# Successes and Challenges for the Modern Sea Surface Temperature Observing System: The Group for High Resolution Sea Surface Temperature (GHRSST) Development and Implementation Plan (GDIP)

[1]C. J. Donlon, [2]K. S Casey, [3]C. Gentemann, [4]P. LeBorgne, [5]I. S. Robinson, [6]R. W Reynolds, [7]C. Merchant, [8]D. Llewellyn-Jones, [9]P. J. Minnett, [10]J. F. Piolle, [11]P. Cornillon, [12]N. Rayner, [2]T. Brandon, [13]J. Vazquez, [13]E. Armstrong, [14]H. Beggs, [15]I. Barton, [16]G. Wick, [17]S. Castro, [18]J. Hoeyer, [19]D. May, [20]O. Arino, [5]D. J. Poulter, [9]R. Evans, [21]C. T. Mutlow, [13]A. W. Bingham and [22]A. Harris.

[1]ESA/ESTEC, Noordwijk, The Netherlands, [2]NOAA National Oceanographic Data Center, USA, [3]Remote Sensing systems, USA, [4]Meteo France, France, [5]National Oceanography Centre, Southampton, UK, [6]NOAA National Climate Data Center, USA, [7]University of Edinburgh, UK, [8]University of Leicester, UK, [9]University of Miami, USA, [10]IFREMER, Brest, France, [11]University of Rhode Island, USA, [12]The MetOffice, UK, [13]NASA- JPL PODAAC, USA, [14]Bureau of Meteorology, Melbourne, Australia, [15]CSIRO, Hobart, Tasmania, Australia, [16]NOAA ESL, USA, [17]University of Colorado, USA, [18]Danish Meteorological Institute, Denmark, [19]NAVOCEANO, USA, [20]ESA/ESRIN, Frascati, Rome, Italy, [21]Rutherford Appleton Laboratory, UK, [22]University of Maryland, USA

# Executive Summary

This paper provides a core contribution to address the OceanObs09 theme “Developing technology and infrastructure” under Session 4B “Satellite data integration and products”. It has been prepared for those implementing and working with the modern sea surface temperature (SST) observing system. The purpose of the paper is threefold: (1) to highlight key developments of the modern era SST observing system over the last 10 years (2) to discuss the principal challenges for the observing system in the next 10 years and (3) to propose an ideal plan for the global integrated high resolution SST observing system. The paper also constitutes the Group for High Resolution Sea Surface Temperature (GHRSST) Development and Implementation Plan (GDIP).

The following recommendations are made:

1. We recommend that steps should be taken to increase the accuracy and spatial resolution of passive microwave SST data sets through innovative satellite design. These improvements would enable better coverage of the coastal zones and better discrimination of mesoscale structures that are closely coupled to the surface wind field improving SST climate data records and numerical weather forecasts.
2. A concerted effort should be made to ensure the sustained continuity of passive microwave SST using a ~6.9 GHz channel in an operational (redundant) context. Steps should be taken to ensure that better accuracy and high spatial resolution are key design goals for future passive microwave satellite radiometers.
3. A capability for satellite infrared dual view along track scanning radiometry should be sustained in an operational context with redundancy.
4. Future single view satellite infrared radiometer systems developed for accurate SST retrieval should include channels at ~8.6, 3.9 and 4.05 $\mu$m.
5. The capability for high-quality satellite infrared multi-frequency radiometry in geostationary orbit should be sustained. Ideally a constellation of 6 geostationary instruments is required to provide global coverage at satellite zenith angles amenable to quantitative SST estimation. We note that continuity of both IR and passive microwave SST observations on polar and geostationary platforms is considered essential for an accurate and robust SST Climate Data Record (CDR).
6. All SST observations are reported together with a depth of observation and that adequate transmission codes are provided for in situ radiometers and other instruments to report their SST data in near real time via operational communication systems. Special efforts should be made to improve the quantity and documented calibration of all in situ SST measurements – especially those made by ships – so that they can be used in satellite validation activities. Finally improvements to the sampling of the global drifting buoy array in the Southern Ocean and Arctic Ocean should be implemented, building on existing legacy systems, to achieve better sustained sampling in these regions.
7. Observing system experiments (OSE), sampling studies and, error analyses (such as the Potential Satellite Bias Error, PSBE) should be an integral part in the design, development and operation of an integrated SST observing system suited to both near real time operations and climate data record production.

8. We strongly recommend that all Argo floats are equipped with a capability to make high vertical resolution measurements of SST in the upper 10m of the ocean surface and that shallow-water Argo floats be developed and deployed. Evidence should be gathered and presented to Argo float manufacturers highlighting the applications and benefit of providing high resolution temperature sensors on all Argo floats for the investigation of SST diurnal variability.
9. A concerted effort is required to develop a framework to provide robust uncertainty and bias estimates for *in situ* SST data sets based on the use of satellite, model, and climatological reference data sets in collaboration with JCOMMOPS and other Agencies.
10. We fully endorse the Global Climate Observing System (GCOS, 2006) climate monitoring principles, which are all relevant to an integrated (satellite plus in situ) SST observing system, and urge agencies to take steps to implement these principles in their short and long term planning for the future of satellite SST data sets.
11. We recommend that a systematic framework in which satellite SST data sets can be re-processed is developed and operated. The system should foresee multiple re-processing of L0 (engineering) data through to L2 (geophysical) products to produce the best Climate Data Record for each satellite sensor. Furthermore, developments achieved for climate re-analyses should be quickly integrated into operations.
12. The Group for High Resolution SST/GCOS SST & Sea Ice Working Group modern-era SST inter-comparison activity is sustained and enhanced to include other satellite data sets in order to identify systematic biases between complementary SST Climate Data Record.
13. Further steps should be taken to strengthen data stewardship activities for the modern era SST data record including tools that provide easy access to multiple source satellite and in situ data. Tools should also be provided for the regular processing and analysis of combined data sets (e.g., the ESA Grid Processing on Demand (G-POD) system).
14. A consensus approach is obtained and documented to produce sensor specific SST climate data records (one per sensor/series) in an optimized end-to-end high-resolution global coverage SST measurement system. Also, that an appropriate international cooperative data service is implemented that will develop, maintain and provide SST Climate Data Records.
15. We recommend that an optimal level of diversity in approach and competition to achieve the best result is encouraged when implementing SST Essential Climate Variable processing systems.
16. Further research must be conducted to develop and refine SST optimal estimation techniques that maximize the error reduction from having multiple complementary observing systems in space.
17. Satellite developers should include better (two calibration targets) on-board calibration systems for satellite infrared and passive microwave sensors, ensure that adequate data are collected on board the spacecraft to characterise instrument calibration, and that these data are telemetered to ground. We also recommend, based on 17 years of experience, that the along-track scanning technique provides an excellent approach to measuring accurate SST from satellites and should be more widely adopted for operational systems.
18. Future satellite and in situ SST observing systems should provide a measurement of contemporaneous wind stress in order to understand the context of SST measurements (e.g., cool skin and thermal stratification), to help blending of SST measurements made by different methods under moderate wind stress and to facilitate the application of SST measurements in scientific and operational systems.
19. A systematic data gathering exercise should be performed to identify and catalogue diurnal SST warming events across the global ocean using satellite and in situ measurements in order to better understand the spatial and temporal variability of such events and their impact at daily, monthly seasonal, annual and multi-annual timescales.
20. Steps should be taken to secure high temporal resolution (ideally at an hourly resolution) wind fields over the global ocean for use in diurnal SST variability modelling.
21. Single Sensor Error Statistics (SSES) describing SST uncertainty for individual satellite sensors on a pixel-by-pixel basis should be stratified by season and region and include a variety of auxiliary data that can be used to derive the SSES in a hypercube (multi-dimensional) matchup database approach.
22. We note the need for better in situ SST observations and supporting data and metadata (including depth of observation and calibration history) for use in Single Sensor Error Statistic derivation schemes. We recommend that (1) operators of in situ radiometer systems are encouraged to share their observations in order to maximize the benefit to the satellite SST community and (2) all in situ observations when reported should include depth, calibration history, measurement method and a meaningful estimate of uncertainty.
23. The Committee on Earth Observation Satellites (CEOS) Quality Assurance for Earth Observation data (QA4EO) approach should be used by Space Agencies working together with inter-Governmental agencies, in situ data providers and the scientific SST community when defining, implementing and operating a sustained SST validation program.
24. We recommend that a concerted effort to develop better Single Sensor Error Statistics uncertainty and error estimates for all satellite and in situ data sets should be a priority and urge agencies and scientists to devise new schemes that can be implemented in an operational manner. We recommend that uncertainty estimation follows and, where appropriate, contributes to the Quality Assurance for Earth Observation data (QA4EO) approach both for satellite data and for in situ data.

25. We endorse the approach in which well characterised and accurate satellite sensors may act as a reference data source for bias correction and Single Sensor Error Statistics derivation but note that steps should be taken to assure the sustained provision of such data (e.g., the ESA Sentinel-3 Sea and Land surface Temperature Radiometer (SLSTR) that will follow the Advanced along Track Scanning Radiometer (AATSR) due for launch in 2012). We further recommend that new satellite infrared radiometer systems make full use of additional channels in the 3-12 $\mu$m spectral region.
26. Space Agencies working together with inter-Governmental agencies, in situ data providers and the scientific SST community take steps to define, implement and operate a sustained validation program for all satellite radiometers producing SST for the lifetime of the mission. We also recommend that Space Agencies and in situ data providers include sustained validation of their SST products as a fundamental component of each satellite mission and in situ platform for the entire mission/instrument duration.
27. We recommend that an optimum number of in situ high-quality infrared radiometer systems are maintained for the purpose of infrared and passive microwave satellite validation building on the demonstrated capability during the last 10 years. Such systems should be augmented by better aircraft measurements that have the additional benefit of covering large transects although care should be taken to ensure that atmospheric attenuation of the sea surface signal does not compromise the accuracy of data when flying at altitude.
28. An international reference match-up database that applies the same control and process methodologies to the match-up process between satellite and in situ SST measurements and supporting input data is developed, maintained and operated in order to allow a fair, consistent and up-to-date estimation of errors on satellite SST retrieval. We note such a common database is best constructed following the systematic re-processing of satellite data as part of a Climate Data Record development activity and should be consistent with the CEOS QA4EO process.
29. We recommend that the SST community of producers and users establish and maintain:-
    a. A programme of *in situ* measurements, both thermometers on buoys, ships and subsurface vehicles and radiometers on ships and platforms that can be used for validating the different products. For the sake of efficiency It is desirable that this be a fully collaborative programme shared between all the agencies responsible for SST products. There is also a need to specify the requirements for new *in situ* data acquisition systems to support data integration, including wider coverage by ship-based radiometers, diurnally resolving moorings, Argo with additional sensors for near-surface sampling and the OceanSites approach.
    b. A forum in which sets of quality metrics relevant to specific uses of SST data are defined and regularly reviewed (consistent with QA4EO). Different metrics will need to be defined for different applications, for example timeliness and reliability of error statistics may be more important than accuracy for assimilation into forecasting models. whereas absolute accuracy and stability are most important for climate datasets.
    c. Tools such as the high-resolution diagnostic dataset (HRDDS) described above, that facilitate comparisons between different SST data products. New developments should include automated comparisons between different SST products to alert producers when inconsistencies are detected.
    d. Match-up databases (MDB) which relate satellite-derived SST products with in situ measurements of SST. It is essential to preserve independence between those *in situ* observations used for calibration of SST products and those used for validation of the sensor specific error statistics attached to each data product, for which a new set of tools needs to be developed. Moreover, for long term climate monitoring it will be important to establish a “protected MDB” that can be preserved as an independent validation reference for future generations of SST product re-analyses. Ideally such data should not be allowed to influence the evolution of product algorithms in order to preserve their independence.
    e. A generic system to acquire better feedback from data assimilation systems (e.g., observation weights rejection statistics) in order to ascertain how much influence different SST observational products have on the observing system models. Such information will identify targets for the quality of satellite SST data products, necessary for them effectively to constrain ocean models.
30. The ocean SST (GHRSST) community work effectively with the Global Space Based Satellite Inter Calibration System (GSICS) and establish if and how it may assist GSICS for the benefit of improved SST measurements by working together (e.g., using the High Resolution Diagnostic Data Set (HR-DDS) and GHRSST Multi Product Ensemble (GMPE) tools).
31. SST analyses producers should provide a user oriented document (to a standard template) that describes each analysis product and the choices and assumptions that have been made for the analysis procedure. The document should highlight the strengths and weaknesses of the analysis output in order to help users use the data in the most appropriate manner.
32. The GHRSST Multi Product Ensemble (GMPE) capabilities should be extended to provide quantitative data outputs in a GHRSST format for use by the general ocean community, consider higher spatial resolutions for the system grid specification, the inclusions of automated procedures for statistical analysis of the ensemble and the creation of regional GMPE tools for specific and challenging regions (e.g., Tropical Pacific, Bay of

Bengal and Indian Ocean, Indonesia, Arctic Ocean, Great Lakes). These services could be distributed across multiple centres.

33. The GHRSST High Resolution Diagnostic Data Set (HR-DDS) should be extended to include user defined quantitative analysis and event driven monitoring capabilities in near real time to monitor the quality and performance of the SST observing system. Reports should be automatically sent to data providers and users when an event of interest is encountered by the HR-DDS system. Furthermore, the HR-DDS should be expanded to include long time series of SST data sets to be visualised for reanalysis purposes.
34. A high resolution inter-comparisons group should be established under GHRSST that would be similar to the SST inter-comparison group already established for historical SSTs within the framework of GCOS. This group would include representatives of each of the major centres creating high resolution SST analyses with terms of reference established by the GHRSST science team.
35. As cloud screening of infrared satellite data is remains a significant challenge, despite nearly 30 years of activity, and failure to detect sub-pixel clouds remains the source of substantial uncertainty in satellite data sets, we recommend that a systematic review of cloud clearing approaches is undertaken with the purpose of properly documenting the strengths and weakness of each approach. Such a review should identify and prioritise a set of activities that should be undertaken to ensure that the best possible methods of cloud clearing are identified and used by satellite data providers. The review should also assess the severity of cloud detection to the degree of tolerable SST impact for unscreened clouds in the context of user applications.
36. We recommend that further development of probabilistic cloud detection and flagging algorithms should be undertaken to potentially improve the quality of satellite SST data sets from infrared sensors.
37. More effort must be given to the definition and implementation of ice masking procedures and techniques in polar regions for infrared satellite observations.
38. Satellite SST data providers using infrared systems should review the performance of their atmospheric correction algorithms in polar atmospheres and take steps to develop more appropriate algorithms for these regions.
39. Further efforts should be undertaken to continue the SST time series with more reliable in situ and radiometer observations in the Arctic Ocean.
40. Producers of L2 SST products discuss and agree on a common definition of global lakes for which Lake Water Surface Temperature (LWST) shall be retrieved routinely from satellite data ideally as part of the SST processing systems.
41. Satellite SST navigation techniques should be improved to allow confident delineation of lake and coastal shores, tidally exposed wet/dry areas and improvement in the retrieval of LSWT and SST in complex coastal regions. Pixel classification flags should be revised and improved in a consistent manner so products use/share a common flag set (at a minimum the definitions should be agreed) to assist in these activities.
42. The netCDF Climate Forecast (CF) convention is used by satellite and in situ data providers to manage metadata and interoperability of the modern-era SST data record. We urge other satellite SST data providers to adopt the GHRSST L2P, L3, and L4 netCDF format approach to providing SST data to user communities.
43. Satellite SST data providers should take steps to make their L1b data available for use in the SST CDR re-analysis community and as part of NRT GHRSST products.
44. The adoption of a standards based approach to metadata and file formats has been one of the foundations of success for GHRSST and we recommend that the approach is strengthened and extended to other satellite and in situ SST data providers using appropriate mechanisms.
45. We recommend that the GHRSST Data Specification (GDS), in a revised and updated format, and the GHRSST Regional/Global Task sharing Framework (R/GTS) are strengthened and sustained in order to maintain and develop the strong user-producer collaboration that has enabled a new generation of SST data products and services to develop over the last 10 years.
46. GHRSST Regional Data assembly Centres (RDAC) should be sustained and strengthened following the subsidarity principle in order to provide the maximum benefit to each region, to ensure that feedback from national and regional users is gathered, and that the modern era SST record and services develop according to user requirements.
47. The GHRSST Global Data Assembly Centre (GDAC) service should continue to develop and collaborate with RDACs and the Long Term Stewardship and Re-analysis Facility (LTSRF) to actively monitor the status of the R/GTS, data products and work with users.

# 1. SST observations and their societal benefit

Sea Surface Temperature (SST) at the ocean-atmosphere interface is a fundamental variable for understanding, monitoring and predicting fluxes of heat, momentum and gas at a variety of scales that determine complex interactions between atmosphere and ocean. The ocean stores heat from the sun and redistributes it from the tropical regions to higher latitudes and to the less dense atmosphere regulating global weather and climate. Through the hydrological cycle the coupled system controls terrestrial life by redistributing fresh water over the land surface. From large ocean gyres and atmospheric circulation cells that fuel atmospheric depression systems, storms and hurricanes (all with their attendant wind waves and storm surges), to local scale phenomena such as the generation of sea breezes and convection clouds, SST at the ocean-atmosphere interface has a significant societal impact.

Mapping of SST is required for a range of tasks that have become the responsibility of operational monitoring and forecasting agencies in the 21st century (Robinson et al., 2009) and since 1981 an operational stream of satellite SST measurements has been sustained together with in situ measurements and collectively these form the modern-era SST observing system. Accurate knowledge of global SST distribution and temporal variation at finer spatial resolution is needed as a key input to forecasting and prediction systems to constrain the modelled upper-ocean circulation and thermal structure at daily, seasonal, decadal and climatic timescales, for the exchange of energy between the ocean and atmosphere in coupled ocean-atmosphere models and as boundary conditions for ocean forecasting models. Such models are increasingly being used operationally for various applications including maritime safety, military operations, ecosystem assessment, fisheries support and tourism. Well defined and error quantified measurements of SST are also required for climate time series (in the form of climate data records) that can be analysed to reveal the role of the ocean in short and long term climate variability.

In 1998 SST data production was considered a mature component of the observing system with demonstrated capability and data products. However SST product availability was limited to a few data sets that were large, scientific in format and difficult to exchange in a near real time manner. But, product accuracy was considered insufficient to the emerging ocean model and data assimilation community while at the same time the number of applications requiring an accurate high resolution SST data stream was growing. Considering these issues, the Global Ocean Data Assimilation Experiment (GODAE) defined the minimum data specification required for use in operational ocean models, stating that SST observations with global coverage, a spatial resolution of 10 km and an accuracy of 0.2ºC need to be updated every six hours (Smith and Koblinsky, 2001). Although there is a network of SST observations from ships and buoys, the only way to achieve this demanding specification is to use an integrated approach built on four principles:

(1) Respond to user SST requirements through a consensus approach,
(2) Organise activities according to principles of subsidarity[1] and shared responsibility,
(3) Develop complementarity between independent measurements from earth observation satellites and in situ sensors and,
(4) Maximise synergy benefits of an integrated SST measurement system and end-to-end user service.

These are the foundations on which the international ocean remote sensing community, marine meteorologists, Space Agencies and ocean modellers combined their energies to meet the GODAE requirements by establishing the GODAE High-Resolution Sea Surface Temperature Pilot Project (now called the Group for High Resolution SST, GHRSST[2]). GHRSST has four main tasks that are relevant to the development of the SST observing system: (1) Improved SST data assembly/delivery (2) Testing of SST data sources (3) perform inter-comparison of SST products (4) Develop applications and data assimilation of SST to demonstrate the benefit of the improved observing system. GHRSST has successfully demonstrated that the requirements of GODAE can be met and has been instrumental in defining the shape and form of the modern-era SST measurement system and user service over the last 10 years (Donlon et al, 2008). However, while progress has been made, much work remains to be done especially with respect to the 3$^{rd}$ and 4$^{th}$ principles of the integrated approach outlined above. Against this background, the paper also constitutes the GHRSST Development and Implementation Plan (GDIP)

[1] Subsidiarity is an organizing principle that calls for matters to be handled by the smallest, lowest or least centralized competent group.
[2] GHRSST will be used in the remainder of this paper and includes all activities of the current Group as well as the (?) Pilot Project.

This paper will first review key developments in SST over the past 10 years focusing on the GHRSST experience reporting the successes and challenges of the last decade. The second part of the paper will focus on the primary challenges for SST in the next 10 years, and the third part will propose an ideal high resolution SST observing system for the community to build toward over the longer term.

