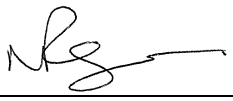
Customer	**:**	**ESRIN**	**Document Ref**	**:**	**SST_CCI-CAR-UKMO-001**
WP No	**:**	**40630**	**Issue Date**	**:**	**24 January 20143**
			Issue	**:**	1

Project : **CCI Phase 1 (SST)**

Title : **ESA SST CCI Climate Assessment Report**

Abstract : This document summarises the climate assessment of the ESA SST CCI LT products. This includes comparison of the products to other climate data sets of SST, their use in climate model evaluation and comparison of some aspects to OC CCI products. It also includes preliminary feedback from volunteer test users.

Author :



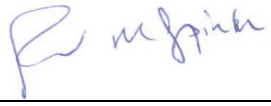
Nick Rayner, John Kennedy, Chris Atkinson, Tim Graham, Met Office Hadley Centre

Emma Fiedler, Alison McLaren, Met Office

Gary Corlett, University of Leicester

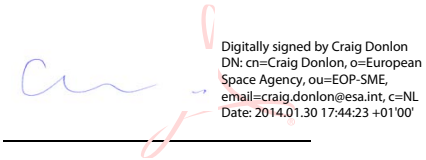
And the trail blazer users

Checked :



P. Spinks, Project Manager
Space ConneXions Ltd.

Accepted by ESA :



Craig Donlon
ESA Technical Officer

Acceptance date:

Distribution : SST_cci team members

ESA (Craig Donlon)

EUROPEAN SPACE AGENCY
CONTRACT REPORT

The work described in this report was done under ESA contract. Responsibility for the contents resides in the author or organisation that prepared it.

AMENDMENT RECORD

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

AMENDMENT RECORD SHEET

ISSUE	DATE	REASON FOR CHANGE
A	06 Dec 2013	Initial Issue
B	20 Jan 2014	Amended following ESA review

TABLE OF CONTENTS

1. INTRODUCTION

1.1 Purpose and Scope

This document summarises the climate assessment of the European Space Agency sea surface temperature Climate Change Initiative (ESA SST CCI) long term (LT) products, version 1.0. This includes comparison of the products to other climate data sets of SST and an assessment of their stability. It documents their use in climate model evaluation and comparison of some aspects to Ocean Colour (OC) CCI products. It also includes preliminary feedback (on a pre-release version which does not differ from version 1.0 in any way that impacts the studies undertaken, except where reported otherwise) from volunteer test users, including an assessment of the scientific utility of the products and comments on their ease of use and the quality of their documentation.

We assess all three ESA SST CCI LT version 1.0 products:

- ATSR. SSTs from Along Track Scanning Radiometer (ATSR) instruments in L3U format at 0.05° latitude by 0.05° longitude resolution covering 1991 – 2010. (Shortened to SST CCI ATSR.)

- Analysis. Satellite-only SST-depth L4 daily analysis created by the Operational Sea Surface Temperature and sea Ice Analysis (OSTIA) system from SST CCI ATSR and SST CCI Advanced Very High Resolution Radiometer (AVHRR) products at 0.05° latitude by 0.05° longitude resolution covering 1991 – 2010. (Shortened to SST CCI analysis.)

- AVHRR. SSTs from AVHRR instruments in L2P format at Global Area Coverage (GAC) resolution covering 1991 – 2010. (Shortened to SST CCI AVHRR.)

In future, we will use a mailing list to advertise the availability of the products and invite all potential users to download and test the data. Feedback from this document and subsequent user interactions will be folded in to assessments of User Requirements in Phase 2 of the CCI.

1.2 Structure of the Document

After this introduction, the document is divided into a number of major sections that are briefly described below:

Section 2 gives an Executive Summary of the key scientific results, feedback on scientific utility of the products and ease of use of the products and the documentation in the view of the preliminary users, along with recommendations by the users.

Section 3 presents an assessment of trends and variability in the SST CCI LT version 1.0 products and comparison to other SST products and an assessment of their stability. In order to assess the multi-annual and decadal behaviour of the long-term products, comparisons are made to existing SST data sets including those used by the Intergovernmental Panel on Climate Change (IPCC) in their 5th Assessment Report as well as data sets used in other high profile monitoring reports. Differences between the SST CCI products and the comparison datasets are highlighted. The stability of the SST CCI datasets has been assessed through comparisons to Global Tropical Moored Buoy Array (GTMBA) moorings at 1m depth. The stability of the SST CCI analysis has also been compared to that of other analyses in the Group for High Resolution SST (GHRSST) Multi Product Ensemble.

Section 4 details the use of the SST CCI products in model evaluation by the SST CCI Climate Research Group at the Met Office Hadley Centre, along with voluntary reports from trail blazer users describing their applications and what they have discovered from using the products.

Section 5 examines the relationship between SST CCI and OC CCI products in their representation of fronts and compares the results to those derived from a precursor product.

Section 6 lists any further reported issues identified by the trail blazer users and any other recommendations they have made for future SST CCI products.

1.3 Referenced Documents

The following is a list of documents with a direct bearing on the content of this report. Where referenced in the text, these are identified as RD.n, where 'n' is the number in the list below:

RD.326	Atkinson, C.P., N.A. Rayner, J. Roberts-Jones, R.O. Smith (2013), Assessing the quality of sea surface temperature observations from drifting buoys and ships on a platform-by-platform basis. Journal of Geophysical Research - Oceans, 118, 3507–3529, doi:10.1002/jgrc.20257.
RD.341	Good, S.A., M.J. Martin and N.A. Rayner (2013) EN4: Quality controlled ocean temperature and salinity profiles and monthly objective analyses with uncertainty estimates, Journal of Geophysical Research: Oceans, doi: 10.1002/2013JC009067.
RD.72	Rayner, N.A., P.Brohan, D.E.Parker, C.K.Folland, J.J.Kennedy, M.Vanicek, T.Ansell and S.F.B.Tett 2006: Improved analyses of changes and uncertainties in sea surface temperature measured in situ since the mid-nineteenth century: the HadSST2 data set. Journal of Climate. 19(3) pp. 446-469
RD.330	ESA SST CCI Product Validation and Intercomparison Report
RD.332	Woodruff, S.D., S.J. Worley, S.J. Lubker, Z. Ji, J.E. Freeman, D.I. Berry, P. Brohan, E.C. Kent, R.W. Reynolds, S.R. Smith, and C. Wilkinson, 2011: ICOADS Release 2.5: Extensions and enhancements to the surface marine meteorological archive. Int. J. Climatol. (CLIMAR-III Special Issue), 31, 951-967 (doi:10.1002/joc.2103).
RD.334	Rayner, N. A., J. J. Kennedy, R. O. Smith and H. A. Titchner (2013) The Met Office Hadley Centre Sea Ice and Sea Surface Temperature data set, version 2, part 3: the combined analysis. In prep. for JGR Atmospheres
RD.210	Kennedy J.J., Rayner, N.A., Smith, R.O., Saunby, M. and Parker, D.E. (2011b). Reassessing biases and other uncertainties in sea-surface temperature observations since 1850 part 1: measurement and sampling errors. J. Geophys. Res., 116, D14103, doi:10.1029/2010JD015218
RD.211	Kennedy J.J., Rayner, N.A., Smith, R.O., Saunby, M. and Parker, D.E. (2011c). Reassessing biases and other uncertainties in sea-surface temperature observations since 1850 part 2: biases and homogenisation. J. Geophys. Res., 116, D14104, doi:10.1029/2010JD015220
RD.216	Casey, K.S., T.B. Brandon, P. Cornillon, and R. Evans (2010). "The Past, Present and Future of the AVHRR Pathfinder SST Program", in Oceanography from Space: Revisited, eds. V. Barale, J.F.R. Gower, and L. Alberotanza, Springer. DOI: 10.1007/978-90-481-8681-5_16.

RD.74	Rayner, N. A.; Parker, D. E.; Horton, E. B.; Folland, C. K.; Alexander, L. V.; Rowell, D. P.; Kent, E. C.; Kaplan, A. (2003) Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century J. Geophys. Res.Vol. 108, No. D14, 4407 10.1029/2002JD002670
RD.342	Kaplan, A., Y. Kushnir, M. Cane, and M. Blumenthal (1997) Reduced space optimal analysis for historical data sets: 136 years of Atlantic sea surface temperatures, J. Geophys. Res., 102, 27,835– 27,860
RD.79	Smith, T.M., R.W. Reynolds, Thomas C. Peterson, and Jay Lawrimore, 2008: Improvements to NOAA's Historical Merged Land-Ocean Surface Temperature Analysis (1880-2006). Journal of Climate, 21, 2283-2296.
RD.81	Kaplan, A., et al. (1998), Analyses of global sea surface temperature 1856-1991, Journal of Geophysical Research-Oceans, 103(C9), 18567-18589.
RD.85	Ishii, M., Shouji, A., Sugimoto, S. and Matsumoto, T. (2005), Objective analyses of sea-surface temperature and marine meteorological variables for the 20th century using ICOADS and the Kobe Collection. Int. J. Climatol., 25: 865–879. doi: 10.1002/joc.1169
RD.103	Berry, David I.; Kent, Elizabeth C. 2011 Air–Sea fluxes from ICOADS: the construction of a new gridded dataset with uncertainty estimates. International Journal of Climatology, 31 (7). 987-1001. 10.1002/joc.2059
RD.233	Karspeck, A. R., Kaplan, A. and Sain, S. R. (2012), Bayesian modelling and ensemble reconstruction of mid-scale spatial variability in North Atlantic sea-surface temperatures for 1850–2008. Q.J.R. Meteorol. Soc., 138: 234–248. doi: 10.1002/qj.900
RD.212	Reynolds, R.W., N.A. Rayner, T.M. Smith, D.C. Stokes, and W. Wang, 2002: An improved in situ and satellite SST analysis for climate. J. Climate, 15, 1609-1625
RD.239	Roberts-Jones, Jonah, Emma Kathleen Fiedler, Matthew James Martin, 2012: Daily, Global, High-Resolution SST and Sea Ice Reanalysis for 1985–2007 Using the OSTIA System. J. Climate, 25, 6215–6232. doi: http://dx.doi.org/10.1175/JCLI-D-11-00648.1
RD.76	Reynolds, Richard W., Thomas M. Smith, Chunying Liu, Dudley B. Chelton, Kenneth S. Casey, Michael G. Schlax, 2007: Daily High-Resolution-Blended Analyses for Sea Surface Temperature. J. Climate, 20, 5473–5496. doi: http://dx.doi.org/10.1175/2007JCLI1824.1
RD.343	Lanzante, J.R. (1996), Resistant, Robust and Non-Parametric Techniques for the Analysis of Climate Data: Theory and Examples, Including Applications to Historical Radiosonde Station Data. Int. J. Climatol., 16: 1197–1226. doi: 10.1002/(SICI)1097-0088(199611)16:11<1197::AID-JOC89>3.0.CO;2-L
RD.226	MacCallum and Merchant (2012), SST CCI Algorithm Selection Report, http://www.esa-sst-cci.org/
RD.356	Knight, J.R., R.J. Allan, C.K. Folland, M. Vellinga and M.E. Mann (2005), A signature of persistent natural thermohaline circulation cycles in observed climate, Geophys. Res. Lett., 32, L20708, doi: 10.1029/2005GL024233.
RD.339	Ohring, G., Wielicki, B., Spencer, R., Emery, B. & Datla, R. (2005). Satellite Instrument Calibration for Measuring Global Climate Change: Report of a Workshop. Bulletin of the American Meteorological Society, 86, 1303-1313
RD.296	Merchant, C. J., O. Embury, N. A. Rayner, D. I. Berry, G. Corlett, K. Lean, K. L. Veal, E. C. Kent, D. Llewellyn-Jones, J. J. Remedios, and R.

	Saunders (2012), A twenty-year independent record of sea surface temperature for climate from Along Track Scanning Radiometers, J. Geophys. Res., 117, C12013, doi:10.1029/2012JC008400.
RD.355	Oka, E. and Ando, K. (2004). Stability of Temperature and Conductivity Sensors of Argo Profiling Floats. Journal of Oceanography, 60, 2, 253-258.
RD.337	Wimmer, W., Robinson, I. S. & Donlon, C. J. (2012). Long-term validation of AATSR SST data products using shipborne radiometry in the Bay of Biscay and English Channel. Remote Sensing of Environment, 116, 17-31.
RD.338	Minnett, P. J. & Corlett, G. K. (2012). A pathway to generating Climate Data Records of sea-surface temperature from satellite measurements. Deep Sea Research Part II: Topical Studies in Oceanography, 77–80, 44-51.
RD.317	Merchant, C. J., Mittaz, J. and Corlett, G. K. (2013), Climate data evaluation framework, GHRSST document reference CDR-TAG_CDEF/Version 1.0.
RD.344	Brasnett, B. (2012). A 20-year reanalysis of sea surface temperature. Report, CMC.
RD.345	Kennedy, J. J., Rayner, N. A., Millington, S.C. & Saunby, M. (2013). The Met Office Hadley Centre sea ice and sea surface temperature data set, version 2, part 2: Sea surface temperature analysis. In prep.
RD.237	Kurihara, Y., Sakurai, T. and Kuragano, T. (2006). Global daily sea surface temperature analysis using data from satellite microwave radiometer, satellite infrared radiometer and in situ observations. Weather Bulletin, 73, 1-18. (In Japanese)
RD.357	Merchant C J, Embury O, Le Borgne P and Bellec B (2006), Saharan dust in night-time thermal imagery: detection and reduction of related biases in retrieved sea surface temperature, Rem. Sens. Env., 104, 15-30.
RD.358	Legeckis, R., 1977. Long waves in the eastern equatorial pacific ocean: A view from a geostationary satellite. Science 197, 1179–1181.
RD.359	Menkes, C.E.R., Vialard, J.G., Kennan, S.C., Boulanger, J.P., Madec, G.V., 2006. A modelling study of the impact of tropical instability waves on the heat budget of the eastern equatorial pacific. J. Phys. Oceanogr. 36, 847–865.
RD.360	Jochum, M., Cronin, M.F., Kessler, W.S., Shea, D., 2007. Observed horizontal temperature advection by tropical instability waves. Geophysical Research Letters 34.
RD.361	Brodeau, L., Barnier, B., Treguier, A., Penduff, T., Gulev, S., 2010. An ERA40-based atmospheric forcing for global ocean circulation models. Ocean Modelling 31, 88–104.
RD.362	O'Dea, E. J., A. K. Arnold, K. P. Edwards, R. Furner, P. Hyder, M. J. Martin, J. R. Siddorn et al. "An operational ocean forecast system incorporating NEMO and SST data assimilation for the tidally driven European North-West shelf." Journal of Operational Oceanography 5, no. 1 (2012): 3-17.
RD.363	Blockley, E. W., Martin, M. J., McLaren, A. J., Ryan, A. G., Waters, J., Lea, D. J., Mirouze, I., Peterson, K. A., Sellar, A., and Storkey, D.: Recent development of the Met Office operational ocean forecasting system: an overview and assessment of the new Global FOAM forecasts, Geosci. Model Dev. Discuss., 6, 6219-6278, doi:10.5194/gmdd-6-6219-2013, 2013.
RD.364	Dee, D. P., S. M. Uppala, A. J. Simmons, P. Berrisford, P. Poli, S. Kobayashi, U. Andrae et al. "The ERA□Interim reanalysis: Configuration and performance of the data assimilation system." Quarterly Journal of the Royal Meteorological Society 137, no. 656 (2011): 553-597.

RD.365	Waters, J., M. Martin, J. While, D. Lea, A. Weaver, I. Mirouze. Implementing a variational data assimilation system in an operational 1/4 degree global ocean model. Submitted to Q. J. R. Meteorol. Soc
RD.66	Ingleby, Bruce, and Matt Huddleston. "Quality control of ocean temperature and salinity profiles—Historical and real-time data." Journal of Marine Systems 65, no. 1 (2007): 158-175.
RD.366	Taylor, K.E., D. Williamson, and F. Zwiers: 2000. The Sea Surface Temperature and Sea-Ice Concentration Boundary Conditions for AMIP II Simulations. PCMDI Report No. 60, 28 pp.
RD.367	Kosaka and Xie, 2013, Nature Research Letters, doi:10.1038/nature12534
RD.368	Rossow, WB and RA Schiffer, 1999: Advances in understanding clouds from ISCCP. Bull. Amer. Met. Soc., 80, 2261-2287. doi: 10.1175/1520-0477(1999)080<2261:AIUCFI>2.0.CO;2
RD.346	Cayula, J-F., and Cornillon, P. (1992) Edge detection algorithm for SST Images. Journal of Atmospheric and Oceanic technology, 9, 67-80.
RD.347	Stegmann, P. M., and Ullman, D. S. (2004) Variability in chlorophyll and sea surface temperature fronts in the Long Island Sound outflow region from satellite observations. Journal of Geophysical Research, doi:10.1029/2003JC001984.
RD.348	Castelao, R.M., Mavor, T. P., Barth, J. A., Breaker, L. C. (2006) Sea surface temperature fronts in the California Current System from geostationary satellite observations, Journal of Geophysical Research, doi:10.1029/2006JC003541.
RD.349	Belkin, I.M., Cornillon, P.C. (2005) Bering Sea thermal fronts from Pathfinder data: Seasonal and interannual variability. Pacific Oceanography 2, 6-20.
RD.350	Belkin, I.M., Cornillon, P.C., and Sherman, K. (2009) Fronts in Large Marine Ecosystems. Progress in Oceanography, 81, 223-236.
RD.351	Wall, C. C., Muller-Karger, F. E., Roffer, M. A., Hu, C., Yao, W., and Luther, M. E. (2008) Satellite remote sensing of surface oceanic fronts in coastal waters off west-central Florida. Remote Sensing of Environment, 112, 2963-2976
RD.352	Obenour, K., Cornillon, P., and Buckingham, C. (2010) Detection and Probability of Ocean Fronts in AMSR-E Satellite Data. Poster. Ocean Sciences 2010, Portland, Oregon, USA. [copy at ...]
RD.353	Kahru, M., Lorenzo, E. D., Manzano-Sarabia, M., and Mitchell, B. G. (2012) Spatial and temporal statistics of sea surface temperature and chlorophyll fronts in the California Current. Journal of Plankton Research, doi:10.1093/plankt/fbs010.
RD.247	Ullman, D. S., and Cornillon, P. (2000) Evaluation of front detection methods for satellite-derived SST data using in situ observations. Journal of Atmospheric and Oceanic technology, 17, 1667-1675.
RD.354	Storm, T. and M. Boettcher (2012) OC CCI Product User Guide Issue 1.0. http://www.esa-oceancolour-cci.org/?q=webfm_send/231

1.4 Definitions of Terms

The following terms have been used in this report with the meanings shown.

Term	Definition
AATSR	Advanced Along-Track Scanning Radiometer
AMIP	Atmospheric Model Intercomparison Project
AMM	Atlantic Margin Model implementation of FOAM
AMSR	Advanced Microwave Scanning Radiometer
ARC	ATSR Reprocessing for Climate
ATSR	Along-Track Scanning Radiometer
ATSR-1	First ATSR instrument
ATSR-2	Second ATSR instrument
AVHRR	Advanced Very High Resolution Radiometer
CAR	Climate Assessment Report
CCI	Climate Change Initiative
CDAF	Climate Data Assembly Framework (of GHRSST)
CDR	Climate Data Record
CDRP	Climate Data Research Package
chl-a	Chlorophyll-a
CMC	Canadian Meteorological Centre
CMIP5	Coupled Model Intercomparison Project (Phase 5)
CMUG	Climate Modelling User Group
COBE	Centennial in situ Observation-Based Estimates of the variability of SSTs
CRG	Climate Research Group
DMI	Danmarks Meteorologiske Institut (Danish Meteorological Institute)
DMI	Dipole Mode Index
ECMWF	European Centre for Medium-Range Weather Forecasts
ECV	Essential Climate Variable
EN3 or 4	Met Office Hadley Centre dataset
ENSO	El Nino Southern Oscillation
ENVISAT	ENVIronmental SATellite
EOF	Empirical Orthogonal Function
ERA	ECMWF Re-Analysis
ERA-40	ECMWF Re-analysis covering 40 years
ERA-CLIM	ECWMF Reanalysis Archive for Climate
ERA-interim	ECWMF Interim Reanalysis Archive for Climate
ERS-2	Second European Remote Sensing satellite
ERSEM	European Regional Seas Ecosystem Model

Term	Definition
ERSSTv3	Extended Reconstruction SST V3
ESA	European Space Agency
FCDR	Fundamental Climate Data Record
FOAM	Met Office short-range operational ocean forecasting model
GAC	Global Area Coverage
GC	Global Coupled (MOHC model versions)
GCM	General Circulation Model
GCOS	Global Climate Observing System
GDS	GHRSST Data Specification
GEWEX	Global Energy and Water cycle Experiment (of the WCRP)
GHRSST	Group for High Resolution SST
GMPE	GHRSST-PP Multi-Product Ensemble
GPCP	Global Precipitation Climatology Project
GTMBA	Global Tropical Moored Buoy Array
GTS	Global Telecommunication System (of WMO)
HadGEM	Hadley Centre Global Environmental Model
HadISST	Hadley Centre Global sea-Ice coverage and SST
HadSST	Hadley Centre Sea Surface Temperature
ICOADS	International Comprehensive Ocean-Atmosphere Data Set
IDL	Interactive Data Language
IOD	Indian Ocean Dipole
IPCC	Intergovernmental Panel on Climate Change
ISCCP	International Satellite Cloud Climatology Programme
ITCZ	Inter-Tropical Convergence Zone
K	Kelvin
L2P	Level 2 (Pre-processed)
L3U	Level 3 uncollated
L4	Level 4
LT	Long term products
MD	Match-up Dataset
MERIS	Medium Resolution Imaging Spectrometer
MGDSST	Merged satellite and in situ data Global Daily SST
MIZ	Marginal Ice Zone
MJO	Madden-Julian Oscillation
MMD	Multi-sensor Match-up Dataset
MODIS	Moderate Resolution Imaging Spectroradiometer
MOHC	UK Met Office Hadley Centre

Term	Definition
MOHSST	Met Office Historical Sea Surface Temperature data set
MSLP	Mean Sea-Level Pressure
N216	Dynamical model at 60km resolution
NASA	National Aeronautics and Space Administration (USA)
NCAR	National Center for Atmospheric Research (USA)
NCDC	National Climate Data Center (US)
NCEO	National Centre for Earth Observation
NCEP	National Center for Environmental Prediction (US)
NEMO	Ocean modelling framework
NEMOVAR	NEMO variational data assimilation system
NEODC	NERC Earth Observation Data Centre
NERC	Natural Environment Research Council
netCDF	Network Common Data Format
NOAA	National Oceanic and Atmospheric Administration (USA)
NOAA-NCDC	NOAA National Climatic Data Center
NOAA-PMEL	NOAA Pacific Marine Environment Laboratory
NOCS	National Oceanographic Centre Southampton
NOCL	National Oceanographic Centre at Liverpool
NODC	National Oceanographic Data Center
NOOS	North West European Shelf Operational Oceanographic System
NWP	Numerical Weather Prediction
OC	Ocean colour
OI	Optimum interpolation
OI.v2	Reynolds et al (2002) Optimal Interpolation analysis
ORCA025	0.25° ocean model developed under the NEMO framework
ORCA1	1° ocean model developed under the NEMO framework
OSI SAF	Ocean and Sea Ice Satellite Application Facility
OSISAF	EUMETSAT Ocean and Sea Ice Satellite Application Facility
OSI-SAF	Ocean and Sea Ice Satellite Applications Facility
OSTIA	Ocean Surface Temperature and Ice Analysis
PCMDI	Program for Climate Model Diagnosis and Intercomparison
PDO	Pacific Decadal Oscillation
PGM4	Persian Gulf Model
PIRATA	PIlot Research moored Array in the Tropical Atlantic
PMEL	Pacific Marine Environment Laboratory
PUG	Product User Guide
PVIR	Product Validation and Inter-comparison Report

Term	Definition
RAMA	Research Moored Array for African-Asian-. Australian Monsoon Analysis and Prediction
RD	Reference Document
RMS	Root Mean Square
RSMAS	Rosenstiel School of Marine and Atmospheric Science (U Miami)
RSOI	Reduced Space Optimal Interpolation
RSOS	Reduced Space Optimal Smoother
SD	Standard Deviation
SeaWIFS	Sea-viewing wide field-of-view Sensor
SIED	Single Image Edge Detection
SOI	Southern Oscillation Index
SMHI	Swedish Meteorological and Hydrological Institute
SST	Sea Surface Temperature
SSTA	Sea Surface Temperature Anomaly
SSTdepth	Sea Surface Temperature (depth)
SSTfnd	Foundation SST
SSTskin	Sea Surface Temperature (skin)
SST-TR-n (n=1-48)	SST Technical Requirement
STD	Standard Deviation
TAMG	Tropical Atlantic Meridional SST Gradient
TAO	Tropical Atmosphere Ocean
TIW	Tropical Instability Wave
TOGA	Tropical Ocean and Global Atmosphere
TRITON	Triangle Trans Ocean Buoy Network
VOS	Voluntary Observing Ship
WMO	World Meteorological Organisation

2. EXECUTIVE SUMMARY

Here we provide a bullet point summary of the key points from this Climate Assessment Report.

The three ESA SST CCI LT version 1.0 products assessed are:

- ATSR. SSTs from ATSR instruments in L3U format at 0.05° latitude by 0.05° longitude resolution covering 1991 – 2010. (Hereafter, SST CCI ATSR.)
- Analysis. Satellite-only SST-depth L4 daily analysis created by OSTIA system from SST CCI ATSR and SST CCI AVHRR products at 0.05° latitude by 0.05° longitude resolution covering 1991 – 2010. (Hereafter, SST CCI analysis.)
- AVHRR. SSTs from AVHRR instruments in L2P format at Global Area Coverage (GAC) resolution covering 1991 – 2010. (Hereafter, SST CCI AVHRR.)

These are utilised over the period 1991-2010.

New User Requirements identified in this Climate Assessment Report will be analysed as part of the User Requirements assessment for SST CCI Phase II. Potential users should be aware of some features of the SST CCI products before choosing to use them for particular studies. These features are highlighted here and will also be flagged up in the Product User Guide.

Comparison of SST CCI products to other climate SST data sets (Section 3.1):

- The SST CCI LT products are comparable to the main population of comparison data sets in terms of their general evolution of SST and variability. They show smaller differences from HadSST3 on the global average than some comparison datasets (e.g., SST CCI AVHRR [0 K to -0.2 K, with some larger excursions] compared to AVHRR Pathfinder v5.2 [~ -0.4 K]; and SST CCI analysis [~ 0 K] compared to Daily OI [~ -0.1 K]).
- The AVHRR Pathfinder data set is an outlier amongst the data sets compared, with generally relatively cool-biased SSTs.
- There are some anomalous periods of larger bias in the SST CCI AVHRR (four individual months between 1991 and 1994) and SST CCI ATSR (one month in 1996) products.
- SST CCI AVHRR products are inconsistent in their overall trends relative to SST CCI ATSR (and hence the SST CCI analysis). The SST CCI ATSR and SST CCI analysis products warm more rapidly than the majority of the comparison data sets, but the SST CCI AVHRR products tend to warm less rapidly over the period 1992-2010.
- For some climate indices, peak-to-peak variability of SST CCI LT products is higher than in the comparison data. This appears to be due to better feature resolution in the SST CCI LT products, arising from better sampling.
- The period prior to 1995 is characterised by larger cool biases in the SST CCI LT products relative to in situ based data sets, especially in the mid latitudes. As this is seen in comparisons both to the GTMBA and to the wider in situ network, we conclude that these differences arise from the SST CCI products.
- The SST CCI AVHRR product in the northern Indian Ocean exhibits spurious variation on annual timescales.

- Multi-year variability in the SST CCI LT products is generally consistent with that seen in the comparison data.
- Month-to-month correlation is higher in the SST CCI LT products and indicates they are less noisy than in situ only products (due to better sampling). Lag correlations at greater than one month lag are lower than in comparable products in places.

Stability of SST CCI products (Section 3.2):

- The GCOS stability requirement (of better than 3 mK/year, estimated here by evaluating the trend in the difference between SST CCI products and GTMBA measurements) is met in the tropical Pacific over the period 1995-2010 by the SST CCI AVHRR night-time product (-2.0 < trend < 2.0 mK/year). The SST CCI ATSR daytime (0.7 < trend < 3.2 mK/year) and SST CCI analysis (0.1 < trend < 3.2 mK/year) products have confidence intervals that go marginally outside this target over the same time period. Stabilities of other products fall outside of the GCOS requirement. [SST CCI AVHRR day time: -12.3 < trend < -7.4 mK/year; SST CCI ATSR night-time: -1.4 < trend < 6.4 mK/year]
- Tropical Pacific SST CCI ATSR day time stability over 1995-2010 is comparable to that of the ARC data set (a precursor ATSR-based analysis). Night time stability over the same period is less good.
- Tropical Pacific SST CCI ATSR day (-13.6 < trend < 60.1 mK/year) and night-time (-7.4 < trend < 36.8 mK/year) stability over the period 1991-1995 is improved relative to that of the ARC data set.
- Stability outside the tropical Pacific cannot be evaluated using known techniques because insufficient stable reference data exist away from this region.
- The SST CCI analysis is apparently less stable in the tropics with respect to GTMBA measurements than the other GMPE analyses. However, the non-SST CCI analyses ingest GTMBA data, and this lack of independence means the results obtained are likely to exaggerate their true stability in non-GTMBA locations.
- The stability of the SST CCI analysis improves over successive ATSR missions, reflecting the improvements in stability of the successive instruments in the series, and/or improvements in the stability of GTMBA over similar time periods.
- Methods to assess stability over wider geographical areas using, for example, Argo and improvements to the moored buoy network, are required.
- A consistent approach to step-detection techniques is required across several CCI ECVs and domains. Preferably, specialist statistical expertise in this area would be embedded in a tool or methodology that can be confidently used by CCI science teams.

Impact of the use of SST CCI products in model evaluation (Section 4.1):

- The SST CCI analysis product is a suitable tool for the evaluation of coupled model mean state.
- Differences between observational estimates are smaller than differences between simulated and observed SST. However, they are of similar size to differences between model versions. Therefore, in the future, it should be possible to meaningfully use observational uncertainties to assess significance of differences from observations.

