

Customer	: Space ConneXions Ltd	Document Ref	: ISAR-DEFRA-FREP01
:	acting for Defra	Issue Date	: 10 March 2006
		Issue	: 1B

Title : **Validation of AATSR Sea Surface Temperature Products using the shipborne ISAR Radiometer - Final Report**

Abstract : This document presents to DEFRA the final report of the subcontract with Space ConneXions for AATSR validation using the ISAR radiometer on the Pride of Bilbao.

Author :

Ian Robinson, Werenfrid
Wimmer, Craig Donlon

Distribution : Hugh Kelliher, Space ConneXions Ltd
Dr Craig Donlon, UK Met Office
AATSR Validation Scientist, Dr. Gary Corlett, Uni. Leicester

© The Copyright of this document is the property of National Oceanography Centre, Southampton. It is supplied on the express terms that it be treated as confidential, and may not be copied, or disclosed, to any third party, unless authorised by NOCS in writing.

National Oceanography Centre, Southampton
European Way, Southampton, SO14 3ZH, UK
Tel: +44 (0)23 8059 3438
Fax: +44 (0)23 8059 3059

Contents

Executive Summary.....	4
1. Introduction.....	5
2. The ISAR System.....	7
2.1 The need for autonomous radiometers on ships of opportunity.....	7
2.2 The ISAR instrument.....	9
2.3 Operational deployment of ISAR.....	11
2.4 Calibration and validation of the ISAR instrument.....	13
2.5 Post deployment data processing.....	18
2.6 Experience with autonomous operation.....	18
3. Measurements made by the ISAR System.....	21
3.1 Data collected from ISAR deployments.....	21
3.2 The variability of skin SST in the Bay of Biscay.....	25
3.3 Match-ups between ISAR and AATSR data.....	27
4. Validation of AATSR against ISAR data.....	34
4.1 Comparisons between AATSR and ISAR.....	34
4.2 Discussion – significance of results.....	36
5. Conclusion.....	38
5.1 Achievements of the project.....	38
5.2 Requirement for ongoing work.....	38
Appendix A Data acquired by ISAR.....	40
A.1 Data records for Dep1: ISAR-002. 17 March - 20 April, 2004.....	40
A.2 Data records for Dep2: ISAR-002. 27 April - 15 June, 2004.....	40
A.3 Data records for Dep3: ISAR-003. 16 July - 16 Oct, 2004.....	41
A.4 Data records for Dep4: ISAR-002. 16 th Oct - 21 st Dec, 2004.....	42
A.5 Data records for Dep5: ISAR-003, 1 st Feb - 15 th March, 2005.....	43
A.6 Data records for Dep6: ISAR-003, 15 th March - 22 nd June, 2005.....	44
A.7 Data records for Dep7: ISAR-003, 22 nd June - 20 th Sep, 2005.....	45
A.8 Data records for Dep8: ISAR-003, 20 th Sep - 13 th Dec, 2005.....	46
A.9 Data records for Dep9: ISAR-003. 13 th Dec, 2005 - 6 th Jan, 2006.....	47
Appendix B Maps of observed skin SST.....	48
B.1 Ship track maps by Deployment.....	48
B.2 Latitude - date distribution of skin SST, by deployment.....	53
Appendix C ISAR-AATSR comparisons and maps of match-up locations.....	58
C.1 Dep 1.....	58
C.2 Dep 2.....	59
C.3 Dep 3.....	60
C.4 Dep 4.....	61
C.5 Dep 5.....	62
C.6 Dep 6.....	63
C.7 Dep 7.....	64
C.8 Dep 8.....	65
C.9 Dep 9.....	65
Appendix D Pre & post cruise validation data.....	66
D.1 Calibration plots for Dep1 and Dep2.....	66
D.2 Calibration plots for Dep3.....	69
D.3 Calibration plots for Dep4.....	71
D.4 Calibration plots for Dep5.....	73
D.5 Calibration plots for Dep6.....	75
D.6 Calibration plots for Dep7.....	77
D.7 Calibration plots for Dep8.....	79
D.8 Calibration plots for Dep9.....	81
Appendix E ISAR Instrument Service Logs.....	83
Appendix F Copies of publications.....	86

AMENDMENT RECORD SHEET FOR THIS DOCUMENT

ISSUE-draft	DATE	REASON
1a	20 Feb 2006	Initial draft issue by ISR for discussion with Space ConneXions
1b	10 Mar 2006	Completed draft including WW alterations and additional tables

Executive Summary

Since March 2004, the National Oceanography Centre, Southampton has undertaken a programme of work to measure sea surface temperature (SST) at the level of the surface skin, in order to provide a set of baseline data for validating the skin-SST products retrieved from the Advanced Along-Track Scanning Radiometer (AATSR), a sensor procured by Defra and flown on ESA's Envisat satellite. The particular role of this subcontract within the wider AATSR validation programme is to provide a match-up dataset that monitors the ongoing stability of the AATSR performance.

The work has been performed using a unique instrument, the Infrared SST Autonomous Radiometer (ISAR). An ISAR instrument is mounted on the upper bridge of the P & O ferry *Pride of Bilbao*, measuring skin-SST every few minutes when there is no precipitation. It therefore observes the temperature of the sea surface in exactly the same way as the AATSR does. Because it operates autonomously day and night, it can acquire *in situ* observations to match the satellite data whenever the AATSR passes over the ferry in cloud-free conditions. During the two-year project an ISAR was operating on the ferry for 608 days. Laboratory calibrations confirmed that the ISAR measures skin SST to within 0.1 K

During the project a total of 143 validation pairs of dual-view satellite retrievals and *in situ* observations were acquired, measured within 1 km and 30 minutes of each other, corresponding to 28 different overpasses. Using a less strict criterion of coincidence, within 20 km and 2 hours, 817 validation pairs were acquired from 76 overpasses. About twice as many matches were obtained for the AATSR nadir products. The comparisons showed that the ISAR measurements of skin SST were mostly within 0.1 K of the dual view SST products from AATSR, implying that AATSR is operating well within its specification. However, the 2-channel nadir-only SST products were biased more than 0.6 K higher than the ISAR, consistent with other validation activities that have prompted a revision of the AATSR nadir-only SST algorithms.

The successful two year operation of this validation exercise has demonstrated the effectiveness of the ISAR for acquiring validation match-ups of skin SST for comparison with satellite data products. It has proved that an autonomous radiometer on a ship of opportunity can reliably deliver high quality measurements. Lessons were learned from occasional instrument malfunctions that were detected and remedied, and we expect to improve reliability even further in future operation. The validation activity is continuing with ongoing deployments of ISAR on *Pride of Bilbao*. It is also planned to use the available data record to make further studies of the process of validation, determining the effect of regional variability of SST on the reliability of the matchups, and characterising the bulk-skin temperature differences.

1. Introduction

In March 2004, the National Oceanography Centre, Southampton (NOCS: formerly the Southampton Oceanography Centre) commenced a two-year programme of work to measure sea surface temperature (SST) at the ocean surface skin, in order to provide a set of baseline data for validating the skin-SST products retrieved from the Advanced Along-Track Scanning Radiometer (AATSR). The AATSR is an Earth Observation sensor procured by Defra, presently operating on ESA's Envisat satellite, in order to provide a global record of SST of the highest quality and with an absolute accuracy that leads the world. Its primary objective is to supply reliable evidence about the Earth's changing climate as revealed by decadal trends in SST.

One of the essentials for achieving the desired stability and accuracy of global SST measurements, in addition to the fundamental design of the AATSR instrument itself, is to be able to confirm the quality of its measurements by independent observations made at sea level. This requirement is met by the AATSR validation programme that is supported by Defra through the Data Exploitation Contract and coordinated by the AATSR Validation Scientist. The particular role of the NOCS subcontract within the wider AATSR validation programme is to provide a match-up dataset of coincident AATSR and ship-based radiometer measurements of the sea surface skin temperature. In order to be able to monitor the ongoing stability of the AATSR performance, the acquisition of matching *in situ* data must be sustained by repeatedly sampling transects across the same region of sea over the AATSR's lifetime.

This goal has been achieved by deploying the Infrared SST Autonomous Radiometer (ISAR), a unique instrument designed by Dr Craig Donlon to measure skin SST to within ± 0.1 K from ships of opportunity without operator supervision. Two ISARs constructed at NOCS were made available for the AATSR validation work. One of these at a time is mounted on the upper bridge of the P & O ferry *Pride of Bilbao*, which makes regular crossings of the Bay of Biscay and English Channel. The ISAR measures skin-SST every few minutes when there is no precipitation, observing the sea surface in exactly the same way as the AATSR does. Because it operates autonomously day and night, ISAR can acquire *in situ* observations to match the satellite data whenever the AATSR passes over the ferry in cloud-free conditions.

During the two-year project several hundred matched pairs of dual-view satellite retrievals and *in situ* observations have been acquired for validation, along with a suite of ancillary measurements including the SST measured by conventional (sub-surface) thermometry, meteorological parameters and radiative heat fluxes. The analysis of the matched pairs of *in situ* and AATSR-derived skin SST have shown close agreement with the dual-view SST retrievals from the AATSR, while confirming

the evidence from other validation experiments that the nadir-only SST retrievals displayed a bias of a few tenths of a Kelvin. Overall, the ISAR system has proved its ability to deliver regular match-ups against satellite data throughout the year, at a rate which far exceeds the occasional measurements hitherto made by manned radiometers on research cruises.

This document provides a final report summarising the activities and achievements of the first, two-year, phase of the ISAR-AATSR validation contract. Section 2 describes the ISAR measurement system in more detail and discusses how it has performed during nearly two years of autonomous deployments. Section 3 presents the results that have been obtained, considering both the total observational dataset of all the measurements made by the ISAR system including ancillary sensors, and the number of matches made with coincident AATSR data. Section 4 examines what the matches reveal about the quality of the AATSR performance during 2004 and 2005. Finally Section 5 draws broad conclusions about what the project has achieved, and points towards the ongoing work that has been recently approved to continue the AATSR validation activities for a further two years. The Appendices contain detailed tables defining the data acquired during the project, maps of match-up locations and figures displaying the data graphically.

It should be noted that the scope of this report is limited to the AATSR validations made using the ISAR observations. A more complete view of AATSR validation during the same period, based on *in situ* data from a number of sources including ISAR, can be found in [make formal reference to Gary's final report]

2. The ISAR System

2.1 The need for autonomous radiometers on ships of opportunity

Before describing the ISAR instrument and the operating system developed for its deployment, it is important to establish the significance of the two requirements that dictated its design; the ability to measure the skin temperature of the ocean and the capacity for autonomous operation. Both of these requirements are inherent in the design of an effective validation programme for a satellite sensor intended to deliver skin SST data products specified to an accuracy of better than ± 0.3 K.

