum Gliders (Schofield et al., 2007), both of which use the same underlying principles, implemented in significantly different ways. Differences in manufacturer design lead to a wide range of vehicle capabilities, constraints, procedures, and interfaces, even for vehicles of a common type. Many vehicle interfaces were designed considering traditional single-vehicle mission formats and, as a result, lack the interoperability needed to facilitate multi-vehicle operations.

2.2. Deployment-Specific Configurations

Each vehicle type has different sensor suite configurations, enabling the vehicle to be customized to meet the science requirements and energy constraints of the mission. This leads to complexity in the command, control and data processing associated with each vehicle, as different physical components have different commands, operating constraints, failure-modes and data formats. For instance, an altimeter on a glider enables the vehicle to measure its altitude above the seafloor, reducing the likelihood of collisions. However, the altimeter significantly increases power consumption, so to maximize vehicle endurance, pilots may be trained to switch the altimeter off whilst away from the seafloor. This creates complexity for the data and metadata systems, as the number of datastreams from the vehicle is not fixed throughout the mission. For the analysis of vehicle reliability, it is also important to record whether the loss of a datastream mid-mission was operator commanded, or as a result of a vehicle/sensor fault. The inclusion of the altimeter or other scientific sensors within the vehicle configuration also increases the number of components which may fail and thus should be routinely checked by the pilot throughout a deployment to ensure correct operation.

Consequently, piloting interfaces and procedures, as well as data and metadata formats, need to be adaptable and robust to changes in sensor suite.

2.3. Availability of Near-Real-Time Data Mid-Mission

As long-range vehicles are typically at sea for weeks to months at a time, pilots receive data throughout the mission as well as fully curated research-quality data-sets following recovery of the vehicle. Following the definitions of Roemmich et al. (2010), we refer to mid-mission data as “near-real-time” (NRT), as whilst there is a delay due to satellite communications only being possible whilst on the surface, this data is available significantly faster than with traditional devices, where data is only accessible on recovery of the instrument (referred to as “delayed-mode” data). Major advantages of NRT data include the ability to use this data to inform piloting and scientific decisions made mid-mission, for example in adjusting the location of survey locations, or the sampling parameters of on-board sensors etc to increase data quality.

The C2 system will allow the opportunities presented by NRT data to be fully exploited to benefit MAS deployment

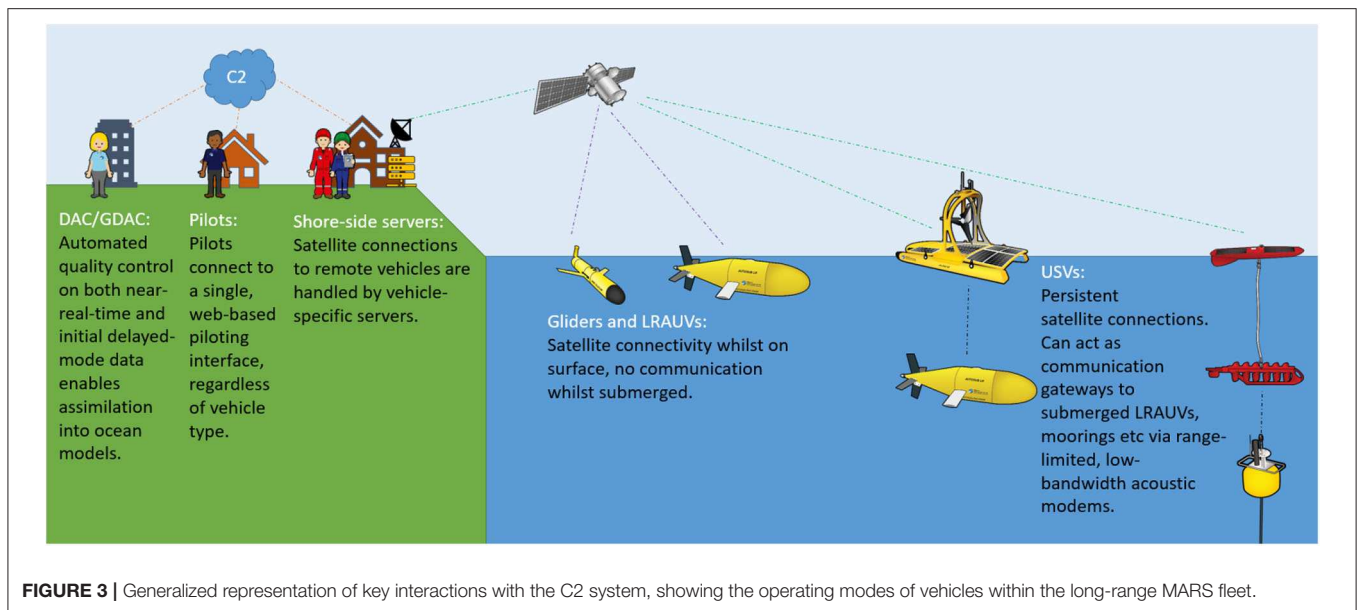


FIGURE 3 | Generalized representation of key interactions with the C2 system, showing the operating modes of vehicles within the long-range MARS fleet.

scientific objectives. The data pathway of the C2 system will automatically process both delayed-mode data and NRT data, with priority given to the processing of NRT data into standard scientific data formats. Historically, within BODC the bulk of the data in the archives is delayed-mode, and much effort has been spent to log and include the necessary metadata to ensure the data is discoverable. However, BODC also has experience of NRT data management which, in the case of Argo (Roemmich and the Argo Steering Team, 2009) dates back to 2001, and for gliders from 2011. The advent of the C2 data system with its emphasis on metadata assembly prior to deployment promises to revolutionize the role of the data center by: (1) significantly reducing the effort required from data center staff by, for example, automating the accessioning process, (2) speeding up processing and reducing time-to-delivery by two orders of magnitude or more, and (3) providing a highly automated processing service for scientists. In other words, users can expect to see calibrated and quality controlled data one or two minutes after the relevant MAS vehicle has surfaced.

DACs, such as BODC, submit data to Global data assembly centers (GDACs) and other end-users for ingestion and assimilation into models and data products. Such users may require the data to be delivered within a specific time-window for inclusion. For example, the World Meteorological Organization's Global Telecommunication System (which coordinates the collection, exchange and distribution of observation data for ocean and weather forecasting) requires data to be submitted within a 19 h window. As such, the data system within the C2 must be able to meet these time-constraints. There are also data quality levels to be considered—automated quality control methods can be applied to both NRT data and the initial delayed-mode data collected on recovery of the vehicle. This process enables a quality assurance that is sufficient for assimilation into operational ocean models.

2.4. Data Standardization

In order to maximize the exploitation of the data by end users, it is crucial that metadata and data are easily accessible and presented in a well-defined standard format. Without a standard, there is a significant learning curve for end-users to read and analyse a data set from a new vehicle or data provider, whilst also complicating the comparison of data between multiple platforms, or over different time periods. Therefore it is important that MAS missions are supported by an appropriate data management strategy that archives and disseminates collected data and associated metadata in a sustainable format that is exchangeable and discoverable between stakeholders. For underwater gliders the European community has recently standardized on the Everyone's Gliding Observatories (EGO) Network Common Data Form (NetCDF)² format (EGO gliders data management team, 2017) which documents the naming conventions and metadata content. However, adoption of a new format requires buy-in from all users and the time and resources to convert from existing processing methods, so further work is still required before this can be considered industry-standard. Whilst different standards have been defined within the US (IOOS) (U.S. IOOS National Glider Data Assembly Center, 2018) and Australian (IMOS) (Australian National Facility for Ocean Gliders, 2012) glider communities, these formats have many similarities with EGO NetCDF and there is ongoing communication between the communities to standardize into a global format. There are base standards such as NetCDF, Climate and Forecast (CF), Attribute Convention for Data Discovery (ACDD)³, and common vocabularies for metadata that underpin the common

²Network Common Data Form (NetCDF). Available online at: <https://www.unidata.ucar.edu/software/netcdf/> (accessed November 14, 2018).

³Attribute Convention for Data Discovery. Available online at: http://wiki.esipfed.org/index.php/Attribute_Convention_for_Data_Discovery (accessed November 14, 2018).

elements of interoperability between IMOS, IOOS, and EGO formats. Efforts to fully harmonize the formats will simplify access to users and the development of common software tools for working with ocean glider data. For metadata, applicable exchange formats are SeaDataNet Common Data Index (CDI)⁴ records, Marine Sensor Web Enablement⁵ profile SensorML (Open Geospatial Consortium, 2016) and Semantic Sensor Network (Haller et al., 2017). A recently introduced metadata standard is schema.org which will enable data discovery via Google dataset search. There are also community data access tools, such as ERDDAP⁶, that aid interoperability and data exchange. As LRAUVs are a relatively new technology that is still heavily under development, community standardized data formats have yet to emerge. Therefore, there is a requirement within our C2 system to ensure that data from LRAUVs sent to the DAC is in a suitably adaptable, modular format, to facilitate later adoption of a standard or unification with EGO NetCDF, for instance.

2.5. Scalability and Resilience

It is essential that the C2 system is resilient and scalable to increases in vehicle connections, data streams and vehicle types. The system must be resilient to failures or interruptions at any stage of the processing of data or piloting commands. As a result, critical services and databases, for example those which contain metadata and control information, must have failover in place. The architecture of the C2 system must be sufficiently modular, flexible and portable to support expansion and enable the potential use of Cloud resources where required.

2.6. Vehicle Reliability Requirements

Deploying and operating robots in the extreme environment of the world's oceans is very challenging. The inherent uncertainty of the highly dynamic operating environment adds significant risk to MAS deployments, both to the physical safety of the vehicle itself and to the successful delivery of the scientific data. If a vehicle is lost without first sending data back via satellite, any data which has been collected is also lost. Whilst this has clear implications for the delivery of the mission's scientific objectives, the loss of the engineering data also has significant impact for analysing the reliability of the fleet. If a vehicle is lost or experiences a fault whilst at sea, it is crucial that we analyse the available data in an attempt to identify the root cause, design effective mitigation, and ascertain whether the fault is common to other vehicles of that model or type. At present, human pilots are required to supervise the operation of complex vehicles throughout the day and night (constantly, in the case of LRUSVs). Under such working conditions, a degree of human error is inevitable (Stokey et al., 1999)—it is very easy within the

existing manual process for a pilot to fail to spot an emerging trend for example, or to fail to record key information whilst communicating with the vehicle under time pressure. To both increase the amount of engineering data available for reliability analysis and to reduce the emphasis on individual pilots to record key events and information, the new system should log all data and communications with the vehicle automatically. This data can then be analyzed in near-real-time by the server, with anomalous data brought to the user's attention.