# 2. SST successes of the last decade through the GHRSST experience

The GHRSST project was a significant (but not the only) contribution to progress in SST over the last decade as it nurtured a community of scientists from the scientific and operational agencies and institutions. A series of international workshops were held during 2000-2009 that established a set of user requirements for all GHRSST activities in five areas: (1) scientific development and applications, (2) operational agency requirements, (3) SST product specifications, (4) programmatic organization of an international SST service and (5) developing and sharing scientific techniques and insight to improve data products and exploit the observing system. These requirements were critical to establishing a framework and a work plan and formed an essential part of the GHRSST evolution. By establishing and documenting clear requirements in a consultative manner at the start of the project and through all stages of its development, GHRSST was able to develop confidently and purposefully to address the needs of the international SST user community. A consensus GHRSST Data Processing specification (GDS) was developed that described how distributed international satellite data providers should process satellite data streams, the format and content of data products and the basic approaches to providing uncertainty estimates and auxiliary data sets that should be included in products to help users interpret the SST measurements. GHRSST also conducted scientific research and developed a data management framework including long term stewardship of all products. It developed a thriving international community that has revolutionized the way in which satellite and in situ SST observations are provided and used by a large user community including process scientists, operational ocean and NWP forecast centres and climate scientists. GHRSST realised the benefits of creating modular data processing architectures in which many partners around the world can contribute to improve global monitoring and the GHRSST service has encouraged the development of new SST monitoring and forecasting initiatives by operational agencies that use the new data products. In particular the end-to-end service reveals how a system which enables the complementary use of data from different sources reinforces the importance of each, as it leads to new records of SST with enhanced accuracy and improved spatial and temporal resolution. A full discussion of GHRSST success over the last 10 years is reported in Donlon et al., (2008, 2009). The key developments for SST since 1999 include:

- The successful development, launch, and operation by Space Agencies of accurate, high resolution satellite radiometers with complementary polar low earth and geostationary orbits having infrared and passive microwave channels and multiple views (CEOS, 2009).
- Wide and open access in near real time to many satellite SST data products has been established in an operational-like manner using existing data user-driven distribution protocols, tools and services. Over 26 Gb of data are provided in NRT every day by GHRSST Services, and over 25,500 international users have accessed GHRSST products (Donlon et al., 2009).
- International agreement on the definition of different SST parameters in the upper layer of the ocean that distinguish between measurements made by infrared radiometers, passive microwave radiometers, in situ sub-surface observations and SST merged analysis outputs. These definitions have been registered in the Climate Forecast (CF) standard name table for wide application (Donlon et al., 2008).
- Diverse satellite SST data product formats and product content have been homogenised according to international consensus and user requirements to include measurement uncertainty estimates for each derived SST value and supporting auxiliary data sets to facilitate their use by data assimilation systems.
- Development and sustained deployment of ship mounted autonomous in situ infrared radiometer systems for satellite SST validation.

- A significant increase in the number of in situ SST measurements from a variety of complementary sources are now available including Argo, drifting buoys, moored buoys and ships. In situ operators (VOS-Clim) are now reporting observation depth with SST measurements.
- 10 international GHRSST SST community workshops have been held attended by an average of 45 delegates to ensure that the SST user-producer community has been involved at all stages of GHRSST service and product development and evolution.
- GHRSST technical advisory groups have successfully conducted extensive research to ensure that SST diurnal variability (DV) is properly flagged within observational data, developed methods to correct for bias in different satellite data sets, provided uncertainty estimates on a measurement by measurement basis, developed high resolution sea ice data sets and accurate SST products in the marginal ice zone.
- New cost effective approaches to an integrated and optimised SST measurement system have been developed and are now used operationally to reduce bias error in AVHRR data using targeted global deployment strategies for drifting buoys (Zhang et al., 2009).
- New SST analysis products using new methods to merge in situ data with complementary microwave and infrared satellite data have been developed and implemented operationally.
- Inter-comparison frameworks (e.g., the GHRSST Multiproduct Ensemble (GMPE) see http://ghrsst-pp.metoffice.com/pages/latest_analysis/sst_monitor/daily/ens/index.html) have been developed at resolutions of 10km or better for the global ocean and other regions of interest. An operational high Resolution Diagnostic Data Set (HR-DDS see http://www.hrdds.net) has been established for real time inter-comparisons and validation/verification of GHRSST products allowing real time monitoring of satellite and in situ SST data streams.
- A delayed-mode intercomparison framework has been established in conjunction with the GCOS SST and Sea Ice Working Group to understand the linkages between the modern era satellite-based SST record and historical primarily ship-based SST reconstructions (see http://ghrsst.nodc.noaa.gov).
- Methods to convert between radiometric ‘skin’ SST and the SST at depths measured by ships and buoys have been developed (e.g., Donlon et al, 2002) that are now used by operational SST analysis systems (e.g., Stark et al., 2007).
- An internationally distributed suite of user focussed services are now provided in a sustained Regional/Global Task Sharing (R/GTS) framework that addresses international organisational challenges and recognises the implementing institutional capacities, capabilities and funding prospects. Long term stewardship, user support and help services including standards-based data management and interoperability have been developed that are manned and operated within the R/GTS on a daily basis.
- Methods to manage long-term satellite SST data sets for use in a reanalysis program that considers SST data for the entire satellite era have begun.

GHRSST has earned broad recognition as the international authority for modern-era SST activities because it has successfully built and nurtured a framework in which the exchange of satellite SST data has flourished and given new life to the study and application of high-resolution SST using satellite and in situ data. Applications have demonstrated positive impact in ocean and atmospheric forecasting systems and a new generation of data products and services to serve these and other users have been built and are operated on a day-to-day basis. The success of GHRSST stems from the Agencies and Offices that have supported the activities of the Pilot Project allowing a dedicated group of scientists and operational entities to successfully work together and bridge the gap between operations and science. All good operational systems are underpinned by excellent science and GHRSST has endeavoured to provide a forum in which operational systems and scientists can meet and discuss problems and solutions to address the real-world challenges associated with the application of high-resolution SST data sets.

# 3. Progress and challenges

While GHRSST has injected new vigour into the SST community it is not the only aspect of progress over the last 10 years. The following sections take a broader view of progress beyond the GHRSST activities.

## *3.1. Increased number and diversity of satellite observations*

According to the Committee for Earth Observation Satellites (CEOS) on-line database (CEOS, 2008, 2009) over 30 satellite missions capable of measuring SST in a variety of orbits (polar, low inclination and geostationary) have been launched since 1999. Of these, four missions (~15%) have a passive microwave SST retrieval capability (Oceansat-1 MSMR, ADEOS-1 and 2 AMSR, EoS AQUA AMSRE, TRMM-TMI, WindSat) with the remainder being infrared systems. None of these missions are part of a sustained operational framework or constellation.

### 3.1.1. New passive microwave SST capability

Infrared space sensors are demonstrably capable of achieving high levels of accuracy because their intrinsic characteristics allow for the introduction of sophisticated and reliable calibration procedures into the observing cycle. However, the inability to penetrate cloud not only severely limits the temporal and spatial continuity of observations but also, when used for climate monitoring, it has been speculated could introduce a potential cloud-free bias into the long-term SST record. Further work on this issue is required. By contrast, microwave sensors suffer only minimal absorption in the presence of clouds, but the intrinsic properties of the longer wavelengths means that there are major engineering challenges in providing high-performance and compact on-board calibration systems, especially if reasonably high spatial resolution, leading to large antenna surfaces, is sought.

Measurements of SST are made by satellite microwave radiometry in all weather conditions except rain. Microwaves penetrate clouds with little attenuation, giving an uninterrupted view of the ocean surface. This capability is a distinct advantage over infrared measurements of SST, which are obstructed by clouds and contaminated by atmospheric aerosol. The temperature of the sea surface measured using the 6-10-GHz band depends primarily on SST and wind speed. The wind effect is largely removed from the measured brightness temperatures using information in both the horizontally and vertically polarized channels providing a unique relation between the measured brightness temperatures and SST although improved retrievals are made when using independent wind vector estimates. Additional problems of side-lobe contamination occur when islands or a coastline reach into the antenna footprint of the microwave instrument (Wentz et al. 2000). The Tropical Rainfall Measuring Mission (TRMM) satellite Microwave Imager (TMI) launched in November 1997 (and still operating) is a well-calibrated passive microwave radiometer having a 10.7 GHz frequency channel for sea surface temperature retrievals. The imager provides 25 km gridded data coverage between 40N and 40S in a low inclination orbit although at an SST of < ~185 K the sensitivity of the TMI 10.7 GHz channel to SST is lost. The Earth Observing System (EOS) Aqua platform was launched on May 4, 2002 and carries the Advanced Microwave Scanning Radiometer for EOS (AMSR-E). AMSRE extends microwave SST capability into high latitudes using a 6.9 GHz channel and an 85GHz channel that is used to retrieve high resolution sea ice concentration/edge. AMSR-E can effectively monitor SST at a 25 km grid resolution in high latitudes with temporal resolution sufficient to capture small time- and space scales for the first time overcoming the confounding effect of near-permanent cloud cover. Passive Microwave SST has dramatically improved sampling of global SST with >90% global data coverage each day with minimal seasonal bias compared to infrared measurements (Guan eta la, 2003). Comparisons with ocean buoys show a root mean square difference of about <0.6 K (Stammer et al, 2003), which is partly due to the satellite-buoy spatial-temporal sampling mismatch and the difference between the ocean skin temperature and bulk temperature. The development of passive microwave SST retrievals has been one of the main highlights of the modern-era SST observing system over the last 10 years.

Current results (Vazquez et al., 2009) have shown that in areas of large SST gradients, high resolution data sets are necessary to adequately represent the dynamics. Chelton et al. (2007) demonstrate that summertime coupling between wind stress curl and SST gradients was not well represented in the NOAA North American Mesoscale Model, most likely due to the poor resolution of the model. Thus a major challenge in the future will be creating ultra-high resolution (< 5 km) SST data sets that can be incorporated into climate and forecast models. Vazquez et al. (2009) show that SST gradients in the Gulf Stream are effectively doubled when derived from 4 km rather than 9 km data sets. Daily SST data sets at 1-2 km resolutions can only be achieved by merging infrared and passive microwave data. **We recommend that steps should be taken to increase the accuracy and spatial resolution of passive**

**microwave SST data sets through innovative satellite design. These improvements would enable better coverage of the coastal zones and better discrimination of mesoscale structures that are closely coupled to the surface wind field improving SST climate data records and numerical weather forecasts.**

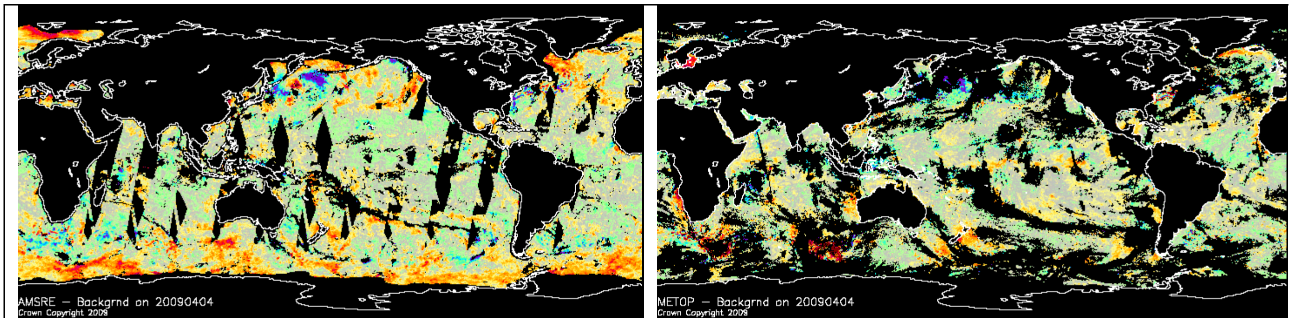


Figure 1 (a) AMSRE-OSTIA SST analysis (background) showing coverage of Aqua AMSRE 4th April 2009.

Figure 1 (b) METOP AVHRR - OSTIA SST analysis (background) showing coverage of METOP AVHRR April 4th 2009. Note reduced coverage in the Southern Ocean compared to AMSRE.

Between 2009 and 2030, about 40 separate satellite instruments capable of measuring SST are planned or approved of which ~15% are planning to provide a passive microwave capability (GCOM1, 2, and 3; AMSR-2; HY2A; NPOESS C3 and C4; and MTVZA). A major concern is the sustained provision of passive microwave SST using a 6.9 GHz channel following AMSRE to avoid any gap in the passive microwave SST record (NRC, 2008). While the GCOM series of satellites developed by JAXA plan to provide continuity over 3 missions using an upgraded AMSR-2 radiometer, only the first of these missions is approved and under development.

GHRSST has led to the development of daily SST analyses (e.g., the Met Office UK's OSTIA system) that utilize the complementary characteristics and advantages of several different SST sensors. The combined use of this disparate SST measurements means that operational users now have access to cloud-penetrating SST data from Microwave sensors which are bias-corrected to the levels of accuracy achievable by an AATSR-type sensor. Two important conclusions can thus be drawn: firstly microwave SST observations have quite quickly graduated from the status of experimental products from instruments which are produced and flown on an *ad hoc* or opportunistic basis, to products which are now routinely used in operational systems. This operational reliance creates a problem of continuity and a poses a challenge to Space Agencies. Secondly, if providers of operational data do accept this challenge and incorporate microwave SST sensors into their operational payloads, these microwave sensors should each be co-located with an infrared sensor that would enable the data-processing system to achieve the necessary bias corrections in the most accurate way possible. **We recommend that a concerted effort is made to ensure the sustained continuity of passive microwave SST using a 6.9 GHz channel in an operational (redundant) context. Steps should be taken to ensure that better accuracy and high spatial resolution are key design goals for future passive microwave satellite radiometers.**

## 3.1.2. New infrared SST capability

In March 2002, the Advanced Along-Track Scanning Radiometer (AATSR) was launched onboard the European Space Agency's Envisat satellite. This highly successful satellite, including the AATSR, is still functioning well and is producing high-quality data on a continuous basis. AATSR was the third in a series of experimental sensors designed to provide SST observations to the levels of accuracy, coverage and stability required for modern climate research. The first two ATSR sensors, which had been gathering data since the launch of ATSR-1 on ERS-1 in 1991, were experimental sensors and the developmental nature of the data was not well suited to many applications especially operational ones. For the ENVISAT mission ESA completely revised the data-processing scheme, taking full advantage of the experience gained during the two experimental missions and, as a result, AATSR now produces SST data which are shown by numerous validation activities, to achieve global accuracies in the region required by Climate applications. The (A)ATSR instruments are sensitive imaging radiometers designed to achieve this through rigorous calibration procedures, both pre-launch and in flight, combined with a new dual view of the Earth's surface which provided a superior atmospheric correction those which are possible using the conventional multi-channel method, which the (A)ATSR sensors also use in addition to the dual view observing technique. The dual-view capability is particularly important during periods of elevated stratospheric aerosol following major volcanic eruption, which occur with a repeat time of about a decade. SSTs from single view sensors suffer significant biases in such circumstances, but with appropriate SST retrieval design, dual view sensor SSTs are robust to such events (Merchant and Harris, 1999). Comparisons between the AATSR and drifting-buoy measurements made by the UK Met Office have shown that AATSR is capable of achieving biases in Global SST which typically < 0.2 K (O Carroll et al 2008). The SST data from AATSR can reach a level of accuracy needed for the climate monitoring and research and AATSR data are currently undergoing rigorous assessment at Climate Monitoring Centres around the world. Also, during the Envisat mission GHRSST has been able to take advantage of this stable and high-quality data-stream from ESA as part of the ESA Medspiration Project and provide AATSR to operational applications. Perhaps the most significant of these applications is the development by the Met Office UK of OSTIA, the Operations Sea Temperature and Ice Analysis (Stark et al, 2007), which is a daily global SST analysis using SST products from numerous data-sources, both satellite and in situ. This development enables the final SST product from OSTIA to combine the complementary advantages of different observing systems. In the case of OSTIA, the accuracy of AATSR is combined with the coverage of AVHRR, MODIS and geostationary sensors and the cloud-penetration capabilities of microwave sensors. Error analyses show clearly that AATSR SST data act as a 'benchmark' of accuracy, against which data from other sources can be bias-corrected. An additional and important attribute of AATSR SST data is the fact that the data-set is independent of other SST observations. This is because the retrieval schemes are based on physical principles and do not rely on regression techniques or empirical bias-corrections derived from in situ measurements. This independence means that, when introduced to either a climate record or to an NWP system, AATSR data introduce truly independent information. **We recommend that a capability for satellite infrared dual view along track scanning radiometry should be sustained in an operational context with redundancy.**

Sea surface temperature (SST) products have also been derived from the MODIS (MODerate Resolution Imaging Spectroradiometer) sensors onboard the NASA Terra and Aqua platforms since November 2000 (Terra) and 2002 (Aqua). These SST products are derived from the MODIS mid-infrared (IR) and thermal IR channels and are available in various spatial and temporal resolutions. The Terra/Aqua satellites are in a sun-synchronous near polar orbit at an altitude of 705 km with a descending node of 10:30 a.m. The MODIS sensor detects emitted and reflected radiance in 36 channels spanning the visible to IR spectrum (0.4 - 14.4 um). The MODIS SST products are distributed at various resolutions separated in ascending and descending orbits. The highest resolution for the Level 3 mapped products is 4.88 km. All Level 3 mapped products are derived from gridding the 4.63 km Level 3 binned data. The 4.88 km observations are themselves derived from binning and averaging (not subsampling) the nominal 1 km observations. Validation work conducted at the University of Miami for MODIS using 3-band algorithms incorporating 8.6, 11 and 12 μm for day-time and 3.75, 3.95 and 4.05 μm for night-time yields single swath RMS differences approximately equivalent to those of AATSR. Very little residual influence from water vapour, aerosol and cloud fringes is observed. Adding the 8.6, 3.95 and 4.05 μm bands improves performance across wide swath single-view sensors providing a substantial coverage gain with good performance, especially when coupled with the narrow dual swath AATSR-type reference coverage. **We recommend that future single view satellite IR radiometer systems developed for accurate SST retrieval should include channels at ~8.6, 3.9 and 4.05 μm.**

Geostationary satellite radiometers such as the EUMETSAT SEVIRI, Japanese MTSAT and USA GOES images provide significant source of SST data, which was not obvious in the early 90's due to challenges with atmospheric correction due to the lack of appropriate multi-spectral channels in the thermal infrared. The present generation of geostationary radiometers provides measurements that are comparable to that of radiometers aboard polar orbiting platforms such as the AVHRR (even at latitudes over 50° i.e. satellite zenith angle larger than 60°) having typical spatial resolution of 3-5km at nadir. The advantage of geostationary platforms is that observations can be made as often as 15 minutes over the same area allowing (1) detection of diurnal warming events and (2) increased chance of viewing the ocean surface as clouds move between measurement times to reveal the ocean surface. **We recommend that the capability for high-quality satellite infrared multi-frequency radiometry in geostationary orbit should be sustained. Ideally a constellation of 6 geostationary instruments is required to provide global coverage at satellite zenith angles amenable to quantitative SST estimation. We note that continuity of both IR and passive microwave SST observations on polar and geostationary platforms is considered essential for an accurate and robust SST Climate Data Record (CDR).**

## *3.2. Changing state and role of the in situ SST observing system*

Satellite SST observations have been calibrated and validated using *in situ* SST data since the early 1980's. Originally, mooring observations at around 1 m depth were used, with useful drifting buoy SST observations being employed from the mid 1980's complemented by ship observations. Since then, the number of drifter deployments has grown until they have dominated over the number of SST observations from moorings and ships (~80% of all SST observations). In contrast, the number of ship observations has halved since 1980 and the number of moorings remains roughly constant. Although there are typically 20,000 to 30,000 buoy observations of SST accessible in real-time via the Global Telecommunications System (GTS) each day, these do not cover all areas of the oceans. Regions particularly lacking in moored or drifting buoy observations are coastal regions, island archipelagos (e.g. Indonesia) and the Arctic and Southern Oceans (Figure 2). The changing shape of the in situ SST observing system has important implications for the climate record (see white paper by Kent et al): while the number of reported observations has increased, more data are originating from drifting buoy platforms which provide data that are reported more frequently in time and with smaller spatial separation compared to ships. Typically, engine intake or bucket SST observations from the ships of opportunity program (SOOP) are of uncertain accuracy, and therefore are not normally used for near real-time validation of satellite SST observations. Although thermosalinograph and radiometer SST observations from SOOPs and research vessels are used for validation, not all transmit data to the GTS in near real-time: in the case of radiometers due to a lack of approved coding standards. In addition, the oceans of the southern hemisphere are under-represented. **We recommend that all SST observations are reported together with a depth of observation and that adequate transmission codes are provided for in situ radiometers and other instruments to report their SST data in near real time via operational communication systems. Special efforts should be made to improve the quantity and documented calibration of all in situ SST measurements – especially those made by ships – so that they can be used in satellite validation activities. Finally improvements to the sampling of the global drifting buoy array in the Southern Ocean and Arctic Ocean should be implemented, building on existing legacy systems, to achieve better sustained sampling in these regions.**

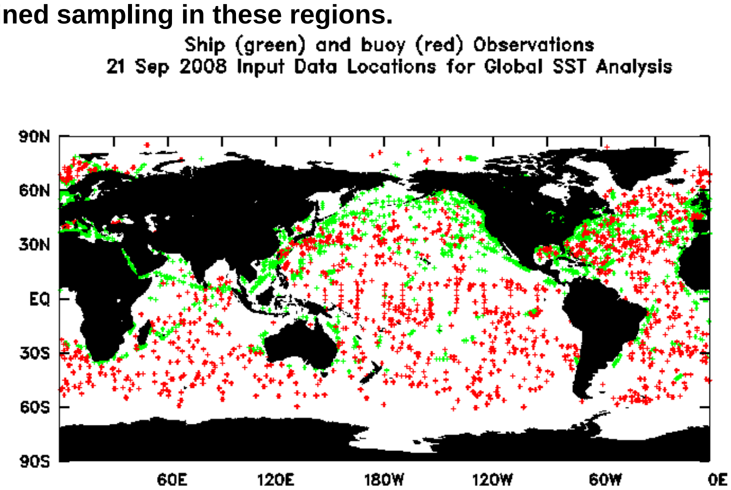


Figure 2. The locations of SST observations from drifting and moored buoys (red) and underway temperature sensors on ships (green) for 21 September 2008, available by 0030 UT on 22 September 2008 on the GTS.

