Customer	**:**	**ESTEC**	**Document Ref**	**:**	**FRM4SST-WR-SCL-001**
Contract No	**:**	**3-15990/19/NL/IA**	**Issue Date**	**:**	**19 October 2020**
WP No	**:**	**50**	**Issue**	**:**	**1**

Project : **FRM4SST (ships4sst)**

Title : **ISFRN Workshop Report**

Abstract : This document contains the proceedings of the ISFRN workshop which took place on 17 – 18 September 2020.

Compiled by : ______________________

Ruth Wilson
Space ConneXions Limited
(Project Manager)

Approved by : ______________________

Werenfrid Wimmer
University of Southampton
(Technical Manager)

Accepted by ESA: ______________________

Digitally signed by Craig James Donlon
Date: 2020.11.06 10:39:32 Z

Craig Donlon, ESTEC
(ESA Technical Officer)

Distribution : FRM4SST team members

Workshop participants

ESA

AMENDMENT RECORD

This document shall be amended by releasing a new edition of the document in its entirety. The Amendment Record Sheet below records the history and issue status of this document.

AMENDMENT RECORD SHEET

ISSUE	DATE	REASON FOR CHANGE
A	09/10/2020	Draft for first internal review
B	12/10/2020	Draft issued to presenters for review
1	19/10/2020	Issued to ESA following comments

TABLE OF CONTENTS

Executive Summary

The latest in a long series of UK and ESA-funded contracts to support the ISAR and SISTeR deployments is known as FRM4SST or "ships4SST". The aim of the ships4sst service is to validate satellite SST such as the Sentinel-3A and Sentinel-3B SLSTR SST data products and to promote and evolve the International SST FRM Radiometer Network (ISFRN or "Network"). To this end, an ISFRN workshop was held virtually on the 17 – 18 September 2020. The aim of the workshop was to bring together scientific and operational users and producers of in situ radiometer SST data from around the world to review progress, achievements and potential developments within the radiometer community, with several presentations given by world experts in several fields, including radiometer operators, data users and validation scientists.

In this workshop report we summarise the key points from the presentations and use participant feedback to comment on the current and future state of the shipborne radiometer network in assessing the accuracy of satellite-derived SST_{skin} and encouraging best practise in the collection, formatting and validation of SST_{skin} data.

We would like to thank and acknowledge the important contribution of all the participants and presenters in support of this workshop.

1. Introduction

1.1 Background

Satellite remote sensing of the Earth has become an essential tool in increasing our understanding of the climate, weather patterns and the impact of climate change. It continues to assist scientists in their analysis of the Earth's climate and policy makers in the formation of policies to adapt to or mitigate the effects of climate change. For this reason, remote sensing data must be as accurate as possible as well as long-term; i.e. they must be suitable for contributing to a reliable data series of linked satellite sensors, which requires that they be validated by comparison to common reference standards. To this end, *in situ* Thermal Infrared (TIR) radiometers are deployed on vessels across the globe to collect SST_{skin} data, which are then used to validate and verify the SST_{skin} data derived from the measurements of satellite radiometers. Ensuring the accuracies needed for climate research sets very stringent accuracy requirements[1].

Shipborne radiometric measurements provide the high accuracy surface temperature measurements (standard uncertainty <0.1 K) necessary to validate high accuracy satellite SST sensors such as the Sea and Land Surface Temperature Radiometer (SLSTR). Shipborne radiometers also provide a traceability route to SI (International System of Units) standards for satellite measurements and therefore a pathway to generate Climate Data Records (CDRs) from satellite SST_{skin} retrievals[2].

To achieve robust traceability to the SI temperature scale (ITS-90), the real-time calibration of shipborne radiometers derived from their internal blackbodies is regularly verified against SI-traceable laboratory calibration targets. The traceability of both the shipborne radiometers and the laboratory calibration targets are confirmed on a regular basis through inter-comparison exercises such as the ESA-funded Fiducial Reference Measurements for SST (FRM4STS) campaign[3] held in 2016 and the planned campaign for 2021/2022 (see section 6.3).

Whilst the protocols and procedures for maintaining robust traceability to SI standards are now well established within the Network, it is important to keep regular contact and have regular feedback from scientists and data users who share a common interest in the SST_{skin} data and satellite data validation results. This is the purpose of the ISFRN and the annual ISFRN workshop.

[1] Ohring, G., B. Wielicki, R. Spencer, B. Emery, and R. Datla, 2005: Satellite Instrument Calibration for Measuring Global Climate Change: Report of a Workshop. *Bull. Amer. Meteor. Soc.*, **86**, 1303–1314, https://doi.org/10.1175/BAMS-86-9-1303.
[2] Minnett, P. J., & Corlett, G. K. (2012). A pathway to generating Climate Data Records of sea-surface temperature from satellite measurements. *Deep-Sea Research Part II: Topical Studies in Oceanography*, *77-80*, 44-51. https://doi.org/10.1016/j.dsr2.2012.04.003
[3] Theocharous, E., and Coauthors, 2019: The 2016 CEOS Infrared Radiometer Comparison: Part II: Laboratory Comparison of Radiation Thermometers. *J. Atmos. Oceanic Technol.*, **36**, 1079–1092, https://doi.org/10.1175/JTECH-D-18-0032.1.

The ISFRN was set up to develop and support a community of in situ radiometer builders, operators and data users and to:

- Promote good practice in the construction and operation of *in situ* radiometers
- Agree and establish protocols, formats and standards for quality assurance of data
- Provide a single access point for the collection and dissemination of radiometer data
- Support satellite radiometer operators in the long-term validation of satellite products
- Share knowledge and coordinate activities between Network members
- Inform the wider community about the Network's activities

The aim of the annual workshop is, amongst other things, to understand the Network's progress against these objectives.

1.2 Workshop Structure

Due to COVID-19 travel restrictions, the ESA-sponsored workshop was held virtually over two days, at different times of the day to ensure that participants from around the globe could attend. The workshop consisted of online presentations and a poster, designed to review progress, results and advances in deployments, calibration and validation as well as to look at how the data from shipborne radiometers are used in practice. Time was also allowed for discussions between participants. The workshop consisted of the following sessions spread over two days:

- Session 1: Experiences of Radiometer Operators
- Session 2: Validation of Satellite SST Measurements
- Session 3: SST Data in Practice
- Session 4: The ISFRN Network
- Session 5: Radiometer Performance and Uncertainties

This sequence of topics also forms the framework of this report. A detailed agenda is included in the Appendix and can also be viewed, along with recordings of the two sessions, at www.ships4sst.org.

2. Experiences of Radiometer Operators

2.1 ISAR UK

The Infrared Sea Surface Temperature Autonomous Radiometer (ISAR) is a single channel (9.5 – 11.5µm) radiometer with a multi-angle sky and sea scan mirror. Routine deployments on the *Pride of Bilboa* began in 2004, after which it was moved onto the *Cap Finistere* between 2010-2012, before it was installed onto the *Pont Aven*, which has proved a successful route. Between these deployments approximately 960,000 SST measurements have been made. See Figure 1 for the full record of deployments. The ISAR has also been deployed on ad hoc cruises over the past few years for additional experiments including oceanic, land, lake and ice side by side comparisons.

Originally, the mirror surface was gold on glass substrate, however the company who produced this changed their cleaning process which resulted in poor results. Now a gold on copper mirror is used which works well. The instrument is autonomous and works in most environments; however careful maintenance is needed to keep the ISAR working well. Protocols are available for users via the ships4sst website.

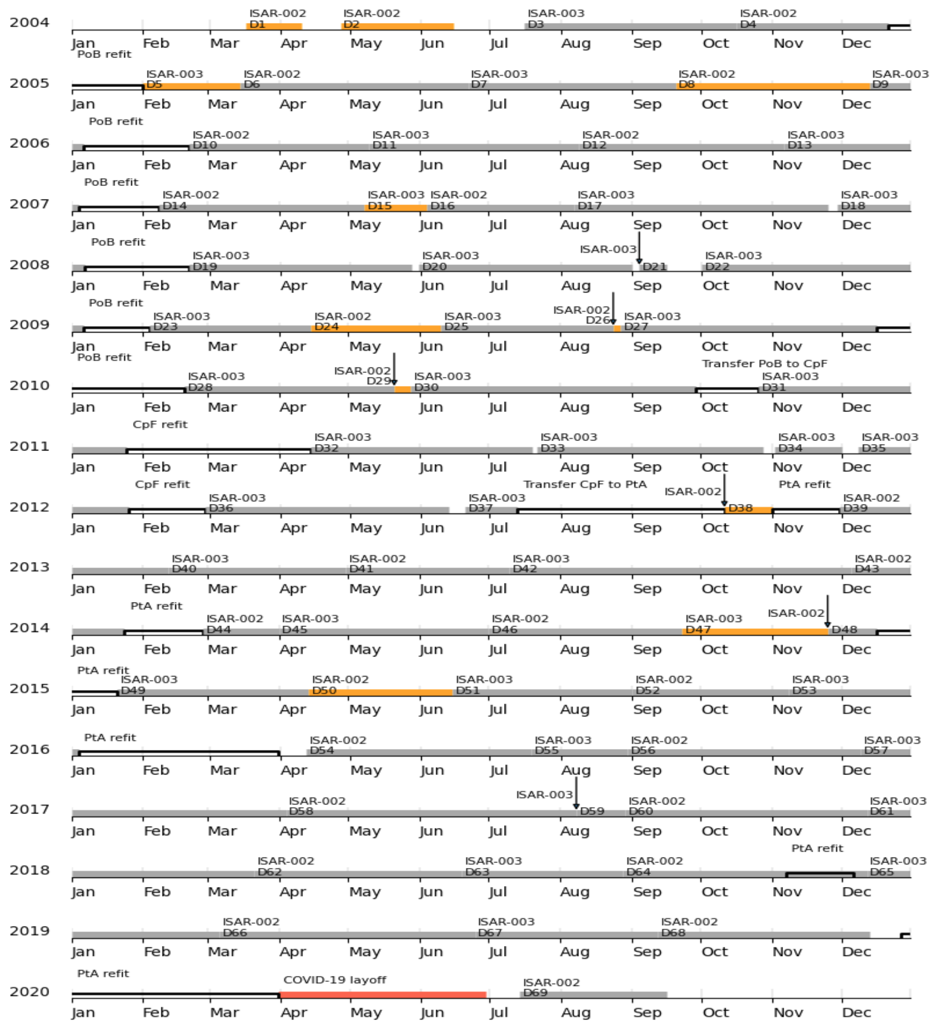

Figure 1: Bay of Biscay and English Channel deployments between 2004 – 2020. There are 69 deployments, approximately 5,000 days at sea and 200 SST_{skin} measurements a day. The orange parts show failures during deployments and red shows the break during the COVID-19 lockdown in the UK.