2.7. Machine-to-Machine Interfaces

With the recent research focus and subsequent popularization of deep learning tools and the wider field of Artificial Intelligence (AI), there is the clear potential to apply such techniques to the remote piloting of MAS vehicles. At present, human pilots follow standard procedures of checks and manual optimization of vehicle performance. Some of these tasks, such as the regression of glider flight variables, are well-suited to AI and machine learning methods. Furthermore, the online availability of external data sets such as weather forecasts and ocean models, opens up the opportunity to use algorithms to perform larger-scale optimization across a wide-range of piloting tasks, from waypoint definition to minimize vehicle resource usage, to sensor parameter adjustment to optimize data quality. To enable the use of intelligent algorithms and automation programs, it is necessary to have well-defined machine-to-machine interfaces which:

- Allow bidirectional communications.
- Prioritize the safe operation of MAS vehicles, preventing the violation of vehicle, or mission constraints.
- Ensure secure access to remote vehicles and associated infrastructure.
- Follow industry standards, enabling software developers to quickly understand the interfaces they need to use to generate their applications.

However, many existing vehicles and associated piloting interfaces were designed for a human pilot to operate a single vehicle via a user interface, and as such there was never a previous need for interfaces to enable machine access and control. The development and standardization of machine-to-machine interfaces to unify piloting across different vehicle types is a key objective of the C2 project.

2.8. Range of Stakeholders

The Oceanids C2 project aims to serve a wide number of stakeholders as a UK national infrastructure. In **Table 1** we present the identified stakeholders.

3. LITERATURE REVIEW

3.1. Autonomous Ocean Sampling Networks

There have been multiple successful demonstrations of the potential for Autonomous Ocean Sampling Networks (AOSN) (Ramp et al., 2009; Leonard et al., 2010; Haworth et al., 2016). AOSN deployments typically consist of a heterogeneous fleet of MAS vehicles, ships, and/or sensor moorings, operating within

⁴SeaDataNet Common Data Index (CDI). Available online at: <https://www.seadatanet.org/Metadata/CDI-Common-Data-Index> (accessed November 11, 2018).

⁵Open Geospatial Consortium, Sensor Web Enablement. Available online at: <http://www.opengeospatial.org/domain/swe#initiative> (accessed November 14, 2018).

⁶ERDDAP. Available online at: <https://coastwatch.pfeg.noaa.gov/erddap/index.html> (accessed November 14, 2018).

TABLE 1 | Key users of the oceanids C2.

Stakeholder	Role
Principal investigator (PI)	Scientist/engineer who instigated the deployment and is ultimately responsible for designing missions to meet their scientific objectives.
Pilot	Highly-trained operator who sends commands to one or more vehicles to achieve the PI's mission. Concerned with the day-to-day operation and health of the vehicle.
Fleet manager	The person in charge of the use of the vehicles within an organization. Concerned with ensuring vehicles, sensors and pilots are able to meet the requirements of the PI. The Fleet Manager may also be a Pilot.
Data assembly centre (DAC) e.g., BODC	The data flow within the C2 data system is delivered by three stakeholders, the PI, Data assembly centre (DAC) and Global Data Assembly Centre (GDAC). The PI provides the data and all metadata to the DAC, where it is assembled into EGO-compliant files where it is forwarded to the GDAC to be made publicly available. Formerly for NERC funded deployments this has been a highly manual and time-consuming operation.
Global data assembly Centre (GDAC)	GDAC are the global data aggregation and community delivery centers that national data centers submit data to. The Ocean glider programme GDAC is hosted by Ifremer. Concerned with ingestion of data in standard accepted formats.
Scientist	Expert end-user of scientific data. May have PI's permission to access NRT data. Concerned with scientific data quality and coverage, metadata availability, standardization of data formats for ease of analysis.
Engineer	Expert end-user of engineering data. Responsible for the maintenance and operation of the MAS system at the vehicle's home institution. Concerned with engineering data quality and coverage, metadata availability, standardization of data formats for ease of analysis.
Reliability analyst	Expert end-user of engineering data. Typically concerned with identifying trends in vehicle reliability across deployments and/or vehicle types from delayed mode data, rather than NRT data from a single mission.
C2 Collaborator	Third-party software developer and end-user of C2 APIs, for example from academic or industrial partner. Concerned with ease of access, standardized formats and representative schemas.
Member of the Public	Under the NERC data policy, data collected by NERC vehicles is ultimately a public good. End-user of NERC data for a wide variety of reasons. Assumed to be non-expert concerned with ease of access and discovering information about data availability.
Policy maker	Requires data products suitable for decision making and regulatory monitoring

the same area of interest, providing significant advances in temporal and spatial resolution compared to traditional single vehicle or ship-based surveys. Ramp et al. (2009) deployed a wide range of gliders, AUVs, aircraft, and ships in Monterey Bay, assimilating NRT data from these platforms into ocean models. The models were then used to provide forecasts for the following day and inform the adaptation of planned survey strategies. The C2 seeks to ease the deployment and management of such large fleets by unifying the command, control and data infrastructure of MAS vehicles within the NMEP. Through well-defined application programming interfaces (APIs), web-based connectivity and quality control of NRT data, the C2 system will enable assimilation of data into ocean models for decision making mid-mission.

Paley et al. (2008) present experimental results of their Glider Co-ordinated Control System (GCCS), which generates waypoints using environmental models to coordinate a fleet of Slocum and Spray gliders. A remote I/O component handles communication to and from the vehicle shore-side servers, enabling a human pilot to monitor the status of a mission, and notifies the user of any software or operational issues. The GCCS was later used on the large scale Adaptive Sampling and Prediction field experiment, which continuously coordinated six gliders for 24 days (Leonard et al., 2010). Whilst this work has similarities with our C2 system, Paley et al. (2008) focused on lower-level control and the prediction of gliders motion, rather than on the development of a common piloting and data infrastructure. Through the development of an Automated Piloting Framework (see section 8.1), the C2 will enable external collaborators to integrate control algorithms which generate generic commands (such as waypoints) which are verified before

translation to vehicle-specific instructions and transmission to the shore-side server for execution on the vehicle. We seek to create a closed loop between data collection, quality control, model assimilation, and adaptive data-driven decision making. Through the definition of industry-standard APIs, we aim to enable the C2 to interface with a wide range of externally-hosted models and control systems developed by C2 collaborators.

3.2. Web-Based Piloting Infrastructure

It is becoming increasingly commonplace for organizations operating fleets of long-endurance MAS vehicles to display NRT data from these platforms via a public webpage, e.g., PLOCAN⁷, MARS⁸, GEOMAR⁹, and ANFOG¹⁰. Such portals are beneficial for both public and stakeholder engagement, enabling end-users to view NRT data and details about different vehicle types. However, it is unclear how many of these portals also enable pilots to interact with the remote vehicle via the web-browser.

The LSTS Toolchain (Pinto et al., 2013; Ferreira et al., 2017) was developed to support the control of heterogeneous fleets of maritime robots, from underwater through to aerial vehicles, communicating across multiple channels for a range of applications including oceanographic survey (Ferreira et al., 2019). The toolchain comprises: *Dune*, a framework for the

⁷PLOCAN Data Portal. Available online at: <http://obsplatforms.plocan.eu/vehicle/USV/test/> (accessed November 11, 2018).

⁸National Oceanography Centre - NMF vehicles. Available online at: <https://mars.noc.ac.uk/> (accessed November 11, 2018).

⁹GEOMAR Navigator. Available online at: <https://waveglider.geomar.de/navigator/> (accessed November 11, 2018).

¹⁰Australian National Facility for Ocean Gliders - ANFOG Glider Fleet. Available online at: <http://anfog.ecm.uwa.edu.au/index.php> (accessed November 11, 2018).

development of embedded software onboard the robot; *IMC*, a common control message format; and *Neptus*, a distributed command and control desktop application for mission planning, execution, monitoring and post mission analysis. In addition, Dias et al. (2006), Faria et al. (2014), and Pinto et al. (2017) describe a Cloud-based piloting server named Ripples, which acts as a communication hub, forwarding data and commands between disparate groups of deployed vehicles and maintaining a global state. The LSTS Toolchain shares the goal of the C2 to provide a common piloting interface to multiple vehicle types, enabling the operation of heterogeneous fleets of vehicles. However, the LSTS is built around a desktop piloting tool, Neptus, with the Ripples layer providing global situation awareness between multiple Neptus workstations and their associated vehicles. The C2 is a fully web-based microservice architecture, where all piloting is performed via a web-interface. The C2 assumes a many-to-many relationship between pilots and vehicles, which may be deployed for months at a time. The Norwegian University of Science and Technology Applied Underwater Laboratory have developed a Mission Control System (MCS) (Buadu et al., 2018) in collaboration with the LSTS toolchain. The main purpose of the MCS is to control formations of fleets of vehicles. MCS can either work in combination with or as a replacement for Neptus in the LSTS toolchain. In a similar way, MCS or similar systems which provide capabilities such as formation movement, could in theory be interfaced with the C2 via the third-party interfaces provided within the Automated Piloting Framework (discussed in section 8.1).

3.3. Ocean Data Delivery Systems

The data processing and delivery of ocean glider data builds on that of more established infrastructures and observation networks. The AtlantOS Strategy for knowledge management, protection and exploitation of results (Reitz et al., 2016) outlines the architecture and stakeholders within the marine community bordering the Atlantic. Buck et al. (2019) outline potential new data users and stakeholders along with the enabling technologies for data standardization. Rudnick et al. (submitted) describe the future vision for ocean glider data over the next decade including alignment of data delivery with the FAIR principles (Wilkinson et al., 2016) and across existing international infrastructures.

4. METHODOLOGY

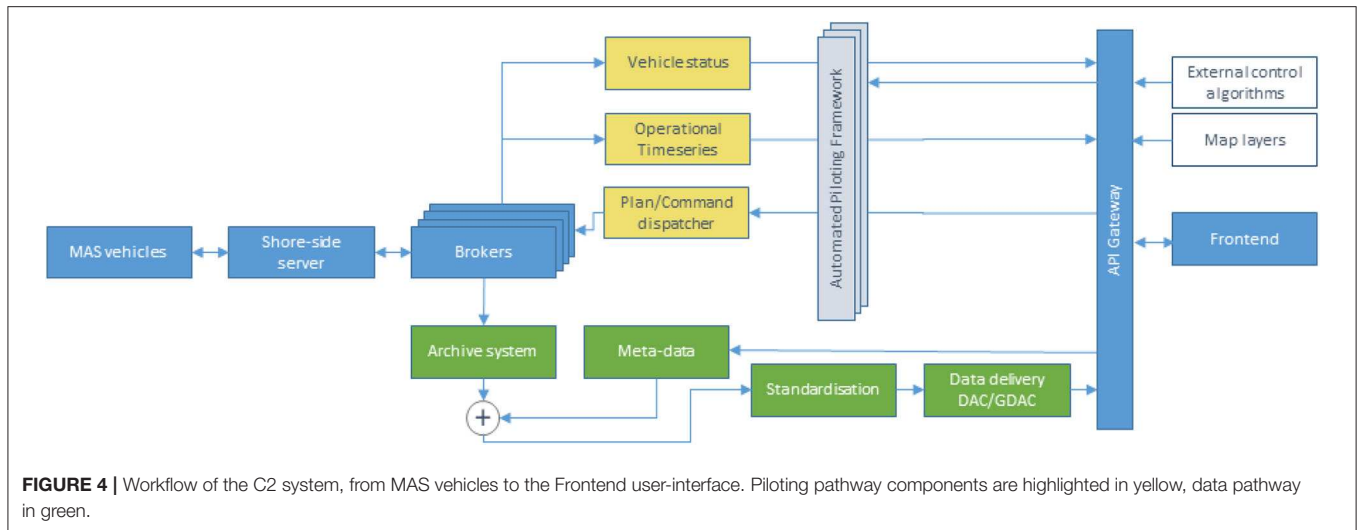
The development ethos of the C2 system centers around the need to capture the continually developing requirements of the diverse range of end-users and project stakeholders (see **Table 1**). To perform cutting-edge research at the forefront of marine science and maximize the use of the available observation technology, PIs and science users require MAS vehicles to be operated in new and innovative ways in increasingly challenging or remote environments. As a result, the development of the C2 system is a multi-institutional effort, in which the user and stakeholder community is engaged throughout the development of the system, using releases of the new system in parallel with existing piloting software and acting as “product owners” for

