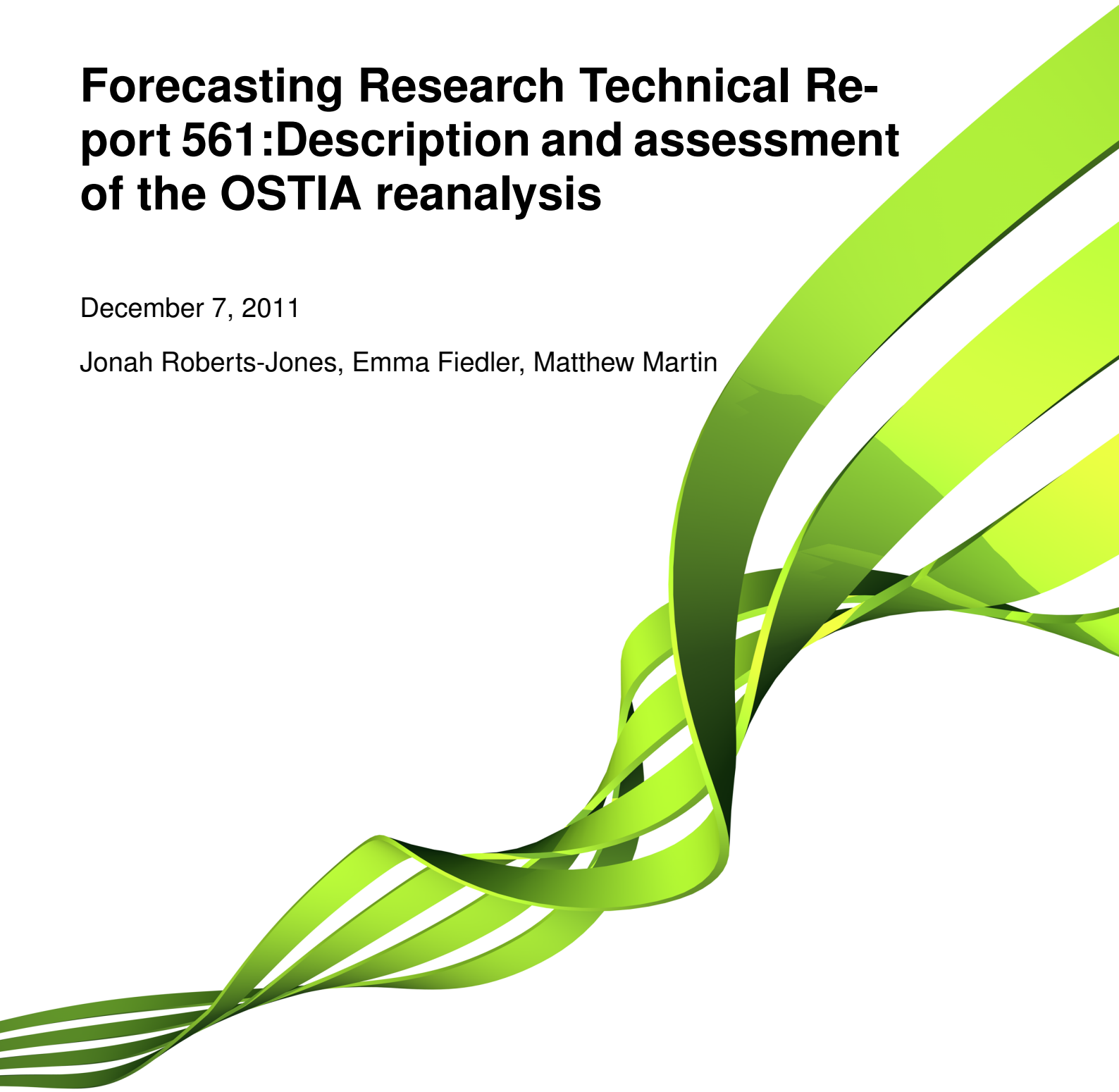
Forecasting Research Technical Report 561:Description and assessment of the OSTIA reanalysis

December 7, 2011

Jonah Roberts-Jones, Emma Fiedler, Matthew Martin



Abstract

A sea surface temperature (SST) and sea ice reanalysis has been produced at the Met Office based on the Operational SST and Sea-Ice Analysis (OSTIA) system. The OSTIA reanalysis produces daily, global foundation SST and sea ice concentration fields on a $1/20^{\circ}$ grid from the 1st January 1985 to the 31st December 2007. The OSTIA reanalysis uses SST satellite observations from the AVHRR Pathfinder and (A)ATSR multi-mission series data sets and in situ SST observations from the ICOADS data set. The sea ice concentration fields are interpolated from the OSI-SAF sea-ice concentration reprocessing. The SST reanalysis is obtained using a multi-scale optimal interpolation (O.I.) type scheme similar to the near real time OSTIA system but differences in both the data used and system design mean that the two analysis systems are distinct. Validation of the SST analysis using assimilated in situ observation minus background statistics showed that the accuracy of the analysis increases throughout the reanalysis time period to approximately 0.50 K by 2007. Independant validation using ARGO data showed the OSTIA reanalysis to have a global cold bias of 0.10 K and standard deviation error of approximately 0.55 K. Assesment of the OSTIA reanalysis at high latitudes shows that the SST and sea-ice fields are more consistent in the Southern than in the Northern Hemisphere. Comparison to similar SST reanalysis products have shown global warm bias in the OSTIA reanalysis compared to the HadISST reanalysis of approximately 0.50 K but good agreement with the Reynolds SST reanalysis.

Contents

Glossary

(A)ATSR (Advanced) Along Track Scanning Radiometer

AMSR-E Advanced Microwave Scanning Radiometer for Earth Observing System

AVHRR Advanced Very High Resolution Radiometer

CCI Climate Change Initiative

ESA European Space Agency

ICOADS International Comprehensive Ocean-Atmosphere Data Set

IPY International Polar Year

ITCZ Inter-Tropical Convergence Zone

PFSST Pathfinder Sea Surface Temperature (algorithm)

MODIS Moderate Resolution Imaging and Spectralradiometer

NAR Near Atlantic Regional

NASA National Aeronautic and Space Administration

NEODC NERC Earth Observation Data Centre

NERC Natural Environment Research Council

NESDIS National Environmental Satellite, Data and Information Service

NLSST Non-Linear Sea Surface Temperature (algorithm)

NOAA National Oceanic and Atmospheric Admininstration

OLS Operational Linescan System

OSI-SAF Ocean and Sea Ice - Satellite Application Facility

OSTIA Operational Sea Surface Temperature and Sea-Ice Analysis

QC Quality Control

RSS Remote Sensing Systems

SEVIRI Spinning Enhanced Visible and Infra-Red Imager

SMMR Scanning Multichannel Microwave Radiometer

SSES Single Sensor Error Statistic

SSM/I Special Sensor Microwave/Imager

SST Sea Surface Temperature

TMI Tropical Microwave Instrument

VOS Voluntary Observing Ship

1. Introduction

1.1 Background

The need for information about global sea surface temperature (SST) comes from a wide range of users. A sub-set of these users require information about both the historical SST record and about the current SST in near real time (NRT). For instance, numerical weather prediction (NWP) centres require information about the SST to provide a lower boundary condition for their models in both forecast and reanalysis modes, seasonal forecasters require SST data to constrain the air-sea interface within their coupled ocean-atmosphere models, and climate groups monitor the current state of the SST within the context of historical SST records. Most of these users require the SST information as global gap-free analyses (otherwise known as Level 4 data).

A large number of NRT Level 4 SST data-sets exist, coordinated through the Group for High Resolution SST (GHRSST) Inter-Comparison Technical Advisory Group (https://www.ghrsst.org/The-Inter-Calibration-TAG-(IC-TAG).html). These are usually produced on a daily basis with high horizontal resolution (at least $1/4^\circ$ resolution), and use the large number of satellite data currently available. A number of inter-comparisons of these NRT products have been performed (Reynolds and Chelton, 2010, for example).

There are also a number of long-term historical SST data-sets, with inter-comparison carried out through the Global Climate Observing System (GCOS) SST inter-comparison project. These are generally of lower temporal and spatial resolution (typically of the order of weekly 1° resolution), and use in situ SST data (with some products using the satellite data in the more recent period). Of particular relevance here is the HadISST analysis (Rayner et al., 2003) produced at the UK Met Office, as it is important that a reanalysis of the recent period is made within the historical SST record. Until now only one SST analysis has been available covering both the historical satellite period and as a NRT product, produced at NOAA's National Climatic Data Centre (Reynolds et al., 2007).

The Operational Sea Surface Temperature and Sea-Ice Analysis (OSTIA) system (Donlon et al., 2011), developed and run at the UK Met Office, was originally designed for NRT applications. For those purposes it is widely used, particularly by NWP centres (including the UK Met Office, ECMWF, Meteo France and others), operational ocean forecasting systems, and climate monitoring groups.

In order to meet the needs of the users described above, a homogeneous reanalysis using OSTIA has been produced as part of the MyOcean European project (http://www.myocean.eu.org), covering the period 1st January 1985 - 31st December 2007. The reanalysis system is based on the NRT OSTIA system with a number of changes to input data and methods. The OSTIA system provides an estimate of the foundation SST, which is defined to be the SST free of diurnal warming (Donlon et al., 2002). The foundation SST is representative of the temperature throughout the mixed layer just before sunrise measured at a nominal depth of 0.2-1 m.

This report provides a description of the OSTIA reanalysis system as well as a thorough assessment of the outputs. The production of a level 4 reanalysis relies on a number of external projects aimed at producing lower level data. The various data inputs provided by these projects are described in section 2. The method used to produce the reanalysis is described in section 3 with a focus on the differences to the NRT OSTIA system, together with a description of the additional quality control performed. Various aspects of the reanalysis data are investigated in section 4, including an assessment of the accuracy of the reanalysis in various regions, the impact of aerosols from the Mount Pinatubo eruption on the reanalysis, assessment of the high-latitude regions where sea-ice is important, comparison with other longer-term SST reanalyses (particularly the HadISST and Reynolds products), comparison with the NRT OSTIA system, and an assessment of the new OSTIA climatologies which have also been produced. Section 5 summarises the results and describes the work which is planned to improve the OSTIA reanalysis in the future.

1.2 OSTIA reanalysis product suite

Several products are available from the OSTIA reanalysis to fulfil different users requirements. Products are produced in netcdf GHRSST L4 format, see (GHRSST-PP, 2005). High resolution refers to products on the native $1/20^{\circ}$ OSTIA grid whilst reduced resolution refers to a $1/4^{\circ}$ regular latitude-longitude grid.

- High resolution daily OSTIA reanalysis. Files contain the foundation SST, SST analysis error, sea-ice concentration and land/sea/ice mask fields.
- Reduced resolution daily anomaly. Files contain the SST anomaly from the Pathfinder climatology (REF), foundation SST and the sea-ice concentration fields.
- High resolution daily OSTIA reanalysis climatology. Files contain the daily mean foundation SST, SST standard deviation, maximum SST, minimum SST, mean sea-ice concentration, sea ice standard deviation and land/sea/ice mask fields.
- Reduced resolution daily OSTIA reanalysis climatology. Files contain the daily mean foundation SST, SST standard deviation, maximum SST, minimum SST, mean sea-ice concentration, sea ice standard deviation and land/sea/ice mask fields.

- High resolution monthly OSTIA reanalysis climatology. Files contain the monthly mean foundation SST, SST standard deviation, maximium SST, minimum SST, mean sea-ice concentration, sea ice standard deviation and land/sea/ice mask fields.
- Reduced resolution monthly OSTIA reanalysis climatology. Files contain the monthly mean foundation SST, SST standard deviation, maximum SST, minimum SST, mean sea-ice concentration, sea ice standard deviation and land/sea/ice mask fields.

2. Input data sources

2.1 Overview

Figure 2.1 presents a timeline showing the observational datasets included in the OSTIA reanalysis v1.0. The AVHRR Pathfinder (Advanced Very High Resolution Radiometer) and ICOADS (International Comprehensive Ocean-Atmosphere Dataset) in situ datasets have been used throughout the entire period. In the case of the ATSR (Along Track Scanning Radiometer) series, different versions of the data become available with the introduction of new instruments. In the overlap periods between these different missions the newest version of the data has been used and any gaps in the newer timeseries within the overlap periods have not been filled with data from the old timeseries. This method allows for easier investigation into the effect on the reanalysis of the different data sources. The OSI-SAF sea ice concentration reprocessing dataset was also used in the reanalysis, where data from the SMMR instrument was replaced with SSM/I from 1987. The ftp data download locations and dates the data were obtained are given in table 2.1 and full descriptions of each dataset are provided in the following sections. In order to maintain consistency of the reanalysis, new satellite datasets have not been included as they become available through the reanalysis period.

Table 2.1: Data sources (all acronyms given in glossary)

Data type and version number	Source location	Latest download date
AVHRR Pathfinder v5.0	data.nodc.noaa.gov	September 2010
(A)ATSR multi-mission series version 2.0	ftp.neodc.rl.ac.uk	September 2010
ICOADS in situ Release 2.1 updated from 1998	Met Office Hadley Centre (after processing)	January 2010
OSI-SAF sea ice reprocessing	osisaf.met.no	July 2010

Table 2.2 highlights the key differences between the data types used in the reanalysis and oper-

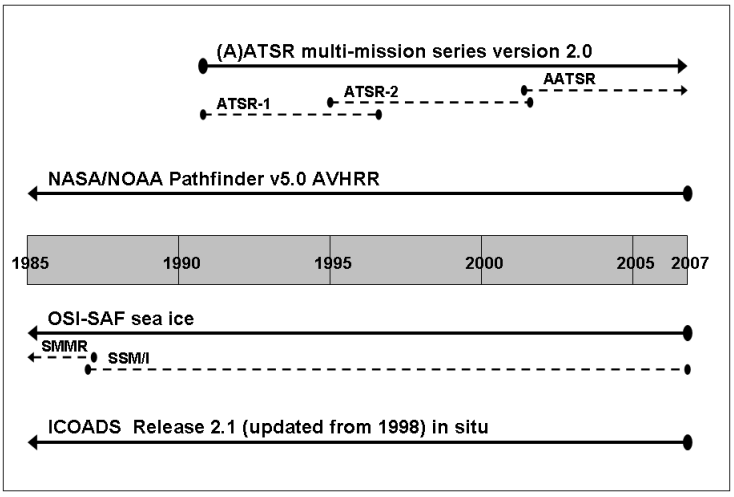


Figure 2.1: Timeline of availability of observational datasets for the OSTIA reanalysis period. All acronyms given in glossary.

ational versions of OSTIA. Timeseries of the number of observations for each of the SST datasets used in the reanalysis are shown in figure 2.2. Clearly the AVHRR Pathfinder has the greatest number of observations and, in comparison to the satellite datasets, in situ data volumes are very small. The in situ dataset is however very important as it is used as a reference dataset for bias-correction of the AVHRR and ATSR-1 datasets. A small proportion of days in the reanalysis period have limited or no data available for one or more data types. This is due to missing data files, missing data within files or days when a large number of observations failed the QC procedures (which are described in section 3.2). A complete list of days without data for each data source during the reanalysis period can be found in Appendix A.

Figure 2.3(a) gives the total number of observations for each $1/4^o$ grid box for the whole reanalysis period (1985-2007). Areas which are predominantly cloud-free, such as the equatorial regions, the Mediterranean, and the subtropical high pressure zones, have considerably more observations than those regions where cloud cover is more prevalent, such as the ITCZ (inter-tropical convergence zone), and the subpolar low pressure zones. Owing to the differing magnitudes of the number of satellite-based and in situ observations (figure 2.2), figures 2.3(a), (c) and (d) represent in effect the spatial variation of the satellite data only.

The global coverage of the ICOADS in situ dataset, which includes observations from drifting buoys, moored buoys and ships, also shows spatial variation (figure 2.3(b)). The greatest number of in situ observations are found in the North Atlantic Ocean and thus the number of observations in the northern hemisphere is larger than for the southern hemisphere. Shipping tracks can be seen in, for example, the northern part of the Indian Ocean and the mid-Atlantic and provide increased coverage in these in situ data sparse regions. There is also in general a relatively greater coverage

Table 2.2: Input data differences between the reanalysis and operational OSTIA systems (see glossary for acronyms)

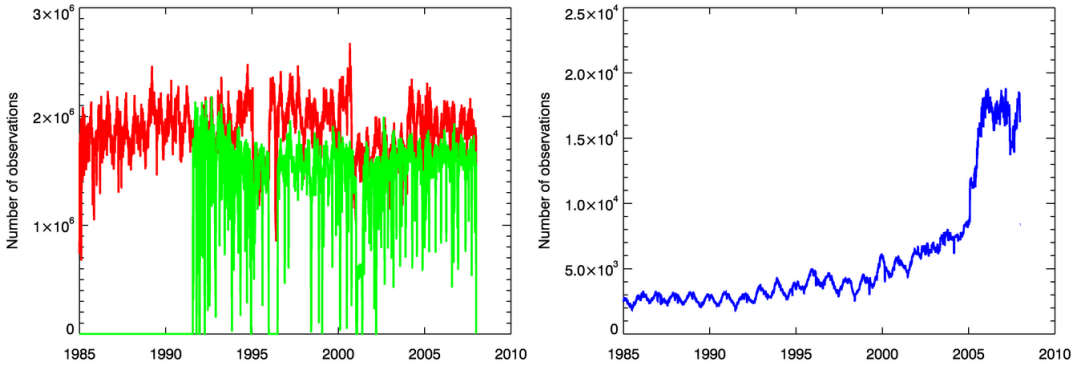
Reanalysis	Operational
ATSR-1, ATSR-2, AATSR (not subsampled)	AATSR (subsampled)
AVHRR Pathfinder (not subsampled)	NOAA-18 AVHRR, MetOp AVHRR (subsampled)
No other satellite data types	SEVIRI, NAR, TMI, AMSR-E (subsampled)
ICOADS in situ, nighttime only	In situ via GTS, 24 hrs
OSI-SAF reprocessed ice concentration (SMMR and SSM/I)	OSI-SAF operational ice concentration (SSM/I)

in the higher latitudes than for the satellite data, although coverage close to the ice edge is poorer (figure 2.3).

Figures 2.3(c) and 2.3(d) provide a comparison of the total number of observations for each grid box for two different years: one year from before the introduction of ATSR series data to the reanalysis (1990) and one afterwards (2004). The ATSR series data increases the number of observations globally, but particularly in the equatorial regions and mid-latitudes. The number of observations at higher latitudes is also increased, see for example the north Atlantic Ocean and Southern Ocean.

Figures 2.3(e) and 2.3(f) show the total number of in situ observations for 1985 and 2007 respectively. These figures illustrate the improved global coverage of in situ observations at the end of the reanalysis period compared to the start, when the number of observations in the northern hemisphere was significantly greater than in the southern hemisphere. It is interesting to note the change in the pattern of observations, from the straight lines characteristic of ship tracks in 1985 (figure 2.3(e)) to the curved lines of drifters following current structures in 2007 (figure 2.3(f)). It is also notable that in 2007 the number of in situ observations in the Atlantic Ocean is greater than in the Pacific Ocean (figure 2.3(f)). There is also a relatively large number of in situ observations in 2007 around the Antarctic Peninsula and in the Arctic north of Scandinavia (figure 2.3(f)), presumably related to IPY (International Polar Year) activities at this time.

Figure 2.4 shows the total number of observations for each $1/4^o$ grid box by season for the entire reanalysis period (1985-2007). There is a marked seasonal difference in the number of observations in the polar regions as all observations under ice are masked out (section 2.2) and the ice extent varies seasonally. The southern hemisphere region of multi-year ice in the Weddell Sea can be seen in the southern hemisphere summer and autumn (December-February and March-



(a) Number of observations for satellite data. Red is AVHRR Pathfinder and green is (A)ATSR multi-mission series.

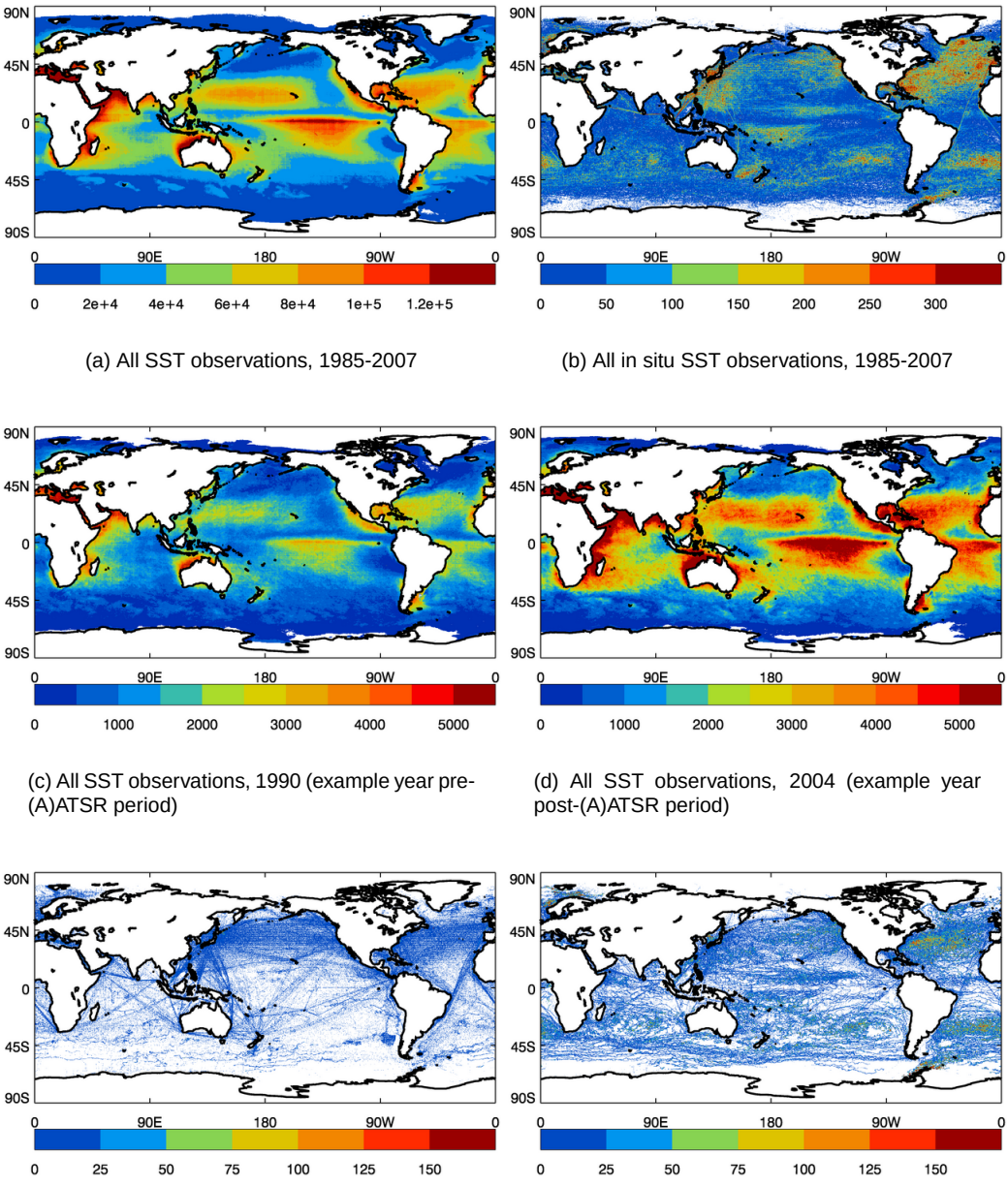
(b) Number of observations for ICOADS in situ data.

Figure 2.2: Timeseries of number of observations for each SST dataset used in OSTIA reanalysis v1.0 (see text). Smoothed using a 5 day rolling mean for clarity. Note the difference in scales between figures (a) and (b).

May, figures 2.4(a) and 2.4(b)) as the only region to not have any SST observations. The Ross Sea polynya can also be identified in the southern hemisphere melt season as a region within the ice pack where SSTs are accepted (figure 2.4(d)). The northern part of the Indian Ocean also displays significant variation between seasons in the number of observations available. In the northern hemisphere winter (figure 2.4(a)) there are several thousand more observations available in this region for each grid box than in the summer months (figure 2.4(c)). This is related to the behaviour of the Indian Monsoon, as the change in prevailing winds from north-easterly in winter, bringing dry continental air, to south-westerly in summer, bringing humid maritime air, results in a significant increase in precipitation for this region. As satellite infra-red radiometers (e.g. AVHRR, (A)ATSR) are unable to 'see' through cloud, the number of SST observations in the summertime is markedly reduced.

A seasonal effect on the number of observations available in each grid box can be observed in other regions, e.g. north-east Australia, where increased rainfall in the austral summer (December to February, figure 2.4(a)) is related to a prevailing maritime flow. Other regions have consistently few numbers of observations, for example off the coast of Peru, owing to cold upwelling which leads to the formation of marine stratocumulus cloud, thus preventing satellite infra-red radiometer SST measurements.

