National Oceanography Centre
NATURAL ENVIRONMENT RESEARCH COUNCIL

National Oceanography Centre

Research & Consultancy Report No. 45

The use of gliders for oceanographic science: the data processing gap

S C Painter & A P Martin

2014

National Oceanography Centre, Southampton
University of Southampton Waterfront Campus
European Way
Southampton
Hants SO14 3ZH UK

Author contact details:
Tel: +44 (0)23 8059 6342
Email: adrian.martin@noc.ac.uk

DOCUMENT DATA SHEET

AUTHOR PAINTER, S C & MARTIN, A P	*PUBLICATION DATE* 2014
TITLE The use of gliders for oceanographic science: the data processing gap.	
REFERENCE Southampton, UK: National Oceanography Centre, 34pp. (National Oceanography Centre Research and Consultancy Report, No. 45)	

ABSTRACT

Autonomous gliders represent a step change in the way oceanographic data can be collected and as such they are increasingly seen as valuable tools in the oceanographer's arsenal. However, their increase in use has left a gap regarding the conversion of the signals that their sensors collect into scientifically useable data.

At present the novelty of gliders means that only a few research groups within the UK are capable of processing glider data whilst the wider oceanographic community is often unaware that requesting deployment of a glider by MARS does not mean that they will be provided with fully processed and calibrated data following the deployment. This is not a failing of MARS – it is not in their remit – but it does mean that a solution is needed at the UK community level. The solution is also needed quickly given the rapidly growing glider fleet and requests to use it.

To illustrate the far from trivial resources and issues needed to solve this problem at a community level, this document briefly summarises the resources and steps involved in carrying glider data through from collection to final product, for the glider owning research groups within the UK which have the capability.

This report does not provide a recommendation on whether such a community facility should be the responsibility of NOC, BODC or MARS but does provide information on possible protocols and available software that could be part of a solution.

This report does, however, recommend that, to support the growing use of the MARS gliders, a permanently staffed group is needed as a priority, to provide data processing and calibration necessary to allow the translation of glider missions into high impact scientific publications.

KEYWORDS:

ISSUING ORGANISATION **National Oceanography Centre**
University of Southampton Waterfront Campus
European Way
Southampton SO14 3ZH
UK

PDF available at http://nora.nerc.ac.uk/

Page intentionally left blank

Contents

Introduction

The use of gliders to collect oceanographic data is increasingly popular due to the perceived low cost of data collection and the longevity of a typical glider deployment. The establishment of MARS and the subsequent funding to expand the fleet of gliders available to the UK marine community will rapidly accelerate this, both by raising the profile of gliders and by providing resources to allow wider access to the UK glider fleet.

It is evident however that there is a skills gap in the chain leading from MARS to scientific result. MARS has a clear, and defensible, view that its remit is to physically deploy, pilot and recover gliders and to ensure the raw data collected are passed to the relevant scientists. However, many scientists requesting gliders for projects are unaware that data cannot be used straight from the glider: it has to be quality controlled and calibrated. Like all remotely sensed data, there are spikes and glitches that need to be removed and experience of extant glider researchers in the UK indicates that factory calibrations seldom perform well against independent field data. This is perhaps unsurprising given the considerable effort (and cost) expended on research cruises to calibrate salinity/conductivity and oxygen sensors even on traditional CTD rosette packages.

As a result of both the skills gap and the lack of awareness amongst some scientists of the need to calibrate sensors, a number of projects do not request sufficient resources to process and analyse glider data. This situation has arisen not just because there appears to be little appreciation of the considerable work necessary to carry out the important task of calibration but because there may be little appreciation that it is even needed.