features relevant to their expertise. At the start of the project, whilst the need for a unified command, control and data system was clear (see section 2), the exact form of the optimal solution was not known. Consequently, the development of the C2 has followed agile software design principles, prioritizing modularity, and extensibility in architecture design, which has enabled us to welcome the evolution of user requirements throughout development and to prioritize features in response to the needs of both users and scheduled autonomous vehicle deployments. As a result, features and releases to-date have focused on underwater gliders and long-range autonomous underwater vehicles, rather than autonomous surface vessels (ASVs). Periodic release cycles enable the user community to benefit from additional functionality, e.g., web-based piloting tools, as soon as these are ready, which provides ongoing user feedback, additional requirements and maintains close communication with the stakeholder community. The C2 development ethos represents a significant shift, from standalone desktop piloting applications, designed for a one-to-one relationship between pilot and MAS vehicle, toward a software-as-a-service model in which the C2 forms a central common component in the operation of over-the-horizon MAS vehicles within the NMEP. This holistic view of MAS piloting and data infrastructure prioritizes maintainability and scalability through identifying and exploiting commonality between vehicle and data types. As an example, gliders from different manufacturers are controlled through different interfaces using different workflows. Whilst these vehicles have some key differences, they share the underlying principles and model of glider operation. Through the identification of shared concepts and the development of common design patterns, we seek to develop a single web-based piloting front-end user interface to enable a pilot to focus on making high-level control decisions to optimize operation of a heterogeneous fleet without first needing training in the many implementation differences between vehicle types. Likewise, data are harmonized from the variety of manufacturer prescribed formats to a single community EGO NetCDF and metadata to Marine SWE profile SensorML format. This simplifies user data access and enable interoperability of data and metadata.

To aid future collaboration with external software developers, e.g., the integration of a new decision making algorithm or data analysis tool, and ensure that users of the MARS fleet continue to benefit from advances in technology made within the wider marine, computing, and robotics communities, the C2 will use open standards, software, and APIs wherever possible.

5. SYSTEM OVERVIEW

Through the creation of well-defined APIs and associated microservices, the C2 infrastructure provides an abstraction on top of custom vehicle control systems, enabling a common web-based user-interface to be used for piloting across the MARS fleet, regardless of vehicle type. By creating a common piloting framework on top of the underlying machine-to-machine APIs, the C2 infrastructure will provide the necessary safeguards and constraints to allow the creation of multi-vehicle co-operative



survey strategies by collaborators and stakeholders, such as the wider robotics and marine science communities.

The solution implements the workflow illustrated in **Figure 4**. The system is an event-driven architecture, which reacts to connections from vehicles, inputs from users and internal/external systems e.g., scheduled data processing. For brevity, we discuss the workflow as initiated by the connection of a remote vehicle, e.g., a glider:

1. The vehicle connects to its associated shore-side server. Shore-side servers are typically manufacturer-provided and are responsible for communicating with the vehicle via its native protocol, which may be proprietary.
2. The brokers sit between the shore-side servers and the C2 infrastructure and provide a common interface to all shore-side servers. A key role of the broker is to distribute data streams, coming to and from the shore-side server, to relevant microservices within the wider C2 infrastructure.
3. The workflow then separates into two logical backend infrastructure “pathways” that appear to end-users as a single unified API gateway and web-based frontend. These pathways are:
 - **Piloting pathway**—encompasses all the infrastructure required for a user or external system to interact with the MAS vehicles and associated shore-side servers. The Vehicle Status service discretizes health information (e.g., battery and positional data) into a state representation for the vehicle reducing the complexity of decision making. The Operational Timeseries service stores the engineering and scientific data from the vehicle within a continuous time-indexed database. Finally, the Plan/Command Dispatcher generates vehicle-specific commands from abstract representations (e.g., waypoints, survey behaviors etc.), shared between vehicle types.
 - **Data pathway**—encapsulates the flow of data to the DAC (BODC). Data from scientific sensors is received from the piloting pathway (based on site at NOC Southampton),

standardized and fused with operational metadata (on site at NOC Liverpool). By developing the data pathway on site at the DAC, the data pathway is able to interface directly with BODC’s existing archive tools, ahead of dissemination to the relevant scientific communities and networks, e.g., via GDACs. Metadata are entered into the metadata system via web forms with values for terms based on controlled vocabularies on the NERC vocabulary server¹¹. The metadata are exposed to the web in SensorML and SSN formats, with snapshots of metadata exposed to the C2 data system in JSON format. Data are archived via a push to the API gateway by the piloting application and, after automated virus checking, are placed in a secure archive. Data are ingested to an intermediate internal NetCDF 4 format, named RXF, for subsequent delivery in a range of standard formats. Data are automatically pushed to the UK Met Office for operational assimilation if the data sets have an open data access policy. Data are delivered in a range of community formats (currently EGO with a desire to include SeaDataNet NetCDF later). Future delivery of data will include API endpoints such as ERDDAP and automated push of files to the Ocean Glider Programme.

5.1. Architecture and Back-End Infrastructure

As a multi-institution project supporting the national infrastructure of the NMEP, it was essential that components of the C2 could be easily portable and reconfigurable between partner institutions, scaling with the addition of new vehicles and data streams. For example, pilots from the Scottish Association of Marine Science (SAMS) control Seagliders within the MARS fleet, so commands and data from the C2 system need to be sent to the Seaglider shore-side server at SAMS, and vice versa. It was considered that the easiest way to communicate between

¹¹NERC Vocabulary Server. Available online at: https://www.bodc.ac.uk/resources/products/web_services/vocab/ (accessed November 11, 2018).

institutions given their independently-managed IT networks was to develop a web-based system, as it is rare for outgoing web traffic to be blocked by institutional firewalls, thus lowering the barrier of entry for participating institutions.

5.2. Microservices

In contrast to traditional monolithic systems, the architecture of our C2 system is highly modular and extensible, comprising a collection of microservices which interact through HTTP REST APIs and the event-driven Advanced Message Queuing Protocol (AMQP) via RabbitMQ¹². Most of our microservices provide an API which can be consumed by anyone with access to a HTTP client and has the requisite permissions, while the event-driven component of the system (the “event-bus”) allows the microservices and external API consumers to listen for and react to specific events as soon as they happen, which is more reactive and efficient than polling the HTTP APIs for changes. One example of this is a service which sends SMS messages to pilots as soon as one of the vehicles connects to the system, allowing pilots to respond quickly to intermittent events when these occur, rather than having to constantly watch a computer. By exposing the event-bus using the EventSource (Hickson, 2015) and WebSocket standards (Fette and Melnikov, 2011), we have developed a web-based front-end to the system which can react to changes instantaneously.

Most microservices are running on top of the Docker containerization software, which creates a consistent and immutable environment for each service regardless of the host hardware and operating system used. Docker images are portable by nature, and allow third-party services to be easily included in the system, as many popular server software packages provide Docker images on Docker Hub¹³. This reduces the complexity of installing and configuring the constituent parts of the system and the supporting software, so that it is easier to port the system to different environments, such as different Cloud services. The Microservices architecture separates the system into small lightly-coupled, highly reconfigurable functional units. This provides a number of key advantages for the C2 system:

- **Increased portability**—Component microservices can be deployed to run either on the Cloud or on servers based within partner institutions (or a mixture of the two, known as Hybrid-Cloud). This is powerful for the C2 because it does not restrict the system to a particular infrastructure or service provider, enabling us to adapt to evolving technical and financial requirements. Cloud-based infrastructure has clear advantages in terms of power and maintenance: professional Cloud platforms such as Microsoft’s Azure and Amazon Web Services provide infrastructure at a scale generally not achievable by any single academic organization.
- **Increased scalability and resilience**—The division of functionality into microservices makes it straightforward to add or update component services without first having to take

other parts of the system down. This reduction in down-time is important as MAS operations take place around the clock throughout the year. The modularity of the microservices architecture also eases integration of the C2 with other external systems, as the system can be split across many processes which can be started and stopped independently of each other.

- **Increased flexibility**—As independent components of the system, microservices can make use of different underlying technologies, selected to best-suit their individual function rather than needing to be supported by, and relevant to, all components. For example, within the C2 system the Timeseries microservice uses the open-source Timescale extension¹⁴ to PostgreSQL, which optimizes the underlying database technology for indexing and searching large time-series datasets. However, due to the modularity of the microservice pattern, this database extension does not need to be used by all other services within the C2.
- **Ease of development**—The modularity of the microservice architecture is well-suited to distributed development teams, as is common within academic or research consortia such as the C2 project, because each service has a separate code-base and thus is smaller and easier to develop, debug and understand.

The majority of the C2 ecosystem is managed by the Kubernetes orchestration software¹⁵, which is responsible for managing resources given to each service, monitoring health and acting on service failures, providing an increased level of resilience to the system.

5.3. API Gateway

To reduce the apparent complexity of a growing collection of microservices, we utilize an API Gateway to present an end-user or external system with a single coherent access point to the C2. This pattern abstracts the user from the underlying architecture and technologies within the C2—reconfiguration and addition of microservices can occur without impacting the end-user and ensures the system remains portable, i.e., the user does not need to change their access route in response to a change to on-site or Cloud hosting. The gateway secures and controls access to different end-users, allowing them to see relevant parts of the system and to meet the needs of the NERC data policy¹⁶.

All APIs being served via the gateway have publicly available Swagger¹⁷ definitions. These definitions are independent of the code and allow both internal and external developers to see all available methods along with a list of the required and optional input variables. The definitions include notes on method implementation, descriptions of method inputs for both required and optional variables, and reason for HTTP status codes. We

¹²RabbitMQ - open source message broker. Available online at: <https://www.rabbitmq.com/> (accessed November 12, 2018).

¹³Docker Hub - Dev-test pipeline automation. Available online at: <https://hub.docker.com/> (accessed November 12, 2018).

¹⁴Timescale—Open-source time-series database powered by PostgreSQL. Available online at: <https://www.timescale.com/> (accessed November 12, 2018).

¹⁵Kubernetes. Available online at: <https://kubernetes.io/> (accessed November 14, 2018).

¹⁶Natural Environment Research Council (NERC) data policy. Available online at: <https://nerc.ukri.org/research/sites/data/policy/data-policy/> (accessed November 12, 2018).