The following sections provide details of the different observational datasets used in the OSTIA reanalysis. The data sources have been discussed in some detail as the accuracy of the reanalysis is highly dependent on the input data.



(a) All SST observations, 1985-2007

(b) All in situ SST observations, 1985-2007

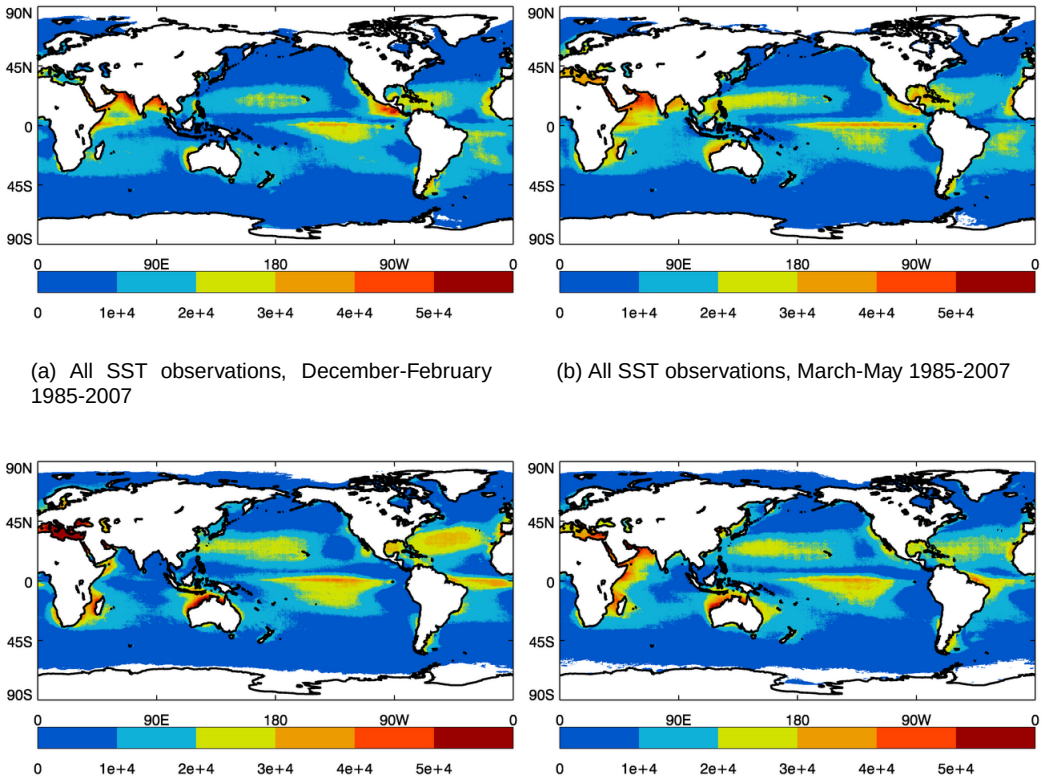
(c) All SST observations, 1990 (example year pre-(A)ATSR period)

(d) All SST observations, 2004 (example year post-(A)ATSR period)

(e) All in situ observations, 1985

(f) All in situ observations, 2007

Figure 2.3: Total number of observations for each 1/4^o grid box (note differing scales).



(a) All SST observations, December-February 1985-2007

(b) All SST observations, March-May 1985-2007

(c) All SST observations, June-August 1985-2007

(d) All SST observations, September-November 1985-2007

Figure 2.4: Total seasonal number of observations for each 1/4^o grid box.

2.2 AVHRR Pathfinder

The NOAA/NASA Oceans Pathfinder SST dataset used in the OSTIA reanalysis is a daily global composite of AVHRR data compiled from individual satellite overpasses (NODC, 2009; Kilpatrick et al., 2001). The measurements over the timeseries are derived from different AVHRR instruments onboard NOAA 7, 9, 11, 14, 16, 17 and 18 polar orbiting satellites (NODC, 2009). The SST algorithms derived for AVHRR have undergone various modifications over the time of these instruments and hence Pathfinder was developed to reprocess the database using a consensus SST algorithm to produce a consistent timeseries (Kilpatrick et al., 2001). A consensus calibration procedure was also used to minimise spurious trends and to ensure homogeneity across the sensors aboard different spacecraft (Kilpatrick et al., 2001).

The Pathfinder SST algorithm (PFSST) is based on the NOAA/NESDIS nonlinear operational algorithm (NLSST) (Kilpatrick et al., 2001) and given in full by Robinson (2004). The PFSST algorithm is the most established for reprocessing of the AVHRR archive (Evans and Podesta, 1998). It was developed to achieve relatively uniform performance over a wide range of atmospheric and oceanic conditions (Kilpatrick et al., 2001) and provides a reliable long term record of SST (Robinson, 2004).

There are a variety of other satellite-derived global SST products with different grid sizes, integration times, data selection and gap-filling strategies for cloud cover but, unlike Pathfinder, they are derived for short-term operational requirements (Robinson, 2004).

Subsequent versions of the PFSST algorithm have undergone improvements (Kilpatrick et al., 2001). The most recent version at the time of writing, Pathfinder v5.0, has been used for the OSTIA reanalysis on a daily time period with a spatial resolution of 4 km on an equal-angle longitude-latitude grid (NODC, 2009).

The actual definition of a 'day' employed in these files not only encompasses a temporal component (e.g. 24 hours) but also a spatial one (e.g. full global coverage) (Kilpatrick et al., 2001). This is to avoid large discontinuities at the time or space boundary. The spatial component is defined as the "time at which the satellite orbit track crosses the 180 ^{o} meridian nearest to the equator" (Kilpatrick et al., 2001). Using only the satellite orbit to define the start and end times of the data day would result in large discontinuities at the meridian boundary as the orbit precesses (Kilpatrick et al., 2001). To avoid this issue, the temporal component of the data day includes any data taken within 2 hours before or after the crossing of the meridian. Hence a data day in the Pathfinder files may include data over a 24 to 28 hour period (Kilpatrick et al., 2001).

The AVHRR instrument uses three infrared channels to estimate SST located in wavelength regions where the atmosphere is relatively transparent (Kilpatrick et al., 2001). The accuracy of the brightness temperature measurements from individual channels is between $0.1\ K$ and $0.2\ K$ (Robinson, 2004). The entire surface of the Earth is viewed at least twice a day; once from the ascending (daylight) overpasses and once from the descending (night) passes (Robinson, 2004). The swath of adjacent orbits overlap significantly away from the tropics, meaning at high latitudes locations are observed several times a day (Robinson, 2004). Robinson (2004) gives full details of AVHRR instrument channels and NOAA satellite orbits.

In the PFSST algorithm the SST assigned to each cell of the composite grid is based on the average of all the samples obtained within the cell during the integration period, after cloud screening and atmospheric correction (Robinson, 2004). Robinson (2004) also notes that for pixels with scarce observations, the composite value may be biased by overdependence on too few overpasses or that the observations may not be spatially representative of the full width of the cell. The time of observation for each pixel is not provided as part of the dataset since the final SST value is a combination of multiple observations from that particular data day, and only those of the highest quality level available (see below). In addition, drift in the orbit times makes it difficult to associate an exact time with a particular observation throughout the measurement period (Kilpatrick et al., 2001).

Version 5.0 of the PFSST algorithm includes the use of empirical coefficients which are revised every month throughout the data record to correct for slow changes in the detector response, drifts in the satellite orbit and variations of stratospheric aerosol. The coefficients are estimated using a regression of remotely sensed brightness temperatures on in situ temperature measurements

from drifting and moored buoys from a match-up database (Kilpatrick et al., 2001). Two different sets of coefficients are used according to the levels of atmospheric water vapour (Robinson, 2004). The same algorithm and sets of coefficients are applied to both daytime (ascending) and nighttime (descending) data, although the daytime and nighttime files are produced separately (Kilpatrick et al., 2001). OSTIA is an analysis of the foundation SST and thus descending pass files (nighttime) only were used in order to minimise the diurnal warming signal.

The global accuracy of the Pathfinder algorithm is $0.2 \pm 0.5^{o}C$ on comparison with in situ buoys, after atmospheric correction (Kilpatrick et al., 2001). The relatively large uncertainty is partly explained by the method of tuning the algorithms to buoy temperature measurements, as the difference between bulk and skin SSTs is variable (Robinson, 2004). The performance of the PFSST algorithm is dependent on how close the atmospheric and oceanic conditions are to the mean state and is therefore likely to depend on the geographical location and season of the measurement (Kilpatrick et al., 2001). The algorithm is also unable to completely correct for the presence of high-level atmospheric aerosols injected during major volcanic eruptions (Robinson, 2004). The impact on the OSTIA reanalysis of the effects of the 1991 Mount Pinatubo eruption (Reynolds, 1993; Merchant et al., 1999) are discussed in section 4.2. Biases related to aerosol contamination from Saharan dust storms are also a problem (NODC, 2009). However, as discussed in sections 3.4 and 2.1, in the OSTIA reanalysis the Pathfinder data is bias-corrected to in situ data (and, post June 1995, to ATSR series data) which should reduce the impact of these errors on the reanalysis.

The processing of the Pathfinder data includes the use of decision trees to objectively determine cloud contamination of pixels and the quality level of the SST retrieval (Kilpatrick et al., 2001). One of eight possible quality flags between 0 (low) and 7 (high) is assigned to each pixel following a series of quality tests. Full details of these tests are given by Kilpatrick et al. (2001), who note the quality flags are provided only as guidance and are not associated with specific error levels for estimates of SST. SST values to be composited into a single grid cell may be associated with different quality flags. In this case, those values with the highest flag are averaged, which sets the flag for the cell, and the other values are discarded (Robinson, 2004). SST estimates with a flag of 4 should be the minimum quality threshold used for most ocean scientific applications (Robinson, 2004). These pixels will have passed a cloud contamination test and a reference test, which are failed if the data have a difference of greater than 2 ^{o}C from a 3 week running mean reference SST field (Kilpatrick et al., 2001). This method is specified so that departures of more than 2 ^{o}C from climatology are not excluded but short-lived issues are eliminated (Robinson, 2004).

It was found through investigations of a preliminary run of the OSTIA reanalysis that using flags of only the highest quality meant the Pathfinder dataset was too sparse, particularly at high latitudes. The introduction of a minimum quality flag of 4 did not greatly affect the OSTIA reanalysis global error statistics and thus the decision was taken to include all Pathfinder data with a flag of 4 and above. Section 4.1 gives a full discussion of the use of these error statistics for validation purposes. It should also be noted here that, as discussed in section 2.1, the data also underwent our own

quality control procedures including a maximum and minimum value check and a background check against the previous day's analysis after relaxation to climatology (section 3.3), before inclusion in the OSTIA reanalysis. The standard deviation provided for each cell with the Pathinder data was used as the observation error in the OSTIA reanalysis. For instances where the value in the cell was derived from few observations and hence had a standard deviation close to 0, an error of 0.33 K was assumed, following Kearns et al. (2000).

Pathfinder v5.0 includes some information on sea ice. This is used in the Pathfinder data processing procedure to exclude pixels from the internal 3 week reference SST check mentioned above. The aim of this approach is to reduce the likelihood of mis-classified SST pixels, for example those with high quality levels but falling on the ice mask (NODC, 2009). The ice data is not used to specifically flag the data as poor quality, although there are plans to include this in future versions of the Pathfinder dataset (NODC, 2009). Thus ice-covered pixels should only be assigned quality flags of 3 or below. However, it was found during investigations of a preliminary run of the OSTIA reanalysis that this processing was unsuccessful in the Arctic in the boreal summer. A considerable number of pixels in the ice-covered region are flagged as high quality in the summertime, see section 3.2 for a description of how the OSTIA system quality controlled these erroneous SST observations.

In addition, NODC (2009) notes the presence of erroneously cold SST values in high latitudes in the Pathfinder dataset and recommends masking them out. As mentioned in section 3.2, a minimum SST of -2.0^oC was set for all datasets used in the OSTIA reanalysis which removes the spurious values.

2.3 (A)ATSR multi-mission series

The (A)ATSR multi-mission archive is a ESA/NEODC project to produce a 15-year archive of sea surface temperatures using the (A)ATSR series of instruments at levels of accuracy of 0.3 K or better (NEODC, 2011). The dataset begins in 1991 with the launch of the ATSR-1 instrument, onboard the ERS-1 satellite. This was followed in 1995 by the ATSR-2, onboard ERS-2 and AATSR onboard Envisat in 2002.

The ATSR series of instruments are multichannel radiometers which have three wavebands in the infra-red part of the spectrum (Robinson, 2004). All three ATSR instruments employ the same onboard calibration procedure for the thermal channels, whereby two high precision blackbodies are viewed every scan cycle (NEODC, 2011; Robinson, 2004). The temperatures of the blackbodies approximately straddle the ground temperature range to precisely calibrate the brightness temperatures for each channel (Robinson, 2004). A cooling procedure is used to lower the detector temperature, thereby improving the stability of the calibration procedure (Robinson, 2004). The instruments employ a unique conical scanning geometry, using a single detector to view the same area of ground twice; once at a forward angle of around 55^o and once at nadir (around

0^o) (Robinson, 2004). This means that, as well as a multispectral approach, atmospheric correction procedures can use the dual view method (Robinson, 2004). Unlike the atmospheric correction method used for AVHRR (section 2.2), in situ calibration measurements from ships and buoys are not required as radiation transfer models are used (Robinson, 2004). A consistent algorithm for both day and night observations is produced (Robinson, 2004).

Separate brightness temperature products are provided for the forward and nadir views. The nadir image produces the most sensitive measurement of SST, where the radiometric resolution is better than 0.1 K (Robinson, 2004). A major drawback however of the scanning arrangement is that the width of the swath is severely limited to about 500 km (Robinson, 2004). Complete global coverage therefore takes at least 4 days (Robinson, 2004). Validation using bulk SST measurements or aircraft radiometry has shown the original algorithms account for the variability of atmospheric effects to better than 0.3 K (Robinson (2004) and references therein).

Archive AATSR, ATSR-2 and ATSR-1 products have been reprocessed by ESA/NEODC using the current operational processing procedures to produce a common format archive. Further information on changes and improvements made to the datasets during the reprocessing can be found in NEODC (2011). A correction for the overlap of satellite orbits was applied for the AATSR data but at the time of running the OSTIA reanalysis v1.0 this had not yet been applied to the ATSR-2 and ATSR-1 datasets.

When available, ATSR-2 and AATSR data were used as a reference dataset in the analysis bias correction scheme along with in situ data (section 3.4). The calibration procedures and the dual view sensors described above make this dataset ideal for use as a reference dataset over other satellite instruments. However, owing to the problems of stratospheric aerosol resulting from the Mount Pinatubo eruption, after investigation the ATSR-1 data were not used as reference data but were included in the reanalysis and bias-corrected against in situ data, as for the AVHRR data. Section 4.2 provides more information on this investigation. When used as a reference dataset, only data flagged as the highest quality (flag 5) were used but, for the ATSR-1 period data with flags of 3, 4 and 5 were included in the reanalysis, and subsequently bias-corrected against the in situ data. Further information on the use of flags for the ATSR-series data is given in section 4.2. Occasional outages in the (A)ATSR multimission series data are listed in Appendix A. The bias and standard deviation statistics provided with the (A)ATSR series data SSES (Single Sensor Error Statistic) were used in the OSTIA reanalysis; the bias as a correction prior to the assimilation of the data and the standard deviation as the observation error.

2.4 ICOADS in-situ

ICOADS (International Comprehensive Ocean-Atmosphere Data Set) is the largest and most extensive set of in situ marine observations available, containing observations from ships and buoys from diverse data sources and spanning several centuries (Worley et al., 2005; Woodruff et al., 2010).

The dataset has provided input to many climate products and research applications and has been used in previous reanalysis products as the (or a significant part of) reference dataset for the ocean surface (Worley et al., 2005; Woodruff et al., 2010, and references therein). Thus the dataset is ideal for use as a reference for the OSTIA reanalysis.

The ICOADS QC (quality control) procedure includes consistency checks, checks against climatologies, trimming of outliers, position errors against a land-sea mask and addition of QC flags to the record (Worley et al., 2005). Additional QC was performed by the UK Met Office Hadley Centre, comprising of basic QC checks for valid date, time and position, a day/night check based on the solar zenith angle, a positional QC ship track check, a comparison to climatology, a freezing check and a buddy comparison with other nearby observations. Before inclusion in the OSTIA reanalysis the data also underwent a maximum and minimum temperature check and a background check using the previous day's analysis.

Each ICOADS release has been developed incrementally from previous releases and has been successively improved since the release of the initial version in 1985 (Worley et al., 2005). Although Release 2.5 is the most recent version at the time of writing, a version of Release 2.1, updated from 1998 by the UK Met Office Hadley Centre using NCEP GTS data, was used for the OSTIA reanalysis as this was the most recent version to have undergone the full QC procedures.

To produce the best possible set of observations a complex and robust blending and duplicate elimination process is used in ICOADS to remove overlaps and duplication of observations between some of the sources, as well as to correct issues such as new versions of datasets which only partially replace older versions, and where some parts of sources are of inferior quality (Worley et al., 2005; Woodruff et al., 2010). ICOADS does not include adjustments to account for observing system changes (e.g. instrumentation, sampling procedures and measurement coding changes) or measurement biases (Worley et al., 2005). However, in comparison with the variation in observation methods across the entire ICOADS Release 2.1 dataset (from 1784), heterogeneities in observing systems during the OSTIA reanalysis period of 1985-2007 are small. In addition, owing to denser, more accurate sampling, the uncertainties in the dataset should decrease closer to near-current dates (Worley et al., 2005).

Figure 2.5 shows the number of nighttime in-situ observations used in the reanalysis throughout the period. In order to produce the foundation SST, nighttime data only were used for the OSTIA reanalysis. For ICOADS, this is set by the solar zenith angle. If the sun was above the horizon one hour previous to the measurement time, the observation is classified as daytime and is not used. The number of ship-based observations has decreased over time (figure 2.5) while the contributions from moored and drifting buoys have increased, particularly rapidly for drifters from the mid-1990s onwards. The step increase in the number of drifter observations in 2005 is possibly due to an increase in the frequency of reporting times as well as a potential increase in the number of drifters themselves. As noted by Woodruff et al. (2010), each observing platform has its advantages. VOS (Voluntary Observing Ships) provide regular sampling along the major shipping routes and research

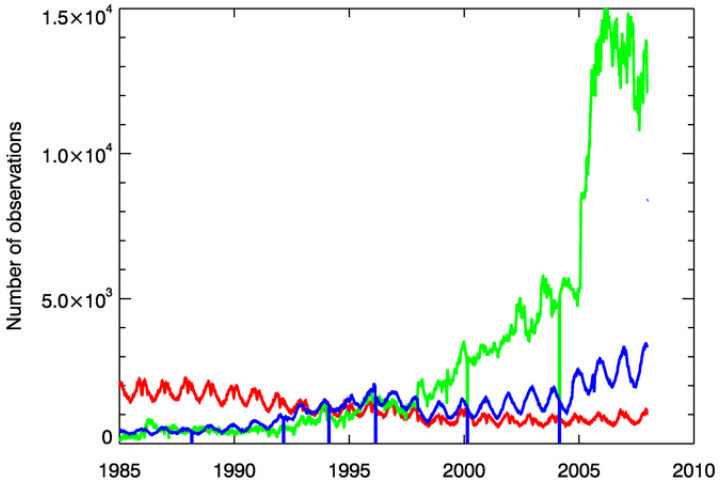


Figure 2.5: Daily number of in situ observations included in the OSTIA reanalysis, split by observing platform, where green is drifters, red is ships and blue is moored buoys smoothed by 5 day rolling mean.

vessels visit a wide range of areas, though return to the same areas less regularly than VOS. Only small regions of the ocean are sampled by moored buoys but the measurements tend to be of high frequency and accurate and the number of instruments is increasing over time. Coverage by drifting buoys is good, but drift in the measurement sensor can result in inaccuracies as, in general, they are not recovered and thus no calibration is carried out (Woodruff et al., 2010).

Automated platforms such as moored or drifting buoys are able to provide higher frequency observations, at 3-hourly or hourly intervals, than ships which report at 6-hourly synoptic periods (Worley et al., 2005). Therefore, from 1990 onwards the total number of observations increases as the dominant observing platform gradually changes from ships to drifting buoys.

There is a distinct seasonal cycle in the number of nighttime in situ observations available (figure 2.5), i.e. there are fewer measurements available during the northern hemisphere summer than the winter. This is due to the shorter nighttime period in the summer and, as there are many more northern hemisphere observations (figure 2.3), the global trend is biased towards the northern hemisphere seasonality in number of observations. Owing to an error in the date QC, the data for 29th February in leap years has erroneously been flagged as bad and was not used in the reanalysis. This will be corrected in future versions of this dataset.

All in situ data types were used as a reference for the OSTIA reanalysis although data from ships are of lower quality than those from moored or drifting buoys (section 4.1.1). This use of ship data is consistent with the operational OSTIA system and in addition, prior to the mid-1990s, few drifting and moored buoy observations are available (figure 2.5). No information was available to bias correct the in situ data but it is possible this could be provided by ICOADS or the UK Met Office Hadley Centre in the future.

2.5 OSI-SAF sea-ice concentration reprocessing

A daily sea ice concentration field is provided with the OSTIA reanalysis. This is used in to update the background SST field (see section 3.3) with the aim of improving the consistency between the SST and the sea ice. The ice concentration field is also used in the quality control procedure to mask out observations in all of the datasets for regions which appear to be under ice (section 3.2). This is particularly important for the Pathfinder data owing to an issue of erroneous data in the Arctic in the boreal summer, discussed in section 2.2.

The OSTIA ice field is derived from the OSI-SAF sea ice concentration reprocessing dataset between 1985 and 2007 (Eastwood et al., 2010) by regridding from a 10 km resolution polar stereographic grid on to the OSTIA $1/20^{o}$ equal-area grid. A QC procedure was performed by eye on the ice concentration files before inclusion in the reanalysis to remove files with regions of missing or poor data. For days when the ice concentration field was not available due to satellite malfunction or planned maintenance (Tonboe and Nielsen, 2010), or did not pass our QC checks, the previous day's ice concentration field was persisted. Appendix A gives a complete list of missing days.

For large gaps in the data (7 days or more), the first file available after the outage has been copied to the date in the middle of the outage. This is because (making the assumption that the growth/retreat of ice extent can be approximated as linear) after the halfway point the ice concentration field will be closer to that in the later file, and less close to that persisted from the file at the start of the outage. This aims to reduce large jumps in the SST and sea ice consistency. This method could be improved, for example by using an interpolation method for the sea ice concentration in the data gap to produce a daily file. Analysis of the effectiveness of this procedure is included in section 4.4.2.

Areas of missing data in the ice concentration field due to missing scan lines, missing orbits and the polar observation hole were filled by OSI-SAF. This was accomplished using an algorithm employing both spatial and temporal interpolation (Eastwood et al., 2010). However, for this version of the reprocessed ice concentration data there still remained missing data within some sea ice files. For missing regions enclosed within the ice concentration field a spatial 2-D bilinear interpolation was used during the OSTIA regridding process. For missing regions which extended beyond the ice edge the region could not be filled using this method and the file was rejected. Over open water the effect of atmospheric noise on the passive microwave sensor increases the ice concentration above 0% (Tonboe and Nielsen, 2010). Therefore data with a concentration of less than 15% was set to 0% for the OSTIA reanalysis.

The OSI-SAF sea ice concentration dataset is comprised of data from the SMMR (Scanning Multichannel Microwave Radiometer) instrument between January 1985 and July 1987 and the various SSM/I (Special Sensor Microwave/Imager) instruments between July 1987 and December 2007. Lists of the satellites carrying the instruments and overlap periods between instruments are given by Eastwood et al. (2010). The SMMR instrument, aboard the Nimbus 7 satellite, operated

between October 1978 and August 1987 (Gloersen et al., 1997). Owing to power supply limitations the instrument was operated every other day. There is no coverage poleward of 84 degrees owing to the swath width and satellite orbit inclination. Eastwood et al. (2010) and references therein provide a full description of this instrument. The OSI-SAF dataset uses version 6 of the Remote Sensing Systems (RSS) SSM/I dataset which covers the period of available satellites between 1987 and 2007. The RSS data incorporates sensor calibration, quality control procedures and geolocation corrections. Intercalibration between the different satellites during overlap periods has also been performed (Eastwood et al., 2010). Therefore the SSM/I data is expected to be more accurate than the SMMR data used for the earlier period. During the overlap period between the two instrument types in July and August 1987 the SSM/I data has been used for the OSTIA reanalysis, following the recommendations of Eastwood et al. (2010).