Adequate in situ observations are needed to correct systematic biases associated with satellite retrieval algorithms. These biases occur for both infrared and microwave retrievals. Infrared retrievals are impacted by cloud and aerosol contamination, while microwave retrievals are impacted by precipitation and land contamination. Many of the processes (e.g. clouds, anomalous atmospheric state compared to that used to derive atmospheric correction parameters, atmospheric aerosol, and rainfall) that introduce biases into satellite SST cannot be controlled or predicted. An optimised (Zhang et al., 2006) in situ network is needed to correct large-scale satellite biases. One approach called the Potential Satellite Bias Error (PSBE) has been defined as the residual satellite bias that cannot be further reduced at a given buoy density (Zhang al. 2009). In this procedure, developed only for the operational AVHRR satellite sensor, the number of equivalent buoy (EB) observations in a 500 km grid box must be 2 < EB < 7 per week in order to reduce bias error from ~1K to <0.5 K. SST measurements from ships have greater uncertainty and ~7 ship measurements equal one buoy measurement in the PSBE. Since 2003 PSBE has been calculated operationally and reported quarterly as a global indicator map showing where additional in situ observations are required to reduce AVHRR bias. This map provides NOAA with an objective means of systematically evaluating the effectiveness of the network in achieving an integrated global ocean observing system. The number of drifting buoys reporting SST has increased from ~400 in 1998 to ~1300 today and calculations since 1998 show a systematic decrease is the SST bias from ~0.8 K to ~0.5K. **We recommend that observing system experiments (OSE), sampling studies and, error analyses (such as the Potential Satellite Bias Error, PSBE, Zhang et al., 2009) should be an integral part in the design, development and operation of an integrated SST observing system suited to both near real time operations and climate data record production.**

A relatively new source of *in situ* SST data for satellite validation and calibration are the Argo profilers, delivering vertical temperature measurement profiles from both ascending and descending trajectories to an upper limit of ~5 m depth. Since 2003 the Argo array has developed rapidly and currently has over 3100 active floats deployed in the global ocean providing approximately 300 temperature and salinity profiles, daily (Roemmich et al., 2009). Recent developments have included the use of high resolution temperature and salinity sensors that can be switched on in the upper 25 m of the ocean surface providing SST observations suitable for (a) validation of foundation SST analysis maps, (b) characterisation of near surface stratification (diurnal warming and freshwater surface layers) and (c) the investigation of surface to subsurface thermal coherence. Matches between Argo 5 m SSTs with drifter 1 m SSTs within 1º latitude and longitude and 10 days during 2004 to 2008 have shown a mean difference of 0.002ºC and a standard deviation of 0.36ºC, indicating that Argo data are a good approximation of nighttime subskin SST (Roemmich et al., 2009). The relatively low density of Argo SST observations in any one day compared to drifting buoy observations mean that although Argo profilers may provide a useful augmentation to a calibration/validation (cal/val) SST data set, they cannot fully represent those regions not covered by buoys. The lack of these profiling Argo floats in relatively shallow waters and enclosed basins further limits their use as a cal/val data source. **We strongly recommend that all Argo floats are equipped with the capability to make high vertical resolution measurements of SST in the upper 10m of the ocean surface and that shallow-water Argo floats be developed and deployed.**

New ways of using the satellite SST observations to improve the *in situ* network must also be considered since there is a clear need for high quality *in situ* SST observations with greater timeliness, spatial, and temporal coverage than is currently available. Validation and quality control of the ship SST observations currently being sent to the GTS, additional instrumentation of vessels with reliable, calibrated SST sensors, and ensuring that as much high quality ship SST as possible is transmitted within 24 hours to the GTS are all needed. For example, since 2008, the Australian Integrated Marine Observing System (IMOS: http://www.imos.org.au) Project has enabled accurate, quality controlled, SST data to be supplied in near real-time (within 24 hours) to the GTS from four SOOPs and research vessels in the Australian region, with a further nine vessels planned to be instrumented by the end of 2010, including two Antarctic research vessels.

Enhancing the *in situ* network using satellite-based SSTs is now possible. For example, satellite SSTs are now used to detect erroneous buoy measurements. A network has been implemented under the auspices of the GHRSST STVAL TAG and a tool has been built at the Norwegian Met Service to monitor the comparison of each individual *in situ* measurements against the OSI SAF METOP/AVHRR 1 km SST, worldwide. Large differences are analyzed and, if the in situ measurement is doubtful, it is reported to JCOMMOPS. Such monitoring is complementary to the classical monitoring by comparison with model analysis, and its sensitivity appears to be improved over the classical approaches. Where measurements are shown to be persistently of poor quality, public greylists and for erroneous data blacklists of platform identifiers must be maintained and published on a regular basis. **We recommend that a concerted effort is made to develop a framework to provide robust uncertainty and bias estimates for *in situ* SST data sets based on the use of satellite, model, and climatological reference data sets in collaboration with JCOMMOPS and other Agencies.**

## *3.3. The modern era SST climate data records and link to the historical SST climate record*

We are fortunate to have satellite SST records since 1981 from the AVHRR, from 1992 the (A)ATSR series and from 1997 the TMI data set. These have been complemented by more recent sensors such as MODIS (1999) and AMSR-E (2002) in and several geostationary radiometers. However, even for the long time series sensors, characteristics of the system change. For example, instruments may be replaced, orbits may drift through the diurnal cycle, and instrument channels may fail. Human-induced climate changes are small, when taken over the globe as a whole, compared to the effects of such drifts and discontinuities. In order to minimise the impact of these changes on the climate SST record, “continuity of satellite measurements through appropriate launch and orbital strategies should be ensured” (GCOS, 2006) with at least a year’s overlap between successive sensors in order to develop homogeneity adjustments. In addition, constant sampling within the diurnal cycle should be maintained. Of course, adjustments to maintain homogeneity with other satellite-borne sensors is not all that is required. All SST data prior to 1980 that compose the 150-year record of SST measurements are derived from observations on board ships and from drifting and moored buoys. In addition to “maintenance of complementary baseline *in situ* observations” (GCOS, 2006), which might imply some minimal set to ensure provision of ground truth, comes the need to maintain the full spectrum of this *in situ* record and understand how SST retrieved from satellite instruments relates to it. Indeed, where we have instruments with “rigorous pre-launch instrument characterisation and calibration” (GCOS, 2006), such as the (A)ATSR series, we can use these to understand the limitations of the *in situ* record itself (O’Carroll et al, 2008). All of these activities allow us to identify “random errors and time-dependent biases in satellite observations and derived products” (GCOS, and in *in situ* measurements), which is essential for all applications of the data.

The main problems and progress in the development of multi-decadal SST analyses has been discussed in the white Paper prepared by Rayner et al (OcenObs 09). The focus of the modern-era SST CDR is the preparation of accurate bias free satellite CDR with a full uncertainty specification. The main problems and progress in the development of multi-decadal integrated satellite and in situ SST CDR are:

(i) Quantification of systematic biases and their correction including the characterisation of the changing nature of biases through the multi-satellite record
(ii) Quantifying "random" uncertainty estimates in satellite data using in situ data and how better understanding of high resolution physical processes (diurnal variability, vertical profile of temperature in first few metres) has also helped.
(iii) Understanding how satellite observations since 1981 can be used to help identify and correct various satellite and in situ platform-dependent biases
(iv) Developing better SST measurements in the marginal ice zone SST
(v) Estimation of trends and the effect of analysis methodology,
(vi) Implementation of the GCOS satellite data monitoring principles (GCOS, 2006).

Reprocessing of satellite data sets to include the latest SST retrieval algorithms and understanding of calibration, uncertainty and bias based on inter-comparison and scientific research is an essential step in creating CDR for individual satellite missions. Reprocessing is effective in maximizing the benefit from missions to the widest possible cross section of users. There are several reasons for this. First, consolidated and archived data can be more complete that those available for processing in near-real time. Second, there are always lessons learned that can improve products when applied retrospectively, sometimes very significantly. In the context of SST, lessons may range from re-appraisal of calibration to innovative approaches to cloud detection (e.g. Merchant et al., 2005) or retrieval (e.g., Merchant et al., 2008a). Third, in a reprocessing context, homogeneity of treatment throughout the time series can be improved. Fourth, cross-characterization of sensors via periods of overlap can improve consistency for multi-mission time-series in adherence to the GCOS satellite monitoring principles (GCOS, 2006). These principles provide basic but essential guidance regarding the planning, operation and management of observing networks and systems, including satellites, to ensure that high-quality climate data are available and contribute to effective climate information. **We fully endorse the Global Climate Observing System (GCOS, 2006) climate monitoring principles, which are all relevant to an integrated (satellite plus in situ) SST observing system, and urge agencies to take steps to implement these principles in their short and long term planning for the future of satellite SST data sets.**

Applications of SST for climate are those that most obviously benefit from the additional completeness and homogeneity that reprocessing efforts can supply. Comprehensive error estimation is a complex challenge in its own right that is essential for many climate applications, and which is often only possible retrospectively, since understanding of the error characteristics of a given instrument usually continues to accumulate long after initial cal/val activities are complete**.** The Pathfinder-SST project (Kilpatrick et al., 2001; Casey et al., 2009) is a high-profile example of reprocessing to obtain greater accuracy and consistency in the area of sea surface temperature, applying refined techniques for empirical regression of AVHRR observations to buoy SSTs over the period 1981 to the present day. Microwave SSTs from the TRMM microwave imager and AMSR-E have undergone several reprocessing cycles to take advantage of better modelling of temporal variations in the calibration of the instrument and of the dependence of microwave sea surface emissivity on wind speed and salinity. The ATSR Re-analysis for Climate (ARC) is exploiting improvements in understanding of the infra-red spectroscopy of the sea surface atmosphere and new techniques in cloud detection and retrieval, to create a time-series of SST independent of in situ observations (Merchant et al, 2008b). The EUMETSAT Ocean and Sea Ice Satellite Application Facility have developed new SEVIRI SST processing chains in preparation for reprocessing the SEVIRI archive from 2003 to the present. However several sensors have yet no firm plans for a comprehensive re-processing capability. **We recommend that a systematic framework in which satellite SST data sets can be re-processed is developed and operated. The system should foresee multiple re-processing of L0 (engineering) data through to L2 (geophysical) products to produce the best fundamental Climate Data Record (fCDR) for each satellite sensor. Furthermore, developments achieved for climate re-analyses should be quickly integrated into operations.**

Identification of bias and uncertainty in satellite data is a challenging scientific endeavour that requires careful design and implementation rigour. Global and near-global sea surface temperature analysis products are created using a wide range of statistical reconstruction and interpolation techniques that are applied to data sets from a variety of input platforms. These data sets are subjected to quality control processes, bias corrections, and input from sea ice data as well as *a priori* assumptions. The result of these different analysis routines is a collection of products that can say subtly or significantly different things about the changing climate. Working together with the Global Climate Observing System (GCOS) SST and Sea Ice Working group, the GHRSST Long-term Stewardship and re-analysis facility has developed an initial set of SST analyses covering the modern satellite era (1981-present). These have been incorporated into an online inter-comparison system for the modern satellite era (see http://ghrsst.nodc.noaa.gov/intercomp.html). The initial data set includes analyses derived from satellite-only, in situ-only, and combined input observations. The framework provides value-added access to the analyses as well as several standard diagnostics, including RMS differences, standard deviation, and global and hemispheric anomalies (see Figure 3). To maximize user compatibility, all of the SST analyses included in the intercomparison system are reformatted to GHRSST standards (CF-compliant netCDF with GHRSST-defined content and metadata attributes) as well as Matlab-formatted files. An ultimate goal of the SST/SI Working Group is to connect these modern, satellite-based analyses with the longer historical SST record. Accordingly, the group has incorporated several data sets and SST reconstructions beginning in the 1800s into the inter-comparison framework and computed the standard diagnostics for them as well.

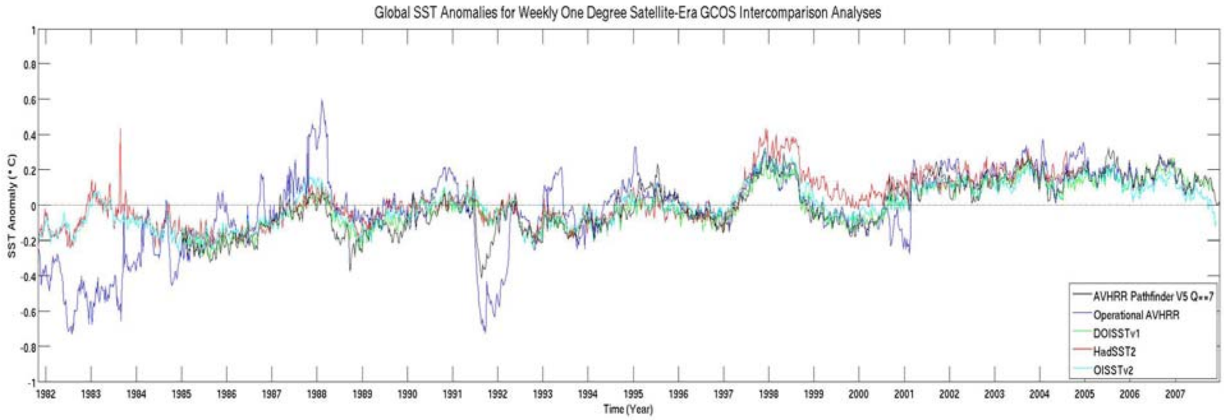


Figure 3. Global SST anomalies for weekly one-degree averaged AVHRR satellite data and various SST analyses highlighting the considerable variability in the operational AVHRR data stream.

In addition to observing and describing differences between SST analyses through the intercomparison framework, the group has undertaken a series of tests designed to understand the origin of these differences. This new understanding will be achieved by performing various quality control and analysis routines on a common set of input SST observations. These tests will result in a set of recommendations of actions and criteria necessary to ensure quality and consistency among SST analyses in the future. Future efforts will also focus on using inter-comparisons to identify and evaluate relative trends among SST analyses. Ultimately this information will be used to gain, and share with users, a better understanding of the differences between SST analyses and the implications of those differences. **We recommend that the GHRSST/GCOS SST & Sea Ice Working Group modern-era SST inter-comparison activity is sustained and enhanced to include other satellite data sets in order to identify systematic biases between complementary SST Climate Data Records.**

Ensuring that high-quality SST climate data records are collected, retained and made accessible for use by current and future generations of scientists and decision-makers is a key objective. The flow of data to the user community and to the International Data Centres is not adequate and, while improvements have been made (e.g. GHRSST), the lack of national engagement and/or resources, restrictive data policies, and inadequate national and international data-system infrastructure are still limiting. It is essential that all SST data be properly archived and managed with simple data access and supporting metadata as they will be reused many times in the future, often as a part of reprocessing or reanalysis activities. Good stewardship of the data also requires that data be migrated to new media as technology changes, be accessible to users, and be made available with minimal incremental costs. **We recommend that further steps are taken to strengthen data stewardship activities for the modern era SST data record including tools that provide easy access to multiple source satellite and in situ data. Tools should also be provided for the regular processing and analysis of combined data sets (e.g., the ESA Grid Processing on Demand (G-POD) system http://eogrid.esrin.esa.int/).** This approach recognizes that the SST Essential Climate Variable (ECV) in the modern era will be constructed from a composite system of satellite and *in situ* measurements and that no single technology or source can provide all the measurements required. Data from ships, buoys, floats, ocean profilers, balloons, samplers, and aircraft, as well as from many satellites instruments are required. Meta-data (e.g.., information on where and how the observations are taken) are absolutely essential to the success of this approach.

**We recommend that a consensus approach is obtained and documented to produce sensor specific SST climate data records (one per sensor/series) in an optimized end-to-end high-resolution global coverage SST measurement system. Also, that an appropriate international cooperative data service is implemented that will develop, maintain and provide SST Climate Data Records.** This is particularly important as the planning horizon for Space Agencies is on the order of 5-10 years. **We also recommend that an optimal level of diversity in approach and competition to achieve the best result is encouraged when implementing SST Essential Climate Varaible processing systems.**

## *3.4. The prospects for global coverage high resolution SST at 1 km accurate to <0.3 K and stable to ~0.05 K decade$^{-1}$ (the SST CDR)*

The traditional infra-red imager for SST having a single-view and channels around 3.7, 11 and 12 μm can demonstrate global 1 km accuracy approaching 0.3 K at night-time (i.e., when all three channels are used), although instability of calibration, cloud-detection failures, and episodes of atmospheric aerosol can each degrade this potential significantly. With only the 11 and 12 μm channels available (as in day time), coefficient-based retrievals have been limited to accuracies of about 0.4 K, even in the absence of such problems. However, recent work on METOP-A shows that optimal estimation techniques can drive accuracy down to ~0.3 K for SST estimates where the retrieval cost is low (Merchant et al., 2008a). All these quoted errors have a random element, but are in large part correlated on the synoptic scales of the atmosphere. However, there is even greater potential in using the system of sensors jointly to reduce errors, rather than privileging a single sensor (which may fail) as a reference. This reduction can be achieved if the differential error characteristics of different sensors can be thoroughly characterized in an optimal and automated way. **We recommend that further research is conducted to develop and refine SST optimal estimation techniques that maximize the error reduction from having multiple complementary observing systems in space.**

The Along Track Scanning Radiometer series and future Sea and Land Surface Temperature Radiometer have an important role as a reference sensor for SST (stark et al., 2007), owing to their stable two-point calibration, low noise detectors and their dual-view, which permits robustness to atmospheric aerosol biases (Merchant and Harris, 1999) that otherwise have a deleterious effect for ~1 yr following major volcanic eruptions. The ATSR series in particular appears to have extremely stable calibration such that re-analysis of the time-series may be able to deliver the challenging target for climate research of a stability of observation of 0.05 K over a decade (Merchant et al., 2008b). In the future, even greater confidence in the absolute stability of calibration may be possible, via periodic in-flight re-calibration of sensors. One approach proposed is to use phase transitions, occurring at precisely determined temperatures, of masses embedded in calibration targets. Such technology will support space-based SST as a climate data record, which is of critical importance. **We recommend that satellite developers include better on-board calibration systems (two managed calibration targets) for satellite infrared and passive microwave sensors, ensure that adequate data are collected on board the spacecraft to characterise instrument calibration, and that these data are telemetered to ground. We also recommend, based on 17 years of experience, that the along-track scanning technique provides an excellent approach to measuring accurate SST from satellites and should be more widely adopted for operational systems.**

## *3.5. Improving SST observations by understanding complementary Wind measurements.*

Although there are many in situ observations from moored and drifting buoys and ships, a truly global coverage is only obtainable from an analysis incorporating satellite borne instruments. Each instrument has its own strengths and weaknesses, depending on the sensor type, platform, orbit, accuracy, delivery timeliness and coverage. The production of an accurate SST is dependent on merging the data from these sources, accounting for the differing measurement methods and accuracies, while maintaining as much of the information in the observations as possible. This fusion of data is a significant challenge for the SST community. GHRSST has made great progress in developing a near real-time system that delivers quality controlled data including uncertainty estimates to the SST analysis community which has reacted positively and developed several new SST analysis systems and/or upgraded existing systems to take full advantage of the GHRSST data streams.