¹⁷Swagger. Available online at: <https://swagger.io/> (accessed November 14, 2018).

have also used Swagger to model the JSON formatted data we pass around the C2 system.

5.4. Security

The design of the C2 system seeks to minimize the risk of external agents disrupting or interfering with either the vehicle, mission or data:

- *Security of the vehicle*—For commercially-designed MAS systems, the security of both onboard systems and the satellite link between the vehicle and its shore-side servers are often proprietary and limited to what is provided by the manufacturer or satellite provider. This creates challenges for securing the entire data chain from vehicle to data assembly center, as end-to-end encryption, for example, is not possible if not supported by the vehicle. We are currently investigating options to increase security at this point in the chain, for example by creating a VPN between the shore-side servers and the satellite service provider.
- *Integrity of the mission*—The C2 reduces the risk of unauthorized interference with the mission and control of the vehicle through authentication¹⁸ and authorization of users¹⁹. In addition, we are implementing different pilot user-profiles, which limit the access of the user to only the vehicles and commands they have been trained to use.
- *Security of data interfaces*—In order to ensure all transmitted data comes from a trusted source, the API gateway requires a token on all API calls for the pushing of data to the C2 system. These tokens are then decoded and stored as part of an audit trail allowing the system to link all incoming data to a user account. The same is true for the metadata APIs. All connections to the gateway are encrypted using industry-standard encryption certificates.

6. PILOTING PATHWAY

The piloting pathway facilitates interaction between the user, the vehicle shore-side server and the vehicle itself. The piloting pathway consists of multiple services which convert from vehicle-specific data formats to normalized abstract representations and vice versa. For example, Seaglidors and Slocums report their positions or accept waypoints in different vehicle-specific formats, but an event or plan within the C2 system captures that information in a generalized representation which is vehicle-agnostic. Farley et al. (2019) provides a comparison of the existing piloting workflows of Slocum gliders, Seaglidors, and Autosub Long Range, alongside their resulting unification into a single common workflow.

At the time of writing, there are 22 microservices within the C2 architecture providing APIs for piloting and data processing functionality. Of these, there are several which are crucial to the piloting pathway:

- **Vehicle status**—On receipt of communication streams from a remote vehicle via the brokers (see **Figure 4**), the Vehicle Status service extracts key information from these communications including: the vehicle's GPS position, the time according to its internal clocks, any engineering data provided by the vehicle and relevant information about vehicle behavior, e.g., the current waypoint, or mission state. This data stream is then discretized into vehicle events, representing a change in vehicle or connection state. For example, a vehicle connecting to the C2 or aborting whilst connected to the C2 would result in a change in Vehicle Status, whilst continuous changes to vehicle battery would not unless this fell below a pre-defined threshold, triggering a low-battery vehicle state. The formats and protocols containing vehicle communications differ for each type of vehicle: Iridium Short-Burst Data (SBD) messages are generally received via email, which contains the Iridium geolocation as well as the vehicle-specific formatted data, usually in a concise binary format; Slocum gliders output information in a continuous stream of text over a long-running satellite modem connection, where the order in which text is received is not fixed. Other devices attached to a vehicle, such as third-party GPS trackers, e.g., Argos²⁰, may present their information via other web APIs in more common formats such as JSON, CSV, or XML. The Vehicle Status service parses these disparate, heterogeneous streams of information as they arrive within the C2 and generates a timeline of communication events per vehicle in a normalized format. This can then be used by the rest of the system without having to decode the unique formats provided by different vehicle types. For instance, a position report will provide latitude, longitude, timestamp, and radius of GPS accuracy in the same format, regardless of which vehicle or communications medium the original message was sent from. Wherever possible engineering data is normalized across vehicle types.
- **Operational timeseries**—The Timeseries service provides an efficient database in which to store and query continuous numerical engineering and science data arriving from each vehicle, such as battery voltages, altimetry ranges and uncalibrated CTD data. Similar to the Vehicle Status service, the Timeseries accepts data files from the vehicles in their native format and normalizes observation points into PostgreSQL database rows, which can then be output by the API in whichever format is suitable for an end-user (such as comma-separated values, CSV)²¹. The raw data coming from the entire MARS fleet across all deployments has the potential to be stored in this database, allowing us to perform complex aggregate queries involving one or many variables across one or many vehicles. Whilst this service provides a similar functionality to the data pathway, it should be noted that as this service is designed to facilitate operational access to data mid-mission, it only stores the raw data from vehicles without

¹⁸ OAuth Community Site. Available online at: <https://oauth.net/> (accessed April 23, 2020).

¹⁹ Open Policy Agent. Available online at: <https://www.openpolicyagent.org/> (accessed April 23, 2020).

²⁰ Argos - Worldwide tracking and environmental monitoring by satellite. Available online at: <http://www.argos-system.org/> (accessed November 12, 2018).

²¹ Timescale - Open-source time-series database powered by PostgreSQL. Available online at: <https://www.timescale.com/> (accessed November 12, 2018).

delayed-mode data ingestion or quality-control pipelines in place. This service is useful for reacting to engineering and preliminary scientific data as soon as it is received from the vehicles, for quick access and analysis by pilots or to inform planning algorithms whilst the vehicle is still on the surface. The raw engineering data stored within Timeseries may also be used to inform reliability studies post-deployment.

- **Plan/Command dispatcher**—This is comprised of three services: a planner service stores vehicle-agnostic plans (abstract lists of behaviors, such as lists of waypoints or instructions to switch sensors on and off). The compiler service can then be used to convert these plans into the formats required by individual vehicle types. As microservices are designed to be stateless and provide individual functionalities, to manage the flow of execution required to first generate a plan and then send it to a vehicle, a third service, the Conductor, accesses these services in turn and sends the generated output to the data brokers. The data brokers then forward the vehicle-specific plans to the required shore-side servers. The Conductor monitors the state of the plan throughout creation and dispatch, informing users as each event takes place. The Conductor can halt and revert the process to an earlier step should an error occur at any point. This transforms a complex sequence of operations into what appears to external users as a single operation.

6.1. Piloting Frontend—User interface

At present, each of the different models of long-range vehicles within the MARS fleet are piloted with a different bespoke user-interface as provided by the vendor. These solutions range from proprietary desktop applications to sending modified files via FTP/SSH to shore-side servers. The significant variation in piloting interface results in a steep learning-curve for pilots. Consequently, a new pilot has to undergo training for each of the bespoke tools and operating procedures for each model of vehicle within the fleet.

The C2 system will unify piloting by implementing a single web-based user interface for the piloting of all over-the-horizon vehicles within the MARS fleet. Whilst each vehicle has different capabilities and modes of operation, preventing the use of an identical piloting interface for all platforms, we have identified commonality between vehicle workflows, developing design patterns to be used throughout the user interface and shared by all vehicles.

Where possible, we are following a user-interface design language, Material Design, developed by Google²² and used throughout their products including Google Maps, Docs, and Gmail. As these products are incredibly widely-used by the general public, the behavior and association of many user interface elements, from buttons and sidebars to icons and use of color, will already be very familiar to many users of the C2. Through this consistency with existing widely-used products, we aim to reduce the learning-curve associated with the C2 system, enabling trainee pilots to focus on learning the

technical operation and control of the vehicle itself, rather than how to navigate the website and interact with online forms, plots etc.

During deployments of MAS vehicles, pilots are on-call to check and ensure the safety of the platform at all times. Consequently, the ability to monitor mission progress from a mobile device was a requirement identified following engagement with pilot users. Material Design components scale and rearrange in consistent ways, familiar to users of Google apps, enabling use of the piloting interface on mobiles and tablets without the need for the development of a separate mobile-optimized front-end interface.

Due to the broad nature of the C2 web-based frontend as a portal for piloting, data discovery and analysis, there have been occasions where the existing Material Design standards did not cover our use-cases and therefore we have had to extend the principles. For example, we extended the concept of the top-level navigation bar to accommodate nested layers of navigation within different parts of the app. This is illustrated in **Figure 5**, where the second-level navigation bar represents functionality within the “Pilot” app, currently selected in the Top-level navigation bar. Equally, Material Design only provides a specification for simple in-browser pop-ups to notify the user of an event. However, as piloting is a high-risk activity requiring active user-engagement, our system needs to be able to interrupt the user in the event of an abort, for instance. Consequently, we have made use of native system notifications to notify the user of any events whilst the piloting interface is minimized or behind another application window (e.g., whilst the user is currently using a different application on their device). However, the underlying philosophy has been maintained and is in-line with the recent general broadening of the Material Design standard.

Figure 5 shows a labeled wireframe of the C2 user-interface, illustrating key design patterns, common to the piloting interfaces of all vehicles within the C2. In the remainder of this section, we discuss these key concepts in more detail.

6.2. Timeline

Whilst data collected by MAS vehicles is continuous and time-varying, vehicles typically connect to shore-side servers at discrete points in time. Within the C2, the continuous data from the vehicle is labeled according to these communication events, enabling a pilot to easily evaluate the data over the period since they last communicated with the vehicle, interpreting current health, and data quality.

To enable the pilot to explore data from earlier stages in the mission and to analyse trends across a period of data, the C2 front-end uses a Timeline design pattern (see **Figure 5**). The Timeline is a slider component allowing a user to move the state of the C2 to a particular point in time, or to specify a particular historic time range. This makes analysis of events and associated causality faster. Markers for specific events, such as changes in mission or platform errors, are also able to be added to the timeline.

²²Material Design. Available online at: <https://material.io/design/> (accessed November 11, 2018).

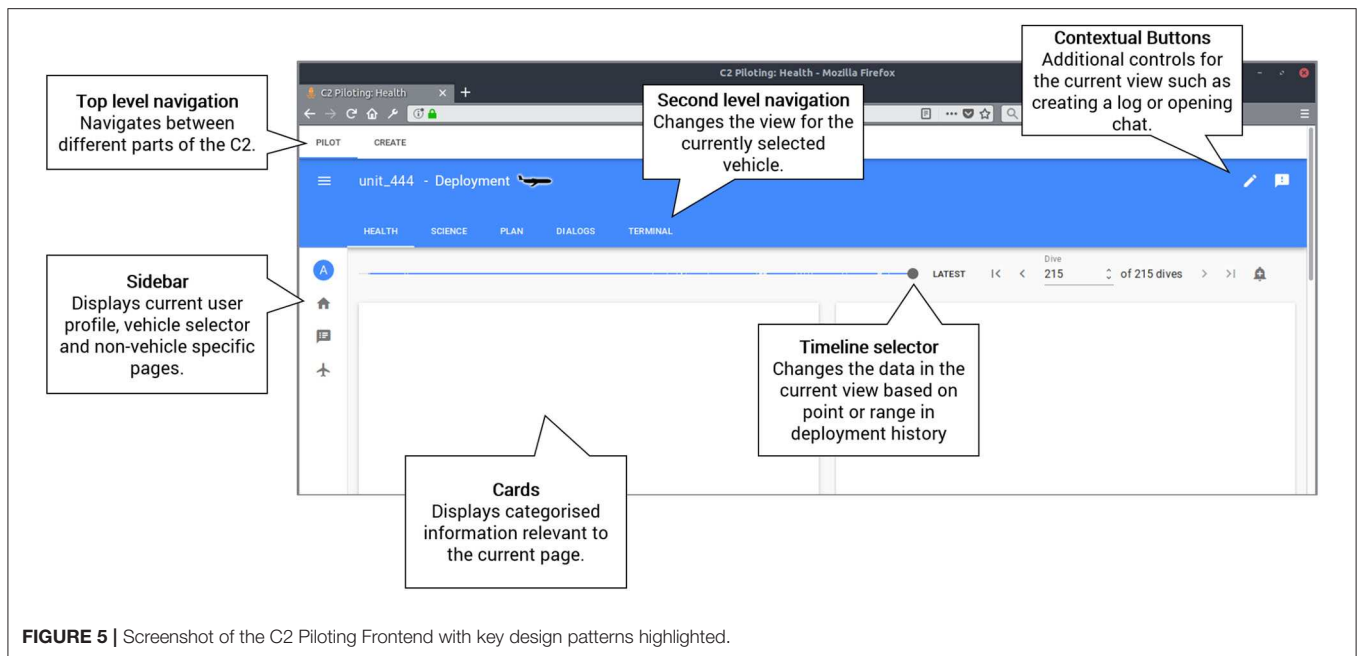


FIGURE 5 | Screenshot of the C2 Piloting Frontend with key design patterns highlighted.