2.2 M-AERI

The Marine-Atmospheric Emitted Radiance Interferometer (M-AERI) is a very well-calibrated and stable sea-going Fourier transform infrared (IR) interferometer. It is calibrated before and after each deployment and contains two internal blackbodies for at-sea calibration. It can also be run autonomously with daily checks so can be deployed for months without maintenance. Deployments first began in 1996, and now 3 Mk2 M-AERI are usually deployed on Royal Caribbean International ships. COVID-19 caused a hiatus in 2020 but deployment is expected to resume in early October 2020. M-AERI has, like the ISARs, also taken part in inter-comparison workshops.

An uncertainty model has been made for each of the M-AERI instruments. The accuracy of the instrument's air-temperature measurements were validated during a comparison of radiometric and research grade conventional measurements between M-AERI and ATOMIC (Atlantic Tradewind Ocean-Atmosphere Mesoscale Interactions Campaign) during a cruise in 2005[4]. Here, an estimate of the near-surface air temperature was made from the measured spectra of the atmospheric CO^2 emission where the photon e-folding path length is ~7m. Radiative transfer modelling was used to decide on the best, 7m, path length. Too close and heat from the instrument could affect the results and too long a path length may subject the results to ship rolling. The results of difference in air-temperature between the instruments as a function of relative wind speed and of solar flux showed a high of a tenth of a degree Celsius warming at high wind speeds and solar flux, showing the excellent calibration of the M-AERI instrument.

M-AERI data between 2013 – 2020 are available on the ships4sst website and at https://doi.org/10.17604/bswq-0119. There are 212,437 SST data points over 1,529 days.

2.3 ISAR Australia

CSIRO (Nicole Morgan) is responsible for the calibration, maintenance and repairs of the ISAR on the *RV Investigator* vessel, a blue-water sea research ship. The Australian Bureau of Meterology (Helen Beggs, Janice Sisson and Joel Cabrie) are the primary data users of the ISAR data. *RV Investigator* is funded for 300 days at sea a year and carries a number of scientific instruments including a weather radar, barometer and ISAR. A temperature drop probe is located under the vessel on a drop keel which can move and so the depth of the probe varies throughout the voyage – this is important to consider as the probe is used to validate the ISAR SST data. The current ISAR was installed on the

[4] *Minnett, P.J. et al. (2005). Infrared interferometric measurements of the near surface air temperature over the oceans. Journal of Atmospheric and Oceanic Technology 22, 1016-1029.*

RV Investigator in 2014 and has completed 44 voyages to date, and provided 829 days of data. A new ISAR has been ordered and the expected delivery is mid-2021 when it will be installed on the *RSV Nuyina* (for the Australian Antarctic Division).

Some issues that need to be taken into account include:

- The mounting location of the ISAR is not ideal (port bridge wing); access is bad and can be wet and slippery when out at sea (and sometimes crew are not even allowed out). However, there are no other locations that are not significantly affected by the sea conditions.
- There is a short turnaround time between voyages, in some cases only a day, and so it isn't always possible to remove, calibrate and reinstall the ISAR before the vessel leaves again. This is why another ISAR has been ordered.
- Inability to calibrate for cold climates (a water temperature of approx 0º) – this requires a room with a cold ambient temperature to prevent condensation forming on the blackbody (BB), however, as the BB temperature increases over time, the ambient temperature remains cold so the calibration isn't as good and doesn't reflect a real-life situation.
- Noise on thermistor measurement circuit – there have been quite a few problems with noise, which is coming from the instrument power supply. It has affected the blackbody thermistor on all 43 voyages to date. The new ISAR that has been ordered addresses these issues.

Future plans:

- An environmental test chamber has just been funded to allow for better calibrations; it will allow the temperature to be increased slowly as the instrument BB temperature increases.
- When the new ISAR arrives in 2021 an ISAR side by side comparison will be done on the *RV Investigator*.
- Domestic collaborations – the Australian National Measurement Institute has shown some interest in participating in a comparison exercise
- After a 4 month shutdown, ISAR has just started up again. The following link shows where the Investigator is at any point on a voyage: www.cmar.csiro.au/data/underway

2.4 ISAR China

The ISAR-5C is deployed on the *Dong Fang Hong* 2 and 3 research vessels, at approximately ~13m in height on the ships. The instrument first took part in the Committee on Earth Observation Satellites (CEOS) inter-comparison of infrared radiometry in support of satellite calibration and validation for measuring SST for studies of climate change in 2009 before it went on its first shipborne deployment on 21 September 2009. ISAR was recently moved onto the new vessel in October 2019 and is currently on its 76th cruise (although only 2 cruises have been done on the new ship so far and analysis of the best instrument position is still underway). The external blackbody BB-ASSIST II (LR TECH INC) is used to calibrate ISAR before and after each campaign. It will be interesting to see how the ISAR measurements change due to the difference in character of the old and new ships.

Evaluation of Suomi NPP VIIRS SST: the ISAR data uses a temporal window of 1 hour and a spatial window of 0.01º. This equates to 853 match-ups. As part of the quality control, the match-ups that are close to the edge of cloud are removed to eliminate the cloud influence, which improves the standard deviation results to approximately 0.3 K with a bias of 0.14K.

2.5 ISAR Denmark

DMI has deployed their ISAR on the *MS Norröna* ferry, which travels between Denmark, the Faroe Islands and Iceland, since December 2017. It has an incidence angle of 25º and is deployed on round trips of 1 week. ISAR gets serviced and calibrated every 2-3 months due to the harsh weather in the North Atlantic. It also undergoes routine pre- and post- deployment calibration as an FRM and you can use the internal calibration BB to help correct performance e.g. correction when the mirror gets dirty.

A thermal camera has been used to assess the spatial variability of SST in various positions around the ship. This involves measuring the SST field around the ISAR for 10 minutes, during clear sky and cloudy conditions and then again in other locations around the ship with broken water. The camera was deployed several times on one cruise in day/night and cloudy/not cloudy conditions; there was significantly lower spatial variability when there were cloudy conditions (especially during broken water conditions) compared to clear sky (most likely an emissivity effect on the water). There is a plan to buy another thermal camera to run on more campaigns (this is handheld so not autonomous). The image data can be made available.

An FRM4STS field campaign called IST FICE was done in 2016 and proved successful with 3 research teams and 6 TIR radiometers. All the instruments were mounted on sea ice for the inter-comparisons. DMI go to Greenland 3 times a year but these are minimal exercises compared to the 2016 inter-comparison.

DMI are working on a microwave (MW) and IR radiometer inter-comparison; IR and MW measure fundamentally different temperatures (skin vs. subskin) and have different geophysical dependencies, therefore a simultaneous deployment could help compare IR and MW observations. This is important for existing CDRs and for homogenisation of future reference missions (e.g. SLSTR vs. CIMR). The focus will be particularly on cold waters. Within the ships4sst project, the Technical University of Denmark (DTU) are currently refurbishing their MW radiometers for DMI, specifically the C and X-band, with a performance <0.1 K for 1 second integration. A static deployment will then be done this year over 1-2 days between a pilot MW and IR radiometer and recommendations will be provided to guide constructions of future MW radiometers. There are plans to continue this work in more detail next year (pending funding).

Future Plans include:

- Continue operational deployments on *MS Norröna*

- Two more ISARS are expected within the next month and will be used to minimise the gaps in the data records, as well as to have a spare for additional campaigns.
- Conduct MR and IR ship inter-comparison
- Conduct basic ice campaign
- Calibrate Thermal IR camera and use UAV for spatial temperature variability assessments.

A clear specification of what is expected if an ISAR/radiometer is taken into different scenarios/environments, particularly in the arctic climate has been requested, given the effect sub-zero/cold temperatures have on power, insulation requirements etc. There is a configurable heater in the ISARs now which does not fix all the cold issues but it does help. A one page document with some requirements for a future generation radiometer based on the expected issues of instruments in different climates could be written.

2.6 SISTeR

SISTeR (Scanning Infrared SST Radiometer) is a chopped, autonomous, self-calibrating infrared filter radiometer that can measure IR brightness temperatures to high accuracy (~30mK). It measures the upwelling radiance from the sea surface and corrects for the reflected sky component with measurements of the downwelling sky radiance. The blackbody thermometer calibrations are traceable to ITS-90. SISTeR generates level 0 data and a dedicated processor unpacks this data. The SISTeR processor is coded in IDL and all higher level products are encoded in netCDF. Level 2 and level 3 SST products follow the L2R in-situ radiometer data format.

SISTeR was first deployed in 1997 and since 2010 has been deployed regularly on the Cunard *Queen Mary 2* (*QM2*) liner (north Atlantic between May - January and annual world cruise between January - May) where it is mounted on a dedicated platform above the starboard bridge wing. A data logger laptop is connected to the ship's Ethernet network and emails daily level 0 products back to the UK. In the past couple of years there have been some issues, including a degraded scan mirror and failed thermometry, both of which have been fixed. The instrument is still on 'cruise 22' as STFC have not been able to retrieve their instrument due to COVID-19 restrictions; the *QM2* is at the south coast of the UK but is currently inaccessible. Cruising is due to restart after a ship refit in spring 2021 (no around the world cruise is planned).

SISTeR is mounted approximately 30m high; to factor in this height difference, a small correction term (of order 100th degree Celsius) can be used as the radiometer is using part of the IR spectrum where the atmosphere isn't as transparent as at ~10µm. So the height of a radiometer onboard a ship should be modelled but overall, it has a very small effect on the measurements.

SISTeR participates periodically in radiometer inter-comparisons organised by the national metrology laboratories to validate the calibration chain. It has also made direct in-situ inter-comparisons with the

UoS ISAR, which could be extended to other instruments. Scene radiances are referred to two on-board blackbodies and instrument calibration is validated against an external CASOTS (Combined Action for the Study of the Ocean Thermal Skin) blackbody before and after every deployment.

Future plans include retrieving SISTeR from the ship at the earliest safe opportunity, assessing the current mirror durability and success, and producing a next generation of mirror using in-house manufacturing capabilities (diamond-turned solid copper stub with directly-deposited gold).