One key issue facing the community was how could infrared satellite data measuring SST in the top ~10μm of the ocean and/or passive microwave SST measurements representing the SST in the top ~1 mm of the ocean be merged with data at 1-5 m depth obtained from buoys and ships? Addressing this problem is critical when thermal stratification occurred in low wind speed conditions, typically with a diurnal cycle. As a result of the absorption of solar radiation in the upper ocean, the sea surface is generally warmer than the overlying atmosphere, with the consequence that heat flows from the ocean to the atmosphere by molecular conduction through a thermal skin layer on the aqueous side of the interface. At infrared wavelengths, the electromagnetic skin depth, the region just below the interface from which the radiation originates, is within the thermal skin layer so the temperature of the emitted radiation is cooler than that of the water beneath the thermal skin layer. The thermal skin layer thickness is generally less than a millimetre, and depends on the turbulent energy density in the water immediately beneath. At microwave frequencies the electromagnetic skin depth is greater than that of the thermal skin layer, and the temperature associated with the microwave emission is a weighted average of the temperature profile beneath the surface. The thickness of the thermal skin layer, and the temperature drop across it, is dependent on several variables, but it appears that a simple parameterization of the skin temperature difference on the surface wind speed is an effective approach (Donlon et al, 2002). It is assumed that wind stress induces sufficient turbulence at the air-sea interface allowing a free flow of heat through enhanced surface renewal from the ocean to the atmosphere with minimal thermal gradients. Observations made by radiometers aboard ships show that the skin temperature deviation from that at depth asymptotes to -0.17 K for wind speeds > ~6 $ms^{-1}$. Making a small adjustment to satellite skin temperature retrievals then allows data to be merged with confidence and is the approach taken by some operational systems (e.g., Stark et al, 2007). The application of a wind-speed dependent skin temperature correction requires the availably of reliable wind speeds at the time of the SST measurement, and since the winds are highly variable, and the skin effect extremely responsive, the time interval between the wind speed estimate and the SST measurements has to be short. However, the dependence of the skin layer temperature difference on wind speed is weak for speeds > ~6 $ms^{-1}$ and the requirement for accurate determination of wind speed becomes less stringent for stronger winds. **We recommend that future satellite and in situ SST observing systems should provide a measurement of contemporaneous wind stress in order to understand the context of SST measurements (e.g., cool skin and thermal stratification), to help blending of SST measurements made by different methods under moderate wind stress and to facilitate the application of SST measurements in scientific and operational systems.**

## *3.6. Improving current and future SST measurements through better understanding of diurnal variability*

Diurnal SST variability complicates the assimilation of SST observations by ocean and atmospheric forecast and prediction systems. It also complicates the derivation of atmospheric correction algorithms for satellite radiometers based on empirical regression against in situ measurements. It is also a significant complication for SST data merging and analysis systems. In many cases today, SST observations are quality controlled for 'diurnal errors/contamination' and yet thermal stratification of the upper ocean on a diurnal time scale is a reality and the challenge is to learn how to identify, quantify and apply such knowledge.

SST changes on hourly time-scales when observed with the accuracy of a few tenths of kelvin. On the global average, there is a diurnal cycle of amplitude ~0.5 K with a minimum around dawn, a rise to a maximum around 2 pm local solar time, and a more gradual reduction in temperature through the late afternoon and evening. This reflects the average response of the ocean surface to the diurnal cycle of solar heating – heat which is predominantly deposited in the upper 10 m of the ocean. However, such an average view masks enormous variability. This is because wind-driven mixing is a major control over the depth and degree of diurnal stratification, and surface winds are, of course, highly variable in space and time. Very large amplitude diurnal warming events have long been reported (Saunders et al., 1982; Flament et al., 1994) from satellite data but remain unobserved in situ (to our knowledge). Recent geostationary satellite retrievals of SST from SEVIRI have suggested diurnal warming events in excess of 6 K at the ocean skin while no in situ observations have been found to directly validate these observations.

Their occurrence has been put beyond doubt by analysis of hourly SSTs from SEVIRI, over the Mediterranean region (Merchant et al., 2008c) and also in a multi-sensor study of the open Atlantic ocean (Gentemann et al., 2008). Much more common is the diurnal warming of 1 or 2 K in low-wind (< 2 $ms^{-1}$) areas during the day. Such warming features can have spatial coherence over scales of 10 to 1000 km, see Figure 4. Due to their high repeat cycle, geostationary satellite SST fields are indispensable to study diurnal warming. It should be recognised that stratification effects also impact in situ observations in the upper few meters of the ocean surface but due to the lack of information on the depth of measurement and the more complex thermal stratification in the upper 1 m layer, the impact on these measurements is largely undetermined. **We recommend that a systematic data gathering exercise is performed to identify and catalogue diurnal SST warming events across the global ocean using satellite and in situ measurements in order to better understand the spatial and temporal variability of such events and their impact at daily, monthly seasonal, annual and multi-annual timescales.**

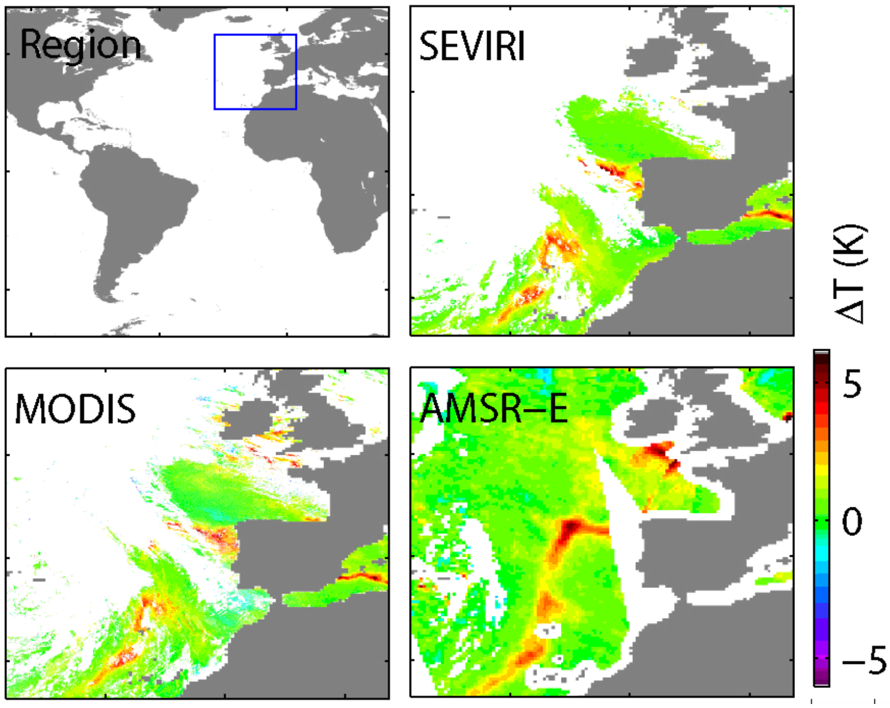


Figure 4. Examples of a SST diurnal warming event as viewed from different satellite sensors in the NE Atlantic Ocean.

Sub-daily variability may modify air-sea fluxes to a climatologically significant degree (Curry et al., 2004). It is also a source of asymmetric (and intermittently large) "error" in satellite SSTs when those SSTs are used as an estimate of foundation, mixed-layer or bulk SST. Estimating diurnal SST variations is significantly limited by imperfect knowledge of the locations and persistence of wind speed minima (winds <4 $ms^{-1}$) and surface heat fluxes. However, the sampling of SST through the day from satellite also contains direct information of the diurnal SST variability (within some error). This implies that the most fruitful approach to estimating the daily cycle will be analysis, in the sense of using both models and observations together to create a time-resolved optimum estimate. One limitation to this approach is the availability of high temporal resolution wind estimates from NWP systems. **We recommend that steps are taken to secure high temporal resolution (ideally at an hourly resolution) wind fields over the global ocean for use in diurnal SST variability modelling.**

### 3.6.1. The need for improved measurements to characterize and resolve global diurnal variability on a daily basis

With the emergence of diurnal variability as a significant source of uncertainty in the production of SST analyses, the desire to provide detailed error estimates with all satellite SST retrievals, and the need to resolve diurnal variations in SST for applications such as flux estimation, it is important to obtain improved in situ measurements of diurnal warming. Models and analyses of diurnal warming in the near-surface temperature are being developed and implemented to estimate the amount of diurnal warming present in individual SST observations and to enable daily analyses of the foundation SST to be referenced to specific depths and time of day. To this point, however, there are still relatively few detailed in situ observations of significant instances of diurnal warming and the corresponding environmental conditions with which to completely validate the accuracy of these models and approaches. To confidently apply these techniques, additional observations are critically required.

Between the skin SST to which satellites are sensitive and the depth at which diurnal variations are small, there may be significant stratification over depths of 0.1-1 m. To understand the depth-related differences between satellites, drifters and moored buoys, it is important to have at least the nominal depth of observation routinely associated with reported in situ SSTs. Buoy observations, while relatively extensive spatially, may be at depths too great to observe extreme warming. Cruise-based radiometric and near-surface observations capable of resolving this warming are too limited in extent and number to reliably capture such relatively infrequent events. Potential solutions include increased numbers of near-surface temperature observations at well-characterized depths such as being pursued with modified Argo floats and increasing the number of radiometric sea surface observations from ships and aircraft. Coincident estimates of the winds and heat flux are also important to enable the validation of methods to globally predict and analyze these warming events. Experimental modifications to Argo floats are demonstrating the ability to sample the SST and salinity profile within the upper 10 m. If such observations become widespread and reliable, this could give a secure observational basis for understanding diurnal variability as seen by satellites relative to foundation SST and SST as observed by in situ means. **We recommend that evidence is gathered and presented to Argo float manufacturers highlighting the applications and benefit of providing high resolution temperature sensors on all Argo floats for the investigation of SST diurnal variability.**

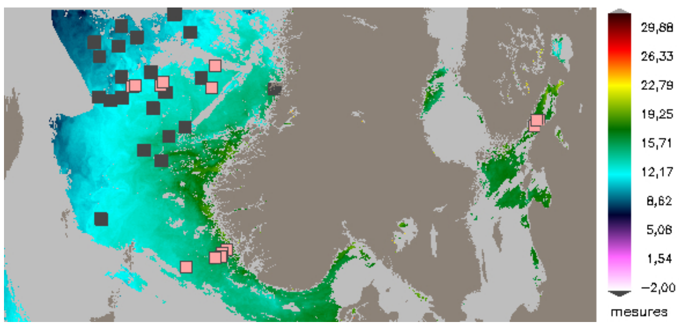


Figure 5. In situ measurements (pink: ship, black : drifters superimposed over a METOP SST full resolution granule during a highlighting poor quality data in July 2008 in the Norwegian Sea.

Figure 5 represents the in situ measurements coincident with a probable diurnal warming event recorded by METOP in July 2008 (on going OSI-SAF study). It illustrates the difficulty of monitoring such events with in situ measurements, even in an area with a dense network of drifters or ship measurements.

## 3.7. *Improving current and future SST measurements through better uncertainty and error estimation*

A key user request from the user community is the provision of uncertainty estimates to be attached to each pixel in SST satellite measurements, to each observation for in situ measurements (including the depth of measurement) and for each grid point in SST analysis products. The GHRSST Science Team agreed that error (bias) and uncertainty (rms variability) should be the basic quantities provided with each measurement or grid point. GHRSST adopted a framework of Single Sensor Error Statistics (SSES) designed to take into account uncertainties for specific instrument/platforms (e.g. AATSR has a specific parallax effect between forward and nadir views inducing an error on the back side of cloudy areas in the flight direction and passive microwave data are contaminated by rain and contamination in close proximilty to land (side-lobe contamination)). Bias and uncertainty estimates have been largely derived from near contemporaneous match ups between satellite and in situ SST measurements which are periodically analysed to provide SSES. Furthermore, SSES are currently sub-divided by objective assumptions (e.g., subdivision based on the clear skies, proximity to the nearest cloud, or proximity to land and rain flags, features of the dual view technique and features specific to microwave sensors) to provide further structure for analysis of MDB data (although in future more analytical approaches are foreseen). Users need an estimate of the satellite SST estimate at pixel levels. The EUMETSAT OSI-SAF developed a statistical method to derive, from their MDB, the SSES bias and standard deviation associated to each of the confidence level assigned to the retrieved SST estimate. The confidence levels assigned to each pixel are based on tests to the reliability of the cloud mask and the SST algorithm conditions. Figure 6a shows an example SSES map for the MSG SEVIRI and accompanying SST image data. Figure 6b shows the relationship between confidence level, bias and uncertainty for the METOP sensor in the North and South Pacific Ocean where SSES are shown to have a different character. The confidence levels from 2 to 5 reflect in general the level of the SST quality (increasing from 2 to 5) but as shown the confidence level to bias and standard deviation characteristics has also a regional (and seasonal) meaning that needs to be accounted for.  An alternative approach called the Hypercube has also been developed based on a match-up data base for the Aqua and Terra MODIS sensors. In this case, the MDB includes near-contemporaneous, co-located satellite brightness temperatures, in-situ buoy and radiometer SST, auxiliary data from model or satellite observed fields, and the satellite viewing geometry. All satellite data have been matched using drifting buoy and the Marine Atmosphere Emitted Radiance Interferometer (MAERI) observations. A series of quality tests is been applied during processing of the MODIS data to identify cloud and dust aerosol contaminated retrievals and assign pixels to one of four different quality levels with quality 0 being the best quality possible. The relative immunity of the MODIS 3.95 μm and 4.05 μm bands to both water vapour and aerosol aerosols as compared to the increased sensitivity to both in the MODIS 11μm and 12 μm bands is used to identify aerosol data. After eliminating records with quality levels greater than 1, each match-up database is partitioned into a multi-dimensional array with the following 7 dimensions: time by season (4 values), latitude bands (5 steps in 20 degree increments from 60S to 60N), surface temperature (8 increments in 5 degree steps), satellite zenith angle (4 increments), brightness temperature difference as a proxy for water vapour (4 intervals for 4 μm and 3 intervals for 11-12 μm SST), retrieved satellite SST quality level (2 intervals) and day/night selection (2 intervals). The bias (satellite-in situ) and standard deviation are then computed for each element to define a hypercube look up table (LUT).  The LUT is then used during satellite data processing to predict the SSES bias and standard deviation of the SST retrieval.  **We recommend that Single Sensor Error Statistics (SSES) should be stratified by season and region and include a variety of auxiliary data that can be used to derive the SSES in a hypercube (multi-dimensional) matchup database approach.**

The role of an optimised network of in situ high quality SST observations with attendant metadata describing their own in situ error and uncertainty characteristics that are reported in near real time in an operationally sustained manner is obvious.  The network should ideally include high resolution (at least 0.5 m) vertical profiles of the upper 25 m of the water column, ship mounted radiometers in optimised transects to maximise satellite overpass collocation, improved SST observations from ships and the drifting buoy network. **We note the need for better in situ SST observations and supporting data and metadata (including depth of observation and calibration history) for use in Single Sensor Error Statistics (SSES) derivation schemes.  We recommend that (1) operators of in situ radiometer systems are encouraged to share their observations in order to maximize the benefit to the satellite SST community and (2) all in situ observations when reported should include depth, calibration history, measurement method and a meaningful estimate of uncertainty.**

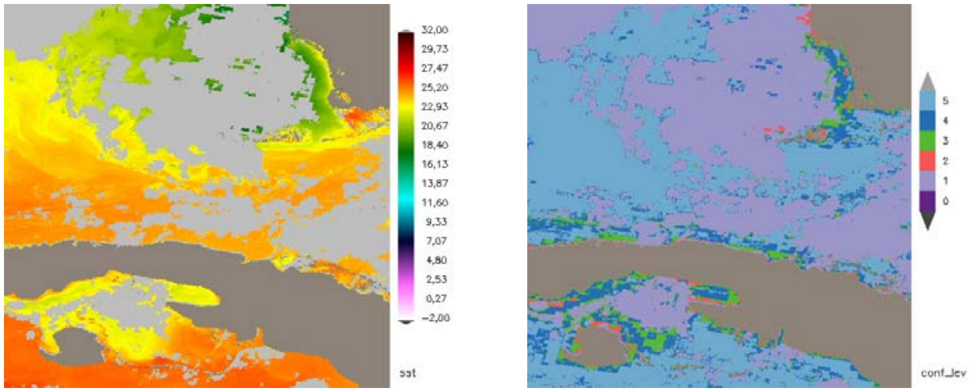


Figure 6a Left: SST over the Florida Strait on the 2nd of February 2009 at 0249 UTC , right: the corresponding confidence levels reflecting the reliability of the SST calculation.

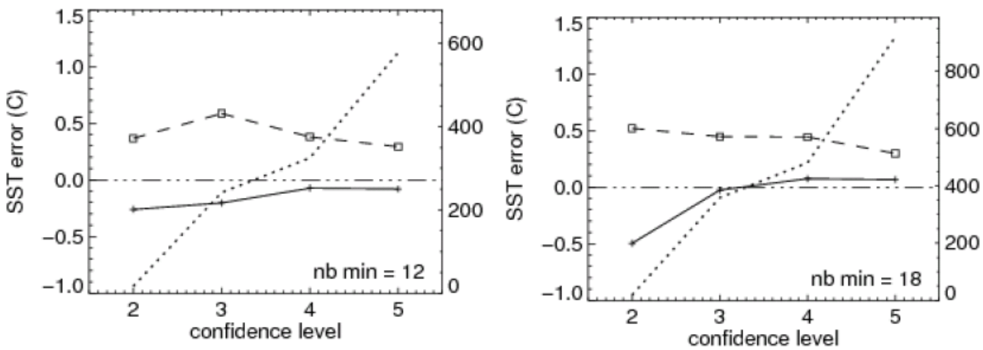


Figure 6b. Left: Error characteristics of METOP/AVHRR SST in January 2009 as a function of the confidence levels; left: over the north Pacific; right: over the south Pacific. Continuous line: bias; long dots: standard deviation; small dots: number of cases

GHRSST established several activities in partnership with international satellite data providers to develop and implement SSES and has made excellent progress in achieving a basic approach. However, many of the SSES retrieval schemes are limited by the availability of in situ data matched to satellite data and almost exclusively focus on the SST value. In a more holistic and analytical approach the SSES scheme should be extended to consider and report specific uncertainties based on:

1. Knowledge of instrument measurement uncertainty (channel noise determined from on-board calibration blackbodies)
2. Knowledge of geo-location uncertainty describing registration errors and geo-location errors provided in a useful format.
3. Knowledge of geophysical retrieval uncertainty (Error, precision)
4. Knowledge of long term stability
5. Selection and use of an appropriate geophysical reference data sets (e.g., in situ observations from radiometers, spectrally homogenous target areas etc.) linked to international standards where possible.
6. Knowledge of the data processor system and associated data (e.g. LUTs, correction schemes, filter function definitions, calibration/model versions) performing SST data processing.

Furthermore, uncertainty values should be, where possible, traceable to accepted international reference standards and SI units. For example, the SLSTR on-board calibration blackbodies should be traceable to the same reference standards that are used by the in situ validation teams. Satellite instruments should be traceable to the same reference standards. Ground truth instrumentation and follow on missions should all reference the agreed international standard.

An important and recent development relevant to the standardisation and evolution of 'best practice' for satellite validation, initiated by the Committee on Earth Observation Satellites (CEOS) Working Group on Calibration and Validation, is an initiative to introduce a more formal level of Quality Assurance for Earth Observation data (designated QA4EO). The objective is to introduce guidelines for the collection of data both from space and from in *situ* instrumentation based on the adoption / adaptation of existing "best practises". QA4EO recognises that in order to achieve harmonisation and facilitate interoperability between data type and sources, two key elements need to be considered:

- A communication / data policy
- A means to ascribe a Quality Indicator (QI) (based on a quantifiable metrological / statistical based measure) to a Knowledge Information (KI) product.