- Comparing daily SST variability in a HadGEM3 model control run with two SST analyses (SST CCI analysis and Reynolds et al [2007] Daily OI), shows that the model and observed variability broadly compares well in terms of both magnitudes and spatial patterns.

- The magnitude of HadGEM3 model SST anomaly variability shows a better general agreement with the SST CCI analysis product than with the Daily OI, demonstrating the potential value of the SST CCI dataset for this kind of analysis.

- Assumptions made about the correlation structure of analysis errors have an impact on the details of conclusions drawn when comparing simulated and observed variability, so more work is needed to understand the correlation structure of SST CCI analysis errors.

- Variability in the single-view AVHRR SSTs may be exaggerated in the Gulf of Arabia/Arabian Sea owing to biases arising from intermittent desert dust. There is evidence that there may be a positive impact of the use of passive microwave data in this region (as proposed in an option to SST CCI Phase II by DMI). The Gulf of Arabia/Arabian Sea is a particularly complicated region for the retrieval of SST, but also very important for climate modelling. In the Arabian Sea the SST CCI analysis is up to 0.6 K cooler than some other datasets in June-July-August (JJA) period. This cooler SST causes a significant reduction in the simulated Indian monsoon rainfall when used to drive a 25 km model, magnifying an existing model bias. A detailed exploration of the representation of these regions in SST CCI products should be undertaken in discussion with climate modellers.

Key points arising from use of the SST CCI products in trail blazing applications (views of trail-blazer users):

- Daily mean SST data provided by the SST CCI analysis is much more useful than foundation SST data (such as in MyOcean OSTIA Reanalysis v1) for the purpose of evaluating model simulations of heat transport by tropical instability waves. This is because the daily mean SST CCI analysis is more comparable to simulated daily mean SSTs (Section 4.2.1).

- The standard deviation of the surface temperature error in a reanalysis of shelf seas (assimilating only SST) was markedly lower in the period when SST CCI products were used, relative to the earlier period when Pathfinder data were assimilated (Section 4.2.2).

- The use of the newly developed SST CCI uncertainties gave a significant reduction in the RMS errors of the shelf seas reanalysis when compared to in situ observations (Section 4.2.2).

- Different SST datasets were used to force the same global atmosphere-land general circulation model at low (~130 km) and high (~25 km) resolutions. Several differences in the simulated mean state seem to be influenced by differences between the SST datasets, including the Indian monsoon rainfall and surface temperature differences over North America. Whether these impacts are due to the superior resolution of the SST CCI analysis, or due to other differences between data sets (e,g, in their relative biases, of order tenths of a degree C), is not clear but will be explored further in Phase II of the ESA CCI. Tropical cyclone climatologies are also affected, particularly in the Eastern and Western Pacific regions (Section 4.2.3).

- It is important to have a range of forcing datasets, e.g. through provision of an ensemble, as proposed as an optional activity for SST CCI Phase II, so that model biases can be put into context compared to the model response to different forcing datasets (Section 4.2.3).

- The SST CCI analysis is a significant improvement over the pre-January 2013 Operational OSTIA SST dataset, both in terms of accuracy and spatial resolution (Section 4.2.4).
- The SST CCI AVHRR product is better in the Arctic than Pathfinder 5.1, 2006-2008. However, the SST CCI AVHRR product showed cold summer bias relative to reference data (seasonally varying, Section 4.2.5).
- An Arctic analysis showed improvements (having lower RMS difference from reference data) when SST CCI data were used for 2006-2008 instead of precursor data (Section 4.2.5).
- Inter-annual variability in and EOF patterns calculated from the SST CCI analysis are geophysically convincing at ~1 degree resolution, more so than in-situ-only aggregated data (Section 4.2.6).
- The Reynolds et al (2007) Daily OI exhibits a discontinuity (to relative warmth) at the switchover between use of AVHRR only and inclusion of AMSR-E data (Section 4.2.7).
- AMIP, Daily OI and SST CCI analysis are all very similar in terms of their representation of anomaly variability and teleconnections (Section 4.2.7).
- AMIP, Daily OI and SST CCI analysis exhibit generally consistent relationships with precipitation and total cloud amount. However, SST CCI analysis seems more strongly related to precipitation and cloud in some locations. These differences could impact conclusions drawn by studies on small regional scales (Section 4.2.7).

Consistency between SST and OC CCI products (Section 5):

- SST and OC CCI products contain information on SST and chl-a concentration off the south coast of Mexico that appears to be largely consistent on daily and 4km scales.
- SST and OC CCI products, when analysed together, provide a credible source of information for the analysis of coincident SST and chl-a fronts.
- Combining sensors in the OC and SST CCI products provides sufficient coverage of the area south of Mexico to perform daily analyses of frontal positions and to produce summary statistics over a year (2003).
- There may be merit in providing lower quality data alongside best quality data, as it appears to contain useful information for the analysis of ocean fronts.
- Coincident SST and chl-a fronts are a common feature of the ocean south of Mexico and are strongly organised in certain locations, although the exact locations appear to vary from year to year.
- Coordinated analysis of strongly connected variables can highlight potential deficiencies in aspects of the processing of either variable.

Feedback on ease of use of the products (Section 6):

- The SST CCI product has been generally easy to use as the output files are CF compliant, (following the GHRSST GDS specification) making them easy to read into to IDL and python.
- The Climate Research Group found that the uncertainties were easy to use

- The processing of the SST CCI products was more straight-forward than the corresponding Pathfinder v5.2 products.
- The instructions on how to download the data, as well as the documentation are understandable and relatively easy to follow. However, the Product User Guide should be made available in the data directory.
- However, data download speeds have been reported to be a problem. This issue will be explored when developing updated User Requirements for Phase II. The data are provided by a dedicated data centre, so this is not likely to be the source of the problem.

Further recommendations for the future by the trail blazer users (Section 6):

- Collaboration between modellers and the SST CCI team on the use of uncertainty estimates to improve climate model evaluation would be useful. [A workshop dedicated to the discussion of the use of uncertainty information is planned for SST CCI Phase II.]
- It would be useful to understand the discrepancy in the SST during the 1997-1998 El Niño between the SST CCI analysis and the Reynolds et al (2007) Daily OI and MyOcean OSTIA reanalysis v1. [In general in the analyses in this document, Pathfinder data and the analyses that incorporate them tend to be cooler than other comparison data. The relative coolness of the OSTIA v1 and Daily OI in the 1997/8 El Niño is likely in part due to the general bias in the Pathfinder data, although this bias seen during the 1997/8 El Niño is a larger relative bias than generally expected.]
- For future applications, it would be useful to also have the observations supplied with sufficient information to reverse the adjustment to a constant reference time-of-day. (SST CCI AVHRR and ATSR products currently have skin SST at satellite time and SST_{20cm} at a reference time, but not SST_{20cm} at the satellite time.)
- Future reanalyses would benefit from the SST CCI AVHRR being extended back in time to 1982 to reduce the need to use Pathfinder products and ensure a more consistent use of observations.
- Climate modellers are interested in how the diurnal cycle of SST may affect simulated variability in the climate model and would like to be able to provide updated SST information to atmospheric models every 3 hours. Over land, it is known that the diurnal cycle is poorly simulated, mainly due to the representation of convection. This may be important for the Madden-Julian Oscillation, and possibly for Indian monsoon active/break cycles.
- No quality 5 observations with fill values for SST should be in the data set. [This was a bug in the pre-release data.]
- A list of known gaps in the data could be useful.
- In the persistent cloudy regions of the high latitudes, an SST CCI product based upon Passive Microwave observations would be very valuable.
- The option to download a lower resolution version of the data set, or to download the data of sub-regions would be helpful as global fields are time-consuming to manipulate.
- It would be very helpful if both day and night (ascending/descending) values would be available in the SST CCI analysis.

- Estimate of the long-term correlated/uncorrelated uncertainties would be useful in order to estimate the uncertainty in the trend.
- A regridding tool which runs in real time would be helpful.

3. ASSESSMENT OF TRENDS, VARIABILITY AND STABILITY IN SST CCI PRODUCTS AND COMPARISON TO OTHER PRODUCTS

This section is split into two sub-sections. Section 3.1 assesses the trends and variability in the SST CCI products and compares them to other SST products to determine to what extent the new SST CCI products are credible Climate Data Records. Section 3.2 assesses the stability of the SST CCI products and that of other products in the GHRSST Multi Product Ensemble to determine to what extent they meet the GCOS requirement on stability.

3.1 Trends and variability in SST CCI and other products

3.1.1 INTRODUCTION

In order to assess the multi-annual and decadal behaviour of the long-term SST CCI products, comparisons are made to existing lower resolution SST data sets. These data sets include those used by the IPCC in their 5th Assessment Report as well as data sets used in other high profile monitoring reports such as the Bulletin of the American Meteorological Society's annual State of the Climate Report. Differences between the SST CCI products and the comparison datasets are highlighted.

3.1.2 DATA SETS

The three ESA SST CCI LT version 1.0 products assessed are:

- ATSR. SSTs from ATSR instruments in L3U format at 0.05° latitude by 0.05° longitude resolution covering 1991 – 2010. (Hereafter, SST CCI ATSR.)
- Analysis. Satellite-only SST-depth L4 daily analysis created by OSTIA system from SST CCI ATSR and SST CCI AVHRR products at 0.05° latitude by 0.05° longitude resolution covering 1991 – 2010. (Hereafter, SST CCI analysis.)
- AVHRR. SSTs from AVHRR instruments in L2P format at Global Area Coverage (GAC) resolution covering 1991 – 2010. (Hereafter, SST CCI AVHRR.)

These are compared to the following data sets over the period 1991-2010.

3.1.2.1 Gridded reference data

This data set comprises quality-improved in situ observations from ships and buoys from Atkinson et al. (2013, RD.326) with near-surface observations from Argo profiling floats from EN4 [Good et al (2013), RD.341]. The ship and buoy data are a subset of those used to create HadSST3 (Section 3.1.2.2). Data from these two sources were averaged onto a regular 5 degree latitude by 5 degree longitude grid at monthly resolution using the method of Rayner et al. (2006, RD.72). The data were not bias adjusted.

This data set will be referred to as the "gridded reference data set" since it comprises information from the reference data set utilised in the PVIR [RD.330, their Section 4.2.4, information reproduced below] together with measurements from ships that have undergone additional quality checking.

The ship and buoy data are a blend of observations taken from ICOADS 2.5 (Woodruff et al., 2011; RD.332) and Met Office Hadley Centre QC flags. The QC flags provided have been produced by the HadISST2 QC system. The general QC procedures are described

in Rayner et al. (2006; RD.72) and the high-resolution background climatology and land-sea mask used by this system is described in Rayner et al. (2013; RD.334). This system carries out the following suite of checks: (i) observations are checked for a meaningful location, date and time and that they are not surrounded on all sides by land, (ii) each platform with an individual callsign is tracked to verify its reported position, speed and direction (those without a callsign or with a generic callsign, e.g. SHIP, are passed unchecked), (iii) each SST observation is checked that it is above the freezing point of seawater and within ± 8°C of the 1961-1990 background climatology interpolated to that day, (iv) each SST observations has a "buddy check" applied which compares the value of an individual SST anomaly to the mean anomaly from neighbouring observations; individual observations differing too much from their neighbours are flagged as bad. The HadISST2 QC flags have been supplemented as follows:

1. Drifting buoy SST observations from ICOADS deck 715 have been blacklisted as investigation suggests they are of variable quality.

2. Drifting buoy and ship SST observations have an additional QC flag set which follows the procedures described in Atkinson et al. (2013; RD.326). This flag is generated by tracking the quality of observations made by individual drifting buoys and ships over time using the Met Office Operational Sea surface Temperature and sea Ice Analysis (OSTIA) as a reference (a globally complete satellite based analysis). It differs from the SST checks described above in that observation quality is tracked over time to detect biases/instrument failures etc., rather than assessing observations individually. Drifting buoys observations are flagged where they are deemed to be too biased or too noisy, or a buoy is deemed to be out of water having run aground or been picked up. Ship observations are flagged when observations from a particular ship (identified by its callsign) are deemed unreliable (i.e. if a ship callsign is blacklisted all observations from this ship are flagged). In general, ship observations are of variable quality and this flag is intended to reduce ship observations to a higher quality subset; the extra QC excludes between 50 and 60% of ship observations. The development of this extra QC step was funded by the FP7 ERA-CLIM project. This is the first time the information has been used in this way.

3.1.2.2 HadSST3

HadSST3 [Kennedy et al. 2011b, RD.210; Kennedy et al. 2011c, RD.211] is an in situ-only data set. Individual observations from International Comprehensive Ocean-Atmosphere (ICOADS) release 2.5 [Woodruff et al. 2011, RD.332] are averaged onto a 5 degree latitude by 5 degree longitude grid at monthly resolution using the method of Rayner et al. (2006, RD.72). The ship and buoy data used in the gridded reference data set (see above Section 3.1.2.1) are a subset of the ship and buoy data in HadSST3. Adjustments are applied to the gridded data to account for the effect of systematic errors associated with changes in measurement methods over time. In the period examined here, the principal change is the switch from mostly ship-based observations in the 1980s to a mixture of ship and buoy observations in the late 2000s. Because the exact size of the systematic errors is not known, the data set is presented as an ensemble of 100 different versions (realisations), which are indicative of the uncertainty in the adjustments applied. Parameters used in the statistical modelling of the biases are varied within their likely ranges to produce an ensemble of bias adjustments. These are then used to create the ensemble of adjusted SST anomaly fields. In addition to the uncertainty associated with the adjustments, there are uncertainties arising from other kinds of measurement error. Estimates have been made of these and they are also used.

3.1.2.3 AVHRR Pathfinder

The Advanced Very High Resolution Radiometer instruments measure radiances in the infra red part of the spectrum. AVHRRs have flown on US National Oceanographic and

Atmospheric Administration (NOAA) polar orbiting satellites since the early 1980s. The measured radiances can been used to estimate SSTs. This is usually done using coefficients estimated from a regression against drifting buoy or ship data that also vicariously adjusts the AVHRR instrument calibration. Pathfinder v5.2 [Casey et al. 2010, RD.216] is a consistent reprocessing of the AVHRR data from 1981 to 2007. The version used in this analysis was presented on a 1/24 degree equidistant cylindrical grid.

3.1.2.4 HadISST

Met Office Hadley Centre sea Ice and SST (HadISST) is a globally complete analysis of sea-surface temperature and sea-ice concentrations. It is based on in situ and satellite (AVHRR) measurements of SST. Gaps in the data coverage are filled using a statistical technique known as Reduced Space Optimal Interpolation (Kaplan et al. 1997, RD.342). In areas of the ocean where there is an estimated non-zero sea ice concentration, the SST is inferred from the sea ice concentration. HadISST1.1 [Rayner et al. 2003, RD.74] is presented on a 1 degree latitude by 1 degree longitude monthly grid, although the anomaly analysis is performed on a 2 degree latitude by 2 degree longitude grid (and then added to a 1 degree latitude by longitude climatology) in the modern period (1949 onwards) including the whole period covered in this report.

3.1.2.5 ERSSTv3b

Extended Reconstruction SST (ERSSTv3b, Smith et al. 2008, RD.79) is another globally complete SST data set. It is based on in situ measurements of SST and uses a combination of Empirical Orthogonal Teleconnections and a low-frequency smoothing to reconstruct SSTs globally. ERSSTv3b is presented on a 2 degree latitude by 2 degree longitude grid. The grid is offset slightly such that the equator passes through the centre of one of the grid boxes rather than forming the boundary of one of the grid boxes.

3.1.2.6 Kaplan

The Kaplan data set [Kaplan et al. 1998, RD.81] is a near-globally complete SST data set based on an old version of the Met Office Historical SST data set (MOHSST5). The data are reconstructed using Reduced Space Optimal Smoothing (RSOS). RSOS is similar to RSOI (used in HadISST, Section 3.1.2.4) and differs in that it uses information about the temporal evolution of SST to reduce the uncertainty in the estimated fields. MOHSST5 has not been update since the late 1990s, so more recent updates have been calculated by processing NOAA SST fields using the same RSOS technique. The data are presented on a 5 degree latitude by 5 degree longitude monthly grid. The data set is not globally complete. Regions of the southern oceans and Arctic ocean – regions where there were too few data to estimate reliable Empirical Orthogonal Functions (EOFs) – are not reconstructed.

3.1.2.7 COBE SST

The Centennial in situ Observation-Based Estimates of the variability of SSTs (COBE SST, [Ishii et al. 2005, RD.85]) is a globally complete SST data set based on in situ data. The data are reconstructed using optimal interpolation. In the absence of data the reconstruction relaxes to the climatological average. In marginal ice zones, SSTs are inferred from sea ice concentrations. The data are presented on a 1 degree latitude by 1 degree longitude monthly grid.

3.1.2.8 NOCS Surface Flux data set v2.0

The National Oceanography Centre, Southampton Surface Flux data set [Berry et al. 2011, RD.103] is based on ship data from ICOADS release 2.1. The daily data are reconstructed using optimal interpolation. In the absence of data the reconstruction

relaxes to the climatological average. The data are presented on a 1 degree latitude by 1 degree longitude daily or monthly grid. The monthly grids were used herein.

3.1.2.9 Karspeck et al. 2011

Karspeck et al. [2011, RD.233] present an analysis of north Atlantic sea surface temperatures. Their analysis is based on a large scale reconstruction of SST from Kaplan et al. [1998, RD.81] which is then augmented using a mid-scale optimal interpolation. The data set is based largely on in situ data, although satellite data are used to estimate some of the parameters of their reconstruction scheme. The data are presented on a 1 degree latitude by 1 degree longitude monthly grid for the north Atlantic.

3.1.2.10 OI.v2

The Reynolds et al. [2002, RD.212] weekly and monthly OI.v2 is based on in situ and satellite data. The satellite SST retrievals come from the AVHRR series of instruments and biases in the data are adjusted to more closely match the in situ data. The resulting fields are interpolated using optimal interpolation and the data are presented on a 1 degree latitude by 1 degree longitude weekly and monthly grid. The monthly grids are used herein.

3.1.2.11 MyOcean OSTIA reanalysis

The Operational Sea Surface Temperature and Sea Ice Analysis (OSTIA) reanalysis v1.0 [Roberts-Jones et al. 2012, RD.239] from 1985 to 2007 is based on reprocessed satellite and in situ measurements. AVHRR Pathfinder SSTs and retrievals from the ATSR instruments are combined with data from ICOADS release 2.1 using a multiscale optimal-interpolation scheme. The data are presented on a 0.05 degree latitude by 0.05 degree longitude daily grid.

3.1.2.12 Daily OI

Two varieties of the Daily optimum interpolation (OI) version 2 [Reynolds et al. 2007, RD.76] data set were used. One incorporates AVHRR and in situ SSTs and runs from 1981 to present. The other incorporates AVHRR, Advanced Microwave Scanning Radiometer (AMSR-E) and in situ SSTs and runs from 2002 to 2011 covering the period of operation of the AMSR-E instrument. Both varieties use an optimal interpolation scheme to reconstruct missing values and are presented on a 0.25 degree latitude by 0.25 degree longitude daily grid.

3.1.3 METHODS

The SST CCI products and the comparison data sets are presented on a range of different grids and also, in those cases where the data are presented as anomalies, relative to different climatological averages. In order to make a direct comparison, the data were first converted into anomalies relative to the MyOcean OSTIA reanalysis climatology for the period 1985-2007. The climatology was averaged from 0.05 degree latitude by 0.05 degree longitude daily to have the same resolution as the data set being processed. Secondly a common land/sea mask was applied to the data. Again this was based on the MyOcean OSTIA reanalysis climatology. Data were not masked to exclude sea-ice covered areas. Prior to some of the comparisons, the data sets were regridded to a common grid. Where this is done it is noted in the text.

The following sub-sections detail the metrics used to compare the data sets.

3.1.3.1 Linear trends

Time series of area-average temperature anomalies were calculated from each of the data sets for the regions shown in Figure 3-1 and listed in Table 3-1. Area averages were calculated as a weighted average of all non-missing grid box values within the area. The weights were proportional to the area of ocean within the grid box. In coastal grid boxes which were not entirely covered by ocean, the area of ocean was estimated using the OSTIA reanalysis climatology. Grid boxes in the original 0.05 degree climatology which had an assigned SST were assumed to be 100% ocean. Areas without an assigned SST were assumed to be 0% ocean. The 0.05 degree map of land and sea was than averaged to the target resolution.

Area average anomalies were calculated for each data set with its native coverage and also after the coverage had been reduced to that of HadSST3. To reduce the coverage to that of HadSST3 the anomalies for each dataset were first averaged to 5-degree monthly resolution.

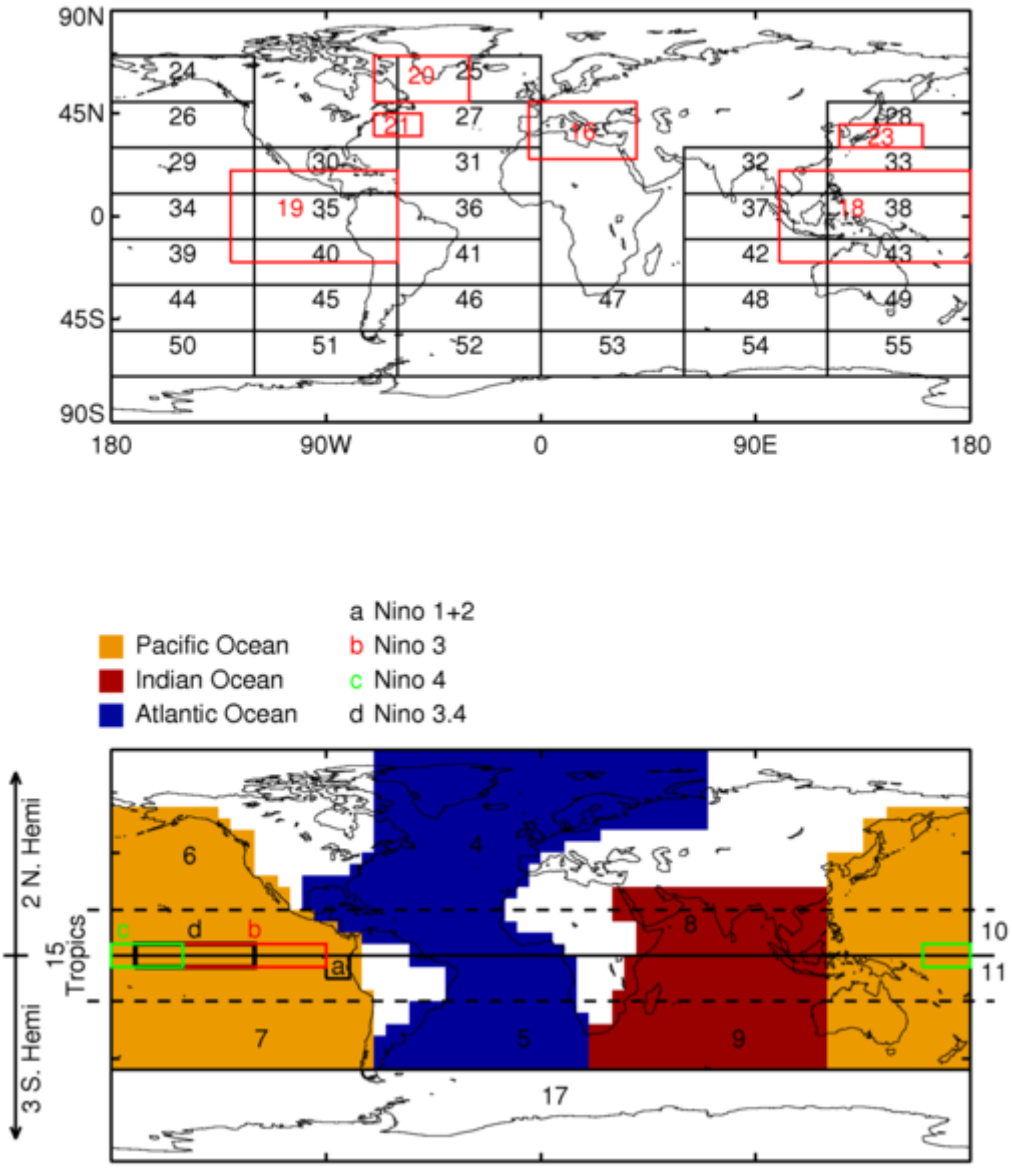


Figure 3-1. Maps showing regions used in the analysis of trends and indices. The regions are also described in Table 3-1.

Table 3-1. List of regions used for the analysis of trends. The regions are shown in Figure 3-1.

Region Name	Region Number	Region Name	Region Number
Globe	1	Area 30°-50°N, 120°-180°E	28
Northern Hemisphere	2	Area 10°-30°N, 180°-120°W	29
Southern Hemisphere	3	Area 10°-30°N, 120°-60°W	30
North Atlantic Ocean	4	Area 10°-30°N, 60°-0°W	31
South Atlantic to 50°S5	5	Area 10°-30°N, 60°-120°E	32
North Pacific Ocean	6	Area 10°-30°N, 120°-180°E	33
South Pacific to 50°S	7	Area 10°N-10°S, 180°-120°W	34
North Indian Ocean	8	Area 10°N-10°S, 120°-60°W	35
South Indian Ocean to 50°S	9	Area 10°N-10°S, 60°-0°W	36
Northern Tropics	10	Area 10°N-10°S, 60°-120°E	37
Southern Tropics	11	Area 10°N-10°S, 120°-180°E	38
Atlantic Ocean to 50°S	12	Area 10°-30°S, 180°-120°W	39
Pacific Ocean to 50°S	13	Area 10°-30°S, 120°-60°W	40
Indian Ocean to 50°S	14	Area 10°-30°S, 60°-0°W	41
Tropics (20°N-20°S)	15	Area 10°-30°S, 60°-120°E	42
Mediterranean	16	Area 10°-30°S, 120°-180°E	43
Southern Ocean, 50°S Southwards	17	Area 30°-50°S, 180°-120°W	44
Western Tropical Pacific	18	Area 30°-50°S, 120°-60°W	45
Eastern Tropical Pacific	19	Area 30°-50°S, 60°-0°W	46
Greenland 50°-70°N, 30°-70°W	20	Area 30°-50°S, 0°-60°E	47
Gulfstream 35°-45°N 50°-70°W	21	Area 30°-50°S, 60°-120°E	48
Southern Hemisphere and Northern Indian Ocean minus rest of NH	22	Area 30°-50°S, 120°-180°E	49
Kuroshio 30°-40°N, 125°-160°E	23	Area 50°-70°S, 180°-120°W	50
Area 50°-70°N, 180-120W	24	Area 50°-70°S, 120°-60°W	51
Area 50°-70°N, 60°-0°W	25	Area 50°-70°S, 60°-0°W	52
Area 30°-50°N, 180°-120°W	26	Area 50°-70°S, 0°-60°E	53
Area 30°-50°N, 60°-0°W	27	Area 50°-70°S, 60°-120°E	54
		Area 50°-70°S, 120°-180°E	55

Linear trends in the area averages were calculated from all non-missing monthly values using the ordinary least squares method [as implemented in the POLY_FIT routine in IDL v7.1]. This was the version used in the Figures. A resistant method for estimating the trends – median of pairwise slopes [Lanzante 1996, RD.343] - was also used (not shown) to check that outliers did not have a strong effect on the results. Trends were calculated

over two periods 1992-2010 and 1997-2010. The former period covers all complete years in the SST CCI LT products. The latter period excludes the problematic ATSR1/Pinatubo period.

Results are shown in Figures 3-3, 3-4, 3-5 and 3-6 and discussed in Section 3.1.4.

3.1.3.2 Indices

In addition to the time series for the regions described in Figure 3-1 and Table 3-1, indices for certain standard modes of variability were also calculated. These were:

1. Niño 1+2 [0°-10°S, 90°-80°W]
2. Niño 3 [5°N-5°S, 150°W-90°W]
3. Niño 4 [5°N-5°S, 160°E-150°W]
4. Niño 3.4 [5°N-5°S, 170°W-120°W]
5. Dipole Mode Index (DMI) calculated as the difference between the area-average SST anomalies for the regions [50°-70°E, 10°S-10°N] and [90°-110°E, 10°S-10°N]
6. Tropical Atlantic Meridional SST gradient (TAMG) calculated as the difference between the area-average SST anomalies for the regions [60°W-African Coast, 5°-28°N] and [60°W-20°E, 20°S-5°N]

These six indices are all based on area-averages which were calculated in the same way as detailed in Section 3.1.3.1. Linear trends were also calculated in the same way as in Section 3.1.3.1. Results are shown in Figures 3-3, 3-4, 3-5, 3-6, 3-7 and 3-8 and discussed in Section 3.1.4.

3.1.3.3 Multi-annual and decadal averages

Multi-annual and decadal averages were calculated for the periods 1991-1995, 1996-2000, 2001-2005, 2006-2010, 1991-2000 and 2001-2010. An average was calculated for each period for a grid box when at least 30% of monthly values in that period were non-missing. In order to highlight multi-annual variability, five-year averages for each data set were also expressed as differences from the 1991-1995 average for that data set. Results are shown in Figure 3-9 and 3-10 and discussed in Section 3.1.4.

10 year average differences between data sets were also calculated. In order to do so, all data sets were averaged onto a common 5-degree monthly grid and differences were calculated at those locations where both data sets had data. An average was calculated when at least 30% of monthly values in the difference maps were non-missing. Results are shown in Figure 3-11 and discussed in Section 3.1.4.

3.1.3.4 Autocorrelations

Lagged correlations were calculated within each data set. In order to make a direct comparison, all data sets were re-gridded to 5 degree monthly resolution. Lag correlations at lags of 1, 2, 3 and 4 months were calculated in all grid boxes for which at least 30% of monthly values were non-missing. Results are shown in Figure 3-12 and discussed in Section 3.1.4.