2.7 Skin Temperature Measurements for the Saildrone

There are a number of challenges for making measurements from a saildrone; currently there is a large effective incidence angle and only a downlooking radiometer so there are plans to add a sky view (i.e. a separate uplooking radiometer) at some point in the future. A bow-mounted sensor was added in 2019 to test out mitigation of an additional boom/wing angle effect. An incidence angle of 50 degrees was chosen. The range of effective incidence angles for the hull incidence angle and wing incidence angles were compared; the wing incidence angle covers a range of ~40 degrees but is in a range where the emissivity is approximately constant (~1.0) whereas the hull incidence angle covers ~10 degrees but the emissivity varies a lot more. Comparing the SST skin data from each section shows good agreement between the wing and hull data. The greatest differences occur at large roll and high winds. The differences at high winds seem due to the modulation of the large incidence angle of the hull measurements; so wing location is a preferable location as it minimizes the heel effect due to a smaller incidence angle.

There is no external calibration for a saildrone and the power supply is small which means that there is not enough power for a hot blackbody, so a simplification of the calibration is possibly required. During a 2016 SPURS-2 cruise with ROSR (Heitronics KT15), the performance using ambient BB only was examined, which suggest an improved KT15 stability; the calibration was stable for rates of change of solar radiation below +- 4.5 ºC per hr. Better instrument (temperature) insulation will help reduce the effect that changes in solar radiation (i.e. cloudy vs. not cloudy) have on the radiometer. Sensor stability experiments suggest that the ambient blackbody calibration is adequate.

3. Validation of Satellite SST Measurements

3.1 M-AERI validation of MODIS, VIIRS and SLSTR SST_{skin}

Target accuracies and decadal stability requirements are demanding and challenging to verify. Both buoys and radiometers are traditionally used to validate satellite SST_{skin} data, each having their pros and cons. For example, buoys are numerous, started in the early 1980s and so have a long time series, whilst radiometers are fewer in number with a shorter time series but they have very good calibration and are a comparison of like-with-like with satellite IR radiometers. Comparison with shipborne radiometers ensures that satellite SST_{skin} retrievals have an SI-traceable reference, and enable SST_{skin} CDRs to be generated.

M-AERI is one such radiometer. As noted previously, M-AERI is a very well-calibrated and stable sea-going Fourier Transform Infrared Interferometer. Two internal blackbody cavities with thermometers with NIST-traceable calibration are used for at-sea calibration. M-AERI is also calibrated before and after deployments using NIST-designed water-bath blackbody calibration target at RMSAS; this uses SI-traceable thermometers with mK accuracy and undergoes periodic radiometric characterisation by the NIST TXR and NPL AMBER radiometers.

3.1.1 SLSTR

Since 2017, M-AERI has been deployed on 4 cruises; the *Equinox*, *Allure*, *Adventure* and Ronald *H Brown*, where the SST_{skin} data obtained has been used to validate Sentinel-3a SLSTR SST[5]. Only the best (QL = 5) data was used to validate the SLSTR WCT L2P format data, which was retrieved from the Eumetsat Copernicus online data access server (https://codarep.eumetsat.int/). The WCT files comprise of the 'best' retrievals from four possible algorithms; N2 (across track single view day time retrieval), N3 (across track single view night time), D2 (dual view day time) and D3 (dual view night time). 5216 comparisons were used, the majority at warm (~300 K) temperatures. The standard deviation (STD) is a little high, largely due to some 'cold tails' and this may be an issue with cloud screening.

There are no data points that are common to all algorithms; so when comparing the SLSTR SST_{skin} with M-AERI SST_{skin} by retrieval type, the N2 channel came out best (though there are a small number of samples as N2 is considered to be the least accurate channel). The means and medians are very good on average but there is a lot of scatter and it is not clear what is causing this. The robust standard deviations (RSD) are a bit disappointing in some places – some of this may be due to proximity to coasts (off-shore wind interference and taking a while for the boundary layer of the atmosphere to adjust to the marine conditions). In general however, the comparison statistics are good with a median of -8mK and a RSD of 0.296mK.

[5] Luo, B. et al. (2020). Validation of Sentinel-3A SLSTR derived Sea-Surface Skin Temperatures with those of the shipborne M-AERI. Remote Sensing of Environment 244, 111826. https://doi.org/10.1016/j.rse.2020.111826

3.1.2 MODIS

MODIS on Terra and Aqua; Terra was launched on 18 December 1999 and Aqua was launched on 4 May 2002 and was recently revitalised after a data formatter failure caused 2 weeks of loss of data. During Q3-4 of 2019 all missions were re-processed by Goddard Space Flight Center (GSFC); this reprocessing is called R2019. A full set of match-up databases is not available yet (COVID-19 shut down GSFC and they are now running on a reduced set of servers so progress is slow). The major changes in R2019 include:

- Replacing the NOAA OI 'Reynolds' SSTs with the CMC as the reference field.
- New cloud screening – alternating decision trees[6]
- Night-time aerosol correction – additive term to atmospheric correction algorithm if an aerosol index threshold passed[7]
- High Latitude coefficients[8]
- Improvement to cloud-ice discrimination.

So far, the results when using the partial R2019 MODIS (11 and 12 micron) SST_{skin} data against the M-AERI SST_{skin} data show median and RSD values that are not as good as that for the previous R2014 dataset.

Regarding the MODIS mission, there is a plan to possibly continue another ~3 year period of support for the MODIS missions. It is likely that in late 2021-2022 the Terra orbit will not be maintained and will be allowed to drift with a controlled re-entry. Aqua would presumably follow a year or so later.

3.1.3 VIIRS

Suomi-NPP VIIRS was launched on 28 October 2011 and has fewer channels than MODIS as it is missing the SST4 pair. The NASA SST_{skin} atmospheric correction algorithm is comparable to MODIS NLSST, and the NASA night-time-only algorithm is SST_{triple} based on at lambda = 3.70, 10.8 and 12.0 microns.

Overall, SLSTR, MODIS and VIIRS are producing very good SST_{skin}, but there is room for improvements. Outstanding issues include better cloud screening and atmospheric correction algorithms, developing full error and uncertainty budgets for satellite-derived SST_{skin}, assessing sampling errors of *in situ* SST_{skin} measurements, improved modelling of thermal skin effect and the Sensor Specific Error Statistics (SSES) for each SST_{skin} product should be revisited.

---

[6] Kilpatrick. K.A. et a., (2019) Alternating Decision Trees for Cloud Masking in MODIS and VIIRS NASA Sea Surface Temperature Products. Journal of Atmospheric and Oceanic Technology 36, 387 – 407 DOI: 10.1175/jtech-d-18-0103.1

[7] Luo, B. et al. (2019). Improving Satellite Retrieved night-time infrared sea surface temperatures in aerosol contaminated regions. Remote Sensing of Environment 223, 8-20. https://doi.org/10/1016/j.rse.2019.01.009

[8] Jia, C., Minnett, P.J. (2020). High Latitude SST Derived from MODIS Infrared Measurements. Remote Sensing of Environment – Accepted.

3.2 The Evaluation of the *in situ* SST quality control applied in iQuam

The *in situ* SST Quality Monitor (iQuam) system was established at NOAA in 2009[9] to support the accurate and consistent calibration and validation of satellite and blended SST products. Its objectives are to:

1. Pull together a comprehensive set of *in situ* data from various sources covering full satellite era 1981-onwards,
2. Perform the advanced, flexible and unified community consensus QC,
3. Monitor quality controlled SST online, and
4. Distribute to users in near-real time.

When comparing the performance of iQuam QC against external data sources such as ICOADS (the International Climate Ocean-Atmosphere Data Set), Argo and IMOS (Integrated Marine Observing System) ship data, iQuam QC shows robust performance for all platforms under various situations. Data that passes iQuam QC but fails external QC appear of consistently good quality, whereas data that pass external QC and fail iQuam QC frequently show unstable behaviours with degraded statistics and appear as large outliers in the corresponding time series.

During a recent study by Haifing Zhang and his colleagues, several improvements to iQuam QC were identified. The main one being more frequent screening of large diurnal signals; all the reference fields currently available (e.g. CMC, Reynolds) are foundation products that are reported only once a day and do not resolve the diurnal cycle. In certain dynamic regions (e.g. where there are strong currents, steep temperature gradients) the current L4 analyses may under-represent detailed spatial features. Future work to improve iQuam QC may therefore focus on an update of the reference field, incorporation of a DV (diurnal variability) model and exploring QC enhancements from other systems, such as the Met Office. There is also a plan to incorporate data from shipborne radiometers into iQuam in the future.

3.3 Validation work at Lake Tahoe and Salton Sea

Work at Lake Tahoe began in 1999 and has involved multiple buoys with *in situ* box sensors and radiometers. The site was chosen due to its large size, high location (2km latitude), freshwater, easy access and large annual temperature range of between 5 – 25ºC. The site is therefore available all year round for validation measurements. There are currently 4 solar-powered buoys at Lake Tahoe, each containing a radiometer and instruments that measure bulk water temperature at different depths, air temperature, wind speed and relative humidity. The measurements are taken every 2 minutes, every day and sent to the lab via a modem.

[9] Xu & Ignatov, (2014). www.star.nesdis.noaa.gov/socd/sst/iquam

In 2008 work also began at the Salton Sea which sits below sea level and is much warmer than Lake Tahoe, so between the two sites, there is the entire range of *in situ* liquid water temperatures to work with. Simon Hook and colleagues have been using both sites to validate a whole range of instruments.

The process used in order to perform validation is to:

1. Extract the bulk temperatures and radiometric temperature, which is corrected SST_{skin} using the measured emissivity of water,
2. Propagate the SST_{skin} to the satellite using a radiative transfer model and interpolated atmospheric profile (usually ~tenths of a degree adjustment to the measurement in the window channels),
3. Convolve the propagated at-sensor radiance to the instrument response function to obtain the Vicarious Radiance (VR),
4. Extract the image radiance derived using the On Board Calibrator (OBC), and
5. Compare and contrast the OBC and VR Radiance values.