An important component of QA4EO is the inter-comparison of instrumentation used for validation and to introduce a degree of traceability to the often diverse validation observations which are used. **We recommend that the Committee on Earth Observation Satellites (CEOS) Quality Assurance for Earth Observation data (QA4EO) approach is adopted by Space Agencies working together with inter-Governmental agencies, in situ data providers and the scientific SST community when defining, implementing and operating a sustained SST validation program.**

**We recommend that a concerted effort to develop better SSES uncertainty and error estimates for all satellite and in situ data sets should be a priority and urge agencies and scientists to devise new schemes that can be implemented in an operational manner. We recommend that uncertainty estimation follows and, where appropriate, contributes to the Quality Assurance for Earth Observation data (QA4EO) approach both for satellite data and for in situ data.**

There are many SST data products, now available under the common GHRSST formats (L2P, L3, and L4). Users often want more than one source for their application over a given region, and quite often observe inter-sensor biases which result from the variety of sources and processing methods. The inter-sensor bias correction is a mandatory step for the analysis schemes (such as ODYSSEA and OSTIA), but this step is also useful for other users merging data of various origin. It will be also part of any global reprocessing effort. A first inter-sensor bias correction method has been developed based on correcting all sensor biases against the ENVISAT AATSR derived SST (e.g., Stark et al, 2007). As already noted, the AATSR is well characterised by design and concept to provide and accurate and consistent SST reference data source. This method has proven efficient at the Atlantic scale although working over smaller basins such as European seas in winter raises another challenge as cloud-free AATSR data are rare, and cloud coverage and proximity of land/ice poses many problems to the other sensors. Comparison/correction with AATSR data offers also a synoptic view of each sensors regional biases, that gives also guidelines to correct the algorithms or the production scheme right at the producer level. **We endorse the approach in which well characterised and accurate satellite sensors may act as a reference data source for bias correction and SSES derivation but note that steps should be taken to assure the sustained provision of such data (e.g., the ESA Sentinel-3 Sea and Land surface Temperature Radiometer (SLSTR) that will follow the Advanced along Track Scanning Radiometer (AATSR) due for launch in 2012). We further recommend that new satellite infrared radiometer systems make full use of additional channels in the 3-12 μm spectral region.**

## *3.8. Improving current and future SST measurements through better validation*

In the case of SST observations, *in situ* measurement are and always will be needed, to be used in conjunction with the satellite data. This is a fundamental characteristic of the modern era SST record. The reason for this is that validation is actually a continuous obligation. Obviously, in the early days of a satellite mission, an intense and, it is hoped, short validation campaign is essential in order to verify that the sensor is performing according to its design requirements. However, when data are used operationally for daily weather predictions, or when they are applied to the long-term monitoring of climate processes the demand for both accuracy and stability can only be met if validation is undertaken as a continuous and necessary commitment. This is the only way in which the effects of, for example, anomalous instrumental drift or unusual geophysical situations, can be detected in a timely manner. Also, it is only through long-term validation programmes that subtle limitations in the treatment of the data can be detected and the appropriate refinements undertaken. Thus, validation not only serves to assure the current levels of quality but it is also an essential prerequisite to developments and refinements that lead to the improvements which are always needed and are generally beneficial when the user needs for accuracy are always increasing. There is a growing requirement from the community that validation should be treated as a generic activity in its own right and, rather than be conducted on an instrument-specific basis, carried out on the basis of the parameter observed. In other words, for a parameter such as SST, there are multiple data sources which increasingly are being merged in a highly productive way and there is therefore a need for a continuous quality-assured programme of validation, independent, ideally, of the priorities and needs of individual sensor operators. The problem of organising and funding such generic programmes are an important challenge for inter-governmental agencies, Space Agencies and the scientific community. **We recommend that Space Agencies working together with inter-Governmental agencies, in situ data providers and the scientific SST community take steps to define, implement and operate a sustained validation program for all satellite radiometers producing SST for the lifetime of the mission. We also recommend that Space Agencies and in situ data providers include sustained validation of their SST products as a fundamental component of each satellite mission and in situ platform for the entire mission/instrument duration.**

It has long been apparent that the calibration of satellite-derived SST by comparison with in situ near surface thermometric measurements of SST from ships and buoys cannot avoid bias and uncertainties associated with the variable thermal structure in the upper meters of the ocean (Donlon et al 2002; Robinson & Donlon, 2003). During the last decade it became imperative to provide ground-level radiometric measurements of skin SST to enable “like-with-like” validation of new satellite products such as (A)ATSR that explicitly retrieve the skin SST. A number of radiometric systems have been developed to meet this need (see for example Minnett *et al*., 2001; Jessup *et al*.,2002), some of them capable of autonomous operation on ships of opportunity (Donlon *et al*., 2008). Ship-mounted radiometers are now deployed routinely in operational validation exercises, contributing a unique element of the validation of AATSR SST products (Corlett *et a*l., 2006). Moreover a series of inter-calibration experiments for ship-mounted SST radiometers have been held (e.g. Barton et al, 2004) and plans are in place to ensure that ship radiometers and the laboratory black-body systems used to validate them (Donlon et al 1999) are regularly referenced to the same internationally recognized thermal infra-red radiometric standards as the satellite sensors themselves. These data can be used to bridge across gaps in the satellite SST record (e.g. between AATSR and Sentinel-3 SLSTR) if required (Kelliher et al, 2007). **We recommend that an optimum number of in situ high-quality infrared radiometer systems are maintained for the purpose of infrared and passive microwave satellite validation building on the demonstrated capability during the last 10 years. Such systems should be augmented by better aircraft measurements that have the additional benefit of covering large transects although care should be taken to ensure that atmospheric attenuation of the sea surface signal does not compromise the accuracy of data when flying at altitude.**

One of the challenges facing the satellite SST community in terms of validation is a wide variety of approaches, techniques and data that are used to validate SST data products. Numerous sources exist for in situ data at global scale (from e.g., GTS to CORIOLIS) or local scale (private instrumentation) making it possible for satellite SST providers to monitor and validate the quality of their data, relying on satellite to in situ match-up databases. However building such a database is challenging as:

- sources of in situ data differ, resulting in a very different sampling or comparing to a different reference (e.g. different depth)
- correction, processing (space or time averaging) and quality control of these in situ data may also be very different from one centre to the other
- the matching process may also differ, using different time/space constraints, preferring time proximity over distance, etc...
- the delay for the update for these databases may extend from delayed mode (once a year, every 6 months,...) to real-time, resulting sometimes in not up-to-date error statistics or late detection of sensor problems

These discrepancies between the match-up databases existing at SST providers premises may result in inaccurate or heterogeneous sensor error estimations. Yet no single quality controlled source of in situ data matched to satellite data using a common approach has been achieved complicating the comparison of quoted accuracy and uncertainty between different satellite data sets. GHRSST has initiated a common satellite and in situ matchup database approach to address this issue but progress has been slow to date – largely due to the lack of resources required to translate data providers heterogeneous matchup databases into a common format. **We recommend that an international reference match-up database of satellite and in situ observations that applies the same control and process methodologies to the match-up process and input data is developed, maintained and operated in order to allow a fair, consistent and up-to-date estimation of errors on SST retrieval. We note such a common database can only be constructed following the systematic re-processing of satellite data as part of a CDR development activity and should be consistent with the CEOS QA4EO process.**

Within the modern-era SST record, there is already a wide variety of global spatially-mapped SST products available for scientific use and operational applications. The different types can be summarized as (a) those retrieved at level 2 from individual satellite sensors, (b) level-3 products created by averaging level 2 data over space-time, (c) level 4 products generated by optimal interpolation analyses of multiple observation sources, and (d) the dynamically consistent SST outputs from ocean GCMs that assimilate satellite SST observations at levels 2, 3 or 4. We may expect that continued scientific effort will strive to improve the quality of each of these SST product types in the future, to achieve higher accuracy and/or finer spatial and temporal resolution. In order that development effort for new SST products is managed efficiently and effectively so that it meets the needs of the data users, it is important to be able to readily assess the quality of the different data products in relation to a set of metrics such as accuracy, timeliness of delivery and the validity of the error statistics assigned to each observation. **We recommend that the SST community of producers and users establish and maintain:-**

- **A programme of *in situ* measurements, both thermometers on buoys, ships and subsurface vehicles and radiometers on ships and platforms that can be used for validating the different products. For the sake of efficiency It is desirable that this be a fully collaborative programme shared between all the agencies responsible for SST products. There is also a need to specify the requirements for new *in situ* data acquisition systems to support data integration, including wider coverage by ship-based radiometers, diurnally resolving moorings, Argo with additional sensors for near-surface sampling and the OceanSites approach.**
- **A forum in which sets of quality metrics relevant to specific uses of SST data are defined and regularly reviewed (consistent with QA4EO). Different metrics will need to be defined for different applications, for example timeliness and reliability of error statistics may be more important than accuracy for assimilation into forecasting models. whereas absolute accuracy and stability are most important for climate datasets.**
- **Tools such as the high-resolution diagnostic dataset (HRDDS) described above, that facilitate comparisons between different SST data products. New developments should include automated comparisons between different SST products to alert producers when inconsistencies are detected.**
- **Match-up databases (MDB) which relate satellite-derived SST products with in situ measurements of SST. It is essential to preserve independence between those *in situ* observations used for calibration of SST products and those used for validation of the sensor specific error statistics attached to each data product, for which a new set of tools needs to be developed. Moreover, for long term climate monitoring it will be important to establish a “protected MDB” that can be preserved as an independent validation reference for future generations of SST product re-analyses. Ideally such data should not be allowed to influence the evolution of product algorithms in order to preserve their independence.**
- **A generic system to acquire better feedback from data assimilation systems (e.g., observation weights rejection statistics) in order to ascertain how much influence different SST observational products have on the observing system models. Such information will identify targets for the quality of satellite SST data products, necessary for them effectively to constrain ocean models.**

We note the recent development of the Global Space Based Satellite Inter Calibration System (GSICS, see http://www.star.nesdis.noaa.gov/smcd/spb/calibration/icvs/GSICS/) which has been developed by operational agencies to assure high-quality, inter-calibrated measurements from the international constellation of operational satellites to support the GEOSS goal of increasing the accuracy and interoperability of environmental products and applications for societal benefit. GSICS will improve the use of space-based global observations for weather, climate and environmental applications through operational inter-calibration of the space component of the WMO World Weather Watch (WWW) Global Observing System (GOS) and Global Earth Observing System of Systems (GEOSS). The basic GSICS strategies to achieve this goal are:

- To establish a GSICS Virtual Library to efficiently share information, software and data relevant to calibration;
- To build collaborations ensuring that each satellite instrument meets specifications by making pre-launch tests traceable to SI standards;
- To improve on-orbit calibration of satellite instrument observations by means of an integrated cal/val system, including instrument performance monitoring, inter-satellite/inter-sensor calibration, lunar and stellar calibration, vicarious calibration and validation with reference sites;
- To establish a distributed research component and a plan for research to operations transition;
- To build collaborations to retrospectively re-calibrate archive satellite data using the operational inter-calibration system in order to make satellite data archives worthy for NWP forecasts and climate studies.

At present GSICS is focussed on the operational atmospheric community and **we recommend that the ocean SST (GHRSST) community work effectively with the Global Space Based Satellite Inter Calibration System (GSICS) and establish if and how it may assist GSICS for the benefit of improved SST measurements by working together (e.g., using the High Resolution Diagnostic Data Set (HR-DDS) and GHRSST Multi Product Ensemble (GMPE) tools).**

### 3.8.1. Improving current and future SST measurements through better monitoring

The increase in satellite and in situ SST measurements has facilitated an increase in the number of additional data products and analyses which are operational or under development. Analyses are computed over a variety of regions and time periods with different spatial and temporal resolutions. Users have a choice of analyses that was never possible before GHRSST was established. The problem for any potential user today is not to obtain an SST analysis but to find the *best one for that particular user*. An analysis is a gridded field produced from irregularly spaced data. Analyses can be produced at any temporal resolution for any spatial grid. However, as the resolution increases, the noise also increases. Of course as more high resolution satellite data becomes available, the analysis resolution can be increased. However, in monthly in situ only analyses differences exceeding 1 K were significant; recent daily analysis differences can exceed 5 K. An example of the daily difference can be seen in Figure 7 which shows regional SST maps for 1 January 2007. This region was selected just off the Carolina Coast because in winter the warm Gulf Stream is found off shore while colder shelf water is present between the Gulf Stream and the coast. The colder shelf water is evident in the daily OIs, the FNMOC and the RTG-HR analyses, but weaker in the RSS and OSTIA analyses. In addition the scale of the details is highest in the RSS analysis and lowest in the OSTIA. Each group providing analyses has many choices. One of the first is which type of satellite and in situ data should be used. Then choices have to be made on the frequency of the analysis and the grid type and resolution. These are just some of the choices. All the choices are not always easy to make. Furthermore, they are not always well documented. The important point is that the choices are nonlinear. Thus, differences among analyses can potentially be large. **We recommend that SST analyses producers provide a user oriented document (to a standard template) that describes each analysis product and the choices and assumptions that have been made for the analysis procedure. The document should highlight the strengths and weaknesses of the analysis output in order to help users use the data in the most appropriate manner.**

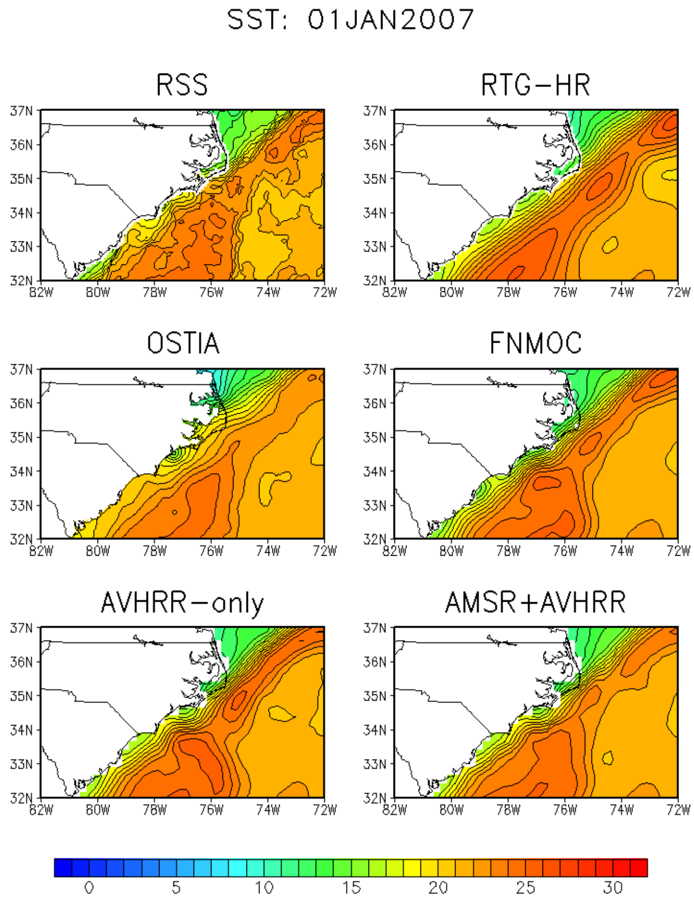


Figure 7. Six daily SST analyses for 1 January 2007. Note in particular differences off the South Carolina Coast (33°, 80°W).

Several solutions are already being pursued. One method is to simply take an ensemble of the products and select a best estimate based on medians. Another is a web-based functional interface which allows inter-comparisons of SST products at selected points. The GHRSST Multi-Product ensemble (GMPE) This system uses ensemble techniques to investigate SST analysis differences using both analyses and observational products. The GMPE allows operational agencies to assess the relative performance of their L4 product against a 'consensus' standard L4 SST product. Each day a median average SST map from 10 international operational global coverage SST analysis products is calculated after their differing analysis grids have been homogenised by area averaging onto a standard 0.5° lat/lon grid. The GMPE ensemble is updated each day and output data sets are available together with a variety of diagnostic and anomaly plots on a daily basis. An example of the GMPE output is given in Figure 8. The GMPE service provided by the Met Office is an extremely useful to users and producers of SST analysis systems based on the number of registered service users and user feedback. **We recommend that GHRSST Multi Product Ensemble (GMPE) capabilities are extended to provide quantitative data outputs in a GHRSST format for use by the general ocean community, consider higher spatial resolutions for the system grid specification, the inclusions of automated procedures for statistical analysis of the ensemble and the creation of regional GMPE tools for specific and challenging regions (e.g., Tropical Pacific, Bay of Bengal and Indian Ocean, Indonesia, Arctic Ocean, Great Lakes). These services could be distributed across multiple centres.**

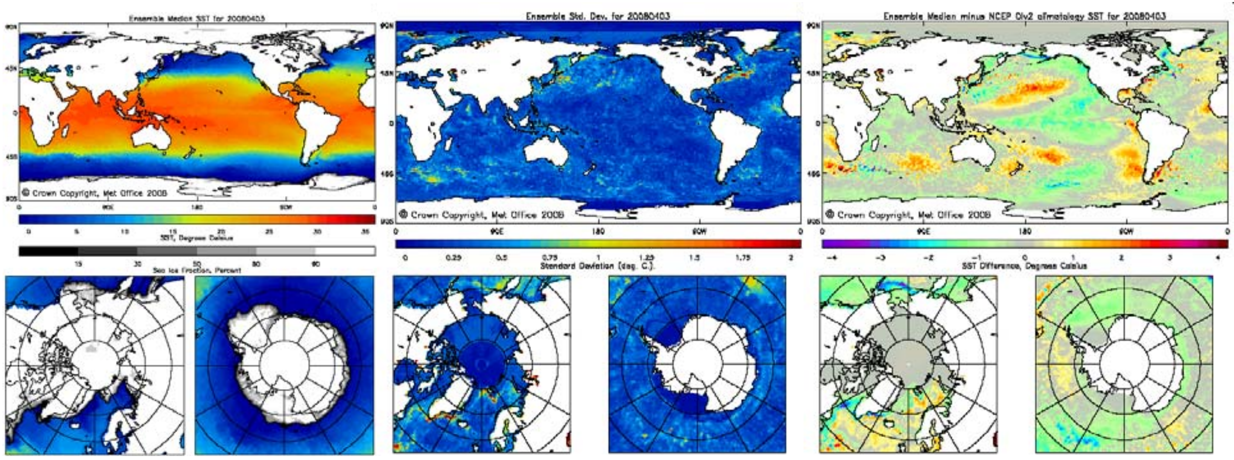


**Figure 8** The GHRSST Multi-Product Ensemble (GMPE) Left median SST ensemble, centre: median Ensemble STD, Right: Ensemble median minus NCEP OIv2.0 climatology for 3rd April 2008. http://ghrsst-pp.metoffice.com/pages/latest_analysis/sst_monitor/daily/ens/index.html

The GHRSST High Resolution Diagnostic Data Set (HR-DDS) is a component service conceived within GHRSST allowing users to interactively view, compare and analyse observational and merged SST data products, ocean model data sets and auxiliary data sets from the various data streams within GHRSST. The HR-DDS system consists of regularly gridded subsets of all available GHRSST SST within predefined small sites (typically 2º by 2º in size). Examples outputs of the HR-DDS web site are shown in Figure 9. There are approximately 250 sites chosen to provide evenly distributed global coverage and allow detailed examination of the effects of specific atmospheric or oceanic conditions. Sites are included to represent regions affected by Saharan dust aerosol, or areas of high spatial and temporal SST variability, such as western boundary currents. The HR-DDS ensures that both operational users and scientists have ready access to information, in a well defined format tuned to specific areas and issues, that can be used to diagnose faults and data problems immediately. Also included in the HR-DDS are model SST outputs so that direct comparison between model and measurement data can be made. Work is underway to link the HR-DDS with the GHRSST Match-up Data-Base (MDB) system to create a complete set of web based interactive diagnostic tools that better characterise both the magnitude and sources of errors not only in satellite derived SST fields but model and analysis system outputs as well. Tools are provided with the HR-DDS portal which allow for download of comma separated value (CSV) files or TAR archives containing only information chosen by the user. The HR-DDS is being extended to include capabilities that allow the service to be properly integrated into the daily work of SST scientists and operational systems by providing a set of user-configured event-driven indicators that can provide an automated response to a given event. For example, a user may wish to monitor the performance of a given satellite sensor and may request the HRDDS system to automatically analyse the data set for events when the sensor SST values exceed the median average of all others available within 6 hours of the measurement. In a different scenario, a user may wish to monitor for diurnal variability events and request the system to analyse for all events where the surface wind speed is less than a threshold value and elevated SST is observed during the local afternoon. The potential for such an event-driven system is large if performed in near real time as this will alert SST scientists to potential problems with the observing system as well as identify problems with analysis systems. **We recommend that the GHRSST High Resolution Diagnostic Data Set (HR-DDS) is extended to include user defined quantitative analysis and event driven monitoring capabilities in near real time to monitor the quality and performance of the SST observing system. Reports should be automatically sent to data providers and users when an event of interest is encountered by the HR-DDS system. Furthermore, the HR-DDS should be expanded to include long time series of SST data sets to be visualised for reanalysis purposes.**

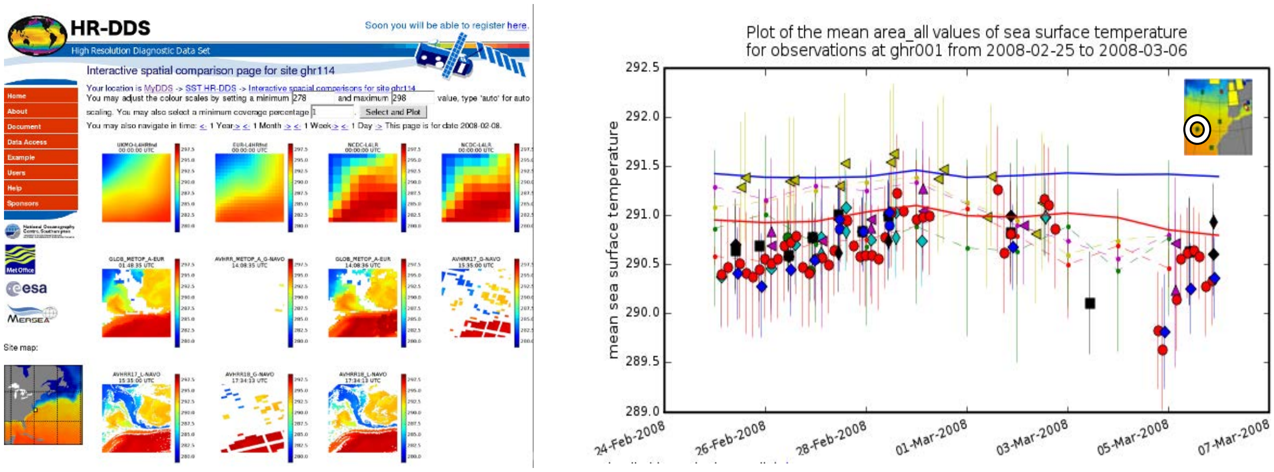


**Figure 9** Examples of DDS images(left) and time series of observational data for a variety of satellites and model SST outputs (right). The HR-DDS provides updates to all data sets in near real time. Solid lines indicate model outputs, dotted lines indicate analysis data sets and large symbols measurements. See http://www.hrdds.net

A second capability of the DDS data base is the ability to compare high spatial resolution images of SST from different satellites that have been obtained at the same location and close in time. An example of this is described by Barton (2007b) in which SST images over Australia's Gulf of Carpentaria were obtained from a range of satellites. By comparing carefully co-located images detailed information of the differing performances of the satellites can be extracted. An example is shown in Figure 10 that provides collocated SST images form AVHRR, GLI and AATSR obtained within the same four hours. The Figure also gives a histogram of the differences between the SST values at each 0.01 degree location (note that the HRDDS allows direct access to common gridded data and difference maps on-line at a resolution of 0.05 degrees as standard). In this example, the AVHRR data are for approximately 0300 local time with AATSR and GLI being four hours earlier. All three images show cooler water in the south-east of the area. The two difference histograms show a sharp peak indicating low variations in the SST differences. The spread of the histograms is consistent with the inherent spread in the SST values from applying the SST algorithms to the digitized brightness temperatures in each instrument channel. The AVHRR-AATSR difference peak is between 0.2 and 0.3 K which is expected as the AATSR algorithms are derived to give the radiometric (skin) temperature while the AVHRR algorithm gives the bulk temperature. The minor striping in GLI that is aligned close to lines of latitude are due to scan mirror and detector array effects in the instrument. Some secondary diagonal striping is thought to be due to electronic noise in the instrument electronics. Both striping effects are minor anomalies and the quality of the GLI SST images is very high – especially when averaged over a small array of 1-km pixels.