Before being used for the calculation of ice concentration, the surface brightness temperature data for both SMMR and SSM/I were corrected by OSI-SAF for the influence of wind roughening over open water and atmospheric water vapour using a radiative transfer model and NWP data. This procedure is described in Andersen et al. (2006b). Three different ice concentration algorithms were used in the processing of the passive microwave data record: the TUD (Technical University of Denmark) from 1991 onwards (Andersen et al., 2006a) used on the high frequency channels, the Bootstrap algorithm in frequency mode (Comiso, 1986; Comiso et al., 1997) and the Bristol algorithm (Smith, 1996). The Bootstrap and Bristol algorithms were used in a linear combination as a hybrid algorithm throughout the data record for the lower frequency channels. To retain the advantages of each of the two algorithms, the Bristol algorithm was given little weight at low ice concentrations and the Bootstrap frequency mode greater weight, and vice versa over high ice concentrations. Hence over open water the Bootstrap algorithm is used and the Bristol over ice. The ice concentration at intermediate concentrations up to 40% is a linearly-weighted average between the two algorithms (Eastwood et al., 2010).

The daily ice concentration files contain all observations obtained within 24 hours, centred on 12:00 UTC (Eastwood et al., 2010). The observations within one grid cell are averaged and weighted according to the distance between the observation and the centre of the gridpoint, and the radius of influence, dependent on the sampling radius which varies for different channels. Observations to be averaged are available from multiple satellite missions. During the SSM/I period, observations from different satellites are averaged and potential inter-satellite calibration differences are handled by the dynamical tie-point approach described in Eastwood et al. (2010). This method also minimises the effects on the ice concentration values of sensor drift and climatic trends in atmospheric emission and surface emissivity (Eastwood et al., 2010). A coastal correction step was also performed and a monthly maximum ice extent climatology was applied to remove erroneous data (Eastwood et al., 2010). During the overlap between SMMR and SSM/I the data are not mixed as there is a greater difference between these two instruments than between different SSM/I instruments (Eastwood et al., 2010).

The OSI-SAF processing flags provided with the dataset were not used in the OSTIA reanalysis, as although they provide an indication of corrections applied or questionable data they are not quality estimates nor are they used to modify the ice concentration values themselves. The 2 m air temperature check, indicating the possibility of false ice due to a surface air temperature above +10^oC was investigated but had a negligible impact on the OSTIA reanalysis. Ice concentration uncertainty estimates are not provided with the OSTIA reanalysis v1.0 data. OSI-SAF have provided an error estimate for the sea ice concentration data in the form of a standard deviation for each data point, derived from different error sources described in Eastwood et al. (2010), including representativeness, temporal sampling and random instrument noise. The OSI-SAF error estimates cannot be directly interpolated on to the OSTIA grid as the representativeness error is a function of grid size. However, it is planned to include an estimate of the ice concentration errors in future versions of the OSTIA reanalysis.

A validation of the OSI-SAF ice concentration reprocessing dataset against weekly NIC (National Ice Center) ice charts was carried out by Tonboe and Nielsen (2010). The accuracy of the OSI-SAF dataset based on this comparison was shown to be within 10-20% on a yearly average, although this figure has a strong seasonal dependence. The accuracy of sea ice concentration retrieval algorithms is known to be seasonally dependent (Robinson, 2004). For the best case, when there are only one or two ice types present, the typical error of sea ice concentration retrieval algorithms is an underestimate of ice concentration of 4% (Robinson, 2004). The concentration may be underestimated by up to 30% in summer when meltponds are present, or in early autumn when it is difficult to distinguish young ice from open water (Robinson, 2004).

During the brief overlap period between the SMMR and SSM/I instruments the data was compared by Tonboe and Nielsen (2010). The bias between the two instruments was found to be small (within a few percent) although SMMR estimates were higher along coastlines. Further validation of the OSI-SAF/OSTIA sea ice concentration is included in the results section (section 4.4).

3. Description of the OSTIA reanalysis system

The OSTIA reanalysis system produces a daily, global SST and sea-ice analysis on a high resolution $1/20^{\circ}$ regular latitude-longitude grid for the period from 1st Jan 1985 to 31st Dec 2007. OSTIA provides an estimate of the foundation SST, which is defined to be the SST free of diurnal warming (Donlon et al., 2002). The foundation SST is representative of the temperature throughout the mixed layer just before sunrise. See section 3.2 for a description of the quality control carried out to try and get an estimate of foundation SST by the OSTIA system for the different data sources.

The OSTIA reanalysis has been designed to provide a continuous, homogenous, daily SST analysis by combining the different sources of observations. The day to day variability of the reanalysis fields should reflect where possible the daily variability of ocean SST features. It is also intended that the SST and sea-ice fields for a given day will be consistent with one another. The OSTIA reanalysis is intended for use by amongst others, seasonal and decadal forecasting applications, ocean and atmosphere reanalysis producers and also by the climate prediction community to validate climate models and for climatic monitoring. These users require a SST analysis to be as homogenous as possible which led to the use of only those satellite data sources that were available for a large proportion of the reanalysis time period. This has led to the OSTIA reanalysis being a distinct product from the operational near real time OSTIA analysis.

The resolution of the OSTIA reanlaysis grid (approximately 6km) is one limit on the ability of the SST analysis to resolve fine scale SST features. The feature resolution of any SST analysis is limited by the availibility of the high resolution observations. In the OSTIA reanalysis the infrared satellite data can provide observations to a resolution of up to 1km to supplement the relatively sparse in-situ network (see section 2.4). However these observations do not provide daily global coverage due to the width of the swaths and cloud cover. As no physical model is used the analysis relies on statistical methods to spatially propogate the observational information. This is carried out via the background error covariance matrix, described in section 3.5. The shortest correlation length scale of which is 10km which limits the feature resolution of the OSTIA reanalysis further. Due to limits on the number of iterations of the assimilation scheme carried out (currently 10) the analysis doesn't converge to the optimal solution, this will also affect the feature resolution. The

specification of these statistical parameters is a compromise between being able to resolve fine scale ocean features and the daily variability of the analysis reflecting that in ocean features rather than changes in the observational coverage. Within the OSTIA reanalysis ocean regions which remain unobserved for long periods of time will eventually revert back to climatological SST values, see section 3.3.

3.1 System architecture and overview

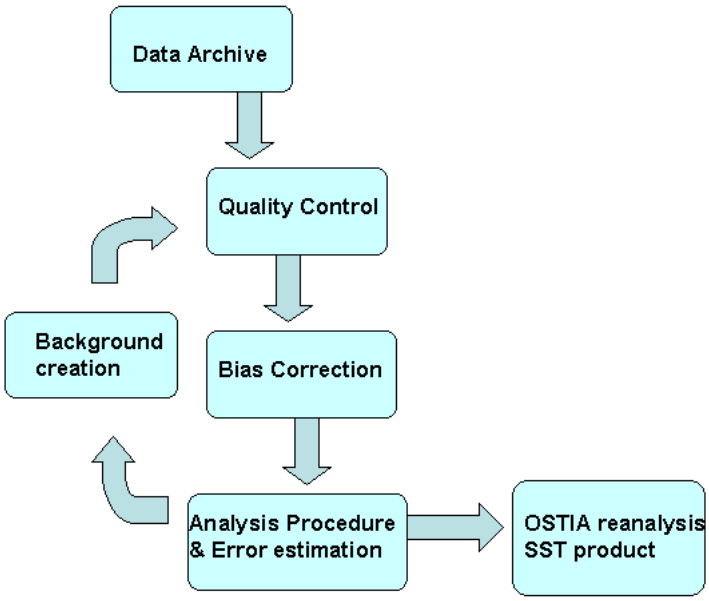


Figure 3.1: Schematic of the OSTIA reanalysis system

Figure 3.1 shows a schematic of the steps involved in the production of the daily SST fields in the OSTIA reanalysis system. Each of the steps in the system are discussed in turn within the following sections. Observations are extracted from files held in an archive and the quality control (QC) described in section 3.2 is carried out. These observations are assimilated onto a background field which is persisted from the previous day's analysis with a relaxation to climatology. Prior to assimilation selected satellite observations are bias corrected to a reference data set. An objective analysis then produces the daily SST field.

In order to run the reanalysis system efficiently it was run in a leap-frogging fashion in which observations were extracted and quality controlled in parallel with the production of a new SST field. The 3 day assimilation window, described in section 3.5, meant that observations were extracted 2 days ahead of the SST analysis day. As one of the QC checks described in section 3.2 is a background check this implies that the most recent SST field available as a background will be 3 days older than the day of the observations. The OSTIA reanalysis was run in 2 year sections to

reduce the running time. The sections were overlapped by 2 months and the homogeneity of the reanalysis was confirmed by checking the differences in SST and SST increments fields at the end of the overlap periods.

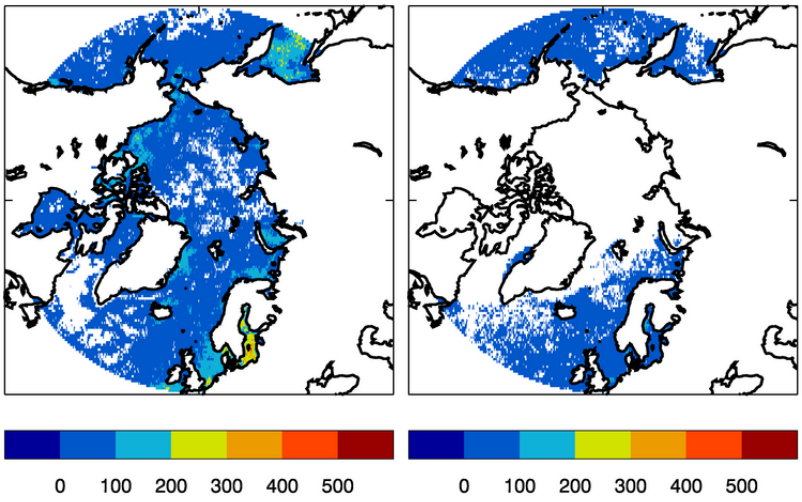
3.2 Quality control procedure

To minimise the risk of erroneous observations being assimilated in the OSTIA reanalysis the system carries out a series of quality control checks on the input obsevations. These are in addition to the quality control procedures performed by the data producers described in chapter 2. The (A)ATSR data which are observations of skin SST are adjusted to compensate for a skin temperature bias by adding 0.17 K to the SST observation, (Donlon et al., 2002).

As the OSTIA analysis is an estimate of foundation SST observations that are likely to be contaminated by diurnal warming are not used by the system. For the (A)ATSR multi-mission data observations are flagged as being at risk of diurnal warming if the wind speed is less than 6 m/s and the sun is above the horizon. In such cases it is likely that the satellite is observing a thin stratified layer of warm surface water that is not representitive of the foundation SST. The in-situ ICOADS and AVHRR Pathfinder data didn't contain the ancillary information required to perform this check. So to minimise the possibility of diurnal signal in the analysis, only night-time observations were used by the system. "Night-time" was defined in different ways by the two data producers as described in sections 2.4 and 2.2.

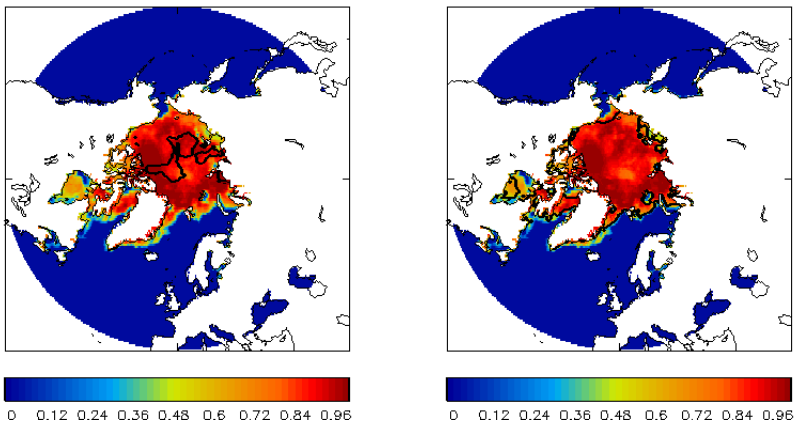
A Bayesian background check is carried out on the observations against the analysis from three days previous. This SST file was the most recent file available when the observations were extracted due to the 72 hour assimialtion window and the reanalysis being run in leap-frogging fashion. For a full description of the background check algorithm see (Ingleby and Huddleston, 2007). The scheme uses the background and observation errors described in section 3.5 and calculates the probability of gross error for each observation. If this is greater than 0.5 the observation is rejected.

Observations under ice and over land are rejected by the QC process. The sea-ice field interpolated from the OSI-SAF data has been used to mask out SST observations in grid boxes with a sea-ice concentration greater than 15%. This check was introduced to eliminate the erroneous AVHRR Pathfinder observations discussed in 2.2. In a preliminary version of the OSTIA reanlaysis AVHRR Pathfinder observations over the sea ice, as shown in figure 3.2, led to anomalously warm OSTIA SSTs in the Arctic. Figure 3.3 shows that the OSTIA SST does not drop below freezing in areas of high ice concentration. The ice masking described above greatly improves the consistency of the SSTs with the ice concentration field as shown in figure 3.3. A minimum SST check is also carried out which rejects SST observations of less than -2.0^oC.



(a) June 1989 (b) December 1989

Figure 3.2: Number of AVHRR Pathfinder observations in OSTIA for each 1/4^o grid box over 1 month, quality flags 4 to 7 (where 7 is highest quality). Spurious observations over sea ice regions are seen in all summer months of the Pathfinder dataset (e.g. (a) June image) but this is not an issue at other times of the year (e.g. (b) December image).



(a) No ice mask (b) Observations under ice masked out

Figure 3.3: Sea ice concentration and -1.7^oC SST contour for 21 June 1989, without masking of SST observations under ice (a) and with mask (b). Use of mask greatly improves consistency between SST and sea ice concentration. -1.7^oC contour used to indicate freezing temperatures rather than -1.8^oC as relaxation to -1.8^oC (section 3.3) means SST never reaches this value. -1.7^oC contour encompasses all values close to or below freezing and therefore regions which should be under sea ice.

3.3 Background creation

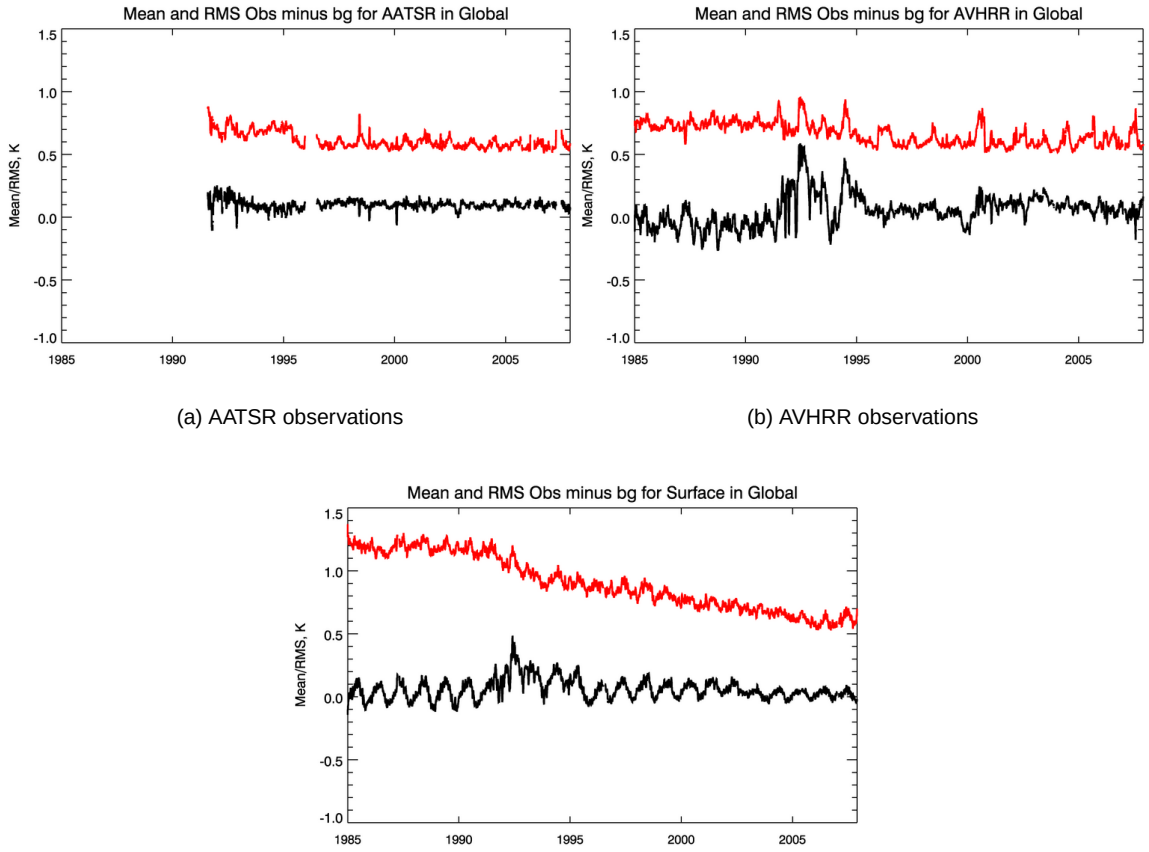
The SST analysis assimilates the satellite and in-situ observations onto a background SST field. This background is based on persisting the previous day's SST analysis with a relaxation to climatology. To ensure consistency between the SST and sea-ice concentration reanalysis fields there is a relaxation to -1.8^oC under areas of sea-ice. Each grid point of the previous SST analysis is determined to be ice-free or under ice by comparison with the sea-ice concentration field for the current analysis day. A grid point is determined to be under ice if the sea-ice concentration is greater than 50% otherwise it is labelled as ice free. Ice free grid points are relaxed towards a climatology for the same time of year with an e-folding relaxation time scale of approximately 30 days. The climatology is weekly and was constructed in house from the Daily Pathfinder Climatology (REF??). For grid points under ice the SST field is relaxed towards -1.8^oC with a relaxation time scale which decreases linearly depending on the sea-ice concentration, from approximately 17.5 days at concentration of 50% to approximately 5 days at concentration of 100%.

3.4 Bias correction

By nature satellite observations suffer from both random measurement errors and systematic biases due to such factors as poor cloud detection and aerosol contamination. These biases must be removed from the satellite observations prior to their assimilation in the OSTIA analysis as the objective analysis scheme assumes the observations used to be unbiased. To this end a bias correction is applied to the AVHRR and ATSR-1 data using the in-situ data and the ATSR-2 and AATSR data, when they become available after July 1995, as a reference data set. The inclusion of the ATSR-2 and AATSR data in the reference data set is possible due to the high accuracy of the data due to the along-track scanning dual-view approach to the SST retrivals detailed in section 2.3. (Stark et al., 2007) discuss in detail the inclusion of the (A)ATSR observations in the reference data set used in the OSTIA bias correction. The main benefit is to provide accurate SST measurements in regions of the globe poorly sampled by the in-situ observation network such as in the Southern Ocean.

Prior to the bias correction carried out by the OSTIA reanalysis system the SSES bias value, if provided by the data producers, is subtracted from the measured SST observation. Then a further bias correction is applied to all the relevant observations on each analysis day in which observations from a 3 day time window (analysis day plus and minus 1 day as described in section 3.5) are used to generate bias fields for the AVHRR and also the ATSR-1 data when applicable. Collocations in space (25km) and time (24 hrs) between the reference and the individual satellite data are found and differences are calculated and stored. These differences along with the SSES standard deviation, which supplies an error estimate for the observation and is provided by the data providers, are assimilated using a large scale version of the optimal interpolation used in the OSTIA SST analysis.

A background of the previous day's bias field, scaled by 0.9, is used together with a longer correlation length scale of 700km. This bias field is then interpolated to the satellite observation locations and the bias value is removed from the measurements prior to assimilation.



(a) AATSR observations

(b) AVHRR observations

(c) All in-situ observations

Figure 3.4: Global observation minus background RMS (red) and Mean (black) statistics for preliminary run of the OSTIA reanalysis

The decision not to use ATSR-1 data as a reference data set was taken following results from a preliminary version of the reanalysis in which all (A)ATSR missions were used as reference. Figure 3.4 shows the observation minus background statistics for (A)ATSR, AVHRR Pathfinder and in-situ observation (see section 4.1 for description of how the statistics where calculated) for the preliminary version of the reanalysis for the full reanalysis time period. It is evident that the use of the ATSR-1 data as a reference data set in the bias correction has had a detrimental impact on both the mean and the RMS of the observation minus background values for both the in-situ and AVHRR Pathfinder observations. This is clearly seen in the disruption to the seasonal cycle in both. The statistics recover towards the end of the ATSR-1 period and once the ATSR-2 data comes online in June 1995.

Figure 3.5 is the similar to figure 3.4 but for 1990-1992 only. On the left are statistics from the

preliminary version of the reanalysis and on the right are statistics from an experiment in which the ATSR-1 data were assimilated but not used as a reference data set in the bias correction and were bias corrected to the in-situ observations. Figure 3.5 shows that the detrimental impact on the seasonal cycles of the AVHRR Pathfinder and Surface observation statistics no longer occurs between August 1991 and December 1992, this informed the decision not to include the ATSR-1 data in the reference data set.

Experiments were performed to decide what quality level of ATSR-1 data to use. It is desirable to use lower quality data as the highest quality observations have poor spatial coverage during this period, particualrly in the tropics, due to increased levels of stratospheric aerosol from the Mount Pinatubo eruption (discussed in detail section 4.2). Merchant and Corlett (personal communication) also questioned the applicability of the algorithm used to deermine quality flags during this period of magnified aerosol which also suffers from a low number of drifting buoys available for assessment.

Two experimental runs of the reanalysis were carried out for August and September 1991. In one run the ATSR-1 observations with quality flag of 3,4 and 5 were assimilated, in the other only ATSR-1 observations of quality flag 5 were assimilated. In both runs the ATSR-1 observations were bias corrected to the in-situ data.

Figure 3.6 shows the monthly averaged observation minus background mean for the ATSR-1 observations for August 1991. Figure 3.6(a) is the run in which ATSR-1 observations with quality flag of 3,4 and 5 were assimilated. Figure 3.6(b) is for the run in which ATSR-1 observations of quality flag 5 only were used. Figure 3.6 illustrates the lack of data of quality level 5 in the tropics during this period and demonstrates the similar levels of observation minus background values in regions of level 3 and 4 data.

Tables 3.4 and 3.4 show the observation minus background statistics globally and regionally for the ATSR-1 and in-situ observations respectively for two runs of the reanalysis averaged over August and September 1991. In conjunction with Figure 3.6 table 3.4 illustrates how using data of quality level 3 and above increases the number of observation by more than a factor of three globally. The data void in the tropics is now filled and the number of tropical observations is increased by nearly an order of magnitude.

Table 3.4 shows that including the lower quality level 3 and 4 data in the analysis in addition to the quality level 5 data increases the ATSR-1 observation minus background global bias from -0.06K to -0.24K. This increase in bias is seen over all of the ocean regions apart from the Mediterranean. It is worth noting that the bias correction employed by the assimilation scheme should remove some of this bias. However the RMS is slightly decreased globally from 0.86K to 0.84K. This decrease in RMS is regionally varying and is seen in approximately half of the ocean regions. The in-situ observation minus background statistics (see table 3.4) show that including the lower quality ATSR-1 data has marginally decreased the in-situ observation bias globally from 0.02K to 0K. Again this decrease in bias is seen over almost all the ocean regions apart from the North Atlantic and the Mediterranean. The Surface RMS is almost identical for both runs both globally and regionally.

	Quality level 3,4 and 5		Quality level 5	
Regions	Mean (RMS)	No. of obs	Mean (RMS)	No. of obs
Global	-0.24(0.84)	1579712	-0.06(0.86)	414936
North Atlantic	-0.08(0.83)	245197	-0.05(0.93)	61849
North Pacific	-0.24(0.86)	361846	-0.04(0.91)	102795
South Pacific	-0.29(0.76)	344913	-0.04(0.70)	74871
South Atlantic	-0.24(0.77)	100920	-0.02(0.71)	20272
Mediterranean	0.08(0.78)	36333	0.14(0.77)	22344
Arctic	-0.15(1.21)	31508	-0.15(1.12)	20037
Southern Ocean	-0.34(0.88)	160925	-0.07(0.78)	36732
Indian	-0.30(0.82)	271524	-0.18(0.73)	63824
Tropics	-0.32(0.78)	847599	-0.16(0.73)	95589

Table 3.1: Average regional ATSR-1 observation minus background bias(RMS) in K and no. of observations for experimental runs using ATSR-1 quality level 3,4 and 5 and 5 data

It is worth noting that in the tropics mean bias is decreased from 0.12K to 0.07K and the RMS decreases from 0.88K to 0.87K. As the in-situ observation statistics are used to quantify bias and accuracy for validation purposes and these showed improvement in the run with the lower quality data both globally and in the Tropics where the additional ATSR-1 data was located, the decision was made to use data of quality level 3,4 and 5 in the reanalysis.