Processing covered by MARS

Taking Seagliders as an example, the basic process of working with MARS gliders during their deployment is carried out by MARS. This involves downloading dive files from the glider to the basestation via the Iridium satellite system at the end of every dive (an automatic process), and then passing these dive files through a series of manufacturer supplied Matlab scripts (a manual process) for the purposes of piloting the gliders. The dive files contain data in engineering units only (counts or voltages) and the primary purpose of the manufacturer supplied Matlab scripts is to inform the pilot of the health and orientation of the Seaglider. The secondary purpose of these scripts (following modification) is to allow preliminary investigation of the data, which can be undertaken following application of the manufacturer provided instrument calibrations to the raw engineering data. This can produce a dataset with scientific units that is useful for quick interpretation but not for scientific analysis and publication.

The need for calibration

Rigorous calibration against in-situ data, as is required standard practice for other oceanographic data sources, remains a major problem for AUV's. As AUV's operate remotely, AUV's usually suffer from a lack of in-situ data against which to calibrate sensors (for discussion of the problems applicable to SeaGliders see Perry *et al*., 2008). A common procedure currently used is to calibrate the instruments against a CTD cast at the start (deployment) and end (recovery) of each mission to provide a 2-point calibration (implicitly making significant assumptions over instrument stability/biofouling in between). The problem of calibrating instruments on AUV's is non-trivial and has previously prevented publication of research (*e.g.* the study by Sackmann *et al*., (2008) submitted to Biogeosciences Discussion was blocked from further revision by Reviewers who strongly disagreed over attempts to sidestep the calibration process). Publications from the most comprehensive biogeochemical glider study to date (North Atlantic Bloom Experiment 2008; NAB08) give prominence to procedures for sensor calibration. Considerable time is needed to calibrate data from gliders following every deployment, even by experienced glider users, and the novice glider user is therefore the most disadvantaged in this regards.

UK interests in glider deployments for long-term statutory monitoring purposes (e.g. with DEFRA, CEFAS, SEPA etc) may in some cases be undertaken with lower quality data requirements, though every effort should be made to acquire the best quality data possible.

The gap: post-deployment, pre-science data processing

It is hoped that glider data processing will harmonise around community agreed "best-practice" procedures (e.g. GROOM Deliverable 5.3[1]) in the same way that Argo float, ADCP and CTD data procedures have largely been harmonised for hydrographic data. At present, however, protocols and software are being developed independently with obvious duplication of effort.

Although experienced individuals are sparsely scattered across the UK (e.g. Mark Inall at SAMS, Karen Heywood and Jan Kaiser at UEA, Matthew Palmer and David Smeed at NOC), a common theme within all current users of AUV's is the development of small teams of individuals dedicated to using and exploiting glider data. There is no precedent for an individual researcher to deploy, calibrate and exploit glider data without significant support. Despite several high profile research programmes utilising AUV's (e.g. Pine Island Glacier, OSMOSIS) the bulk of data processing has to date been undertaken by established glider groups (i.e. SAMS, UEA) and the expertise has not been widely disseminated.

BODC are engaged in international efforts to harmonise the quality assurance procedures of raw glider data within the data management community. However, they are not engaged in facilitating glider data processing / calibration and are instead, like MARS, leaving this to individual PI's to undertake. The advantage of BODC's effort, however, is that a unified data format, regardless of

[1] Groom Deliverable 5.3 *Protocols for sampling, sample analysis, inter-calibration of missions, and data analysis for the recommended parameters*

glider type, will be produced. From this starting point, routines to calibrate the data should hopefully become more standardised and therefore easier to use.

Current approaches to data processing within the UK

To provide a quick, rough estimate of the resources and issues associated with linking glider data collection to scientific use, a questionnaire was sent to the main glider groups in the UK. Details can be found in Appendix B but summaries are given here…

University of East Anglia (UEA)

Karen Heywood led UEA as early adopters of gliders within the UK and they have developed a good track record of glider use, particularly within the Southern Ocean, for physical oceanographic research. A small, dedicated research group now exists consisting of Principle Investigators, post-docs, PhD students and technicians many of whom primarily focus on glider-based science. This group has a growing international reputation for glider use and has developed a series of in-house procedures for dealing with glider data. However, despite regular glider deployments the process of handling data remains non-trivial, often taking several months or longer for each glider deployment. As this group has a more physical perspective their efforts have focussed on attaining the best salinity calibrations and also on the best estimates of current velocities and transports. Biogeochemical work with gliders is increasing with Jan Kaiser in particular active in this direction. The group at UEA are currently in the process of preparing a Matlab based toolbox that may be of wider interest and have previously provided data processing scripts to SAMS.