3.1.4 DISCUSSION OF RESULTS

3.1.4.1 Linear trends and Indices

Although the SST CCI long term (LT) products are long-term data sets in one sense, the period covered is relatively short when considered in the context of data sets based on in

situ data. One consequence of having only 20 years of data is that the long-term warming during this period is comparable to the year-to-year variability, which is dominated by the period of high El Niño/La Niña activity around 1997-2000 and again from 2007-2010. Therefore, linear trends are not a good 'model' for temperature change over this period. Nonetheless they do highlight differences such as 'drifts' (or relative instabilities) between the data sets.

The upper panel of Figure 3-2 shows the global-average SST anomaly for each of the SST CCI products and for the comparison data sets. The first thing to note is that there is apparent disagreement between the comparison data sets. There are three particular outliers. Two are warmer – NOCS and Karspeck – and one is generally cooler – Pathfinder. Inspection of the NOCS fields shows large areas of erroneous SST anomalies in data sparse regions. The Karspeck analysis is only for the North Atlantic and so its average is not globally representative. The Pathfinder data set has a persistent cold bias relative to the other data sets. Removing these three outlying datasets gives a much neater comparison (Figure 3-2 middle panel). The MyOcean OSTIA reanalysis and Daily OI data sets are also cool relative to the other comparison data sets, but the disagreement is not as marked and the variability is similar.

The three CCI products show very similar year-to-year variability at a global average scale to that seen in the comparison data sets; post-1995, after the period of notable instrument-related biases in both the AVHRR and ATSR time series. However, there is an offset of order 0.2 K between them, with the ATSR product being warmer (this is discussed further below). The SST CCI AVHRR data are generally cooler than the comparison data sets, although the relative cool bias is much less marked than in the Pathfinder data set, which is also based on AVHRR data. Early in the record – between 1991 and 1994 – there are three pronounced spikes in the SST CCI AVHRR product that affect four months of the record. This has been traced to periods of anomalous calibration in a particular AVHRR sensor, not used in the Pathfinder data set. The warm spikes contrast with the general coolness of the SST CCI AVHRR products during this period which, until around 1994, may reflect residual effects of the post-Pinatubo stratospheric aerosol (which also cooled the climate at this time). There is also a pronounced cold spike in the SST CCI ATSR ATS1 product in 1996, which has an opposite sign in the northern (-) and southern hemispheres (+). This anomaly in ATSR calibration is not linked to the known calibration and detector temperature issues affecting ATSR, and this requires further study. Within Phase II, we will explore whether or not ATSR-1 can be cross-referenced to a relatively stable AVHRR over this period.

Aside from anomalous spikes in the data, there are more subtle differences at a global scale between the SST CCI products and the comparison data sets, and between the SST CCI products themselves. The SST CCI ATSR and SST CCI analysis products warm more rapidly than the majority of the comparison data sets (Figure 3-3), but the SST CCI AVHRR products tend to warm less rapidly over the period 1992-2010 (this holds when the resistant estimator that is less sensitive to outliers is used to calculate the trend). This can be seen in Figure 3-2, with the SST CCI analysis and SST CCI ATSR products near the middle or lower end of the comparison data set distribution around 1991 and near the top or middle of the distribution around 2010. This difference persists even after the data sets are reduced to a common coverage (Figure 3-2 lower panel and Figure 3-4) suggesting that this is not simply related to changing coverage in the in situ data sets.

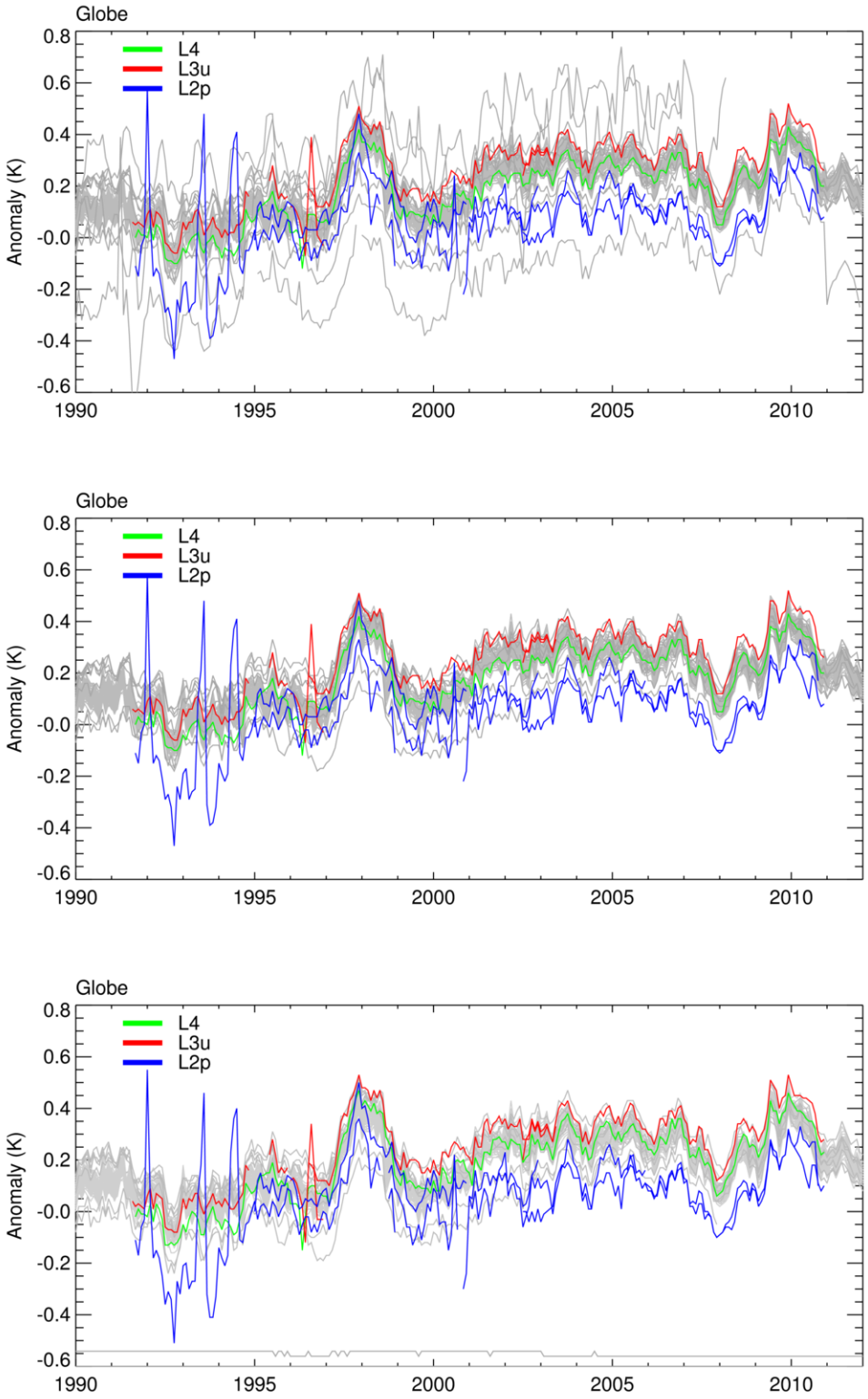


Figure 3-2. Top: Global average SST anomaly (K, relative to MyOcean OSTIA reanalysis climatology) for each of the comparison data sets (grey) and the SST CCI products [SST CCI ATSR (L3U): red, SST CCI analysis (L4): green, SST CCI AVHRR (L2P): blue]. Middle: as for (top) but excluding the NOCS, Karspeck and Pathfinder comparison data sets. Bottom: collocated comparison. For SST CCI ATSR and SST CCI AVHRR there is one line for each individual satellite.

Since both the SST CCI AVHRR and ATSR products represent the same SST depth and phase of the diurnal cycle (SST_{20cm} at 10.30h is used), their relative bias (of order 0.2 K) represents a difference outside the target for the project of biases <0.1 K. The size of this relative bias is unlikely to be explained by SST retrieval algorithm issues. All the SST retrievals were harmonised (ATSR to ATSR) or cross-referenced (AVHRR to ATSR) as part of the algorithm development and were independently tested, as reported in the Algorithm Selection Report [MacCallum and Merchant, 2012; RD.226]. For all the instruments, the combined (day plus night) global discrepancy relative to the independent drifting buoy data in the selection data set was within the range 0.0 to +0.1 K. The relative coolness of the SST CCI AVHRR product compared to the SST CCI ATSR product may arise from: (i) the aggregated effect of residual cloud contamination being more significant for the AVHRR record than for the ATSRs and/or (ii) a systematic effect in the diurnal cycle adjustment of the “afternoon” AVHRR instruments to the standard 10.30h reference time. Improving AVHRR cloud detection is a major focus in Phase II of the SST CCI. From 1997, the trends in the SST CCI ATSR and SST CCI analysis products are closer to those of the comparison data sets, sitting within the typical range.

Figure 3-3 to 3-6 show trends for the SST CCI products and for the different comparison data sets for the various regions shown in Figure 3-1. Figure 3-3 shows trends for January 1992 to December 2010 for each of the data sets. The comparison data sets are shown as grey lozenges with the three “outlier” data sets shown as small grey pips. The coloured lozenges show the trends from the SST CCI products. The blue shaded area is an estimate of the measurement and sampling uncertainty as estimated from the HadSST3 data. This indicates the likely spread in trends attributable to weakly-correlated or uncorrelated measurement errors. The spread is generally small for large area averages (e.g. global and hemispheric averages) or for well-sampled regions such as the North Atlantic. Because some of the differences between the SST CCI products and the comparison data sets (and various combinations of the same) might arise from differences in large-scale sampling, Figure 3-4 shows the trends over the same period after the coverage has been reduced to that of the *in situ*-only HadSST3 data set.

Figure 3-5 shows trends for the shorter period 1997-2010. Figure 3-6 shows trends for the same period after the data sets have been reduced to a common coverage, as in Figure 3-4.The period 1997-2010 was chosen because it excludes the problematic ATSR-1 period which includes unusual spikes in the SST CCI ATSR and AVHRR products. Over this shorter period, the effect of year-to-year variability (and noise) on the trends can be seen, most clearly in the Niño 1+2 region where the effect of the 1998 El Niño leaves a strong negative trend over the period.

The divergence in trends between the SST CCI AVHRR and the SST CCI ATSR/SST CCI analysis products can also be seen in many of the sub regions shown in Figure 3-1 and is apparently more pronounced in the Southern Hemisphere than the in the Northern (Figure 3-3, Figure 3-4), although at the highest southern latitudes (50-70S) the SST CCI AVHRR warms relative to the SST CCI ATSR/SST CCI analysis.

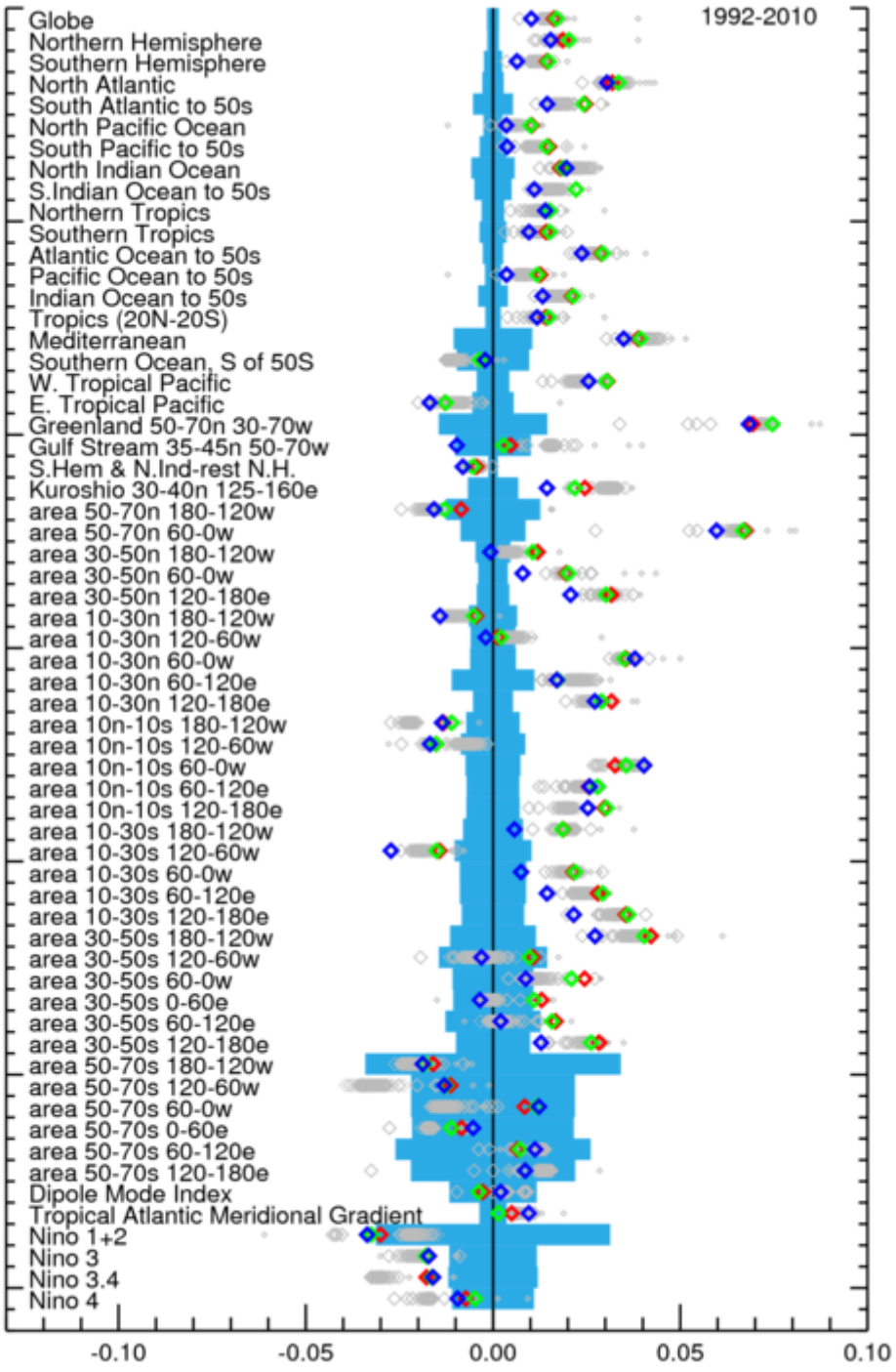


Figure 3-3. Linear trends (K/year) from January 1992 to December 2010 for each of the 61 regions and indices and each of the comparison data sets (grey) and the three CCI data sets: (red) SST CCI analysis, (green) SST CCI ATSR and (blue) SST CCI AVHRR. The comparison data sets shown by grey dots are Pathfinder, NOCS, Karspeck, OSTIA reanalysis and Daily OI using AMSR. The pale blue area is an estimate of the uncertainty in the trend arising from measurement and sampling errors in the HadSST3 data.

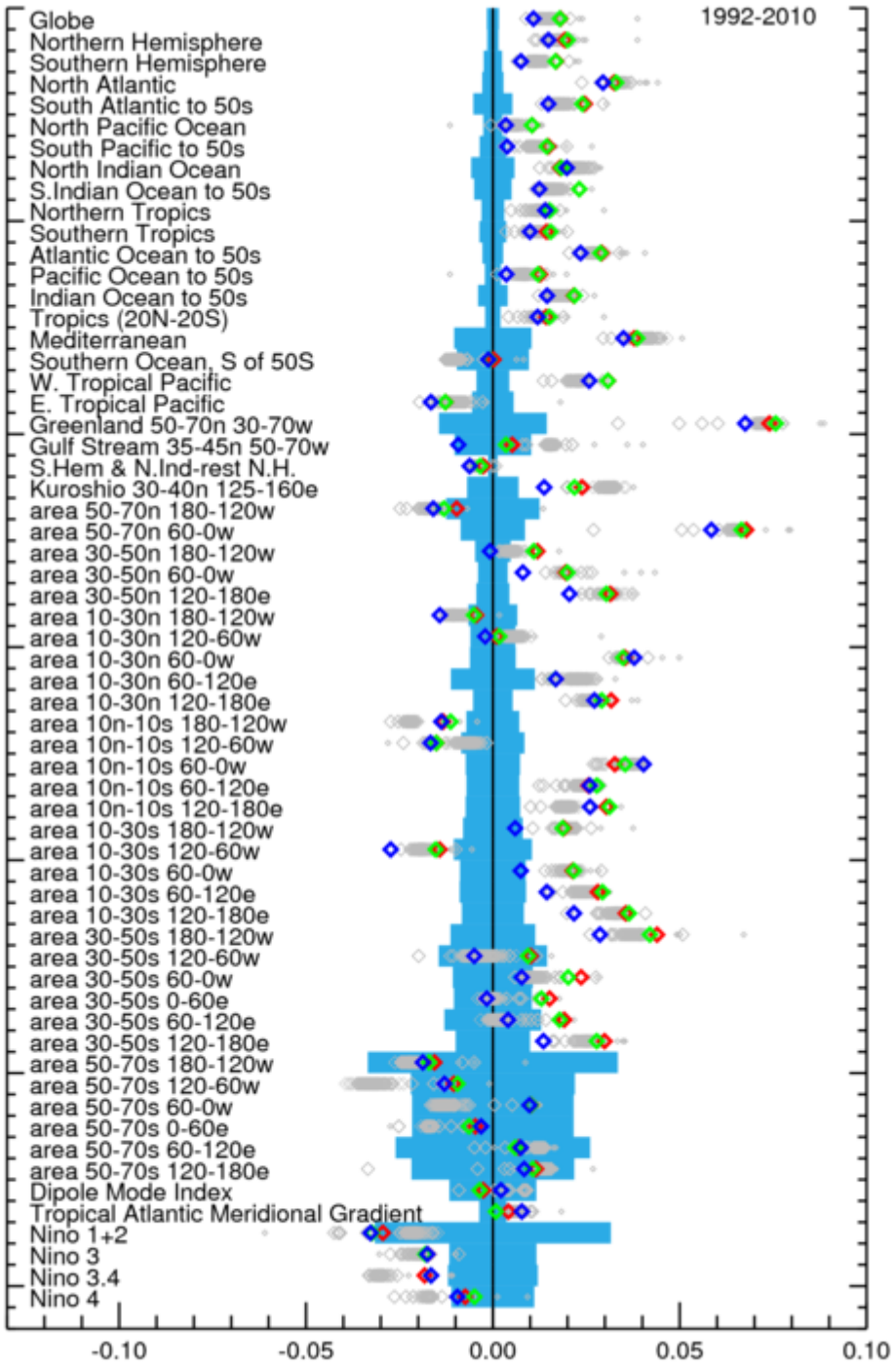


Figure 3-4. As for Figure 3-3 but each data set has been reduced to the coverage of HadSST3.

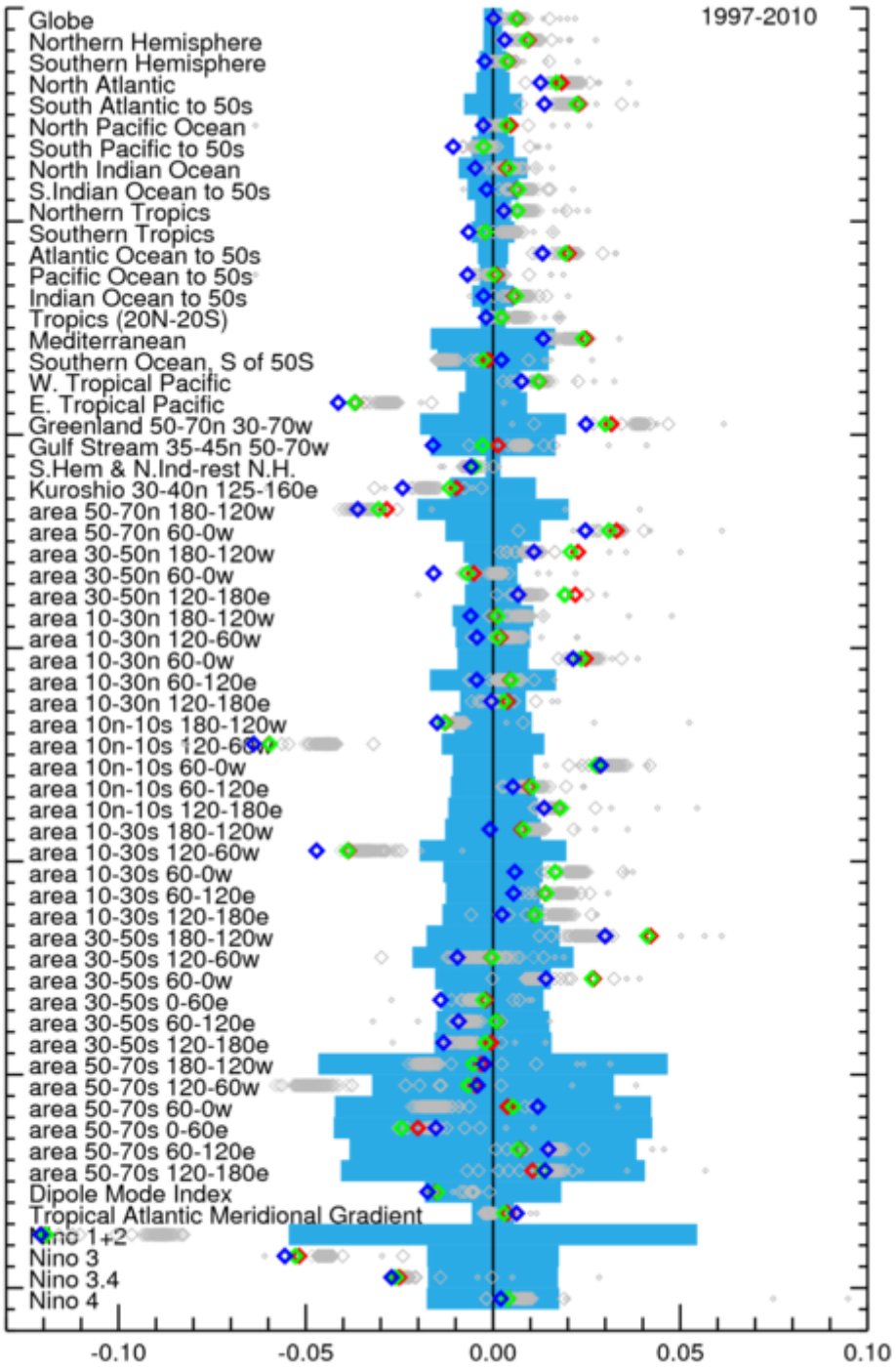


Figure 3-5. Linear trends (K/year) from January 1997 to December 2010 for each of the 61 regions and indices and each of the comparison data sets (grey) and the three CCI data sets: (red) SST CCI analysis, (green) SST CCI ATSR and (blue) SST CCI AVHRR. The comparison data sets shown by grey dots are Pathfinder, NOCS, Karspeck, OSTIA reanalysis and Daily OI using AMSR.

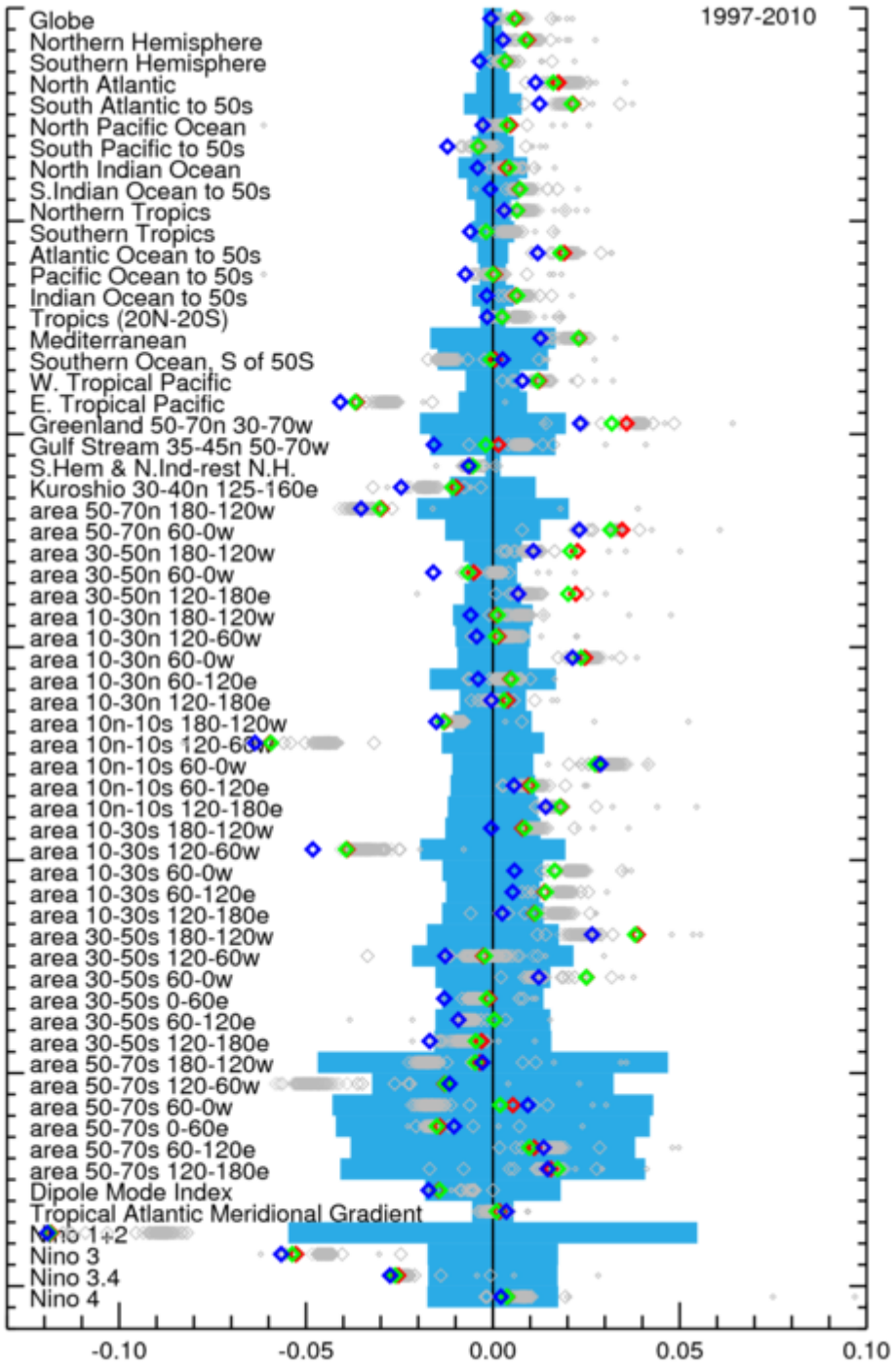


Figure 3-6. as for Figure 3-5 but each data set has been reduced to the coverage of HadSST3

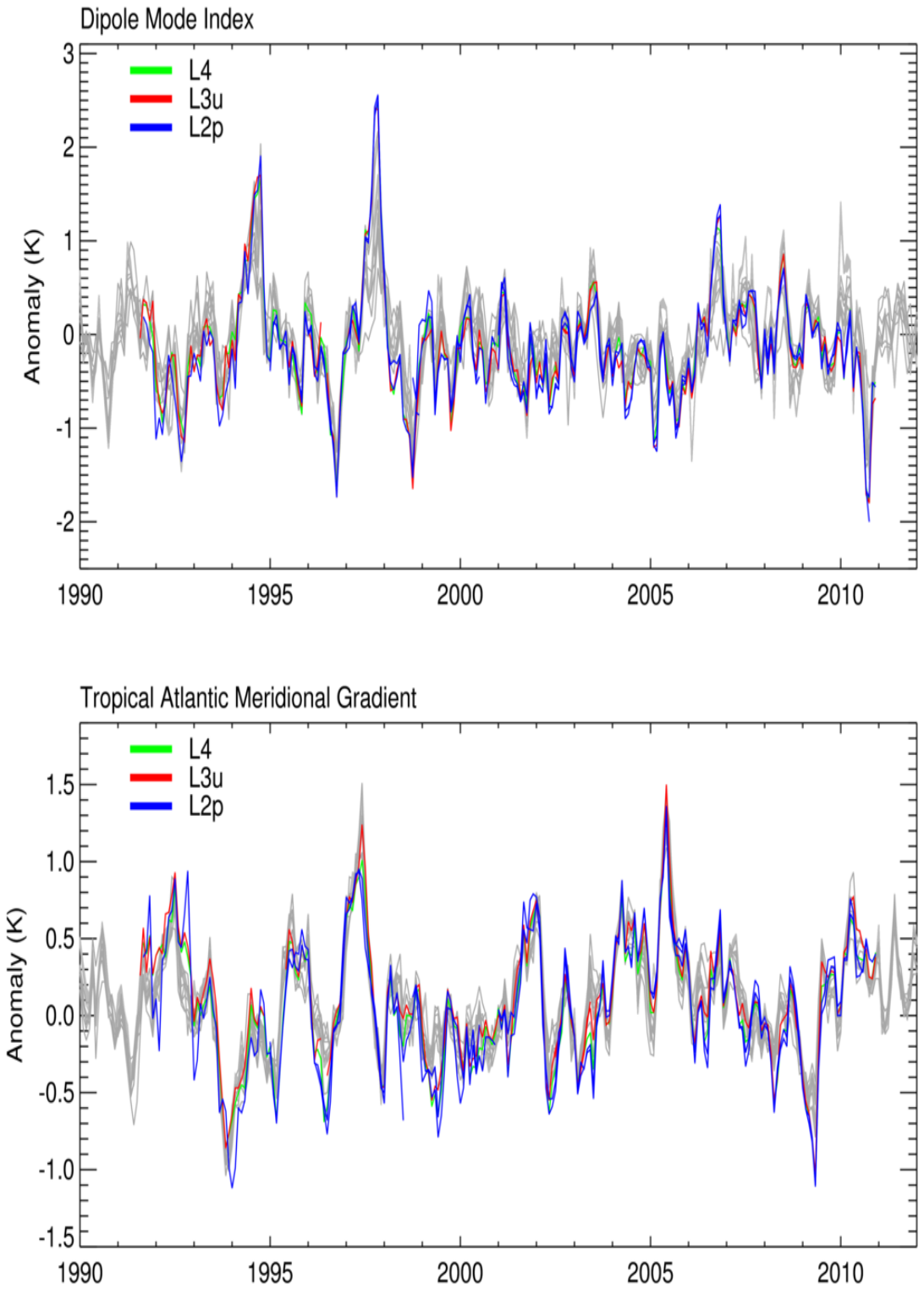


Figure 3-7. Dipole Mode Index (K, top) and Tropical Atlantic Meridional Gradient (K, bottom). SST CCI ATSR (L3U): red, SST CCI analysis (L4): green, SST CCI AVHRR (L2P): blue. Units are in K because the indices are calculated as the difference between two simple area averages of temperature.