A further set of experiments was carried out to ascertain whether the ATSR-1 data (levels?) should be bias corrected to the in-situ observations or assimilated without any bias correction. Two experimantal runs were carried out for August and September 1991. Tables 3.4 and 3.4 show the observation minus background statistics globally and regionally for the ATSR-1 and in-situ observations averaged over August and September for runs with and without bias correcting the ATSR-1 data. The tables show the expected statistical outcome of the bias correction to the insitu data with the ATSR-1 observational bias increasing globally from -0.11K to -0.24K when the bias correction is carried out, whilst as expected the in-situ bias decreases from 0.05K to 0K. The global ATSR-1 observation minus background RMS increases from 0.76K to 0.84K whilst the in-situ observation minus background RMS remains unchanged when the bias correction is carried out. The same pattern is seen across almost all the regions. Taking into account the in-situ results and the desire to make the different sources of observations as consistent as possible the decision was made to

	Quality level 3,4 and 5		Quality level 5	
Regions	Mean (RMS)	No. of obs	Mean (RMS)	No. of obs
Global	0.00(1.11)	2545	0.02(1.11)	2586
North Atlantic	-0.03(1.13)	844	-0.01(1.13)	859
North Pacific	0.00(1.20)	978	0.00(1.20)	996
South Pacific	0.01(0.78)	247	0.04(0.77)	250
South Atlantic	0.02(1.29)	50	0.06(1.30)	51
Mediterranean	0.01(1.21)	36	0.00(1.22)	37
Arctic	-0.03(1.49)	27	-0.05(1.48)	28
Southern Ocean	0.07 (0.73)	198	0.09(0.73)	200
Indian	0.08(0.99)	117	0.15(1.00)	119
Tropics	0.07(0.87)	707	0.12(0.88)	717

Table 3.2: Average regional in-situ observation minus background bias(RMS) in K and no. of observations for experimental runs using ATSR-1 quality level 3,4 and 5 and 5 data

bias correct the ATSR-1 data to the in-situ observations.

The final decision was to use ATSR-1 data with quality level of 3 or higher and to bias correct it using the in-situ data as a reference.

3.5 Analysis procedure

The background field and bias corrected observations are then used to produce an SST analysis using a multi-scale optimal interolation (O.I.) type scheme. The O.I. equation is solved in the OSTIA assimilation scheme using the Analysis Correction Method (Lorenc et al., 1991). This method uses an iterative procedure which approximates the O.I. solution. The OSTIA reanalysis assimilation scheme uses 10 iterations following the same configuration as the operational OSTIA system. See (Donlon et al., 2011) for a detailed description of the assimilation scheme.

As no physical model is used in the OSTIA analysis system it is the background error covariance matrix which solely determines how the observation increments are spread onto the background field. The covariances are split into two components, one which is intended to represent the errors due to mesoscale ocean features and the other which represents larger scale errors such as those introduced by synoptic atmospheric features. These two components are then summed to give

	ATSR-1 with bias correction		ATSR-1 without bias correction	
Regions	Mean (RMS)	No. of obs	Mean (RMS)	No. of obs
Global	-0.24(0.84)	1589370	-0.11(0.76)	1608688
North Atlantic	-0.08(0.83)	244515	-0.01(0.78)	248564
North Pacific	-0.24(0.86	365001	-0.15(0.80)	372102
South Pacific	-0.29(0.76)	348601	-0.11(0.62)	349907
South Atlantic	-0.24(0.77)	100648	-0.08(0.68)	102118
Mediterranean	0.08(0.78)	36071	0.09(0.77)	36123
Arctic	-0.15(1.21)	31523	-0.04(1.17)	32327
Southern Ocean	-0.33(0.88)	162946	-0.22(0.81)	164688
Indian	-0.30(0.82)	273634	-0.13(0.72)	277170

Table 3.3: Average regional ATSR-1 observation minus background bias(RMS) in K and no. of observations for experimental runs with and without bias correcting ATSR-1 quality data to in-situ reference data

the total covariance used by the assimilation scheme. The background error covariance matrix used in the OSTIA reanalysis is the same as that used in the operational OSTIA system and has been calculated using output from a 3 year hindcast of the Met Office FOAM system (Storkey et al., 2010; Bell et al., 2000) which resulted in error correlation length scales of 10km and 100km. The effective correlation scale is the sum of both components and varies spatially so that in areas of high mesoscale variability the errors associated with that variability dominate, whilst in other regions larger spatial scale errors will dominate.

The observation error covariance matrix determines the relative weight given to each observation within the assimilation. The observation errors comprise both a measurement error and a representivity error. Within the OSTIA system an assumption is made that the high resolution grid implies that errors of representivity are small and measurement errors dominate. For both the AVHRR Pathfinder and (A)ATSR satellite data each observation comes with an estimate of the measurement error provided by the data producers. For the AVHRR Pathfinder data these errors are provided by the standard deviation of the observation values that contributed to the SST value within each particular grid box. For grid boxes with a single observation only a minimum standard deviation of 0.33 K is used (Kearns et al., 2000). For the (A)ATSR data this information was provided in the form of the SSES standard deviation (REF). These are used to form the observation error covariance matrix. For the in-situ data the observation errors vary spatially and are obtained from

	ATSR-1 with bias correction		ATSR-1 without bias correction	
Regions	Mean (RMS)	No. of obs	Mean (RMS)	No. of obs
Global	0.00 (1.11)	2549	0.05(1.11)	2546
North Atlantic	-0.03(1.13)	847	-0.01(1.12)	846
North Pacific	-0.01(1.20)	980	0.04(1.20)	981
South Pacific	0.01(0.78)	247	0.12(0.78)	246
South Atlantic	0.01(1.30)	50	0.10(1.30)	51
Mediterranean	0.02(1.21)	37	0.01(1.22)	36
Arctic	-0.02(1.49)	27	-0.05(1.47)	27
Southern Ocean	0.07(0.73)	197	0.18(0.76)	195
Indian	0.08(0.99)	118	0.21(1.00)	117

Table 3.4: Average regional in-situ observation minus background bias(RMS) in K and no. of observations for experimental runs with and without bias correcting ATSR-1 quality data to in-situ reference data

a static 2D field which is specified a priori, and is identical to that used in the operational OSTIA system. These in-situ observation errors do not vary for the different types of in-situ measurement and were obtained from the 3 year FOAM hindcast. It is assumed that after large scale satellite errors are bias corrected there are no spatial correlations in the observational errors leading to a diagonal observation error covariance matrix.

The assimilation scheme is run using a rolling observation window of 72 hours centred on 1200 UTC on the analysis day, thus giving an overlap of plus and minus 1 day. This is different from the operational OSTIA system which employs a 36 hour rolling window. The preliminary version of the OSTIA reanalysis used a 24 hour assimilation window. The lack of an overlapping assimilation window resulted in temporally noisy increments, with large positive increments often followed by large negative increments on the next day. In increasing the assimilation window the lack of time of observation information in the AVHRR Pathfinder data limited options for the length of the overlap period to be full days only.

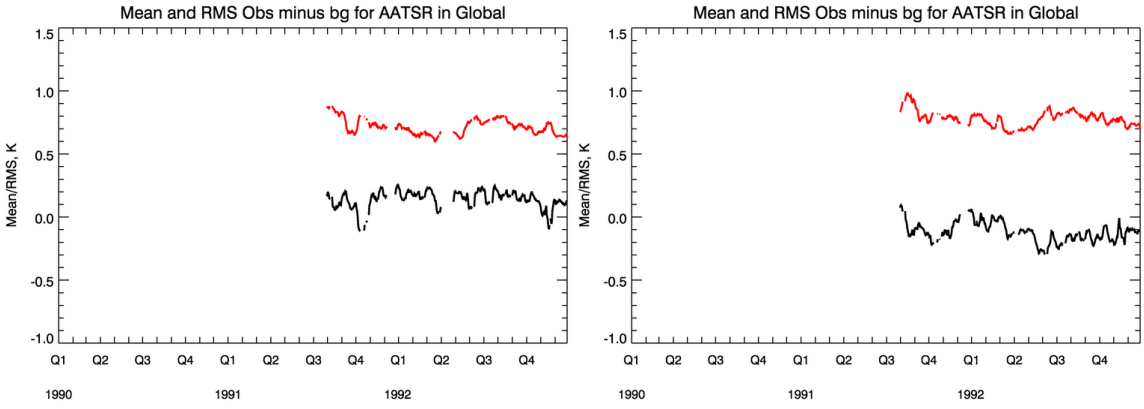
Despite the fact that a 72 hour assimilation window is used within the reanalysis scheme the analysis produced is valid for a single day only. Thus it is desireable to give observations closest to the centre of the analysis a higher weight in the assimilation scheme. This is achieved through increasing the observation error, thus giving the observation less weight in the assimilation, as the time of observation from 1200 UTC on the analysis day increases. The observation error is scaled

by a factor which increases linearly from 1 at 1200 UTC on the analysis day to 1.5 at plus/minus 36 hours. The magnitude of the scaling required was estimated by assuming that observation errors propagate with ocean features, thus the temporal correlations in the observation error can be obtained by calculating the temporal correlations in the SST's or in SST anomalies (which removes the seasonal cycle from the correlations). A linear fit to the SST and SST anomaly correlations resulted in scaling factors of 1.3 and 1.7 respectively. Analysis of the impact of the different scaling factors on the analysis increments and observation minus background fields resulted in the mid-way value of 1.5 being chosen.

3.6 Error estimation

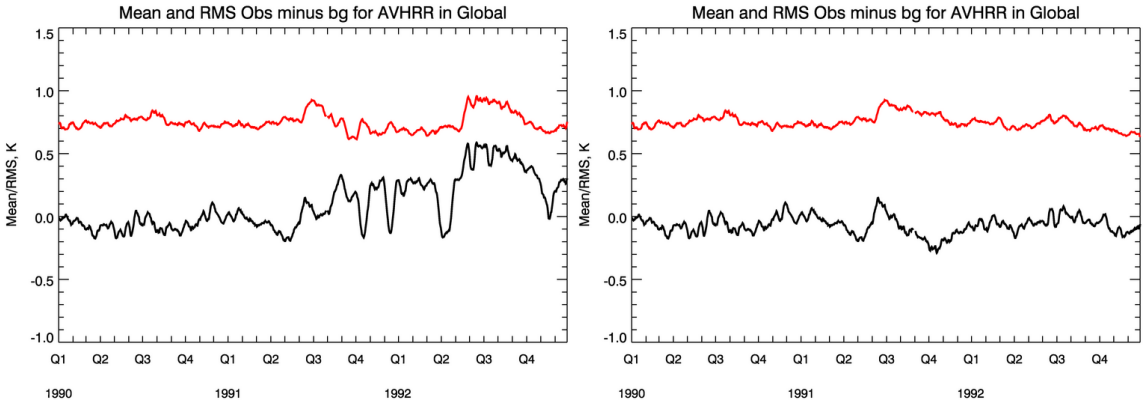
Each SST analysis is accompanied by an uncertainty estimate in the form of the estimated error standard deviation of the analysed SST. Computing the analysis error within the optimal interpolation scheme requires much more computational resources than the analysis itself so a simpler approach is used to estimate the analysis error. This is carried out in the OSTIA system using an analysis quality (AQ) optimal interpolation, as described in (Donlon et al., 2011).

The analysis error is calculated by carrying out a second optimal interpolation identical to that described in section 3.5 except all observations are given a value of 1.0, the background is set to 0 and the background and observation errors are the same as those used in the SST analysis. New observations add information to the AQ field and observation gaps result in the field decaying. The analysis error field provides information on the data coverage and observation errors that have been input into a given analysis plus information on the analysis background errors.



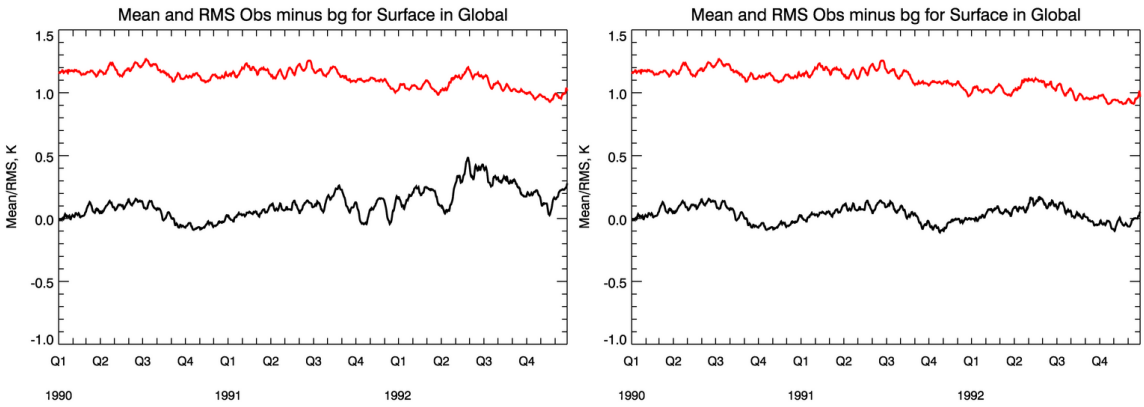
(a) ATSR-1 observations, ATSR-1 used in reference

(b) ATSR-1 observations, ATSR-1 bias corrected



(c) AVHRR observations, ATSR-1 used in reference

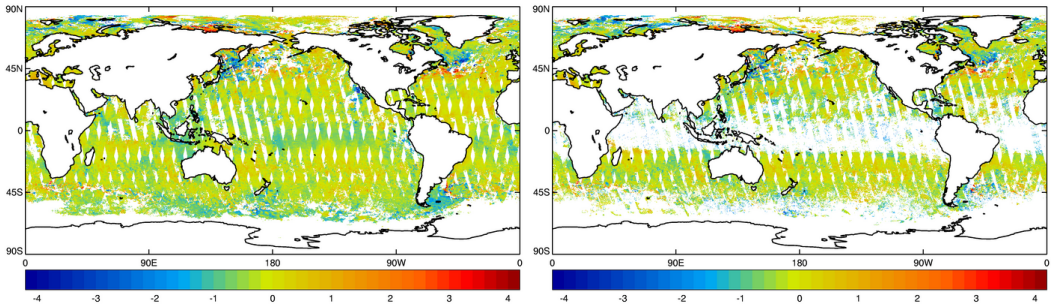
(d) AVHRR observations, ATSR-1 bias corrected



(e) In-situ observations, ATSR-1 used in reference

(f) In-situ observations, ATSR-1 bias corrected

Figure 3.5: Global observation minus background RMS (red) and Mean (black) statistics for the OSTIA reanalysis for runs in which ATSR-1 data is used in the reference data set (Left) and ATSR-1 data is bias corrected to the in-situ data (Right)



(a) ATSR-1 observations of quality level 3,4 and 5

(b) ATSR-1 observations of quality level 5

Figure 3.6: Monthly average observation minus background field for August 1991 for ATSR-1 observations of quality level 3, 4 and 5 and quality level 5 only

4. Assessment of OSTIA reanalysis v1.0

4.1 Validation statistics

To truly assess an SST reanalysis product statistical comparisons of the analysis fields with an independent reference data set are required. However in the reanalysis framework the scarcity of long-term measurement records dictates that as many observations as possible are assimilated. It is not feasible to withhold observations from the assimilation scheme for validation alone, particularly for the full reanalysis period. If independent validation data had been withheld from the analysis this would be detrimental to the analysis product. Within the OSTIA reanalysis there are no independent observations that cover the full reanalysis period, thus a comprehensive, objective assessment of the assessment skill is not possible (Gregg et al., 2009). Validation using an independent data set has been carried out using Argo data when they are available for the more recent period as detailed in section 4.3. Validation of the SST reanalysis over its full time period is required and in the absence of an independent reference data set, assimilated data shall be used to verify the OSTIA SST reanalysis.

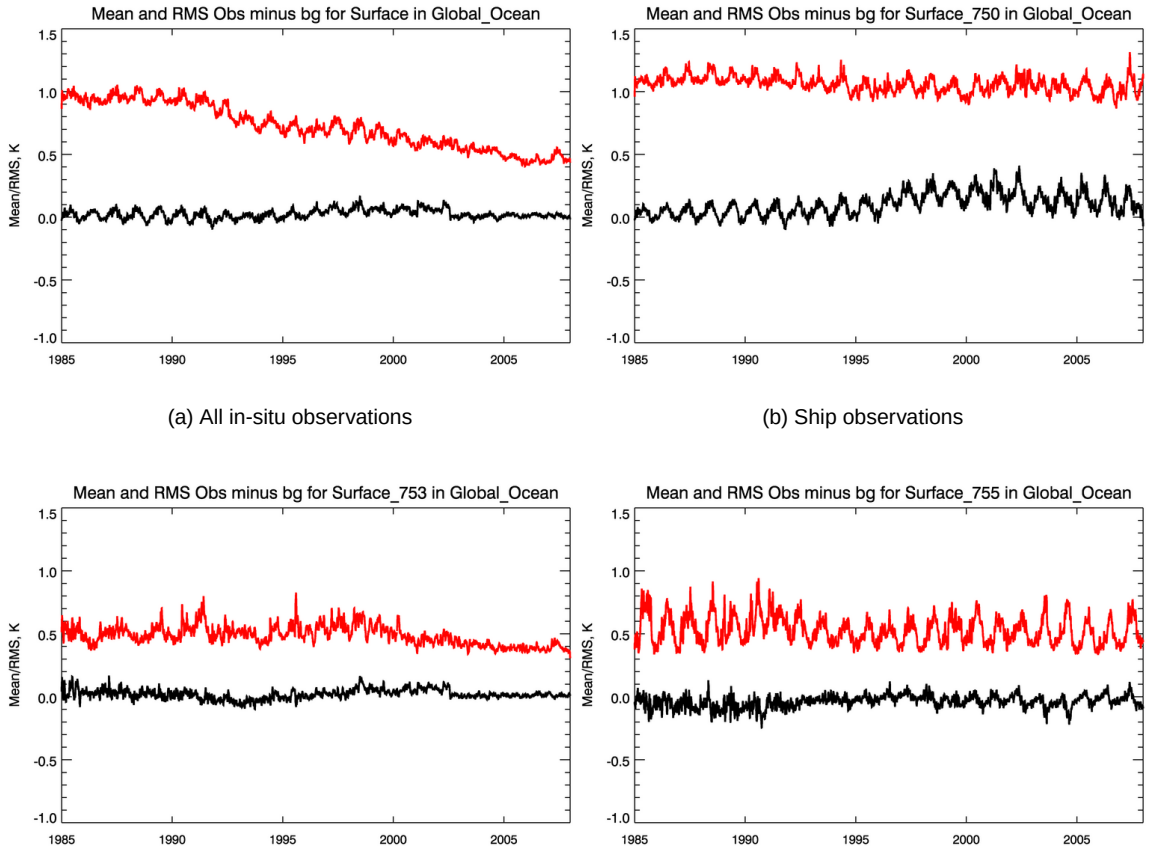
Within the reanalysis the 3 day assimilation window means that 2/3 of the SST observations are shared between successive days analyses, however the scaling of observation errors which increase with time from the analysis validity time imply that the error characteristics of the same observations will differ from one day to the next. This use of observations implies that the background field is not independent of the observations from any particular day. However validation using independent Argo data produced similar validation results to those presented here for the recent period, 4.3, which supported this approach. This along with the fact that the use of observation minus background statistics is widely accepted as a verification and validation technique by operational agencies. Bearing in mind the caveat that the background data is not truly independent and in the absence of a viable alternative for validation over the full reanalysis period, observation minus background statistics shall be used.

The background SST field used by the OSTIA reanalysis system is calculated from the previous day's analysis. The differences between each SST measurement data type and this background

field are calculated by bi-linearly interpolating the background to each of the observation locations for that day. The mean and root mean square (RMS) of these differences is then calculated within each grid box of a $1/4^{\circ}$ latitude-longitude grid. The mean value of all the differences provides an estimate of the bias and the RMS of the differences can be used to assess the accuracy.

These statistics are calculated for all grid boxes and then averaged both over the global ocean and regionally to provide statistical information on how well the observations or analysis perform spatially.

4.1.1 In-situ statistics



(a) All in-situ observations

(b) Ship observations

(c) Drifting buoy observations

(d) Moored buoy observations

Figure 4.1: Global observation minus background RMS (red) and Mean (black) statistics for in-situ observations for the OSTIA reanalysis

Figure 4.1(a) shows a time series of the daily statistics averaged over the global ocean for all in-situ observations over the full reanalysis period. Data has been filtered with a five day rolling mean to smooth out the daily variability. Figure 4.1(a) shows an increase in the accuracy of the observations throughout the reanalysis as the observation minus background RMS decreases from approximately 1K in 1985 to below 0.5K in 2007. This is an encouraging result and indicates that

by the end of the reanalysis the system has similar levels of accuracy as the NRT OSTIA system. Comparable statistics were calculated for the NRT system for the period 1st Jan 2007 to 31st Dec 2010 and resulted in an RMS of 0.57K (Donlon et al., 2011). A seasonal cycle can be seen in the RMS with an increase in Northern Hemisphere Summer. This may be due in part to the Northern Hemisphere bias of the observation network, see figures 2.3(e) and 2.3(f), which due to the QC of suspected diurnally warmed SSTs means less data are available due to shorter nights in Northern Hemisphere summer.

The full in-situ observation network comprises observations from ships, drifting buoys and moored buoys and each of the different observation types have different measurement errors which manifest themselves in the observation minus background statistics. Also the observation network itself evolves over time with the relative impact of the different observation types changing as their proportion of the full observation network changes. To determine what effect this has on the above statistics the observation minus background statistics were calculated for each of the in-situ observation types individually.

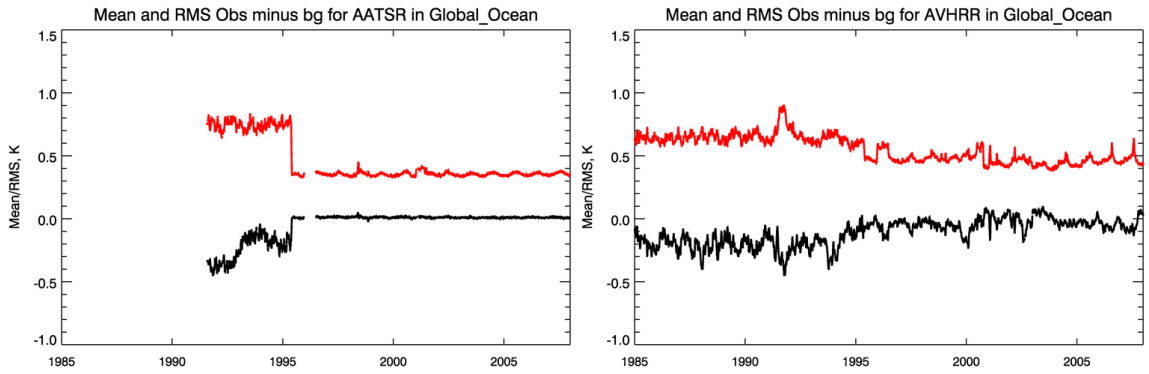
Figure 4.1(b), 4.1(c) and 4.1(d) are the same as figure 4.1(a) except that the statistics have been calculated separately for ships, drifiting buoy and moored buoy observations. They show that each of the in-situ observation types have different error characteristics. The seasonal cycle described above in the RMS can be discerned in all three observation types. In the drifting buoy statistics this decreases as the drifter network matures to provide a truly global in-situ observation network after 2002. The largest seasonal variation is apparent in the moored buoys which is comprised of both coastal and tropical moorings. Further work to explain this magnified seasonal cycle is ongoing.

It is also striking that for individual observation types the validation statistics remain relatively consistent throughout the reanalysis period with observation minus background RMS of 1.1 K, 0.5 K and 0.55 K for ship, drifting buoy and moored buoy observations respectively. One of the main factors contributing to the decrease in the full in-situ observation minus background RMS during the reanalysis period is the change from a ship dominated observation network in 1985 with high RMS values to a drifting buoy dominated observation network in 2007 with more accurate observations. Section 2.4 provides information as to the relative number of observations from each of the observation types.