Scottish Association for Marine Science (SAMS)

SAMS have independently developed a glider capability that shares many similarities with that developed by UEA. A small team of researchers have, over a number of years, established a series of procedures for handling glider data and have borrowed and modified procedures developed at UEA. They have a dedicated glider pilot / data processor who works alongside the PI's to undertake both jobs of piloting and data processing. The main focus of this group has also been on physical oceanography with more emphasis on salinity calibrations and application of gliders to hydrographic questions than to biogeochemical questions, though as with UEA this is changing.

National Oceanography Centre (NOC)

Two researchers at NOC (Mathew Palmer and David Smeed) have developed extensive capabilities for using Slocum glider data, but in both cases this has been through the judicious appointment of engineers/interns who have written extensive software routines to exploit the data.

British Antarctic Survey (BAS)

BAS have a developing glider capability (http://swallow.nerc-bas.ac.uk/slocum/) in support of their research activities at Rothera. Their approach to data processing is based on self-written scripts and calibration against the Rothera CTD timeseries.

A software option outside the UK - SOCIB

The international research community has yet to settle upon basic data processing procedures (but GROOM 5.3. Deliverable is imminent). Nevertheless groups have been developing software. As an example of this, the Balearic Islands Coastal Observing and Forecasting System (www.socib.es) based in Mallorca has spent considerable time developing protocols for processing glider data for operational purposes. This was originally designed for Slocum gliders but has now also been done for Seagliders. Within 1 day of receipt, level 1 data are available from the publically accessible web-page, having had QC and basic corrections (e.g. temperature lag) applied. This first stage is essentially automated. For level 2 data a final salinity calibration is applied, either by comparison to simultaneous CTD etc data or else from historical/climatological data. The main time constraint here is the wait for the necessary simultaneous data to be available. Once again the software has already been written to carry out the necessary processing. In summary, SOCIB have a suite of software, already publically available (www.github.com/socib/glider_toolbox), written in Matlab (but being made compatible with Octave) which follows clear protocols to take glider data from receipt from glider through to fully processed and publically available.

Summary of what is required

The successful model used by all glider owning research groups is for small groups of researchers, numbering between 4 and 20, to be heavily involved in end-to-end aspects of glider missions on a full time basis. MARS covers the deployment through to recovery but, particularly giving the rapidly increasing MARS fleet, the questionnaires reveal that a permanent team of several people is required to provide data processing and calibration to the growing UK glider user community. This may seem costly, but the cost of individual scientists repeating and reinventing the same steps in isolation will be of significant greater cost to NERC.

Such efforts have successfully been introduced into international programmes such as ARGO (and handled via BODC), whilst many international field programmes seek a basic level of accuracy and comparability in their measurements (e.g. WOCE, Geotraces) regardless of the precise methodology employed.

A common data processing system would (if sufficiently widely supported) provide a strong platform upon which the UK can develop a leading capability in glider usage. However, the diversity of data processing procedures for even long-established common oceanographic instrumentation such as CTD’s or ADCP’s indicates two things: there will always be a need for bespoke solutions for particular situations and sensors; there will be no community solution unless a high level national lead is taken.

Appendix A

Instrumentation

The two varieties of glider owned and operated by MARS are the Slocum and the Seaglider. The default configuration of both gliders is the same and typically consists of sensors to measure...