The Dipole Mode Index is shown in Figure 3-7. As before, the year-to-year variability is largely similar in the SST CCI products to what is seen in the comparison data sets. Correlations between the SST CCI analysis product and the comparison data sets lie between 0.70 and 0.96. Correlations are highest when compared to the higher resolution data sets such as Daily OI and the OSTIA MyOcean reanalysis. One thing to note is that the spread of estimates is relatively narrower for the DMI because the variability is higher and because the index is calculated as the difference between two areas. Taking differences will tend to reduce the effect of systematic offsets between data sets. However, there are some interesting differences between the DMI calculated from the SST CCI products and the comparison data sets. The strong peak in 1997 is larger in the SST CCI products than in the comparison data sets. The same is true for smaller peaks seen in 1994 and 2006. In contrast a peak in late 2009/early 2010 that is seen in most of the comparison data sets is absent in the SST CCI products. In this case, the spike is most prominent in those data sets using in situ data and arises from unusually warm observations in the western pole of the DMI. It is most prominent in HadSST3. In contrast the peaks in late 1997, 1994 and 2006 are stronger in the SST CCI products. In this case the eastern pole of the DMI is much cooler in the SST CCI products. This cold anomaly, which is strongest along the coast of Sumatra and Java, is not well resolved in the comparison data sets, particularly those based solely on *in situ* data. This shows up in the trend comparisons for this index for the period 1997-2010 (Figure 3-5 and 3-6). The satellite data provide the best feature resolution in these areas. However, they also suffer from dust-related biases that the *in situ* data do not. It is difficult to categorically state that the SST CCI data have the best representation from this evidence alone.

Figure 3-7 also shows the Tropical Atlantic Meridional Gradient (TAMG). Once again, because the variability is high and the index is calculated by taking the difference between two area averages the relative spread of the data sets is narrower. There is some suggestion that the peak-to-peak variability of the SST CCI products is higher than for the comparison data sets. Note in particular that the negative peak in 2009 is stronger in the SST CCI products. As for the DMI, this appears to be due to a cold anomaly in the northern pole of the TAMG that was not as clearly resolved in the comparison data sets, but in addition, the strength of the peak is increased by a relatively small scale positive anomaly at the northern edge of the southern pole of the TAMG that is resolved in the SST CCI products. Interesting differences between the SST CCI products and the comparison data sets can be seen in 1996, 1999 and 2004 when the SST CCI products are consistently below the comparison data sets. In each case the SST CCI ATSR product is closest to the comparison series.

Figure 3-8 shows time series of SST anomalies for the Niño regions. All data sets resolve the strong year-to-year variability associated with ENSO, which is much larger than the spread between the data sets. Consequently, correlations between SST CCI analysis and the comparison data sets are higher than 0.91 for all Niño regions. Excluding the NOCS comparison data set raises the minimum correlation to 0.94 for Niño 1+2 and 0.97 for all other Niño regions. The SST CCI AVHRR product shows large inconsistencies prior to 1995 with a general cool offset and erroneous variability in the Niño 4 region. The relative cool bias in the SST CCI AVHRR SST CCI product is most obvious in the Niño 1+2 region in the eastern Tropical Pacific.

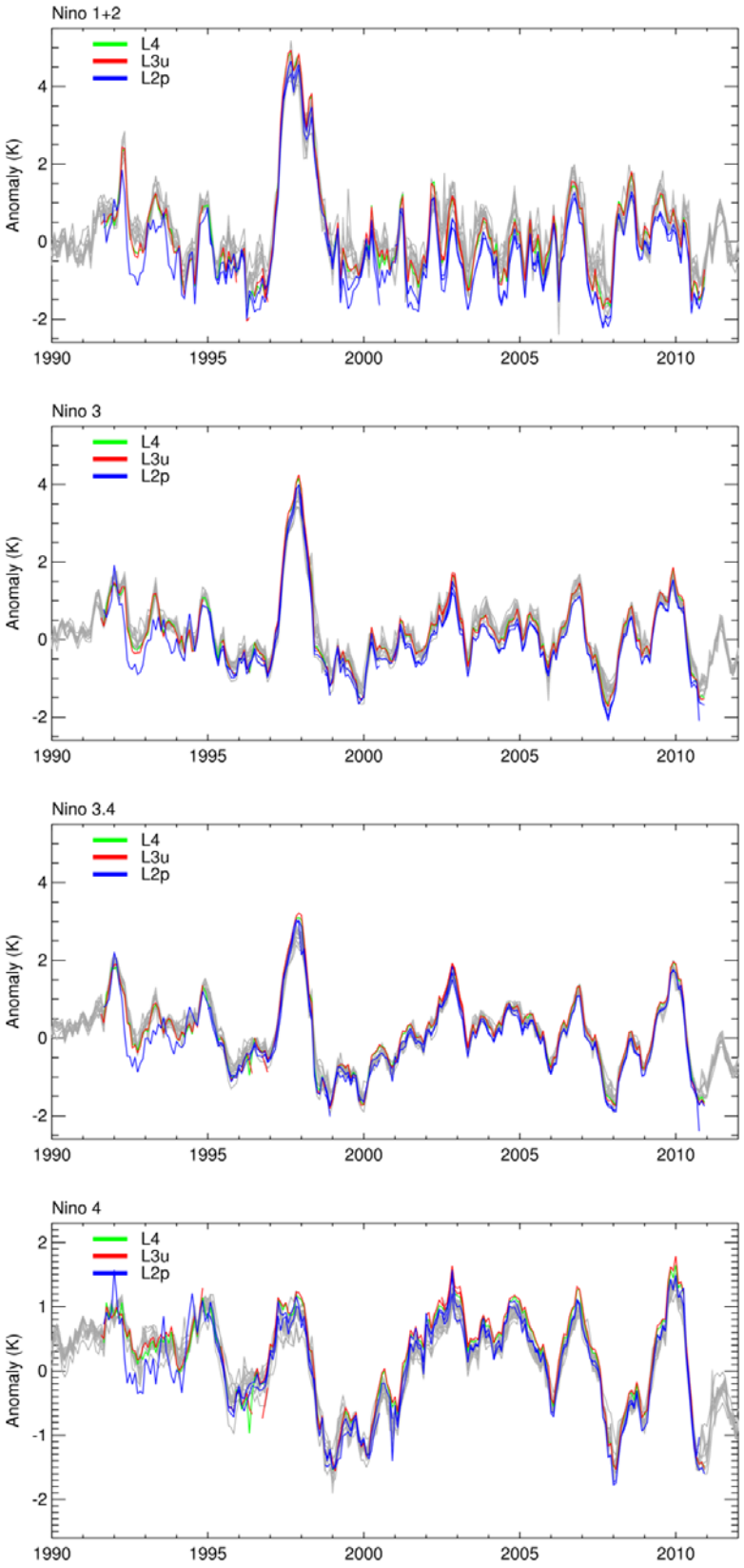
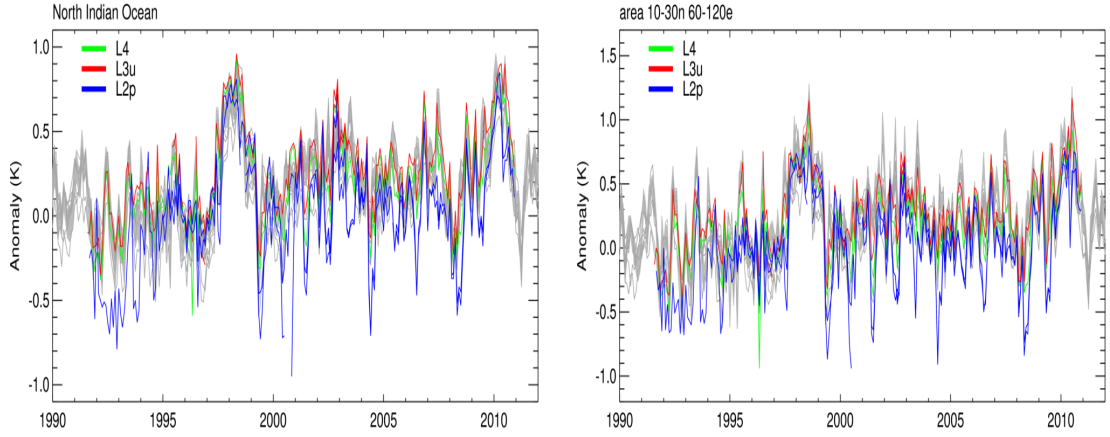
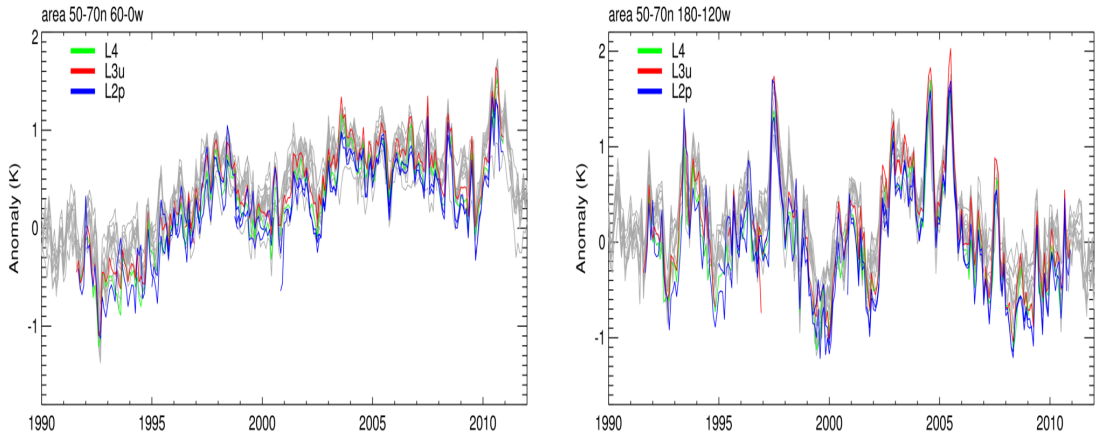


Figure 3-8. Niño indices. Area-average SST anomalies (K) for the four Niño regions described in Figure 3-1. SST CCI ATSR (L3U): red, SST CCI analysis (L4): green, SST CCI AVHRR (L2P): blue.

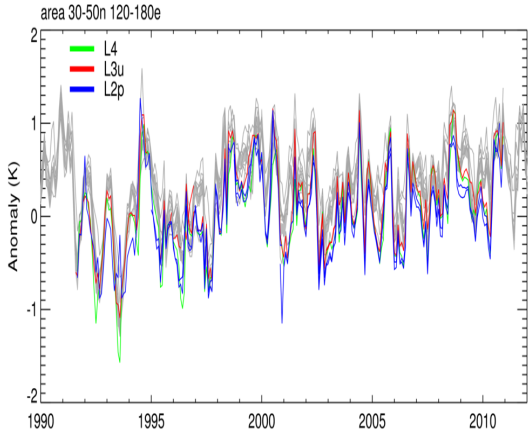
Interesting features were also evident in some of the smaller comparison regions. In region 8, the North Indian Ocean, the SST CCI AVHRR product has a spurious annual cycle, with periodic cool 'dips' (typically between May and July) relative to the other data sets. This can also be seen in region 32 (10-30N 60-120E, which is a subset of region 8). In the western tropical Pacific for the period 1992-2010, all three SST CCI products warm faster than the majority of comparison data sets. In regions 24 (50-70N, 180-120W) and 25 (50-70N, 60-0W), there is a higher peak-to-peak variability in the SST CCI data sets. Regions 24 and 25 contain areas with different characteristics. Region 24 is in the North Pacific, relatively well-observed by *in situ* sources at its southern edge, but poorly observed at its northern extent. The region also straddles the Bering Strait, which separates areas with quite different climates. Likewise, region 25 includes the Labrador Sea at its western extent, the Norwegian Sea at its north-eastern corner and the Irish Sea at its south-eastern corner. The *in situ* sampling and climatic character of these three regions are quite distinct. It is likely that the satellite sampling is more uniform than *in situ* sampling in these regions, leading to better representation of variability. In region 28, around and to the east of Japan, the SST CCI products are consistently cooler than the comparison data sets. In regions 48 and 49, which are in the southern hemisphere, the SST CCI products warm relative to the comparison data sets through the period 1992-2010.



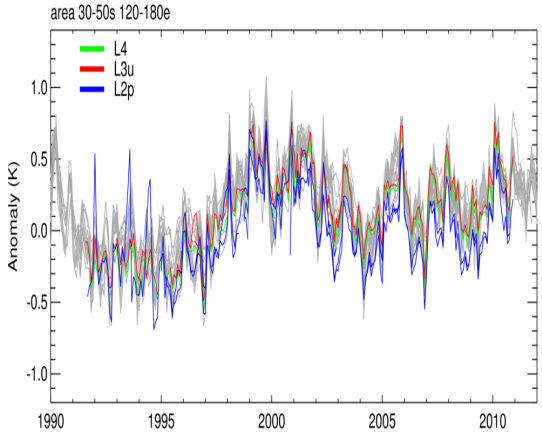
Regions 8 and 32

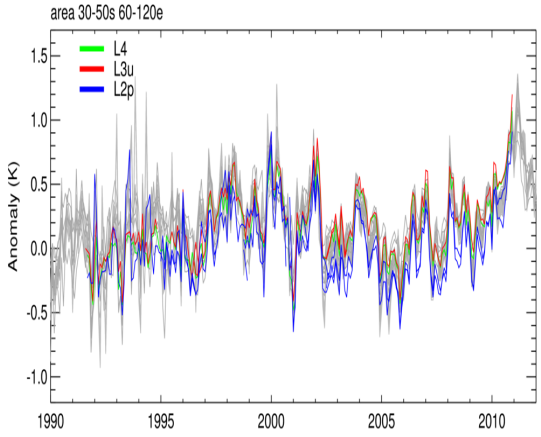


Regions 24 and 25



Region 28





Regions 48 and 49

3.1.4.2 Multi-year averages

The multi-year variability, which is typically larger than the relative biases seen in the large-scale averages, is generally consistent between the SST CCI products and the comparison data sets.

From the 1991-1995 baseline shown in Figure 3-9, the period 1996-2000 is characterised by a cooling in the central Pacific and southern Pacific and warming in most other areas. The north Atlantic is especially warm, indeed it has warmed faster (Figure 3-3, 3-4) than the other ocean basins since 1991 reflecting the switch to the warm phase of the Atlantic Multidecadal Oscillation [Knight et al., 2005; RD.344]. The cool South Pacific anomaly, which can be seen clearly in the in situ only HadSST3 and the gridded reference data sets, is consistently strong in the CCI data sets, but is weaker in COBE and HadISST. There is also a strong cool band in the southern Ocean eastern hemisphere which is present in the Pathfinder and Daily OI data sets, but less pronounced, or absent in others.

In the period 2001-2005 the central Pacific is again of interest. A cool anomaly east of the dateline can be clearly seen in most of the comparison data sets, but is absent from the SST CCI data sets and the Pathfinder data set. Cooling seen south of Africa around 45S is more muted in the SST CCI datasets than in the *in situ*-only-based comparison data sets. There is insufficient evidence to say why.

In the period 2006-2010, there are, again, interesting differences in the Pacific. A cool region extends from the Central Pacific to the west coast of South America in the SST CCI products. This cool region is seen in most, but not all of the comparison data sets. It is notably absent in the Pathfinder, Daily OI and MyOcean OSTIA reanalysis series. Given that the Pathfinder data are included in the latter two, this appears to be a feature of the Pathfinder data. There is no evidence to explain why they might have this feature.

The decadal averages are shown in Figure 3-10. In this case the decadal average anomalies relative to the MyOcean OSTIA reanalysis climatology are shown. The relative coolness of the SST CCI AVHRR and SST CCI analysis products can be seen in the period 1991-2000, particularly in the north Pacific where in situ based data sets like HadSST3, the reference data set, COBE and HadISST are warmer than the SST CCI products. This is a general pattern. Taking the decadal average of collocated differences for 1991-2000 shows (Figure 3-11) that the SST CCI analysis SST CCI product is consistently cooler at mid latitudes (in both hemispheres) than the comparison data sets (with the exception of Pathfinder, which is coolest by far). At low latitudes, there is a warm or neutral bias relative to the comparison data sets. The net effect is that the tropics to poles gradient is slightly stronger.

In terms of the differences between the decadal averages (1992-2000 and 2001-2010), both SST CCI ATSR and SST CCI analysis products warm relative to the majority of comparison data sets (see also Figures 3-3 to 3-6). The strongest warming relative to the comparison data sets between the two periods is seen along the equator and at mid latitudes, around 50-60N and 35-50S (roughly connecting the southern tips of the Americas, Australia and Africa). Between these latitudes there is either relative warming or cooling depending on the data set. One area in which the comparison data sets have warmed relative to the SST CCI ATSR and SST CCI analysis products is in the eastern Tropical Pacific. This change is seen more strongly in HadSST3 than the gridded reference data set suggesting that it is an artefact of warm biased ships in the later period. The association with the shipping lane from the Panama Canal to New Zealand is strongest in the NOCS data set, but also appears to affect ERSSTv3 and the Daily OI.

In general, the SST CCI AVHRR product has a tendency to cool relative to the comparison data sets (see also Figures 3-3 to 3-6). The relative cooling is most pronounced around 20-40S and 20-40N. There are regions in the central and western Pacific and large areas of the Southern Ocean where the Pathfinder shows relative warming compared to the comparison datasets.

3.1.4.3 Auto-correlations

Auto-correlations for the SST CCI products are comparable to those seen in other data sets (Figure 3-12). The SST CCI analysis product has somewhat higher correlations at one month lag than comparable data sets such as Daily OI and the MyOcean OSTIA Reanalysis in most areas. The lag-1 correlations are highest relative to other data sets in the Indian Ocean and South Atlantic. Higher correlations suggest a lower level of noise, but could also indicate excess smoothing. Although the lag-1 correlations are generally higher in the SST CCI analysis, the lag-2, lag-3 and especially lag-4 correlations are lower in some areas. Note in particular the Caspian Sea.

The SST CCI products provide higher lag correlations than do the purely in situ data sets, HadSST3 and the gridded reference data set suggesting that the SST CCI fields are far less noisy than equivalent fields based solely on in situ data. This is mostly due to better sampling in the SST CCI products. (However, comparing the gridded reference data to HadSST3, we see that the extra quality control applied increases the lag correlations despite reducing the numbers of observations substantially.) The differences are largest where there are fewer in situ observations and smallest in well-trafficked areas such as the north Atlantic and North Pacific. Slightly higher average correlations persist out to 4 month lag, but at lag-3 and lag-4 the correlation is lower in the SST CCI analysis product

than in either of HadSST3 or the reference data set in parts of the north West Pacific, the Gulf of Mexico and around the UK. Comparisons with the in situ based SST reconstructions (HadISST, ERSST and COBE) are also interesting. In HadISST and ERSST, the lag-1 correlations are higher than the lag-1 correlations for the SST CCI analysis in some area. For ERSST, these areas include much of the Indian Ocean and the western and southern Pacific as well as parts of the North Atlantic and Tropical Pacific. Both ERSST and HadISST are reconstructed using large-scale, relatively-smooth SST patterns, which implies that they are not effectively reconstructing small scale variability and thereby overestimating the lag-1 correlation. COBE which uses a local interpolation scheme behaves in a similar fashion to the combined satellite and in situ data sets such as Daily OI.

The lag-correlations for the SST CCI AVHRR data are generally lower than the equivalents from the SST CCI analysis data and, at lag-2 lag-3 and lag-4 lower than the equivalent SST CCI ATSR data. In comparison to the Pathfinder data, the lag-1 correlations are typically slightly higher, although parts of the southern Indian Ocean, Kuroshio extension and southern Pacific are lower. At lag-2, lag-3 and lag-4 a distinctive pattern of differences emerge with the correlations being higher in the SST CCI AVHRR product in the tropical Indian Ocean and Western Pacific and lower in the SST CCI AVHRR product in the northern and southern Indian Ocean, the South Pacific, central north Pacific, south Atlantic and parts of the north Atlantic.

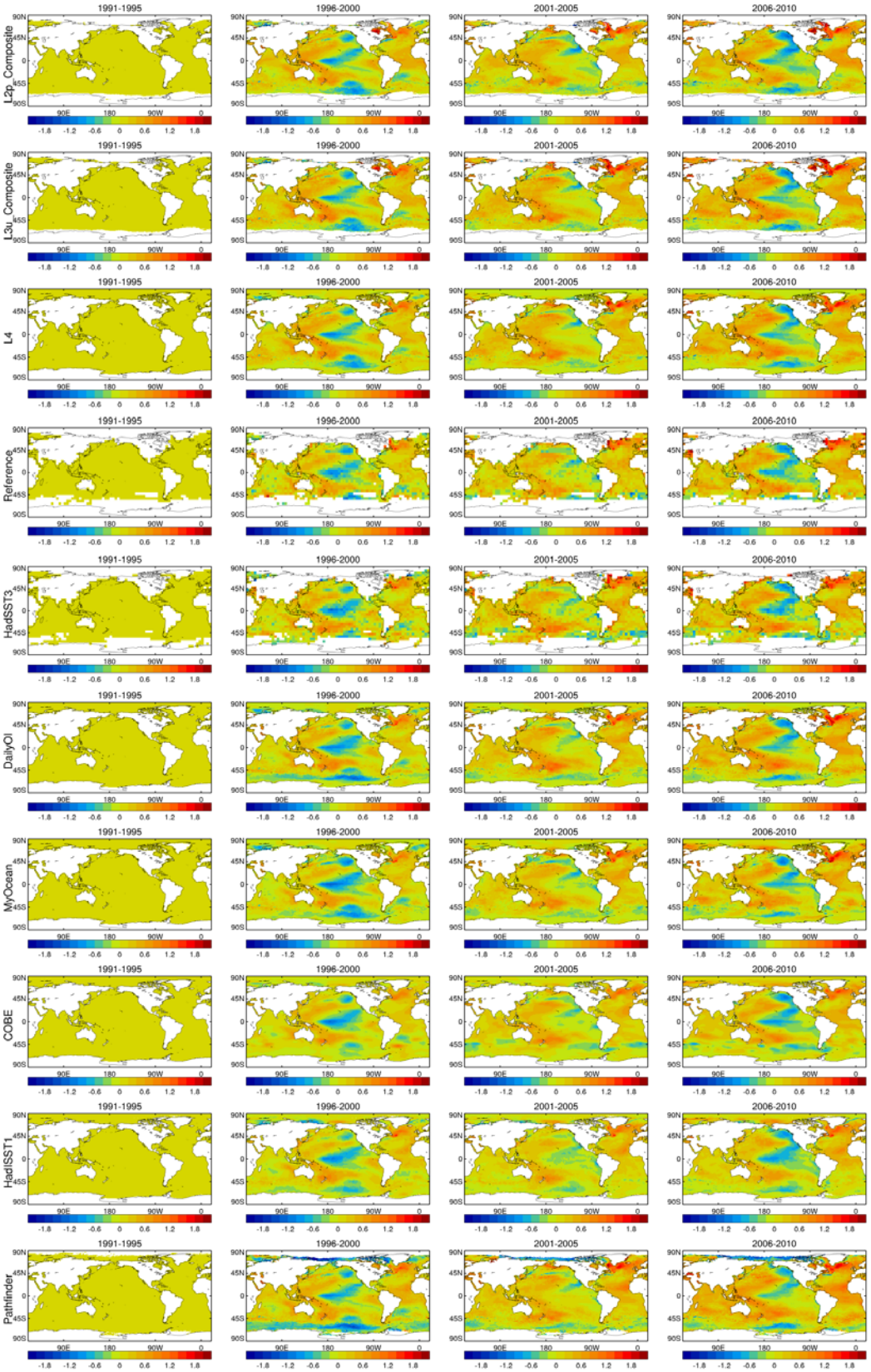


Figure 3-9. 5-year averages (K) relative to average for 1991-1995 for a selection of data sets. From left to right the periods are: 1991-1995, 1996-2000, 2001-2005, 2006-2010. For a pixel to be filled more than 30% of months need to have a valid SST.

("Reference" here refers to the gridded in situ reference data set. L4 to the SST CCI analysis, L3U to the SST CCI ATSR and L2P to the SST CCI AVHRR.)

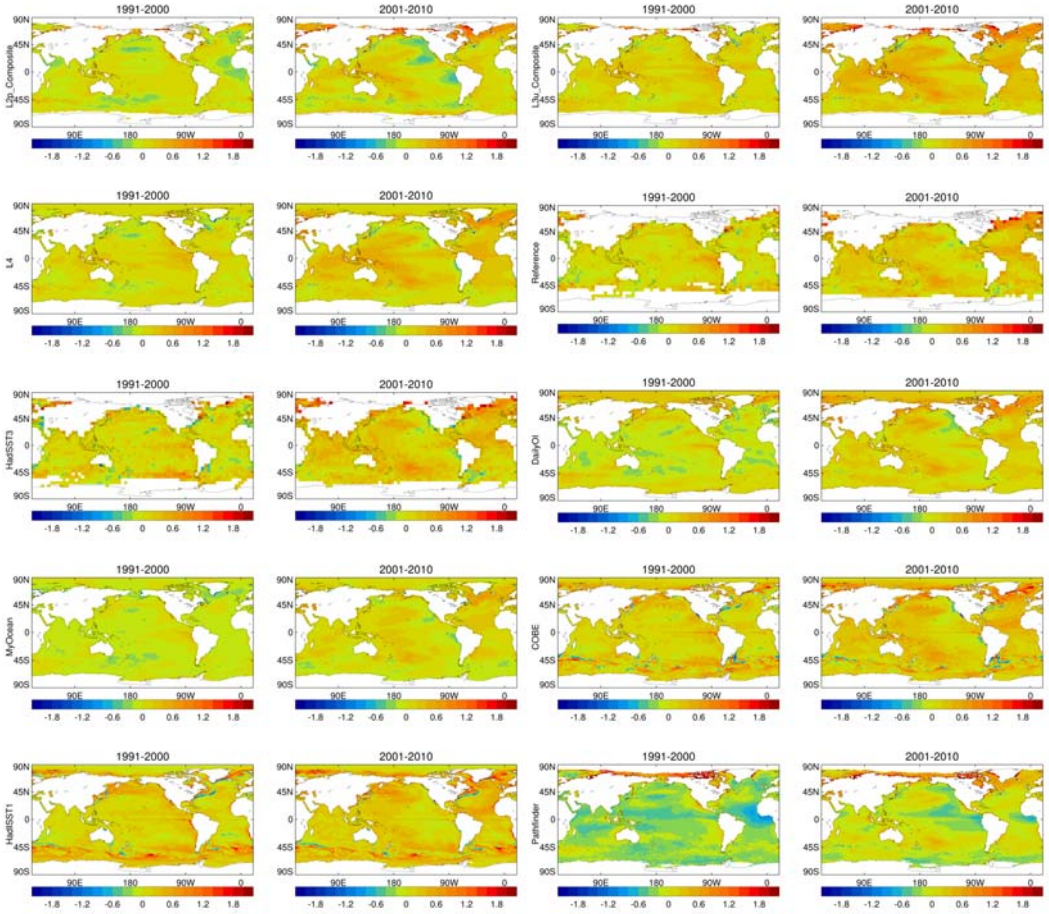


Figure 3-10. Decadal average SST anomalies (K, relative to MyOcean OSTIA Reanalysis v1.0 climatology) for a selection of data sets. From left to right the periods are: 1991-2000, 2001-2010. For a pixel to be filled more than 30% of months need to have a valid SST. ("Reference" here refers to the gridded in situ reference data set, L4 to the SST CCI analysis, L3U to the SST CCI ATSR and L2P to the SST CCI AVHRR.) Note that each row contains two data sets.

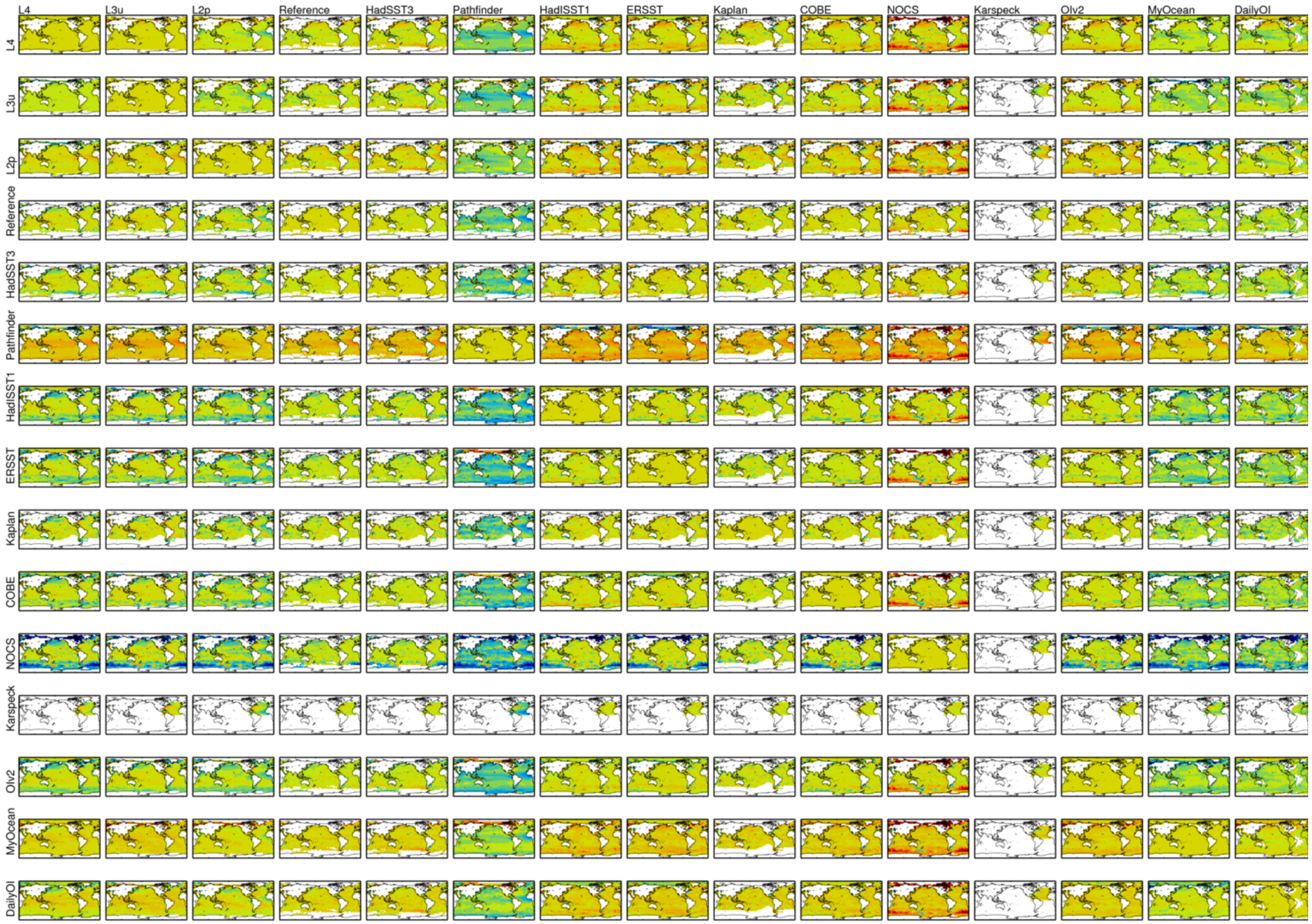


Figure 3-11. (a) Colocated decadal (1991-2000) average differences (K) between all pairs of data sets. Scale is same as for Figure 3-10. The difference is X minus Y, so for example the Pathfinder column is consistently negative because the Pathfinder has a cold bias relative to all other datasets (i.e. column contains Pathfinder minus other data sets). ("Reference" here refers to the gridded in situ reference data set. L4 to the SST CCI analysis, L3U to the SST CCI ATSR and L2P to the SST CCI AVHRR.)