The changes in the satellite observations during the reanalysis have also had an effect and can be discerned in the bias of the in-situ observations. Figures 4.1 all show that there is an increase in the observational bias after July 1995 once ATSR-2 satellite data is used to bias correct the AVHRR Pathfinder data. Prior to this all satellite data was bias corrected to the in-situ observations so this is not suprising, but does highlight possible disagreement in the ATSR-2 and in-situ observations. Figure 4.1(b) shows that the initial increase in the in-situ observation bias is most marked in the ship observations where the bias goes from approximately 0.1 K to approximately 0.25 K during the ATSR-2 period. The subsequent drop in full in-situ bias once AATSR data comes online in July 2002 highlights a possible dissparity in the ATSR-2 and AATSR observations and can be discerned in all

three observation types. The drifting buoy statistics, which are widely viewed as the least biased of the in-situ measured types, had a drop in bias from 0.06K to 0.02K.

4.1.2 (A)ATSR multi-mission and AVHRR Pathfinder statistics



(a) (A)ATSR observations

(b) AVHRR Pathfinder observations

Figure 4.2: Global observation minus background RMS (red) and Mean (black) for satellite observations for the OSTIA reanalysis

Figure 4.2(a) is similar to figure 4.1 but for the (A)ATSR observations, statistics are shown from the three missions, ATSR-1, ATSR-2 and AATSR. The impact of bias correcting to the ATSR-2 data is evident by the step change in the RMS in June 1995 when the data comes online, with the RMS decreasing from approximately 0.75 K to approximately 0.35 K. Figure 4.2(a) shows a discernable seasonal cycle in the RMS with highs during Northern Hemisphere summer, the magnitude of this cycle decreases once ATSR-2 data comes online. The noisier daily RMS and bias values for the ATSR-1 observations relative to the ATSR-2 and AATSR observations may in part be due to the ATSR-1 data volumes being less stable, with a greater proportion of days with reduced numbers or no observations, than the later missions. The peaks observed in the RMS in April 1998 and in Febuary-June 2002 correspond to periods during which there was a significant reduction in the numbers of observations. The gap in the time series during which there were no observations between 1st Jan 1996 and the 1st July 1996 is due to a known scan mirror problem with the ATSR-2 satellite. See figure 2.2 for the number of (A)ATSR observations.

The initial large bias in the ATSR-1 observations minus background difference of approximately -0.4 K is due to the fact that the satellite was launched in to the aftermath of the Mount Pinatubo eruption and retrievals were contaminated by the elevated levels of aerosol in the atmosphere. The effect of the Mount Pinatubo eruption is discussed in detail in section 4.2. As the aerosol disperses the bias decreases and averaged globally the bias stabilises to approximately -0.15 K by Jan 1994.

It is worth noting that the observation values used in these calculations are those before any bias correction is carried out and so any observation bias that we are attempting to correct for will

show in these results. The transition from an analysis procedure which bias corrects the ATSR-1 observations to the in-situ observations to a procedure which treats the ATSR-2 observations as a reference data type and thus pulls the analysis closer to ATSR-2 observations. This is evident in the observation minus background bias decreasing from approximately 0.24 K to approximately 0.00 K in June 1995 once the ATSR-2 data is used as a reference dataset in the bias correction. The observations remain unbiased for the remainder of the analysis with the bias remaining close to 0.00 K. Again occasional increases in the bias are seen due to days of low data volumes as discussed earlier.

Figure 4.2(b) shows observation minus background statistics for the AVHRR Pathfinder data for the full reanalysis time period. A seasonal cycle can be observed in the RMS with peaks occuring in the Northern Hemisphere Summer. Unlike the (A)ATSR observations a similar pattern is also seen in the mean bias with largest magnitude of bias occuring in Northern Hemisphere Summer. The impact of the Mount Pinatubo eruption can be discerned in figure 4.2(b) in both the RMS and the mean bias timeseries soon after the eruption in June 1991. The errors peak in 1992 when the levels of aerosol in the atmosphere have the most effect on the AVHRR retrievals. The differences in the impact of the Mount Pinatubo aerosol on the two types of satellite observations are discussed in section 4.2. The positive effect of being able to include the ATRS-2 data in the bias correction can be observed in the AVHRR observation minus background statistics. The RMS decreases from approximately 0.65 K to approximately 0.50 K when the ATRS-2 data is used in the bias correction. This impact is most evident in the period of ATSR-2 data outage described previously between 1st Jan 1996 and 1st Jul 1997 where the RMS almost returns to its pre-ATSR-2 values. A further reduction from approximately 0.50 K to approximately 0.45 K occurs when the AATSR data is used as a reference.

4.1.3 Regional statistics

Tables 4.1.3, 4.1.3, 4.1.3 and 4.1.3 show the average observation minus background mean and RMS differences for the MyOcean regions for all three observation types. Values have been temporaly averaged into three distinct time periods, 1985/01/01-1991/08/01 (pre-ATSR-1), 1991/08/02-1995/05/31 (ATSR-1), 1995/06/01-2002/07/21 (ATSR-2) and 2002/07/22-2007/12/31 (AATSR). Splitting up the time period into the pre-ATSR-1, ATSR-1,ATSR-2 and AATSR periods was necessary to make comparisons between the different observation types valid, given the changes in how the OSTIA reanalysis system processes the (A)ATSR multi-mission observations that has been described previously.

Table 4.1.3 shows that in the Black Sea the in-situ observations have a large bias which persists throughout all four time periods, this region has very few in-situ observations and so the resulting statistics are not robust. The in-situ and the AVHRR observations have large biases relative to the global value in Tropical regions, which is most evident in the AVHRR observations in the Tropical

Regions	In-situ	AVHRR Pathfinder	pre-ATSR-1
Global Ocean	0.013 (0.945)	-0.199 (0.648)	NA
Arctic Ocean	0.030 (0.926)	-0.204 (0.661)	NA
Baltic Sea	0.026 (0.975)	-0.127 (0.580)	NA
North West Shelf	0.003 (0.721)	-0.250 (0.475)	NA
IBI ROOS	0.035 (0.804)	-0.262 (0.520)	NA
Mediterranean Sea	0.001 (0.985)	-0.186 (0.575)	NA
Black Sea	0.109 (1.193)	-0.042 (0.736)	NA
North Atlantic	-0.003 (0.999)	-0.240 (0.614)	NA
Tropical Atlantic	0.054 (0.899)	-0.350 (0.699)	NA
South Atlantic	0.027 (0.851)	-0.174 (0.718)	NA
North Pacific	0.010 (0.970)	-0.224 (0.630)	NA
Tropical Pacific	0.064 (0.731)	-0.202 (0.612)	NA
South Pacific	0.034 (0.649)	-0.145 (0.614)	NA
Indian Ocean	0.039 (0.884)	-0.208 (0.673)	NA
Southern Ocean	0.031 (0.644)	-0.171 (0.664)	NA
Actual NW Shelf	-0.005 (0.741)	-0.211 (0.451)	NA

Table 4.1: Average regional observation minus background bias (and RMS) in pre-ATSR-1 period for all observation types

Atlantic where this bias persists throughout the reanalysis. Table 4.1.3 shows that the introduction of the ATSR-1 data resulted in the in-situ statistics remaining relatively consistent. The AVHRR Pathfinder statistics show a decreased bias in the Arctic and Baltic but this may be due to under-sampling in these regions rather than an improvement in the observations. Table 4.1.3 also shows that the ATSR-1 observations have large bias values in regions not well sampled by the in-situ observations such as the Arctic Ocean, Baltic Sea, North West Shelf, Black Sea and the Southern Ocean. The RMS is also large in these regions apart from the North West Shelf.

The same pattern cannot be observed in the regional statistics for the AVHRR observations. Table 4.1.3 shows the impact of bias correcting to the ATRS-2 data in addition to the in-situ observations. Across most of the regions the in-situ biases are worse as the analysis is now pulled to the ATSR-2 data in addition to the in-situ observations. It is expected that due to the larger number of

Regions	In-situ	AVHRR Pathfinder	ATSR-1
Global Ocean	-0.000 (0.772)	-0.199 (0.653)	-0.242 (0.747)
Arctic Ocean	0.032 (0.788)	-0.026 (0.637)	-0.547 (0.970)
Baltic Sea	0.029 (0.859)	-0.071 (0.621)	-0.529 (0.996)
North West Shelf	0.004 (0.607)	-0.180 (0.484)	-0.445 (0.763)
IBI ROOS	0.013 (0.563)	-0.203 (0.524)	-0.322 (0.666)
Mediterranean Sea	-0.020 (0.901)	-0.101 (0.571)	-0.310 (0.706)
Black Sea	0.009 (1.075)	-0.269 (0.646)	-0.484 (0.861)
North Atlantic	-0.010 (0.860)	-0.199 (0.637)	-0.210 (0.722)
Tropical Atlantic	0.031 (0.823)	-0.398 (0.746)	-0.167 (0.698)
South Atlantic	0.002 (0.772)	-0.165 (0.729)	-0.264 (0.865)
North Pacific	0.002 (0.759)	-0.230 (0.649)	-0.243 (0.722)
Tropical Pacific	0.025 (0.501)	-0.312 (0.641)	-0.203 (0.647)
South Pacific	0.019 (0.500)	-0.168 (0.599)	-0.228 (0.695)
Indian Ocean	0.002 (0.844)	-0.255 (0.709)	-0.206 (0.742)
Southern Ocean	0.028 (0.581)	-0.085 (0.612)	-0.457 (0.890)
Actual NW Shelf	-0.000 (0.657)	-0.139 (0.468)	-0.420 (0.743)

Table 4.2: Average regional observation minus background bias (and RMS) in ATSR-1 period for all observation types

ATSR-2 observations these will dominate the bias correction applied. The AVHRR Pathfinder statistics are improved both in terms of bias and RMS across most of the ocean regions. This however is not true in the Arctic Ocean which shows an increased bias and RMS, and remains problematic through the AATSR period. As expected using the ATSR-2 data as a reference has improved the statistics compared to those for ATSR-1 across all regions. Those regions poorly sampled by the in-situ observations described earlier continue to be those with the poorer statistics relative to the global average.

Table 4.1.3 shows a reduction in in-situ bias and RMS once the AATSR data are assimilated. The Tropical Atlantic and Tropical Pacific are regions which, despite showing a reduction in error, are still poor relative to the global average. Those regions that are poorly sampled by the in-situ observations, such as the Arctic, Baltic, Black Sea, North West Shelf and Southern Ocean, still show

Regions	In-situ	AVHRR Pathfinder	ATSR-2
Global Ocean	0.055 (0.650)	-0.045 (0.484)	0.012 (0.354)
Arctic Ocean	0.065 (0.875)	0.144 (0.604)	0.025 (0.483)
Baltic Sea	0.171 (1.053)	0.014 (0.519)	0.028 (0.578)
North West Shelf	0.083 (0.523)	-0.047 (0.373)	0.018 (0.379)
IBI ROOS	0.067 (0.454)	-0.111 (0.410)	0.013 (0.336)
Mediterranean Sea	0.037 (0.827)	-0.013 (0.450)	0.022 (0.394)
Black Sea	0.149 (0.571)	0.027 (0.484)	0.050 (0.517)
North Atlantic	0.020 (0.731)	-0.087 (0.507)	0.017 (0.394)
Tropical Atlantic	0.119 (0.619)	-0.202 (0.544)	0.011 (0.328)
South Atlantic	0.075 (0.595)	-0.032 (0.489)	0.009 (0.363)
North Pacific	0.064 (0.649)	-0.060 (0.497)	0.010 (0.363)
Tropical Pacific	0.108 (0.492)	-0.061 (0.487)	0.011 (0.324)
South Pacific	0.080 (0.482)	-0.007 (0.446)	0.010 (0.324)
Indian Ocean	0.076 (0.587)	-0.027 (0.486)	0.015 (0.341)
Southern Ocean	0.082 (0.514)	0.026 (0.451)	0.009 (0.373)
Actual NW Shelf	0.107 (0.594)	-0.045 (0.373)	0.020 (0.391)

Table 4.3: Average regional observation minus background bias (and RMS) in K during ATSR-2 period for all observation types

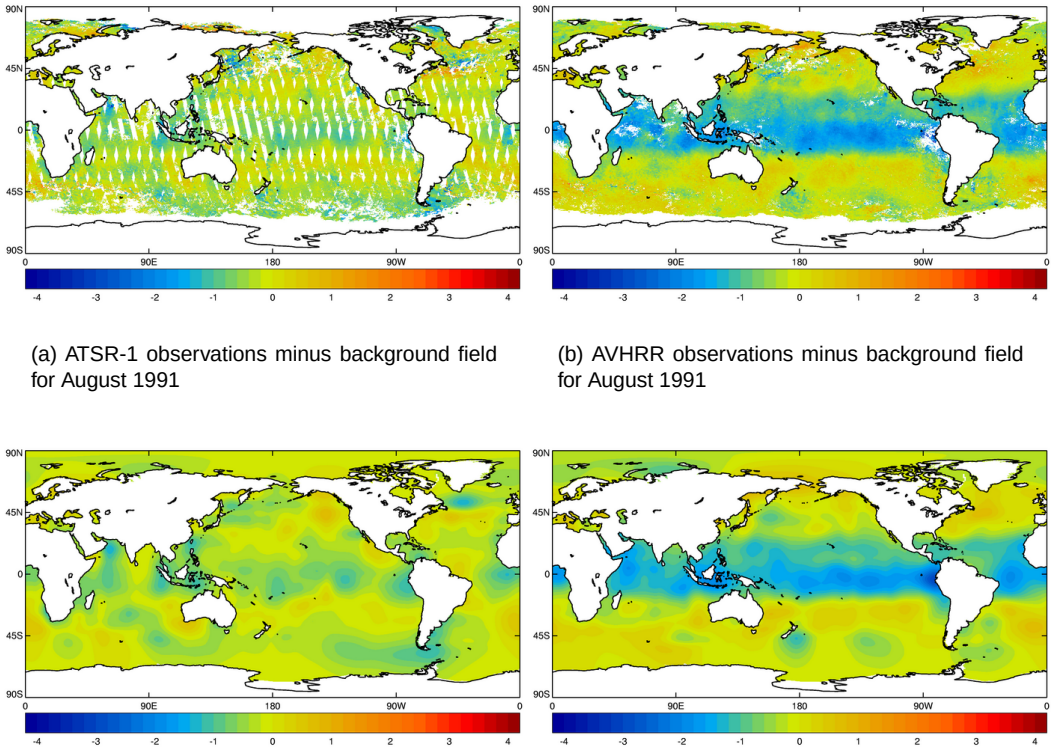
worse bias and increased RMS than the global statistics. The AATSR data has had a positive effect on the statistics for the AVHRR Pathfinder data throughout most of the ocean regions, although this is not true in regions poorly sampled by the in-situ observation network. A slight decrease is apparent in most regions in the AATSR statistics compared to those for the ATSR-2 observations.

Regions	In-situ	AVHRR Pathfinder	AATSR
Global Ocean	0.019 (0.496)	-0.023 (0.452)	0.010 (0.355)
Arctic Ocean	0.056 (0.435)	0.073 (0.606)	0.027 (0.488)
Baltic Sea	0.348 (0.958)	0.112 (0.524)	0.015 (0.579)
North West Shelf	0.080 (0.493)	0.009 (0.356)	0.014 (0.370)
IBI ROOS	0.057 (0.378)	-0.043 (0.384)	0.011 (0.345)
Mediterranean Sea	0.034 (0.733)	0.059 (0.454)	0.020 (0.401)
Black Sea	0.090 (0.615)	0.122 (0.475)	0.045 (0.505)
North Atlantic	0.003 (0.627)	-0.061 (0.490)	0.014 (0.393)
Tropical Atlantic	0.062 (0.413)	-0.212 (0.528)	0.008 (0.330)
South Atlantic	0.021 (0.428)	-0.013 (0.440)	0.007 (0.361)
North Pacific	0.013 (0.532)	-0.036 (0.472)	0.010 (0.369)
Tropical Pacific	0.045 (0.361)	-0.057 (0.464)	0.009 (0.327)
South Pacific	0.034 (0.336)	0.008 (0.401)	0.007 (0.324)
Indian Ocean	0.009 (0.453)	-0.020 (0.455)	0.013 (0.345)
Southern Ocean	0.032 (0.409)	0.075 (0.385)	0.009 (0.369)
Actual NW Shelf	0.100 (0.581)	0.013 (0.364)	0.012 (0.380)

Table 4.4: Average regional observation minus background bias (and RMS) in K during AATSR period for all observation types

4.2 Impact of the Mount Pinutubo eruption on the OSTIA re-analysis

The volcanic eruption of Mount Pinatubo in the Philippines in June 1991 resulted in a substantial quantity of aeorsol being ejected into the stratosphere. This changed the radiative properties of the atmosphere and thus impacted the retrieval of SSTs from satellite instruments. Atmospheric aerosol contamination leads to low biased satellite SST retrieval as the infrared radiation emitted from the sea surface is absorbed by the aerosol and re-emitted at the lower temperature of the aerosol (Reynolds, 1993). Of relevance to the OSTIA reanalysis is the effect of the Mount Pinatubo eruption on the accuracy of the satellite observations used and the ability of the OSTIA system to correct any biases associated with the Mount Pinatubo aerosol. During this period both the ATSR-1 and AVHRR Pathfinder data were bias corrected to the in-situ observations before being used in the analysis. The global distribution of the two satellite observations minus background fields can be used as a proxy for the biases in the satellite data assuming the in-situ data to be unbiased. The observation minus background statistics were calculated in the same way as described in section 4.1 for each grid box and monthly averages were then calculated.



(a) ATSR-1 observations minus background field for August 1991

(b) AVHRR observations minus background field for August 1991

(c) ATSR-1 bias correction for August 1991

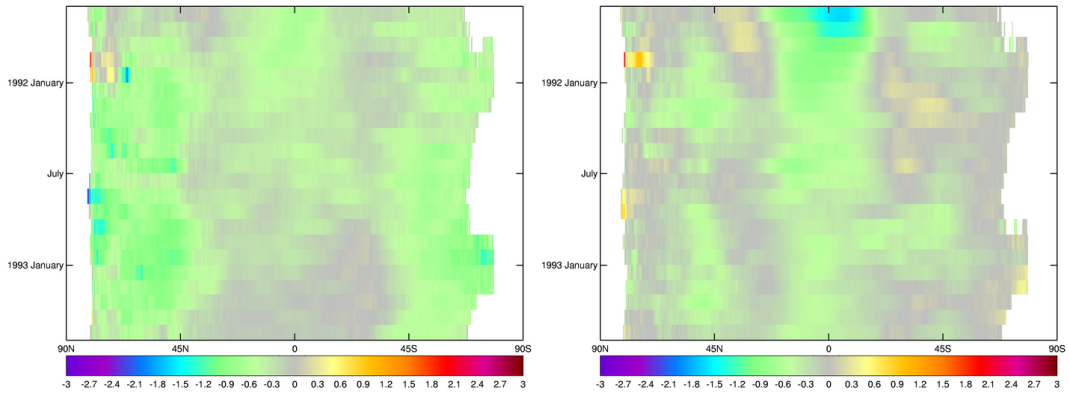
(d) AVHRR Pathfinder bias correction for August 1991

Figure 4.3: August 1991 monnthly averaged observation minus background field (top) and bias correction field (bottom) for ATSR-1 and AVHRR Pathfinder observations.

Figures 4.3(a) and 4.3(b) show the mean average monthly observation minus background plots for August 1991 for both the ATSR-1 and the AVHRR Pathfinder observations. The effect of the aerosol ejected into the atmosphere from the Mount Pinatubo eruption can be observed as it leads to a cold bias in the IR retrievals. At this time the aerosol is latitudinally confined and can be discerned as a zonal band of cold bias around the Tropics in both plots. The zonal pattern of the observation minus background biases are similar for both the ATSR-1 and AVHRR data. The fine scale structure of the bias is also similar with regions of increased negative bias, such as those off the western coast of South America and those off the western coast of North Africa, is comparable for both the sources of data. The magnitude of the bias in the AVHRR Pathfinder data is much larger than that of the ATSR-1 data. Regions of increased bias can be discerned in the AVHRR Pathfinder observations very close to the equator.

Figures 4.3(a) and 4.3(b) also illustrate that the latitudinal extent of the bias is greater for the AVHRR Pathfinder observations than the ATSR-1 observations. As the SST retrievals of the brightness temperatures of the two types of satellite observations use different methodologies it appears that the dual-view capability of for the ATSR-1 instrument is more robust to increased atmospheric aerosol than the retrieval carried out for the AVHRR Pathfinder SST observations. There are differences in the biases between the two data sources in the Western Tropical Atlantic where the ATSR-1 observations are warm biased compared to the background whilst the AVHRR Pathfinder observations are cold biased. There is also a difference in the Southern Ocean where the ATSR-1 data generally have a colder bias than the AVHRR Pathfinder data. Not all the biases observed in the figures will be due to the Mount Pinatubo eruption. This may obviously be the case in regions away from the tropically confined aerosol but there may be additional biases that are compounded by the aerosol bias in the tropical region such as those due to Saharan dust. Figure 4.3(a) also illustrates the shortage of ATSR-1 data when the satellite was initially launched as the available observations in a full month do not give full global coverage. This necessitated the decision to include data with quality level of 3 and 4 in addition to data of the highest quality level 5 of which there were very few tropical observations in this period.

Figures 4.4(a) and 4.4(b) are zonally averaged monthly observation minus background plots calculated from 1st August 1991 to 30th June 1993. Figures 4.4(a) and 4.4(b) show that the cold tropical bias due to the Mount Pinatubo aerosol persists for both the ATSR-1 and AVHRR Pathfinder observations. The magnitude of the AVHRR Pathfinder bias decreases with the largest biases close to the equator no longer discernable by October 1991. Throughout the time period shown the tropical bias is greater in the AVHRR Pathfinder data than in the ATSR-1 data. The continued dispersal of the atmospheric aerosol can be observed as the magnitude of the bias decreases with time for both ATSR-1 and the AVHRR Pathfinder observations. Figure 4.4(a) shows that by January 1993 the tropical band of bias can no longer be observed in the ATSR-1 observations, although figure 4.4(b) shows a persisting tropical bias in the AVHRR observations after January 1993 this was present prior to the Mount Pinatubo eruption. What is also more evident in figure 4.4(a) for



(a) ATSR-1 observations minus background meridional Hovmuller plot

(b) AVHRR observations minus background meridional Hovmuller plot

Figure 4.4: Meridionally averaged observation minus background Hovmuller plot for ATSR-1 observations (Left) and AVHRR Pathfinder observations (right).

the ATSR-1 data than in 4.4(b) for the AVHRR data is a greater cold bias at high latitudes which is discussed shortly.

During this time period the OSTIA reanalysis system attempts to correct the biases in both the ATSR-1 and the AVHRR Pathfinder data using the in-situ data as a reference data set as described in section 3.4. Figures 4.3(c) and 4.3(d) show the average bias correction applied to the satellite observations from both sources during the OSTIA reanalysis assimilation during August 1991. Comparison of these figures with figures 4.3(a) and 4.3(b) illustrates that the correction applied to the data resembles the spatial pattern of the bias observed and discussed previously. This is consistent throughout the period of the ATSR-1 data. The spatial coverage of the in-situ reference data set is particularly poor at high southern latituides during this time period. The bias correction methodology is dependent on a significant number of match-ups being found betwen the observations and the reference data set which cannot be achieved in the Southern Ocean during this time period, thus the validity of the bias corrections applied at this time in the Southern Ocean and in other undersampled regions may be in doubt. This problem may be exacerbated by possible cold biases in the ATSR-1 data due to cloud contamination (Merchant, 2010, pers. comm.).

4.3 Comparison with Argo data and the near-real time OSTIA system

In order to provide an independent assessment of the accuracy and bias of the reanalysis data-set in the more recent period, a comparison has been made with near-surface Argo data (which are not used in the OSTIA analysis). The Argo data have been taken from the EN3 database (Ingleby and Huddleston, 2007) at version 2a, which contains quality controlled sub-surface ocean temperature and salinity profiles. For each Argo temperature profile, the top value which has passed the quality control falling within 3-5m depth is taken. This data has been shown to be a good estimate of foundation SST by Merchant and Corlett (pers. comm., 2011) who performed a three-way comparison between Argo, surface drifter and AATSR data. The overall number of observations available within each month for the period January 2003 to December 2007 is shown in figure 4.5 together with examples of their geographical distribution at the start and end of that period. The geographical coverage of the Argo data in January 2003 is far from global, and the statistics from this early period are less robust than those from later in the time-series.