1. Conductivity

The standard conductivity cell on a glider is unpumped and thus prone to significant and sometimes rather serious temporal lags, which offset the simultaneous measurements of conductivity and temperature. If left uncorrected such offsets impact salinity and density calculations.

2. Temperature

The temperature sensor on gliders is prone to a sampling delay, known as the thermal lag, which ultimately decouples the measurements of conductivity and temperature. This requires correction and suggestions are that delays approaching 100 seconds may be common, though any such delay is likely to be variable.

3. Dissolved Oxygen

Standard procedures are to i) Apply the manufacturers calibration and then ii) undertake a secondary calibration to in-situ data. Consideration of sensor drift or lack of stability are largely ignored due to the lack of in-situ calibration data to confirm the extent of the problem.

4a. Wetlabs Ecopuck – Chlorophyll fluorescence

Chlorophyll fluorescence is widely measured as a means of assessing algal biomass but is also widely recognised for its limitations. Photochemical and non-photochemical quenching are both important factors impacting near-surface fluorescence and ultimately estimates of chlorophyll concentration. There is no widely accepted correction for quenching.

Standard procedures are to i) Apply the manufacturers calibration, which is likely to overestimate chlorophyll concentrations and then ii) undertake a secondary calibration to in-situ data. Developing techniques to calibrate chlorophyll fluorescence in the absence of in-situ data are being developed at NOC, but require appropriate peer-review before they can be considered viable.

4b. Wetlabs Ecopuck – Optical backscatter

The optical backscatter sensor provides information of water column turbidity (particle loading) and methods to use this data stream to estimate particulate organic carbon distributions exist.

4c. Wetlabs Ecopuck – CDOM fluorescence

Although it is considered possible to monitor CDOM (chromophoric dissolved organic matter, yellow substances or 'gelbstoff') in seawater, results from CDOM sensors are poorly understood. Firstly, CDOM is a complex pool of organic compounds the exact composition of which is not known. Secondly, whilst a few CDOM compounds have been isolated and identified the vast majority are unknown and consequently there is no artificial standard that can be used to calibrate CDOM sensors. Originally CDOM sensors were developed to detect hydrocarbon sources or leaks, and have only lately been marketed as a means of tracking CDOM concentrations. Thirdly, the current best practice for CDOM sensor calibration is to calibrate against a series of quinine sulphate standards which can be made to precise concentrations, and which fluoresce in a similar way to CDOM, but the result is that the investigator is reduced to reporting quinine sulphate or QS units – which is a qualitative rather quantitative indicator of CDOM concentration. For these reasons results from CDOM sensors are still largely viewed as qualitative (and questionable by some parts of the community) indicators of dissolved organic matter pools. However, such data do bear some resemblance to expected patterns and distributions.

Other instrumentation

There is a growing appetite for additional sensors to be fitted to AUV's. Such examples include the ISUS nitrate sensor, Acoustic Current Doppler Profilers, turbulence sensors and PAR sensors. All come with their own problems.

Appendix B - Questionnaire

The following set of questions were sent to glider users at SAMS, BAS, UEA, NOC(L), NOC(S)

PEOPLE

- Do you have a dedicated glider pilot or is the piloting shared amongst several people?
- Do you employ staff dedicated to assisting glider missions? (i.e. it is their primary role) or are people co-opted on an ad-hoc basis?
- Do you utilise short-term contract staff/students to develop your capabilities? If so, what do they do?
- For a hypothetical 4-month glider mission how many people would be involved from the initial deployment right through to the production of a final calibrated dataset?
- How many years experience do you and/or your group now have of glider operations?
- Does that experience make dealing with each new glider dataset easier or do you still encounter new problems?