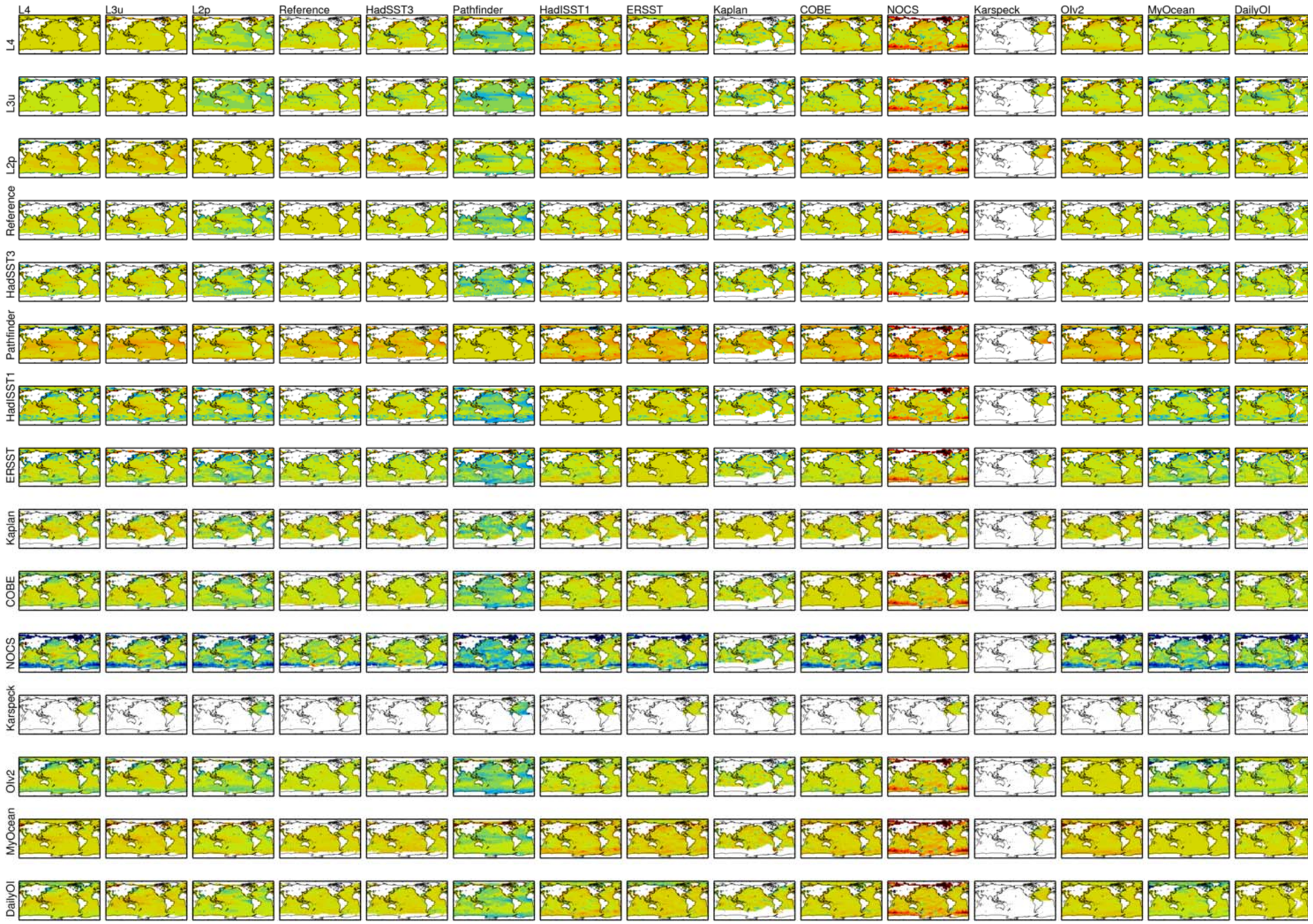


Figure 3-11 (b) As for Figure 3-11(a) but for 2001-2010.

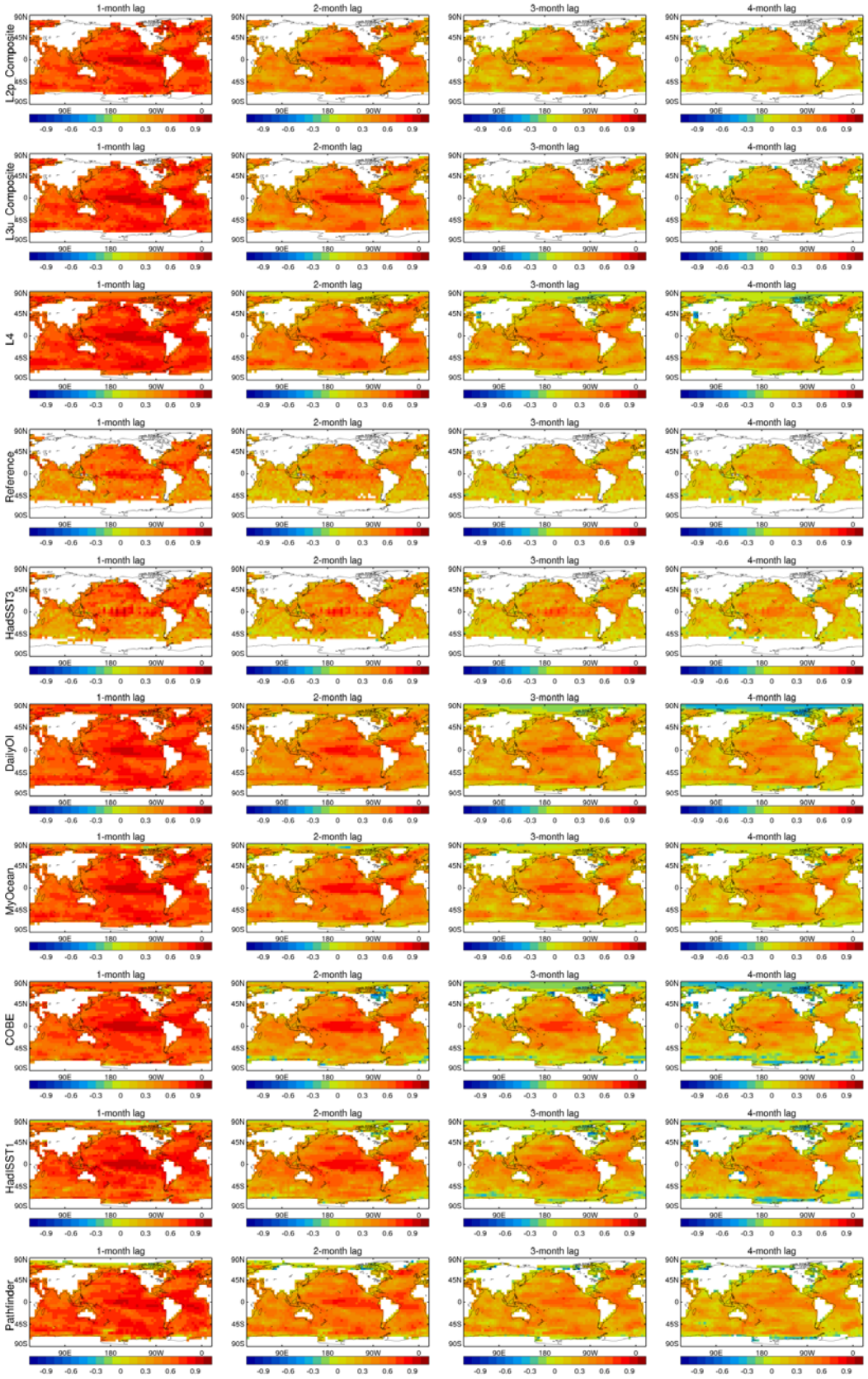


Figure 3-12. Lag autocorrelations for lags from one to four months for a selection of data sets. (L4 to the SST CCI analysis, L3U to the SST CCI ATSR and L2P to the SST CCI AVHRR.)

3.1.5 SUMMARY OF KEY POINTS

- The SST CCI LT products are comparable to the main population of comparison data sets in terms of their general evolution of SST and variability. They show smaller relative biases on the global average, relative to HadSST3, than some (cf SST CCI AVHRR [0 to -0.2 K, with some larger excursions] to Pathfinder v5.2 [~-0.4 K] and SST CCI analysis [~0] to Daily OI [~-0.1 K]).

- In general, the Pathfinder data set is an outlier amongst the data sets compared, with generally relatively cool-biased SSTs.

- There are some anomalous periods in the SST CCI AVHRR (four individual months between 1991 and 1994) and SST CCI ATSR (one month in 1996) products.

- SST CCI AVHRR products are inconsistent in their overall trends relative to SST CCI ATSR (and hence the SST CCI analysis). The SST CCI ATSR and SST CCI analysis products warm more rapidly than the majority of the comparison data sets, but the SST CCI AVHRR products tend to warm less rapidly over the period 1992-2010.

- For some climate indices, peak-to-peak variability of SST CCI LT products is higher than in the comparison data. This appears to be due to better feature resolution in the SST CCI LT products, arising from better sampling.

- The period prior to 1995 is characterised by larger cool biases in the SST CCI LT products relative to *in situ* based data sets, especially in the mid latitudes. As this is seen in comparisons both to the GTMBA and to the wider *in situ* network, we conclude that these differences arise from the SST CCI products.

- The SST CCI AVHRR product in the northern Indian Ocean exhibits spurious variation on annual timescales.

- Multi-year variability in the SST CCI LT products is generally consistent with that seen in the comparison data.

- Month-to-month correlation is higher in the SST CCI LT products and indicates they are less noisy than in situ only products (due to better sampling), however lag correlations at greater than one month lag are lower than in comparable products in places.

3.2 Stability assessment

We define stability as the degree of invariance over time of the mean error from systematic effects in SST. Ideally for a stability assessment of the SST_CCI products we require a long time series of reference measurements with known stability. For the SST CCI datasets we are targeting the Global Climate Observing System (GCOS) requirement of "<0.03 K over 100 km scales". (The time dimension is missing from this statement, but later text shows that absence of trend artefacts greater than 0.03 K per decade (i.e. 3mK/year) is the intended requirement (Ohring et al., 2005, RD.339).)

So far only one assessment of stability capable of being informative at the level of the GCOS requirement has been published (Merchant et al., 2012, RD.296). In Merchant et al. (2012, RD.296) the stability of the long-term ATSR Reprocessing for Climate (ARC) record was assessed relative to components of the Global Tropical Moored Buoy Array (GTMBA) over a 20 year period. Merchant et al. (2012) concluded that, over the period 1994 to 2010, collocated ARC and GTMBA SSTs are stable, with better than 95% confidence, to within 0.005 K yr^{-1}. As ARC and GTMBA are two independent datasets it is

reasonable to assume that the stability of 0.005 K yr^{-1} determined by Merchant et al. (2012) is an upper limit on the stability of the datasets individually for at least 1994 to 2010. As such, we have high confidence in the stability of the GTMBA for assessment of the SST_CCI datasets across this period. We attribute the high stability of the GTMBA buoys to their routine maintenance, and, crucially, their pre- and post-deployment calibration. However, as noted by Merchant et al. (2012) stability can only be directly assessed for equatorial latitudes using the GTMBA.

Other options in the reference dataset such as the drifting buoy network are not known to be stable to this level (there are no published assessments) and do not have spatial distributions that are stable in time. Therefore, drifting buoys cannot be used with confidence at the level of the GCOS stability requirement – although time-series of satellite-buoy differences should still be calculated in order to preclude major instability in the satellite record outside of tropical latitudes (e.g., see PVIR, RD.330).

The Argo network of profiling floats has sensors stated to be of very high accuracy and stability (Oka and Ando, 2004, RD.355), but covers too brief a period to allow a rigorous assessment of decadal stability: the earliest the network is adequately complete is 2004, and the techniques for and potential of the Argo network for stability assessment requires further study owing to (a) its relatively short lifetime and (b) its poor coverage over time at each location (so the data cannot be deseasonalised for example). Indeed, Argo is likely to provide the first global assessment of SST stability and consequently research into how to maximise the benefit of Argo data for stability will become more urgent in the coming years. (We have proposed an option for SST_CCI Phase II to address this, which is under consideration at time of writing.).

The utility of ship-borne radiometers (1998 onwards) in areas of repeat ship tracks for stability assessment has not yet been established (Wimmer et al., 2012, RD.337; Minnett and Corlett, 2012, RD.338) and is another area for on-going research. Radiometers may have a unique advantage above other reference datasets for stability assessment as uncertainties should be available per measurement thus avoiding the need for large numbers to reduce some of the uncertainty of the comparison between a point measurement and a satellite footprint. The issue of limited coverage is extremely acute with radiometers.

The GTMBA moorings provide consistent SSTs across the whole time period of the SST CCI datasets (August 1991- December 2010), although the number of observations available over this period increases with time (see Figure 3-13) due to (a) changes in reporting frequency (e.g. from every hour to every minute) and (b) further deployments (e.g. the addition of the PIRATA and RAMA arrays). The analysis reported here uses GTMBA locations that are present for >75% of the full SST CCI record (i.e. a single LT record from 1991 to 2010) so the number of locations does not greatly change over time (see Figure 3-15). Such moorings are mainly situated in the tropical Pacific in the TAO/TRITON arrays as these offer the longest records.

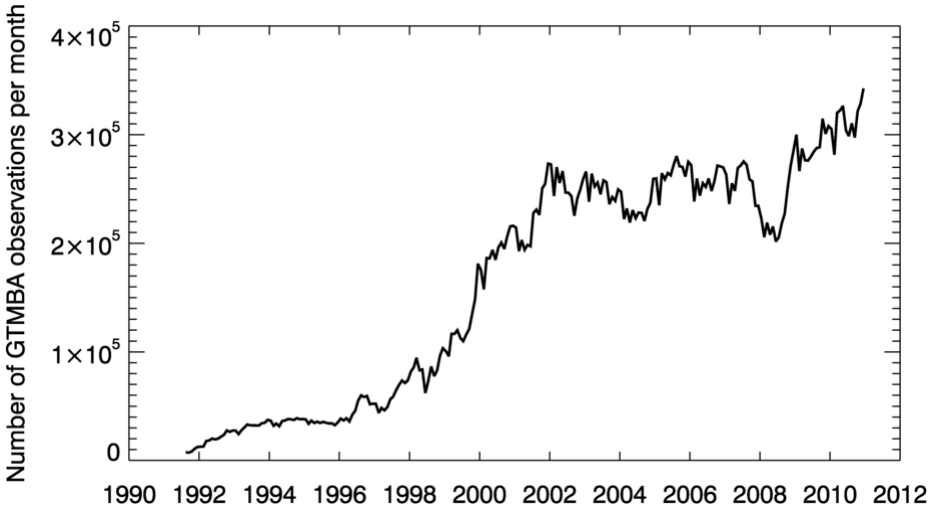


Figure 3-13. Monthly total number of GTMBA observations (August 1991 – December 2010). The analysis presented here only uses mooring locations that provide data for > 75% of the SST_CCI period (1991 – 2010) and so the real number of match-ups analysed per month is reasonably consistent over time.

3.2.1 STABILITY OF THE SST CCI PRODUCTS

3.2.1.1 Introduction

The stability of the SST CCI products has been assessed through comparisons to GTMBA moorings at 1 m depth using the approach and methodology provided in the GHRSST Climate Data Assembly Framework (CDAF, Merchant et al., 2013, RD.317).

3.2.1.2 Methodology

Briefly, the three SST CCI datasets (SST CCI AVHRR, SST CCI ATSR and SST CCI analysis) were individually matched to GTMBA data for the full time period (1991 – 2010). The first set of matched data was generated from the Multi-sensor Matchup System (MMS), which was built on pre-matched ATSR data to GTMBA data extracted from a combination of ICOADS v2.0 and National Centers for Environmental Prediction (NCEP) real-time feeds. The resulting number of match-ups was not sufficient to carry out a meaningful stability assessment and so the set of high-temporal resolution GTMBA provided by NOAA's Pacific Marine Environmental Laboratory (PMEL) was downloaded from the web (http://www.pmel.noaa.gov/tao/disdel/disdel.html) and new Matchup Dataset (MD) files were created. The high-temporal resolution GTMBA data had a sampling resolution of either 5, 10 or 60 minutes and the highest available temporal resolution was always used if multiple resolutions were available. The GTMBA data were matched to the nearest SST CCI pixel centre and a maximum time difference of 30 minutes was used as a threshold. For the SST CCI analysis the GTMBA data was a mean of the nearest measurements to 10:30 AM and 10:30 PM as the SST CCI analysis is a daily mean. No further quality control or filtering was applied to the data prior to analysis.

The methodology for performing the stability assessment defined in the CDAF is given below. This methodology has been internationally agreed for trial implementation within the context of GHRSST at the present time, and may be subject to future revision.

- Following the initial match-up process the monthly median SST CCI minus GTMBA difference for each GTMBA location was calculated. This considers each location independently and avoids aliasing by periods with larger number of match-ups or changes in GTMBA reporting.
- Then for each month of the year and location, the multi-year average of the monthly median SST CCI minus GTMBA differences was calculated. This considers each location independently and allows us to create an SST anomaly at each location.
- For each month the data were then deseasonalised by subtracting the multi-year average for the appropriate month of the year from each month of the time series. For the SST CCI AVHRR and ATSR products, separate multi-year averages were used for day and night. The data were deseasonalised to minimise any potential aliasing of an annual cycle in residual time series following the approach of Merchant et al. (2012, RD.296).
- Subsequently, the match-ups were further reduced by retaining only locations where buoy data were available for >75% of the period to be assessed, i.e. > 15 years within the 1991-2010 period. This allows us to use a consistent number of match-ups per month throughout the time series and avoids aliasing through additional GTMBA locations.
- The monthly mean difference across all locations was determined to end up with a single SST CCI minus GTMBA time series for each SST CCI product (as shown in Figure 3-14 for SST CCI AVHRR, ATSR and analysis products).
- A least squares linear fit to each time series of monthly mean differences was calculated to determine any trend in the data and 95% confidence intervals were estimated for the final analysis.

3.2.1.3 Results

The results from the stability assessment are shown in Figure 3-14 for all SST CCI products, ATSR (top), AVHRR (middle) and analysis (bottom). For SST CCI ATSR and AVHRR, results are presented separately for day time (red) and nighttime (blue). For the SST CCI analysis there is a single daily mean (green). Over plotted in Figure 3-14 are linear least square fit (solid line) and its associated 95% confidence bands (dashed lines) for each time series. An objective of this analysis is to identify step changes over times that are most likely due to issues with the data quality. Although step-detection techniques were not used, a clear step-change is apparent from 1995 onwards, which is most likely due to the change between ATSR-1 and ATSR-2 as the two sensors were not harmonised and different retrieval methods were used for each instrument. In addition to the step change, there other notable features in Figure 3-14, including: (i) a period of instability at the end of life of the first ATSR (early 1996) and (ii) a trend, prior to 1995, associated with diminishing retrieval artefacts following the eruption of Mt Pinatubo.

The main reason for not using step-detection techniques is that these could not be included as baseline activities in SST CCI Phase I. Such techniques require significant resource and expertise to implement and were provided as an option for Phase II. The need for this topic to be addressed at a cross-ECV level that benefits all variables has been raised both in SST project meetings and cross-ECV meetings during Phase I.. So far, no mechanism to do this has been found. A consequence, of not using step-detection techniques is that step changes (such as the change between ATSR-1 and ATSR-2 evident in Figure 3-14) have to be identified visually/subjectively, with a corresponding chance both of step changes being missed and being falsely imputed.

As the ATSRs were used to bias correct the radiances for the AVHRRs the feature is also apparent in the SST CCI AVHRR time series and, as no further bias correction was done prior to ingestion into the SST CCI analysis system, the step-change is also present there as well. Consequently, the 95% confidence interval on the slope of the fit was calculated

for two separate periods, 1991 to May 1995 covering the ATSR-1 period and June 1995 to 2010 covering the ATSR-2/AATSR period.

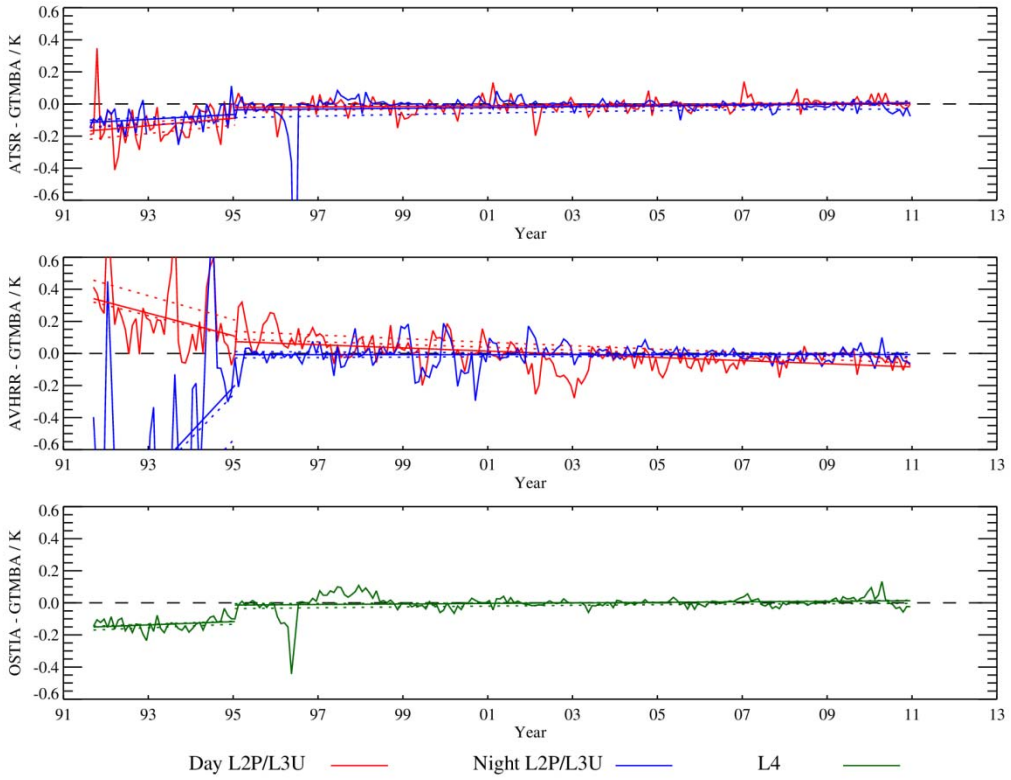


Figure 3-14: Time series of deseasonalised composite monthly mean differences (K) between the SST CCI products and the GTMBA. Separate day and nighttime series are provided for the SST CCI AVHRR and ATSR products. Also plotted are the results of a least squares linear fit (solid lines) for the 1991 to May 1995 and June 1995 to 2010 periods and their 95% confidence bands (dashed lines). Please see text for further discussion; OSTIA/L4 refers to SST CCI analysis, L2P to SST CCI AVHRR and L3U to SST CCI ATSR.

The resulting confidence intervals for the least squares linear fits to the time series of SST CCI minus GTMBA differences are summarised in Table 3-2.

SST CCI 95% confidence interval (mK year^{-1}) for 1991 – 1995			
	Day	Night	Both
SST CCI AVHRR	-137.9 < trend < -2.4	105.9 < trend < 462.3	
SST CCI ATSR	-13.6 < trend < 60.1	-7.4 < trend < 36.8	
SST CCI analysis			-1.8 < trend < 22.1
SST CCI 95% confidence interval (mK year^{-1}) for 1995 – 2010			
	Day	Night	Both
SST CCI AVHRR	-12.3 < trend < -7.4	-2.0 < trend < 2.0	
SST CCI ATSR	0.7 < trend < 3.2	-1.4 < trend < 6.4	
SST CCI analysis			0.1 < trend < 3.2

Table 3-2: Summary of 95% confidence intervals for least squares linear fits to SST CCI minus GTMBA monthly mean difference time series for 1991 to May 1995 and June 1995 to 2010.

For the SST CCI ATSR product, the nighttime trend in the differences to the GTMBA measurements for the 1995-2010 period is comparable to that calculated by Merchant et al. (2012, RD.296). However, the day time stability confidence interval doesn't include zero, and relative to RD.296 is somewhat less stable; nonetheless, the true stability is still likely to be within the GCOS requirement. For the ATSR-1 period, both the day and night trends calculated for the SST CCI ATSR product have improved stability (based on the most likely relative trend) compared to that reported in Merchant et al. (2012, RD.296), although there is nonetheless likely to be a positive trend artefact that is outside the GCOS target as the 95% confidence limits of the estimated trend are outside the GCOS requirements and are not symmetrical (i.e. are dominated by a positive value).

Regarding the SST CCI AVHRR product there is no comparable analysis in the literature for pre-cursor datasets such as Pathfinder, i.e. long-term SST records generated from the AVHRRs. We note that, as for the SST CCI ATSR product, the day time stability is poorer than for nighttime. This may reflect the greater amplification of error in two-channel relative to three-channel SST retrieval that is common to all infra-red sensors and retrieval methods. The SST CCI AVHRR time series in Figure 3-14 appears to have a step improvement in inter-annual stability from 2003/4. The cause of this is not clear, as this does not correspond to a changeover between sensors within the time series. Also, we note the day time trend in the SST CCI AVHRR minus GTMBA time series is negative in both periods. Since this is a time series from multiple instruments, each with their own calibration issues, differential quality of the CLAVR-x cloud masking and pattern of orbit drift, a trend at this subtle level is likely to be a complex mix of these factors.

Being tied to the SST CCI ATSR calibration, the SST CCI analysis product has a stability over the period 1996 to 2010 that likely meets the GCOS requirement.

As noted earlier, this stability assessment is somewhat limited as it only covers a region in the tropics. Consequently, the resulting deseasonalised trends are only representative of

this area and we will not necessarily see the effect of the large scale regional biases reported in the SST CCI PVIR (RD.330). The requirement that > 75% of months were available over the period resulted in primarily the tropical Pacific Ocean GTMBA buoys being used, as these are the longest records. The locations of all GTMBA moorings (109 locations; indicated by the blue and red dots and the reduced set used here (65 locations; indicated by the blue dots only) are shown in Figure 3-15. It is clear from inspecting Figure 3-15 that the SST CCI stability assessment reported here is only directly applicable to the tropical Pacific Ocean.

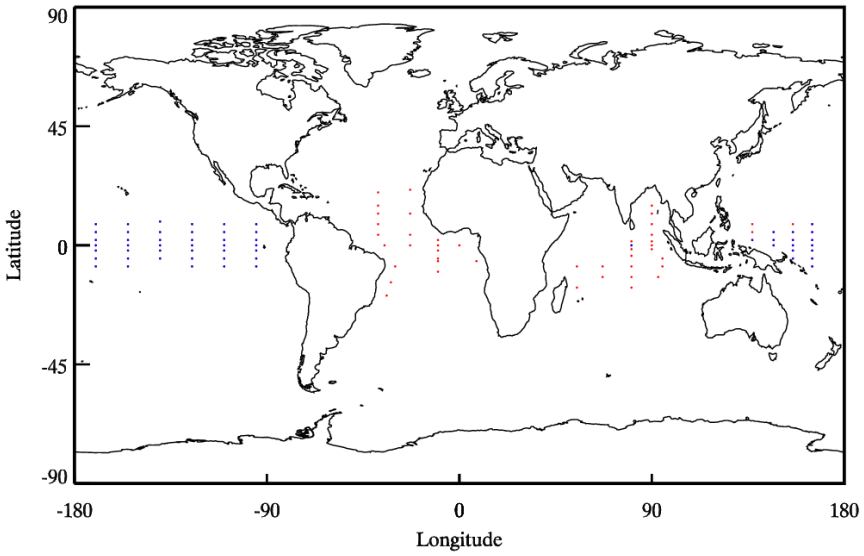


Figure 3-15: Nominal reference location of GTMBA buoys (red & blue dots) and the reduced set of locations (blue dots) used for the SST CCI stability assessment. The plot is valid for the period 1991-2010.