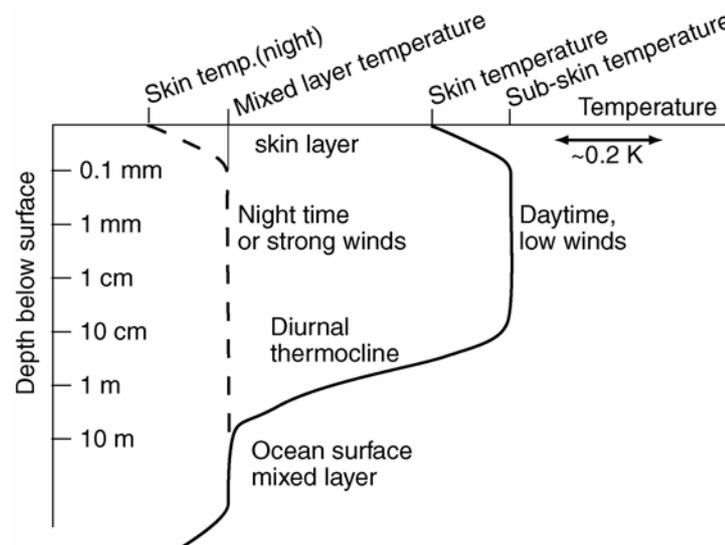


Figure 2-1 Schematic of the typical thermal structure near the surface of the ocean. The dashed line corresponds to conditions at night-time or of low insolation and strong winds. The solid line represents conditions of strong solar heating and low winds.

The need for in situ observations of skin SST can be appreciated by reference to Figure 2-1 which shows the typical structure of temperature in the top few metres of the sea. This shows two phenomena which cause the skin temperature of the ocean to differ from the temperature sampled by thermometers on buoys or ships which normally measure at a depth of between 1 and 8 m. Firstly under all conditions the surface microlayer, the upper 10 - 20 μm of water which determines the temperature detected by an infrared radiometer, is cooler by approximately 0.15 - 0.2 K compared to the sub-skin temperatures within 1 cm of the surface. This is a consequence of the suppression of turbulence by the air-sea interface and the net flow of heat out of the sea through the surface. This is true even under strong insolation since most solar flux penetrates deeper than this layer and delivers heat to the sea below the surface. Secondly solar heating tends to raise the temperature of the upper few metres during the day and, if the wind is too weak to mix this heat deeper, a diurnal thermocline

develops. In this case the water temperature at 1 cm depth may be up to 1 or 2 K warmer than that measured at 5 or 10 m. Unlike the skin temperature deviation, which is believed to be fairly predictable within 0.1 K, the diurnal thermocline can vary rapidly over hundreds of metres and within tens of minutes, depending on the wind and cloud variability.

Both these effects, and particularly the latter, introduce considerable uncertainty when relating the skin temperature as measured by an infrared radiometer on a satellite to the *in situ* temperature measured by sensors on a buoy or the hull of a ship. This is clearly unsatisfactory if the *in situ* measurements are used to validate satellite skin SST products. In that case they are not comparing like with like and such validations must be considered spurious unless there is a strong reason to expect that there is no diurnal thermocline present, as at night or under very strong winds. The most insidious problem is that under moderate winds there may still be a warming of a few tenths of a Kelvin that is not obvious but is nonetheless unacceptable when *in situ* data for validation are required to estimate the skin temperature to an accuracy of 0.1 K, in order that the AATSR products specification of better than ± 0.3 K can be tested.

This explains the essential requirement for *in situ* measurements of the skin temperature, which can be achieved only by making radiometric measurements of the sea similar to those made from the satellite. Because radiometry from ships is difficult and there is no long heritage of skin temperature measurements in contrast to a century of "bucket" measurements of the SST at a depth of a metre or more, there are very few skin SST observations available for the AATSR validation programme. Although there are one or two high quality radiometers available they have required close operator supervision, which makes them too expensive to operate apart from occasional research cruises, with a disappointingly low return of AATSR match-up points because cloud cover prevents matches from many potential coincidences between ship and satellite. There is one other programme of continuous high quality skin temperature monitoring that uses the M-AERI run by RSMAS, Miami, which makes data available for AATSR validation, but these measurements are mainly from tropical waters. To meet the need for continuous long-term monitoring of skin SST in temperate latitudes, at a reasonable cost, it became evident some years ago that ship-borne radiometers needed to be designed to be deployed autonomously on ships of opportunity. In this way they would be acquiring data at sea most of the time and thus able to add to the AATSR match-up database whenever the satellite passed over the ship in cloud-free conditions. This is what motivated the design and prototyping work which led to two ISAR instruments being provided for the work described in this report.

2.2 The ISAR instrument

Ship-borne radiometry for measuring the skin SST is a very recent technology. The radiometric measurement of temperature to within ± 0.1 K is just as challenging at sea level as it is from space. Because the internally calibrated optical components of a radiometer must be identical to those in the target-viewing optical path, it is not possible to protect the fore-optics behind a protective front window. The marine environment is much less benign than that in space, and shipborne radiometers must be designed with great care if they are to maintain their calibrated accuracy for several weeks at a time in all weathers, including rain and salt spray. A primary objective in designing the ISAR was to provide a weather-proof shutter that closes as soon as any precipitation is detected, in order to protect the critical optical elements such as mirrors and black body surfaces from exposure to rain and sea spray.

Nonetheless, the design philosophy also acknowledged that autonomous operation cannot offer complete protection in the most extreme conditions, and so components should be able to withstand occasional ingress of salt water past the storm shutter without catastrophic loss of radiometric performance. Because some degradation is inevitable, and because rigorous pre- and post-deployment calibration tests are an essential part of the operating regime, continuous operation throughout the year requires the use of two instruments, switched every two to three months for calibration, maintenance and (when necessary) refurbishment of critical optical parts.

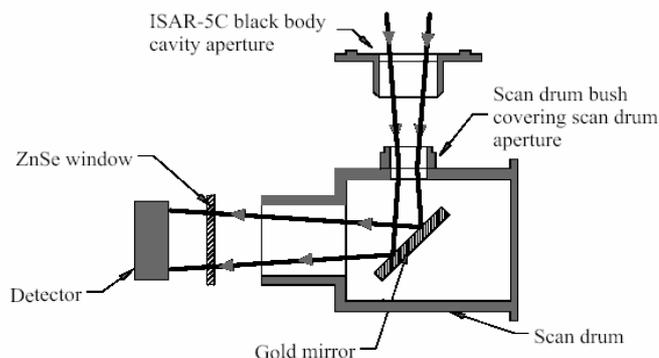


Figure 2-2: The fore-optics components of the ISAR, consisting of a ZnSe window through the weatherproof housing and the scan mirror inside its drum housing, shown here pointing into one of ISAR's black bodies.

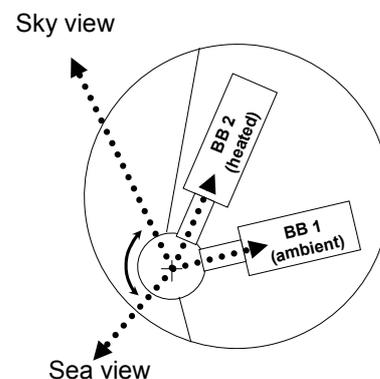


Figure 2-3: Section through ISAR and the scan drum, showing the arrangement for viewing the black bodies, the sea and the sky.

The ISAR instrument consists of several subsystems described in outline as follows. The **fore-optics system** which routes the optical path from different targets to the detector is shown in Figure 2-2. The scan mirror is a gold 3 mm glass substrate front-surface mirror housed in a rotating protective scan drum driven by a computer controlled shaft encoder. Figure 2-3 shows how the scan mirror points the field of view successively at the different targets (sea, sky and both black bodies), viewed via

a small aperture cut into the scan drum wall that is the only place that water may enter into the instrument. The optical path includes a plane ZnSe window set deep within the ISAR instrument that completely seals the detector and instrument electronics from the external environment. It is coated for a high infrared transmission of >90%. The scan drum is designed so that when the radiometer view field points outwards the black body cavities are closed, and when it points towards one of the black bodies, the scan drum seals the black bodies from ingress of water or salt spray.

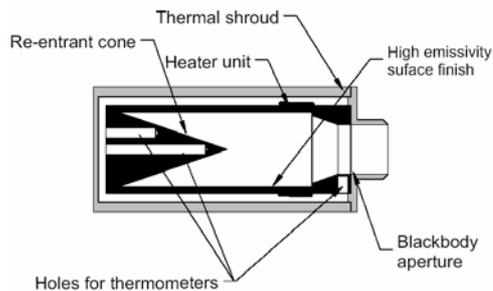


Figure 2-4: Section through one of the ISAR black body cavities



Figure 2-5: External view of the ISAR showing the shutter open. The optical rain sensor is in front

The **detector and black body calibration system** consists of a Heitronics model KT15.85D which delivers an analogue output dependent on incident radiation within a spectral band-pass of 9.6-11.5 μm over a temperature range of 173- 373 K. During an operating cycle it views sea, sky and two calibration black bodies housed within the ISAR. The black bodies (see Figure 2-4) were designed with a re-entrant cone base and a partially closed aperture to have an infrared emissivity > 0.999. Three precision thermistors (having a NIST traceable calibration to ± 0.05 K) are used to monitor the temperature of each black body.

The **environmental protection subsystem** incorporates a storm shutter and an optical rain detector. Whenever the rain detector output rises above a threshold indicating precipitation (or dust particles) the scan drum immediately points inside the ISAR to protect the fore-optics and an external shutter rotates circumferentially to cover the 150° viewing port (shown open in Figure 2-5). When the rain detector output falls below the threshold for a sufficient period (normally 10 min) the shutter re-opens and monitoring resumes.

The **internal control and data acquisition system** is an on-board computer system that manages the viewing cycle, controls the shutter operation, performs the analogue-to-digital conversions for the radiometer and thermistor outputs and logs the data. It also monitors a number of other variables logged in the ISAR, such as

GPS location and time which uniquely identify every data record, pitch and roll, power supply voltages and the ambient temperature inside the ISAR. It operates the measurement cycle which, for this validation work, is set at 40 samples viewing the sea, 30 samples each viewing the two black bodies and 10 samples viewing the sky, a cycle which takes about 140 s to complete. It uses the black-body views to calibrate the detector and then uses the sea and sky views to calculate the skin temperature of the sea, making allowance for sky reflection and the non-blackness of the emission from the sea. It logs the average values for each scan cycle on internal flash memory and sends the full resolution data to an external logging computer. However, it is capable of autonomous operation independently of the external computer.

The **external interface** uses RS232 protocol to communicate with the external logging computer. An RS485 interface also allows for ancillary atmospheric and ocean monitoring sensors mounted on the vessel to be powered and to submit their data into the ISAR's internal data logging system.

The subsystem components described above are housed in a compact (570 x 220 mm) cylinder shown in Figure 2-5. Figure 2-6 shows the ISAR in its operating position, viewing the sea surface at an incidence angle of 25° when mounted on the bridge wing of the P&O vessel *Pride of Bilbao*.



Figure 2-6 The ISAR and rain detector mounted on the *Pride of Bilbao*.

2.3 Operational deployment of ISAR

The ISAR is installed on the P & O Ferry *Pride of Bilbao*, which sails regularly from Portsmouth to Bilbao and back, crossing the English Channel and the Bay of Biscay.