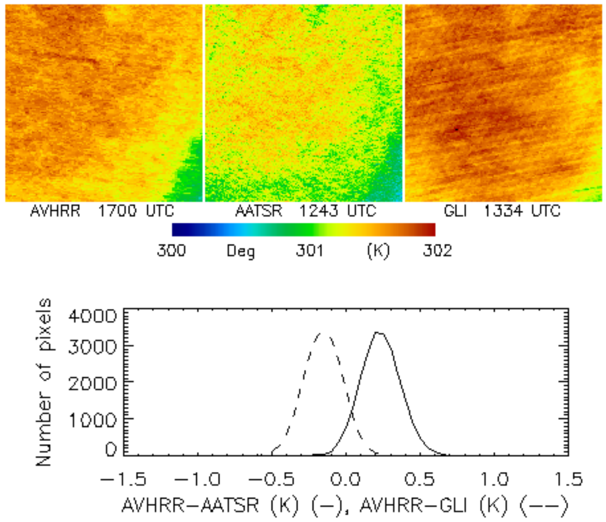


Figure 10. Upper panel: Collocated SST fields derived from AVHRR, AATSR and GLI for the night of May 25, 2003 in the Gulf of Carpentaria. Lower Panel: Histograms of the differences between the derived SST fields displayed. The instrument and time of data collection are included below each image.

At this time there are also inter-comparisons done by one or several investigators for individual technical or scientific publications resulting in a situation where there are several groups working on different systems. A more optimal solution would be to combine or link and extend these comparisons so that they are mutually beneficial and **we recommend that a high resolution inter-comparisons group is established under GHRSST that would be similar to the SST inter-comparison group already established for historical SSTs within the framework of GCOS. This group would include representatives of each of the major centres creating high resolution SST analyses with terms of reference established by the GHRSST science team.**

## *3.9. Better cloud detection in satellite data without loss of data at strong SST gradients*

Satellite-derived SST fields obtained from infrared sensors are corrupted by clouds, with the temperature of cloud contaminated pixels generally colder than the actual sea surface temperate at those pixel locations. Inclusion of these pixels renders the data inaccurate and difficult to use. For these reasons, the flagging of cloud contaminated pixels in SST fields based on satellite-borne infrared sensor observations has received a great deal of attention over the past 30 years. Despite the effort devoted to such algorithms, significant problems exist with the current state of cloud detection. This difficulty stems from the broad range of uses made of satellite-derived SST fields. At one extreme are applications in which the absolute accuracy of the retrieved SST values is central to their use; e.g., when the fields are used as boundary conditions for meteorological models or in global change studies. For these applications it is important to exclude any pixel that is even slightly contaminated. At the other extreme are applications in which the location of oceanographic features is important. In these fields it is the relative accuracy of SST values that is important; e.g., studies of surface ocean fronts or the identification of a feature associated with an anomalous event observed in mooring data. For these applications it is the relative accuracy of neighbouring pixels that is important. The problem of cloud detection is farther aggravated by the difficulty in determining precisely what constitutes a cloud – the degree of cloud contamination is a continuum from no contamination to completely obscurity. Specifically, the opacity of clouds in the infrared ranges from zero to one and the size of clouds varies from much smaller than the region over which the sensor averages the upwelling radiation, the instrument pixel, to much larger regions dominated by synoptic weather patterns.

To date, development of cloud screening algorithms has focused on applications for which the absolute accuracy of the SST value is paramount. In general, these algorithms rely on contrasts in emissivity, reflectivity, temperature and spatial structure between the ocean surface and clouds. These contrasts work well in identifying cloud-contaminated pixels under most open ocean conditions. However, because the screening is based on thresholds associated with these parameters and because the underlying distributions are in most cases continuous, there will be ambiguity in some cases, when one or more of the parameter values is close to a threshold value. This ambiguity is especially true in the vicinity of clouds. The problem is therefore intrinsically probabilistic, with a trade-off between false alarms and hits, a balance that depends critically on the user's application. In order to deal with this, many SST fields are now provided with a separate 'quality' field, which is often derived from the cloud screening portion of the retrieval algorithm, This field allows users to mask SST values based on the quality threshold that meets their needs. One difficulty associated with these fields is that they are derived differently by different data providers with different meanings and they are not described in sufficient detail to understand these differences making it difficult for the user to apply them consistently. This challenge is something that requires careful attention in the future. In addition to providing quality fields with the SST data, there is a trend toward increasing use of simulations in near-real time from national weather programs to inform the discrimination – either by dynamically calculating thresholds or as input to a probabilistic calculation (Merchant et al., 2005). **We recommend that further development of probabilistic cloud detection and flagging algorithms should be undertaken to potentially improve the quality of satellite SST data sets from infrared sensors.**

Despite the importance of feature-related studies, neither the trend toward improved thresholds nor that of inclusion of quality fields addresses the fundamental problem associated with cloud flags from the perspective of feature-related studies. Specifically, most cloud screening algorithms are sensitive to large gradients in the retrieved fields – pixels in a high gradient region are generally flagged as cloud contaminated. This problem is a result either of a spatial homogeneity test applied to brightness temperatures, the use of coarse reference SST fields, or of the use of temporal information in the screening process. In the latter case, a rapidly moving front results in a large temporal SST gradient at a given location and is flagged as cloud contaminated. Unlike the problem associated with the opacity of clouds, this one is more tractable in that it does not depend on the basic definition of a cloud, but rather it is an introduced problem based on the structure of the underlying field. There are two ways of addressing the problem: (1) increase the spatial and temporal resolution of any reference SST fields used, apply the spatial homogeneity test last, and separate the flag for the failure of this test from the general quality flag, or (2) make use of the structural characteristics of fronts to either reset the quality mask for those pixels that are believed to be frontal pixels that were falsely flagged as clouds (Cayula and Cornillon, 1996) or add a new flag indicating same. One advantage of this test is that it can be applied after the SST retrieval and quality fields have been obtained. C**loud screening of infrared satellite data is remains a significant challenge, and despite nearly 30 years of activity, failure to detect sub-pixel clouds remains the source of substantial uncertainty in satellite data sets, we recommend that a systematic review of cloud clearing approaches is undertaken with the purpose of properly documenting the strengths and weakness of each approach. Such a review should identify and prioritise a set of activities that should be undertaken to ensure that the best possible methods of cloud clearing are identified and used by satellite data providers. The review should also assess the severity of cloud detection to the degree of tolerable SST impact for unscreened clouds in the context of user applications.**

## *3.10. Improving SST provision in the high latitude regions*

There are two specific issues associated with the accurate retrieval of sea surface temperature (SST) at high latitudes using infrared radiometry: a) the discrimination between ice-free and ice-covered water at the resolution (temporal and spatial) of the SST retrieval schemes; and b) the accurate correction of the effects of the atmosphere on the infrared radiation as it propagates from the sea surface to the satellite radiometer. In addition, persistently oblique solar illumination can exacerbate the cloud detection problems mentioned in Section 3.9 for high latitudes.

Microwave radiometers on polar-orbiting satellites provide the most-widely used maps of sea-ice extent, independent of cloud cover and solar-illumination, but the spatial requirement of 1-10 km for the derived

SST maps can not be achieved using current passive microwave technology. Enhanced resolution retrievals from the SeaWinds scatterometer provide an alternative source of ice extent information at a spatial resolution of up to 2.2km on a daily basis (Haarpaintner et al, 2004). However, the most obvious solution to identify ice for the high resolution infrared SST retrieval is to generate an ice mask based on contemporaneous information from visible bands. Ice masking is very analogous to the identification of cloud, and the same techniques can be used to identify ice as are used to identify cloud with the addition that coarse resolution ice analyses may provide useful background or a-priori constraints. For infrared SST retrievals, during the day, reflected sunlight provides a powerful mechanism for identifying open, cloud-free water. During the polar night the problem of identifying ice becomes more difficult, and while a simple temperature threshold test might be adequate to identify pack ice, this would not be sufficient in the marginal ice zone. Surface temperature retrievals from spacecraft infrared radiometers below -1.8°C, the freezing point of sea water, can be classified as ice cover. However, this is prone to error as a) there is noise in the satellite-derived surface temperature so that ice-free retrievals could fall below the threshold, and ice-covered pixels fall above the threshold; and b) when melting, sea ice, especially if covered by snow, may remain frozen at temperatures above the threshold. **We recommend that more effort be given to the definition and implementation of ice masking procedures and techniques in Polar Regions for infrared satellite observations.**

Considering the corrections for the atmospheric influence on infrared radiative transfer at high latitudes, the polar atmosphere is generally very dry and cold, and is thus an extreme in terms of the climatological distribution of atmospheric properties. As such it represents an anomalous set of conditions for routine atmospheric correction algorithms that are used to retrieve SST from infrared brightness temperatures. It is to be expected that systemic retrieval errors in the derived SSTs will result when they are obtained using standard atmospheric correction algorithms optimized for the global range of atmospheric variability (e.g. Walton et al, 1998; May et al, 1998). Such bias errors, usually resulting in an erroneously warm SST, are routinely observed and can be greater than 1 K (Vincent et al., 2007a). Recent research (Vincent et al., 2007a, b) using AVHRR brightness temperature data collocated with ship-based radiometric skin SST measurements have shown that a simple, single channel retrieval algorithm can produce improved accuracy in the measurement of skin SST and Ice Surface Temperature (IST; Key et al, 1997). This can be explained by the loss of the correlation between the brightness temperatures measured at 10.5 and 11.5 µm with the atmospheric water vapour that occurs in very dry atmospheres. The brightness temperature differences and relationship to the atmospheric state, at the heart of the assumptions behind all multi-channel atmospheric correction algorithms, are not able to identify the effects of the intervening atmosphere. In such situations single-channel algorithm appears to be better suited to the problem than current multi-channel approaches as shown in Figure 11. As with atmospheric water vapour, the air-sea temperature difference in Polar Regions manifests values that are seldom seen elsewhere over the oceans. Very large values are possible for off-ice airflow. The air-sea temperature difference is important in introducing uncertainties in the retrieved SSTs as it is closely related to the temperature difference between the ocean surface and the atmospheric gases that modify the infrared radiation on its passage to the satellite radiometer. Although less important than in moist atmospheres, the wide range of air-sea temperature differences encountered in Polar Regions introduce a source of uncertainty in the SST retrievals. At this point, it is not clear whether single- or multichannel SST algorithms will be best able to account for such variability. **We recommend that satellite SST data providers using infrared systems review the performance of their atmospheric correction algorithms in polar atmospheres and take steps to develop more appropriate algorithms for these regions.**

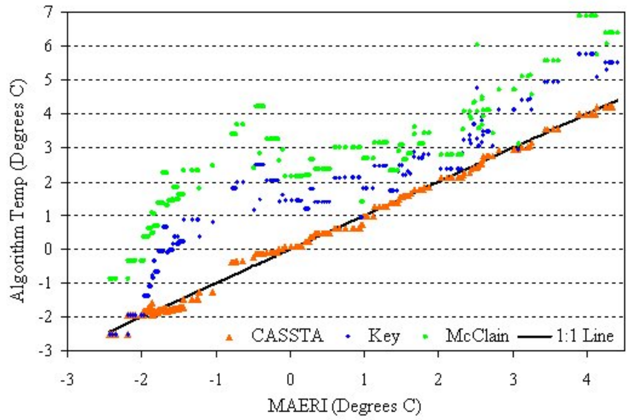


Figure 11. M-AERI ground truth data is compared to CASSTA, McClain SST (1985) and Key IST (Key at al, 1997) estimates. A significant gain in accuracy is evident with CASSTA, which closely follows the 1:1 line. (From Vincent et al, 2007a).

The Marginal Ice Zone between the sea ice and open waters is acknowledged by the GCOS group to be a very important area to observe. However, the conditions in this region are very challenging, as these areas typically area areas with high variability, strong temperature gradients and very few satellite observations from infrared radiometers. In addition, the quality of the observations in this region is questionable and the statistics of the variations when including the marginal ice zone (MIZ) observations in the SST analysis is altered by the presence of ice. An effort is therefore required in order to retrieve more satellite and in situ observations in this area and to quantify the errors associated with the observations. More work is also needed to assess the statistical behaviour of the temperature observations in the marginal ice zone, to obtain a more realistic treatment of the observations in SST analysis products. Figure 12 shows the number of ice-covered days during year 2006 and the average age of the satellite observations (more than 2 days) in the area. The age of the satellite observations as well as the SST variability is large in the marginal ice zone in the East Greenland current.

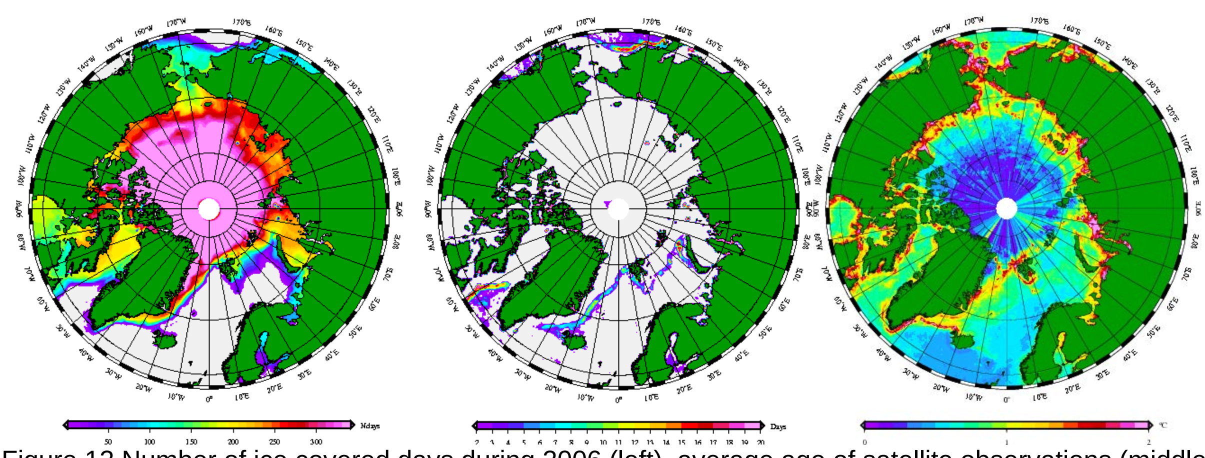

Figure 12 Number of ice covered days during 2006 (left), average age of satellite observations (middle) and day to day variance of the satellite observations (right). The special conditions in the marginal ice zone are clearly identified.

The atmospheric and oceanographic conditions in the high latitudes and the Arctic Ocean are unique, and the ice free satellite SST observations in the Arctic Ocean tend to have higher uncertainty compared to the other regions. However, the number of reliable in situ observations for validation is limited in the ice free waters of the Arctic, and only very few dedicated validation campaigns have been carried out. In order to assess and improve the quality of the SST observations, **we recommend further efforts should be undertaken to continue the SST time series with more reliable in situ and radiometer observations in the Arctic Ocean.**

## *3.11. The challenge of systematic high-resolution SST observations over inland seas, lakes and in the coastal zones*

Current satellite systems are not ideally suited to these applications and important future developments could be that of high spatial resolution geostationary sensors, which are under active consideration by several centres, or constellations of SST imaging sensors which provide observations of any given site several times per day, perhaps not of the accuracy of an AATSR, but with sufficient radiometric discrimination and spatial resolution to monitor the evolution of coastal processes. However, Operational satellite-derived lake water surface temperature (LWST) can improve numerical weather prediction models on local scales Oesch et al (2004). LWST variations are dominated by the annual cycle, however, the varying geospatial attributes of each lake result in specific surface temperature characteristics. Diurnal variability is different from open ocean conditions due to the damped mixing in lakes for a typical calm and clear-sky regime. Lakes located close to each other can display considerable differences in average surface temperatures by as much as 3 K (Oesch 2004).

Several data providers routinely provide LWST in their products. The AVHRR Pathfinder dataset includes inland water temperatures for a large number of lakes and rivers. The existing suite of quality and cloud identification tests, however, is not expected to work well in these areas and the accuracy of the inland water temperatures in Pathfinder has not been assesses (Casey et al., 2009). In terms of LSWT accuracy, Oesch et al (2008) conducted a validation study covering 2 years using data from the AVHRR on NOAA 12, 15, 16, and 17, the MODIS on TERRA, and AQUA and with different method-ingested in situ data from different sized lakes. The best results were found for NOAA 16 night time data at Lake Geneva (bias of 0.18 K and standard deviation of 0.73 K) and TERRA night time data at Lake Constance (satellite-buoy bias of −0.08 K and standard deviation of 0.92 K). For all sensor families an overall scatter ranging from 0.9 to 1.6 K was found. Bias of MODIS is larger, −1.73 to 1.9 K, than the one of the AVHRR (−0.28 to 1.5 K). Hook et al (2003) retrieve lake surface temperature from ATSR-2 over lake Tahoe by regressing the in situ measurements against the average ATSR-2 nadir 11- and 12-$\mu$m channel brightness temperatures and show standard errors for recovering the subskin temperatures of 0.4 K for daytime and 0.18 K for nighttime. The EUMETSAT OSI-SAF provide METOP “SST” values over many lakes and validation results over the North American lakes show small bias values 0.25 K and acceptable SD values 0.54 K (for 588 matchups moored buoys) during April to November 2008. As the main error sources, undetected cloudy pixels, false flagging of cold lake areas as cloud, anomalous atmospheric structure and delineation of coastal boundaries appear to be the most significant. However, many data providers do not retrieve LST as normal operational practice and those that do struggle to define a common set of lakes over which LST should be retrieved. Several users of the GHRSST project have indicated that the lack of LSWT in global products is a limiting feature for both SST analysis and in many cases for the L2 temperature observations themselves. **We recommend that producers of L2 SST products discuss and agree on a common definition of global lakes for which Lake Water Surface Temperature (LWST) shall be retrieved routinely from satellite data ideally as part of the SST processing systems.**

The main challenge to improved coastal zone SST and lake temperature retrieval is that of sampling. The navigation of polar orbiters to 1 pixel (1 km) makes it difficult to locate the actual coast line in full resolution imagery. Not only are high spatial resolutions needed, often of a the order of 1 km, for which it is difficult to achieve the radiometric accuracy needed for water surface temperature measurements, but, more significantly coastal zones and lakes are characterized by rapid changes owing to diurnal and tidal changes where the normal ‘two views per day’ often at the same local time, as offered by low orbiting satellite systems, is not suited to the type of variability seen in these regions. In some areas with large tidal excursion (e.g., surrounding the UK coast) significant areas of mudflats can be exposed at low tide leading to errors in SST retrievals close to the coastline and in sizeable estuarine regions. The basic precaution when processing data coastal zones is to relax the temperature tests not only on the cold side (up-welling, river run off) but also on the warm side (intense diurnal warming, protected shallow waters).