Whilst data has been validated from multiple instruments including AVHRR, AATSR, ASTER, MODIS (Terra and Aqua), Landsat 5 and ETM+, MTI, VIIRS (very good results so far), and ECOSTRESS, here we focus on MODIS validation. Simon Hook et al have approximately 22,000 clear sky match-ups at a range of angles, once the data is whittled down to between 0 - 30 ºC you get about 500 match-ups per year (this is every year from 2000 to present day). IR window bands 29, 31 and 32 are the main bands that are used in ocean applications, and whilst bands 31 and 32 align very nicely when plotting the OBC radiance against MODIS Terra vicarious radiance for v6.1 data, band 29 does not. Looking at collection 6.0, band 29 has excellent calibration until 2009, when there is clear degradation of the calibration which impacted surface retrieval. This issue has now been corrected in version 6.1 of the data by the MODIS team and illustrates that long term measurements need to be made in order to notice these kinds of issues. Interestingly, similar problems in channel band 29 on Aqua-MODIS are now being seen, at a similar length of time after launch. This will undoubtedly have to be fixed as was done in v6.0 of the Terra data. Looking at night-time MIR data of MODIS Aqua and Terra, it looks like the mid-IR in Terra needs a small correction.

A number of papers have been written about the validation results of individual satellite sensors over the years, but a paper summarising the validation results for all the sensors has not yet been written. Participants in this workshop showed a strong interest in such a paper.

3.4 Validation of SST and SSS Gradients Using the Saildrone Baja and Gulf Stream Deployments

Comparisons of SST gradients are critical for applications to coastal regimes where mesoscale-submesoscale dominate. A saildrone is an autonomous surface vehicle with ~1 minute sampling, so it is very high resolution and as such provides an excellent platform for validation and the application of SST gradients.

Care does needs to be taken in the derivation of match up pixels with satellite data and so a new co-location strategy has been developed for the derivation and comparison of SST and Sea surface salinity (SSS) gradients; for every grid point of a L4 SST/SSS product, all saildrone measurements inside that grid are averaged and the average acquisition time of the saildrone measurement is computed. These are then sorted to generate a collocated time series of L4 SST/SSS/Saildrone. Gradients are then derived as difference between successive points of the time series and accounting for the distance in space between points.

When viewing the SST and SSS gradients during the 2018 Baja and the 2019 Gulf Stream deployments, comparisons between GHRSST level 4 datasets (e.g. CMC, OSTIA, MURS) and saildrones look fairly good at >0.90. There are some higher gradients that are associated with coastal upwelling and land contamination. All of the SSTs show very strong correlation and low bias with the matchups (apart from MURS) however, it is clear that this does not necessarily lead to high correlation in SST gradients, which drop significantly with larger differences between the products. The MUR derived SST gradients showed the best correlation for the Baja deployment.

For the Gulf Stream deployment SST showed clear relationships to major frontal features associated with the Gulf Stream. Correlations range from 0.3 to 0.4. Cross correlations between L4 saildrone data and the JPLSMAP products (which is an 8 day running mean) show a maximum that does not appear at 0 lag, this could be because fronts move very quickly and the average is not necessarily picking up when the gradient actually occurs between the saildrone and the satellite products.

3.5 Using ships4SST data to validate SLSTR Data

SLSTR reprocessed data are available between 2016 – 2018 (v006 r1i1) and the near-real time data are available for 2018 onwards. FRM radiometer data are available from ships4sst and the MDB files are produced by Felyx. These are reprocessed to allow multiple matches per match-up location (overpass) and follow the Wimmer et al 2012 approach. They validate WST, D3, D2, N3, N2 and all GHRSST CV levels. The quality indicator method which was developed for AATSR has been adapted for use on SLSTR. The focus of this section is on L2b data (2hr temporal and 1km spatial resolutions).

Validation results show generally good data and the QI flagging (using GHRSST level 5) gets rid of the outliers in the data. The quality level 5 products show good/realistic temperature range and a mean difference and standard deviation that are slightly lower than for all the data, which is to be expected because the quality of the data is improved.

The match–up uncertainty for SLSTR reprocessed data between August 2016 and April 2018 has also been investigated. This method uses 5 indicators that are run on match-up grade 4 (QI 0 is grade 4) and uses 4 levels (0, 3, 4, 5). These are:

1. Satellite variability
2. Satellite trend
3. Radiometer variability
4. Radiometer trend
5. Radiometer Sky BT

When this threshold method is used there are only 2 matches for the best match-up quality level (5) – these are very good matches but obviously a very low number of matches at this level. Some more work needs to be done to work out why there are so few matches at the highest quality level (QI5).

Match-up temporal and spatial windows: the results for GHRSST L5 data show nice results for D3 and N3 (night-time data) with a robust standard deviation around 0.18 and a mean difference of -0.01 and 0.01 respectively. Daytime data is slightly noisier with a slightly greater offset compared to night-time data.

In conclusion, validation of SLSTR with FRM data shows that SLSTR is performing very well with virtually no mean difference at night and only a small difference at daytime. The robust standard deviation is lower at night than daytime and comparable to AATSR. It is worth having the 3 channels of data at night-time as they show varying degrees of match-up accuracy. An investigation into the difference between the AATSR and SLSTR results is planned for the future; possible causes include too stringent/wrong thresholds for SLSTR, or processing differences.

Other results and data are also available and include:

- NRT data is available for 2018,
- Results for GHRSST CV 3 and 4,
- Dependence plots for WST, D3, D2, N3 and N2
- QI method on D3, D2, N3 and N2
- Regional results for each route, split by ships name. These have the same set of statistics and plots as the global results.

When uploading reprocessed skin data from ships onto the ships4sst database, note that the update frequency of the verification is dependent on EUMETSAT. W. Wimmer lets data providers know via email if there is a verification coming up.

3.6 SST CCI Validation

The ESA Climate Change Initiative (CCI) programme produces satellite-based Climate Data Records (CDR) for a range of Essential Climate Variables (ECVs), including SST. The aim of the SST CCI CDR is to be independent of *in situ* SST measurements with context-sensitive uncertainty estimates and to be of useful, quantified accuracy and sensitivity.

SST CCI Phase 2 dataset v2 was released at the beginning of 2019[10][11]; it contains 35 years of SST data (Sept 1981 – Dec 2016) and includes L2P swath, L3U gridded and daily L3C products for all sensors (18 x 10^{12} satellite radiance measurements in total). The SST type used is skin at satellite overpass and SST_{20cm} at 10:30 local time and the uncertainties provided are random, correlated and systematic. The SST CCI uncertainty is an output of the retrieval and is independent of *in situ* data; therefore the *in situ* data can be used to validate the uncertainty. The dataset is currently being extended with SLSTR data via C3S ICDR, which is available until the end of 2019. The 2020 data is currently being processed and should be available within the next few months.

The reference *in situ* data is supplied by the Met Office Hadley Centre (HadIOD 21.2.0.0). Multiple platform types are included but most of the validation uses drifters (ICOADS and CMEMS) as they provide the most 'complete' spatial coverage (although limited before 1995 and best from 2005). The ICOADS drifters drop off from 2016 onwards. GTMBA moorings are also used for stability analysis[12] but since ~2012 the number of observations has decreased substantially. In the early 1990s and the last couple of years the Atlantic and Indian Ocean observations are sparse or non-existent.

The Multi-sensor Match-up System (MMS) developed by Brockmann Consult is used to match the *in situ* to the satellite data. It is similar to Felyx and can produce match-up datasets from various combinations of inputs, e.g. satellite L1b, L2P, *in situ* and NWP.

The L2/L3 validation against drifters (using satellite SST_{20cm} and the Fairall-Kantha-Clayson model for time/depth adjustments) shows a bias for NOAA-7 (referenced to *in situ*) and NOAA-18 (referenced to ATSR). The dual view instruments (ATSRs and SLSTRs) all have very low biases. Regarding the uncertainties, ATSR and NOAA-07 show good uncertainty whilst the MetOp uncertainties are overestimated (also applies to NOAA-12 onwards). ATSR/AATSR (which are fully independent) have a global bias ≤ 0.01 K whilst AVHRR 7, 9, 11 (which are tuned in to *in situ*) and AVHRR 12 onwards (tuned to ATSR) have a global bias of ≤ 0.1 K except for AVHRR 7 and 18.

For applications requiring night-time SLSTR SSTs as close as possible to drifting buoy SSTs, the recommendation is to use the reprocessed (2017-2019) C3S "adjusted" SLSTR-A L3C SST(0.2m), remembering that it is also adjusted in time (assuming 10:30 LDT am or pm).

---

[10] Merchant, C.J., Embury, O., Bulgin, C.E. *et al.* Satellite-based time-series of sea-surface temperature since 1981 for climate applications. *Sci Data* **6,** 223 (2019). https://doi.org/10.1038/s41597-019-0236-x
[11] Data available at http://cci.esa.int/data
[12] Berry, D.I.; Corlett, G.K.; Embury, O.; Merchant, C.J. Stability Assessment of the (A)ATSR Sea Surface Temperature Climate Dataset from the European Space Agency Climate Change Initiative. *Remote Sens.* **2018**, *10*, 126.

3.7 SLSTR Match-up Database

The main components in SLSTR SST validation are the match-ups between satellite and *in situ* data on the Felyx platform. Satellite data is used from SLSTR-A/B, AVHRR-B, IASI-B and VIIRS-NPP, from April 2018 to present for near real time (NRT) data and August 2016 – April 2018 for reprocessed. *In situ* data is used from drifters, Argo, moored (from CMEMS), TRUSTED, radiometers (from ISFRN) and saildrones (currently being introduced into the MDB and will be included in the next reprocessing). For drifters, Argo and moored instruments data is available in less than a week, whereas radiometer and saildrone data is available after >1 month. The MDB runs daily with a 1 week delay to allow all the *in situ* data to be collected. All SLSTR L1 data is processed within a 1 month (it is in a 1 month rolling archive) whilst SLSTR L2 WST/WCT takes a little longer to become available. Data are available to the Sentinel-3 Validation Team (S3VT) and via the link: sftp://s3calval.eumetsat.int. Quality control is not performed when data are ingested as QI is done before and after ingestion (during validation).

The objective of S3VT is "to provide independent validation evidence, experimental data and recommendations to the S3 Mission". There are 5 S3VT sub-groups which are altimetry, land, ocean colour, temperature and atmosphere. To become a S3VT member a proposal needs to be submitted (s3vt.org) and access to the SLSTR MDB requested. Within the 'Temperature' team, activities range from shipborne radiometers to drifting buoys to coastal, fronts and calibration work.