6.3. Piloting Navigation

Within the piloting interface for each vehicle, the user is presented with four tabs, the concept of which is common to all vehicle types—“Health,” “Science,” “Plan,” and “Dialogs”:

- **Health** displays Material Design cards which summarize and plot engineering data crucial to evaluating the current status of the vehicle, including battery voltage, device reports, and GPS position.
- **Science** comprises cards which display interactive plots of data from the science sensor suite onboard the vehicle. Some of these cards are common to all vehicles (e.g., CTD) whilst some will be vehicle or deployment specific.
- **Plan** is a dynamic interface which is customized to the vehicle type and capabilities but ultimately enables the user to alter waypoints or define new missions for the vehicle via a common mission planning and execution interface.
- **Dialogs** summarizes and displays the raw contents of human-readable communication from the vehicle. Whilst the health page provides a single-page overview for assessing the vehicle health, the dialogs page enables the pilot to review the specific details of each connection to further investigate faults.

In addition to the above tabs, displayed for all vehicle types, vehicle-specific tabs may be included where necessary. For instance, the piloting interface for the Slocum glider (Schofield et al., 2007) includes a “Terminal” tab, which implements a bidirectional command-line interface to the glider allowing an expert pilot to take full control of the glider when required. Integrating the terminal into the C2 removes the need for a second interface for interacting directly with the vehicle. However, such functionality is conditional upon user-permissions, so trainee glider pilots will not be able to send commands which may endanger the safety of the platform.

6.4. Map

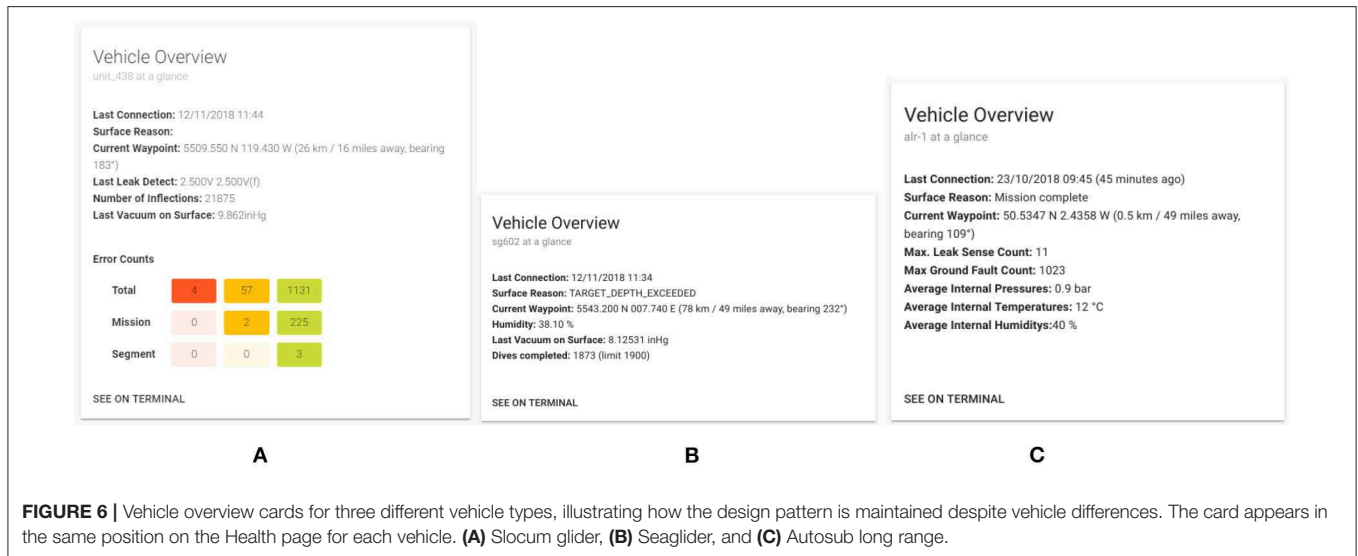
The map card consists of an interactive map implemented with Leaflet²³ which displays the current and previous positions of remote vehicles. Such data will automatically update when a vehicle sends its position to its shore-side server, enabling the pilot to keep track of progress or positioning amongst a fleet.

Users of the C2 are able to add and customize additional map layers to suit specific vehicle and mission requirements, including the uploading of image, video and Geographical Information System (GIS) layers and third-party TileLayers displaying Automatic Identification System (AIS) data. Users are able to set which layers are displayed by default such as isobaths or ice coverage layers depending on the area of operation. With such data sources available to the pilot, the map is also used for designing and verifying mission plans by allowing the pilot to annotate behaviors such as transit behaviors directly on the map before being compiled as a list of waypoints and sent to the vehicle.

6.5. Vehicle Overview Card

The Vehicle Overview card summarizes the most critical vehicle parameters, providing a concise snapshot of vehicle health (see Figure 6). The Vehicle Overview card appears in a consistent position within the piloting interface for each platform. The aim of the card is to alert the pilot to critical events such as water leaks and aborts, which need to be handled urgently. The other cards within the piloting interface then allow the pilot to further investigate the suspected cause of the critical event.

²³Leaflet - an open-source JavaScript library for mobile-friendly interactive maps. Available online at: <https://leafletjs.com/> (accessed November 11, 2018).



6.6. Logs

The accurate logging of pilot actions and metadata associated with MAS deployments is essential for investigating vehicle faults and improving long-term reliability of the MARS fleet.

The creation of log entries needs to be quick and easy, as a pilot is likely to be under time pressure as a result of interacting with the vehicle at the same time, and may be under additional stress if a fault or leak has been identified. As a result, we have developed a consistent logging design pattern throughout the C2 interface, which consists of a button on all pages which opens a floating log entry alongside the current view. The user may continue to navigate between pages and pilot vehicles whilst the log entry is open, without losing the contents of a half-written log.

We are in the process of developing context-specific log functionality which will automatically suggest annotating the log entry with labels relating to the contents of the page the user is currently interacting with. For example, if a log entry is created on the Health tab for a glider whilst the vehicle is displaying an abort and a leak, the log entry will automatically be tagged with the vehicle name, ID, abort, and leak values. The user can then remove these tags with a single-click or add their own. At present, the body of the log-entry is a free-form text box. However, we intend to develop standard templates to prompt the user for key information and to ease subsequent quantitative analysis of log records.

6.7. Interactive Plots

To ensure the platform is navigating and collecting data correctly, pilots must monitor various engineering and scientific data sets that are sent back from the platforms. This can include: multiple flight parameters to allow for the correction of navigation behavior, science data samples to ensure the instruments are working correctly, and long-term engineering data to monitor deterioration in devices and calibrations.

Visualizing this data in plots allows pilots to efficiently analyse the data and make corrections quickly when necessary. However,

with deployments as long as six months, and data sampling as often as one sample every 15 s for as many as 20 variables, the amount of data needed to be visualized, especially while looking for long-term trends, can be very large. Such large datasets can be detrimental to the performance of a web application, especially on mobile browsers, requiring large loading and rendering times. To reduce the amount of data being sent to the client, data is windowed and aggregate statistics (such as 10 min averages) for each window are created for the requested time range. A pilot can then zoom in on narrower time ranges to download more detailed information for those particular ranges, as illustrated in Figure 7.

The `plotly.js` module²⁴ is used to build these plots in the C2 and allows for interaction with the data, such as extracting particular values or toggling the display of variables on the plots, to make analysis easier and reduce visual clutter. Graphs for specific platforms have been replicated in the C2 from existing piloting tools to make the transition to the C2 interface easier for existing users.

7. DATA PATHWAY

The data pathway within the C2 system is being jointly implemented by BODC and SAMS. In the same way as the piloting pathway, the data pathway is accessed via the common API gateway (see section 5.3) which enables the forwarding and delivery of both metadata and data from the vehicles and front-end user interface to the data-assembly center, BODC.

The requirements discussed in section 2 led to key architecture design decisions, namely: a target intermediate NetCDF4-based format devised for the data pathway called RXF, an archive system, exchange of metadata in JSON structures, and a scheduling algorithm to allow prioritization of the most pertinent

²⁴plotly.js - the open source JavaScript graphing library. Available online at: <https://plot.ly/javascript/> (accessed November 11, 2018).

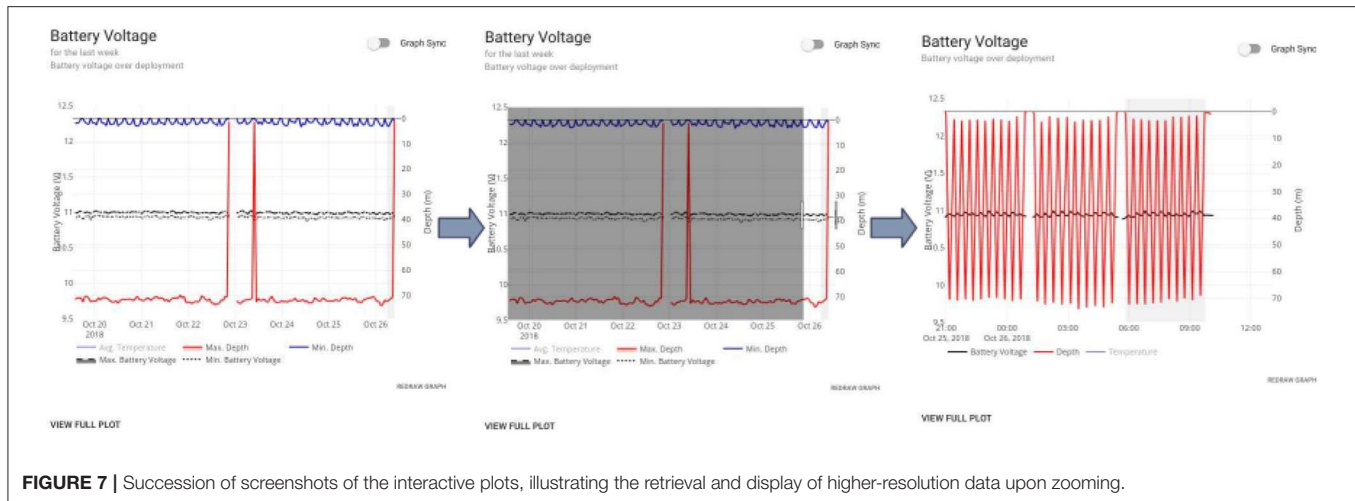


FIGURE 7 | Succession of screenshots of the interactive plots, illustrating the retrieval and display of higher-resolution data upon zooming.

data streams. RXF is a holding format used internally within BODC to allow for long-term stability regardless of updates and changes to external formats. RXF is flexible enough to support subsequent conversion to any required output format, such as EGO NetCDF. **Figure 4** highlights components of the C2 data system in green.