**We recommend that satellite SST navigation techniques are improved to allow confident delineation of lake and coastal shores, tidally exposed wet/dry areas and improvement in the retrieval of LSWT and SST in complex coastal regions. Pixel classification flags should be revised and improved in a consistent manner so products use/share a common flag set (at a minimum the definitions should be agreed) to assist in these activities**. For example, users over several years have requested that lakes and rivers be included in the ESA AATSR data which has not yet occurred. This is in part due to the fact that a L1b 'Lake' flag does not exist. It makes sense to flag the location of lakes and rivers and other generic pixel classification at the L1b processing stage to facilitate the application of an appropriate L2 retrieval process. Preliminary common pixel classification flags are proposed in Table 1 as each of these types impacts the L2 retrieval process and associated uncertainty estimates. Flags should be produced at the highest spatial resolution offered by a particular sensor with L2 teams making decisions on their combination/application at lower resolution.

**Table 1 Common pixel classification for SST data products.**

| Name | Description |
|---|---|
| Ocean | Deep water ocean (>300km offshore Exclusive economic Zone (EEZ) limit ) [*a priori,* static map] |
| Coastal | *n* km [300km EEZ] offshore [*a priori*, static map] |
| Tidal | Areas where seafloor may be exposed on a regular basis according to the state of the tide e.g. coral reef, tidal flats such as Morcecombe Bay, NW England. [*a prior*i, static map but could be dynamic] |
| Lakes/rivers | 'significant' lakes and rivers [*a priori*, static map, significant size TBD] |
| Ice_mask | Ice covered areas [derived from data, could include a classification of ice types] |
| Land | Land areas [*a priori* binary map– further classification is part of L2 process] |
| The flowing flags are not mutually exclusive | |
| Cloud | Cloudy pixels determined by measurements [could also include cloud_type/height in simple high/med/low classes] |
| Snow | Area covered by snow, determined by measurements |
| Sunglint | Sunglint detected in pixel, determined by observations |
| Aerosol present | Aerosol likely or detected in pixel [determined by measurements] |
| Rain | Rain flagged in pixel [determined from passive microwave data] |
| Side Lobe contamination | Data thought to be contaminated by land [derived from data] |

## *3.12. Improving the international SST Service to users*

One of the original requirements for the GHRSST was to foster and increase the interoperability of SST remote sensing products within a 'service' environment so that end users would not be faced with a plethora of varying data and metadata models, file format structures and access methods. Similar requirements have led large earth science satellite programs such as the NASA Earth Observing System Data and Information System (EOSDIS) to adopt "self describing" data formats such as the Hierarchical Data Format (HDF) and HDF-EOS. Although such formats have a learning curve, and specialized software interfaces and libraries associated with them, they are minor disadvantageous in contrast to the benefits of interoperability, platform and operating system independence, and standardization. These formats, although requiring a learning curve, provide advantages by offering interoperability, platform and operating system independence, and standardization of all products. GHRSST evaluated a number of these formats and chose netCDF3 (network Common Data Format) that had a strong history and acceptance in climate modelling and was finding increasing usage in other earth science disciplines including satellite data products. The GHRSST has commissioned a study into the use of netCDF4 for SST products which shall be published in June 2009, however use of the new netCDF4 format is not expected for several year. While self-describing formats such as netCDF excel at documenting the contents of data files in human readable form, standards with regard to attribute conventions, and naming and implementation of discipline-specific data structures still need to be considered. To this end, GHRSST-PP followed the netCDF metadata “model” and implemented Climate Forecast (CF) metadata conventions in all their products. This allowed the Project to leverage a large list of registered conventions related to time, space, and variable attributes and naming conventions. Where the Project found the need of new specific variable names it has petitioned the CF governing community for changes. An example of this is the CF adoption of the sea_surface_foundation_temperature “standard name”. By implementing netCDF in conjunction with CF metadata, GHRSST has met a number of requirements for data interoperability including simplified software readers and other tools for all products, variable documentation through metadata, a format consistent with long term data archiving and stewardship and data distribution via open source protocols such as OPeNDAP. **We recommend that the netCDF Climate Forecast (CF) convention is used by satellite and in situ data providers to manage metadata and interoperability of the modern-era SST data record. We urge other satellite SST data providers to adopt the GHRSST L2P, L3, and L4 netCDF format approach to providing SST data to user communities.**

Today operational meteorological forecast systems use a wide variety of satellite L1b observations as the main product of choice for these mature systems. Model data assimilation systems use directly L1b brightness temperatures and compute geophysical parameters directly. The operational oceanographic community (including physical and bio-geo-chemical communities) do not yet share this approach (partly due to more immature systems) but in the near future, as more operational coupled ocean-atmosphere model frameworks emerge, the need for L1b products form which SST can be derived by the user will increase. This is already the case in the United States where operational ocean forecasting systems already request ocean L1b data streams and within the GHRSST project, several users have requested access to AATSR brightness temperatures for direct assimilation. These users require level 1b brightness temperatures and/or radiances (the latter for VIS/NIR channels) in near real time for the following applications:

- To be able to apply the users own SST retrieval algorithm to make use of new retrieval science quickly and to have control over when the retrieval algorithm is changed.
- To be able to assimilate the radiances directly in NWP models to improve the model surface, total column water vapour and aerosol fields. For NWP assimilation applications the data must be available in less than 3 hours of the measurement time and in a format suitable for NWP centres (currently BUFR).
- To compare satellite radiances with climate model simulated radiances
- To allow satellite radiances to be used as part of the Global Space-Based Inter-Calibration System (GSICS) comparison in real time to inter-calibrate other satellite sensors.

Considering the development of SST CDR for specific sensors, L1b data form the fundamental geophysical record from which all other variables are derived and therefore form a natural long-term archive from which multiple re-analysis may be performed to derive higher order products. **We recommend that satellite SST data providers take steps to make their L1b data available for use in the SST CDR re-analysis community and as part of NRT GHRSST products.**

The modern SST observing system has been successful in generating, delivering, and archiving a large amount of integrated, commonly-formatted SST measurements complete with uncertainty estimates. GHRSST identified the need for a long term data management and user support services including easy access to data sets in an operational near real time context and a long term data stewardship program. GHRSST data flow from the Regional Data Assembly Centres (RDACs) to the Global Data Assembly Centre (GDAC) at NASA's JPL in near real time. 30 days after observation, the data are transferred to the Long Term Stewardship and Reanalysis Facility (LTSRF) at NOAA's National Oceanographic Data Center (NODC). Since large scale GHRSST data production and dissemination commenced in 2006, the GHRSST GDAC and LTSRF have combined to provide over 30,000 users more than 50 terabytes of GHRSST data. Over 18 terabytes of data are in NODC’s LTSRF holdings with another 3/4 of a terabyte added each month. **The adoption of a standards based approach to metadata and file formats has been one of the foundations of success for GHRSST and we recommend that the approach is strengthened and extended to other satellite and in situ SST data providers using appropriate mechanisms.**

At the core of GHRSST's success was the international collaboration on which it was based. In eight years of discussion, debate and planning the main agencies responsible for operating satellite SST sensors and for producing the primary SST datasets have worked with ocean scientists familiar with the processes affecting remote sensing of SST, and with key operational users of SST data, to lay down the rule base for the sharing, indexing, processing, quality control, archiving, analysis and documentation of SST data from diverse sources. These rules are specified in the GHRSST Data Specification document (GDS; Donlon et al., 2006) which defines clearly the input and output data specifications, data processing procedures, algorithms and data product file formats that are common to each GHRSST sub-system. In order for the GHRSST R/GTS framework to function, all GHRSST products must strictly follow the common GDS when generating L2P, L3, and L4 data. As a result, users with tools to read data from one RDAC can draw data from any of the others as well as the GDAC and LTSRF, and will find it is immediately readable by their systems having uniformity within the limits of flexibility permitted by the GDS. Moreover, GHRSST was able to move rapidly from defining the GDS to the present situation in which global L2P, L3, and L4 products are used operationally, because it established by consensus an implementation framework in which the new data products and services are provided. No attempt was made to impose a top-down structure for commissioning data production. Instead, agreement and commitment to the GDS facilitated the existing agencies each to contribute a part of the necessary international effort through the Regional/Global Task Sharing system that is illustrated in Figure 13. This is a distributed modular model with a hierarchical distinction between RDAC, GDAC LTSRF. **We recommend that the GHRSST Data Specification, in a revised and updated format, and the GHRSST Regional/Global Task sharing Framework (R/GTS) are strengthened and sustained in order to maintain and develop the strong user-producer collaboration that has enabled a new generation of SST data products and services to develop over the last 10 years.**

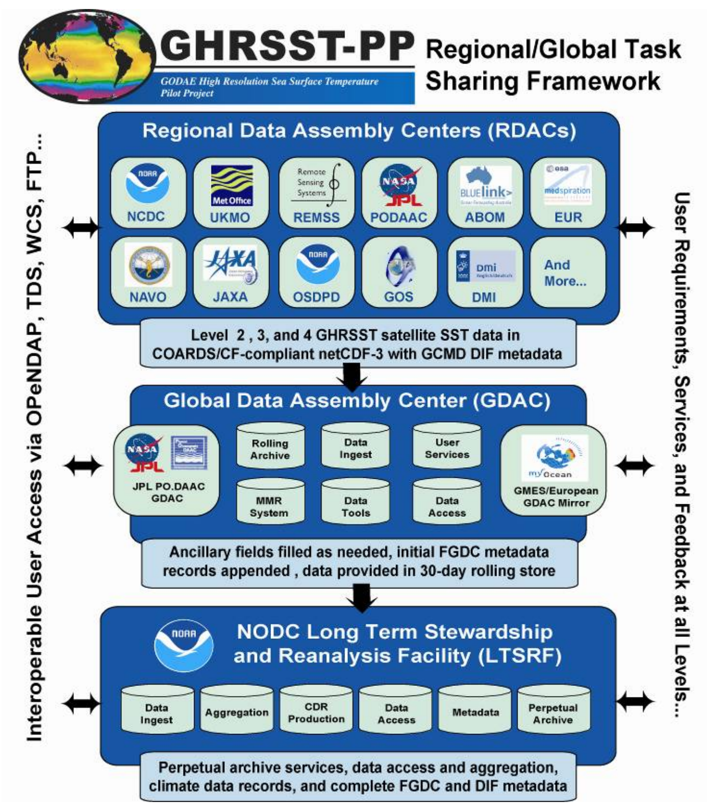


Figure 13 The R/GTS framework that established an international set of RDACs, each of which delivers data to a GDAC (online at http://ghrsst.jpl.nasa.gov) and the regional user community. Data were and continue to be served from the GDAC to near-real-time users and applications for 30 days before the data

are sent to the LTSRF (at http://ghrsst.nodc.noaa.gov) for long-term archive, stewardship, provision to delayed mode users, and future CDR production.

**Regional Data Assembly Centres (RDAC):** Each RDAC has responsibility for one or a relatively small number of satellite products. For example, the US Naval Oceanographic Office RDAC produces global L2P AVHRR SSTs and a global L4 analysis product that incorporates in situ data. A European RDAC produces real time L2P SST products from the AATSR and SEVIRI sensors, a regional high resolution AVHRR product and an ultra-high resolution (~2km) analysis product for the Mediterranean Sea. Another RDAC, generating L2P products from microwave radiometers such as the TMI and AMSR-E is operating at Remote Sensing Systems, Inc. The Japan Aerospace Exploration Agency (JAXA) and the Japan Meteorological Agency (JMA) operate another RDAC in collaboration with the University of Tohoku[3] providing both regional and global measurements and L4 analysis products and L2P products from AMSR-E. GHRSST also has RDAC services at the Australian Bureau of Meteorology (the BLUElink> project[4] and the Integrated Marine Observing System[8]) providing Australian coverage L2P and L4 data, at the NOAA Office of Satellite Data Processing and Distribution (OSDPD) providing GOES-E and GOES-W L2P data and at the NOAA National Climatic Data Center (NCDC) providing L4 analyses. The University of Miami and the Ocean Biology Processing Group, in collaboration with the PO.DAAC at NASA/JPL/Caltech are also serving as the RDAC for MODIS data. Other RDACs exist as well.

Several RDAC systems implement an active user consultation process in order to provide the best possible service to the user community. A GHRSST user support office has been configured at the GDAC facility and provides services to work with the user community and resolve any issues they may raise regarding SST measurements and their specific application. User consultation workshops have been held by RDAC teams to provide a forum in which user communities can feed-back their experience and requirements and a number of key users are engaged at all levels of the project in order to 'pull through' the scientific developments within the RDAC SST activities and to demonstrate the benefit of high-resolution SST data products. User consultation and feedback were essential elements of GHRSST and regular interaction with user communities lay at the heart of its successful implementation. The number of RDAC's is continually growing as more groups choose to adopt the GHRSST standards for data sharing and realise the benefit of the GHRSST R/GTS framework. **We recommend that the Regional Data assembly Centres (RDAC) are sustained and strengthened following the subsidarity principle in order to provide the maximum benefit to each region, to ensure that feedback from national and regional users is gathered, and that the modern era SST record and services develop according to user requirements.**

**Global Data Assembly Centre (GDAC):** The GDAC hosted by the NASA JPL Physical Oceanography Data Active Archive Centre (PO.DAAC[5]) has a central coordination role for RDAC data streams and users. The GDAC lies at the core of the R/GTS providing a data management entity and "clearinghouse" for near real time data ingestion and distribution, metadata management and data search/discovery including responsibility for:

- The ingestion, quality checking and management of the near real-time L2P and L4 products and metadata from RDAC data providers.
- Insertion of model-based meteorological and sea ice ancillary fields into GHRSST-PP L2P products, if not already present.
- Distribution of all GHRSST products through traditional protocols (i.e., FTP, OPeNDAP, WMS) and unique user subscriptions from a 30-day rolling store.
- Delivery of 30 day and older GHRSST products and Federal Geographic Data Committee (FGDC) compliant metadata to the LTSRF for product archiving reanalysis, and long-term access.

[3] See http://www.ocean.caos.tohoku.ac.jp/~merge/sstbinary/actvalbm.cgi
[4] See http://www.bom.gov.au/bluelink
[8] See http://www.imos.org.au
[5] See http://ghrsst.jpl.nasa.gov

- Metadata management and data search services (MMR[6]) and data tools (e.g., sub-setting).
- Customer and Application support for current and future users.
- Integration of user services at the GDAC with the Application and User Services office of the GHRSST-PP at the international level.

User support is coordinated in collaboration with the GHRSST Project Office and the JPL PO.DAAC where customers can submit support requests via a web site monitored by customer service specialists and data engineers. Applications development is currently focused on integrating GHRSST products into coastal decision support systems and marine resource management (e.g., U.S. Integrated Ocean Observing System, IOOS). Measures of effectiveness are tracked and reported at the GDAC to provide overall metrics for the GHRSST-PP program. The GDAC user support service is a natural location to monitor delivery and timeliness of products from RDAC so that users informed of problems when they occur at the earliest possible time. **We recommend that the GDAC service collaborate with RDACs and the Long Term Stewardship and Re-analysis Facility to actively monitor the status of the R/GTS, data products and work with users.**

**Long Term Stewardship and Reanalysis Facility:** At the base of the R/GTS framework is the Long Term Stewardship and Reanalysis Facility (LTSRF[7]) located at the NOAA National Oceanographic Data Center. The LTSRF provides both a long term archive and forms the central hub of the GHRSST reanalysis (RAN) system. Complete scientific data stewardship rather than basic archival storage and backup is required to enable a successful RAN system. Focused stewardship of the data protects and enhances the significant investment made by RDAC and GDAC teams and ensures the long term understanding and usability of the data in a wide range of future applications. The overall stewardship responsibilities of the LTSRF include:

- Data ingest including the receipt, verification, and proper cataloguing, via appropriate file-level and collection-level metadata, of both the near-real time and any delayed mode data.
- Archive activities including offsite backup, media migration, and validation of stored data.
- Data access including the critical role of providing standards based data and metadata to both a diverse user community and to the LTSRF itself. All data are provided in a free and open manner via the internet. Fees for media distribution (CD, DVD, etc.) are limited to cost of production only.
- Application utilities and support to help data providers/developers use and test GHRSST data in a variety of applications. As data providers/developers become much more knowledgeable about the data they are able to quickly identify weakness and rapidly validate the use of GHRSST data in new application areas and provide feedback on the data quality. The latter is essential for proper long-term stewardship of GHRSST data.
- The User Services component of the LTSRF not only provides standard user assistance with questions about the data, but also serves as the vital feedback loop to receive user input on problems, new applications, and directions for future improvements. The GDAC and the LTSRF share statistics on data ingest and access patterns, to better understand user needs and trends in their approaches to acquiring data.
- Reanalysis functionality enabling reprocessing of the entire GHRSST collection at a sufficiently rapid pace to take advantage of new delayed mode data sets or new analysis techniques.

Thirty days after a measurement has been made, GHRSST data are transferred from the GDAC to the LTSRF where stewardship is provided in perpetuity. A large metadata transformation process, which begins at the RDACs and is enhanced at the GDAC through creation of a baseline FGDC compliant record for every daily collection of GHRSST products, is also completed at the LTSRF. Users may access data at any of the GHRSST centres in real time and in delayed mode, using FTP and OPeNDAP protocols. At all levels of the GHRSST international framework, the user community is fully engaged in the development and specification of services and data products. Through the GHRSST R/GTS, observation and analysis SST data sets can be obtained through robust operational data

[6] See http://ghrsst.jpl.nasa.gov/data_search.html
[7] see http://ghrsst.nodc.noaa.gov

servers, in near real time, in the same generic format and with uncertainty estimates. A large amount of SST data has been made available to the community in this way and the GHRSST may serve as a model for other international projects making a contribution to the Global Earth Observation System of Systems (GEOSS).

# 4. Forward-looking community-consensus vision

The activities within the SST community over the last decade have transformed the measurement of SST using a complementary satellites and in situ measurements working in synergy together. The establishment of a framework for the exchange and management of international SST data has been successfully implemented and is operating on a daily basis. A thriving user community has developed in which integrated SST data sets are being used at scientific institution and operational agencies. Tools and data services have been developed and implemented to serve this user community. Through the activities of GHRSST many lessons have been learned that provide the basis for an optimal configuration for the SST observing system in the next 10 years.

Based on the GHRSST experience, the modern-era SST observing system has a new and modern era structure. It is not simply a collection of measurements made by satellites and in situ sensors but a truly integrated system that takes a wide scope end-to-end view. The foundation of the system is embodied in the four guiding principles adopted by GHRSST for an integrated approach to develop the modern-era SST capability:

(1) Respond to user SST requirements through a consensus approach,
(2) Organise activities according to principles of subsidarity and shared responsibility,
(3) Develop complementarity between independent measurements from earth observation satellites and in situ sensors and,
(4) Maximise synergy benefits of an integrated SST measurement system and end-to-end user service.

This paper has highlighted some of the developments that have developed the modern-era SST capability over the past 10 years in and provides a set of recommendations for the next 10 years based on experience and remaining challenges. The major challenges focus on augmenting and maintaining high quality SST measurements from both in situ and satellite instruments, maintaining and developing the scientific and operational SST community, providing robust and sustained methods and tools that provide uncertainty and error estimates in a format that is easy to use by users and developing and maintaining an SST data stewardship and reanalysis program that is able to tackle the development and validation of SST climate data records. Ocean observations themselves provide the substance for the modern era SST capability but it is the community of users and producers that realise their full potential and societal benefit: each is of marginal use without the other.

In terms of developing a consensus for an ideal SST measurement system, Table 2 provides a summary overview of the integrated measurement network. It includes a daily passive microwave satellite capability that delivers global coverage SST measurements every day. This capability is far from secure and in reality only one instrument may be available in the next decade. The complementary infrared satellite system consists of two wide swath polar orbiting satellite instruments with multi-spectral capability and is secure for the next decade. Likewise, the unique dual-view capability of the ENVISAT AATSR is secure although our ideal plan proposes a second satellite placed in an asynchronous orbit to provide a reference SST data set for other satellite systems. We acknowledge the role of Geostationary Observing System maintained for the benefit of weather forecasts and urge providers to ensure that a multi-spectral capability is maintained on all satellites so that SST can be retrieved. We recognise that there is a need for better sea ice data sets derived from a number of satellite and in situ sensors in the polar regions to assist in the provision of accurate SST in the marginal ice zone. We identify the need for surface wind observations that are preferably contemporaneous with SST measurements to allow proper validation and uncertainty estimation using in situ data and to understand diurnal variability – two significant on-going challenges. We recognise that there are different communities working to a common goal such as the

GSISC and note the importance of working together for the common benefit of the ocean community to assure the best quality satellite data sets are provided to the community.