The SLSTR MDB products are produced in netCDF4 format and split into 4 types:

- satellite platform (S3A/S3B),
- satellite data type (a core: WST (L2P) and 4 auxiliary types: WCT (L2 SST algorithms), MET (meteorological information), RBT-i (L1 IR channels), RBT-a (L1 VIS/SWIR channels)),
- *in situ* types (drifters, Argo, moored, radiometers), and
- Assembling period (drifters in NRT = 6hr, all other *in situ* types = 1 day).

A new radiometer dataset version called ships4sst (r1i1) has been created. Updates include a new *in situ* radiometer field, and a fix in Felyx for names with forward slash (e.g. R/V). The reprocessed and NRT MDB have been processed and the full 2019 SLSTR SST MDB is currently in progress. Currently, there are many different processor versions on the database, so it would be useful to simplify the processor versions by having the same processor version for the same instrument-types (i.e. have more consistency). It is not currently possible to reprocess the MDB only for a single radiometer version and it is not convenient to reprocess the whole radiometer MDB for one dataset update. A current limitation of the Felyx system is that the MDB is not designed to have different versions of the same *in situ* type; this will be addressed in the evolution of the MDB.

Future plans include:

- Short term plans: MDB in S3A early commissioning (06-07/2016). S3A/S3B tandem phase (for S3B) – all *in situ* types. RTM (RTTOV) and FKC (SST adjustments). The TRUSTED/HRSST international review workshop is on 2 - 4 March 2021.
- Long term plans: improving the MDB reliability and robustness, and to simplify installation and maintenance.

4. SST Data in Practice

4.1 A Virtual Ride of Discovery: Exploring the Earth's Climate using Drones

The *R/V Falkor* cruise took place in 2019 near Fiji in an area full of trichodesmium and cynobacteria blooms in order to understand how they impact the upper Earth heat budget. The aim of the study is to understand how the ocean is changing on a number of spatial and temporal scales and observe how the oceans ecosystem survives and adapts to changes in the environment. Autonomous aircraft (UAV) were deployed that could map a large area efficiently and with very high resolution (greater resolution than satellite), for example, an underwater pumice was discovered that could then be further analysed by shipborne equipment. The aircraft has complete autonomous takeoff and landing from ships and a dual GPS system determines the aircraft heading. Additionally, the ground station on the ship uses an ALIGN system to send the aircraft data including the ships heading and heave, to allow the aircraft to land autonomously on a moving platform at sea. The aircraft have a long-range (50 nm) capability with a high bandwidth data link that enables real-time mission control and tasking, and sensor payloads can be changed depending on what you wish to measure.

Christopher Zappa showed an example of a cynobacterial patch, showing an estimate of what the thricholosat enhancement is in the skin layer against time. Results from a drifting buoy show that the SST skin and bulk SST are being underestimated by the PWP Model[13]. Another example was of a pumice streak (from an underwater volcano) that was tracked down using dual-UAV aircraft flights and measured in infrared and visible. These achievements during the 2019 *R/V Falkor* cruise demonstrated, amongst other things, 24 hour operations and high endurance flights of > 8 hours by the autonomous aircraft.

ECOSTRESS[14] (a high spatial resolution thermal imager) normally images over land but has the ability to image spots over the ocean. It would be interesting if the autonomous aircraft can locate things such as the pumice, and then a high visual image (of order 60m) could be taken using ECOSTRESS.

4.2 The Sentinel-3 Mission

Anne O'Carroll presented an overview of the Copernicus S-3 mission. Copernicus is funded and managed by the EU. EUMETSAT and ESA both have mission product responsibilities, including Sentinel-3 L1 global data and various L2 products. There are currently two Sentinel satellites in operation; S3A (launched 16 Feb 2016) and S3B (launched 25 April 2018) that work together to optimise coverage with full global coverage in <3 days (OLCI) and <2 days (SLSTR) at the equator,

[13] Wurl et al. (2018), Geophys, Res. Lett., 45 (9), 4230-4237, doi:10.1029/2018GL077946
[14] https://ecostress.jpl.nasa.gov/

and, during the tandem phase, provided inter-calibration opportunities. EUMETSAT perform the routine spacecraft operations and marine data processing/performance/dissemination. There are many applications for ocean research and commercial operations.

SLSTR measures SST with a spatial resolution of 500m for optical and 1km granules for thermal infrared (TIR). It has a number of channels that are used to measure various ECVs, but bands S7 (3.74µm), S8 (10.85µm) and S9 (12µm) focus on measuring SST in TIR. Level 1 data includes the brightness temperatures whilst Level 2 contains GHRSST user product L2P SST with quality flags, meteorological parameters, singe sensor error statistics, algorithm flags plus internal WCT SST (single algorithm). SLSTR has dual-view retrievals that have the ability to provide better atmospheric correction, and 5 algorithms which combine different nadir and oblique views for SST retrieval. Users can consider S3A and S3B as one sensor, and validation results show that SLSTR can be used as a reference sensor against other SST datasets.

Investigating the harmonisation aspects within SLSTR data and together with other satellite SST datasets is something EUMETSAT is interested in and will be working more on in the next few years.

There have been a few updates in 2020 to the SLSTR processing baseline but none that affect SST particularly. Future updates include improvements to Bayesian cloud, especially in the coastal zones, and a revision of the algorithm to produce D2 SST around the 2023 timeframe.

SST validation at EUMETSAT involves inter-comparisons with *in situ* and satellite data.

Projects going on at EUMETSAT include:

- TRUSTED = an improved drifting buoy project. The drifters contain 2 SST sensors, a near surface water pressure sensor, and high frequency data are available. TRUSTED is a 4 year project that began in January 2018. One of the buoys has been recovered; however it is still on a ship and is due back in Brest in November. A workshop will take place 2 - 4 March 2021, most likely online.
- A thermal infrared inter-comparisons experiment that compares TIR *in si*tu and satellite datasets took place on Lake Constance in September 2020 with W. Wimmer.
- Currently preparing a TRUSTED and SLSTR match-up database which will be distributed over the next couple of months.

4.3 The Importance of Radiometric/skin SST on Air-Sea Fluxes

The SST_{skin} is the inter-facial temperature that the atmosphere 'sees' and so it is the temperature that is needed for accurate calculations of the air-sea fluxes. In 1996, a paper[15] was released that showed that in order to calculate more accurate fluxes something that measured inter-facial SST is needed.

[15] Fairall, C. W. et al. (1996), Bulk parameterization of air-sea fluxes for Tropical Ocean-Global Atmosphere Coupled-Ocean Atmosphere Response Experiment, *J. Geophys. Res.*, 101(C2), 3747– 3764, doi:10.1029/95JC03205.

The authors of the paper found that a SST error of 1ºC (including latent and sensible heat flux) can result in a flux uncertainty of the order 40 W/m^2.

The momentum flux can be measured directly if you can measure the turbulent components of the wind, temperature and humidity on the surface – this is difficult to do and requires specialised instruments. Therefore it is not done often and typically an estimate is made in which the relative difference between the wind and the surface currents / difference in temp (at height and at SST skin) / difference in humidity for momentum- / sensible heat- / latent heat- fluxes respectively are calculated. These measurements are combined with others[16] to produce an estimate of the fluxes.

For air-sea feedback from diurnal SST and during suppressed and active Madden-Julian Oscillation (MJO), the difference in the fluxes is much larger during a suppressed MJO where you get a lot of diurnal warming (as is expected)[17]. A SeaFlux satellite data set[18] was used to plot the difference in diurnal vs. non-diurnal fluxes (long wave + sensible heat flux + latent heat flux) over a global map for 10 years worth of data and the resulting map shows the location of the peak solar changing as the day goes by and the impact of the wind so it is not a monolithic variable. The maximum instantaneous difference between the deeper SST vs. SST_{skin} over the 10 years was up to ~400Wm^{-2}. If you average out the data over the 10 years then the mean effect on the flux is, globally around 6Wm^{-2} and in the tropics this is higher (for both day and night).

When you are calculating surface fluxes the correct temperature to use is the sea skin temperature. There can be a 5 - 20% difference in flux measurements at lower wind speeds at different depths. Improvements can be made via more SST skin measurements, ideally by buoys.

Ships are not the ideal platform for making accurate flux measurements as you can get flow distortion and so buoys are more commonly used. However, the sea snake (~1-2cm depth) that is used to get the near SST_{skin} on ships would not work well with buoys so having a radiometer on a saildrone that could make the flux measurements would be useful. The type of equipment needed on buoys to make flux measurements include; DCFS measurements, motion correction packages, light core system, (these have started to be done on the OI for operational uses) , however it is a bit expensive so is not available on every buoy.

4.4 Monitoring the Mediterranean Sea

Jordi Isern-Fontane et al look at global SST but focus mainly on the Mediterranean, which has very good conditions for IR measurements. The Rosby radius is relatively small (~10-20km) and the sea is

[16] Edson, J. B., and Coauthors, 2013: On the Exchange of Momentum over the Open Ocean. *J. Phys. Oceanogr.*, 43, 1589–1610, https://doi.org/10.1175/JPO-D-12-0173.1.
[17] Clayson, C.A. and Roberts, J.: 2016. Diurnal warming impacts on atmospheric and oceanic evolution during the suppressed phase of the Madden Julian Oscillation . American Geophysical Union.
[18] Clayson, Carol Anne; Brown, Jeremiah; and NOAA CDR Program (2016). NOAA Climate Data Record Ocean Surface Bundle (OSB) Climate Data Record (CDR) of Ocean Heat Fluxes, Version 2. NOAA National Center for Environmental Information. doi:10.7289/V59K4885

deep (depths ranging from 1000 m to ~4000 m). It is important to monitor the Mediterranean Sea as it is very variable and responsive to climate change, and there is more than 45 years of measurements near the Catalan coast so there is a long time-series of *in situ* temperature here.

There are various validation and testing sites. The Casablanca oil platform is one such site and provides an excellent platform for deploying instruments for satellite calibration. It has:

1. Very good location for testing for various reasons including, ease for accessing platform and moored instruments, covered by HF radar, it is safe as the site is only accessible when access is requested and it is 'cheap' to access.
2. The platform is located ~40 km from the coast and at a ~200 m depth.
3. A number of instruments have been used on the site; WISE (MIRAS/SMOS) was previously used and the Ocean Colour (OC) AERONET site is currently used to measure ADCP, wind and waves.