The workflow for the processing of both NRT and initial delayed-mode data is as follows:

1. Data are delivered from the shore-side servers to the archive system via the brokers throughout the mission, on the surfacing of the vehicle and completion of each mission.
2. On arrival at the archive system, the data are scanned by a virus checker. The archive system performs automated replication of the data to another site to enable cross-site data system redundancy and secure data archival.
3. Once data are registered in the archive they are ingested into the standardization system in RXF format.
4. On production of the RXF, the result is copied back to the archive to provide resilience. The RXF is built incrementally, with new data added to the existing file at each stage.
5. Once the RXF has been copied to the archive, an EGO file can be produced by the standardization system. This is not done incrementally but restarted for each update of the RXF. This is because EGO is a NetCDF3 format and would need to grow along more than one dimension, something that NetCDF3 does not allow.

Each element of the C2 data system will now be described in detail.

7.1. Metadata Service

The presence of sufficiently detailed metadata, and its subsequent availability on the web, is essential for data discovery and identification, for example by providing key information on sensor variables, units etc. At present, engineers and PIs (see **Table 1**) enter campaign and deployment information prior to the deployment of MAS vehicles, enabling automated data processing to begin when vehicles are at sea. However, the

longer-term aim is to extract metadata directly from both the MAS vehicles as well as NOC calibration and inventory management systems, removing the need for a user to enter this data, reducing error. Consequently, the C2 system has been designed to support this future automated acquisition of metadata and work has started on developing the necessary interfaces between the C2 and existing NOC systems.

The exposure of metadata to the web builds on a prototype developed by the European Commission supported SenseOCEAN project (Kokkinaki et al., 2016; Martínez et al., 2017) and is accessed according to open standards including World Wide Web Consortium (W3C) RDF/XML and the use of the Semantic Sensor Network (SSN) ontology (Haller et al., 2017) and Open Geospatial Consortium (OGC) SensorML (Open Geospatial Consortium, 2016) standard. Metadata exposure via SSN and SensorML is achieved using a database built using pre-existing ontologies and terms from the NERC vocabulary server 2.0 (NVS2) (Kokkinaki et al., 2016), ensuring consistency with other NERC data sources. Data will be delivered in EGO NetCDF format (EGO gliders data management team, 2017) and the OGC Observations and Measurements standard. Additional formats will be served by implementation of endpoints such as the NOAA ERDDAP tool²⁵.

7.2. Archive System

BODC is a ICSU World Data Systems accredited data center²⁶. When data is archived at BODC it is stored in duplicate with a checksum to ensure long-term data integrity within the BODC archive, with a third copy stored offsite as a backup.

Archive initiation is automatic and is triggered by the deployment of a MAS vehicle and the periodic arrival of mission data throughout the deployment. A security token enables data to be pushed to the archive from the shore-side server, via the

²⁵ERDDAP. Available at: <https://coastwatch.pfeg.noaa.gov/erddap/index.html> (accessed November 13, 2018).

²⁶ICSU World Data System. Available online at: <https://www.icsu-wds.org/services/certification> (accessed November 13, 2018).

brokers, and combined with metadata entered in advance of the deployment by the PI.

7.3. Standardization—EGO Production

Ingestion into the standardization system requires the creation of a mapping from vehicle-specific variable names to BODC's P01 vocabulary²⁷, which defines terms for observed properties, e.g., salinity. These mappings are encoded as JSON. To date, we have implemented mappings for both Slocum and Seagliders. Slocum data arrives in a proprietary binary format and is ingested after first being translated into ASCII. Seaglider data is already formatted as NetCDF and therefore no decompression or processing of raw data is required prior to ingestion.

Prior to the production of the EGO formatted data, all data in the RXF must be merged to a common, resolved time channel, as required by the EGO format. The current EGO specification requires NetCDF3 which, as mentioned previously, means that unlike the RXF, the EGO file cannot be grown incrementally. Following the intention of the EGO designers to make the file format largely self-sufficient, EGO contains an abundance of metadata (unlike RXF). Such information is communicated via JSON, utilizing the BODC vocabulary P06, which defines units of measurement²⁸. If variable units require scaling, this is also performed at this stage.

Production of EGO files within the C2 data system has been shown to be relatively fast, taking no more than 30 s for a 100 MB file using existing BODC server infrastructure. Given the slow speed of MAS vehicle missions (gliders and Autosub Long Range move at a slow walking pace) and the time taken to transmit data via Iridium, the time required for EGO production is considered negligible.

7.4. Data Delivery

The standardization of data delivery via the API gateway enables timely NRT data to be served to users of the C2, BODC users, and operational users²⁹. The use of open standards will also enable web interfaces to be rapidly built on the API gateway and the delivery of data to European research infrastructures that include aligned reference models for data infrastructure. The C2 data system contributes data and metadata directly to the OceanGlider (Testor et al., 2019) network which enables the data to be included in Copernicus³⁰ and EMODnet³¹ data products.

The combination of the C2 data system and the underlying database developed within the SenseOCEAN project (Martínez et al., 2017) is aligned with many of the FAIR principles (Wilkinson et al., 2016). Findability, accessibility, and interoperability of data will be archived via dissemination of data to the Ocean Glider Network where unique identifiers

will be assigned. Findability of metadata is partially archived with unique sensor/platform/deployment identifiers at the data centre level, pending the outcomes of the Research Data Alliance (RDA)³² working group on the persistent identification of instruments³³. Accessibility and re-usability of metadata is archived and exposed via the OGC Sensor Observation Service (Bröring et al., 2012). Interoperable metadata is achieved via the use of a the marine Sensor Web Enablement³⁴ profile for SensorML that includes use of the NERC vocabulary server for terms and values. Data reusability is partially obtained via the use of community standards and on-going work to develop a common data access policy.

At the time of writing, the C2 data system metadata is available via the 52° North GmbH variant³⁵ of the OGC SOS service. Future machine-to-machine services include making data available via the 52° North SOS server, addition of the data to an ERDDAP server, a W3C Linked-data data catalogue vocabulary (DCAT) exposure of data with a SPARQL endpoint, and exposure of metadata in W3C Semantic Sensor Network (SSN).

8. ONGOING AND FUTURE WORK

8.1. Automated Piloting Framework

Many aspects of long-range MAS piloting are routine, defined by standard-operating-procedures (SOPs), and require the pilot to check that various vehicle parameters are within the nominal range. Such checks would be straightforward to automate and would reduce the probability of human-error, particularly when piloting out-of-hours or in adverse environments, such as on a moving boat (Stokey et al., 1999). By automating these time-consuming and well-defined checks, we will increase the capacity of expert pilots to focus on meeting the science goals of the deployment and making more complicated fleet-level piloting decisions. By placing this autonomy on servers rather than on the vehicle itself, we enable the entire process to be monitored, interrupted and overridden by a human pilot as required. MAS vehicles will make contact with the C2 via satellite, as at present, receiving commands issued by either the automated piloting system or a human pilot.

Maintaining a level of human oversight and influence over the piloting process is crucial for the accountability of MAS operations. If, at any point, the status of the vehicle is determined to be outside the nominal range (as defined by the SOPs) the automated piloting system will escalate piloting decisions to the supervising human pilot and cease issuing commands. Once the human pilot is satisfied that the status of the vehicle has returned

²⁷NERC Vocabulary Server, P01. Available online at: <http://vocab.nerc.ac.uk/collection/P01/current/> (accessed November 13, 2018).

²⁸NERC Vocabulary Server, P06. Available online at: <http://vocab.nerc.ac.uk/collection/P06/current/> (accessed November 13, 2018).

²⁹OceanGlider Data Assembly Centre (DAC) for UK glider deployments. Available online at: https://www.bodc.ac.uk/data/bodc_database/gliders/.

³⁰Copernicus Marine Environment Monitoring Service. Available online at: <http://www.copernicus.eu/main/marine-monitoring> (accessed November 14, 2018).

³¹EMODnet. Available online at: <http://www.emodnet.eu/> (accessed November 14, 2018).

³²Research Data Alliance. Available online at: <https://www.rd-alliance.org/> (accessed November 14, 2018).

³³RDA Persistent Identification of Instruments working group. Available online at: <https://www.rd-alliance.org/groups/persistent-identification-instruments-wg> (accessed November 14, 2018).

³⁴Open Geospatial Consortium, Sensor Web Enablement. Available online at: <http://www.opengeospatial.org/domain/swe#initiative> (accessed November 14, 2018).

³⁵52° North Sensor Observation Service. Available online at: <https://52north.org/software/software-projects/sos/> (accessed November 14, 2018).

to normal, the human pilot may again delegate control back to the automated piloting system. We therefore require a mixed-initiative approach in which the human pilot may set the level of piloting responsibility to delegate to the automated system. We envisage the use of concepts such as user-adjustable autonomy, as presented by Ai-Chang et al. (2004) in their MAPGEN system for Mars rover missions.

To support the collaborative aims of the C2 development, the framework implements clear interfaces which third-party developers, scientists or engineers may use to develop and integrate their own automated piloting algorithms. The framework ensures that such mission plans created by third-party developers are compliant with regulatory requirements and NOC SOPs, keeping mission risk to within an acceptable level. If any constraints are violated by a proposed plan, the automated piloting framework alerts an expert pilot and ceases sending automated commands. Through these interfaces, additional capabilities such as intelligent COLREGS behaviors for LRUSVs, current-aware piloting algorithms and task scheduling approaches can be integrated into the C2.

We are also working to address the need for ongoing monitoring and detection of adverse performance trends, highlighted by Thieme and Utne (2017), who state that “Unanticipated faults and events might lead to loss of vessels, transported goods, collected scientific data, and business reputation. Hence, systems have to be in place that monitor the safety performance of operation and indicate if it drifts into an intolerable safety level,” a view shared by the Maritime Autonomous Surface Ships Industry Conduct Principles Code of Practice³⁶, section 8.8.1. In collaboration with University College London (UCL), we are currently developing and integrating two systems to address this need:

1. An automated system for determining the optimal flight parameters for Seagliders (Anderlini et al., 2019), which present results to either a human pilot as a recommendation, or can be sent directly to the vehicle, depending on the user-defined level of autonomy. This will enable around the clock operations, reducing the substantial piloting overhead at the start of a mission and allowing more vehicles to be deployed simultaneously. We are planning to test the system at sea in Summer 2020.
2. A monitoring system for gliders in which the data provided by the C2 system is analyzed and compared to the glider flight model using statistical and machine learning methods to identify adverse conditions including biofouling and wing-loss (Anderlini et al., submitted). Early results from the approach have been very promising, with the detection of conditions which have previously been challenging for pilots to manually identify.