We draw special attention to the in situ SST observing system. This component is essential to our long term vision for an integrated SST capability today and for the future. The in situ SST network must be strengthened in several important ways. All SST observations must report their depth of measurement. complete metadata for each platform should be available as an online indexed resource in a central repository (e.g., JCOMMOPS) so that users can look up calibration histories and calibration parameters. Uncertainty and error estimates should be provided for every measurement using QA4EO processes. The network should be strengthened by modern Argo float technologies that measure the SST in the upper 10m of the ocean at a resolution of at least 0.5 m. New ocean moorings (e.g. OceanSites) with high quality reference sensors should be deployed. Radiometers aboard ships measuring the radiometric SST should be more widely deployed. The in situ network is the reference data source on which the satellite community ultimately depends for accurate verification and validation of its SST data products. It must be as accurate as possible if the vast investments made by Governments for SST measurements from space are to be fully realised.

Finally, our vision includes a concerted effort to work closely with the user community to understand and act upon their needs and requirements. The modern era SST capability requires a re-analysis capability that allows users easy and open access to the archives maintained by satellite providers at L2 and at L1b in order to create SST CDRs and ultimately the SST ECV. SST data from the modern era SST system are required by the scientific research community and by the operational ocean and NWP communities. By taking care to obtain a consensus between the space agency producers of satellite SST data, agencies collecting and managing in situ SST data sources, service agencies with an interest in producing SST products and services, and operational users of SST, the future success of the modern-era SST record and capability will be ensured.

**Table 1. Ideal plan for a global high resolution SST integrated observing system: 2009-2025.**

| Integrated SST observing system element | Purpose | Minimum configuration | Radiometer wavebands/data sources | Accuracy | Nadir spatial resolution | Swath width | Coverage /revisit |
|---|---|---|---|---|---|---|---|
| **Ocean in situ SST system** | Derivation of satellite SST uncertainty and error (bias) estimates.<br><br>Satellite SST retrieval algorithm development.<br><br>Monitoring of satellite instrument degradation.<br><br>Validation and verification of satellite data | Global array of drifting buoys, profiling floats with high resolution upper ocean (top 10 m) measurements complemented by ship mounted infrared radiometers. The number of moored fiducial sites with high quality instrumentation should be increased. Uncertainty estimates need to be delivered with all measurements and the depth of SST measurement reported with all measurements. Calibration stability must be assured and ideally demonstrated for all platforms. | Contact thermometry and thermal infrared radiometers. | <0.1K | Point source | N/A | Providing global sample of SST and atmospheric structure and in oceanic areas characterised by strong temperature gradients. There is a need for better sampling in the polar regions and the Southern Ocean<br><br>Reporting hourly together with measurement depth.<br><br>The network should be theoretically optimised for SST applications following appropriate dedicated observing system experiments. |
| **Atmospheric parameters** | Satellite SST algorithm development. Validation of SST retrievals. Providing global sampling of atmospheric variability. Timely NWP access is an essential part of the system as we move towards optimal estimation for SST; and high resolution (in time) NWP (winds and surface fluxes) will assist in diurnal variability analysis. | Atmospheric profiles of temperature, humidity, aerosol loading. Precipitation, surface air temperature, solar irradiance, surface wind speed. | Satellite atmospheric sounders, contact measurements.<br>NWP and ocean model outputs | T: <0.3K, R/hum: <5%<br>For 2m/s <U< 4 m/s accuracy should be 0.5 m/s.<br>For U > 4 m/s accuracy should be < 1.0 m/s. | Point source | N/A | Reporting hourly together with measurement height |
| **Wide swath Infra red imager** | Baseline global high resolution SST observing system. | Two-four low earth orbit sun-synchronous satellites optimally spaced in time with infra-red radiometers | Thermal IR channels within the ~3.7- 12 µm waveband for SST measurement, near-IR and visible channels for cloud flagging. | 0.1-0.3 K | 0.5-1 km (Target 0.25km) | >2000 km | Day and night global coverage by each satellite |
| **Dual view SST** | Baseline global coverage high | At least one LEO satellite | Thermal IR channels | 0.1 K | 0.5-1 km | >500 | Global coverage. |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **reference sensor** | accuracy SST retrieval with long term stability <0.1K decade$^{-1}$ for the SST CDR. Maintaining accuracy of SST/climate observing system during periods of volcanic stratospheric aerosol. Used as a reference data source other satellite SST data | with a dual view radiometer in an 10.00-10.30 LST orbit to continue the (A)ATSR CDR.<br><br>One LEO asynchronous satellite for use as a reference sensor for other satellite data streams. | within the ~3.7- 12 µm waveband for SST measurement, near-IR and visible channels for cloud flagging, each with dual view along track scanning capability. | | (Target 0.25km) | km | |
| **Wide swath passive microwave imager** | Baseline global coverage moderate resolution SST observing system. | Two satellites carrying microwave radiometers optimized for SST retrieval | For global coverage 7GHz is needed. Other channels are required for corrections for wind, precipitation etc. The AMSR2 channels should be considered minimum baseline. | <0.5 K (Target 0.3 K) | ~25km (target: 10km) | >1500 km | Earth coverage in 1 days |
| **Geostationary constellation of infrared imagers** | Baseline non-polar SST observing system providing high temporal resolution SST. | 6 spacecraft equi-spaced in longitude to ensure full coverage from ~70°S to ~70°N | Thermal IR channels within the ~3.7- 12 µm waveband for SST measurement, near-IR and visible channels for cloud flagging. | <0.5 K | 1-5 km (target 1 km) | Earth disk from 36000 km altitude | Sample interval < 30 min |
| **Wide swath ocean surface winds** | Required to characterise the state of the ocean surface for emissivity and skin temperature deviation and diurnal heating and cooling parameterizations | Two satellites carrying passive and/or active microwave systems. | Various | <1 $ms^{-1}$ (Target: 0.25 $ms^{-1}$ between 2-10 $ms^{-1}$) | <25 km (target:10 km) | >2000 km | Several samples per day. Target: 4hirs? |
| **Sea Ice imaging** | Required to determine sea ice concentration and sea ice edge, | At least one LEO sun-synchronous satellite carrying microwave radiometers optimized for sea ice retrieval.<br><br>At least one LEO sun-synchronous satellite carrying Synthetic Aperture Radar (SAR).<br>At least one LEO sun-synchronous satellite carrying a visible imager | Passive Microwave:<br><br>SAR<br><br>High resolution scatterometer<br><br>Vis imager | | PM < 10km<br><br>SAR: 10 m<br><br>Scatt: 2.5 km<br><br>Vis < 1km | > 1500 km | Polar region coverage in at least 1 day. (Target: 6 hours) |

# 5. References

Barton, I.J., 2007a: Satellite-derived sea surface temperatures – a case study of error variability, Proc. ' ENVISAT Symposium 2007', Montreux, Switzerland, 23-27 April 2007 (ESA SP-636, July 2007).

Barton, I. J., P. J. Minnett, K. A. Maillet, C. J. Donlon, S. J. Hook, A. T. Jessup, and T. J. Nightingale, 2004: The Miami2001 infrared calibration and intercomparison. Part II: Shipboard results. J. Atmos. Oceanic Technol., 21, 268–283.

Casey, K.S., R .Evans, K.A. Kilpatrick, E.J. Kearns, T.B. Brandon, R.W. Vazquez, and J. Vazquez (2009). Toward the sea surface temperature climate data record from space: An improved Advanced Very High Resolution Radiometer Pathfinder dataset, in preparation for Journal of Climate.

Cayula, J.-F. and P. Cornillon, 1996, Cloud Detection from a sequence of SST images, *Remote Sens. of Environ.,* **55**, 80-88.

CEOS, 2008, The Earth Observation Handbook: Climate Change Special Edition 2008, available at http://www.eohandbook.com/

CEOS 2009, The Earth Observation Handbook online database available at http://database.eohandbook.com/

Chelton, D. B., M. G. Schlax, and R. M. Samelson, Summertime Coupling between Sea Surface Temperature and Wind Stress in the California Current System, 2007: J. Phys. Oceano, 37 (3), 495-517.

Corlett, G. K., I.J. Barton, C.J. Donlon, M.C. Edwards, S.A. Good, L.A. Horrocks, D.T. Llewellyn-Jones, C.J. Merchant, P.J. Minnett, T.J. Nightingale, E.J. Noyes, A.G. O'Carroll, J.J. Remedios, I.S. Robinson, R.W. Saunders and J.G. Watts, 2006. The accuracy of SST retrievals from AATSR: An initial assessment through geophysical validation against in situ radiometers, buoys and other SST data sets, Adv. Space Res., 37 (4), 764-769.

Curry, J. A. and A. Bentamy, M. A. Bourassa, D. Bourras, E. F. Bradley, M. Brunke, S. Castro,S. H. Chou, C. A. Clayson, W. J. Emery, L. Eymard, C. W. Fairall, M. Kubota, B. Lin,W. Perrie, R. A. Reeder, I. A. Renfrew, W. B. Rossow, J. Schulz, S. R. Smith,P. J. Webster, G. A. Wick, And X. Zeng, Seaflux, Bull. Am. Soc., (3), 85, 2004, DOI: 10.1175/BAMS-85-3-409.

Donlon, C. J., P. J. Minnett, C. Gentemann, T. J. Nightingale, I. J. Barton, B. Ward, and J. Murray, 2002: Toward improved validation of satellite sea surface skin temperature measurements for climate research. *Journal of Climate*, **15,** 353-369

Donlon, C, J., T. J. Nightingale, L. Fiedler, G. Fisher, D. Baldwin, and I. S. Robinson, 1999: The calibration and intercalibration of sea-going infrared radiometer systems using a low cost blackbody cavity. J. Atmos. Oceanic Technol., 16, 1183–1197.

Donlon C, Robinson I S, Reynolds M, Wimmer W, Fisher G, Edwards R, Nightingale T J., 2008. An Infrared Sea Surface Temperature Autonomous Radiometer (ISAR) for Deployment aboard Volunteer Observing Ships (VOS). Journal of Atmospheric and Oceanic Technology 25 (1): 93-113.

Donlon, C. J., I. Robinson, K. S Casey, J. Vazquez-Cuervo, E Armstrong, O. Arino, C. Gentemann, D. May, P. LeBorgne, J. Piollé, I. Barton1, H Beggs, D. J. S. Poulter, C. J. Merchant, A. Bingham, S. Heinz, A Harris, G. Wick, B. Emery, P. Minnett, R. Evans, D. Llewellyn-Jones, C. Mutlow, R. Reynolds, H. Kawamura1 and N. Rayner. The Global Ocean Data Assimilation Experiment (GODAE) high Resolution Sea Surface Temperature Pilot Project (GHRSST-PP), 2009: Oceangraphy Magazine in press.

Donlon, C. J., I. Robinson, K. S Casey, J. Vazquez-Cuervo, E Armstrong, O. Arino, C. Gentemann, D. May, P. LeBorgne, J. Piollé, I. Barton1, H Beggs, D. J. S. Poulter, C. J. Merchant, A. Bingham, S. Heinz, A Harris, G. Wick, B. Emery, P. Minnett, R. Evans, D. Llewellyn-Jones, C. Mutlow, R. Reynolds, H. Kawamura1 and N. Rayner. The Global Ocean Data Assimilation Experiment (GODAE) high Resolution Sea Surface Temperature Pilot Project (GHRSST-PP), 2007: *Bull. Am. Meteorol. Soc*., 88 (8), 1197-1213, (DOI:10.1175/BAMS-88-8-1197)

Donlon, C.J. and the GHRSST-PP Science Team, 2006: *The GHRSST-PP data processing specification version 1.6*., 241 pp. [Available online at www.ghrsst-pp.org.]

Flament, P., J. Firing, M. Sawyer, and C. Trefois (1994), Amplitude and horizontal structure of a large diurnal sea surface warming event during the coastal ocean dynamics experiment, J. Phys. Oceanogr., 24(1), 124– 139.

Gentemann C. L., P. J. Minnett, P. Le Borgne and C. J. Merchant (2008), Multi-satellite measurements of large diurnal warming events, Geophys. Res. Lett, 35, L22602, doi:10.1029/2008GL035730.

Global Climate Observing System (GCOS), 2006: Systematic Observation Requirements for Satellite-based Products for Climate Supplemental details to the satellite-based component of the Implementation Plan for the Global Observing System for Climate in Support of the UNFCCC GCOS-107, WMO/TD No. 1338, available from http://www.wmo.ch/pages/prog/gcos/index.php?name=publications

Guan, L. and H. Kawamura, 2003: *SST* availabilities of satellite infrared and microwave measurements. J. Oceanogr., Vol. 59, No. 2, 201-209.

Haarpaintner, J., R. T. Tonboe, D. G. Long, and M. L. Van Woert, 2004: Automatic detection and validity of the sea-ice edge: an application of enhanced-resolution QuikScat/SeaWinds data. *Geoscience and Remote Sensing, IEEE Transactions on*, **42,** 1433-1443

Hook, S.J., F.J. Prata, R.E. Alley, A. Abtahi, R.C. Richards, S.G. Schladow, and S.O. Pálmarsson, 2003: Retrieval of Lake Bulk and Skin Temperatures Using Along-Track Scanning Radiometer (ATSR-2) Data: A Case Study Using Lake Tahoe, California. *J. Atmos. Oceanic Technol.*, **20**, 534–548.

Jessup, A. T., R. A. Fogelberg, and P. J. Minnett, 2002: Autonomous shipboard radiometer system for in situ validation of satellite SST. Proc. Int. Symp. of Optical and Scientific Technology, Seattle, WA, SPIE.

Kelliher, H., G. corlett and I. S. Robinson, 2007: SST Data Continuity Final Report, ICP2.SCL.REP.001 under contract CSSPA/0705, available from http://www.pinkbox.space.qinetiq.com/icp2/exploitation/Space_Connexions_Final_Web.pdf, 201 pp.

Kilpatrick, K.A., G.P. Podesta and R. Evans (2001). Overview of the NOAA/NASA Advanced Very High Resolution Radiometer Pathfinder algorithm for sea surface temperature and associated matchup database. J. Geophys. Res., 106 (C5), 9179-9197.

May, D. A., M. M. Parmeter, D. S. Olszewski, and B. D. McKenzie, 1998: Operational processing of satellite sea surface temperature retrievals at the Naval Oceanographic Office. *Bulletin of the American Meteorological Society*, **79,** 397-407.

Merchant C J and Harris A R (1999), Toward the elimination of bias in satellite retrievals of skin sea surface temperature. 2: Comparison with in situ measurements, J Geophys Res, 104, C10, 23579-23590.

Merchant C J, Harris A R, Maturi E and MacCallum S (2005), Probabilistic physically-based cloud screening of satellite infra-red imagery for operational sea surface temperature retrieval, Quart. J. Royal Met. Soc., 131, 2735-2755.

Merchant C J, P Le Borgne, A Marsouin and H Roquet (2008a), Optimal estimation of sea surface temperature from split-window observations, Rem. Sens. Env., 112 (5), 2469-2484. doi:10.1016/j.rse.2007.11.011

Merchant C J, D Llewellyn-Jones, R W Saunders, N A Rayner, E C Kent, C P Old, D Berry, A R Birks, T Blackmore, G K Corlett, O Embury, V L Jay, J Kennedy, C T Mutlow, T J Nightingale, A G OCarroll, M J Pritchard, J J Remedios and S Tett (2008b), Deriving a sea surface temperature record suitable for climate change research from the along-track scanning radiometers, Adv. Sp. Res, 41 (1), 1-11. doi:10.1016/j.asr.2007.07.041

Merchant, C. J., M. J. Filipiak, P. Le Borgne, H. Roquet, E. Autret, J.-F. Piolle, and S. Lavender (2008c), Diurnal warm-layer events in the western Mediterranean and European shelf seas, Geophys. Res. Lett., 35, L04601, doi:10.1029/2007GL033071.

Minnett, P. J., R. O. Knuteson, F. A. Best, B. J. Osborne, J. A. Hanafin, and O. B. Brown, 2001: The Marine-Atmosphere Emitted Radiance Interferometer: A high-accuracy, seagoing infrared spectroradiometer. J. Atmos. Oceanic Technol., 18, 994–1013.

National Research Council (NRC), 2008: Ensuring the Climate Record from the NPOESS and GOES-R Spacecraft: Elements of a Strategy to Recover Measurement Capabilities Lost in Program Restructuring Committee on a Strategy to Mitigate the Impact of Sensor Descopes and Demanifests on the NPOESS and GOES-R Spacecraft, National Research Council, ISBN-13: 978-0-309-12184-2 190 pages available from http://books.nap.edu/catalog.php?record_id=12254#orgs.

Oesch, D. C., J.-M. Jaquet, A. Hauser, and S. Wunderle, 2005: Lake surface water temperature retrieval using advanced very high resolution radiometer and Moderate Resolution Imaging Spectroradiometer data: Validation and feasibility study, *J. Geophys. Res.*, 110, C12014, doi:10.1029/2004JC002857.

Oesch, D. C., A. B. Hauser, and S. Wunderle , 2004: Operational application of NOAA-AVHRR data for mapping lake surface temperatures in an alpine environment: feasibility and validation, Remote Sensing for Agriculture, Ecosystems, and Hydrology V. Edited by Owe, Manfred; D'Urso, Guido; Moreno, Jose F.; Calera, Alfonso. Proceedings of the SPIE, Volume 5232, pp. 284-292 (2004).

O'Carroll, A.G., J.R. Eyre, and R.W. Saunders, 2008: Three-Way Error Analysis between AATSR, AMSR-E, and In Situ Sea Surface Temperature Observations. *J. Atmos. Oceanic Technol.*, **25**, 1197–1207.

Robinson, I. S. and C. J. Donlon (2003) Global Measurement of Sea Surface Temperature from Space: Some New Perspectives, J. Atm. Ocean Sci. (previously The Global Atmosphere and Ocean System), 9, (1), 19-37. (Note revised bibliographic reference to make it accessible electronically)

Roemmich D. and the Argo Steering Team, 2009: Argo: the challenge of continuing 10 years of progress, Submitted to Proceedings of the Final GODAE Symposium.

Saunders, R. W., N.R.Ward, C.F.England and G.E.Hunt, Satellite observations of sea surface temperature around the British Isles. Bull. Am. Met. Soc. 63 267-272 (1982).

Smith, N.R. and Koblinsky, C., 2001: The ocean observing system for the 21st Century, a consensus statement. In: Koblinsky, C. and Smith, N.R. (Eds.), Observing the Oceans in the 21st Century, pp 1-25, GODAE Project Office, Bureau of Meteorology, Melbourne, Australia.

Stammer, D, F. J. Wentz, and C. L. Gentemann, 2003, Validation of microwave sea surface temperature measurements for climate purposes, Journal of Climate, 16, 73-87.

Stark, J. D., C. J. Donlon, M. J. Martin and M. E. McCulloch, 2007, OSTIA : An operational, high resolution, real time, global sea surface temperature analysis system., Oceans '07 IEEE Aberdeen, conference proceedings. Marine challenges: coastline to deep sea. Aberdeen, Scotland. IEEE.

Vazquez-Cuervo, J., E. Armstrong, K. Casey, R. Evans, and K. Kilpatrick, 2009: A comparison between the Pathfinder Version 5.0 and Version 4.1 Sea Surface Temperature Data Sets: A Case Study for High Resolution, submitted to J. of Climate.

Vincent, R. F., R. F. Marsden, P. J. Minnett, and J. R. Buckley, 2008b: Arctic Waters and Marginal Ice Zones: Part 2 - An Investigation of Arctic Atmospheric Infrared Absorption for AVHRR Sea Surface Temperature Estimates. *Journal of Geophysical Research*, **113,** C08044.doi:10.1029/2007JC004354.

Vincent, R. F., R. F. Marsden, P. J. Minnett, K. A. M. Creber, and J. R. Buckley, 2008: Arctic Waters and Marginal Ice Zones: A Composite Arctic Sea Surface Temperature Algorithm using Satellite Thermal Data. *Journal of* Geophysical Research, 113, C04021.doi:10.1029/2007JC004353.

Walton, C. C., W. G. Pichel, J. F. Sapper, and D. A. May, 1998: The development and operational application of nonlinear algorithms for the measurement of sea surface temperatures with the NOAA polar-orbiting environmental satellites. Journal of Geophysical Research, 103, 27,999-28,012.

Wentz, F. J.,, C. Gentemann, D. Smith, and D. Chelton, 2000: Satellite Measurements of sea surface temperature through clouds. Science,288, 847–850.

Zhang, H.M., R.W. Reynolds, R. Lumpkin, R. Molinari, K. Arzayus, M. Johnson, and T.M. Smith, 2009: An Integrated Global Observing System For Sea Surface Temperature Using Satellites and in Situ Data: Research to Operations. *Bull. Amer. Meteor. Soc.*, **90**, 31–38.

Zhang, H.-M., R. W. Reynolds, and T. M. Smith, 2006: Adequacy of the in situ observing system in the satellite era for climate SST. *J. Atmos. Oceanic Techol*., 23, 107–120.