3.2.2 COMPARISON OF STABILITY IN THE SST CCI ANALYSIS PRODUCT TO THAT OF THE GHRSST MULTI-PRODUCT ENSEMBLE

Data from the Global Tropical Moored Buoy Array (GTMBA) were again used as a reference to compare the long-term stability of the SST CCI analysis to that of other comparable global, daily SST datasets: My Ocean OSTIA Reanalysis v1.0 (Roberts-Jones et al., 2012, RD.239), Canadian Meteorological Centre (CMC, Brasnett et al., 2012, RD.344), AVHRR-only version of the Daily OI (Reynolds et al., 2007, RD.76), HadISST2 realisation 396 (Kennedy et al., 2013, RD.345; Rayner et al., 2013, RD.334) and Merged satellite and in situ data Global Daily SST (MGDSST, Kurihara et al., 2006, RD.237). A further dataset comprising the median of these six analyses for each grid box, referred to as the “GMPE median”, was also used.

The OSTIA Reanalysis v1.0 and AVHRR-only Daily OI datasets were described previously (Sections 3.1.2.11 and 3.1.2.12 respectively). CMC (Brasnett et al., 2012, RD.344) is a daily global analysis from the Canadian Meteorological Centre, using various AVHRR, ATSR-series and microwave instruments. The analysis is produced on a 1/5 degree grid, and is referenced to ship and buoy data from ICOADS with a typical depth of 1m. MGDSST (Kurihara et al., 2006, RD.237) is a foundation temperature analysis on a ¼ degree grid based on infrared and microwave observations from AVHRR, AMSR-E and WindSat instruments, and in situ data from the GTS. HadISST2 (Kennedy et al., 2013,

RD.345; Rayner et al., 2013, RD.334) is valid at a nominal depth of 20cm, and uses AVHRR and ATSR-series observations (from ARC), with in situ data from ICOADS. A daily, ¼ degree version of HadISST2 was used here, and realisation 396 was randomly selected from the available series of interchangeable realisations.

The stabilities of the SST CCI products were determined previously (Section 3.2.1), evaluated using the mean of the GTMBA observations at 1030 hrs and 2230 hrs local time for each buoy. Here, all available observations, both daytime and nighttime, are used. This provides additional data and allows calculation of the stability of the other SST analyses which, unlike the SST CCI analysis, are not specifically designed to capture the daily mean. Matchups between the GTMBA observations and the SST analyses were produced by interpolating the analyses to the observation locations[1]. The method used for the stability assessment itself is otherwise the same as that used in section 3.2.1.2, and described by Merchant et al. (2013, RD.317).

All the analyses used in the comparison, with the exception of the SST CCI analysis, assimilate moored buoy observations either from ICOADS or received over the GTS. This means the GTMBA dataset used here is not independent from all the other analyses. Although the GTMBA data used here are of a higher temporal resolution than that available in ICOADS (see Section 3.2.1), this nonetheless means that the analyses are locally tuned to GTMBA moorings to a degree dependent on the weighting of the moorings in the analysis. The stability of these datasets in locations away from GTMBA moorings may be significantly different and is unknown.

Table 3-3 shows the results for each dataset, with linear trends in the analysis minus GTMBA differences given in mK per year for the full period (1992 – 2010) and the periods when the different ATSR-series instruments were used in the SST CCI analysis, namely:

- ATSR-1: 01/1992 – 05/1995
- ATSR-2: 07/1996 – 07/2002
- AATSR: 08/2002 – 12/2010

For a product to be considered stable over the whole of its record, sub-periods of that record should also be seen to stable.

Although SST CCI data are available from August 1991, it was decided to begin this stability analysis from January 1992, both for computational efficiency and to allow production of the multi-year monthly average required for deseasonalising from the same number of data points per month, i.e. not including part-years.

In the gap between ATSR-1 and ATSR-2 shown above, the two instruments were being swapped in the SST CCI analysis according to availability of data, so this period is not included in the trend calculations for simplicity and to avoid an outlier in 1996 (arising from problems with ATSR-1, see Figure 3-16a) biasing the results. Not all the analyses use data from the ATSR series of instruments [see PVIR for details, RD.330] but the trends in analysis minus GTMBA difference are still calculated for the same periods to enable intercomparison between datasets. The minimum and maximum linear trends for each period given in Table 3-3 are the 95% confidence intervals. Note that, as OSTIA v1.0 and HadISST2 stop in 2007, the AATSR period for these datasets is 2002-2007.

[1] *Observational uncertainties, e.g. of interpolating analyses to GTMBA locations, are not considered in this process.*

Table 3-3. Linear trends for monthly mean analysis-minus-GTMBA difference for each analysis, in mK/yr. Shown for full time period (01/1992 – 12/2010), ATSR-1 period (01/1992 – 05/1995), ATSR-2 period (07/1996 – 07/2002) and AATSR period (08/2002 – 12/2010, or 12/2007 for OSTIA v1.0 and HadISST2). Max and min trends are the 95% confidence intervals.

	Trend (mK/yr)	Min Trend (mK/yr)	Max Trend (mK/yr)
SST CCI analysis			
Full period	8.2	6.5	9.9
ATSR-1 period	30.1	14.9	45.4
ATSR-2 period	-15.0	-21.0	-9.0
AATSR period	4.2	1.5	6.9
OSTIA v1.0			
Full period	0.6	-1.28	2.5
ATSR-1 period	1.1	-7.4	9.6
ATSR-2 period	-4.3	-8.7	0.2
AATSR period	12.7	9.7	15.6
CMC			
Full period	-0.7	-1.4	-0.002
ATSR-1 period	-7.3	-12.8	-1.7
ATSR-2 period	5.5	0.03	10.9
AATSR period	-1.1	-2.2	-0.04
Daily OI			
Full period	7.7	6.2	9.2
ATSR-1 period	9.7	-9.9	29.4
ATSR-2 period	2.0	-5.5	9.6
AATSR period	17.6	13.1	22.1
HadISST2			
Full period	1.3	0.2	2.3
ATSR-1 period	4.5	-4.6	13.5

ATSR-2 period	-4.0	-8.6	0.6
AATSR period	-2.4	-7.4	2.6
MGDSST			
Full period	0.9	-0.2	2.0
ATSR-1 period	7.1	-8.6	22.7
ATSR-2 period	-17.2	-22.3	-12.1
AATSR period	-4.7	-7.4	-2.1
GMPE median			
Full period	3.7	3.2	4.3
ATSR-1 period	5.1	-2.9	13.0
ATSR-2 period	-5.4	-7.9	-3.0
AATSR period	6.0	4.6	7.4

Trends in analysis minus GTMBA difference over the full time period indicate that CMC and MGDSST are the most stable datasets relative to GTMBA, with trends of -0.7 and 0.9 mK/yr respectively. The uncertainty in the CMC trend is smaller than for MGDSST (Table 3-3). This result may reflect a high degree of stability across tropical regions for CMC and MGDSST, or it may simply reflect that the analyses give a higher weight to the ingested GTMBA data. If the latter, the stability results may only be representative of GTMBA locations. It is impossible to distinguish these possibilities and therefore to make a confident inference of large spatial-scale stability when using non-independent datasets. However, it is useful to compare these stability results with the SST CCI analysis results, which are independent from the GTMBA dataset. The SST CCI analysis shows the largest magnitude trend, 8.2 mK/yr, over the full time-series. Because this dataset is independent from the GTMBA measurements and the satellite algorithms are universal, we can have good confidence that this result is valid for tropical regions in general, and not only GTMBA mooring locations.

The trend over the full time period for OSTIA v1.0 is also small (0.6 mK/yr). However, this is not representative of the trends for different periods in the analysis (Figure 3-16). Trends for OSTIA v1.0 vary between periods corresponding to the influence of the different ATSR instruments on this analysis and the SST CCI analysis. Trends have also been calculated for these same subsections of the other analyses. Results are given in Table 3-3 and time series for each analysis are also shown in Figure 3-16.

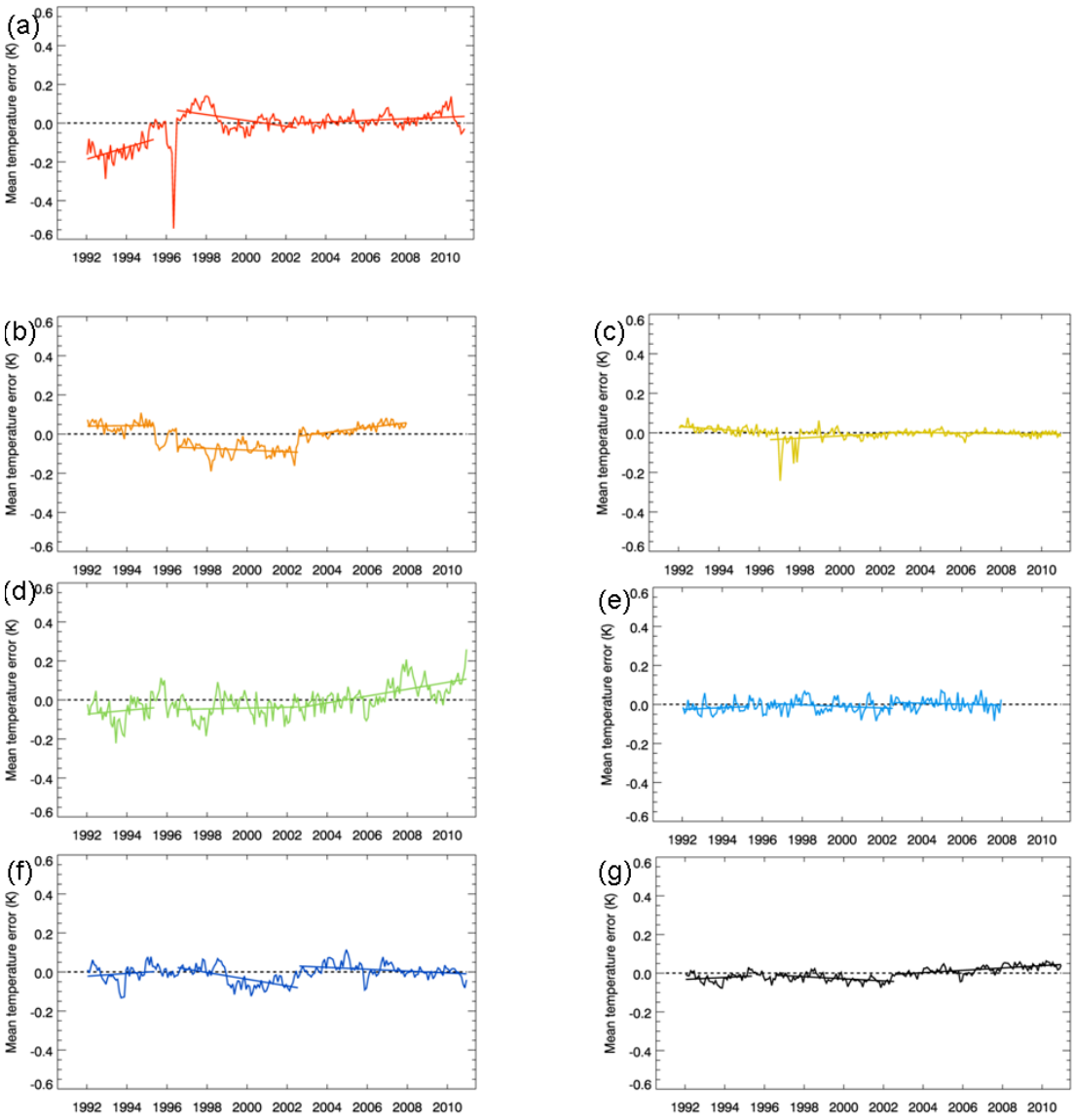


Figure 3-16. Monthly mean deseasonalised analysis-minus-GTMBA difference (K), showing linear trend fits for different ATSR periods (see text). (a) SST CCI analysis, (b) My Ocean OSTIA reanalysis v1.0, (c) CMC, (d) AVHRR-only Daily OI, (e) HadISST2, (f) MGDSST, (g) GMPE median. Differs from Figure 3-14 as all GTMBA data have been used for matchups, not just 1030 AM/PM.

Assessment of the shorter-term trends in analysis-minus-GTMBA differences demonstrates that trends in HadISST2 for each period are very similar to each other and very small (Figure 3-16, Table 3-3). This indicates the HadISST2 analysis is very temporally homogeneous at the GTMBA locations. CMC also shows good results, with fairly small trends in analysis-minus-GTMBA differences which are consistent in magnitude between the ATSR periods, with perhaps a small discernible influence of the ATSR-2 instrument on the time series. The uncertainty on the CMC trend, particularly in the AATSR period is smaller than that for HadISST2. However, as the GTMBA data are not independent from these analyses, the small trends may be related to a high weighting given to them in the analyses. Similarly, the larger trends in analysis-minus-GTMBA difference found in e.g. AVHRR-only Daily OI may indicate a greater weight is given to satellite data relative to in situ data in such analyses. The MGDSST blends in AMSR-E SST from 2003, which coincides with a possible step-change in the systematic difference

to GTMBA towards a relatively less biased MGDSST[2]. Trends in analysis-minus-GTMBA differences for the ensemble median of all the analyses, the GMPE median, are small and fairly consistent between ATSR periods.

The positive trend in the SST CCI analysis-minus-GTMBA differences during the ATSR-1 mission reflects the residual effects on SST retrieval of the Pinatubo eruption. The negative trend during the ATSR-2 mission arises mostly from effects in the first two years of that mission. Over the AATSR period (August 2002 to December 2010), there is a statistically significant positive trend of 4.2 mK/yr. Neglecting the problematic ATSR-1 period and assessing analysis-minus-GTMBA differences over the period June 1995 – December 2010 (as used in section 3.2.1.3) gives a trend of 2.3 mK/yr (0.3 < trend < 4.4, 95% confidence interval). This is comparable to that found for the SST CCI analysis over the same period in section 3.2.1.3 (0.1 < trend < 3.2), despite using all available matchups here and not just those for 1030 hrs and 2230 hrs. Clearly, using only data for the AATSR period produces a larger analysis-minus-GTMBA trend than that calculated over this longer period, illustrating the difficulty of expecting a single statistic to represent the stability of a whole SST time series.

The trend in the SST CCI analysis-minus-GTMBA differences is larger than for the other analyses in the ATSR-1 and ATSR-2 periods, but improves during the AATSR period to be third best (smallest) (Table 3-3). Poorer performance in stability compared to the other analyses is to be expected as SST CCI is the only of these analyses independent from the GTMBA data. However, the improvement in the AATSR period relative to the other analyses indicates that the performance of the SST CCI analysis in the tropics is better than in the earlier ATSR-1 and ATSR-2 periods. This demonstrates the benefit of this intercomparison, despite the lack of independent data. Reduced performance in the ATSR-1 period for the SST CCI analysis is likely a result of the effects of the Pinatubo eruption on SST retrievals, as mentioned above.

3.2.3 SUMMARY OF KEY POINTS

- The GCOS stability requirement (of better than 3mK/year, measured by evaluating the trend in the difference between SST CCI products and GTMBA measurements) is met in the tropical Pacific over the period 1995-2010 by the SST CCI AVHRR nighttime product (-2.0 < trend < 2.0mK/year). The SST CCI ATSR daytime (0.7 < trend < 3.2mK/year) and SST CCI analysis (0.1 < trend < 3.2mK/year) products fall marginally outside this target over the same time period. Stabilities of other products fall outside of the GCOS requirement. [SST CCI AVHRR day time: -12.3 < trend < -7.4mK/year; SST CCI ATSR nighttime: -1.4 < trend < 6.4mK/year]

- Tropical Pacific SST CCI ATSR day time stability over 1995-2010 is comparable to that of the ARC data set. Nighttime stability over the same period is less good.

- Tropical Pacific SST CCI ATSR day (-13.6 < trend < 60.1mK/year) and nighttime (-7.4 < trend < 36.8mK/year) stability over the period 1991-1995 is improved relative to that of the ARC data set.

- Stability outside the tropical Pacific cannot be evaluated because insufficient stable reference data exist away from this region.

[2] *This is evidence that there are potential benefits from the incorporation of microwave data in the SST CCI analysis and, therefore, that there is a need for further work to produce consistent SST retrievals from microwave measurements (as proposed by DMI in an option to SST CCI Phase II).*

- SST CCI analysis is apparently less stable in the tropics with respect to GTMBA measurements than the other GMPE analyses, but since the non-SST CCI analyses are not independent of GTMBA, the figures obtained are likely to exaggerate their true stability in non-GTMBA locations.

- The stability of the SST CCI analysis improves over successive ATSR missions, reflecting the improvements in stability of the successive instruments in the series or improvements in the stability of the GTMBA over similar time periods.

- Methods to assess stability over wider geographical areas using, for example, Argo and improvements to the moored buoy network, are required.

- A consistent approach to step-detection techniques is required across all or most CCI products and domains. Preferably, specialist statistical expertise in this area would be embedded in a tool or methodology that can be confidently used by CCI science teams.

4. USE IN CLIMATE MODELLING AND OTHER APPLICATIONS

Sea surface temperature information is needed, whether directly or indirectly, for the vast majority of climate applications (e.g. for driving atmospheric only climate models and reanalyses, evaluating coupled and ocean model simulations, monitoring surface temperature and initialising seasonal-to-decadal forecasts). In this section, we document the applications in which we and our trail blazer users have tested the pre-release version of the SST CCI products. We also discuss scientific results arising from those test use cases. Feedback arising from these early uses of the (pre-release) products has identified some problems (in some cases, these were simple to rectify before the general product release) and informs the continuing User Requirements analysis.

The first sub-section describes the use of the products in coupled model assessment by the SST CCI Climate Research Group at the Met Office Hadley Centre. The second sub-section then presents the voluntary work of the trail-blazer users.

4.1 Use in model assessment by the SST CCI Climate Research Group at the Met Office Hadley Centre

4.1.1 INTRODUCTION

As coupled climate models are developed, they are tested to assess the realism of the simulated climate. The realism of the SST and its variability is a key aspect of the realism of the climate simulation as a whole. The current version of the Met Office Hadley Centre climate model is HadGEM3 [publication pending]. Present day control simulations are assessed, i.e. unforced simulations with fixed greenhouse gas and aerosol concentrations and start up conditions close to those of the present day to demonstrate the internal variability and unperturbed state of the coupled model. Present day control simulations, rather than simulations of the recent past driven by evolving radiative forcings (such as greenhouse gases and aerosols), are evaluated to later allow simulations of the recent past to be used to detect and attribute the influence of different forcing factors on climate, without concern that these simulations have been tuned to the data to which they are compared.

A standard suite of software is used in routine assessments of development versions of the Met Office Hadley Centre climate models. This process compares the mean state of the model simulation against various observed data sets. New model versions are compared to older model versions to assess their relative performance. Sea surface temperature from coupled model simulations is routinely compared to the Reynolds et al [2007] Daily OI analysis, incorporating data from both AVHRR and AMSR-E. Here we assess the impact of using the SST CCI analysis instead, to determine whether or not the different retrievals, resolution and analysis methodology used affect the assessment of the model simulations. We assess the mean state of two HadGEM3 model versions: the GC2 (newer) and GC1 (older) versions [publication pending], as available in October 2013.

We also utilise 19 years of simulated daily SST from a GC2 control simulation of HadGEM3 to explore the daily variability simulated by the model as compared with observations. There are no standard assessment methods currently used to assess simulated daily variability. Therefore, this study constitutes a subjective exploration of the daily SST variability simulated by HadGEM3, utilising the newly available daily means of the SST CCI analysis product.

4.1.2 ASSESSMENT OF THE MODEL MEAN STATE

Figure 4-1 compares 20-year means of the SST simulated in GC2 and GC1 to 20-year means of the SST CCI analysis. There are large biases in the mean state of both model versions in the Southern Ocean when compared to the SST CCI analysis. Large biases are also seen in the mid-latitude north Pacific, which appear worse in GC2. These biases vary with the seasons, as seen in Figure 4-2 and 4-3. Biases in the Southern Ocean are smaller in June-August and September-November (winter and spring) and larger in December-February and March-May (summer and autumn). Biases in the mid-latitude north Pacific are largest in March-May (spring).

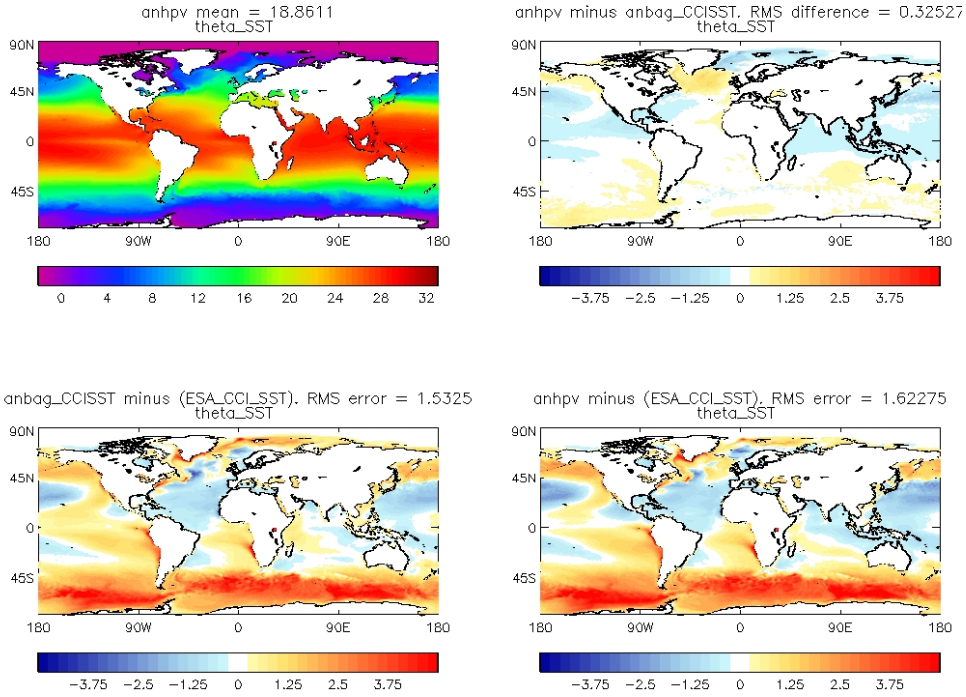


Figure 4-1. Top left: Annual 20–year mean SST (°C) simulated in GC2. Top right: SST in GC2 minus SST in GC1. Bottom left: SST in GC1 minus SST CCI Analysis. Bottom right: SST in GC2 minus SST CCI Analysis.

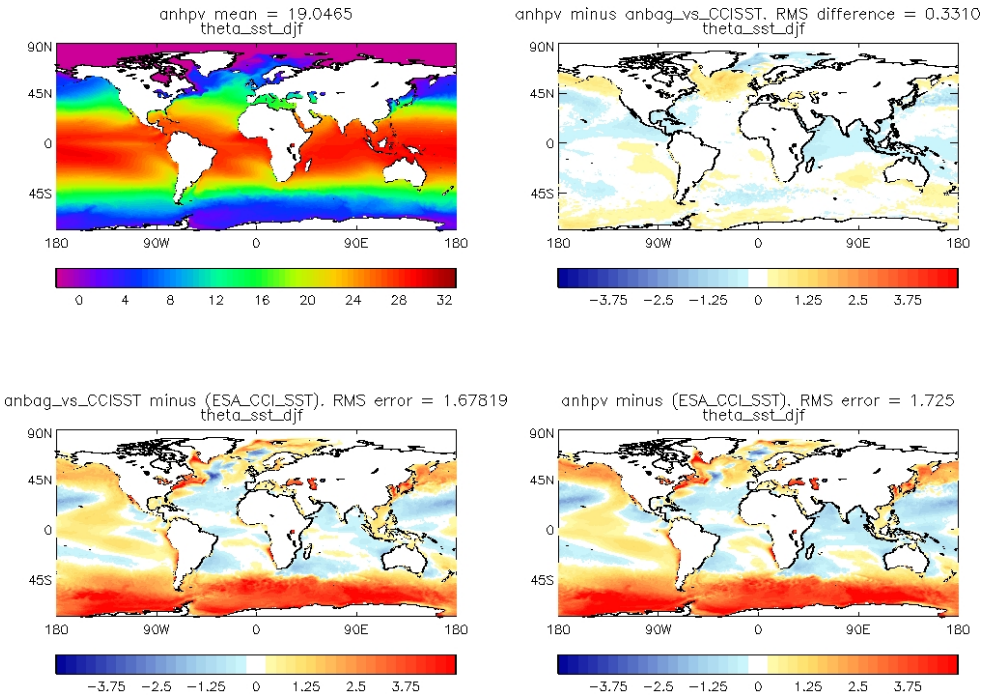


Figure 4-2. Top left: December-February 20-year mean SST (°C) simulated in GC2. Top right: SST in GC2 minus SST in GC1. Bottom left: SST in GC1 minus SST CCI analysis. Bottom right: SST in GC2 minus SST CCI analysis.

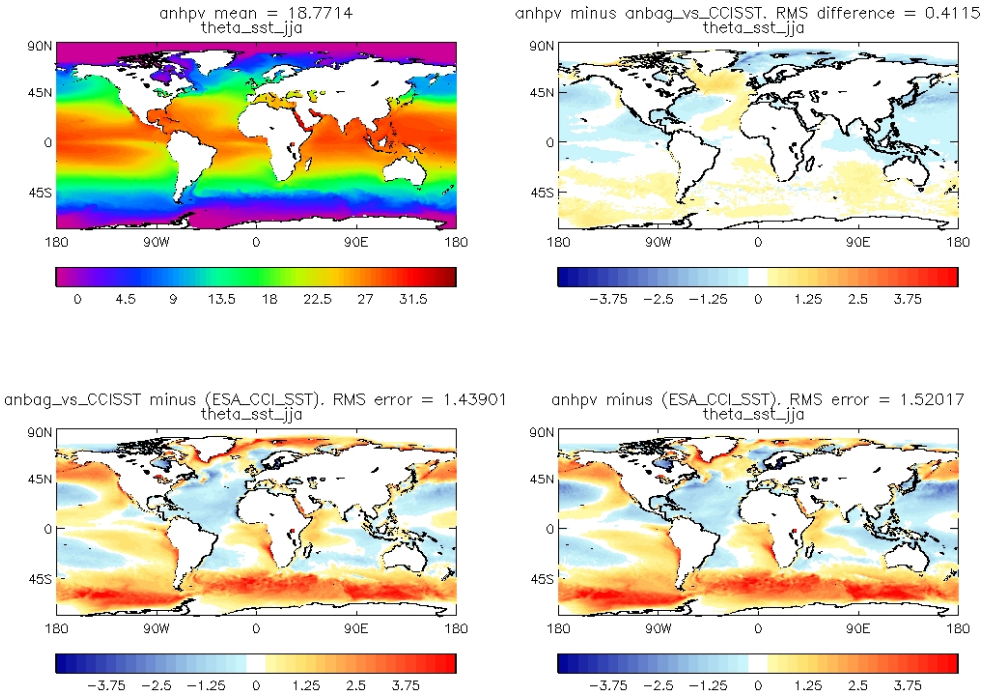


Figure 4-3. Top left: June-August 20-year mean SST (°C) simulated in GC2. Top right: SST in GC2 minus SST in GC1. Bottom left: SST in GC1 minus SST CCI analysis. Bottom right: SST in GC2 minus SST CCI analysis.

Comparing results relative to the SST CCI analysis to those relative to Reynolds et al [2007] Daily OI (Figure 4-4 and Table 4-1) we see a very similar story. In general, GC2 is more different from the observations (larger RMS difference) than GC1. Regionally, whether the Daily OI or the SST CCI analysis is used, GC2 is determined to be worse than GC1 in the same places; the comparison to the SST CCI analysis tends to yield slightly larger errors. The differences between the two observational analyses are very much smaller than the errors in the mean states of the model simulations (Table 4-1). Using a different observational data set makes little difference to the conclusions drawn in this case, but provides increased confidence that the observational estimates are correct.

However, looking to the future, the SST CCI analysis could yield more information for model assessment if its estimate of observational uncertainty could be combined with an estimate of inter-decadal variability in order to give an indication of the bounds beyond which the simulation is deemed to be significantly different from observations. This is because the 20-year means of the model simulation and the observational analysis are not representative of the same period in time, since we are assessing control simulations. In order to correctly propagate uncertainties when calculating time- and area-averages, information about the correlation structure in the uncertainties in the SST CCI analysis product would need to be provided. This will be fed in to the first revision of the User Requirements Document for SST CCI Phase II.

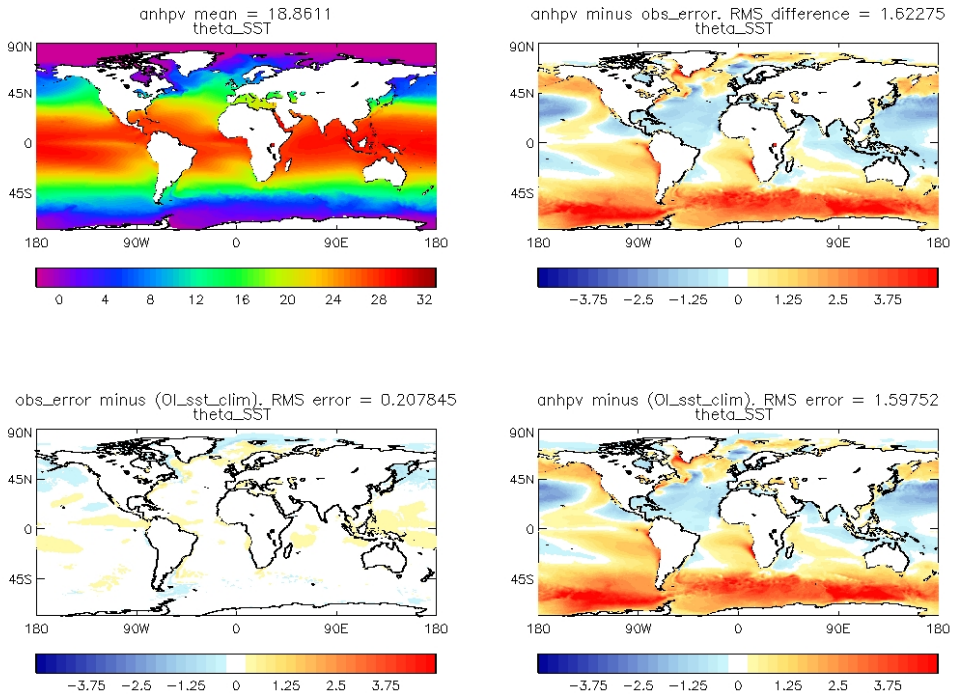


Figure 4-4. Top left: Annual 20–year mean SST (°C) simulated in GC2. Top right: SST in GC2 minus SST CCI analysis. Bottom left: SST CCI analysis minus Reynolds et al (2007) Daily OI. Bottom right: SST in GC2 minus Daily OI.