One of the two available ISAR instruments, ISAR-002 or ISAR-003, is mounted on the top of the starboard Bridge Deck, as shown in Figure 2-7. The two instruments are periodically exchanged, ensuring that a recently calibrated instrument with clean optical components is delivering the required skin temperature data. Removing the instrument after 2-3 months allows post deployment calibrations of the temperature retrievals to be made, following which the instrument is fully serviced, including replacement of any components that have degraded. The *Pride of Bilbao* is scheduled to be at sea most of the time with only brief turnarounds in port, thus maximising the opportunities for match-ups with AATSR, except when the vessel is laid up for refit and maintenance for a few weeks in winter.



Figure 2-7. The P & O Vessel, *Pride of Bilbao* viewed from the port side. The arrow marks the location, on the starboard side, where the ISAR is installed.

There is no need for an operator to travel with the vessel. The ISAR is set running prior to a deployment and then operates autonomously. It is connected by the RS485 interface to an external computer located in the radio room of the *Pride of Bilbao*. ISAR can be manually controlled through this interface, but its continued autonomous operation is independent of the status of the external computer. The logging computer receives the full resolution data stream from the ISAR and any ancillary instruments feeding their data into the ISAR. In addition it collects data from other ancillary sensors, such as the hull mounted thermometers, and logs these together with consistent time stamps across all inputs.

Every 1-2 weeks the vessel is visited when it docks in Portsmouth. Checks are made to ensure that the ISAR and all ancillary sensors are operating properly, and the full resolution datasets are downloaded from the logging computer. During the final weeks of Phase 1 of the project the logging computer was also used to package selected data into a format for experimental direct transmission to NOC every few hours by the Iridium satellite telecommunication system.

In addition to the ISAR measurements of sea surface skin temperature, the ISAR measurement system on the *Pride of Bilbao* records the following ancillary data, each with a date and time stamp from the same source:

- Wind speed and direction relative to the ship.
- Roll and pitch
- GPS position, GPS time and ship velocity derived from GPS
- SST measurements by hull mounted thermometers
- Atmospheric humidity
- Long-wave downward radiation
- Total downward radiation
- Solar azimuth and zenith angle

Further variables such as sea state, wave height and percentage cloud cover can be obtained from UK Met Office as VOSCLIM observations every 6-12 hours, because the *Pride of Bilbao* is now instrumented by Met Office to VOSCLIM standard.

While these ancillary measurements are not directly required for the validation of AATSR data, they are still considered to be essential observations needed to benefit fully from the unique skin SST data delivered by ISAR. For example the comparison between the ISAR skin SST and the SST retrieved at depth from the hull thermometers serves two functions. Broad similarities between them as they move from the warm waters off Spain to the cooler English Channel provides confidence in the capacity of the ISAR to track the basic thermal structure of the sea. On the other hand a detailed study of the absolute differences between skin and depth SSTs will yield new insights into the size of the skin temperature deviation and the occurrence of diurnal warming. This is where the atmospheric ancillary data are important, and should eventually allow us to be able to predict the near surface thermal structure of the sea from the measured meteorological variables - something of importance when interpreting the global SST maps produced from AATSR.

2.4 Calibration and validation of the ISAR instrument

2.4.1 Calibration facilities

Within the global satellite SST research community, there is a growing reliance on the AATSR to provide a “standard” SST product to which the products of other sensors can be compared. Such a reputation for stability and absolute accuracy is justified only as long as the validity of the AATSR products continues to be proven against *in situ* observations. If the ISAR is to provide some of the *in situ* observations by which the AATSR products are judged, the ISAR data in their turn must also be validated to

the same high standard, with traceable calibrations. The ISAR team at Southampton have made considerable efforts to establish laboratory calibration procedures which can test whether the ISAR is delivering radiometric measurements of temperature to within its design specification of ± 0.1 K.

A completely new laboratory black body was designed, constructed and commissioned on 25th June, 2004 (see Figure 2-8). The new black body design (by C. J. Donlon) is based on the National Institute of Standards and Technology (NIST) third generation black body source. It consists of a water bath insulated metal container and an anodized aluminum cavity which is painted with a defusing reflection paint. The water is circulated with a pump, which also acts as the heating source of the water bath. The water-bath temperature is monitored using a Hart Scientific 1504 resistance bridge (Serial:A1B256) with a Thermometrics ES255 100 Ω Platinum Resistance Thermometer (Serial:203). The PRT and resistance bridge were calibrated to NIST standards and have an accuracy of ± 20 mK.

This system has now been adopted as the standard calibration target for external calibration of the ISAR instruments. It is referred to as the CASOTS-2 black body and supercedes the previously used (CASOTS) black body (Donlon et al, 1999)¹. It has been cross calibrated against the Rutherford Appleton calibration black body, and intercalibration tests against a NIST traceable black body are to be performed in Miami in March 2006.



Figure 2-8. The CASOTS-2 Black Body used as the laboratory reference for calibrating or validating the ISAR before and after each ship deployment. **Left:** External view showing the black body cavity aperture and blanking plates. **Right:** Internal view of the water bath revealing the conical end of the cavity. The black object to the left is the pump for stirring.

¹ Donlon, I.S., T. Nightingale, L. Fielder, G. Fisher, D. Baldwin & I.S. Robinson, A low cost black body for the calibration of sea-going infrared radiometer systems. *J. Atmos. Ocean Technology*, 16 (9), 1183-1197, 1999.

2.4.2 Calibration procedures before and after each deployment

Before every deployment of an ISAR on *Pride of Bilbao*, a laboratory test calibration is performed for about 24 hours. This is repeated immediately following each deployment without changing any part of the optical path (i.e. no attempt is made to clean any part of the system until after the post-cruise calibration). A standard procedure has been developed in which the ISAR is placed in a jig which aligns its field of view with the axis of the CASOTS-2 black body cavity, at a distance of 6 cm. Initially the water bath is filled with cold water at a temperature just above the dew point. A powerful pump is used to vigorously stir the water in the thermally insulated container, which has the effect of gradually heating it by 10-15 K during the 24 hours of the test. If necessary the procedure can be repeated using a temperature-controlled room to maintain the ISAR at several different ambient temperatures.

Comparison is made between the independently measured temperature of the target, and the ISAR estimate based on its own internal calibration process. If a small bias is found between the two, a bias correction can be applied to the ISAR data. However, following minor modifications to ISAR's optical path in the early stage of the project and after careful analysis of the ISAR's internal calibration, modelled from first principles and including recommended self-heating corrections for thermistors, it has been found that no bias corrections are normally required by the pre-deployment tests. Fundamental to the absolute accuracy of the radiative temperatures recorded by ISAR are the thermistors which measure the temperature of the ambient black body, and these are specified to an accuracy of 0.05 K.

Although the optical path is likely to degrade during a three month deployment, this should be accommodated by the internal self-calibration procedure inherent in the ISAR's mode of operation. Following most of the deployments the laboratory tests with the CASOTS-2 black body have shown no change in the instrument's performance compared with the pre-deployment tests, at a level greater than the ± 0.1 K specification. Appendix D Figure 2-9 shows such a case from deployment (Dep7) of three months duration. The pre and post-deployment calibrations are in agreement to better than 0.07 K, across a wide range of target temperatures. The pre- and post-cruise calibrations are shown for all the deployments in Appendix D.

If there are difference between the pre- and post-deployment calibration runs which exceed ± 0.1 K, then the record of the internal calibration of the ISAR is examined over the whole deployment to determine the time history of the degradation. Normally this allows bias adjustments to be made within the required confidence interval. If there has been a serious degradation of the optical path leading to a significant change in sensitivity, it is still possible to retrieve useful data within the

required accuracy of ± 0.1 K from most of the record, although in this case some measurements (when there is a large difference between the sea temperature and the ambient instrument temperature) must be assigned error bars greater than ± 0.1 K and these are rejected for match-up purposes.

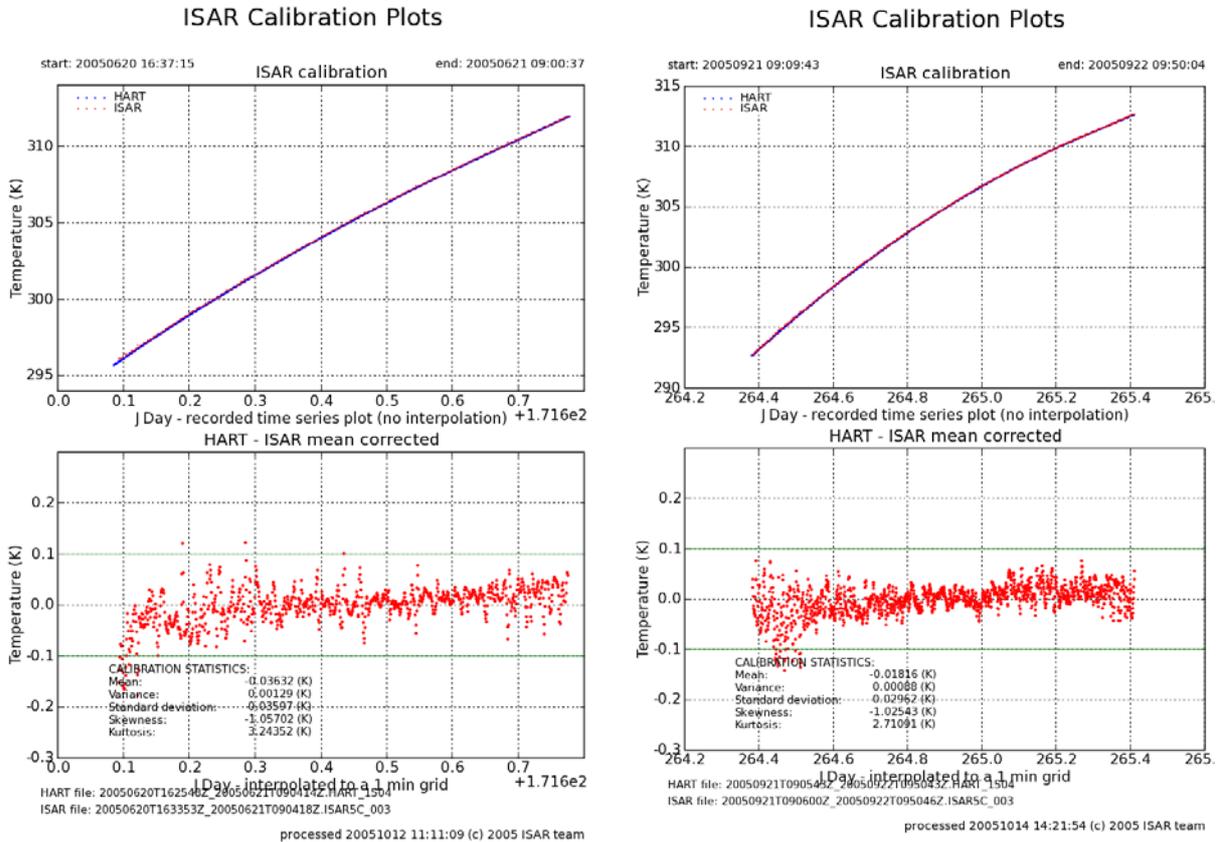


Figure 2-9 Calibration tests for Dep7, showing the comparison between ISAR measurements and the Hart PRT measurements of the CASOTS-2 Black body temperature. Top panels show PRT (blue) and ISAR (red) overplotted. Bottom panels show the ISAR-PRT difference. Left panels are pre deployment (20 June, 2005), right panels are post deployment (21 Sep, 2005).