To fully exploit the potential of SST observations, a dynamical Model is needed. Jordi and his team are currently developing a dynamical framework to characterise the energy cascade from SST by working on the reconstruction of surface currents[19] and monitoring the potential energy cascade. The energy cascade is very important in understanding the model of the ocean as the statistical properties of SST depend on the intensity of the strongest thermal fronts.

Through measuring and modelling the thermal fronts using SST from satellite observations (using L1B (A)ATSR data + SNAP to get L2 SST + GHRSST cloud mask), a connection between the structure functions and the most intense fronts has been (theoretically) found. The plan is to use the same approach using real *in situ* data to monitor the evolution of the global ocean as well as the Mediterranean Sea. A couple of issues, including the role of instrument noise[20], and cloud masks (which mask out strong fronts) have on the thermal fronts also need further investigation.

4.5 Verifying the consolidated theory of atmosphere-ocean CO_2 fluxes and the importance of the skin (AMT4OceanSatFlux)

It is important to understand how much carbon goes into the ocean, not only because of its impact on marine life and our ability to identify regions and ecosystems at risk, but also because we need to be able to quantify global carbon storage and exchange. Whilst the atmospheric and oceanic carbon sinks can be measured, it is currently very difficult to observe or measure the global land carbon sink.

[19] González Haro, Cristina, et. al. (2020). Ocean surface currents reconstruction: Spectral characterization of the transfer function between SST and SSH. Journal of Geophysical Research: Oceans. e2019JC015958. 10.1029/2019JC015958.
[20] Isern-Fontanet, J., and Hascoët, E. (2014), Diagnosis of high-resolution upper ocean dynamics from noisy sea surface temperatures, *J. Geophys. Res. Oceans*, 119, 121– 132, doi:10.1002/2013JC009176.

Instead the land sink is estimated, and all existing methods to estimate the land sink rely upon first being able to accurately quantifying the ocean carbon sink[21] .

To improve our understanding of the exchange across the air-sea interface[22] (i.e. exchange across the top and bottom of the mass boundary layer) a series of ESA projects looked at transferring the theory that is well established in the SST community to that which was applicable to the carbonate system community, specifically consolidating methods for temperature and salinity handling within bulk gas flux calculations[23][24]. This work highlighted the importance of correcting and fully accounting for the near-surface temperature and salinity gradients.

An open source toolbox was then created for community use and to apply the different methods for calculating gas fluxes[25].

Global analyses have identified that ignoring vertical temperature and salinity gradients results in a systematic underestimation in the ocean sink of CO_2, suggesting that the oceanic sink is actually much larger than previously thought. Recent more detailed work[26] supports these early results and shows that the annual difference in ocean uptake can amount to ~10% of annual global fossil fuel emissions. The work suggests that a revision of the global carbon budget is now required; *in situ* SST_{skin} data can be used to verify the consolidated theory by reconciling two sets of independent measurements of gas fluxes. A preliminary study on the impact of mishandling temperature for *in situ* bulk measurements has been done which led to the OceanSatFlux inter-comparison. The OceanSatFlux project is now performing an inter-comparison between direct (eddy covariance) and indirect (bulk) air-sea CO_2 flux calculation techniques to evaluate the theory of the impact of vertical temperature and salinity gradients on air-sea CO_2 exchange, and to advance the uncertainty analyses within air-sea exchange studies (bulk and eddy covariance). The data obtained so far consolidates the theory but more data are needed to be sure. Results found that the largest difference between indirect and direct flux measurements occur across oceanic boundaries where the SST vertical profiles are more variable, reinforcing the idea that vertical temperature gradients should not be ignored when estimating the oceanic CO_2 sink. This work requires accurate SST measurements and their depth, along with temporally and spatially coincident carbonate system measurements and direct gas flux measurements

---

[21] Shutler JD, Wanninkhof R, Nightingale PD, Woolf DK, Bakker DCE, Watson A, Ashton I, Holding T, Chapron B, Quilfen Y, et al (2019). Satellites will address critical science priorities for quantifying ocean carbon. *Frontiers in Ecology and the Environment*, *18*(1), 27-35.

[22] Garbe C.S. et al. (2014) Transfer Across the Air-Sea Interface. In: Liss P.S., Johnson M.T. (eds) Ocean-Atmosphere Interactions of Gases and Particles. Springer Earth System Sciences. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-25643-1_2

[23] Woolf, D. K et al. (2016), On the calculation of air-sea fluxes of CO_2 in the presence of temperature and salinity gradients, *J. Geophys. Res. Oceans*, 121, 1229– 1248, doi:10.1002/2015JC011427.

[24] Goddijn-Murphy, L. M.et al.: The OceanFlux Greenhouse Gases methodology for deriving a sea surface climatology of CO_2 fugacity in support of air–sea gas flux studies, Ocean Sci., 11, 519–541, https://doi.org/10.5194/os-11-519-2015, 2015.

[25] https://github.com/oceanflux-ghg/FluxEngine

[26] Watson, A.J., Schuster, U., Shutler, J.D. *et al.* Revised estimates of ocean-atmosphere CO_2 flux are consistent with ocean carbon inventory. *Nat Commun* **11,** 4422 (2020). https://doi.org/10.1038/s41467-020-18203-3

5. The ISFRN Network

5.1 Status of the ISFRN

The ISFRN is an international network of ocean and remote sensing scientists who share a particular interest in promoting and improving the use of shipborne infrared radiometers for measuring SST_{skin} at the surface of the ocean, comparable to the retrievals made by satellite infrared radiometers. Objectives of the ISFRN are to:

- OBJ-1: Validate satellite SST products to FRM standards
- OBJ-2: Maintain and evolve the International SST FRM Radiometer Network (ships4sst) and deploy on a continuous basis Thermal Infrared Radiometers (TIR) and necessary supporting instrumentation to validate SST products.
- OBJ-3: Process, archive and quality control ships4sst data following documented FRM procedures that approve their use for FRM satellite validation.
- OBJ-4: Deliver approved ships4sst data sets and uncertainty budgets to users.
- OBJ-5: Prepare and submit peer-reviewed journal articles.
- OBJ-6: Conduct communications and outreach material promoting ISFRN activity.

The ships4sst project, which leads the ISFRN, has a number of tasks including:

Task 1 International collaboration: this covers all international developments and partnerships including the ISAR training (e.g. in Korea and China), L2R data upload to the website (e.g. from ISARs, SISTeR and M-AERI), potential collaborations through instrument building and loans and outreach on the webpage, Twitter (@ships4sst) and conferences. A number of papers are being written that should be published in 2021.

Task 2 covers data collection and archive. To date, the UoS have delivered data covering 12 deployments, DMI and RAL have 8 deployments. The COVID-19 pandemic has impacted 2020 deployments with ships being moored for a period of time and no access to some instruments. The data archive is located at ftp/ifremer.fr and can be accessed via the ships4sst website.

Task 3 covers data processing and validation. The Felyx MDB (Taberner et al, 2013) generation is done at Ifremer/EUMETSAT and processes SLSTR L1b and L2 data within 400 x 400 pixels of a match-up, and L2R *in situ* data within 6hrs of a match-up. The MDB analysis tool uses the approach as stated in Wimmer et al. 2012[27] and has recently been upgraded to Python 3. Reprocessed SST fields from 2016, 2017 and 2018 have been used whilst NRT data is currently being worked on.

So far, good results have been obtained in the validation of SLSTR against the shipborne radiometers, similar to those obtained in the past for AATSR (www.atsrsensors.org).

5.2 Status of the Data Archive

The ships4sst data archive has good geographical coverage for IR data. L2R SST_{skin} data on the archive covers 07/2004 to 12/2019 over 5 different datasets from UoS, DMI, CSIRO, RAL and RSMAS.

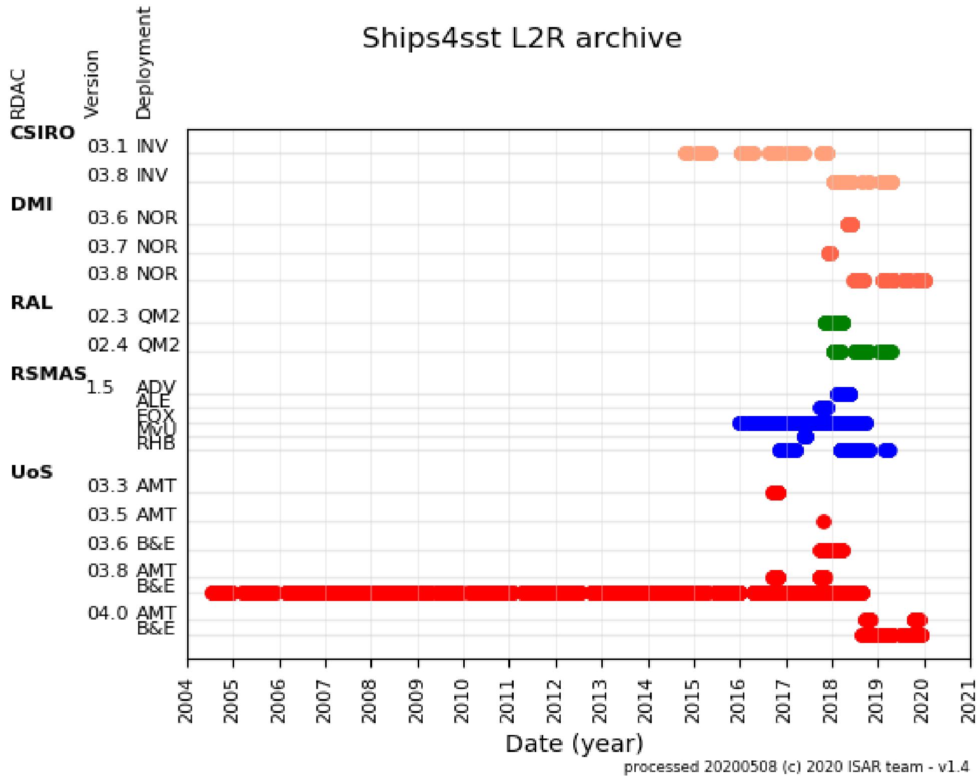

Figure 2: Archive data on the ships4sst archive split by instrument and version number

Issues with data files on the archive can include; wrong file name convention, missing compulsory data fields, new data not matching old data file number on replacement files provided and folder names generation if non alphanumeric characters in instrument name, removing bad quality data and data duplication. In most cases, an email is sent to users to help resolve any problems. There is also an FAQ document on the ships4sst website (user page) to help with the most common problems that users can encounter.