Table 4-1. Mean SST (°C) in GC2 and GC1 simulations, SST CCI Analysis and Reynolds et al (2007) Daily OI and rms differences between combinations of these.

Area	GC2 value	GC1 value	SST CCI analysis	Daily OI	RMS GC2 – SST CCI analysis	RMS GC1 – SST CCI analysis	RMS GC2 – Daily OI	RMS GC1 – Daily OI	RMS SST CCI analysis – Daily OI	RMS GC2 – GC1
Global	18.86	18.93	18.27	18.22	1.62	1.53	1.60	1.51	0.21	0.33
30N-30S	26.14	26.31	26.24	26.09	0.86	0.78	0.81	0.77	0.20	0.29
30N-90N	10.64	10.78	10.61	10.76	1.26	1.17	1.21	1.08	0.29	0.52
30S-90S	9.45	9.301	7.02	7.02	2.78	2.65	2.76	2.63	0.15	0.23
Atlantic (including Arctic)	17.89	17.89	17.95	18.00	0.94	0.93	0.91	0.90	0.22	0.35
Atlantic 30N-30S	25.30	25.32	25.39	25.27	0.91	0.86	0.89	0.84	0.17	0.21
Atlantic 30N-90N	9.12	9.11	9.15	9.23	1.00	1.04	0.96	0.98	0.27	0.47
Pacific	23.13	23.35	23.12	23.05	1.07	0.94	1.02	0.91	0.24	0.37
Pacific 30N-30S	26.21	26.43	26.28	26.12	0.94	0.85	0.88	0.84	0.22	0.30
Pacific 30N-90N	12.66	12.99	12.55	12.79	1.54	1.31	1.47	1.18	0.31	0.59
Indian Ocean	26.39	26.54	26.52	26.39	0.46	0.39	0.44	0.44	0.17	0.31

4.1.3 ASSESSMENT OF MODEL VARIABILITY

4.1.3.1 Data

The model data is a 70 year HadGEM3 N216 ORCA025 present day control simulation[3]. For comparison with the observations, a 19-year block of data was arbitrarily chosen from close to the end of the control run to avoid any problems associated with model spin-up. This analysis uses global surface temperature fields, which are a suitable substitute for assessing SST provided that the data are carefully masked. Model data are presented on a 0.833° longitude by 0.555° latitude grid and comprise 360 day years.

The observational datasets used in this study are the SST CCI Analysis product and the Reynolds et al [2007, RD.76] Daily OI, which are spatially complete daily analyses that are readily comparable with the model data. Data covering 1992-2010 are considered here. SST CCI analysis data from September-December 2001 are not used so as to ensure our analysis is not seasonally biased. Daily OI data are presented on a 0.25° longitude by 0.25° latitude grid.

4.1.3.2 Methods

All datasets were first re-gridded to a 1° longitude by 1° latitude grid to facilitate comparison. SST CCI analysis error fields were also re-gridded under the assumption that the errors are uncorrelated, which is consistent with current SST CCI L4 guidance. Once regridded, a global daily climatology was constructed for each dataset by simple averaging of grid cell values for each calendar day, followed by a smoothing of the resultant seasonal cycles using a 60-day low-pass filter. This latter step was necessary to reduce sub-seasonal noise resulting from the averaging of relatively short, 19-year records. A 60-day cut-off period was chosen on the basis of a short investigation of sample seasonal cycles produced using various low-pass filter cut-off periods. Climatologies were then subtracted from the SSTs to produce daily anomalies. Finally, gridded fields of standard deviations were calculated for each dataset for both anomalies and absolute SSTs so that variability could be assessed. Prior to calculating standard deviation, each grid cell's SST time series was first run through a 5-year high-pass filter to remove longer term trends. This helps to ensure a fair comparison is made between observed and model data, as the latter does not include time-variable greenhouse gas forcing. When inter-comparing observed and model data, model grid cells contaminated by land or sea-ice surface temperatures were masked in both sets of data, as were any observed grid cells containing sea-ice. This ensures a fair inter-comparison is made using only open-ocean grid cells. Any lake temperatures in the observational data were also masked.

4.1.3.3 Discussion of results

4.1.3.3.1 SST Anomalies

Figure 4-5 compares global maps of SST anomaly variability (assessed using standard deviations) for the model, SST CCI analysis and Daily OI datasets. The large scale spatial patterns and magnitude of variability broadly compare well between datasets, with enhanced variability obvious in strong frontal/boundary current regions and their downstream extensions, and in the equatorial Pacific. A closer inspection of the model fields however reveals somewhat damped variability in these (as well as some coastal) regions, when compared with observational data. This is emphasised in Figure 4-6, which shows zonal averages of the variability, and in which reduced model variability in

[3] *This particular model was used in the development of the GC2.0 version of HadGEM3 (publication pending) and includes virtually all the code included in GC2.0.*

the equatorial, and particularly frontal mid-latitudes, is clear. Figure 4-6 also shows that the magnitude of the model variability is generally in better agreement with the SST CCI analysis than with the Daily OI, and that the Daily OI variability is generally enhanced relative to the SST CCI Analysis, with the exception of the equatorial latitudes (which is a consequence of a pool of relatively reduced variability in the central and eastern equatorial Pacific, see Figure 4-7). This enhanced variability in the Daily OI likely arises partly because it is less representative of the daily average than the SST CCI analysis, which assimilates satellite observations processed to a consistent depth and overpass time to produce an estimate of daily average SST. By not accounting for diurnal variability, the Daily OI may include aliased diurnal effects which manifest as extra noise, though other factors such as accuracy of input data and input data types may also be important. An intercomparison of analyses carried out for the SST CCI PVIR (their Section 6.3.1; RD.330) showed that globally the RMS difference between the SST CCI analysis and Argo measurements is smaller than that of the Daily OI (using AVHRR only). This illustrates the value of the SST CCI products for this sort of analysis.

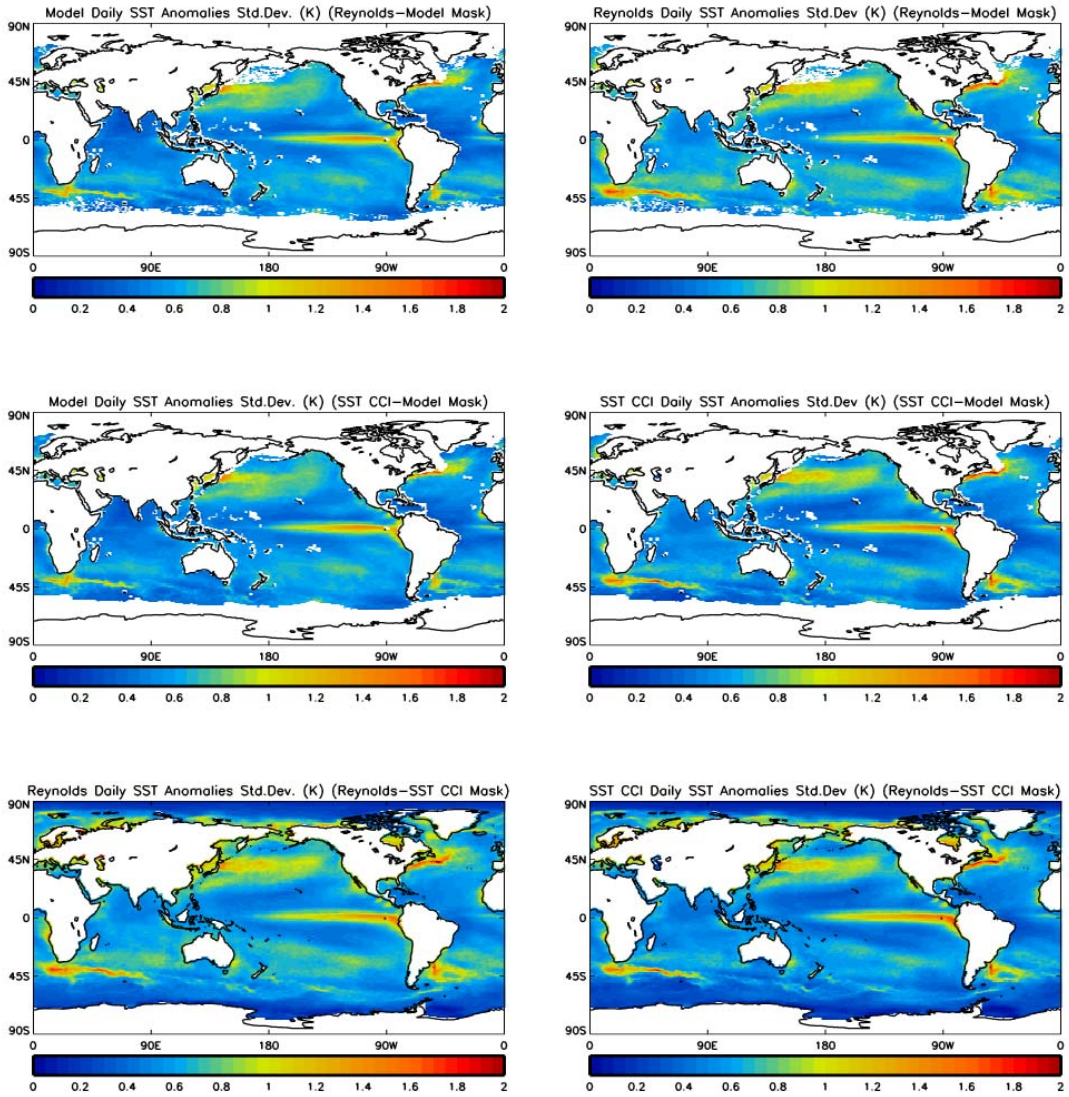


Figure 4-5. Daily SST anomaly variability (standard deviations, K) in a HadGEM3 model control run, the SST CCI analysis, and the Reynolds et al (2007) Daily OI. All datasets have been regridded to a 1° grid. Comparisons are: (top) model with Daily OI; (middle) model with SST CCI analysis; (bottom) Daily OI with SST CCI L4. In each row datasets have been masked as appropriate.

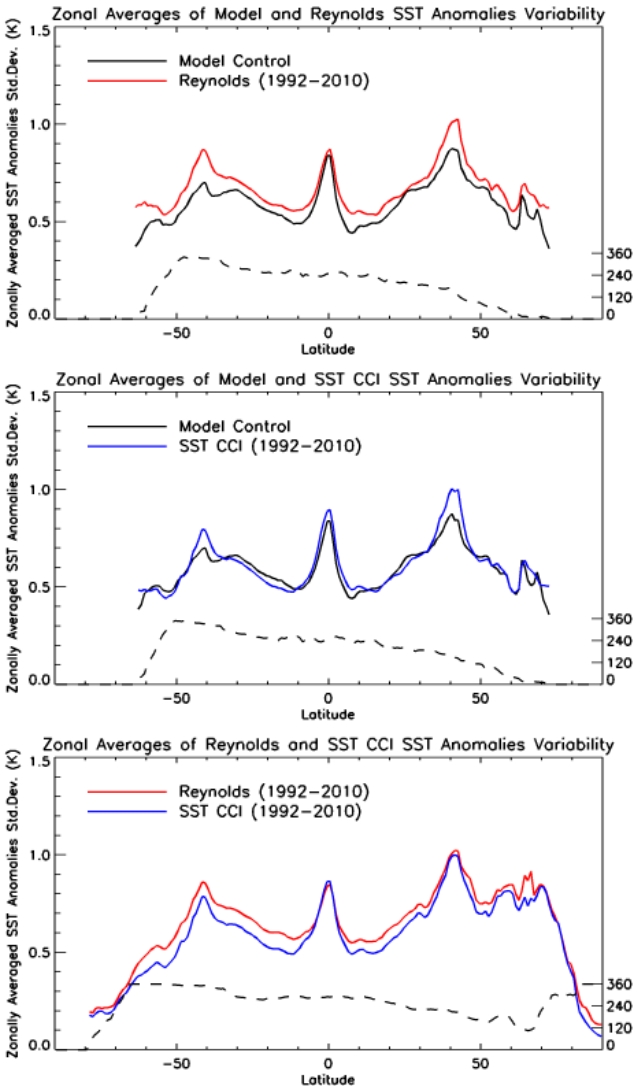


Figure 4-6. Zonal averages of the SST anomaly standard deviations (K) presented in Figure 4-5. Dotted lines denote the number of grid cells used to create the zonal averages.

Figure 4-7 shows the percentage difference in variability between the datasets. The unmasked data in the left column highlight greater spatial structure in the comparisons than is obvious from Figure 4-5. Broadly these plots show that model variability is enhanced relative to observations in the subtropical and high sub-polar and polar latitudes, and relatively reduced in the tropical and mid latitudes. This general pattern is also clear in Figure 4-8, which shows zonal averages of these percentages, and which also highlights again the generally enhanced variability in the Daily OI. When comparing the Daily OI and SST CCI analysis, apart from the enhanced Daily OI variability, the most obvious features of note are patches of relatively reduced Daily OI variability in the coastal regions of the northern sub-polar latitudes and in a pool extending across the central and eastern equatorial Pacific, and patches of relatively enhanced variability in the Indian and Atlantic sectors of the Southern Ocean, and also in the high polar latitudes. A full explanation for these discrepancies is not attempted here but is likely related to the differing treatment of sea ice in the Daily OI and SST CCI analyses (in the case of the Southern Ocean a careful examination of the masked grid cells in Figure 4-7, top left panel, shows that patches of Daily OI sea ice extend further north than for the SST CCI analysis, shown in the panel immediately below). In the case of the equatorial Pacific, a brief investigation of SST CCI analysis and Daily OI anomaly time series suggests that

SST CCI analysis variability is relatively enhanced throughout 1992-2010, which is particularly apparent as higher peak temperatures during the 1998 El Niño event (a similar result seems evident for the Nino 3.4 indices shown in Figure 3-8).

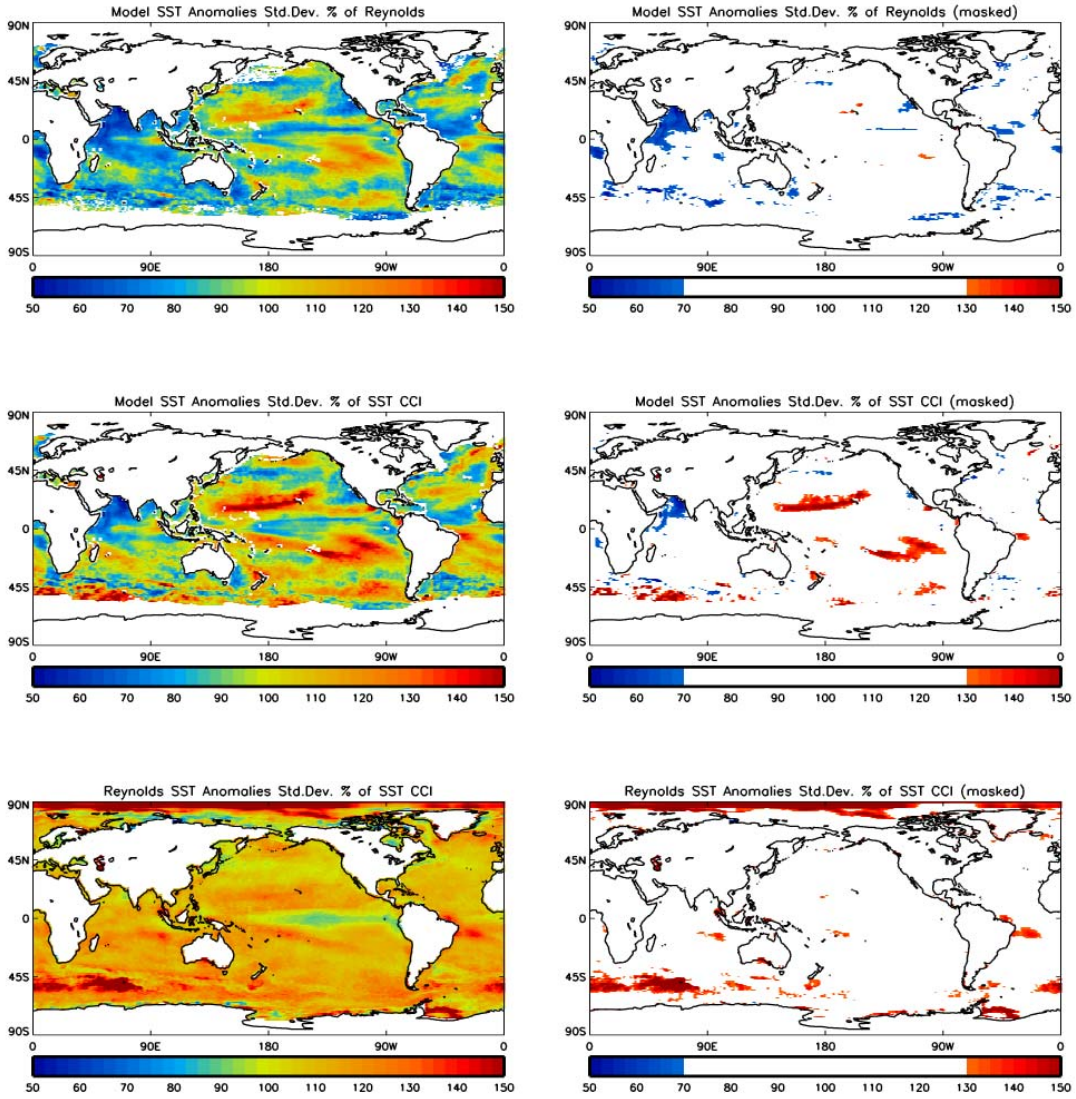
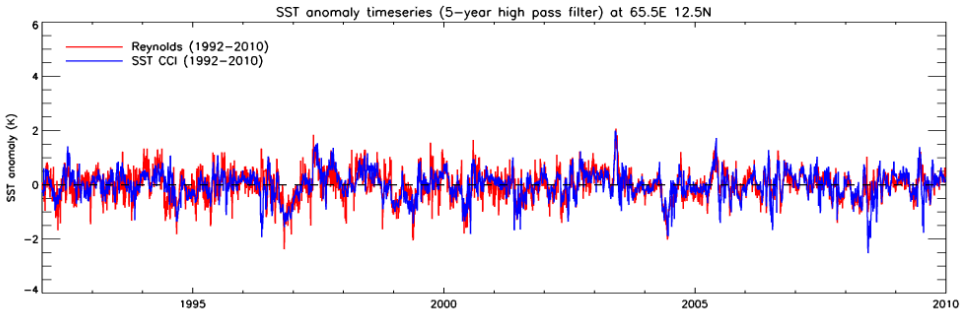


Figure 4-7. Left: Intercomparison of the daily SST anomaly variability shown in Figure 4-5; this is presented as percentage differences between two fields. Top: model variability as a percentage of Daily OI variability. Middle: model as a percentage of SST CCI analysis. Bottom: Daily OI as a percentage of SST CCI analysis. Right: identical to the left column except that percentages falling in the range 70-130% have been masked.

As part of the standard assessment of model simulations against observations, simulated interannual SST variability is compared with observations to see where it falls outside the 70-130% range, which is considered undesirable. As a standard assessment of simulated sub-interannual variability is not currently carried out, it is not clear what equivalent metric should be used for assessing the daily variability presented here, so a 70-130% range is also considered. It is emphasised however that this range may be somewhat arbitrary. The right column in Figure 4-7 is identical to the left column except that grid cells falling in the 70-130% range have been masked. This emphasises that

globally, model variability generally falls within the 70-130% range of the observations. A notable exception is in the Arabian Sea, which is a region where good representation of SST variability may be important for model simulation of the Asian monsoon (Section 4.2.3), and in which model variability appears damped relative to the observations. This may be because the variability in the single-view AVHRR SSTs is exaggerated in the gulf of Arabia owing to biases arising from intermittent desert dust [Merchant et al 2006; RD.357]. Furthermore, an investigation of SST CCI analysis and Daily OI anomaly time series in the central Arabian Sea shows that Daily OI variability is enhanced relative to the SST CCI analysis prior to the introduction of AMSR-E data from 2003, and relatively reduced from 2003 onwards (below). This apparent positive impact of the passive microwave data on the Daily OI relative to the SST CCI analysis suggests that Daily OI variability is exaggerated here prior to 2003, but also that SST CCI analysis variability may be overestimated here too. This region is a particularly complicated one for the retrieval of SST.



Problems with the model simulation are also known to exist here (Section 4.2.3). Given the importance of properly representing monsoonal variability, it is important to undertake further analysis.

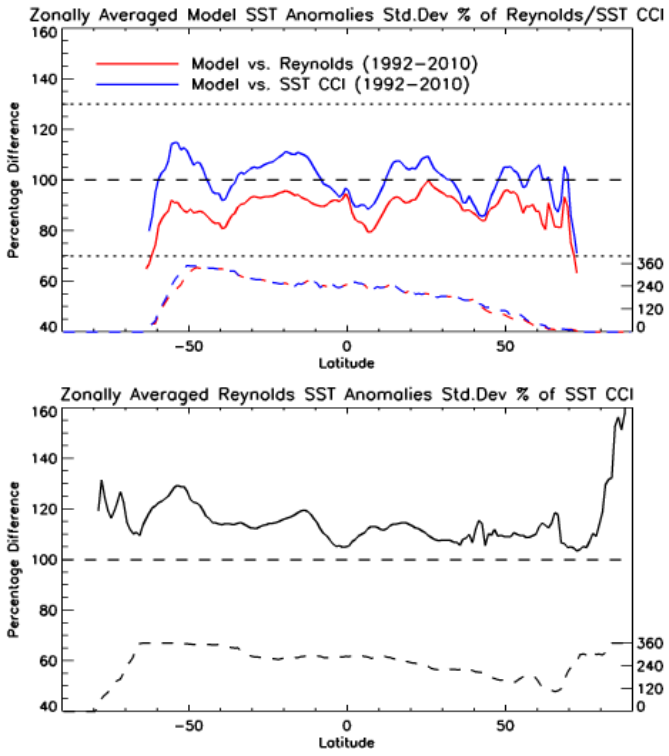


Figure 4-8. Zonal averages of the percentage difference maps presented in Figure 4-7, left column. Dotted lines denote the number of grid cells used to create the zonal averages.

The impact of the SST CCI analysis errors on the above assessment was also considered. In general, the gridded analysis errors are approximately a few orders of magnitude smaller than the variability being considered here, and therefore have no impact on the results described above. This is illustrated in Figure 4-9, top left panel, where zonally averaged SST CCI analysis variability is shown, with and without analysis errors removed. Figure 4-9, top right panel, shows an equivalent result but where analysis errors have been gridded under the assumption that they are entirely spatially correlated in each grid cell. Although current guidance is to treat the analysis errors as uncorrelated, some analysis error correlation may in fact be expected on synoptic and mesoscale length scales, and so it is interesting to explore the limit to which varying the correlation assumption may affect these results. It is clear from Figure 4-9 that for SST CCI analysis SST anomalies, when this correlation assumption is varied, the resulting variability is significantly reduced, and by enough to have an impact on the model-SST CCI comparisons (cf. Figure 4-6, middle panel, noting that the masked grid cells differ in these panels). The details of the model-observation comparison are therefore potentially sensitive to the treatment of analysis errors.

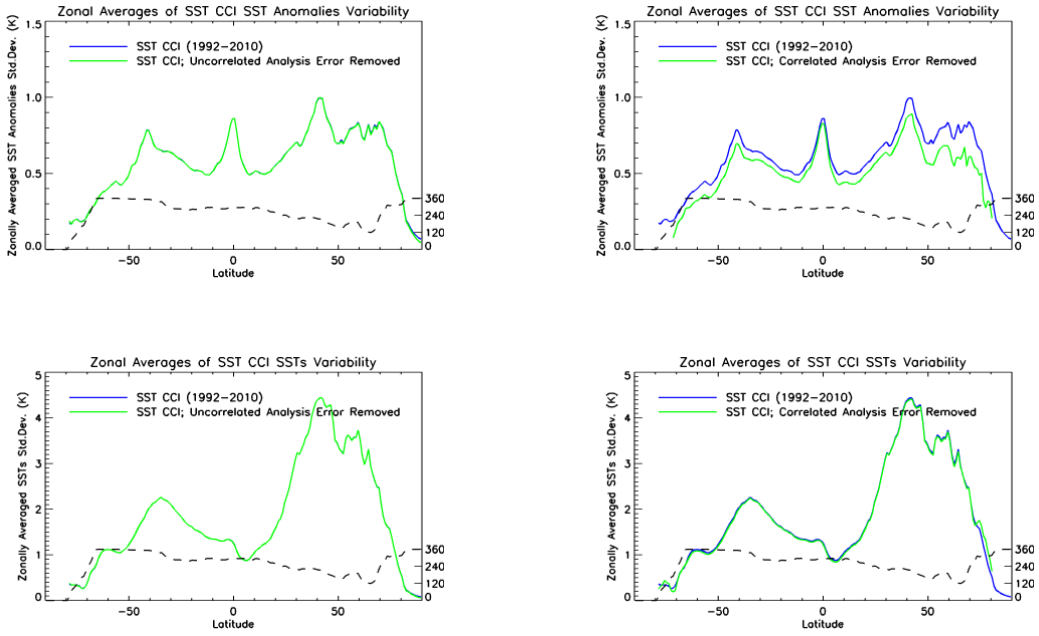


Figure 4-9. Blue lines denote zonal averages of SST anomaly (K, top row) and absolute SST (K, bottom row) variability in the SST CCI analysis, similar to that shown in Figures 4-6 and 4-8. Green lines are equivalent, except that variability associated with the SST CCI analysis error has been removed, under the assumption that grid cell errors are spatially uncorrelated (left column), or entirely correlated (right column). Dotted lines denote the number of grid cells used to create the zonal averages.

4.1.3.3.2 Absolute SSTs

Figures 4-10 to 4-13 are identical to Figures 4-5 to 4-8, discussed above, except showing results for daily absolute SSTs, rather than daily SST anomalies. Initially this investigation was partly motivated by the desire to assess the impact the climatology methodology may have on the results, however this proved difficult to assess away from the tropics where absolute SST variability becomes dominated by a large seasonal signal and so cannot be readily compared with SST anomaly variability (cf. Figure 4-5 and Figure 4-11). As for the SST anomaly results, the spatial patterns and magnitude of the

absolute SST variability broadly compare well between model and observations (Figures 4-11 and 4-13). There is some evidence that the model seasonal cycle may be generally too large, particularly in parts of the subtropical Pacific and Indian Oceans, and also the Southern Ocean (Figures 4-11 and 4-12). This seems consistent with evidence of seasonal biases found in the HadGEM3 mean state in the North Pacific and Southern Ocean described in Section 4.1.2. Also of note is that for absolute SSTs, the generally larger global variability relative to that for SST anomalies means that the enhanced Daily OI variability relative to that of the SST CCI analysis (described above) is no longer evident in the results (cf. Figures 4-8 and 4-11). The only notable exception is in the Arctic (and parts of the Antarctic, particularly the Weddell Sea) where large differences in variability are still evident, presumably due to the differing treatment of SSTs in the presence of sea ice. It is also noted that the greater magnitude of the absolute SST variability means that the SST CCI analysis errors (however their correlation is treated) do not have any meaningful impact on the results (Figure 4-9, bottom row).