2.4.3 Independent verification of the ISAR instrument

Two other ISARs (ISAR-001 and ISAR-004) have been built for our United States collaborators involved in a U.S. National Ocean Partnership Program (NOPP) to measure skin temperatures for validation of US satellite infrared sensors. The work is led by Prof Peter Minnett of RSMAS, Miami and data from the program are made available for validation of AATSR as reported by [Corlett \(2006\)](#). As part of that

activity the ISAR-001 was used in a comparison experiment between M-AERI², CIRIMS³ and ISAR. The three instruments were deployed together on the U.S. research vessel Ronald H. Brown between 31st Oct and 23 Nov 2004 on a research cruise through the Gulf of Mexico from Miami to Panama, and then across the Pacific Ocean from Panama to Santiago, Chile. The results of this intercomparison have now been reported by P. Minnett and are summarised in Table 2-1

Table 2-1 Radiometer inter-comparison (from 31.10.2004 to 23.11.2004 on-board *R/V Ronald H. Brown*). T_i is the SST measured by ISAR, T_c is the SST measured by CIRIMS, T_m is the SST measured by M-AERI and T_2 is a bulk temperature measurement at 2m depth.

	mean	stdev	min	max
$T_i - T_c$	-0.0062	0.1336	-0.6400	0.5200
$T_i - T_m$	-0.0656	0.1755	-0.8400	1.0100
$T_m - T_c$	0.0594	0.1871	-1.1500	1.1000
$T_i - T_2$	-0.1837	0.0971	-0.6400	0.3000

The table shows a very small difference of the measured SST between the three instruments. These results confirm that the ISAR agrees with the other radiometric sensors that are used operationally in USA, within very narrow tolerances (less than 0.1K). This indicates that the quality of the ISAR operational performance is world state-of-the-art.

Since each of the instruments compared differ significantly in critical aspects of their mode of operation and internal calibration, this intercomparison test also validates the basic design of the ISAR. ISAR-001 was the prototype for ISAR-002 and ISAR-003. It was built, initially tested and validated at NOCS and so there is a traceable comparison between the results in Table 2-1 and the performance of ISAR-002 and ISAR-003. Nonetheless, further international ship radiometer intercomparison tests,

²M-AERI (Marine-Atmosphere Emitted Radiance Interferometer) is the primary validation instrument for skin SST retrievals from MODIS (Moderate Resolution Imaging Radiometer). It has an extremely fine spectral resolution that resolves the fine microstructure of atmospheric absorption and its mode of operation in retrieving skin SST is different from that of ISAR or CIRIMS. Its accuracy is +/- 0.1 K, averaged over several minutes.

³CIRIMS (Calibrated InfraRed In situ Measurement System) is based on a standard infrared radiometer like the Heitronics used in ISAR, but it has a different approach to self calibration. It is also used in the NOPP Skin SST Demonstration Project.

involving ISAR-002 and/or ISAR-003 are planned for the future to ensure that all skin measurements from around the world are in agreement to within ± 0.1 K.

2.5 Post deployment data processing

During each deployment the ISAR generates its own estimate of the skin temperature by applying the gain factors derived from the local calibration constructed from viewing the internal black bodies. These are the data that are stored as average records. These are also the values that more recently have been telemetered to NOCS in near-real time via the Iridium link. In future it is intended that these values will be made available through the Project web page (www.isar.org.uk) in near-real time.

However, these must be considered to be no more than provisional values with no confidence value attached to them, until the post-deployment laboratory calibration tests have been performed using the CASOTS-2 black body. A suite of post-processing software has been created which applies the results of the laboratory calibrations to the raw data record to generate a skin SST record with known confidence values. When the final calibration run shows little change from the pre-deployment form, then the final data product will differ very little from the provisional values. However, if bias adjustment is required to match the external black body there will be differences. Moreover, if a serious discrepancy has been detected then careful analysis is required to determine how far into the run the calibration began to degrade. In this case, although skin SST can still be estimated for the whole deployment, some of the values will be assigned a lower confidence value indicating that they are not necessarily within the ± 0.1 K specification.

The post-processing suite operates on the database and is also able to compare the time and location of each ISAR record to the AATSR level-2 ATS_NR__2P data products. When coincidence is found (see 3.3.1) the data are entered into the match-up dataset, as long as the ISAR product is of the highest quality with an expected accuracy of better than 0.1 K.

2.6 Experience with autonomous operation

ISAR is the first, and as far as we know the only, infrared ship-borne radiometer to demonstrate that it can operate truly autonomously. It has proved to be capable of unattended operation for up to three months between servicing. As was expected we have encountered infrequent problems with the ISAR but only occasionally have they necessitated early switching of instruments and caused loss of data. Problems were

anticipated because a ship-based instrument costing about K£40 cannot be expected to offer the same reliability as a satellite sensor costing several orders of magnitude more. However, it was always foreseen in the operational management planning for the ISAR measurement system that some degradation of the sensor would occur in the harsh marine environment, necessitating the use of two instruments that are periodically switched. We have demonstrated that this is a sound and workable approach. Apart from an issue in the early stage of the project calling for a small redesign of the optical path on both instruments, we managed to keep one or other of the ISARs operating on the *Pride of Bilbao* continuously from mid July 2004 to early January 2006, apart from a gap in January 2005 during the vessel refit.

A lot has been learned from the deployments so far. For example we have made minor design changes to the optical path in order to reduce the risk of calibration uncertainties. We have isolated and remedied sources of additional noise in the electronics of the instrument. Most importantly we believe we have identified and eliminated most instrument control software problems and introduced additional checks and diagnostic features to reduce the down-time when hardware problems occur. We have also learned about ISAR's capacity to withstand the elements, and have made one or two minor design changes in this respect, as well as refining the configuration of the instrument management software, which drives the bad-weather closedown process, in order to respond appropriately to the information coming from the optical rain gauge. These issues are discussed in 3.1 in relation to their impact on skin SST data acquisition. More details of the minor modifications made to the instruments during the contract are provided in the Instruments service logs in Appendix E

The pre-and post-deployment laboratory calibrations have taught us two very important lessons. The first is that the basic design of ISAR is fundamentally robust and reliable, as evidenced by the examples of pre and post-deployment calibrations shown in 2-4. Secondly, when there has been a degradation of the optical path, we have discovered that the internal instrument data logged by the ISAR provide enough information to isolate when the problem has occurred and make it possible to assign differential error levels across the data record. Thus we have been able to produce useful observations within the design specification of $\pm 0.1K$ even from those deployments when a significant instrument degradation has occurred.

Nearly two years of operational deployments of the ISAR system during this contract have therefore given us confidence in its performance and its resilience in difficult environmental conditions. Based on our experience we are learning to spot incipient hardware problems as we monitor the internal temperature calibration of the radiometer. Now that ISAR data can be telemetered in nearly real time, the

calibration diagnostics can be analysed more readily. When they indicate a risk of degraded accuracy the instrument can be replaced at the next convenient port call, hopefully before the degradation has reduced their accuracy below the ± 0.1 K specification.

We are sharing our operational experience with the few other groups around the world who are engaged in ship-borne radiometry. A scientific paper describing the ISAR instrument is submitted and another describing its application to the validation of satellite data is about to be submitted. American colleagues are using the two other ISAR instruments, and it is likely that more will be built for other groups, all with the express purpose of validating satellite-derived SST data products. Building on the experience gained from the Defra contract, we intend to work to ensure that all those using ship radiometers follow consistent sound protocols and that the various radiometers are reliably inter-calibrated. This is essential if satellite data are to be calibrated with sufficient reliability for them to contribute to the accurate monitoring of global warming and climate change.

3. Measurements made by the ISAR System.

This chapter reports the measurements that have been obtained during the contract. The first part outlines all the data that were collected using the ISAR system on *Pride of Bilbao*. The second examines the information about skin SST that can be drawn from the ISAR measurements. The third part identifies the number of match-ups that were obtained between *in situ* skin temperature measurements from ISAR and coincident clear-sky SST retrievals from AATSR.

3.1 Data collected from ISAR deployments

The operational acquisition of data commenced on 16th March, 2004 when ISAR-002 was installed on *Pride of Bilbao*. Thereafter over 20 months of operational data were acquired in nine separate deployments, as illustrated schematically in Figure 3-1. Laboratory calibrations were performed before and after each deployment. The instruments were swapped after each deployment, apart from the gap between Dep1 and Dep2 when ISAR-002 was returned to the vessel after laboratory calibration during an AATSR outgassing period.

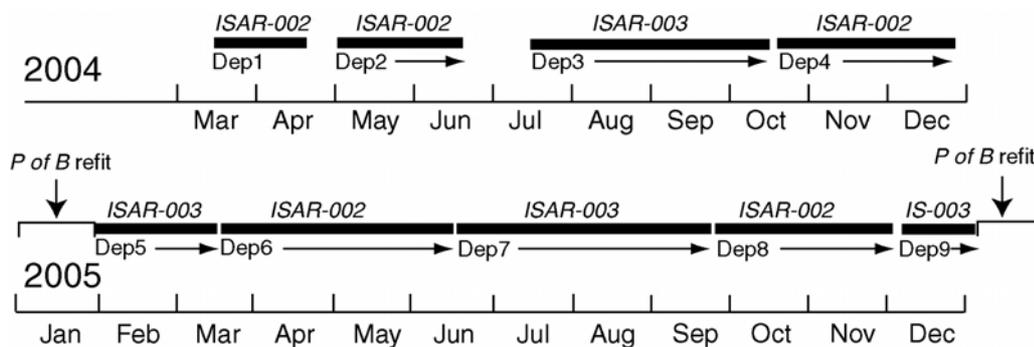


Figure 3-1 Time-line showing when the ISAR system was operating on *Pride of Bilbao* during 2004 and 2005, which of the two instruments was deployed, and the name (Dep1 etc.) assigned to each deployment.

The nominal route taken regularly by *Pride of Bilbao* is shown in Figure 3-2. During 2004, the vessel operated a seven-day cycle, in which it made two return trips between Portsmouth and Bilbao, and one return trip from Portsmouth to Cherbourg. Figure 3-3 presents a map of the skin SST measurements acquired over several weeks in summer 2004, revealing the slight variations in the actual ship track from cycle to cycle, and showing the extra tracks obtained in the English Channel from the Cherbourg run. Throughout 2005 the ferry schedule repeated a strict three-day cycle, purely between Portsmouth and Bilbao.