Investigating why the AUV community has yet to widely adopt adaptive mission planning, Brito et al. (2012) found uncertain

vehicle behaviors to be the largest concern (39.7%) of expert AUV operators. The second largest concern (20.7%) was that the technology is not understood, with the majority of respondents reporting that it has not been well explained. To address this concern and to build confidence in the automated piloting framework amongst both MAS pilots and stakeholders, decisions made by the system must be clearly communicated to the pilot at all times and fully logged. To this end, we are investigating the use of techniques such as explainable planning (Fox et al., 2017) and robot transparency (Wortham et al., 2017).

8.2. Risk and Reliability Modeling

By unifying the piloting interfaces and data flow across the long-range MARS fleet, the C2 offers an opportunity to record, quantify and analyse vehicle reliability on a scale that was previously impossible. Brito and Griffiths' (2009), Brito et al. (2010) previously created a risk model for the Autosub3 AUV. However, due to the lack of objective data and metadata on faults, mission success and vehicle performance, the model was constructed using expert subjective judgment. Whilst this work informed successful mitigation for under-ice work, the authors highlighted the difficulties associated with using expert judgment, stating that a panel of experts provided probabilities of a fault leading to vehicle loss that spanned three orders of magnitude. The C2 system will address this gap in objective data by automatically logging and monitoring vehicle connections, events, and constraint violations, reducing the need for the pilot to accurately record this information manually. The development of data-driven techniques and risk models to extend Brito and Griffiths' (2010) previous work is an active area of our ongoing research. By integrating context-specific logging functionality into the common piloting frontend (see section 6.6), rather than requiring the pilot to log-in to a separate system or resort to pen and paper, we are seeking to minimize the overheads associated with accurate fault and event recording, promoting uptake. Through the C2 APIs, we seek to aggregate vehicle and fault data with data from our inventory management and ship-programme systems, enabling a holistic identification and characterization of the elements involved in faults, such as planned maintenance, equipment history, and deployment location (Dopico-Gonzalez et al., 2019). Once we have sufficient data, we aim to incorporate real-time risk analysis and feedback into the piloting process, alerting the pilot, PI, or automated piloting algorithm when the planned mission is calculated to increase risk to the vehicle or its scientific data cargo.

8.3. Notable Ocean Deployments of the C2 System

To evaluate the design and usability of the Automated Piloting Framework, we created an external path-planning algorithm that uses AI techniques to calculate the optimal trajectory through an area of high-current, using the UK Met Office Forecast Ocean Assimilation Model (FOAM)³⁷ forecasts. This algorithm

³⁶Maritime Autonomous Surface Ships Industry Conduct Principles Code of Practice, Available online at: <https://www.maritimeuk.org/media-centre/publications/maritime-autonomous-surface-ships-industry-conduct-principles-code-practice/> (accessed April 27, 2020).

³⁷E.U. Copernicus Marine Service Information - Atlantic - European North-West shelf ocean physics analysis and forecast. Available online at: <https://tinyurl.com/y82gbodh> (accessed November 12, 2018).

connected to the C2 via the APIs, and was successfully used to command a glider on deployment in the North Sea in Spring 2019 alongside the AlterEco³⁸ project gliders. This deployment proved the concept of the Automated Piloting Framework and engaged with scientific stakeholders, allowing us to identify areas for further development.

During the MASSMO 5 and Summer 2018 AlterECO deployments of Slocum gliders, the C2 data processing system was fully trialed, including the submission of delayed-mode data to the Ocean Glider Programme GDAC, and the submission of NRT data to the UK Met Office. During MASSMO 5, the EGO files produced by the C2 data processing system were visualized by project partners including Plymouth Marine Laboratory (PML) and SAMS. Feedback and lessons learnt during these trials were subsequently fed into requirements for future development.

9. CONCLUSIONS

As the popularity of MAS operations continues to grow, it is essential that factors which currently limit the scalability of piloting and data systems are addressed. Existing MAS piloting interfaces were typically custom-made for each vehicle by the manufacturer, assuming a one-to-one relationship between pilot and vehicle. This has resulted in a significant training overhead for pilots which, combined with limited cross-vehicle standardization, restricts the size and complexity of heterogeneous fleet operations. Likewise, data have historically been primarily “delayed-mode” i.e., available only on recovery of the MAS vehicle. However, advances in long-range MAS vehicles, capable of operating for months at a time, have created opportunities to automate the processing of NRT data, enabling assimilation into ocean models and delivery to the end-user throughout the mission.

³⁸AlterEco - The National Oceanography Centre and the Natural Environment Research Council: <http://altereco.ac.uk/> (accessed July 25, 2018).

REFERENCES

- Ai-Chang, M., Bresina, J., Charest, L., Chase, A., Hsu, J. C. J., Jonsson, A., et al. (2004). Mapgen: mixed-initiative planning and scheduling for the Mars exploration rover mission. *IEEE Intell. Syst.* 19, 8–12. doi: 10.1109/MIS.2004.1265878
- Anderlini, E., Harris, C., Phillips, A. B., Lopez, A. L., Woo, M., and Thomas, G. (2019). Towards autonomy: a recommender system for the determination of trim and flight parameters for seaglidors. *Ocean Eng.* 189:106338. doi: 10.1016/j.oceaneng.2019.106338
- Australian National Facility for Ocean Gliders (2012). *Data Management User's Manual*. Technical report, University of Western Australia.
- Brito, M., Bose, N., Lewis, R., Alexander, P., Griffiths, G., and Ferguson, J. (2012). “The role of adaptive mission planning and control in persistent autonomous underwater vehicles presence,” in *Proceedings of IEEE/OES Autonomous Underwater Vehicles, AUV'12* (Southampton: Institute of Electrical and Electronic Engineers). doi: 10.1109/AUV.2012.6380748
- Brito, M., and Griffiths, G. (2009). “On the use of expert judgment elicitation for autonomous underwater vehicle risk prediction and management,” in *Proceedings of the European Safety and Reliability Conference, ESREL'09* (Prague: Taylor and Francis), 1221–1227. doi: 10.1201/9780203859759.ch169

In this paper, we have presented the Oceanids C2, a unified web-based system for the command, control and data management of MAS vehicles within the NMEP. Through the use of data, metadata, and API standards, and prioritizing the flexibility and scalability of the C2 system, we seek to enable future collaborations with both the marine science and robotics communities to revolutionize the use of long-range MAS vehicles and ultimately increase the delivery of high-quality oceanographic data for cutting-edge science.

AUTHOR CONTRIBUTIONS

CH is leading the Automated Piloting Framework and wrote the paper. AL-L oversaw technical development of the C2 Infrastructure and wrote the paper. OJ developed the backend. JB acted as product owner for the data pathway and wrote the paper. AK was the architect of the metadata system. SL oversaw development of the data standardization workflow and wrote the paper. TG built APIs and integrated the first gateway. AP oversaw the project and wrote the paper.

FUNDING

The development of the C2 infrastructure was funded by the UK's Industrial Strategy Challenge Fund (ISCF)—Natural Environment Research Council (NERC) Oceanids Capital investment in Marine Autonomous Systems.

ACKNOWLEDGMENTS

The authors would like to thank Ashley Morris, Jack Farley, Stephen Fraser, Kay Thorne, Malcolm Hearn, Matthew Cazaly, James Yannaros, Cieran Brazier, Karen Vickers, Jack Gee, Thomas Bailey, Maaten Furlong, Russ Wynn, Carolina Dopico Gonzalez, Paola Arce, Sam Jones, Matt Toberman, and Louise Darroch for their work on the Oceanids C2 project.

- Brito, M. P., Griffiths, G., and Challenor, P. (2010). Risk analysis for autonomous underwater vehicle operations in extreme environments. *Risk Anal.* 30, 1771–1788. doi: 10.1111/j.1539-6924.2010.01476.x
- Bröring, A., Stasch, C., and Echterhoff, J. (2012). *OGC Sensor Observation Service Interface Standard*. Technical report, Open Geospatial Consortium.
- Buadu, S., Schjølberg, I., and Mo-Bjørkelund, T. (2018). “Mission planner for multiple AUVs: verification procedures combining simulations and experiments,” in *2018 IEEE/OES Autonomous Underwater Vehicle Workshop (AUV)* (Porto), 1–6. doi: 10.1109/AUV.2018.8729793
- Buck, J. J. H., Bainbridge, S. J., Burger, E. F., Rio Fernandez, J. d., Delory, E., Fischer, P., et al. (2019). Ocean data product integration through innovation—the next level of data interoperability. *Front. Mar. Sci.* 6:32. doi: 10.3389/fmars.2019.00032
- Cunningham, S. A., Kanzow, T., Rayner, D., Baringer, M. O., Johns, W. E., Marotzke, J., et al. (2007). Temporal variability of the Atlantic meridional overturning circulation at 26.5°N. *Science* 317, 935–938. doi: 10.1126/science.1141304
- Dias, P. S., Gomes, R. M., and Pinto, J. (2006). “Mission planning and specification in the Neptus framework,” in *Proceedings 2006 IEEE International Conference on Robotics and Automation* (Orlando, FL), 3220–3225. doi: 10.1109/ROBOT.2006.1642192