[27] Wimmer, W. et al (2012). Long-term validation of AATSR SST data products using shipborne radiometry in the Bay of Biscay and English Channel. Remote Sensing of Environment, 116, 17-31. DOI: 10.1016/j.rse.2011.03.022

A possible expansion of the ships4sst archive could include saildrone data (contact: Jorge Vazquez).

5.3 Next Generation *In Situ* Radiometer

Tim Nightingale introduced the case study for a next generation *in situ* radiometer and encouraged a discussion on the considerations and requirements from fellow SST_{skin} and radiometer scientists. Current considerations in deciding on the best future generation radiometer design include:

- manufacturability,
- ease of optical alignment,
- Maintainability, e.g. a self-calibrating or a nulling radiometer?
- ease of deployment,
- SST retrieval technique
- Type of measurements, i.e.
  - thermal infrared and/or microwave?
  - Imaging or non-imaging?
  - Single band / multiple bands / spectrometer? 3.7µm, 7.7µm, 10.8µm, or 13+ µm?
- electrical and mechanical interfaces.

All these considerations have pros and cons and many aspects to consider, so feedback and preferences from radiometer operators and SST users is encouraged. Instruments that were used in 2001 have not changed much since then and so it is a good time to plan a next generation radiometer.

Requirements include;

- Skin SST with systematic uncertainty < 100mK 1δ (50mK goal),
- clear traceability route for SST measurements,
- autonomous operation in a maritime environment,
- straightforward user and maintenance,
- compatibility with ISFRN / ships4sst network,
- Power supply (many measurements or some types of measurements need a large power supply which in turn means a heavy ship-bound instrument).

A case study document will be generated towards the end of the year and an additional telecon for anyone interested in discussing the specifics of a future generation radiometer will be convened by the end of October (an email will be sent out to workshop participants and via GHRRST).

6. Radiometer Performance and Uncertainties

6.1 Radiometer Uncertainty Models

FRM are required to determine the on-orbit uncertainty characteristics of satellite measurements via independent validation activities. In order to be classified as an FRM, not only are pre- and post-deployment calibrations required, but also a per-measurement uncertainty model. For ISAR, the model was developed on a first principle bases by analysing the components of the measurement equation (Figure 3), where the measurement equation is shown in yellow. R2T stands for radiation to temperature transformation, R_{sea} is the radiation from the sea, R_{sky} the radiation from the sky, ε the seawater emissivity, $R_{BB1,2}$ the radiation from the two on-board blackbodies, Sig_{Sea}, Sig_{Sky}, $Sig_{BB1,2}$ are the signals from the detector when viewing the sea, sky of the two blackbodies. The ISAR post processor, which was implemented following this model, produces an uncertainty value for each SST_{skin}. A detailed description of the uncertainty model can be found in Wimmer and Robinson (2016) [28].

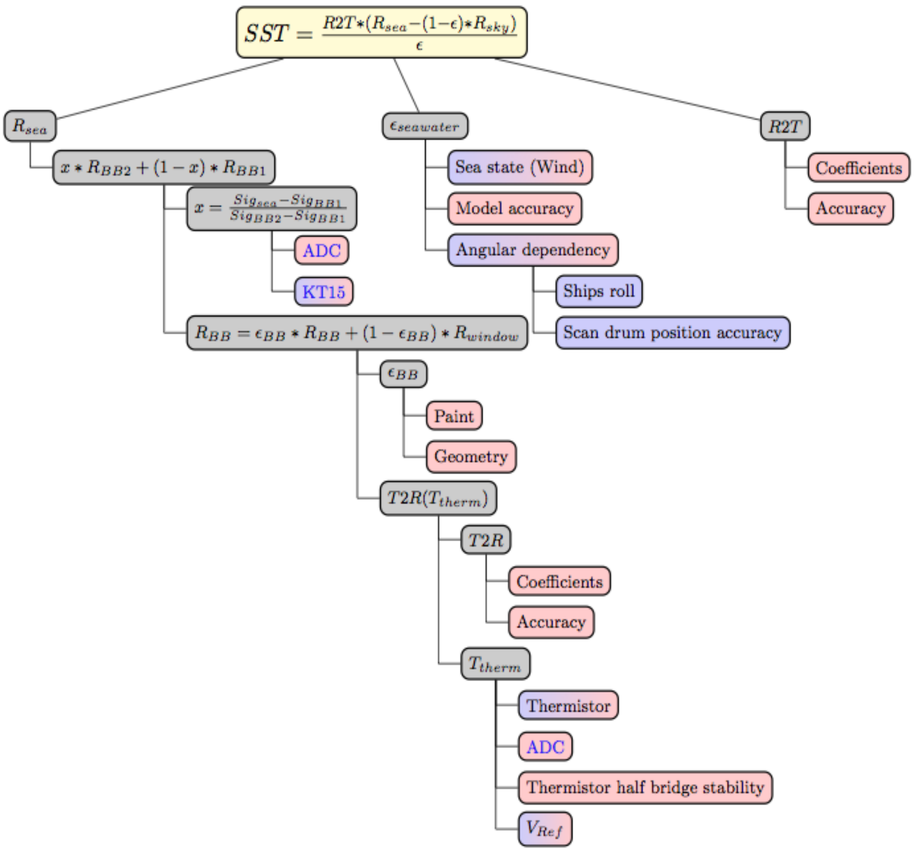

Figure 3: Schematic to illustrate the breakdown of the main elements of the ISAR SST_{skin} processor to reveal the factors that introduce uncertainty. For clarity the R_{sky} branch has not been expanded but is essentially the same as for R_{sea}. Boxes coloured in blue represent type A uncertainties, boxes

[28] Wimmer, Werenfrid & Robinson, Ian. (2016). The ISAR instrument uncertainty model. Journal of Atmospheric and Oceanic Technology. 33. 10.1175/JTECH-D-16-0096.1.

coloured in red show type B uncertainties, and boxes in red and blue contain both type A and type B uncertainties.

The biggest uncertainty for the self-calibrating chopped radiometers is emissivity because it is not measured. Instead, common models are used with knowledge of the view angle which ranges between 25º and 55º. As the ship moves the viewing angle varies and there is also wind dependence. Not having the correct emissivity in the equation can add an uncertainty of up to 50mK to the SST.

The magnitude of certain effects can be minimised using scientific knowledge of the variability in the upper ocean temperature, i.e. SST_{skin} should be retrieved from IR radiometers and the physics of the upper ocean used to compare to reference data at different depths. As the satellite uncertainty increases, the measurement of discrepancy also increases, which essentially means that both the uncertainty model and the validation of uncertainty model are correct, i.e. they are self-consistent.

An uncertainty inter-comparison was done on the *QM2* vessel between ISAR and SISTeR – results are to be analysed and provided with updates to the SISTeR record.

Uncertainty results from an Atlantic Meridional Transect (AMT) 28 and 29 cruises showed that the uncertainty was generally being overestimated, due to the roll dependence of emissivity. Two ISARs were used on these cruises to compare results and the overestimated uncertainty was confirmed. W. Wimmer has started to address this by looking at filtering the emissivity measurements. The higher uncertainties seem to be about right now but work still needs to be done on the lower uncertainties and SST gradients are not well captured. A version 2 of the uncertainty model is currently in progress which is planned for release early next year, and verification on AMT29 data has begun. The SST has not changed in the latest version of the code as the newer version mainly updates the uncertainty.

6.2 Comparison with other *in situ* instruments

The way we consider validation now is in the form of an uncertainty budget. The total uncertainty is made up of 5 components; the satellite, reference, geophysical (surface), geophysical (depth) and geophysical (temporal). Assessment of the uncertainty of satellite measurements also involves comparison to a reference dataset, such as SST_{skin} from shipborne radiometers, subsurface temperatures from drifting buoys, near-surface measurements from Argo profiling floats, and from the Global Tropical Moored Buoy Array (GTMBA). However, validating satellite SST retrievals using reference datasets has many sources of error that cannot easily be corrected. They can, however, be minimised using knowledge of the variability in the upper ocean temperature and comparing SST_{skin} from IR radiometers with reference data at different depths. For example, the difference between daytime and night-time SST data shows a difference in temperature at low wind speeds due to the diurnal variability; the difference arises because the ocean is cooler when the morning satellite measurement is taken 2 hours before the *in situ* measurement and warmer when taken 2 hours later, and vice versa in the evening, during the night the difference is not great (there is also a slight delay in the overlap as it takes some time for the warmth to penetrate to depth).

Overall the results from the model look good. Comparing to independent skin measurements (e.g. radiometer data) you do see a daytime dependence in the NWP wind speed in the data, however, when you apply the FKC time adjustment to the results the wind speed results improve and no dependency is seen. Note that all of the radiometer data needs to be used together to be statistically significant.

Results also show that there may be radiometer measurement uncertainties that are greater than 0.1K, or that the uncertainty model is wrong. Further investigation by all parties is required to find out which it is and operators suspect that there are still components that are contributing to the total radiometer uncertainty that still need to be quantified (including the variability in SST at small scales which is picked up in radiometers).

A paper[29] showing the bridge between AATSR and SLSTR was possible because of the data provided by *in situ* SST reference datasets.

6.3 Plans for radiometer inter-comparison exercises

Plans for the next radiometer inter-comparison exercise are underway with a lessons learnt from the 2016 inter-comparisons document being written and planned to be circulated by the end of the year. The current plan is to do the next inter-comparison exercise in spring 2022, to allow time for participants to obtain funding for travel once the COVID-19 outbreak is over. It will likely take place at NPL for the lab comparisons (radiometer and blackbodies) and then at Wraysbury (near Heathrow Airport) the following week for the field measurements, with possibly wider experiments as for FRM4STS (Land, Ice etc). Unfortunately the docks at Southampton are not really suitable due to the surface being contaminated and disturbed with ships, debris etc.

Benefits from the 2016 inter-comparison workshop include increased confidence in the data, and the confirmation that the calibration of the ambient temperature and target temperature need to be very close, at least when internal blackbody references are introduced (demonstrated by Nicole Morgan for the Australian ISAR).

Many of the lessons learnt from the 2016 inter-comparison involved the field of view not being filled properly or the positioning and footprint of the blackbody apertures. Some blackbodies also had a few temperature issues including the reference BB which should be as spatially uniform as possible. The 30 minutes period that was allocated to participants to set up their instrument was not deemed long enough by many participants, however extending this may increase the cost of the exercise. At

[29] Merchant, C.J.; Block, T.; Corlett, G.K.; Embury, O.; Mittaz, J.P.D.; Mollard, J.D.P. Harmonization of Space-Borne Infra-Red Sensors Measuring Sea Surface Temperature. *Remote Sens.* **2020**, *12*, 1048.