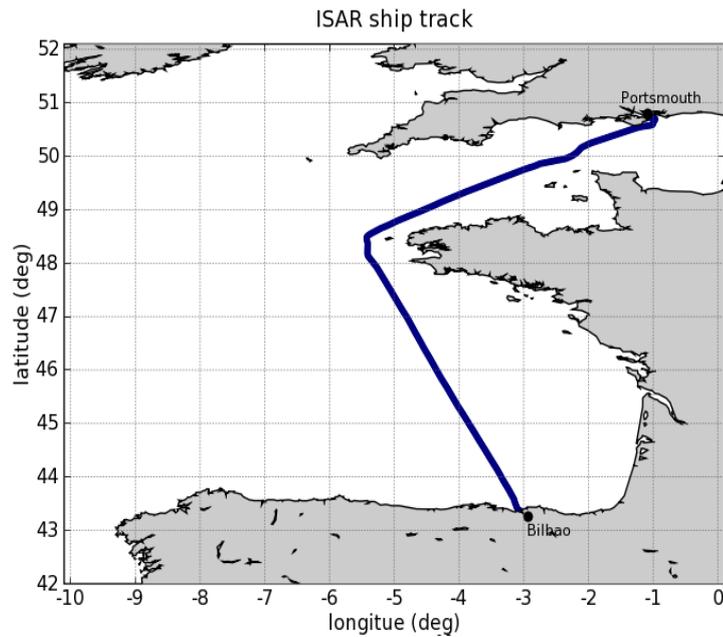


Figure 3-2. Nominal route of *Pride of Bilbao* on which ISAR is installed.

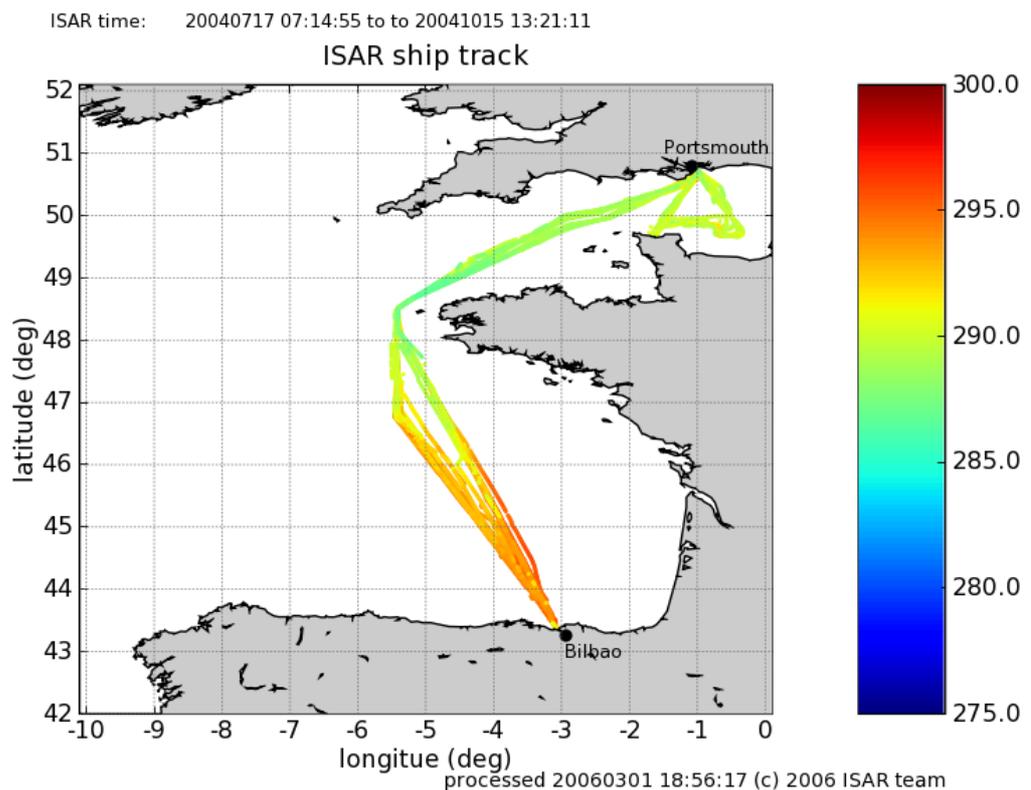


Figure 3-3 Ship track of the M/V *Pride of Bilbao* during the period the period 17th July 2004 to 15th October 2004 (Dep3). Each skin SST temperature measurement obtained by the ISAR is plotted as a dot, colour coded according to the scale below the figure.

A summary of the different deployments, defining their start and end dates, the ISAR instrument used, a list of the ancillary data available and brief notes of the quality of skin data acquired is provided in Table 3-1. ISAR records were acquired on 608 days

over the whole project. The maps of temperature along the ship track for all the deployments are shown in Appendix B.1. The full suite of ancillary instruments consists of:

- An Eppley Precision Infrared Radiometer which records the broadband downwelling incident longwave radiation;
- A Licor sensor for measuring the photosynthetically available radiation (PAR);
- A Kipp & Zonen CM11 pyranometer for measuring the shortwave downwelling incident radiation;
- A sonic anemometer (Gill windmaster 3-axis) for measuring wind speed and direction relative to the ship;
- A Vaisala HMP243 humidity sensor;
- A Seabird hull-mounted thermometer;
- A Minipack multiparameter sensor deployed by the Ferry-box project.

The full suite of sensors was installed on the Pride of Bilbao during the January 2005 refit, and some of the instruments were also deployed in 2004, shown in Table 3-1.

A full list of all the individual data files acquired for each deployment is given in Appendix A. These are held in the project database at NOC and securely backed up in archive storage. Access to the full database can be arranged by contacting W.Wimmer@noc.soton.ac.uk. These are the raw data files. As explained in 2.5, there is a provisional skin SST estimate available before the deployment has ended, which can be accessed by registered users via <http://www.isar.org>. When the post-processing has been performed the skin temperature products can be obtained from the ISAR web page (www.isar.org.uk) by registered users.

Table 3-1 Primary information about each deployment

No.	Dates	ISAR	Ancillary data	Notes
Dep1	17 Mar 04 - 20 Apr 04 (35 days)	002	Pyranometer, Anemometer Minipack	End planned to coincide with AATRS outgassing. Optical components in good condition. Pre- and post-cal showed bias. See notes for Dep 2
Dep2	27 Apr 04 - 15 June 04 (73 days)	002	Pyranometer, Anemometer Minipack	Jammed shutter prompted an early end of this deployment. Calibrations still showed bias. This was traced to the scan drum bush restricting the detector view to the internal BBs, requiring a design mod for both ISARs. Once the problem had been modelled into the calibration equation, it was possible to retrieve temperatures from Dep1 and Dep2 data to within the instrument spec.
Dep3	16 July 04 - 16 Oct 04 (93 days)	003	Pyranometer, Anemometer Hull temp. and Minipack to 17/08	Start delayed while mods made to scan drum bush (see Dep2). Successful deployment leading to a planned changeover. Very good pre-and post-cal agreement. Good condition of optics at the end.

Dep4	16 Oct 04 - 21 Dec 04 (66 days)	002	Pyranometer, Anemometer Hull temp. Minipack	Uneventful deployment. Ended when vessel went to refit. Good pre- and post-cal agreement. However, some very obvious degradation of the scan mirror raised concern about the consistency of these data.
Dep5	1 Feb 05 - 15 Mar 05 (42 days)	003	PIR, PAR , Pyranometer, Anemometer , Dew pt. temp. Hull temp. Minipack	Malfunction in operating software (OS.7.12) resolved on the ship (28.02.05) by upgrade to OS.7.13 . The deployment ended on 15.03.05 when excessive noise was observed in the record. This was traced to a failed voltage regulating chip.
Dep6	15 Mar 05 - 22 June 05 (99 days)	002	PIR, PAR , Pyranometer, Anemometer , Dew pt. temp. Hull temp. Minipack	Deployment ended after it was discovered that the scan drum mirror was damaged by salt water contamination. Having identified from the data record the event when degradation occurred, most of the data from Dep6 could be calibrated within spec.
Dep7	22 June 05 - 20 Sep 05 (90 days)	003	PIR, PAR , Pyranometer, Anemometer , Dew pt. temp. Hull temp. Minipack	No issues. Good agreement between pre and post cruise calibrations. All data of good quality.
Dep8	20 Sep 05 - 13 Dec 05 (85 days)	002	PIR, PAR , Pyranometer, Anemometer , Dew pt. temp. Hull temp. Minipack	Deployment ended after it was discovered that the scan drum mirror was degraded (manufacturing defect). Post cruise calibration allowed most of the record to be calibrated within spec. The rain detector threshold drifted during Dep8 causing excessive shutter closure.
Dep9	13 Dec 05 - 06 Jan 06 (25 days)	003	PIR, PAR , Pyranometer, Anemometer , Dew pt. temp. Hull temp. Minipack	No issues. Good agreement between pre and post cruise calibrations. Deployment ended when <i>Pride of Bilbao</i> entered refit
Dep10	21 Feb 06 -	002	PIR, PAR , Pyranometer, Anemometer , Dew pt. temp. Hull temp. Minipack	Ongoing at end of Phase 1 of the contract on 28 th Feb 06

The ISAR acquisition record summarised in Table 3-1 reveals quite long periods of uneventful operation punctuated by occasional problems. When an issue was detected that was serious enough to risk loss of AATSR matchup opportunities, either because optical degradation caused calibration accuracy to decrease out of specification, or because instrument malfunction prevented acquisition of any data at all, the deployment was terminated and the instrument was swapped. This occurred at the end of Dep2, Dep5, Dep6 and Dep8.

In each of these cases, the team convened an Instrument Performance Review which identified the problem and took action to minimise the risk of the same issue arising again. For example, the restricted view to the internal BBs became evident only during calibration against the newly built CASOTS-2 reference. It demanded a minor

redesign of the scan bushes on both instruments. The failure of an electronics component during Dep5 was at first obscured by a logging software problem which prevented us from examining the logged data. Once the excessive noise was observed the deployment was ended. That experience led to more stringent checks of instrument noise during subsequent deployments.

Visual inspection during routine checks prompted concern about the scan mirror and this led to the stopping of Dep6. In this case, most of the deployment record could be calibrated and few matchup opportunities were lost. On subsequent deployments the internal calibration factor was monitored for excessive variation, characteristic of a degrading mirror. This is the diagnostic test that prompted the premature curtailment of Dep8. A crack in the scan mirror was subsequently detected, and further probing discovered that a batch of spare mirrors was below the design specification and needed to be replaced by the manufacturer. When near-real time transmission of data was enabled during tests of the Iridium system, it became possible to monitor the detector noise and the internal calibration factor on a daily basis. As discussed in 2.6, we consider that the nine deployments of ISAR provided valuable experience for improving our operating practices, promising reduced risk of data losses in future, in addition to meeting their primary purpose of acquiring as large a record as possible of skin SST along the ship track, which is discussed in the next section.

3.2 The variability of skin SST in the Bay of Biscay

The temperatures measured by the ISAR when not closed for protection from precipitation and spray make up an impressive set of measurements defining the variability of the skin temperature of the Bay of Biscay and English Channel over nearly two years. Figure 3-4 shows a latitude-time plot of the complete ISAR skin temperature record. This shows the position along the cruise track in the x axis and the date-time in the y axis. The colours represent the skin temperature. A more detailed set of such plots are presented individually for each deployment in Appendix B.2.