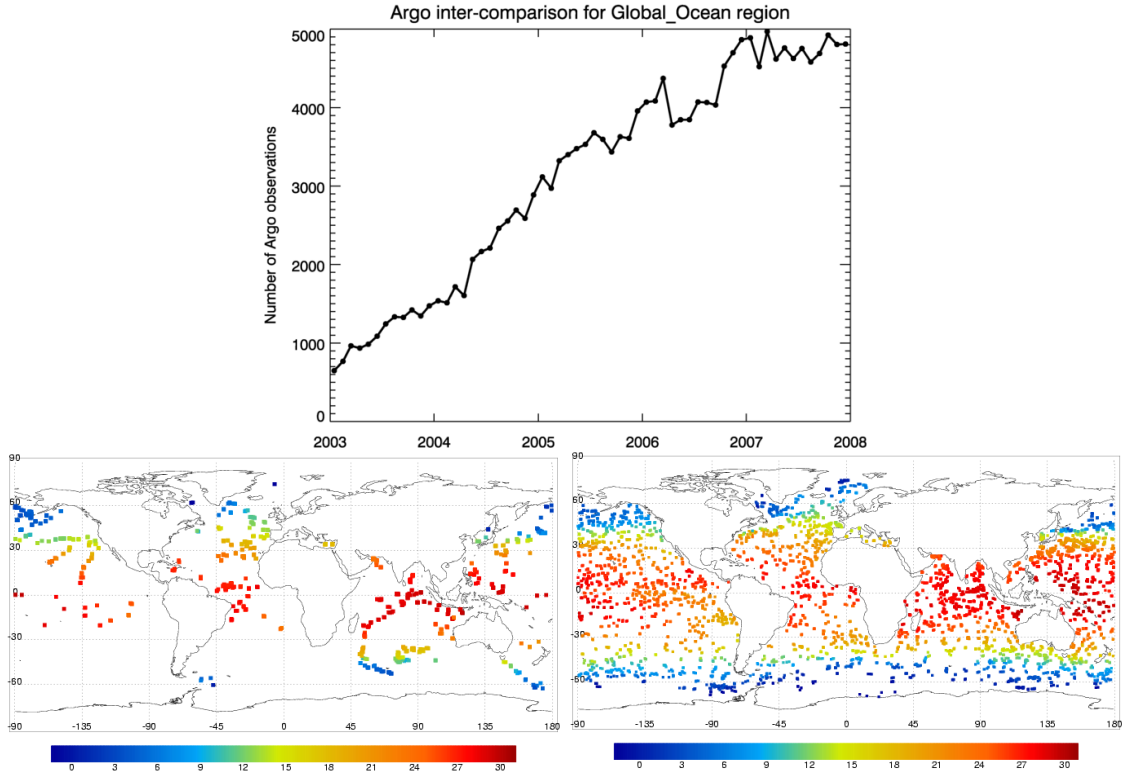


Figure 4.5: Number of near-surface Argo data (a) as a monthly time-series over the global ocean, (b) their locations for Jan 2003, and (c) their locations for Dec 2007.

For each day of the reanalysis the OSTIA data have been bilinearly interpolated to the location of the Argo observations valid on that day. Statistics of the differences between these two datasets have been calculated in various ocean regions for each month, a selection of which are shown in

figure 4.6. The global standard deviation error of the reanalysis is approximately 0.55°C with a bias of 0.1°C. The standard deviation errors are smaller in tropical regions at approximately 0.4°C, as shown in figures 4.6(c) and (d). In areas of higher SST variability such as the North Atlantic, the standard deviation error is also higher at about 0.7°C, as shown in figure 4.6(b).

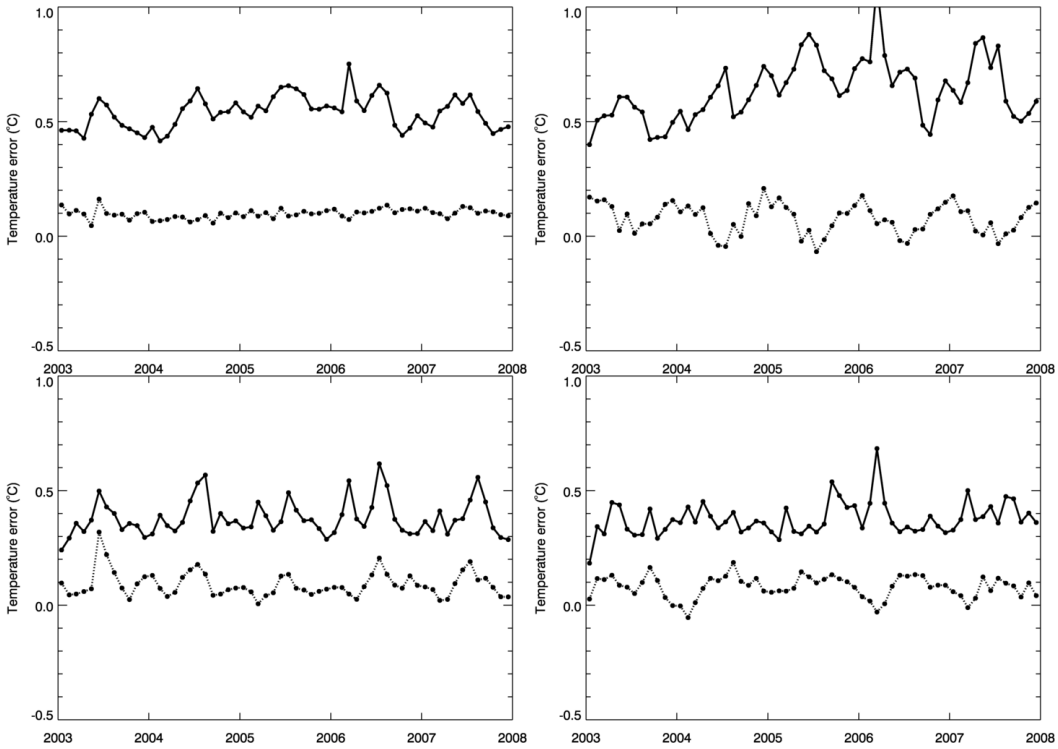


Figure 4.6: Standard deviation (solid line) and mean (dotted line) of differences between OSTIA reanalysis and near-surface Argo data (°C) from January 2003 to December 2007 for the (a) Global Ocean, (b) North Atlantic, (c) Tropical Pacific, and (d) Indian Ocean.

In all regions a cold bias of approximately 0.1°C is evident in the OSTIA reanalysis. The possibility that this could come from a diurnal effect in the Argo data has been tested by performing a similar comparison to Argo data which is valid at night-time only (not shown). The resulting biases are very similar to those shown in figure 4.6 and this cold bias is of the same order of magnitude (and sign) as the differences to the HadISST data-set shown in section 4.6.

A comparison of the errors in the OSTIA reanalysis outputs and those in the near-real time OSTIA system is presented in figure 4.7 for 2007 in the same regions as figure 4.6. Both the standard deviation and mean errors are higher in the reanalysis which implies that the extra data being used in the NRT system (including data from the SEVIRI geostationary satellite, MetOp AVHRR data, and microwave data from AMSR-E and TMI) provide a significant benefit. The decision to use a consistent set of input data (as far as possible) throughout the reanalysis and not to include other satellite data types as they become available has enabled a clean comparison with the real-time system, and enables traceability in the errors of the reanalysis. The observing system used in the reanalysis is only slowly evolving so the errors for the recent period estimated in this section (based on Argo comparisons), and their relation to the NRT system can be expected to be representative

of the whole reanalysis period with some decrease in error throughout the period as described in previous sections.

An example of the spatial distribution of differences between the OSTIA reanalysis and the NRT OSTIA SST fields is provided in figure 4.8 which shows the monthly mean difference between the two products for December 2007. This shows that the small cold bias in the reanalysis is evident in most regions of the globe with the largest differences in the North and South high latitudes around the sea-ice edge, the Sea of Okhotsk, the Eastern Indian Ocean and the Eastern Pacific Ocean. Areas of high SST variability such as the Gulf Stream, the Kuroshio and the Aghulas currents are also evident in the difference plots, although the differences in these regions are not one-signed. This implies that the coverage and resolution of the data being used in the reanalysis is not able to constrain the mesoscale signal in these highly variable regions.

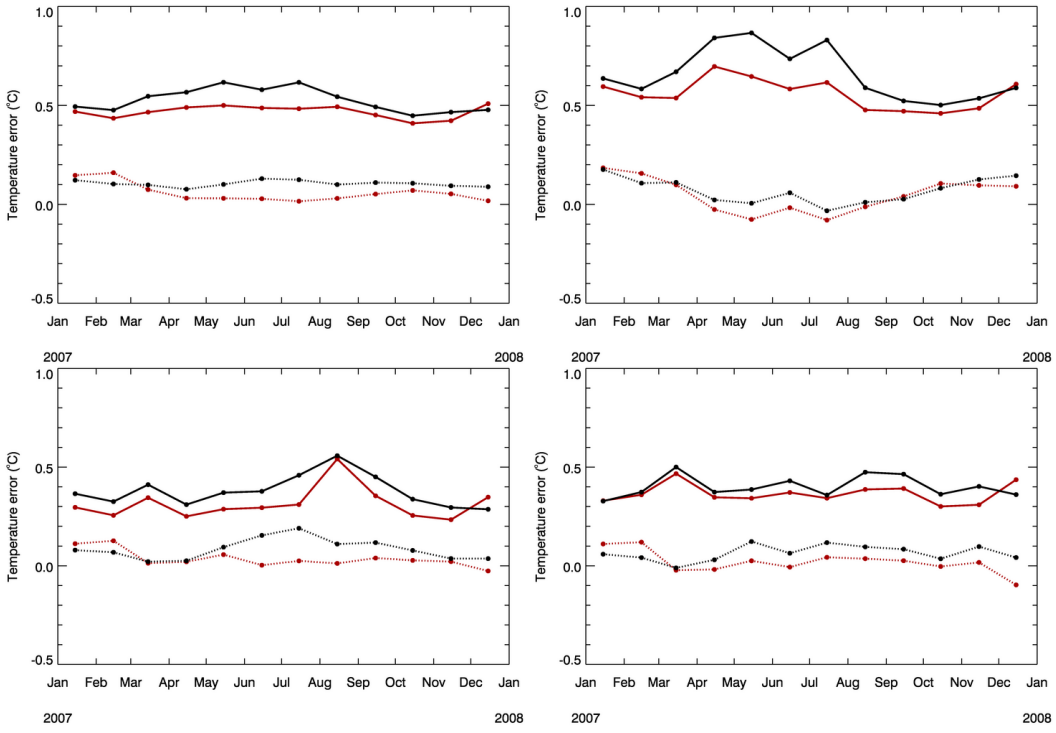


Figure 4.7: Standard deviation (solid line) and mean (dotted line) errors of OSTIA reanalysis (black) and the NRT OSTIA system (red) by comparison with near-surface Argo data (°C) from January to December 2007 for the (a) Global Ocean, (b) North Atlantic, (c) Tropical Pacific, and (d) Indian Ocean.

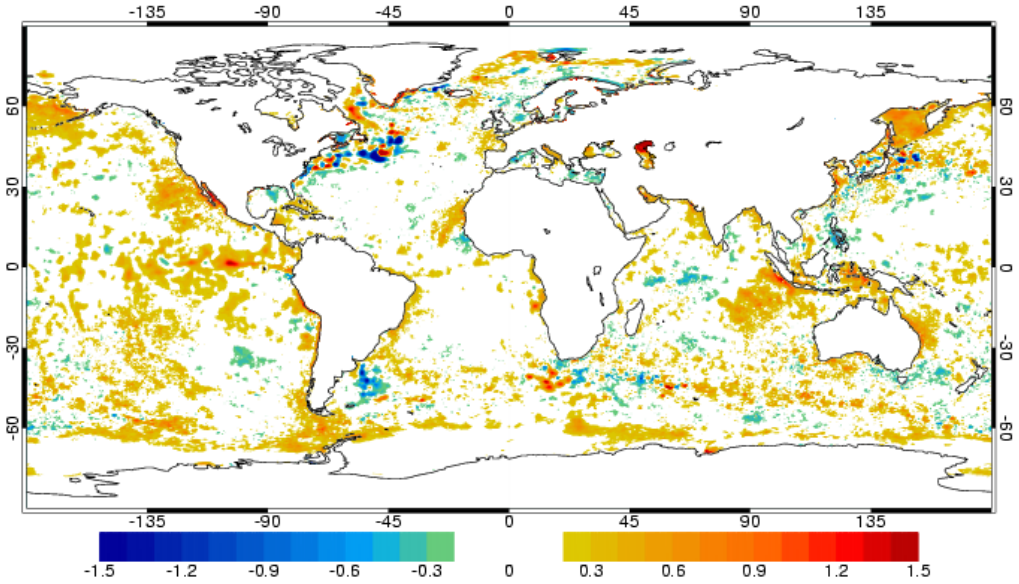


Figure 4.8: Monthly mean difference (°C) between the OSTIA reanalysis and the NRT OSTIA system for December 2007.

4.4 High latitude SST and sea ice

4.4.1 Number of observations

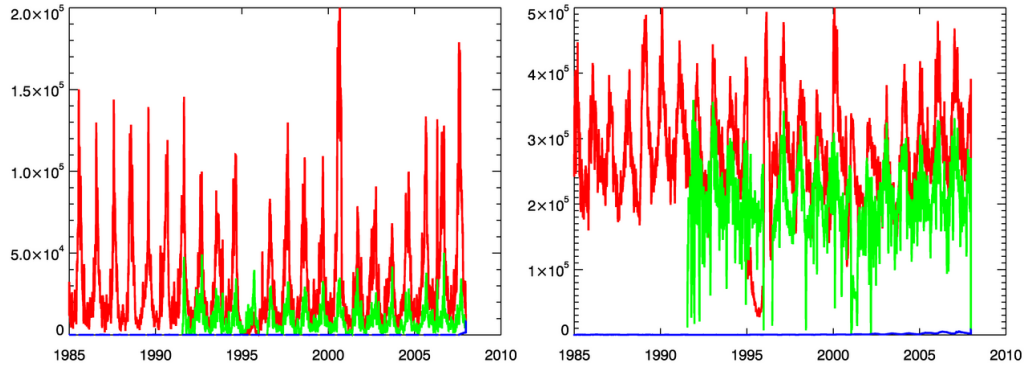


Figure 4.9: Number of observations included in the OSTIA reanalysis in the Arctic (left) and Southern Ocean (right), where AVHRR data are red, ATSR series are green and in situ are blue.

Figure 4.9 shows the number of observations in the Arctic and the Southern Ocean regions as defined by MyOcean for the OSTIA reanalysis time period. The number of observations for the Southern Ocean is greater than for the Arctic as the former region covers a larger area. Similar to the global picture (figure 2.2), the AVHRR data is the most abundant and the in-situ data the most sparse. A strong seasonal cycle in the number of observations can be observed, where there are more data in the summer, which is a result of the masking used to remove observations under sea-ice (section 2.2). When the seasonal ice extent is smallest, the greatest number of observations are available and vice versa.

Figure 4.10 shows the number of in situ observations split by platform into ship, moored and drifting buoy categories. The number of in situ observations in the Arctic decreases to zero in the summer months (unlike the satellite data), as a result of being classified as daytime observations by a solar zenith angle check, rather than the time or satellite pass (ascending or descending) used for the satellite data. This issue does not occur in the Southern Ocean region, as unlike the Arctic, the region includes data outside of the polar circle. The number of drifter observations increases throughout the period, particularly for the Southern Ocean from 1995 (figure 4.10). The impact of the 2007 IPY (International Polar Year) can be seen on figure 4.10 for the Arctic, with a dramatic increase in the number of available drifter measurements for this year.

The biggest issue facing analyses of SST in the high latitudes is the lack of data, both in situ and satellite. Many satellite retrieval algorithms also suffer from questionable accuracy at these latitudes. This means that especially in the polar summer, when there are no in situ data in the OSTIA reanalysis, the SST is not expected to be as reliable as in other regions. Further work to analyse this in detail will be undertaken.

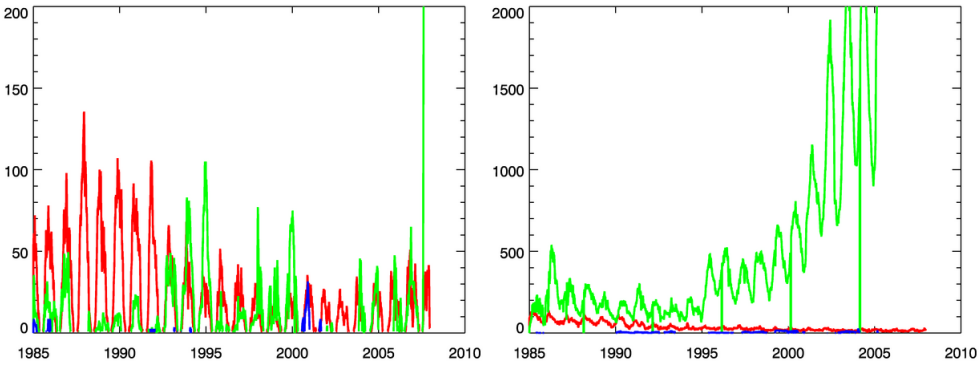


Figure 4.10: Number of in situ observations in the OSTIA reanalysis for the Arctic (left) and Southern Ocean (right) split by data type, where ships are red, moored buoys are blue and drifters are shown in green. The number of drifter observations in the Arctic in 2007 and the Southern Ocean in the later years peak at around 1500 and 5000 respectively, but the axes are truncated for clarity. Note the different scale for figures (a) and (b).

4.4.2 Sea ice and SST consistency

As described in section 3, the sea ice concentration data itself (provided by OSI-SAF, section 2.5) is used to ensure consistency between the OSTIA sea ice and SST fields. The following section provides a discussion of the SST and sea ice consistency for the OSTIA reanalysis.

In general, the match between the sea ice concentration and the OSTIA SST at the freezing point is good (figure 4.11). OSTIA is able to capture well the southern hemisphere breakup of winter sea ice from within the ice pack (figure 4.11). The -1.7^oC contour is shown on these figures rather than the -1.8^oC freezing contour as the relaxation to climatology (section 3.3) draws asymptotically towards -1.8^oC and hence the SST never reaches this value. The -1.7^oC contour includes all locations where the SST is close to or below freezing and therefore under ice.

Figure 4.12 is a timeseries of OSTIA sea ice extent (using 15% ice concentration) and freezing SST extent (using -1.7^oC) for the northern and southern hemispheres, indicating the interannual variability is well captured. The magnitude of the SST and sea ice extents match reasonably well for the southern hemisphere, for example the melt season of summer 1995 as shown in figure 4.11. For the northern hemisphere the consistency between sea ice and SST is poorer than for the southern hemisphere as the ice extent is consistently larger than the freezing extent. Additionally, the rate of change of SST and sea ice extents are more similar in the freezing season than in the melt season. Possible reasons for this poorer matching could be a bias in the ice concentration field or in the SSTs or the lack of data at high latitudes, particularly in the summer months in the Arctic when there is no in situ data. It could also be an issue with the correlation length scales used in the OSTIA objective analysis, which could cause information from SST data to propagate too far into the ice-covered region. In addition, this problem could indicate the need to adjust the relaxation timescales used in the generation of the background field, particularly those for ice concentrations of less than 50%, to

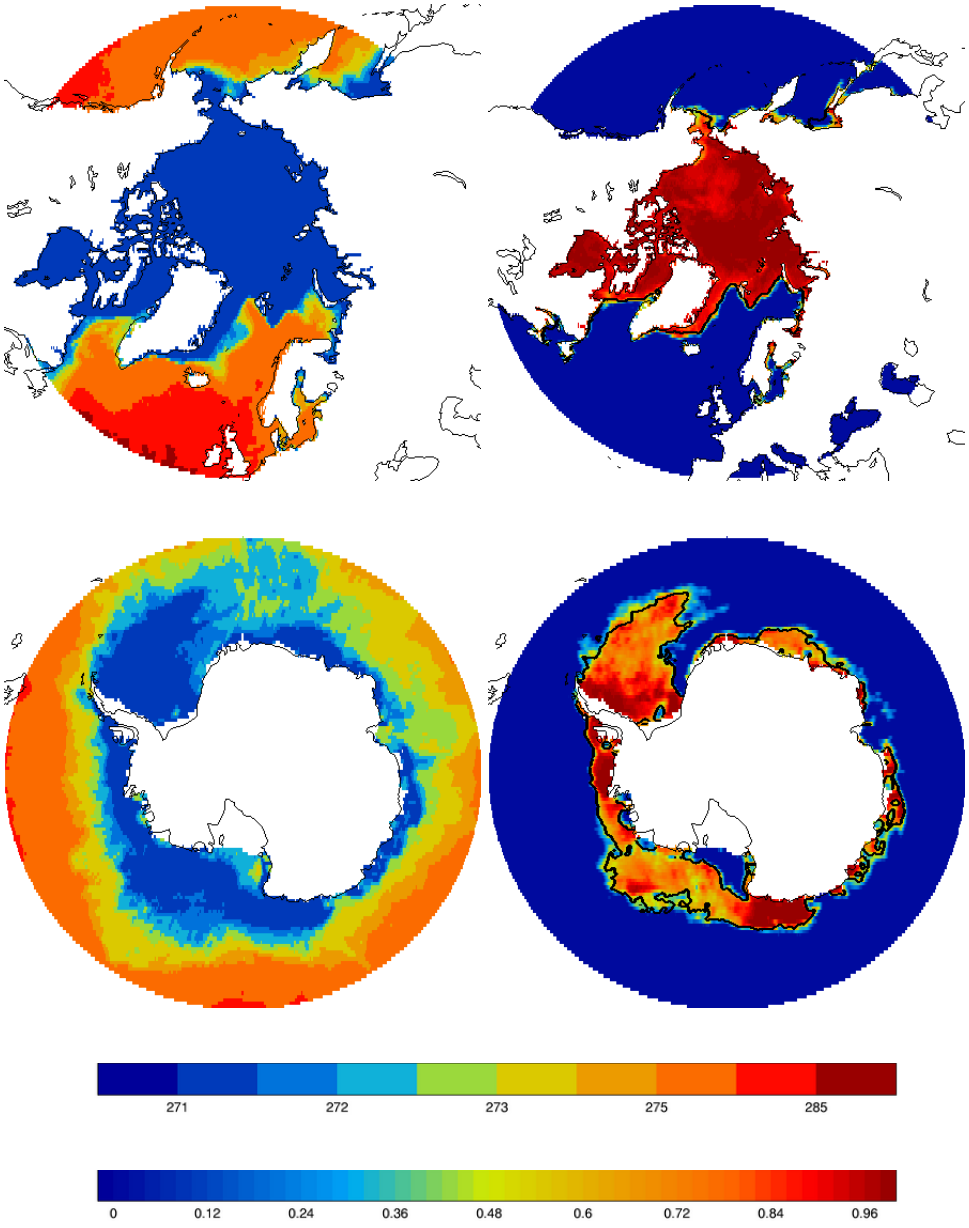


Figure 4.11: SST (left) and sea ice concentration showing -1.7^oC SST contour (right) for 31 December 1995. Northern hemisphere (top) and southern hemisphere (bottom), indicating the SST and sea ice are consistent for this time.

bring them closer in line with those greater than 50%, where a linear relationship to ice concentration is included (section 3.3). It could be possible that biases in the northern hemisphere data which are more significant than in the southern hemisphere are propagating under the ice. However, SST information spreading under the ice can be advantageous, as shown in figure 4.12 where the freezing SST extent was able to continue to grow/decline in the absence of variations in the sea ice data due to missing data. It is not appropriate to vary length scales in the different hemispheres but improvement of these should have a beneficial effect even in the southern hemisphere, as the match between SST and sea ice extent there is not perfect, particularly when ice extent is at its smallest.

In addition, it is uncertain how closely the SST and sea ice extents would be expected to match as the relationship between ice concentration and SST is complex, particularly at low ice concentrations. As the SST and sea ice concentration are two distinct datasets they would not necessarily be expected to match precisely, but the methods used here aim to improve the consistency between the two. Nevertheless, the relationship between ice concentration and SST in both hemispheres would be expected to behave in a broadly similar manner as the methods used to ensure the consistency do not differ between hemispheres. However, as there is a significant difference between the two hemispheres, figure 4.12 indicates this was not the case. It is possible that there is a bias in the northern hemisphere ice concentration field rather than the SSTs (or perhaps a combination of both).

Figure 4.13 provides a comparison of OSTIA and HadISST (Rayner et al., 2003) sea ice extents calculated from the ice concentration field for regions greater than 15% concentration. The OSTIA ice concentration is essentially regridded OSI-SAF sea ice data (section 2.5). The HadISST ice field is comprised of data from a number of different sources including ice chart information and several passive microwave datasets employing different retrieval algorithms. In the northern hemisphere, ice extents are smaller for HadISST than for OSTIA (figure 4.13). In addition, the rate of change in the melt season is faster for HadISST than for OSTIA. In the southern hemisphere, HadISST and OSTIA ice extents are better matched (figure 4.13). This could indicate there is an issue with the northern hemisphere ice extent in OSTIA. Possible sources of error in the OSI-SAF ice concentration data used for the OSTIA reanalysis are discussed in Eastwood et al. (2010). The record September minimum ice extents (where 2007 is the minimum, followed by 2005 (NSIDC)) are well captured by the OSTIA reanalysis ice extents (figure 4.13), despite any possible biases in the ice concentration field.