Wraysbury, the field of view and positioning of the radiometers was again an issue, as was using the same emissivity for each participant. The surrounding environment could be measured to a higher degree in 2022 than that was done in 2016.

The immediate plan is to send out a poll for interested participants to note when they would be likely to take part in another CEOS FRM TIR inter-comparison. More bi-lateral at sea or inter-comparisons in ‘at sea’ conditions that concentrate on radiometer uncertainties have also been requested.

7. Conclusion

There is a clear need for reliable and accurate SST_{skin} measurements for referencing and validating satellite SST retrievals to FRM standards, whether for gap bridging between satellite deployments or to be used in understanding the state of our oceans, for example, in calculating flux in the air-sea boundary and/or modelling the global carbon sink and its effects all over the globe. Work is taking place to gather SST measurements to FRM standards all over the globe. The COVID-19 pandemic has taken a toll on the amount of shipborne radiometer data gathered in 2020, but most deployments are now underway again.

The *in situ* SST measurement field is at a stage of exciting technological progression. The R/V Falkor cruise in 2019 showcased the abilities of autonomous aircraft in discovering notable scientific oceanic anomalies quickly and efficiently and enabling further investigation with shipborne equipment. Saildrone measurements have become increasingly more common over the last few years with ongoing improvements in measurements and designs. And the case for a next generation shipborne radiometer is being developed.

Those members of the ISFRN deploying shipboard radiometers have established protocols, best practices and a recommended data format that is now used by three instruments types (the M-AERI, ISAR and SISTeR). As shipborne radiometers provide a traceability route for satellite SST_{skin} retrievals they are a reference for generating CDRs from satellite SST measurements. The ISFRN has helped to develop and take these practices forward. The Network is also interested in historical records. Further international collaboration is expected with Korea now that ISAR training has taken place, and with South Africa regarding instrument loan. A number of ISARs are being built with an expected delivery date in 2020 and 2021 which should help in creating more and continuous data from shipborne radiometer deployments which, at the moment show statistics that are generally noisier than for other primary *in situ* types due to the reduced number and consequential distribution of shipborne radiometers.

There are a number of areas that have been identified or recommended for further work within the ISFRN community. This includes:

- Improvement in the uncertainty models across all instruments but notably in shipborne radiometers. More bi-lateral experiments would help further understand these uncertainties.
- A possible consolidation of ISFRN instrument data versions to simplify the archive and additional instrument data inputs such as saildrone data.
- A re-visit into the effect of surface emissivity on SST_{skin} measurements.
- Another field campaign investigating FRM TIR/MW for IST experiments (following on from the 2017 field campaign)
- Further understanding of the difference between the TIR SST_{skin} and $SST_{subskin}$ and how waves impact the dynamic temperature of the SST_{skin}, particularly with respect to the flux and ocean carbon sink.

- A specification of what is expected if a radiometer is taken into different environments, particularly in sub-zero climates, was requested. This could come in the form of a one page document with some requirements for a future generation radiometer based on the expected issues of instruments in different climates.
- A suggestion was made to revisit instrument user manuals; there were a few noted occasions where an instrument was not able to work during part or all of a voyage.

With the changing climate and the impact our oceans have on the outcome of these climate changes, it is more important than ever to have reliable and accurate FRM to SI standards. Although there are still advances to be made, particularly with the clarification of uncertainties, shipborne radiometers are able to provide the level of accuracy required for a CDR.

The ISFRN Workshop brought together a number of experts in the radiometry field to present and discuss the latest results in shipborne radiometry and other *in situ* methods such as saildrone and buoys. The latest satellite SST validation activities were discussed and scientists showed how *in situ* SST_{skin} data was being used to research into ocean dynamics. It is encouraging to see the developments within and outside the ISFRN and the international collaborations that have developed over the years. Whilst this report has only summarised the key information from the workshop presentations and discussions, it is clear that shipborne radiometry and *in situ* SST measurement instruments in general are gaining strength and recognition for the consistency, stability and usefulness of the measurements in validating satellite data from instruments including AATSR and SLSTR, and helping scientists understand ocean dynamics and the impacts of climate change.

The presentations, protocols, procedures and reports are all available on the ships4sst website at www.ships4sst.org/documents.

8. Acknowledgements

The FRM4SST project would like to thank and acknowledge the substantial contribution of all the participants and their funding agencies in support of the ISFRN workshop and this review document. In particular, the European Space Agency (ESA), whose funding made the ISFRN workshop possible, and Space Connexions Limited, for hosting the workshop via WebEx. The participation of the group at RSMAS, University of Miami, is supported by funding from the NASA Physical Oceanography Program and the NASA US Participating Investigator Program.

The author acknowledges the work and information provided by all the presenters at the ISFRN workshop, whose data, figures and information have been included in this paper. The authors would also like to thank those attending for their contributions to the service review.

Appendix

Agenda

Thursday 17th September 2020 (1400 – 1730 (UK), 0900 – 1130 (US EDT))

1400 – 1410	Welcome addresses	Craig Donlon, ESA
1410 – 1455	***Experiences of radiometer operators***	
	ISAR UK	Werenfrid Wimmer, UoS, UK
	M-AERI	Peter Minnett, University of Miami, USA
	Skin Temperature Measurements for the Saildrone	Andy Jessup, University of Washington, USA
1455 - 1550	**Validation of satellite SST measurements**	
	Using M-AERI to validate MODIS and SLSTR SSTs	Peter Minnett, University of Miami, USA
	The evaluation of the in situ SST quality control applied in iQuam in NOAA. (Poster)	Haifeng Zhang, NCWCP, USA
	Validation work at Lake Tahoe and Salton Sea	Simon Hook, JPL, USA
	Validation of SST and SSS Gradients Using the Saildrone Baja and Gulf Stream Deployment	Jorge Vazquez, JPL/Caltech, USA
1550 - 1635	**SST Data in Practice**	
	A Virtual Ride of Discovery: Exploring the Earth's Climate using Drones	Christopher Zappa, Columbia University, USA
	The Sentinel-3 Mission	Anne O'Carroll, EUMETSAT, Germany
	The importance of radiometric/skin SST on air-sea fluxes.	Carol Anne Clayson, WHOI,USA
1635 – 1720	***The ISFRN Network***	
	Status of the ISFRN	Werenfrid Wimmer, UoS, UK
	Status of the data archive	Werenfrid Wimmer, UoS, UK
	Next generation In-situ radiometer	Tim Nightingale, STFC, UK
1720 - 1730	Closing remarks	Craig Donlon, ESA
1730	**Close of meeting**	

Friday 18th September (0800 – 1130 (UK) , 1500 – 1830 (Beijing), 1700 – 2030 (Melbourne))

0800 - 0810	Welcome Address	Craig Donlon, ESA
0810 – 0910	***Experiences of radiometer operators***	
	ISAR Australia	Nicole Morgan, CSIRO, Australia
	ISAR China	Prof. Lei Guan, Oceans University China, Qingdao
	ISAR Denmark	Jacob Hoyer, DMI, Denmark
	SISTeR	Tim Nightingale, STFC, UK
0910 – 0955	***Validation of satellite SST measurements***	
	Using ships4sst data to validate SLSTR data	Werenfrid Wimmer, UoS, UK
	SST CCI Validation	Owen Embury, University of Reading, UK
	SLSTR Matchup Database	Igor Tomazic, EUMETSAT, Germany
0955 - 1025	***SST Data in Practice***	
	Monitoring the Mediterranean Sea	Jordi Isern-Fontanet, CSIC, Spain
	Verifying the consolidated theory of atmosphere-ocean CO_2 fluxes and the importance of the skin (AMT4OceanSatFlux)	Jamie Shutler, University of Exeter, UK
1025 - 1110	***Radiometer performance and uncertainties***	
	Radiometer uncertainty models	Werenfrid Wimmer, UoS, UK
	Comparison with other *in situ* instruments	Gary Corlett, Eumetsat, Germany
	Plans for radiometer intercomparison exercises	Werenfrid Wimmer, UoS, UK
1110 - 1120	Closing remarks	Craig Donlon, ESA
1120	**Close of meeting**	

Workshop Participants

Name	Organisation	Country
Ruth Wilson	Space ConneXions Ltd	UK
Hugh Kelliher	Space ConneXions Ltd	UK
Werenfrid Wimmer	UoS	UK
Craig Donlon	ESA	Netherlands
Tim Nightingale	STFC RAL	UK
Jacob Hoyer	DMI	Denmark
Arrow Lee	STFC RAL	UK
Peter Minnett	University of Miami	USA
Nicole Morgan	CSIRO	Australia
Andy Jessup	University of Washington	USA
Owen Embury	UoR	UK
Chris Merchant	UoR	UK
Stéphane Saux Picart	Meteo France	France
Christopher Zappa	Columbia University	USA
Anne O'Carroll	Eumetsat	Germany
Lei Guan	Oceans University China	China
Gary Corlett	Eumetsat	Germany
Jordi Isern-Fontanet	CSIC	Spain
Jamie Shutler	University of Exeter	UK
Ian Ashton	University of Exeter	UK
Helen Beggs	BOM	Australia
Carol Anne Clayson	Woods Hole Oceanographic Institution	USA
Simon Hook	JPL-NASA	USA
Steffen Dransfeld	ESA	Italy
Steinar Eastwood	Met No	Norway
Kyung-Ae Park	Seoul National University	Korea
Rory Scarrott	University of Cork	Ireland
Michael Reynolds	RMRCO	USA
Silvia Scifoni	Serco	Italy
Jorge Vazquez	JPL/Caltech	USA
Ioanna Karagali	DTU Wind Energy	Denmark
Mohamed Hosny Hassan Abdelkader	Soils, Water and Environment Research Institute. Agricultural Research Center	Egypt
Sam Hunt	NPL	UK
Eimear Tuohy	Techworks Marine Ltd	Ireland
Hee-Young Kim	Seoul National University	Korea
Chunying Liu	NOAA ICOADS, NCEI, CCOG	USA
Igor Tomazic	Eumetsat	Germany
Sotirios Skarpalezos	DMI	Denmark
Alexey Kaplan	LDEO of Columbia University, NY	USA
Janice Sisson	BOM	Australia