Figure 3-4 can be used to illustrate the variety of surface temperatures encountered by ISAR over the two years of the project. The range extends between about 5 degC at the coldest to over 22 degC at the warmest. Such a spread of temperatures is essential for a baseline validation programme, since it allows the AATSR algorithms to be proven over that same range. Moreover, since most other skin SST observations for validation come from experiments in tropical and equatorial seas, there is a great need for the low temperature matched pairs which the measurements from this project are providing.

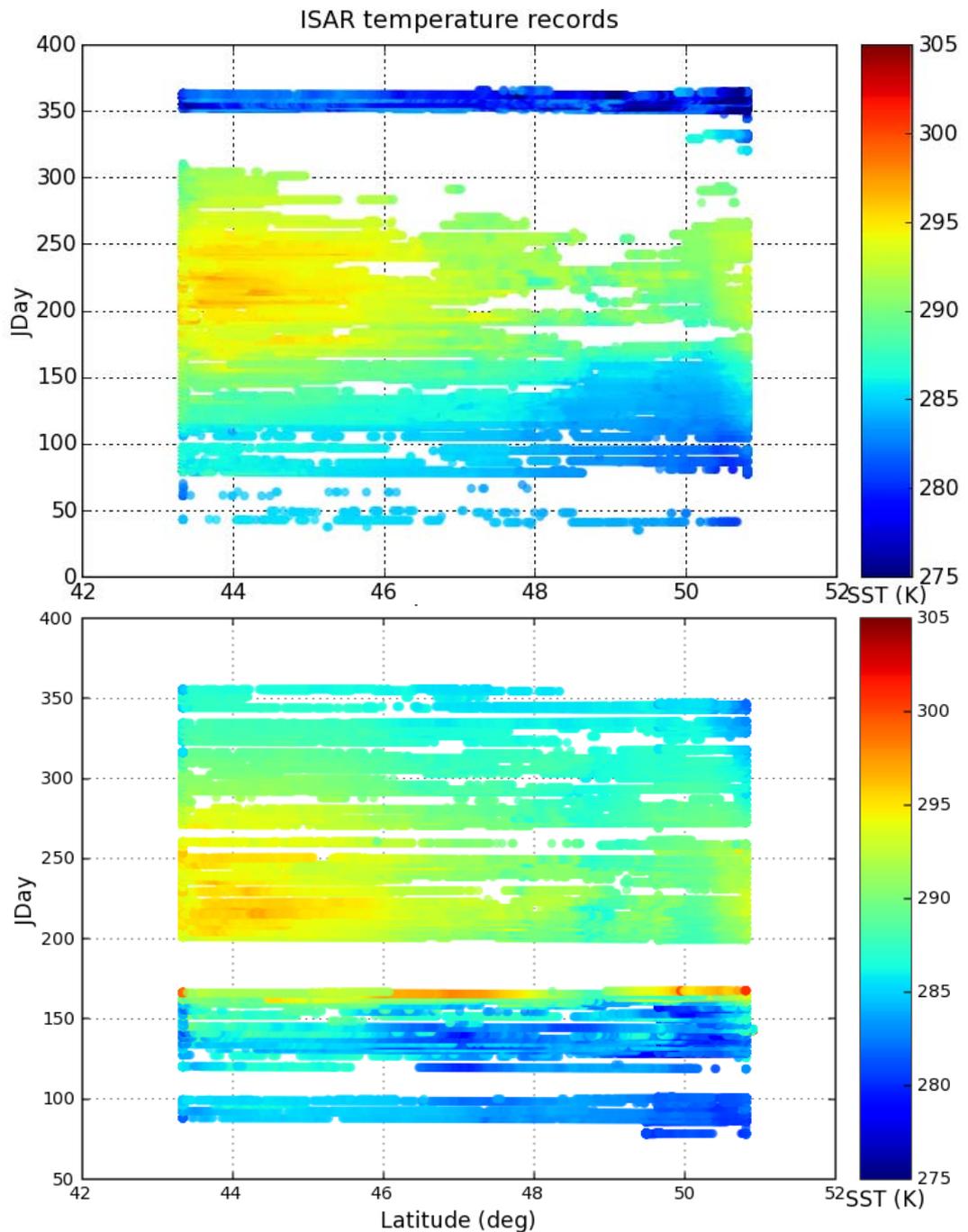


Figure 3-4 Latitude-time plot of the full project record of skin SST from March 2004 to December 2005. The lower panel is for 2004, the upper panel for 2005..

The expected seasonal trends are apparent in the latitude-time plots and it should be noted that the annual range of temperatures is greatest at the southern end of the cruise track close to Bilbao. Moreover, at most times of the year there is a south - to north decrease in temperature. While the ship track samples neither the warmest

temperatures of equatorial waters, nor the coldest polar waters, it nonetheless spans an excellent range of temperature values. While the most productive deployments (in terms of achieving high numbers of matches, see 3.3) have occurred in the warmer summer months, we expect to be able to obtain better match-up returns in the colder months as experience with deploying ISAR grows. It is unfortunate that the coldest month of January has been lost in both 2005 and 2006 because of vessel refits.

One other important factor to note in Figure 3-4 is the tendency for local spatial and temporal variability and patchiness to occur, especially in the warmer summer months. This may in part be due to the dynamic nature of part of the route, and strong seasonal and tidal stratification encountered in parts of the English Channel. However, it is almost certainly also a consequence of the diurnal warming and skin effects which make the skin temperature of the sea differ from what *in situ* contact thermometry in the water would measure. Since this patchiness is very hard to predict, it means that for validation purposes the skin SST observations provided by radiometers such as ISAR are essential.

3.3 Match-ups between ISAR and AATSR data

3.3.1 Basis for defining coincidence between ISAR records and AATSR pixels

Having secured the operational supply of *in situ* skin SST along the Pride of Bilbao cruise track, we were faced with the task of matching these data to the AATSR observations. The challenge is to decide how much flexibility is allowed in the definition of “coincidence” in space and time between ship and satellite samples.

The size of the match-up dataset depends on how stringent is the definition of data coincidence. If the requirement is too severe then there will be very few matches. If it is not severe enough then a large number of the matched pairs will be from quite different parts of the sea. Until we have the experience from which to identify the optimum matching criteria, we have adopted an approach which specifies four different grades of severity. These relate to the search radius, $N\Delta x$, for spatial coincidence and the time window, $M\Delta t$, for temporal matching, as illustrated in Figure 3-5.

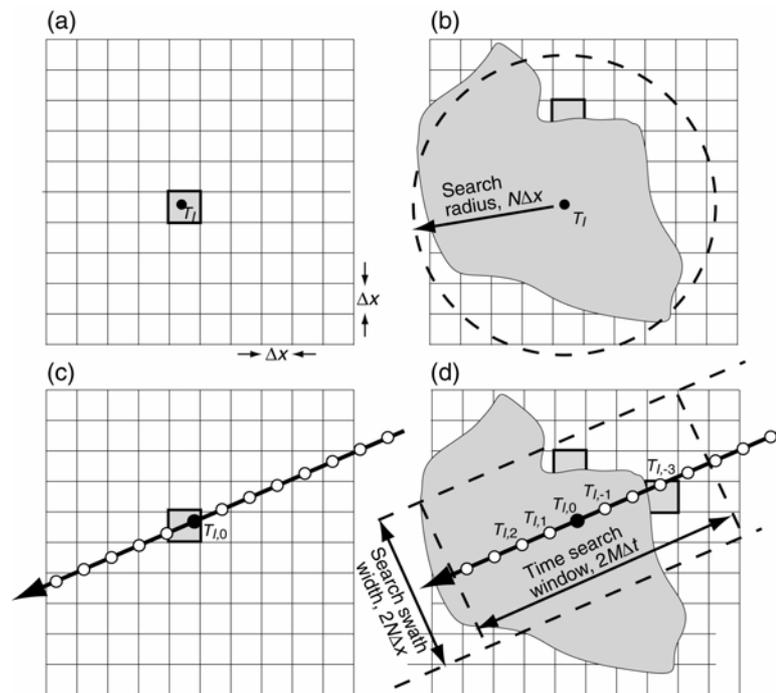


Figure 3-5. Examples of match-up situations encountered in the construction of a match-up database. (a) Point sample when there is no cloud. Match the in situ sample closest to the time of the overpass to the pixel in which it lies. (b) Point sample obscured by cloud. Match the in situ sample closest to the time of the overpass to the closest cloud-free pixel. The search radius needs to be limited to N pixels. (c) Along-track sensor such as ISAR in cloud free conditions. Match the in situ sample closest to the time of the overpass to the pixel in which it lies. (d) Along-track sensor in cloudy conditions.

We match the ISAR record to the corresponding AATSR image datasets according to a set of criteria graded as follows for severity::

<u>Grade</u>	<u>Match-up Criteria</u>
1	Coincidence of ISAR and AATSR sample within ± 2000 s time window and 1 km search radius in space.
2A	Temporal match within ± 2000 s and spatial match within ± 20 km
2B	Temporal match within ± 2 hrs and spatial match within ± 1 km
3	Temporal match within ± 2 hrs and spatial match within ± 20 km.

Grade 1 corresponds to the closest coincidence considered feasible, within 1 km spatially and about half an hour in time, and effectively yields a match to the pixel containing the ship at the time of the overpass.

Grade 2A allows a match to the nearest cloud free pixel to the ship track within a radius of 20 km.

Grade 2B limits the search radius to 1 km but allows a match to an earlier or later part of the ship track within a time window of 2 hours either side.

Grade 3 allows matches with both the wider spatial search radius and the extended time window. This corresponds to typical international practice for the calibration of other SST sensors against, for example, buoy temperatures.

This approach often leads to multiple matches for a given overpass. When this occurs the AATSR-ISAR matches are treated as independent as long as neither the same AATSR pixel nor the same ISAR sample are used for a different matched pair. If this should occur the closest in distance are selected. The matching algorithm rejects any matches found at locations corresponding to the ship being in port, or close to port, where it is unlikely that the ISAR and AATSR are properly viewing the sea surface temperature.

As far as possible, we have sought to produce a processing system which operates automatically without human intervention. Thus given a file(s) containing the calibrated record of an ISAR deployment, and access to the AATSR data files covering the same period of time as the deployment, the program must find the matched pairs for different grades of coincidence criteria, and then evaluate the mean and standard deviation of the differences for all matched pairs corresponding to a particular grade. There are two good reasons for using an automatic system. The first is that, once the system has been developed and validated, it significantly reduces the effort involved, leads to much more rapid matching of ISAR to AATSR data, and thus facilitates near-real time monitoring of AATSR performance as long as the ISAR data can be transmitted to base soon after acquisition. Secondly it avoids the need for subjective decisions to accept or reject match-ups, which might bias the matching and thus the validation process.

3.3.2 Numbers of matches achieved between ISAR and AATSR

Here we report the numbers of matches that have been achieved during the project so far. Discussion of how the ISAR temperatures compare with the AATSR SST products is left to section 4. Our concern is first to demonstrate that the use of the ISAR, deployed autonomously on a ship of opportunity, that spends most of the time at sea on a regular route, has delivered many more validation opportunities than were previously available from occasional dedicated validation campaigns.