- Dopico-Gonzalez, C., Harris, C. A., and Phillips, A. B. (2019). "Risk and reliability management of marine autonomous systems within the UK's national marine equipment pool," in *OCEANS 2019* (Marseille), 1–9. doi: 10.1109/OCEANSE.2019.8867521
- EGO gliders data management team (2017). *Ego Gliders netcdf Format Reference Manual*. Technical report, Ifremer.
- Eriksen, C. C., Osse, T. J., Light, R. D., Wen, T., Lehman, T. W., Sabin, P. L., et al. (2001). Seaglider: a long-range autonomous underwater vehicle for oceanographic research. *IEEE J. Ocean. Eng.* 26, 424–436. doi: 10.1109/48.972073
- Faria, M., Pinto, J., Py, F., Fortuna, J., Dias, H., Martins, R., et al. (2014). "Coordinating UAVs and AUVs for oceanographic field experiments: challenges and lessons learned," in *2014 IEEE International Conference on Robotics and Automation (ICRA)* (Hong Kong), 6606–6611. doi: 10.1109/ICRA.2014.6907834
- Farley, J., Morris, A., Jones, O., Harris, C. A., and Lorenzo, A. (2019). "Marine science from an armchair: a unified piloting framework for autonomous marine vehicles," in *OCEANS 2019* (Marseille). doi: 10.1109/OCEANSE.2019.8867573
- Ferreira, A. S., Costa, M., Py, F., Pinto, J., Silva, M. A., Nimmo-Smith, A., et al. (2019). Advancing multi-vehicle deployments in oceanographic field experiments. *Auton. Robots* 43, 1555–1574. doi: 10.1007/s10514-018-9810-x
- Ferreira, A. S., Pinto, J., Dias, P., and de Sousa, J. B. (2017). "The LSTS software toolchain for persistent maritime operations applied through vehicular ad-hoc networks," in *2017 International Conference on Unmanned Aircraft Systems (ICUAS)* (Miami, FL), 609–616. doi: 10.1109/ICUAS.2017.7991471
- Fette, I., and Melnikov, A. (2011). *Rfc 6455: The WebSocket Protocol*. Technical report, IETF. doi: 10.17487/rfc6455
- Fox, M., Long, D., and Magazzeni, D. (2017). Explainable planning. *CoRR, abs/1709.10256*.
- Furlong, M. E., Paxton, D., Stevenson, P., Pebody, M., McPhail, S. D., and Perrett, J. (2012). "Autosub long range: a long range deep diving AUV for ocean monitoring," in *Autonomous Underwater Vehicles (AUV), 2012 IEEE/OES* (Southampton), 1–7. doi: 10.1109/AUV.2012.6380737
- Garabato, A. C. N., Frajka-Williams, E. E., Spingys, C. P., Legg, S., Polzin, K. L., Forryan, A., et al. (2019). Rapid mixing and exchange of deep-ocean waters in an abyssal boundary current. *Proc. Natl. Acad. Sci. U.S.A.* 116, 13233–13238. doi: 10.1073/pnas.1904087116
- German, C. R., Jakuba, M. V., Kinsey, J. C., Partan, J., Suman, S., Belani, A., et al. (2012). "A long term vision for long-range ship-free deep ocean operations: persistent presence through coordination of autonomous surface vehicles and autonomous underwater vehicles," in *Autonomous Underwater Vehicles (AUV), 2012 IEEE/OES* (Southampton), 1–7. doi: 10.1109/AUV.2012.6380753
- German, C. R., Yoerger, D. R., Jakuba, M., Shank, T. M., Langmuir, C. H., and Nakamura, K.-I. (2008). Hydrothermal exploration with the autonomous benthic explorer. *Deep Sea Res. Part I* 55, 203–219. doi: 10.1016/j.dsr.2007.11.004
- Haller, A., Janowicz, K., Cox, S., Phuoc, D. L., Taylor, K., and Lefrançois, M. (2017). *Semantic Sensor Network Ontology. w3c Recommendation*. Technical report, W3C.
- Haworth, C., Evans, J., McManamon, K., and McNally, K. (2016). "Combined USV/subsea-glider fleets for tidal mixing front tracking and monitoring," in *OCEANS 2016 MTS/IEEE* (Monterey, CA), 1–6. doi: 10.1109/OCEANS.2016.7761332
- Hickson, I. (2015). *Rec-Eventsource-2015020: Server-Sent Events*. Technical report, W3C.
- Hobson, B. W., Bellingham, J. G., Kieft, B., McEwen, R., Godin, M., and Zhang, Y. (2012). "Tethys-class long range AUVs-extending the endurance of propeller-driven cruising AUVs from days to weeks," in *Autonomous Underwater Vehicles (AUV), 2012 IEEE/OES*, 1–8. Southampton: IEEE. doi: 10.1109/AUV.2012.6380735
- Houptert, L., Inall, M., Dumont, E., Gary, S., Johnson, C., Porter, M., et al. (2018). Structure and transport of the north Atlantic current in the eastern subpolar gyre from sustained glider observations. *J. Geophys. Res.* 123, 6019–6038. doi: 10.1029/2018JC014162
- Jones, D. O., Gates, A. R., Huvenne, V. A., Phillips, A. B., and Bett, B. J. (2019). Autonomous marine environmental monitoring: application in decommissioned oil fields. *Sci. Tot. Environ.* 668, 835–853. doi: 10.1016/j.scitotenv.2019.02.310
- Kokkinaki, A., Darroch, L., Buck, J., and Jirka, S. (2016). "Semantically enhancing SensorML with controlled vocabularies in the marine domain," in *Proceedings of GSWC 2016, Geospatial Sensor Webs Conference* (Munster).
- Leonard, N. E., Paley, D. A., Davis, R. E., Fratantoni, D. M., Lekien, F., and Zhang, F. (2010). Coordinated control of an underwater glider fleet in an adaptive ocean sampling field experiment in Monterey Bay. *J. Field Robot.* 27, 718–740. doi: 10.1002/rob.20366
- Liblik, T., Karstensen, J., Testor, P., Alenius, P., Hayes, D., Ruiz, S., et al. (2016). Potential for an underwater glider component as part of the global ocean observing system. *Methods Oceanogr.* 17, 50–82. doi: 10.1016/j.mio.2016.05.001
- Manley, J., and Willcox, S. (2010). "The wave glider: a persistent platform for ocean science," in *OCEANS 2010 IEEE* (Sydney), 1–5. doi: 10.1109/OCEANSSYD.2010.5603614
- Martinez, S., Delory, E., del Rio, J., Jirka, S., Pearlman, J., D'Angelo, P., et al. (2017). *Deliverable d7.8 Policy Document - Sensor Development for the Ocean of Tomorrow*. Technical report, European Union Seventh Framework Programme.
- McPhail, S. (2009). Autosub6000: a deep diving long range AUV. *J. Bion. Eng.* 6, 55–62. doi: 10.1016/S1672-6529(08)60095-5
- McPhail, S., Templeton, R., Pebody, M., Roper, D., and Morrison, R. (2019). "Autosub long range AUV missions under the Filchner and Ronne ice shelves in the Weddell sea, Antarctica-an engineering perspective," in *OCEANS 2019* (Marseille), 1–8. doi: 10.1109/OCEANSE.2019.8867206
- Open Geospatial Consortium (2016). *Sensor Model Language (sensorml)*. Technical report. Available online at: opengeospatial.org.
- Paley, D. A., Zhang, F., and Leonard, N. E. (2008). Cooperative control for ocean sampling: the glider coordinated control system. *IEEE Trans. Control Syst. Technol.* 16, 735–744. doi: 10.1109/TCST.2007.912238
- Phillips, A. B., Salavasidis, G., Kingsland, M., Harris, C., Pebody, M., Roper, D., et al. (2018). "Autonomous surface/subsurface survey system field trials," in *Autonomous Underwater Vehicles (AUV), 2018 IEEE/OES* (Porto), 1–7. doi: 10.1109/AUV.2018.8729740
- Pinto, J., Dias, P., Ribeiro, M., Costa, M., and Sousa, J. (2017). "Networked vehicle systems: from vision to reality," in *OCEANS 2017* (Aberdeen), 1–6. doi: 10.1109/OCEANSE.2017.8084823
- Pinto, J., Dias, P. S., Martins, R., Fortuna, J., Marques, E., and Sousa, J. (2013). "The LSTS toolchain for networked vehicle systems," in *2013 MTS/IEEE OCEANS* (Bergen), 1–9. doi: 10.1109/OCEANS-Bergen.2013.6608148
- Ramp, S. R., Davis, R. E., Leonard, N. E., Shulman, I., Chao, Y., Robinson, A., et al. (2009). Preparing to predict: the second autonomous ocean sampling network (AOSN-II) experiment in the Monterey Bay. *Deep Sea Res. Part II* 56, 68–86. doi: 10.1016/j.dsr2.2008.08.013
- Reitz, A., Visbeck, M., and Fritz, J.-S. (2016). *AtlantOS Exploitation Plan*. Technical Report, European Union Horizon 2020.
- Roemmich, D., Boehme, L., Claustre, H., Freeland, H., Fukasawa, M., Goni, G., et al. (2010). Integrating the ocean observing system: mobile platforms. *Proc. OceanObs* 9:33. doi: 10.5270/OceanObs09.pp.33
- Roemmich, D., Johnson, G. C., Riser, S., Davis, R., Gilson, J., Owens, W. B., et al. (2009). The Argo program: observing the global ocean with profiling floats. *Oceanography* 22, 34–43. doi: 10.5670/oceanog.2009.36
- Roemmich, D., and the Argo Steering Team (2009). Argo: the challenge of continuing 10 years of progress. *Oceanography* 22, 46–55. doi: 10.5670/oceanog.2009.65
- Roper, D. T., Phillips, A. B., Harris, C. A., Salavasidis, G., Pebody, M., Templeton, R., et al. (2017). "Autosub long range 1500: an ultra-endurance AUV with 6000 km range," in *OCEANS 2017* (Aberdeen: IEEE), 1–5. doi: 10.1109/OCEANSE.2017.8084928
- Rudnick, D. L., Davis, R. E., Eriksen, C. C., Fratantoni, D. M., and Perry, M. J. (2004). Underwater gliders for ocean research. *Mar. Technol. Soc. J.* 38, 73–84. doi: 10.4031/002533204787522703
- Salavasidis, G., Munafò, A., Harris, C. A., Prampart, T., Templeton, R., Smart, M., et al. (2018). Terrain-aided navigation for long-endurance and deep-rated autonomous underwater vehicles. *J. Field Robot.* 36, 447–474. doi: 10.1002/rob.21832
- Schofield, O., Kohut, J., Aragon, D., Creed, L., Graver, J., Haldeman, C., et al. (2007). Slocum gliders: robust and ready. *J. Field Robot.* 24, 473–485. doi: 10.1002/rob.20200

- Stokey, R., Austin, T., von Alt, C., Purcell, M., Forrester, N., Goldsborough, R., et al. (1999). "AUV bloopers or why Murphy must have been an optimist: a practical look at achieving mission level reliability in an autonomous underwater vehicle," in *Proceedings of the Eleventh International Symposium on Unmanned Untethered Submersible Technology* (Durham, NC), 32–40.
- Testor, P., de Young, B., Rudnick, D. L., Glenn, S., Hayes, D., Lee, C. M., et al. (2019). Ocean gliders: a component of the integrated GOOS. *Front. Mar. Sci.* 6:422. doi: 10.3389/fmars.2019.00422
- Testor, P., Meyers, G., Pattiaratchi, C., Bachmayer, R., Hayes, D., Pouliquen, S., et al. (2010). "Gliders as a component of future observing systems," in *OceanObs'09* (Venice). doi: 10.5270/OceanObs09.cwp.89
- Thieme, C. A., and Utne, I. B. (2017). Safety performance monitoring of autonomous marine systems. *Reliabil. Eng. Syst. Saf.* 159, 264–275. doi: 10.1016/j.ress.2016.11.024
- U.S. IOOS National Glider Data Assembly Center (2018). *Ngdac Netcdf File Format Version 2*. Technical report, Integrated Ocean Observing System.
- Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., et al. (2016). The FAIR guiding principles for scientific data management and stewardship. *Sci. Data* 3:160018. doi: 10.1038/sdata.2016.18
- Wortham, R. H., Theodorou, A., and Bryson, J. J. (2017). "Improving robot transparency: real-time visualisation of robot AI substantially improves understanding in naive observers," in *IEEE RO-MAN 2017* (Lisbon). doi: 10.1109/ROMAN.2017.8172491

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Harris, Lorenzo-Lopez, Jones, Buck, Kokkinaki, Loch, Gardner and Phillips. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.