DATA PROCESSING (EXCLUDING PILOTING)

- Briefly describe what steps you go through to turn raw glider data (i.e. that recovered from the basestation) into a format useful for scientific applications.
- Do you use your own software to do this? If not, whose do you use?
- How long has it taken to get the software to the state it is in today?
- Do you process any data streams to a final form as they are returned on a dive-by-dive basis or do you wait until the glider mission has finished before starting to process all data streams?
- For the same hypothetical 4-month glider deployment, how long would it take you to produce the final dataset?
- Are you limited by staff numbers, software, or time (complexity of job)?
- Would this be for hydrographic data only (T,S,O_2), biogeochemical data only (O_2, Chl-a, CDOM, backscatter) or both?
- Thinking back to your first glider mission. How long did it take you to produce the final dataset?
- Do you consider your data processing procedures to be easily transferable to new glider datasets? Or are you faced with frequent rewriting of scripts?
- As many potential users of the MARS glider fleet have no previous experience of gliders what do you see as the biggest obstacle(s) to a successful outcome?

CALIBRATION

- Would you consider using data obtained from satellites, climatologies, or models to calibrate glider data?

SCIENTIFIC USE

- Would you trust and use partially processed glider data in your work? (e.g. despiked and smoothed data, but with minimal or no calibration)
- Would you agree with the publication of partially processed glider data for scientific purposes?
- What do you see as the biggest obstacle to wider acceptance of glider-based observations?

	BAS	SAMS	UEA	NOC(S)
PEOPLE				
Do you use a dedicated pilot or is piloting shared?	A single individual is usually responsible but frequent comms problems from Rothera require outside involvement	Piloting is shared between I technician and a small team of scientists	Piloting shared amongst 10 individuals (staff/postdocs/ and students)	Piloting was originally undertaken by 1 individual and/or postdocs. More recently via MARS glider team but with occasional contribution
Do you employ dedicated staff for glider activities?	No, gliders are considered part of a wider job role	1 full-time technician with responsibility for gliders/AUV's (hoping to recruit a second)	Two technicians	No staff employed outside MARS
Do you utilise short-term contract staff/students to develop your capabilities?	No	Yes, external IT contractor for database and website development/maintenance, and data distribution (but not glider data processing) 3 summer students have been used to develop real-	Yes, PhD students and postdocs to pilot gliders, process data, write papers.	MARS has used external IT contractors to develop web interface and piloting tools (but not data processing procedures – this is argued to be the responsibility of science

		time and delayed time data processing routines		users)
For a hypothetical 4-month mission, how many people would be involved from start to finish	In total 5+ base staff support for every mission Testing: 2-4 people Planning: 1-2 people Deployment: 2-4 people Piloting: up to 4 people Recovery: 4 people Data processing: 1 person	Minimum of 3 people at any one time Lab testing prior to deployment: 1 person Water testing prior to deployment: 3 people (2 in field + 1 pilot at base) Deployment: 3 people (2 in field + 1 pilot at base) Piloting: 2 or 3 pilots Recovery: 3 people (2 in field + 1 pilot at base) Post-processing: Minimum of 1-2 people.	Excluding piloting 3-4 people would be needed. Including piloting duties could see up to 10 people involved.	No answer provided
How many years experience do you have?	2 field seasons (+1 years testing)	6 years experience	As a group – 5 years, but individuals experiences range from <2 years to 5 years.	Started in 2007, but not deployed every year. Two most experienced postdocs both left NOC