The number of validation matchups varies not only with the grade of coincidence, but also with the type of AATSR data product being considered. There are four AATSR SST data products generated by different algorithms, labelled as follows:

- D2 The dual view 2-band algorithm using the 11 and 12 μm channels from both the nadir and forward view.
- D3 The dual view 3-band algorithm using the 11 and 12 μm channels from both the nadir and forward views.

- N2 The nadir-only 2-band algorithm using the 11 and 12 μm channels from only the nadir view.
- N3 The nadir-only 3-band algorithm using the 3.7, 11 and 12 μm channels from only the nadir view.

The 3-band algorithm can be used only at night when there is no problem of solar reflection in the 3.7 μm waveband. The 2-band algorithm can be used both day and night, although the standard AATSR products normally provide only 3-band products at night. There tend to be more matches with data produced by the nadir-only algorithms because in patchy cloud the forward views are more frequently flagged as cloudy than the nadir view, hence there are more nadir-only SST pixels than dual-view.

Table 3-2 summarises the number of matches achieved during the whole project. It is broken down by AATSR algorithm type, and the results are presented for each grade of severity of space-time coincidence. Over the full two years of the project for Grade 1, the strictest of coincidence criteria, 143 separate matches have been obtained over day and night for the dual view AATSR skin SST product. When the criterion is relaxed this expands to 827, while for the nadir-only AATSR data the corresponding figures are 248 for Grade 1 and 2048 for Grade 3. Moreover these numbers are well distributed between the products of 2 wavelength algorithms (day overpasses) and 3 channel (night overpasses).

These are extremely encouraging numbers of matches, given the scarcity of radiometrically measured *in situ* SST data available for validating the data products of ATSR, ATSR-2 and AATSR in its first two years. These results justify the approach of using an autonomous radiometer on a ship-of-opportunity, represent a successful outcome of the technical development effort that went into developing the ISAR concept before proven instruments could be made available to Defra through the present contract..

Table 3-2: Total numbers of match-ups during the whole project, between ISAR and AATSR skin SST measurements, for different satellite products and different grades of matchup coincidence.

AATSR product type	Grade 1	Grade 2A	Grade 2B	Grade 3
D3	76	99	274	429
D2	67	82	288	388
N3	111	263	402	915
N2	137	264	576	1133

The spatial distributions of the match-up locations can be plotted for each of the four grades of coincidence criteria. These are presented for each deployment in Appendix C. The locations of Grade 1 match-ups aggregated over all deployments is shown in Figure 3-6. They are spread along the cruise track, with a number in the centre of the Bay of Biscay in deep water up to 200 km offshore, representative of open ocean conditions. Here the SST tends to vary little in both space and time. Consequently the match-up criteria may be relaxed without compromising the ISAR-AATSR comparison too seriously. Figure 3-7 for the Grade 2A matches shows that the extra points gained by the wider spatial window tend to lie in clusters around the grade 1 points. However, by extending the time window for matches, as in Figure 3-8 for Grade 2B and Figure 3-9 for Grade 3, the extra points fill up between the clusters along the cruise track. The reason for the clustering of points in limited zones along the track in the case of the restricted time window of Grades 1 and 2A is to do with the ship's timetable being phase locked to the time of day, to which the AATSR overpasses are also approximately tied by the sun-synchronous satellite orbit.

There are a number of match-up points closer to the coast of Brittany off Ushant, and also in the central English Channel. In both these places SST tends to vary rapidly in space and time, depending on the tidal phase, and so larger validation mismatches may be expected for grades 2A, 2B and 3. This is discussed in the companion Final Report for the Validation Scientist's contract. Care must be taken, especially when matching using the extended coincidence windows, to ensure that any ISAR data acquired in or close to the ports at either end are rejected, even though the AATSR pixel may be up to 20 km offshore in the case of the Grade 2B and 3 datasets.

It is normal for several independent match-up points to be retrieved from a cloud-free overpass. However, it is important that the AATSR data products should be validated across many samples of different atmospheric conditions. Table 3-3 shows the number of different overpasses from which the dual view matches in Table 3-2 were obtained. These are considered to represent independent atmospheric conditions.

Table 3-3: The number of Envisat overpasses, over the whole project, for which a match was obtained between an ISAR record and an AATSR pixel containing dual view 2- or 3-channel SST data, for different grades of match-up coincidence.

AATSR product type	Grade 1	Grade 2A	Grade 2B	Grade 3
D3	12	24	21	30
D2	16	29	30	46
All dual view	28	53	51	76

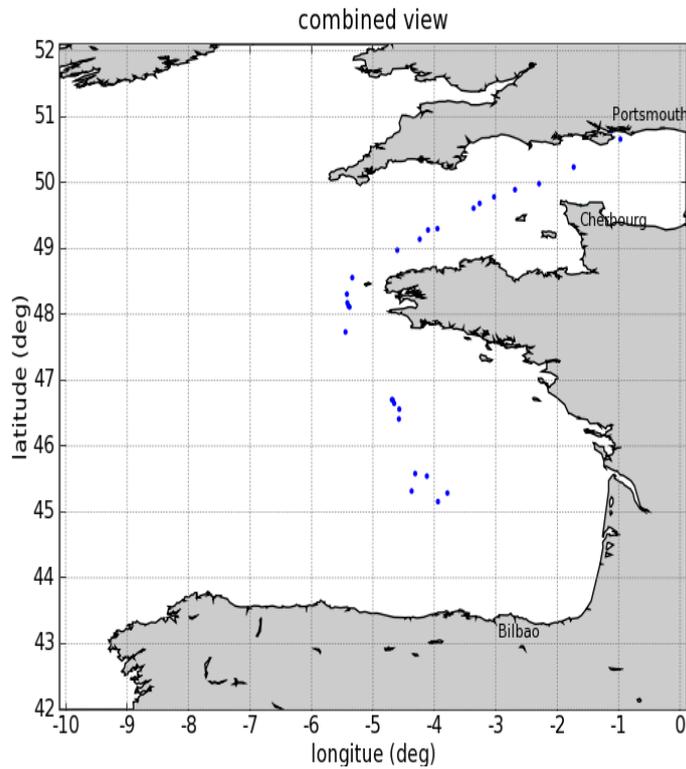


Figure 3-6. Locations of the AATSR pixels in Grade-1 match-ups between ISAR and AATSR dual-view SST, aggregated over all deployments. Grade 1 is within ± 2000 s and 1 km

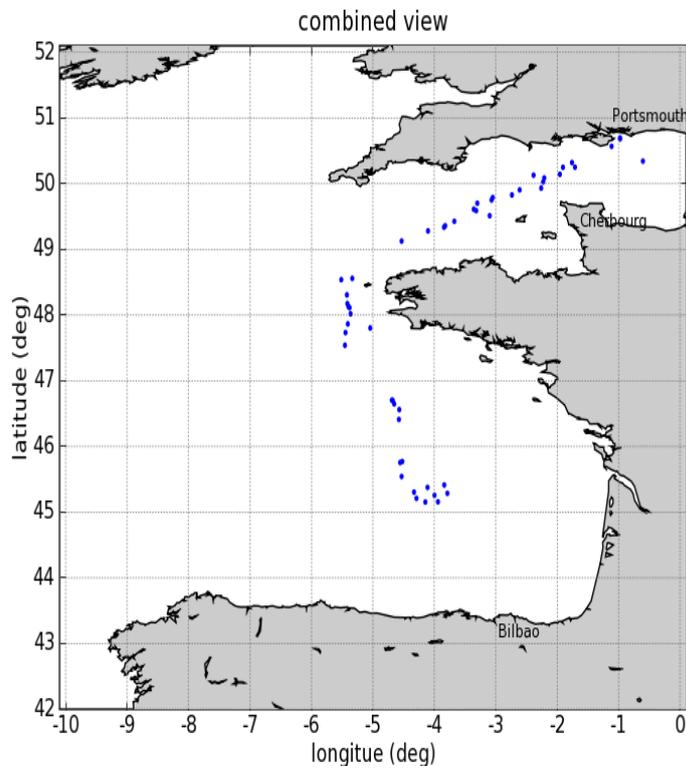


Figure 3-7. Locations of the AATSR pixels in Grade-2A match-ups between ISAR and AATSR dual-view SST, aggregated over all deployments. Grade 2A is within ± 2000 s and 20 km.

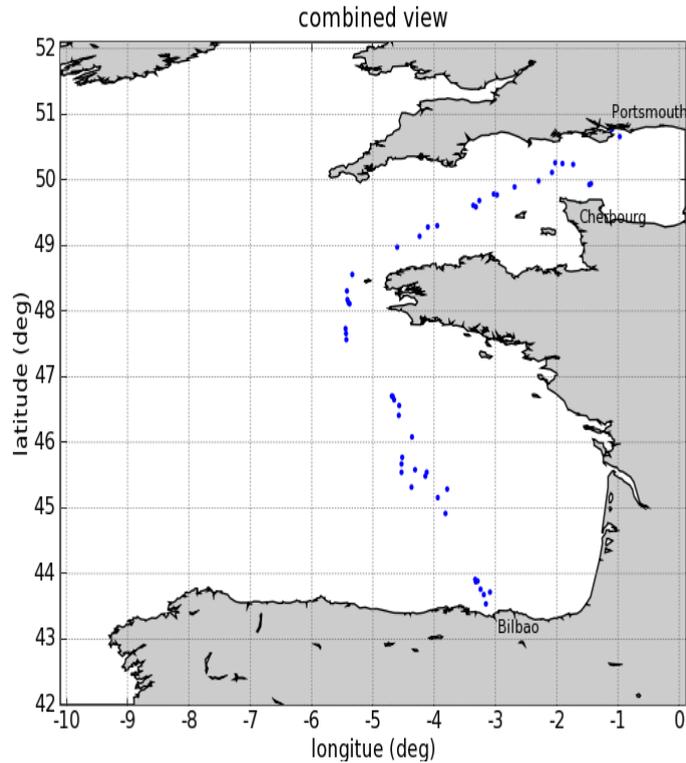


Figure 3-8. Locations of the AATSR pixels in Grade-2B match-ups between ISAR and AATSR dual-view SST, aggregated over all deployments. Grade 2B is within ± 2 hrs and 1 km

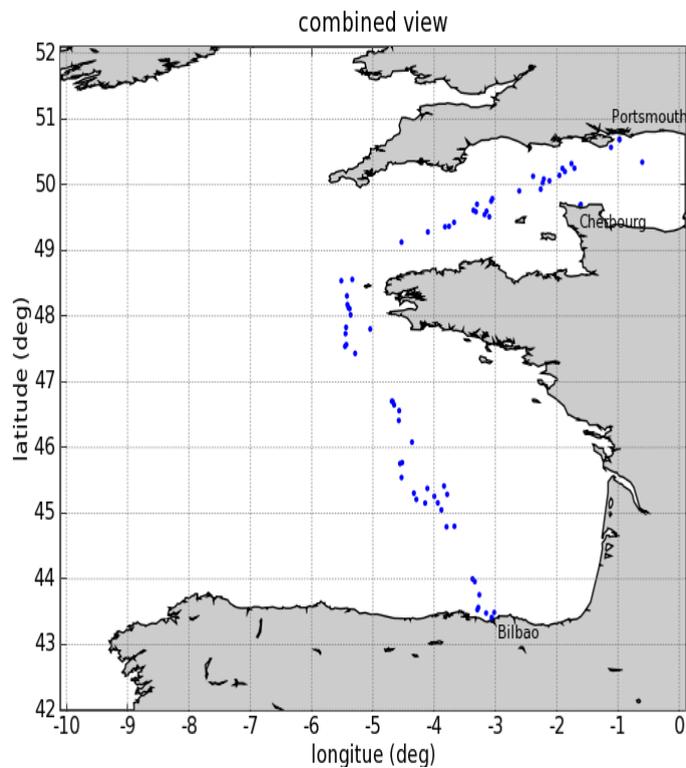


Figure 3-9. Locations of the AATSR pixels in Grade-3 match-ups between ISAR and AATSR dual-view SST, aggregated over all deployments. Grade 3 is within ± 2 hrs and 20 km

4. Validation of AATSR against ISAR data

4.1 Comparisons between AATSR and ISAR

Tables 4-1 to 4-4 show the matchup statistics aggregated over all the deployments for AATSR SST products produced respectively by the D3, D2, N3 and N2 algorithms. In each table the four rows correspond to values obtained using the different matchup criteria. The tables also show the numbers of matchups on which the comparisons are based and, for D2 and D3, the range of SST values encompassed by the match-up dataset.