An issue with the consistency between SST and sea ice in the Arctic appears in the later years of the reanalysis, around 2002 onwards. Figure 4.14 is an example plot of the sea ice concentration and freezing SST contour for 6 July 2007 and illustrates that the SSTs are too warm around the sea ice edge and are thus not consistent with the sea ice extent. This plot is representative of the SST and sea ice relationship in the OSTIA reanalysis for this month. Figure 4.12 shows there can be a difference between the SST and sea ice extents in general but this difference does not usually

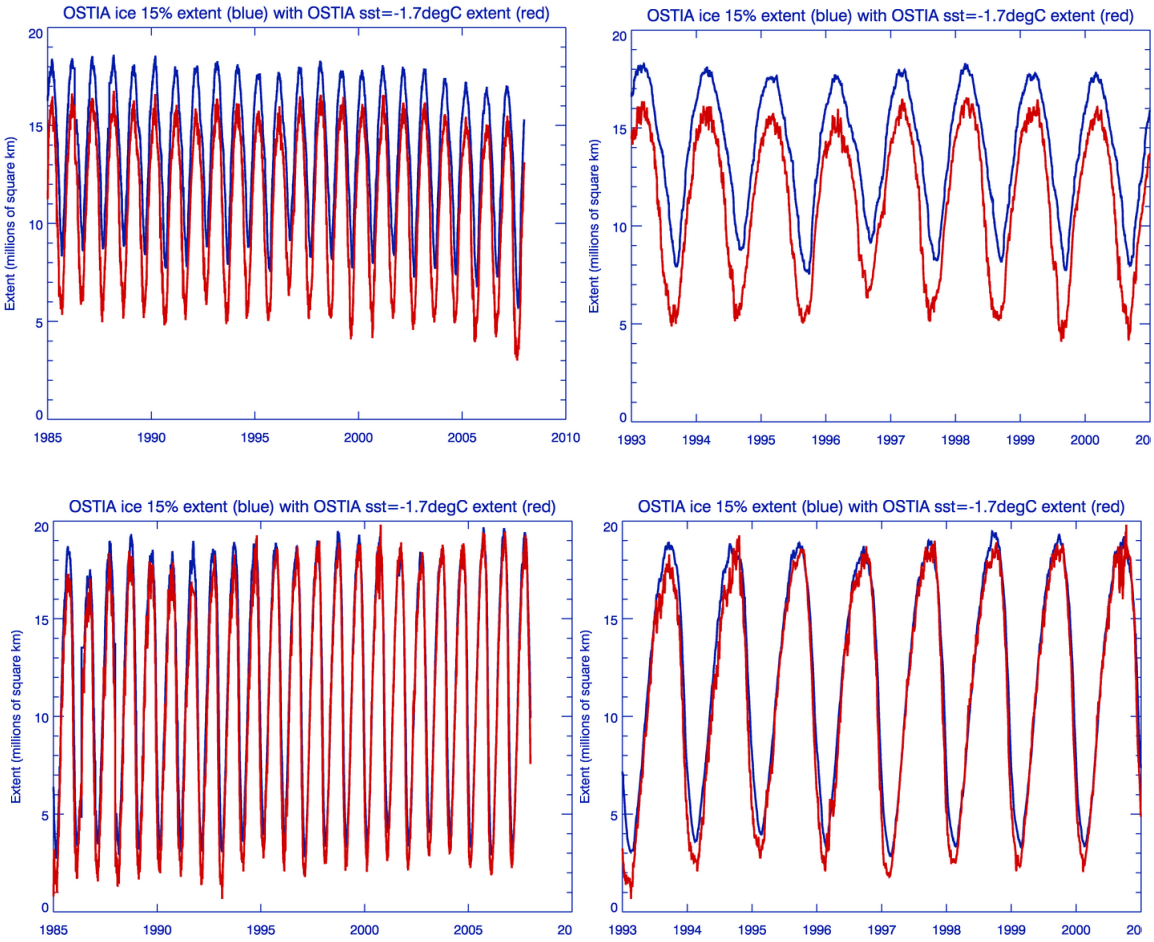


Figure 4.12: Timeseries of ice extent (using 15% concentration) with -1.7^oC SST extent for northern hemisphere (top) and southern hemisphere (bottom). For entire reanalysis (left) and zoom-in of 1993-2000 (right) to show detail.

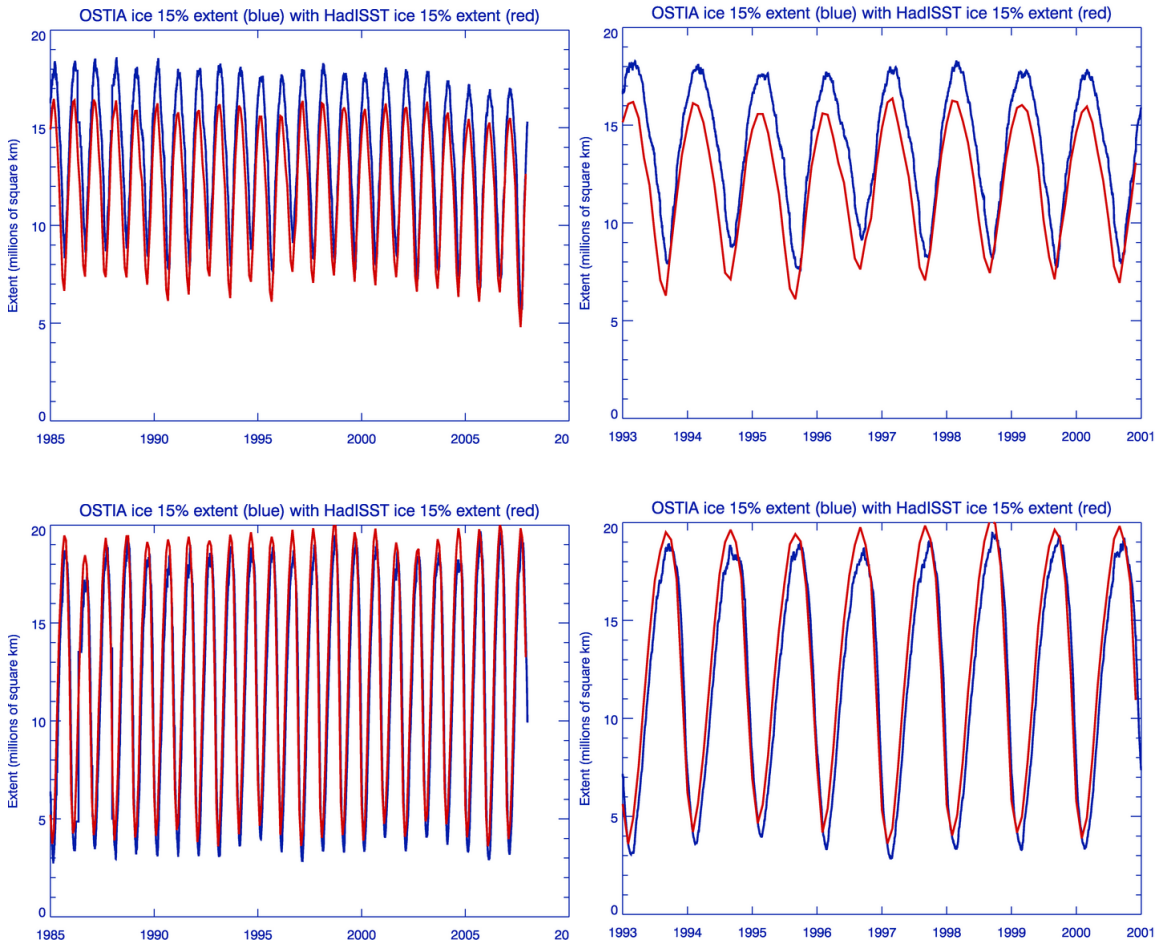


Figure 4.13: OSTIA reanalysis and HadISST ice extents (using 15% concentration) for northern hemisphere (top) and southern hemisphere (bottom). For entire reanalysis (left) and zoom-in of 1993-2000 (right) to show detail.

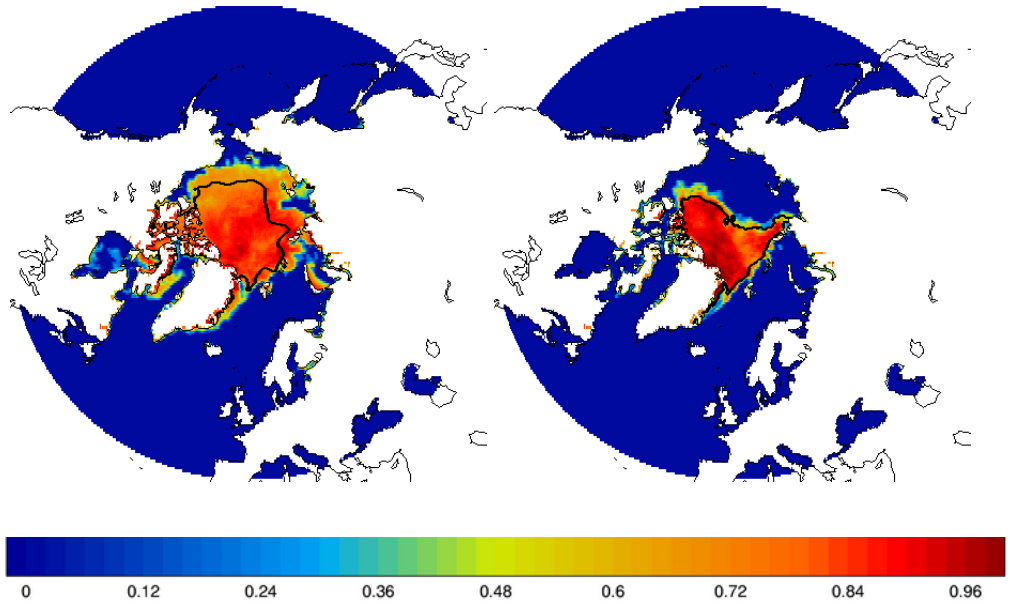


Figure 4.14: Sea ice concentration and -1.7^oC contour for 6 July 2007 (left) and 6 September 2007 (right). The freezing SST contour and sea ice concentration are inconsistent in July but have recovered by September.

manifest itself as such a large and obvious inconsistency, but rather as more subtle differences in extent closer to the ice edge.

During the months when an issue with consistency between the SST and sea ice extents is seen there are no in situ data in the analysis owing to the 24 hour daylight (see section 2.4). Therefore, AATSR only is used as a reference during this time, despite any potential high-latitude biases. Figure 4.15 shows mean and RMS observation minus background statistics for 2007 for AATSR and AVHRR in the Arctic. The errors increase during the summer months when there are no in situ data in the analysis. It is not possible to conclude from figure 4.15 whether it is the AATSR or AVHRR which has a warm bias but future work will include using the in situ data which has not been used in summer as a reference in order to determine this. The indication of biases in the data illustrates that it would have been useful to use the in situ data in the reanalysis even in the summer months, with the constraint that it should be used over the same time window as the satellite data which should minimise the diurnal warming signal.

The summer inconsistencies are not seen during the period when the bias corrections were performed using ATSR-2 as a reference (1995-2002). However, the AATSR period (2002-2007) also coincides with a period of record low ice extents (figure 4.12). Therefore, it is possible that during this period, the reduced ice masking means a greater number of perhaps unreliable high latitude observations were included in the reanalysis. Therefore the effect on the OSTIA reanalysis of any biases in the satellite data during the summer months with no in situ data will be magnified during this period. It should be noted that after the return of the in situ data for the bias correction

procedure the record minimum ice extent in September 2007 (NSIDC) was well captured in terms of the OSTIA SSTs (figure 4.14).

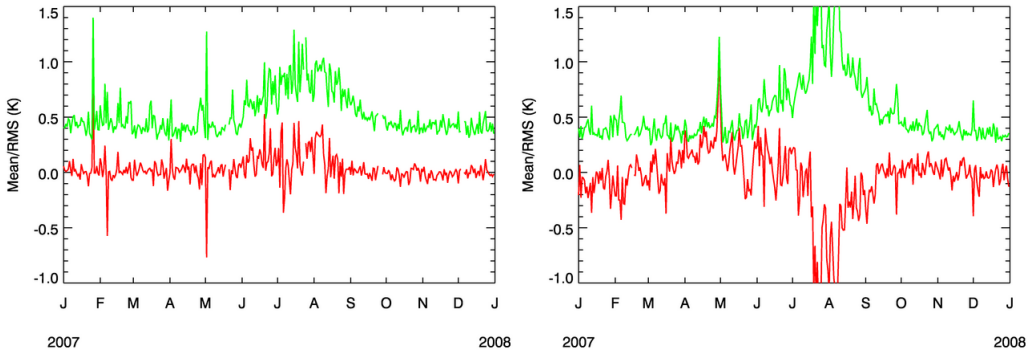


Figure 4.15: RMS and mean observation minus background errors for AATSR (left) and AVHRR (right) for the Arctic.

4.4.3 Effect of sea ice data outages

As mentioned in section 2.5, Appendix A provides a list of data outages in the OSI-SAF sea ice concentration field used to produce the OSTIA ice concentration field. These outages are caused either by a lack of data owing to satellite malfunction or planned maintenance (Tonboe and Nielsen, 2010), or because the data did not pass our QC checks. In the case of an outage, the previous day's file was persisted, sometimes for a number of days. For gaps of a week or more, the first file available after the outage was copied to the date in the middle of the outage. This is because after the halfway point the ice concentration field will be closer to that in the later file, and less close to that persisted from the file at the start of the outage. Persistence of the sea ice field owing to a lack of data can be seen in the sea ice extent timeseries, for example 1986 and 1987 for the southern hemisphere (figure 4.16). SST information from the satellite data is spread under the ice so that the SST field continues to grow/decline even without sea ice data to constrain it.

A closer look at regional issues when ice is persisted reveals that after the ice has been restored, if there is a sudden increase in ice extent, the SSTs are slower to respond to the sea ice than for a sudden decrease in the ice extent. Further investigation of this effect will be carried out and future work will be undertaken to improve the persistence-based approach during long gaps in the sea ice concentration information, perhaps through an interpolation method.

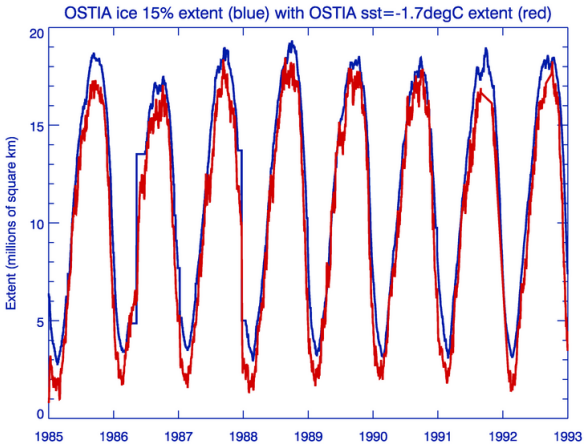
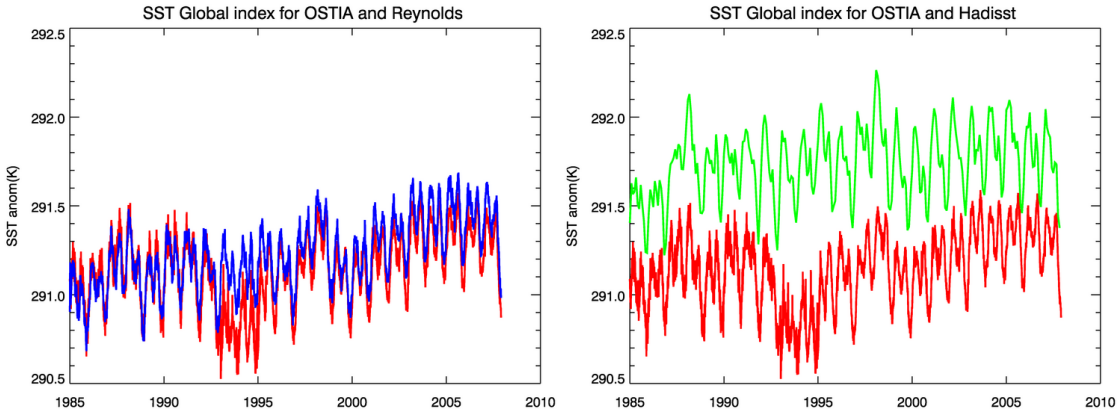


Figure 4.16: Timeseries of southern hemisphere ice extent (using 15% concentration) with -1.7^oC SST extent for 1985-1993. When the sea ice concentration field is persisted owing to lack of data (e.g. 1986), SSTs still continue to grow/decline as information from satellite data is spread under the sea ice region.

4.5 Representation of climatic indices in the OSTIA reanalysis

The temporal evolution of the OSTIA reanalysis SST fields has been validated via the calculation of daily climate indices which have been compared with those calculated from other SST analyses. Here the OSTIA reanalysis has been compared to the HadISST reanalysis and the Reynolds SST reanalysis. The HadISST v1 reanalysis is produced by the UK Met Office Hadley Centre and is a global, monthly product produced on a 1 degree grid from 1871-2007. Data sources used by the analysis are the ICOADS in-situ observations and the AVHRR Pathfinder data. The analysis reconstructs SST using a two-stage reduced space O.I. procedure followed by a superposition of quality-improved gridded observations onto the reconstructions to restore local detail. For a full description see (Rayner et al., 2003). The Daily Reynolds O.I. v2 reanalysis is a global, daily product produced on a 1/4 degree grid 1985-2008. It uses ICOADS in-situ data and the AVHRR Pathfinder and AMSRE microwave data when it is available from 2002 onwards. For a full description see (Reynolds et al., 2007). The SST index is calculated as the daily area averaged SST within a particular pre-defined region. As HadISST is a monthly product the monthly SST index values have been linearly interpolated to produce daily values.

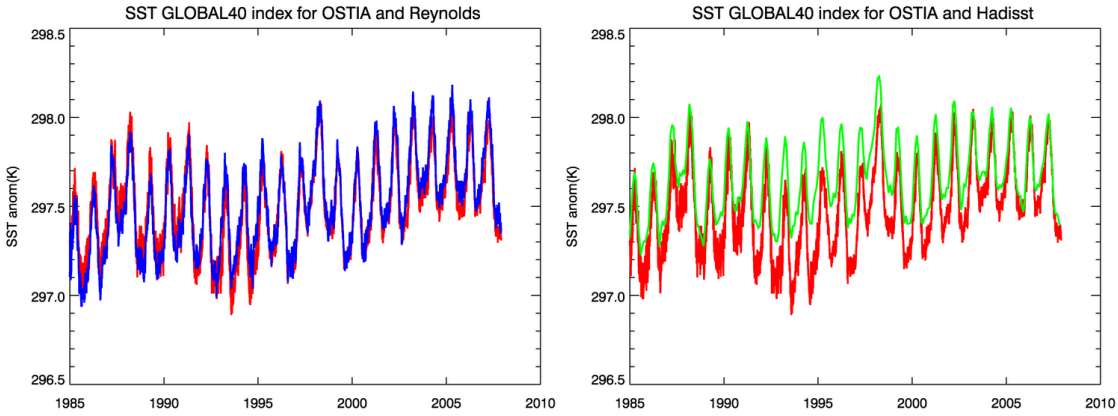


(a) Global climate index for OSTIA (Red) and Reynolds (Blue)

(b) Global climate index for OSTIA (Red) and HadISST (Green)

Figure 4.17: Global climate index for a) OSTIA (Red) and Reynolds (Blue) and b) OSTIA (Red) and HadISST (Green)

Figure 4.17(a) shows the global average SST throughout the OSTIA reanalysis time-period for the OSTIA reanalysis and the Reynolds O.I. SST reanalysis. For the majority of the time period the two analyses are generally consistent with one another both in terms of long term temporal patterns and in terms of inter-annual variability. However the two analyses show a marked divergence in terms of global average SST once the ATSR-1 data comes online in August 1991 which is assimilated by the OSTIA reanalysis but not by Reynolds. This data has led to the OSTIA reanalysis having a colder global average temperature than Reynolds during this period. Prior to this the



(a) Global climate index for OSTIA (Red) and Reynolds (Blue)

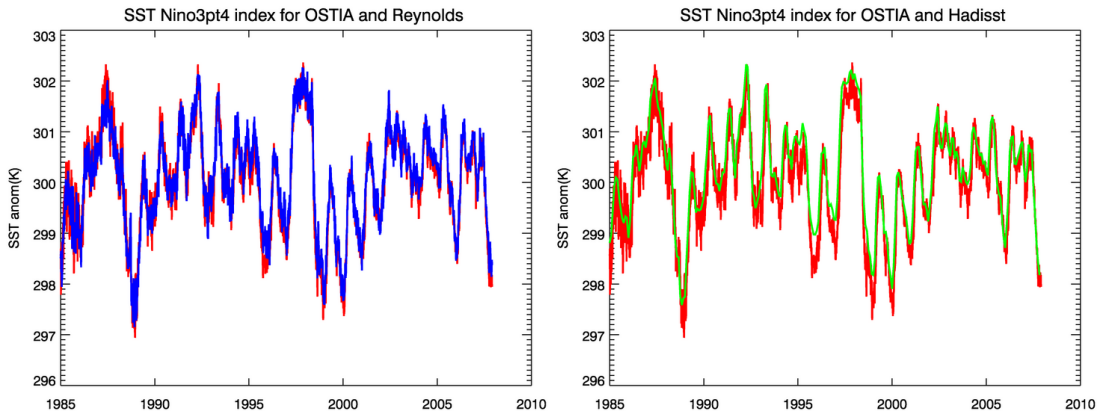
(b) Global climate index for OSTIA (Red) and HadISST (Green)

Figure 4.18: Climate index of zonal band between 40 degrees North and South for a) OSTIA (Red) and Reynolds (Blue) and b) OSTIA (Red) and HadISST (Green)

OSTIA reanalysis was warmer than Reynolds, particularly in Northern Hemisphere summer with good agreement observed in Northern Hemisphere winter. Once the ATRS-2 data is assimilated in July 1995 this brings the two analyses into closer agreement, although OSTIA is still slightly colder. This persists when the AATSR mission replaces the ATSR-2 mission in July 2002 and throughout the remainder of the reanalysis period. Analysis of the differences in the Northern and Southern Hemisphere separately (not shown) indicate that the dip in temperatures associated with the ATSR-1 data being asssimilated is more marked in the Southern than the Northern Hemisphere. A seasonal pattern can be discerned in the hemispheric differences between the two analyses. In the Southern Hemisphere OSTIA is warmer in the pre-ATRS-1 period during the summer and the (A)ATSR data has the effect of bringing the two analyses into closer agreement. The same changes can be observed in the South Pacific and Atlantic regions. In the Northern Hemisphere OSTIA is colder in the winter which persists throughout the analysis, and is observed in both the North Pacific and Atlantic regions. The global differences observed above generally occur at mid to high latitudes in excess of 40° North and South. Figure 4.18(a) is similar to figure 4.17(a) but shows the average SST's for the zonal band between 40° North and South. The differences described above due to the ATSR-1 data are no longer apparent.

Figure 4.17(b) is the same as Figure 4.17(a) but showing the OSTIA reanalysis and the HadISST reanalysis. The global average SST in the two analyses is very different with OSTIA being colder throughout the time period by approximately 0.5 K. The cooling effect of the ATSR-1 data on the OSTIA reanalysis is evident in Figure 4.17(b) as it increases the divergence between the two analyses during this time period. A seasonal pattern can be discerned in the differences in the Northern and Southern Hemispheres individually. Although OSTIA is consistently colder throughout the year the divergence is greatest during the winter in both Hemispheres. Apart from the cold bias there is

a degree of consistency between the two analyses in terms of the long term temporal patterns and the inter-annual variability. The largest differences again exist at mid to high latitudes in excess of 40° North and South. This is shown in figure 4.18(b) which shows closer agreement between the two analyses for SST index calculated over the zonal band between 40° North and South.



(a) Nino 3.4 climate index for OSTIA (Red) and Reynolds (Blue)

(b) Nino 3.4 climate index for OSTIA (Red) and HadISST (Green)

Figure 4.19: Nino 3.4 climate index for a)OSTIA (Red) and Reynolds (Blue) and b) OSTIA (Red) and HadISST (Green)

Figure 4.19(a) plots the average SST in the Nino 3.4 region throughout the reanalysis time-period for the OSTIA and the Reynolds O.I. SST reanalyses and 4.19(b) is the same but for OSTIA and HadISST reanalyses. The Nino 3.4 region was chosen as it is widely used as a climatic index to catagorize whether an ENSO event is occuring (Trenberth,1997). It has been used to investigate how the representation of large scale climatic events such as ENSO within the OSTIA reanalysis compares to those in other analyses. Figure 4.19(a) illustrates that there is good agreement between the OSTIA and Reynolds reanalyses in their representaion of ENSO cycles, to the extent that the daily variability matches well. The strong El-Nino event in 1997 and the strong La-Nina in 1988-1989 can be discerned. Although HadISST as a monthly product is somewhat smoother than OSTIA the long time-scale variability matches reasonably well with large scale ENSO events being captured in both analyses, see figure 4.19(b). There is a slight cold bias in OSTIA relative to HadISST, although this is much smaller than the bias discussed previously for the global comparison.