Does that experience make handling datasets easier? or are you faced with new problems?	Yes, but still encounter data/hardware issues that need fixing	New problems encountered every time, due to lack of standard data processing methodology that is widely accepted and widely used.	Experience does make the job easier, but new problems are always encountered.	Experience is useful but there are always issues as the technology changes.
DATA PROCESSING				
Briefly describe your data processing steps	During deployment: Acquire files from glider, merge files, calculate salinity, density, potential temp, etc, interpolate data, create basic data plots. Sometimes create 1db profiles for up and down dive. Plot data. After recovery: Investigate thermal lag, offset between up and down casts, compare to Rothera CTD timeseries	Raw data (two file formats): *Ascii files* (Oxygen and Wetlab data streams) - Convert engineering units to scientific units using manufacturer instrument calibrations. Adjust oxygen data (Aanderaa Optode) for temperature effects. *Pro files* (CT data) - Convert engineering units to scientific units using manufacturer instrument	During deployment: Acquire files from glider, merge files, calculate salinity, density, potential temp, etc, interpolate data, create basic data plots. This is mostly automated. After recovery: load and merge data into our matlab glider toolbox, and modify toolbox code to accept new sensor	Create NetCDF files from returned data, Apply thermal lag correction for calculation of salinity. Calibrate salinity against independent data (CTD cast), Inspect data and flag periods of fouling. No experience of calibrating/using data from other (Wetlabs/Aanderaa) sensors

data (casts within 1 hour of deployment/recovery) and correct glider data for any problems. Cross check with whatever data available.	calibrations. Remove outliers outside sensor range (does not despike small magnitude outliers). Apply first order lag correction to CT sensor (rough correction only). Calculate underwater lat/lon positions for data. Calculate dive average current and surface drift current. Real time data (Matlab mat file): Group all variables in a single file per dive. Correct oxygen data for salinity and pressure effects Delayed time data (Matlab mat file): Despike all	names (if needed). We believe we're the only ones to adjust for the time offset between sensors that occurs because of the single thread processing on the seagliders (sometimes up to 5 sec offset, so a couple metres) which leads to some very odd spiking in downstream property calculations. Toolbox contains scripts to calculate derived variables (salinity, density, dive-average currents, vertical velocity of water etc) . Also to find corrected pressure and time vectors to account for non-

variables. Calculate and correct sensor drift via cross-comparison to CTD data (or from pre- and post-deployment manufacturers calibration). Realign time stamping on all sensors (Seaglider CPU is single-thread so samples each sensor one after the other, realign all sensors to correct pressure). Correct CT thermal lag to correct salinity (complex and time consuming as glider CT sensor is unpumped). Check compass for drift (important for dive averaged currents)	simultaneity of sensors. Run these. Tune glider flight model. Find all dives with bad temp/salinity data (due to biofouling or sensor failure) – these must be excluded in next step. Correct thermal lag of conductivity cell. Details of method will depend on location/time of year – strong/weak stratification/winter water layers/etc – all can require a slightly different approach. And it's not that we have code for all situations already in existence, so new code development

may be required. This can be quite time-consuming.

Despike and quality control. Some can be automated, but salinity issues near-surface and at mixed layer depth will likely have to be examined dive by dive. This is the most time-consuming step, but it will not be necessary for all applications.

Calibrate salinity against ship CTDs.

If salinity calibration correction is large, re-tune glider flight model (it depends on density).

Hand over to biogeochemists for all their data processing – for chlorophyll, this will involve de-spiking and conversion from engineering to physical units/calibration. (The latter two both involve finding, and applying the 'dark counts' and scale factor. Manufacturer-given dark counts and scale factor tend to be a bit rubbish so these will need to be determined. We have our own Chl a calibration routines with improved dark count determination and regression routines.) For oxygen, de-spiking, tau correction, calibration,

			possibly need to correct for hysteresis. We've implemented Johannes Hahn's methods for O2 calibration and temperature dependent lag correction. Depending on application, some kind of optimal interpolation may be required for gridding purposes. This will again be quite application specific.	
Do you use your own software?	Yes. Custom written software is used but not known if standardised procedures are used	Yes, custom written software in Matlab for all processing steps except sensors time alignment and thermal lag correction (For this we use modified toolbox from UEA, itself based on modified version	Custom written software (Matlab) is used. We've been doing quite a bit of work with other institutes - not so much in the UK, but plenty in the US. We've piloted	Yes. Custom written software is used.