Table 4-1 Statistics of comparison between satellite and in situ skin SST in the match-up dataset, (AATSR - ISAR) for the 3-channel dual view product (D3).

Matchup Grade	Bias, deg C	Standard deviation	Number of matches	SST range, deg C
1	0.19	0.33	76	11.3 - 17.9
2a	0.16	0.39	99	10.1 - 21.3
2b	0.10	0.34	274	7.8 - 20.9
3	-0.03	0.62	429	7.8 - 20.9

Table 4-2 Statistics of comparison between satellite and in situ skin SST in the match-up dataset, (AATSR - ISAR) for the 2-channel dual view product (D2).

Matchup Grade	Bias, deg C	Std deviation	Number of matches	SST range, deg C
1	-0.00	0.36	67	10.1 - 19.1
2a	0.03	0.37	82	9.5 - 19.5
2b	0.11	0.60	288	9.3 - 24.9
3	0.16	0.72	388	9.3 - 24.9

Table 4-3 Statistics of comparison between satellite and in situ skin SST in the match-up dataset, (AATSR - ISAR) for the 3-channel nadir-view only product (N3).

Matchup Grade	Bias, deg C	Std deviation	Number of matches
1	0.31	0.27	111
2a	0.13	0.43	263
2b	0.25	0.29	402
3	0.05	0.58	915

Table 4-4 Statistics of comparison between satellite and in situ skin SST in the match-up dataset, (AATSR - ISAR) for the 2-channel nadir-view only product (N2).

Matchup Grade	Bias, deg C	Std deviation	Number of matches
1	0.71	0.36	137
2a	0.66	0.56	264
2b	0.67	0.48	576
3	0.60	0.67	1133

Taken at face value, assuming the ISAR is measuring the true skin SST which the AATSR products aim to retrieve, the results in Table 4-1 and Table 4-2 suggest that the AATSR is performing well and comfortably within its specification of ± 0.3 K, when the dual view algorithms (3 or 2 channel) are used. For a number of the datasets the bias is less than 0.1 K which is the limit of accuracy of the ISAR. On the other hand, the AATSR SST retrievals using the nadir view only show a warm bias, which for the N2 algorithm appears to be more than 0.6 K. This is consistent with the findings of other validation programmes, which led to a revision of the N2 and N3 algorithms in December 2005. Further observations using the ISAR should be able to confirm whether the new algorithms are better.

What is also encouraging in the tables is that the standard deviation of match-ups is relatively small. It remains around 0.3-0.4 K as the coincidence grade is weakened from 1 to 2A or 2B. Only when the criterion is reduced to grade 3, allowing more mismatch errors, does the standard deviation grow significantly. The standard deviation for the nadir results is comparable with the dual view results, suggesting that the bias for the nadir algorithms is genuine and not an artefact of additional noise in the match-up process. Final conclusions about the validation of AATSR using the ISAR data are presented in the report of the AATSR Validation Scientist in the companion report to this one.

While the above tables show the aggregate matches over the two year project, further match-up statistics presented in Appendix C show the match-ups broken down into bias and standard deviation from each individual deployment. In general these are not so useful since the numbers of samples from each individual deployment are smaller. Nonetheless they help to identify whether a particular deployment is out of line with the others, and should therefore be given less weight in the overall analysis, following the approach adopted by the AATSR validation scientist.

4.2 Discussion – significance of results

Our experience of matching individual ISAR samples of skin SST to AATSR pixels, gained during the present contract, has raised the issue of how to interpret the differences detected between the *in situ* and the satellite data. This difference is often treated as a measure of the accuracy or validity of the satellite measurement, but to do so without careful consideration of other factors could be misleading. The error in the satellite estimate is actually the difference between the true sea surface temperature and the satellite-derived estimate. The “true” SST as observed by the satellite corresponds to the average skin temperature over the field of view of the sensor at the instant of the overpass. In general this is not the same as the *in situ* measurement entered into the match-up database and used for validation.

If validation is to be used as a measure of how well a sensor is performing, and in particular whether the approach to atmospheric correction is successful, it is important to eliminate, or allow for, discrepancies between SST measured *in situ* and what the satellite claims to measure. For example if a satellite product attempts to represent the subsurface or bulk SST, as is the case for the products of some ocean temperature missions, then that is the type of *in situ* observation that should be used to validate it. But the AATSR expressly claims to measure the skin SST averaged over the field of view. Therefore we must take particular care to make *in situ* measurements that do the same.

There are several reasons why *in situ* SST measurements differ from the true satellite observed SST. The most obvious is that all in-water sensors measure near-surface temperatures below the skin. As discussed in 2.1, by using shipborne radiometry we have already eliminated this problem, which would otherwise seriously jeopardise the chance of proving the true accuracy of the AATSR products. However, there remain several reasons why even radiometric *in situ* measurements may not correspond to the “true” SST. The potential discrepancies between the *in situ* measurement of SST and the actual skin SST averaged over 1 km² can be divided into two categories, the instrument or measurement errors and the sampling mismatch discrepancies.

There are two sources of measurement error associated with the skin SST observations undertaken in this work. The first is the accuracy of the ISAR radiometer in measuring the brightness temperature of the infrared radiation reaching it. We believe that the laboratory validation experiments with the CASOTS-2 black body before and after each ship deployment demonstrate that these error are within ± 0.1 K. The intercalibrations in the field against other ship-borne radiometers reinforce this conclusion. The second potential source of error is in the way the sea-viewing brightness temperature is converted into a skin SST, using an assumed sea

surface emissivity and the measure of the sky radiometric temperature. There is no obvious way of evaluating how good the emissivity assumptions are. The ISAR data processing model assumes similar values to those used in the AATSR processing algorithm, so to this extent the ISAR should be observing the same physical property as is seen by the AATSR. The agreement between the ISAR and the M-AERI instrument is also encouraging in that the M-AERI is less dependent on the same emissivity assumptions. It is reasonable to assume that any errors associated with this uncertainty are less than ± 0.1 K, at least in conditions where there are neither organic films nor foam on the sea surface which may reduce its emissivity.

The sampling discrepancies are likely to be larger than the $< \pm 0.1$ K measurement errors, but can also be expected to have a much smaller bias. They arise from the difference between a point sample and an area average (which can be termed the error of unrepresentativeness), and mismatches in space and time preventing perfect coincidence. All of these depend on the natural space-time variability of the SST field and will vary from place to place and time to time. Strictly the capacity of a point sample to represent an area average depends on how uniform the skin SST is at sub-pixel length scales. By using 40 ISAR samples along the ship track acquired over about 1 minute the ISAR is at least performing a line average over a few hundred metres. The mismatch error also depends on whether cloud prevents a perfect spatial coincidence.

The simple consideration in 4.1 of the AATSR performance compared to the ISAR observations did not consider the issues raised in this section. Therefore the conclusions made in relation to AATSR validation must be treated with caution. However, the match-up dataset summarised in 4.1 was sent to the AATSR validation scientist who has examined in more detail the spatial and temporal SST variability surrounding each match-up pair in order to determine how reliable each one is as unbiased validation data. Moreover, the validation of AATSR depends on several other datasets of matched *in situ* observations, although it should be noted that the data presented here offer validation for much lower sea temperatures than any of the other *in situ* matching validation programmes.

5. Conclusion

5.1 Achievements of the project

The successful two year operation on the *Pride of Bilbao* has demonstrated the effectiveness of the ISAR for acquiring validation match-ups of skin SST for comparison with satellite data products. The acquisition of 608 days of data within two years has proved that an autonomous radiometer on a ship of opportunity can reliably deliver high quality measurements. During most of the operational days at sea, the ISAR accuracy was determined to be better than 0.1 K.

During the project a total of 143 validation pairs of dual-view satellite retrievals and *in situ* observations were acquired, measured within 1 km and 30 minutes of each other, corresponding to 28 different overpasses. Using a less strict criterion of coincidence, within 20 km and 2 hours, 817 validation pairs were acquired from 76 overpasses. For the nadir-only view data, the number of matches rose to 248 for the closest coincidence and 2048 for the weakest coincidence criterion.

The comparisons between *in situ* and satellite data showed that the ISAR measurements of skin SST were mostly within around 0.1 K of the dual view SST products from AATSR, implying that AATSR is operating well within its specification. However, the 2-channel nadir-only SST products were biased more than 0.6 K higher than the ISAR. The full implications of the match-up dataset for the validation of AATSR SST products are developed in the final report from the Validation Scientist, since the ISAR results are just a part of the total volume of available validation data. However, the ISAR results are broadly consistent with other validation activity. Moreover it is important to note that without the ISAR data the global validation programme would lack validation data for temperate and cool sea water conditions, and at mid to high latitudes.

5.2 Requirement for ongoing work

Validation of satellite data products is an ongoing process and needs to be maintained for the lifetime of the sensor and its associated product generation. the ISAR programme will therefore continue for at least a further two years in order to contribute to the worldwide effort to validate AATSR SST products. We expect to be able to provide an even more reliable service, with near-real time delivery of ISAR data, building on the operating experience described in this report.

We also propose two particular lines of study which make use of the ISAR data and matchups with AATSR acquired during the last two years. The first is to develop an approach that can assign an error estimate or confidence value to each match-up

data pair entered into the validation database. This would be based on an estimate, of the uncertainties in the matching process, using awareness of the immediate space-time variability of SST associated with the matched data, cloud probability etc. as well as the error on the ISAR measurement itself. The quality flagging of match-up data pairs will allow a more refined definition of error statistics for AATSR. Secondly we plan to analyse the differences measured between the ISAR skin SST and the hull temperature on the *Pride of Bilbao*, and their dependence, if any, on the suite of meteorological parameters that have been collected. Our aim will be to characterise that difference with a view to determining whether, and with what confidence, hull temperatures recorded by other ships can be used to estimate skin SST when no ship-borne radiometer is deployed. If this is feasible it would considerably enlarge the *in situ* data volume available for ongoing AATSR validation.