4.6 Comparison of climatologies

The SST fields from the OSTIA reanalysis have been validated via comparisons to those from the HadISST reanalysis and the Reynolds O.I. SST reanalysis (see section 4.5). The reduced resolution 1/4 degree monthly SST climatologies generated from the full OSTIA reanalysis have

been compared to products generated for the same time period from the other SST reanalyses.

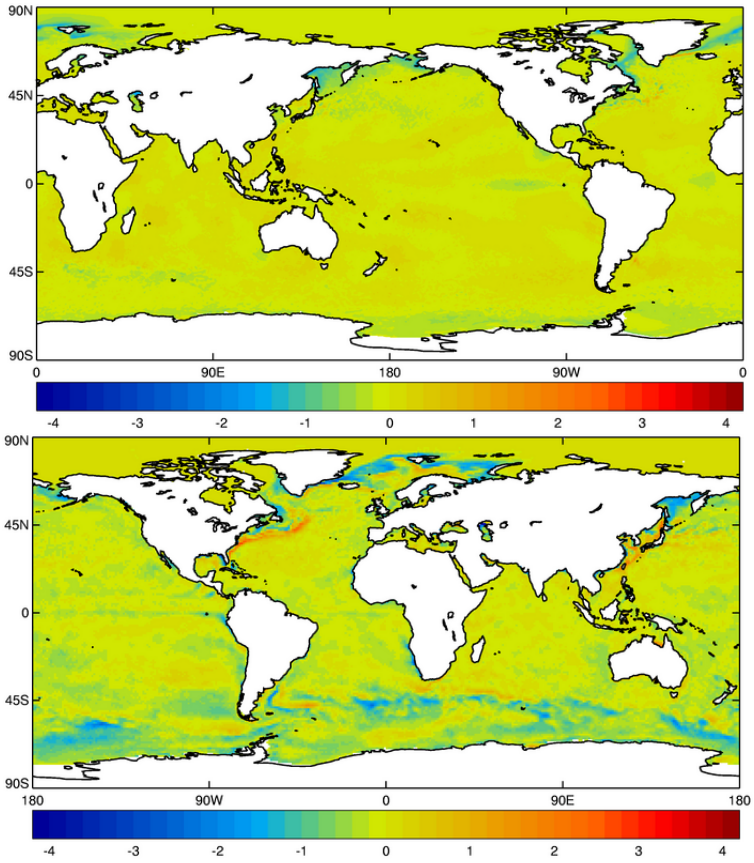


Figure 4.20: January climatological difference between OSTIA and Reynolds (Top) and OSTIA and HadISST (Bottom)

Figure 4.20 shows the differences between the OSTIA and Reynolds and the OSTIA and HadISST climatologies respectively for January. The left hand figure shows the general good agreement between OSTIA and Reynolds. The different ice extents used in the production of the SST fields is illustrated by the cold SST anomaly around the ice edge in the Northern Hemisphere in the OSTIA reanalysis. This is due to the ice extent used in OSTIA being larger than that used in Reynolds. The two climatologies also differ slightly in the high gradient SST regions such as the Gulf Stream and the Kuroshio current. The two analyses place the SST fronts in these regions in slightly different locations which manifests itself in the SST anomaly dipoles that can be observed in figure 4.20. The small cold anomaly in the OSTIA reanalysis in the Southern Ocean is in part due to the (A)ATSR data used in the OSTIA reanalysis but not in Reynolds. This anomaly is most evident during those years of the ATSR-1 mission as described in section 4.5. This is apparent when looking at the average monthly differences between the two analyses (not shown).

Figure 4.20 on shows a much greater disparity between the OSTIA and HadISST reanalyses compared to the OSTIA and Reynolds differences. OSTIA is colder over all ocean regions. The cold anomaly in OSTIA is most evident at mid to high latitudes in both the Northern and Southern

Hemispheres. Differences in the ice extents of the two analyses in the Arctic are discernable through the SST's with OSTIA having a larger ice extent than HadISST. The difference in resolution between the two analyses has manifested itself in a greater disparity in the location of the SST fronts in the high gradient SST regions than those shown for Reynolds. This has led to the dipolar banding which can be observed in the Gulf Stream and Kuroshio current regions as described previously. Similar patterns can be observed at the edge of the Antarctic circumpolar current off the southern coast of Africa and South America and at the western edge of the South Atlantic Gyre.

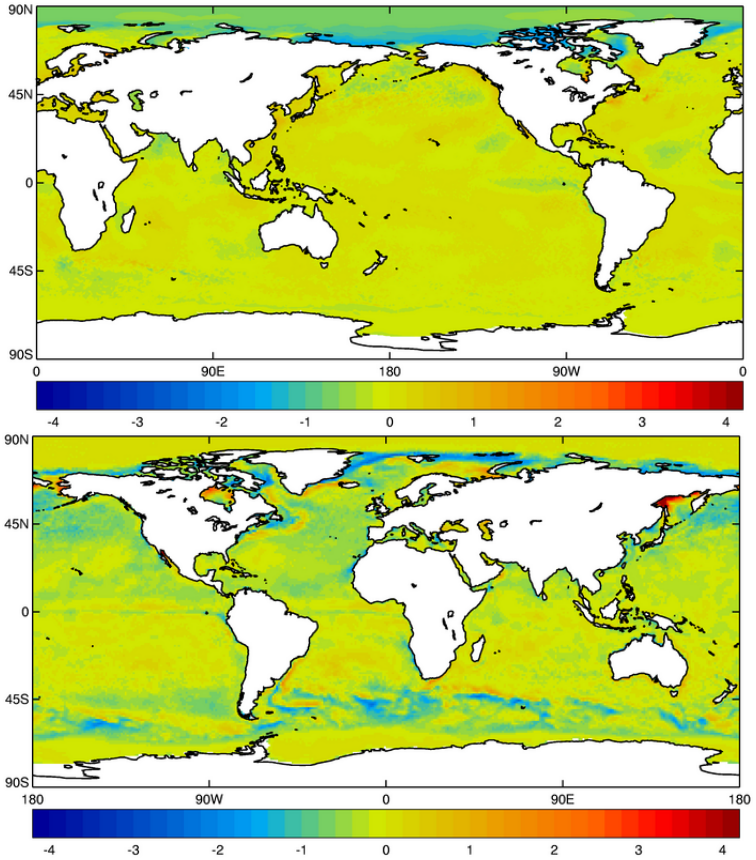


Figure 4.21: July climatological difference between OSTIA and Reynolds (Left) and OSTIA and HadISST (Right)

Figure 4.21 show the differences between OSTIA and Reynolds and OSTIA and HadISST climatologies for July respectively, illustrating the differences in Northern Hemisphere summer. Figure 4.21 on the left is very similar to the left hand of 4.20 with one major exception. A large cold anomaly in OSTIA at high northern latitudes right up to the pole under regions covered by ice during the Northern Hemisphere Summer is apparent. It is postulated that these differences are due to the effect of erroneous AVHRR Pathfinder data that was not used in the OSTIA reanalysis due to the ice-masking of the observations as described in section 2.2. If these suspect observations were used in the Reynolds reanalysis they would have led to the observed Reynolds warm bias. Comparison of figures 4.20 and 4.21 illustrate that the cold anomaly in the OSTIA analysis in the

Southern Ocean is reduced while a cold anomaly appears in the Northern Pacific and North Atlantic. This change indicates that a cold anomaly is apparent at high latitudes in OSTIA in the Summer Hemisphere. The differences in the ice extent and high gradient SST regions described previously for the January comparison are also still evident.

Comaprison of the right hands figures 4.21 to 4.20 shows similar seasonal cycle to that observed between the left hand figures 4.20 and 4.21. Again a cold anomaly relative to HadISST at high latitudes in the in the Summer Hemisphere is observed in OSTIA. This anomaly is larger relative to HadISST than Reynolds as described above. The warm anomaly in OSTIA in the Sea of Okhotsk and in Hudson Bay could be due to differences in the sea ice extent in the two regions between the two analyses as the ice retreats in the Northern Hemisphere summer. The features associated with the ice edge and the high gradient regions described previously are also still evident in July.

5. Summary and Conclusions

The OSTIA reanalysis provides a daily, high resolution, global SST and sea-ice analysis for the satellite period from 1st Jan 1985-31st Dec 2007. The reanalysis has been designed to be as temporally homogeneous as possible given the changes to observation network throughout the reanalysis time period. The data sources used in the reanalysis have been described in detail, as has the OSTIA reanalysis system.

Assessment of the OSTIA reanalysis assuming the in-situ observations to be 'truth' has shown that the accuracy of the analysis improves throughout the reanalysis time period. The in-situ observation minus background RMS decreases from approximately 1.00 K in 1985 to approximately 0.50 K in 2007, a significant factor int his reduction is the improvements in the accuaracy of the in-situ observations themselves. Once the drifting buoy network is mature the accuracy is within the operational target uncertainty of the NRT OSTIA system of 0.50 K. Regional variations exist and change throughout the time period as the observation networks, both satellite and in-situ, evolve. The way in which the OSTIA analysis system makes use of the different sources and missions varies over the reanalysis time period as the accuracy of the data changes. Problems with the ATSR-1 data led to it being bias corrected to the in-situ. These changes manifest themselves in the validation statistics for the satellite observations. Bearing this in mind validation statistics can also be used to asses the observations, assuming the analysis as 'truth'. Bias correcting the AVHRR Pathfinder data using the ATSR-2 and AATSR data has been shown to have a positive impact on the AVHRR statistics. Ship observations have been found to be biased which raises questions as to their inclusion in the reference data set used in the bias correction. Options are limited due to the scarcity of other in-situ observations prior to the drifting buoy network becoming mature after 2002. Future work may involve investigating the possibility of propagating the recent ship to drifting buoy bias back through the reanalysis. The use of an in-situ observation data set which included an estimate of the error information for each observation could also alleviate the problem.

The Mount Pinatubo eruption impacted the satellite observations used in the reanalysis at a global level. Biases in the satellite observations due to the volacnic aerosols are apparent for up to two years after the eruption. These biases were greater in magnitude and in latitudinal extent in the AVHRR Pathfinder data than in the ATSR-1 data. The validity of the bias correction appled to both data sources in the Southern Ocean during this period has been questioned due to the sparsisty

of reference in-situ data in the region. Future work on improving the bias correction scheme during this difficult period will be carried out.

Independent validation using the near-surface Argo data showed the OSTIA reanalysis to be biased cold globally by approximately 0.10 K. Regional variations have been found to exist in the standard deviation error of the analysis which varies between approximately 0.40 K in tropical regions to approximately 0.70 K in the North Atlantic. Comparisons of these independent validation statistics with those from the NRT OSTIA system have shown that both the mean error and the standard deviation error are higher in the reanalysis, demonstrating the benefit of the extra data used in the NRT system. A small cold bias in the OSTIA reanalysis compared to NRT OSTIA system exists across most of the ocean regions.

Assessment of the OSTIA reanalysis at high latitudes has shown that the SST and sea-ice fields are more consistent in the Southern than in the Northern Hemisphere. It was found that after periods in which the sea-ice field is persisted the SSTs are slower to respond to a sudden increase in the sea-ice extent than they are to a sudden decrease in the sea-ice extent once the sea-ice field is restored. The analysis at high latitudes also suffers from issues in Arctic summer due to inconsistencies in how the AVHRR and in-situ data sources flag day and night observations. This resulted in satellite but not in-situ data being used during the Arctic summer. The feasibility of using daytime measurements in the Arctic while minimising any diurnal warming signal should be investigated. Work to improve the consistency between the sea-ice and SST will be undertaken and may involve investigations into shortening of correlation length scales in high latitudes or changes in the relaxation timescales of the SST's under ice.

Comparison to similar SST reanalysis products have shown global warm bias in the OSTIA reanalysis compared to the HadISST reanalysis of approximately 0.50 K but good agreement with the Reynolds SST reanalysis. Similar magnitudes of inter-annual variability exist in all three reanalysis products. The OSTIA reanalysis is most evidently cooler than the other reanalyses during the ATSR-1 data period despite the efforts to mitigate against this. However although the differences are discernable in the global averages they are mostly limited to latitudes in excess of 40° North and South. Climatological differences have been shown to be greatest at mid to high latitudes in the Southern Hemisphere and are apparently due to differences in the ice extents of the three reanalyses. Differences in SST feature resolution of the three reanalyses products in the high gradient SST regions have been observed. Investigations into the cause of the warm bias in the OSTIA reanalysis relative to HadISST reanalysis through comparison of the bias corrections applied to the data in the two analysis systems will be carried out.

The next reanalysis using the OSTIA system will be carried out as part of the ESA SST CCI project using data sources generated during the project.

A. Appendix: Missing Data

The following tables provide a record of the dates with no data for each of the SST and sea ice datasets used in the OSTIA reanalysis. Note this does not include dates with limited or low data volumes in the case of SST data. Dates are written using the convention dd/mm.

Year	Dates
1985	02/01, 12/04-26/04 (SH missing only), 23/07 (SH), 21/09-29/09
1986	01/01, 26/03 (NH missing only), 28/03-25/05, 27/05-04/06 (SH), 06/06-24/06, 30/06 (SH), 02/07 (SH), 29/08 (NH), 02/09 (NH), 12/09 (NH), 16/09 (NH), 22/09 (NH), 12/10 (SH), 07/12-11/12, 15/12-19/12, 31/12 (SH)
1987	02/01-16/01, 23/03 (NH), 06/04-10/04

Table A.1: Dates with no data for SMMR (sea ice). Dates are for both hemispheres unless otherwise stated, where NH = northern hemisphere, SH = southern hemisphere. Note data available every other day. These dates are either a result of missing files, missing data within files or data not passing our QC checks.

Year	Dates
1987	13/07 (SH missing only), 24/08-27/08, 06/10-08/10, 03/12-31/12
1988	01/01-13/01, 26/01 (NH missing only), 05/05 (NH), 06/05-08/05, 09/05 (NH) 23/09, 24/12-28/12
1989	14/01, 15/01 (SH), 06/06-08/06, 21/07-25/07, 22/10-24/10
1990	12/08-14/08, 15/08-23/08 (SH), 24/08-27/08, 28/08-02/09 (SH), 10/09-15/09 (SH), 11/10-15/10 (SH), 20/10 (SH), 21/10-23/10, 25/10-28/10, 29/10 (SH), 02/11-03/11 (SH), 10/11-13/11 (SH), 21/12-27/12
1991	09/03, 27/08-29/08 (SH), 11/11 (SH), 12/11
1992	06/06 (NH), 07/06, 17/06-18/06, 30/12 (NH)
1993	03/01-05/01, 14/01 (NH), 16/01 (NH)
1994	07/03-08/03 (NH), 16/04 (NH), 19/07 (NH), 20/07-21/07, 20/09-01/10 (SH), 19/11-22/11
1995	24/03 (NH)
1996	30/05 (NH), 31/05, 04/06-05/06 (SH), 19/08-25/08 (SH), 13/09-16/09 (SH), 14/10-20/10 (SH)
1997	24/08-01/09 (SH)
1998	-
1999	-
2000	01/12-02/12, 17/09-19/09 (SH)
2001	02/09 (SH), 08/09-10/09 (SH)
2002	22/08-24/08 (SH)
2003	28/09-03/10 (SH), 31/10 (SH), 01/11 (SH)
2004	26/08-10/09 (SH), 16/09-21/10 (SH)
2005	-
2006	-
2007	-

Table A.2: Dates with no data for SSM/I (sea ice). Dates are for both hemispheres unless otherwise stated, where NH = northern hemisphere, SH = southern hemisphere. Data are available every day. These dates are either a result of missing files, missing data within files or data not passing our QC checks.

Year	Dates
1985	01/01-03/01, 09/02-11/02

Table A.3: Dates with no data for AVHRR Pathfinder.

Year	Dates
ATSR-1	
1991	10/08-13/08, 15/10-21/10, 24/10-26/10, 29/10-31/10, 11/12-27/12
1992	18/01, 27/02, 01/04, 05/04-13/04, 12/06-14/06, 26/06, 19/07-22/07
1993	15/02-16/02, 07/04, 31/10, 03/11-07/11
1994	20/04, 02/05-07/05, 13/08-14/08, 18/09-21/09, 16/10, 25/11-27/11, 03/12, 12/12-16/12
1995	20/01, 27/01, 02/02-03/02, 10-02, 15/02-17/02, 23/02
ATSR-2	
1996	01/01-30/06
1997	14/02
1998	04/06-05/06
1999	26/01-31/01, 27/02-31/08
2000	01/01, 08/02-09/02, 01/07-04/07, 08/10-10/10
2001	17/01-06/02, 22/05-23/05, 18/11
2002	05/02, 12/02-13/02, 09/03-19/03
AATSR	
2002	09/09-11/09, 19/11
2003	21/02-22/02, 16/03-18/03, 19/05, 05/09-06/09, 04/12
2004	09/06
2005	-
2006	06/04-07/04, 05/08-07/08, 08/09-10/09, 29/11, 13/12-15/12
2007	24/07, 25/09-27/09, 04/12-05/12

Table A.4: Dates with no data for ATSR multimission series.

References

S. Andersen, R. Tonboe, and L. Kaleschke. *Arctic sea ice thickness: Past, present and future*, chapter Satellite thermal microwave sea ice concentration algorithm comparison. Climate Change and Natural Hazards Series 10. EUR 22416, 2006a.

S. Andersen, R. Tonboe, S. Kern, and H. Schyberg. Improved retrieval of sea ice concentration from spaceborne passive microwave observations using numerical prediction model fields: An intercomparison of nine algorithms. *Rem. Sens. Env.*, 104:374–392, 2006b.

M. J. Bell, R. M. Forbes, and A. Hines. Assessment of the FOAM global data assimilation system for real-time operational ocean forecasting. *Journal of Marine Systems*, 25(1):1 – 22, 2000.

J. C. Comiso. Characteristics of Arctic winter sea ice from satellite multispectral microwave observations. *J. Geophys. Res.*, 91(C1):975–994, 1986.

J. C. Comiso, D. J. Cavalieri, C. L. Parkinson, and P. Gloersen. Passive microwave algorithms for sea ice concentration: A comparison of two techniques. *Rem. Sens. Env.*, 60:357–384, 1997.

C. J. Donlon, P. J. Minnett, C. Gentemann, T. J. Nightingale, I. J. Barton, B. Ward, and M. J. Murray. Toward improved validation of satellite sea surface skin temperature measurements for climate research. *J. Climate*, 15(4):353–369, 2002.

C.J. Donlon, M.J. Martin, J. Stark, J. Roberts-Jones, E. Fiedler, and W. Wimmer. The operational sea surface temperature and sea ice analysis (OSTIA). *Rem. Sens. Env.*, to appear, 2011.

S. Eastwood, K. R. Larsen, T. Lavergne, E. Nielsen, and R. Tonboe. *Global Sea Ice Concentration Reprocessing Product User Manual*. Met Norway/Danish Meteorological Institute, EUMETSAT Ocean and Sea Ice SAF, April 2010.

R. Evans and G. P. Podesta. *NOAA/NASA AVHRR Oceans Pathfinder Sea Surface Temperature Data Set: User's Reference Manual, Version 4.0*, chapter AVHRR Pathfinder Oceans Sea Surface Temperature Algorithm. Appendix G. Jet Propulsion Laboratory, publication D-14070, Pasedena, CA, 1998.

GHRSST-PP. *"The Recommended GHRSST-PP Data Processing Specification GDS (Version 1 revision 1.6)"*. GHRSST-PP, https://www.ghrsst.org/documents/q/category/gds-documents/operational/, April 2005.

P. Gloersen, W. J. Campbell, D. J. Cavalieri, J. C. Comiso, C. L. Parkinson, and H. J. Zwally. *Arctic and Antarctic sea ice, 1978-1987: satellite passive-microwave observations and analysis*. NASA SP-511, Washington D. C., 1997.

W.W. Gregg, M.A.M. Friedrichs, A.R. Robinson, K.A. Rose, R. Schlitzer, K.R. Thompson, and S.C. Doney. Skill assessment in ocean biological data assimilation. *Journal of Marine Systems*, 76: 16–33, 2009.

B. Ingleby and M. Huddleston. Quality control of ocean temperature and salinity profiles - historical and real-time data. *J. of Mar. Systems*, 65:158–175, DOI: 10.1016/j.jmarsys.2005.11.019, 2007.

E. J. Kearns, J. A. Hanafin, R. H. Evans, P. J. Minnett, and O. B. Brown. An independent assessment of Pathfinder AVHRR sea surface temperature accuracy using the Marine Atmosphere Emitted Radiance Interferometer (MAERI). *Bulletin of the American Meteorological Society*, 81:1525–1536, July 2000.

K. A. Kilpatrick, G. P. Podesta, and R. Evans. 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, 2001.

A. C. Lorenc, R. S. Bell, and B. Macpherson. The Meteorological Office analysis correction data assimilation scheme. *Quarterly Journal of the Royal Meteorological Society*, 117(497):59–89, 1991.

C. J. Merchant, A. R. Harris, M. J. Murray, and A. M. Zavody. Toward the elimination of bias in satellite retrievals of sea surface temperature, 1. Theory, modeling and interalgorithm comparison. *J. Geophys. Res.*, 104(C10):23565–23578, 1999.

NEODC. *AATSR factsheet*. NERC Earth Observation Data Centre, http://www.neodc.rl.ac.uk, March 2011.

NODC. *4 km Pathfinder Version 5 User Guide*. NODC, http://www.nodc.noaa.gov/sog/pathfinder4km/userguide.html, June 2009.

NSIDC. *National Snow and Ice Data Center*. NSIDC, http://nsidc.org/.

N. A. Rayner, D. E. Parker, E. B. Horton, C. K. Folland, L. V. Alexander, D. P. Rowell, E. C. Kent, and A. Kaplan. Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *J. Geophys. Res.*, 108(D14):4407, DOI: 10.1029/2002JD002670, 2003.

R.W. Reynolds. Impact of Mount Pinatubo aerosols on satellite-derived sea surface temperatures. *J. Climate*, 6:768–774, 1993.

R.W. Reynolds and D.B. Chelton. Comparisons of daily sea surface temperature analyses for 2007-2008. *J. Climate*, 23:3545–3562, 2010.

R.W. Reynolds, T.M. Smith, C. Liu, D.B. Chelton, and K.S. Casey. Daily high-resolution blended analysis for sea surface temperature. *J. Climate*, 20:5473–5496, 2007.

Ian S. Robinson. *Measuring the Oceans from Space*. Springer-Praxis, Chichester, UK, 2004.

D. M. Smith. Extraction of winter sea ice concentration in the Greenland and Barents Seas from SSM/I data. *Int. J. Rem. Sens.*, 17(13):2625–2646, 1996.

J. D. Stark, C. J. Donlon, M. J. Martin, and M. E. McCulloch. OSTIA : An operational, high resolution, real time, global sea surface temperature analysis system. *OCEANS 2007 Europe*, pages 1–4, 2007.

D. Storkey, E. W. Blockley, R. Furner, C. Guiavarc'h, D. Lea, M. J. Martin, R. M. Barciela, A. Hines, P. Hyder, and J. R. Siddorn. Forecasting the ocean state using NEMO: The new FOAM system. *Journal of Operational Oceanography*, 3(1):3–15, 2010.

R. Tonboe and E. Nielsen. *Global Sea Ice Concentration Reprocessing Validation Report*. Danish Meteorological Institute, EUMETSAT Ocean and Sea Ice SAF, version 1.1 edition, March 2010.

S. D. Woodruff, S. J. Worley, S.J. Lubker, Z. Ji, E. Freeman, D. I. Berry, P. Brohan, E. C. Kent, R. W. Reynolds, S. R. Smith, and C. Wilkinson. ICOADS Release 2.5: Extensions and enhancements to the surface marine meteorological archive. *Int. J. Climatol.*, DOI: 10.1002/joc.2103, 2010.

S. J. Worley, S. D. Woodruff, R. W. Reynolds, S.J. Lubker, and N. Lott. ICOADS Release 2.1 data and products. *Int. J. Climatol.*, 25:823–842, DOI: 10.1002/joc.1166, 2005.

Met Office
FitzRoy Road, Exeter
Devon, EX1 3PB
UK

Tel: 0870 900 0100
Fax: 0870 900 5050
enquiries@metoffice.gov.uk
www.metoffice.gov.uk