		of SLOCUM glider toolbox). UEA toolbox used because UEA developed it first, and logic behind processing widely agreed within Europe glider users. SAMS have modified some elements of toolbox (but disagree internally over some of those changes)	gliders for, and have calibrated data for, CalTech, Virginia Institute of Marine Science and Old Dominion University. Lately, we've been training to glider pilots from VIMS to work with our toolbox and have got them involved in the development.	
How long to develop your software?	-	Work in progress. Started development following first science mission 4 years ago. Constant updating of software.	Work started when gliders first bought and software constantly updated/modified as new problems emerge and as experience and application grows.	Hard to say, as my software is continually changed/updated.
Real-time data processing or delayed	Data processed to final form after mission	Both	Both. Final calibration requires full mission	Both, but generally work on 1 file containing all

mode processing only?	complete, but raw data (or partially) processed is used for mission decisions.		dataset but initial processing of individual dives is often useful for examining data. The toolchain is pretty much automated and we occasionally run it in near-realtime. Less so on multiple glider deployments (e.g. OSMOSIS) because of the need to intercalibrate and delays getting samples analysed - hence the longer turnaround time - but our single glider missions output the data fairly rapidly. This is the Level 1 output. As soon as calibration constants are added to the config script, the level 2 data is	data.

			also output; so technically this can be provided after input of calibration data from a launch CTD.	
How long does it take to produce the final dataset?	Depends on other commitments (weeks-months). Learning curve very steep, and much still to learn from sharing experiences between other groups highly advisable	No answer provided	Depends on application and level of quality control needed on data. Could very easily take as long as the mission or longer. And that would be for one glider only. If multiple gliders deployed each would need the same amount of time.	18 months of data processing after a 3-month mission with one glider.
Are you limited by staff, software or time?	Happy with existing procedures, but much could be learnt from community good practise.	No answer provided	All suggested factors limit the time taken to produce calibrated datasets.	No answer provided

Do you process hydrographic or biogeochemical data?	Both	No answer provided (but hydrography (CT) data is known priority for this lab)	Both, but individual users may take responsibility for individual data channels.	Mostly CTD (hydrographic data). No experience of biogeochemical data
Thinking back your first real mission, how long did it take to generate final dataset?	Unfortunately, not sure as other simultaneous commitments extended time needed.	No answer provided	Currently 18 months since end of last mission, and final datasets still not ready due to quality control requirements.	No answer provided
Are your procedures transferable to new glider datasets or do you need to rewrite scripts?	Generally transferable and procedures also work with data from US gliders.	No answer provided (but clear from above answers that data processing scripts are constantly updated)	Some is transferable, but our code is still under development so we are updating code constantly. A lot of devleopment has been collaborative work with the guys at SOCIB (we now use a common CT lag correction - see the Garau paper).	Mostly transferable

Biggest obstacles for first-time glider users?	Unrealistic plans for deployment/recovery. Poor piloting. Lack of real-time data quality checking (mostly guesswork)	No answer provided	If MARS techs not involved then the issue of deployment/recovery and piloting. If MARS techs are involved then biggest problem is understanding how gliders operate, what they can and cannot do and the data processing. (N.B. Very bad idea to run projects using gliders where no scientist has previous experience)	Deciding how to use the data
CALIBRATION				
Would you consider using satellite, climatology or model output for calibration purposes?	We use Rothera CTD timeseries data, but in extremis would investigate alternatives but this would not be ideal.	No answer provided (but from answer above calibration against CTD data is clearly preferred option)	Our preferred approach is to use CTD data and bottle samples to calibrate gliders. Satellite data is predominately surface	Preference always to calibrate against CTD data. Argo data may be useful. Nothing to gain from models or climatologies for salinity

only and glider data in surface waters often discarded due to spiking so no calibration option. calibration

Models and climatologies are more likely to present averaged conditions so calibrating gliders against these may introduce bias into the data.

We've used models and climatology to calibrate gliders (namely in the Ross Sea, Indian Ocean and Atlantic for GOVARS, Tropical DISGO and GOPINA projects respectively) with relative success - it's very dependent on the local hydrography

			obviously. But this is very mission dependent - OSMOSIS hasn't really relied on these for example.	
SCIENTIFIC USE				
Would you trust and use partially processed but minimally calibrated data?	Depends hugely on application. If relative values or large and reproducible signal is required then possibly. If small-scale structure or important gradients are needed then probably not. Potential for reduced accuracy needs to be stated	No answer provided		For some uses it is acceptable to use data that does not have an absolute calibration. In the case of multiglider deployments inter-calibration between gliders required.
Would you agree with publication of partially processed data?	It should not be the norm that uncalibrated or partially calibrated data be used scientifically but it can have a qualitative	No answer provided	Depends hugely on purpose. Relative comparisons can be made with partially calibrated data, but	Yes

	use (see above). Planning should incorporate the requirement for calibration.		quantified comparisons cannot. I would expect data to be processed sufficiently for the science that is in the same publication	
What do you see as the biggest obstacle for wider acceptance of glider-based observations?	Not sure. Community support will grow as the recognised body of good science grows. Gliders should be seen as part of the normal data collection options (with their own strengths/weaknesses).	Glider data processing is not straight-forward, and users should be made aware of known issues. SAMS are primarily interested in CT data but provided the following information on other sensors Oxygen: We now use Aanderaa optodes, as we found the unpumped Seabird SBE-43 sensor was useless (we are still unsure whether the data	The learning curve of how to deal with gliders and the data they give you. (Also, gliders may not be suitable for some applications particularly if you need sensors which don't exist yet for gliders, or if you need to go deeper than 1000 m.	No answer provided

collected are correctable). Raw Seaglider O_2 data values are only corrected for temperature effects, but they must be corrected for pressure and salinity effects in post-processing.

Chlorophyll: The Wetlabs sensor measures chlorophyll–a fluorescence. As for CTD fluorescence data the chl-a concentration is calculated from the manufacturers calibration constants, which are established using a mono-culture of algae (*Thalassiosira weissflogii*) in the lab which does not match the multi-species composition encountered by the glider.

During cruises discrete sampling for chl-a from CTD casts mitigates this problem, but as this is not an option with gliders the real chl-a values are hard to establish.

Biofouling: this can affect all sensors, but the optical ones are usually worst affected. It is fairly obvious in the data when the Wetlabs is covered by biofouling and unable to see anything, but some questions remain for the data before that point: how do you estimate and correct for the gradual build-up of biofouling? Is it correctable?

[a] See additional information provided

[a] There has been a lot of work going on within the European glider community (namely in the EGO and GROOM projects), with one of the aims being to establish best practices for glider data post-processing (**Deliverable D5.3**, a report on protocols for sampling, sample analysis, inter-calibration of glider missions and data analysis is currently under review). Ultimately, the plan is for all users to follow a set of standard procedures to process glider data (tools are being developed), and output all data in a standard NetCDF file-format (common to Seaglider and Slocum) – basically a system similar to the ARGO floats'. However we are not quite there yet unfortunately, but as the GROOM project is coming to an end this year I would expect to see some results coming out fairly soon.

For Seaglider data, the University of Washington (who invented the Seaglider) has been developing a new version of the basestation software which should provide a new thermal lag correction, more robust than the simple one currently performed by the basestation and possibly better than the one decided on by the EGO/GROOM community… (there may be more community wide discussions ahead in order to decide which processing to use).

Nevertheless, glider data users should soon have data delivered to them in a standard file format, with a stated data quality level. How and who will deliver those datafiles is another issue. For us at SAMS, we operate the gliders as well as use the data so it makes perfect sense that we also do the processing. Same goes for UEA and NOCL. But for users who are requesting gliders from the national pool (MARS), I do not think that MARS will do the data processing so my guess is that the PIs/scientists requesting the data will have to do it.