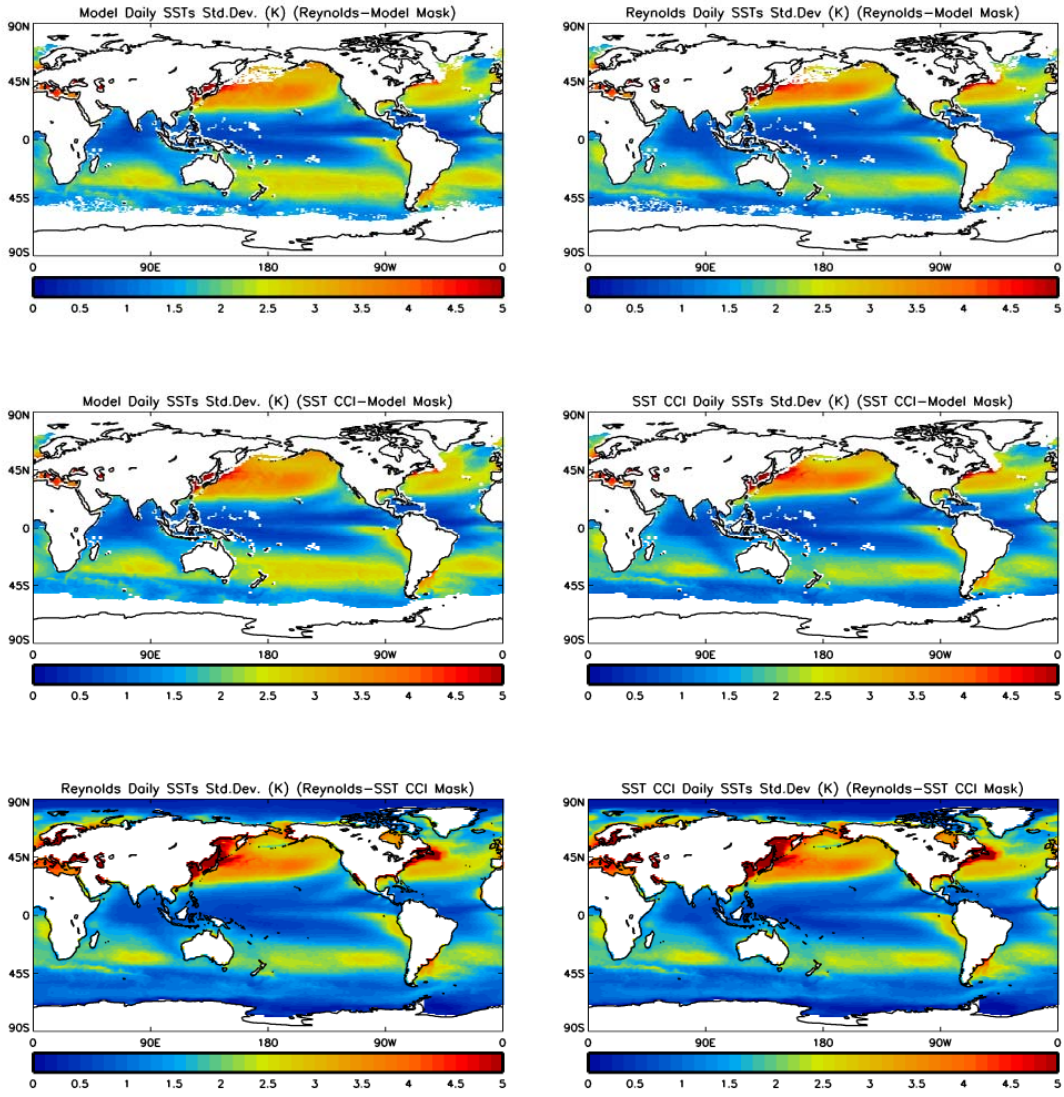


Figure 4-10. As for Figure 4-5 except showing absolute SSTs

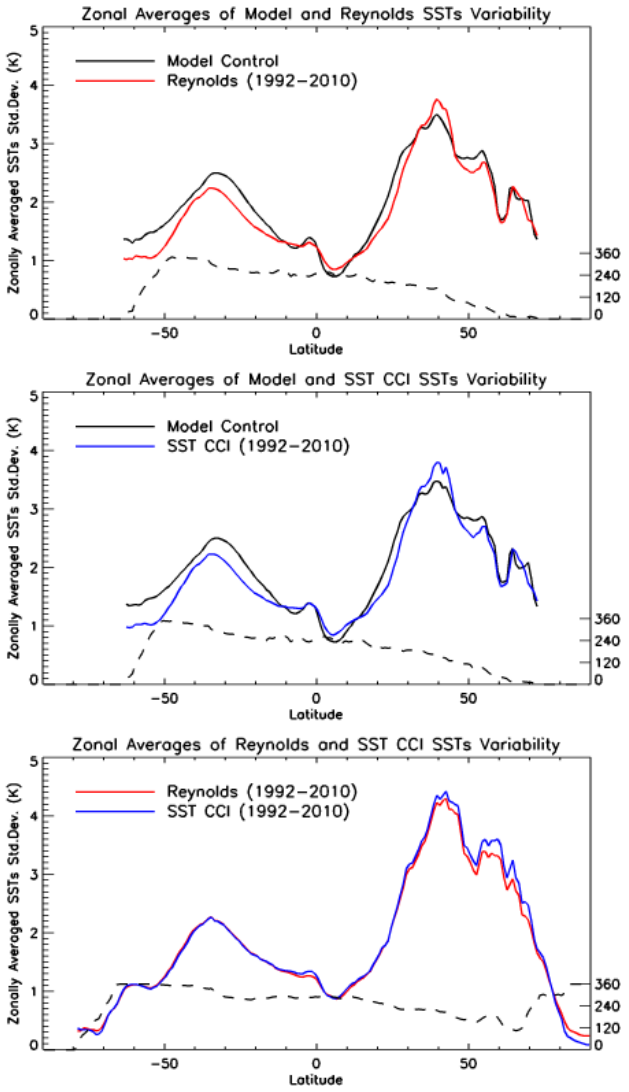


Figure 4-11. As for Figure 4-6 except showing absolute SSTs.

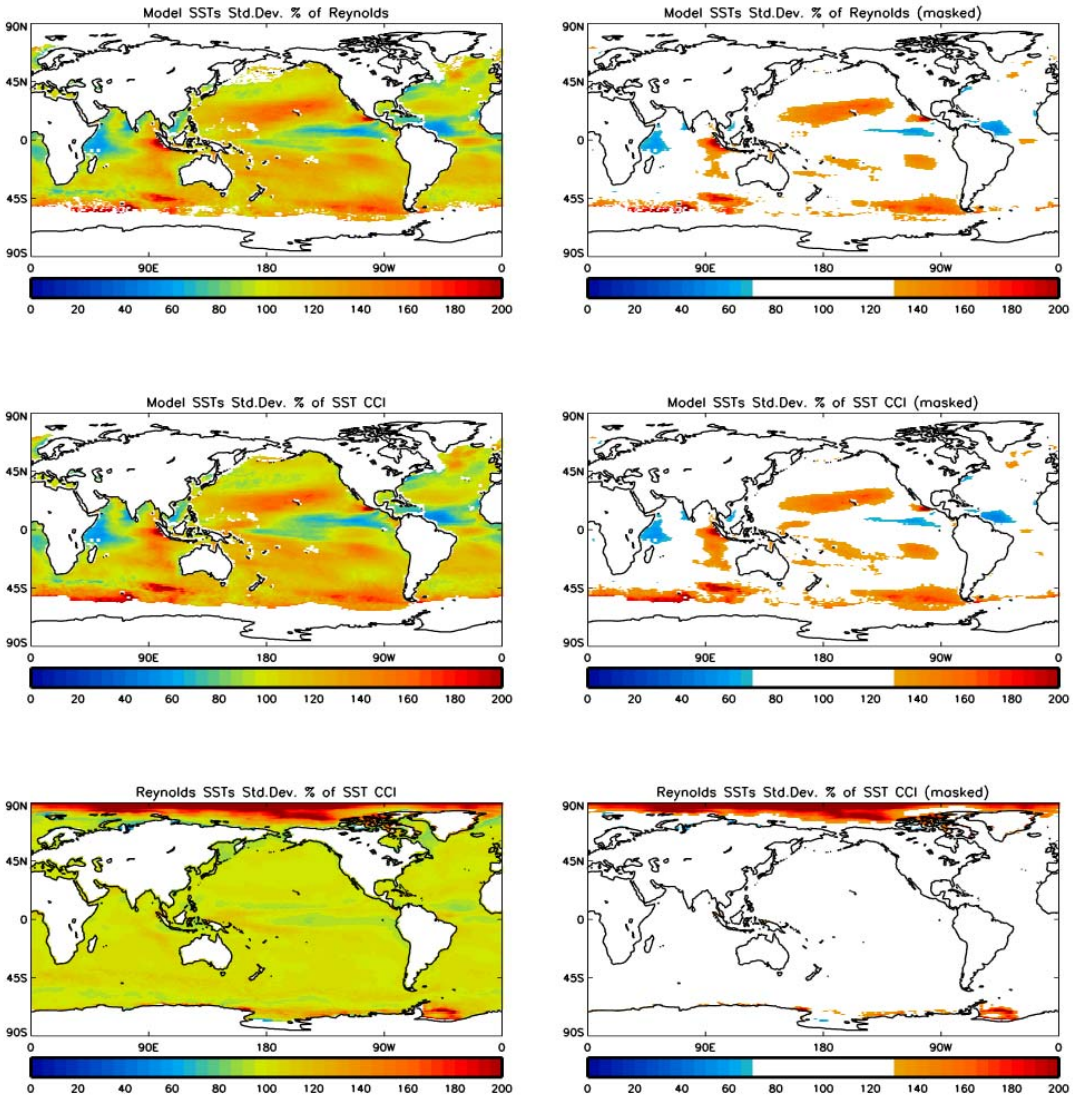


Figure 4-12. As for Figure 4-7 except showing absolute SSTs.

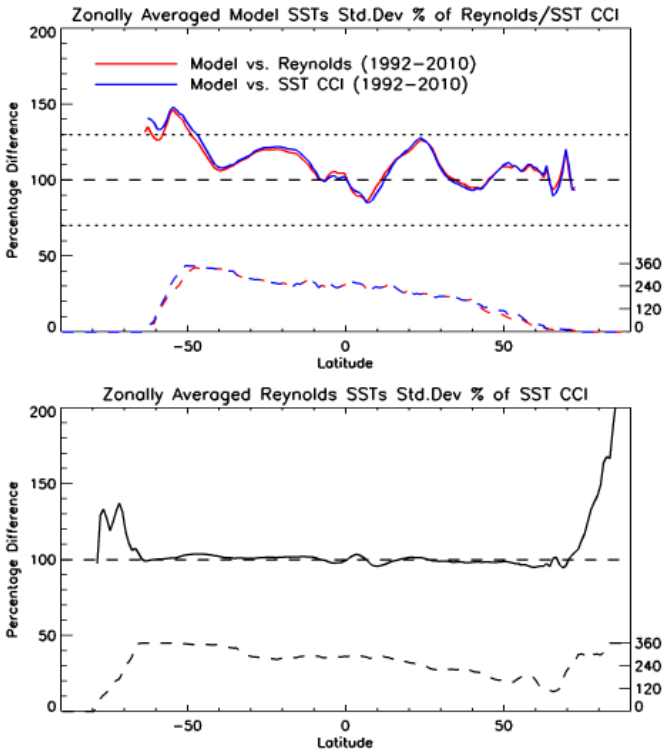


Figure 4-13. As for Figure 4-8 except showing absolute SSTs.

4.1.3.4 Discussion

The results presented here demonstrate a reassuring general agreement between the model and observed daily SST variability, both in terms of magnitude of the variability, and its spatial distribution. Some known model deficiencies are apparent in the results, e.g. the relatively reduced variability of model data in frontal regions because of the non-eddy-resolving grid used by the ocean model. No attempt has been made here to interpret the finer details of the discrepancies in the observed and model variability, which could be related to various sources such as satellite bias corrections, cloud clearing, model biases or decadal variability (e.g. decadal changes in the mean position of ocean fronts in the model control simulation or the ENSO events sampled). This finer detail is therefore not straightforward to attribute, which points to potential benefits from systematic investigation jointly between modellers and observations experts on specific discrepancies. The Arabian Sea is such an example where there may be problems with both model and observed SSTs. The discrepancies in the model-observed SST anomaly variability have also been shown potentially sensitive to assumptions made about how to use the SST CCI analysis uncertainties to adjust variability. Instead, the chief finding of this study is again emphasised, which is that the model and observed fields broadly appear to compare well, with the substantial caveat that standard metrics do not yet exist defining what an acceptable discrepancy between simulated and observed variability is at sub-interannual timescales.

4.1.4 SUMMARY OF KEY POINTS

- The SST CCI analysis is a suitable tool for the evaluation of coupled model mean state.

- Differences between observational estimates are smaller than differences between simulated and observed SST. However, they are of similar size to differences between model versions. Therefore, in the future, it should be

possible to meaningfully use observational uncertainties to assess significance of differences from observations.

- Comparing daily SST variability in a HadGEM3 model control run with two SST analyses (SST CCI analysis and Reynolds et al [2007] Daily OI), shows that the model and observed variability broadly compares well in terms of both magnitudes and spatial patterns.

- The magnitude of HadGEM3 model SST anomaly variability shows a better general agreement with the SST CCI analysis product, demonstrating the potential value of this dataset for this sort of analysis.

- Assumptions made about the correlation structure of analysis errors have an impact on the details of conclusions drawn when comparing simulated and observed variability, so more work is needed to understand the correlation structure of analysis errors.

- Variability in the single-view AVHRR SSTs may be exaggerated in the Gulf of Arabia/Arabian Sea owing to biases arising from intermittent desert dust. There is evidence that there may be a positive impact of the use of passive microwave data in this region. The Gulf of Arabia/Arabian Sea is a particularly complicated one for the retrieval of SST, but also very important for climate modelling. A detailed exploration of representation of this region in SST CCI products should be undertaken in discussion with climate modellers.

4.2 Voluntary reports by the trail blazer users

Recommendations and requests made by the trail blazer users in their reports will be incorporated into the revised User Requirements analysis for SST CCI Phase II,

4.2.1 TIM GRAHAM: EXPLORATION OF HEAT TRANSPORT IN OCEAN MODELS

Tim Graham

Ocean Modelling Group, Met Office Hadley Centre

4.2.1.1 Key messages

- The SST CCI analysis was used as part of a process-based model comparison study looking at heat transport by tropical instability waves in the ORCA025 and ORCA1 ocean models
- Daily mean SST data (as provided by the SST CCI analysis) is much more useful than foundation SST data (such as provided by OSTIA v1) for this purpose because foundation SST is expected (and seen) to be cooler than the daily SSTs simulated by the models.
- Large differences were found between the SST CCI analysis and other analyses during the 1997-1998 El Niño.

4.2.1.2 Scientific analysis

4.2.1.2.1 Aims of the study

Tropical instability waves (TIWs) form in the tropical Pacific and Atlantic Oceans where there is shear between the westward currents on the equator and the eastward currents to the north and south. They can be observed in satellite images of SST [e.g. Legeckis, 1977; RD.358]. It has been shown in models and observations that TIWs transport heat into the equatorial cold tongue [e.g. Menkes et al., 2006, RD.359; Jochum et al., 2007, RD.360].

In this study Graham (submitted to Ocean Modelling) compared the heat transport by tropical instability waves (TIWs) in the tropical Pacific Ocean in the NEMO ORCA025 (1/4°) and ORCA1 (1°) ocean models forced with Drakkar forcing set 4.1 [Brodeau et al., 2010; RD.361]. The aim was to determine whether poor resolution of TIWs could contribute to the worse equatorial cold tongue bias in ORCA1 compared to ORCA025. The difference in the cold tongue was apparent even in these forced model runs where the prescribed surface air temperature prevents SST biases becoming too large.

4.2.1.2.2 Method

Four years of daily mean model output of temperature and velocity was decomposed using a Reynold's decomposition. I.e. for any variable X, $X = \overline{X} + X'$ where $\overline{X}$ is the slowly varying part and X' is the part associated with variability on time scales of 60 days or shorter and therefore dominated by TIWs. The total advective heat transport in the zonal and meridional directions (uT, vT) could then be decomposed in to the part due to TIWs ($u'T', v'T'$) and the slowly varying part ($\overline{u}\overline{T}, \overline{v}\overline{T}$). A time series of SST averaged over the region 110 – 150°W and 4°S-4°N from each model was compared to

the ESA SST CCI analysis long-term product pre-release version. The four years started from 29th September 1994 and therefore covered a La Niña and El Niño period.

4.2.1.2.3 Results

The zonal advective heat transport by TIWs was similar in both model resolutions throughout the four years. During the El Niño period the meridional advective heat transport was also similar between the models but during the La Niña period it was 30% larger in the ORCA025 model than the ORCA1 model. That is more heat being advected into the equatorial cold tongue by TIWs.

The effect of this was seen in the SST time series. During the La Niña period the SST in the ORCA025 model was approximately 0.5°C higher than in the ORCA1 model as a result of the increased advection by TIWs. During El Niño there was little difference between the SST in the two model versions. The mean and RMS errors in SST in ORCA025 were also found to be smaller than in ORCA1.

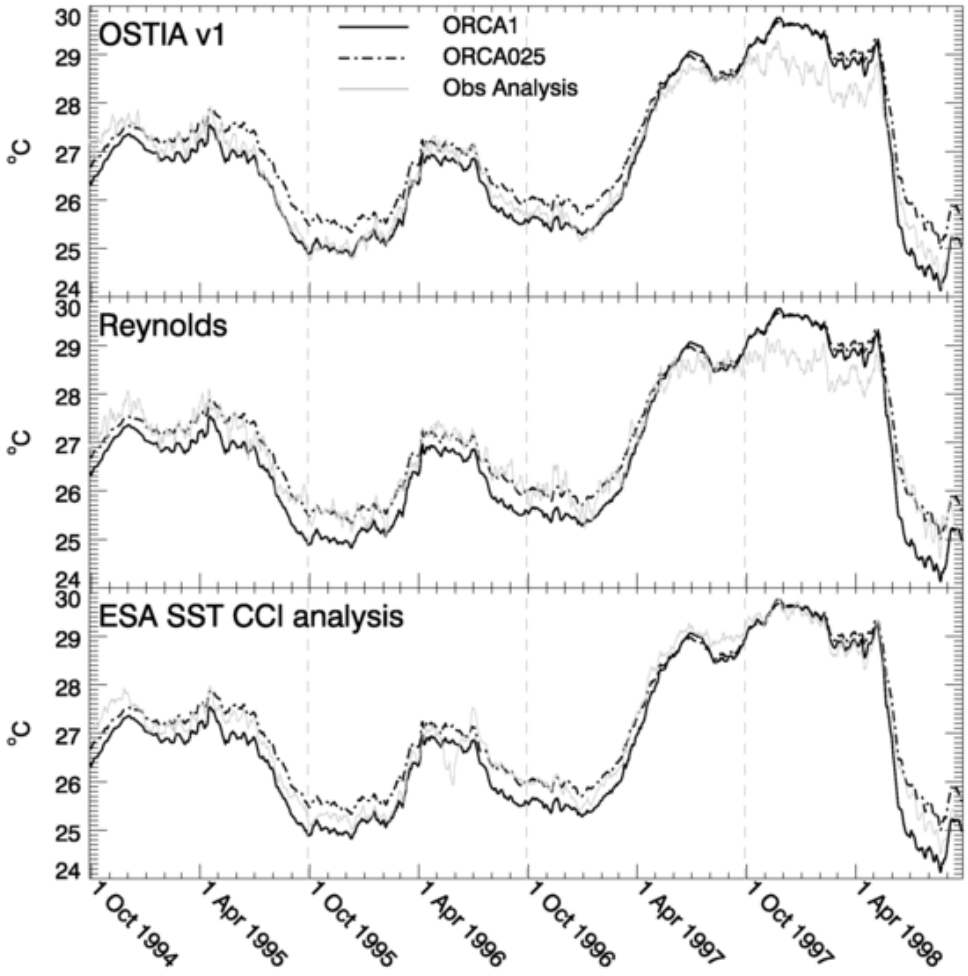


Figure 4-14. Mean SST (°C) over region [110–150°W, 4°S-4°N], 1st October 1994-30th September 1998. Dark lines are repeated in each panel and show the time series as simulated in the ORCA1 and ORCA025 models. Each panel additionally shows the observational time series from: (top) OSTIA Reanalysis v1.0; (middle) Reynolds et al Daily OI and (bottom) ESA SST CCI analysis.

4.2.1.2.4 Conclusions

The study concluded that improved resolution of TIWs plays a role in the reduced equatorial cold tongue bias in the Pacific Ocean in the ORCA025 model compared to ORCA1. This is a complex region, however, and there are many other important processes that could play a role.

4.2.1.3 Feedback on scientific utility of the SST CCI products

The SST CCI analysis was chosen because it was designed to represent the daily mean SST. MyOcean OSTIA Reanalysis v1 was unsuitable for comparison with the model output in this study because it estimates the foundation SST and is therefore expected to be cooler than the daily mean SST output by the model. During the El Niño period it is approximately 1°C cooler than the modelled SST. The Reynolds et al (2007; RD.76) Daily OI was also considered for use in this study. During the La Niña period this produces similar results to the SST CCI analysis but it is also much cooler during the El Niño period. The SST comparison to the models from all three datasets is shown in Figure 4-14. The agreement between the models and the CCI SST analysis during the 1997-1998 El Niño probably indicates that CCI SST is similar to the SST dataset used to force the ERA-Interim reanalysis at that time (the NOAA 2DVAR analysis), since ERA-Interim is used to produce the Drakkar Forcing Set[4].

4.2.2 ROBERT KING ET AL: EUROPEAN NORTHWEST SHELF OCEAN REANALYSIS

Robert King[1], James While[1], Matt Martin[1], Enda O'Dea[2], Rachel Furner[2], John Siddorn[2], Sarah Wakelin[3], Jason Holt[3]

Marine Data Assimilation, OFR&D, Met Office[1]

Ocean Modelling, OCDCC, Met Office[2]

National Oceanography Centre, Liverpool[3]

4.2.2.1 Key messages

- Only SST observations were assimilated in this shelf sea reanalysis.
- The standard deviation of the surface temperature error was markedly lower in the period where ESA CCI SST products were used (relative to the earlier Pathfinder period).
- The use of individual observation error estimates gave a significant reduction in the RMS errors against in situ observations.
- For future applications, it would be useful to also have the observations supplied with sufficient information to reverse the adjustment to a constant reference time-of-day.

---

[4] We believe the 2DVAR analysis used in ERA Interim did not incorporate the Pathfinder data. In general in the analyses in the CAR, Pathfinder data and the analyses that incorporate them tend to be cooler than other comparison data. The relative coolness of the OSTIA v1 and Daily OI in the 1997/8 El Niño is likely in part due to the general bias in the Pathfinder data, although this bias seen during the 1997/8 El Niño is a larger relative bias than generally expected.

4.2.2.2 Scientific analysis

4.2.2.2.1 Aims of the study

We have recently finished running a 27-year reanalysis of the European North-West Shelf (NWS) seas which is being produced for the MyOcean-2 project. This is a joint project between the Met Office and NOC at Liverpool (NOCL). The assessment of the reanalysis products is now underway and is being led by NOCL.

4.2.2.2.2 Method

We have used the Atlantic Margin Model (AMM, O'Dea et al., 2012; RD.362) implementation of the Forecasting Ocean Assimilation Model (FOAM, Blockley et al., 2013; RD.363), used by the Met Office for short-range operational ocean forecasting. This is an s-coordinate model with ~7km horizontal resolution and 50 vertical levels based on the Siddorn & Furner (2013) stretching function. The operational version of the AMM uses NEMO v3.2 (see O'Dea et al. 2012), but for the reanalysis we have upgraded to NEMO v3.4 and included a number of improvements including new European Hydrological Predictions for the Environment (E-Hype, http://www.smhi.se/en/2.575/Hydrology/european-hydrological-predictions-for-the-environment-1.12711) river inputs, and improved Baltic boundary conditions.

The surface boundary conditions are taken from ERA-Interim [Dee et al., 2011; RD.364] while the lateral boundary conditions are provided by existing simulations using our global FOAM configuration (ORCA025). The data assimilation scheme is based upon the 3D-VAR NEMOVAR system [Waters et al., 2014; RD.365] with the inclusion of a new flow-dependent error co-variance specification.

A spin-up run without assimilation was conducted from 1980 to 1983 and a secondary 1-year spin-up with assimilation was then run. The main run covered the period 1985-2012 and a control run without assimilation was run over the same period to allow assessment of model skill.

Unlike in the open ocean, in the shelf seas we currently assimilate only SST, although in future upgrades we plan to include the assimilation of altimeter and profile observations. Here we assimilate SST observations from satellite and in situ sources. We also process profile observations from the EN3 [Ingleby and Huddleston, 2007; RD.66] and NOOS (North-West European Shelf Operational Oceanographic System) data sets to provide independent observations for comparison. Since the reanalysis covers the period 1984-2012, we have used Pathfinder v5.2 SSTs from 1984 to 1991 and ESA SST CCI pre-release version in L2P format from 1991 onward (see Fig. 4-15). However to avoid use of the SST CCI AVHRR data for NOAA 12 (which was a relatively unstable instrument), we have used Pathfinder (which bypassed NOAA 12) until 1996.

Before assimilation all of the SST observations are bias corrected using in situ and (A)ATSR data.

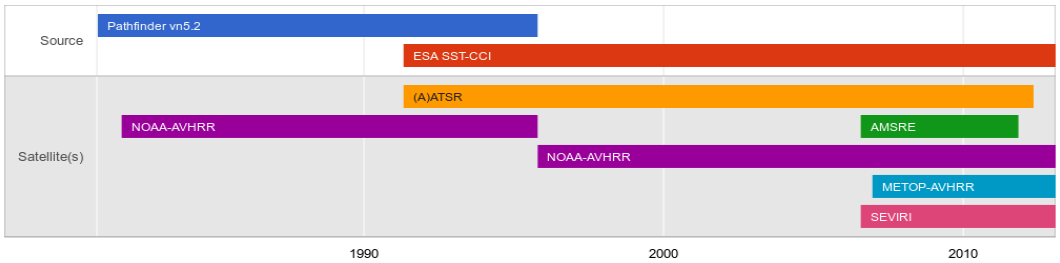


Figure 4-15. SST satellite series coverage and data source over the period of the NWS reanalysis. In the cross-over period (1991-1996) only the (A)ATSR data come from SST CCI.

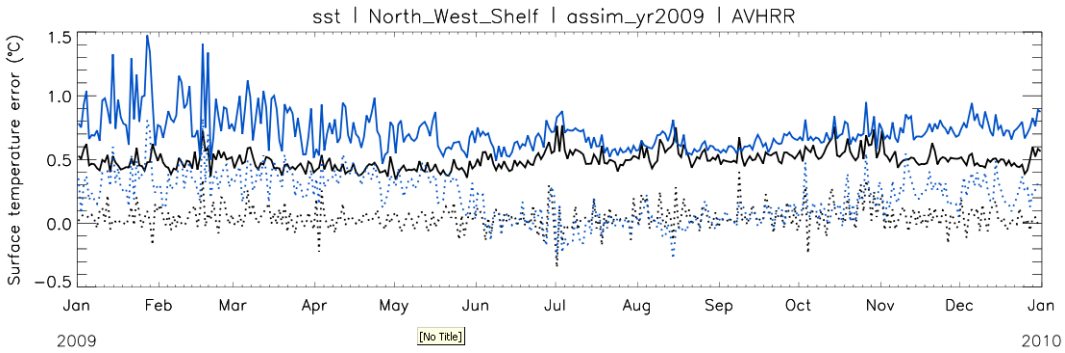


Figure 4-16. Surface temperature errors for AVHRR observations over 1 year. The mean (dashed lines) and standard deviation (solid lines) of the observation-model differences for the assimilative (black) and control runs (blue lines).

4.2.2.2.3 Results

The assessment of the reanalysis runs has only recently begun ahead of the project deadline in March 2014, but our preliminary results indicate some improvements attributable to the SST CCI processing of the observations.

Figure 4-16 shows the surface temperature observation minus model differences for NOAA-AVHRR observations from January to December 2009. These show a significant improvement in both the mean and RMS error from the control (non-assimilating) to the assimilative run. Similarly, Fig. 4-17 shows the temperature errors for a mean profile over the same time period. This indicates an improvement in the temperature error over the upper few hundred metres of the ocean, which is due to the assimilation of SSTs which involves projecting the temperature increments to the bottom of the mixed layer.

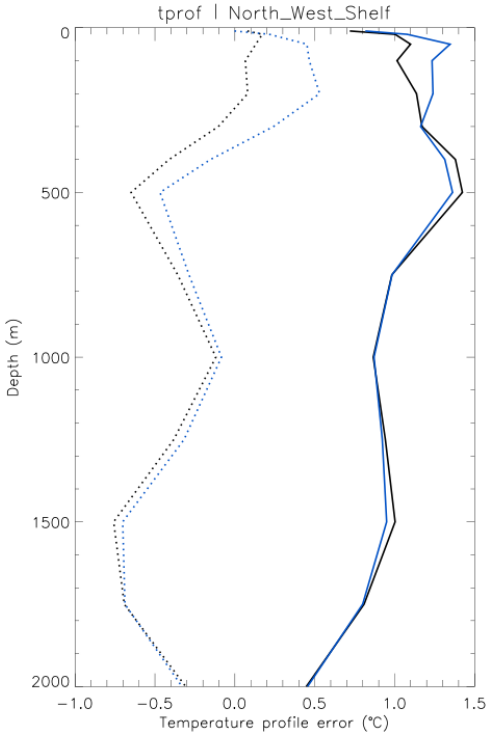


Figure 4-17. Mean profile temperature errors over 1 year. The mean (dashed lines) and standard deviation (solid lines) of the observation-model differences for the assimilative (black) and control runs (blue lines).

4.2.2.3 Feedback on scientific utility of the SST CCI products

Although both correlated and uncorrelated errors on individual observations are supplied, we made use of only the uncorrelated errors. By combining these error estimates with our estimation of the representivity error we found a reduction in the RMS errors against in situ observations compared to the use of a single error estimate for all observations (from 0.27°C to 0.23°C).

Similarly, the inspection of surface temperature error statistics (as estimated by the observation minus model differences) in the reanalysis products showed a reduction in the spread of errors when using the CCI SST products relative to the earlier period where Pathfinder v5.2 was used (see Fig. 4-18: the solid line is closer to zero and more stable after the switch-over, indicating improved standard deviation, and the dashed line has fewer outliers, indicating a more stable mean)[5].

[5] *The number of in situ SST observations assimilated increases gradually over the period of the reanalysis. While this may provide a gradual improvement in the surface temperature error statistics, the change from Pathfinder to SST CCI AVHRR coincides with a marked improvement in the statistics.*

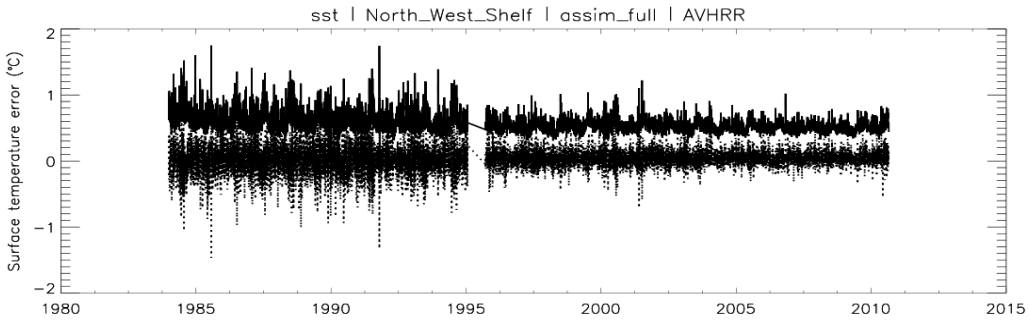


Figure 4-18. Surface temperature error for AVHRR observations over the full reanalysis. The mean (dashed lines) and standard deviation (solid lines) of the observation-model differences for the assimilative run. The observations come from Pathfinder v5.2 from 1984 to 1995 and from ESA CCI SST pre-release version from 1996 onward.

4.2.3 MALCOLM ROBERTS: HIGH RESOLUTION MODELLING

Malcolm Roberts

Met Office Hadley Centre

4.2.3.1 Key messages

- Different SST datasets were used to force the same global atmosphere-land general circulation model
- There are several differences in the simulated mean state that seem to be influenced by differences in the SST datasets, including the Indian monsoon rainfall and surface temperature differences over North America
- Tropical cyclone climatologies are also affected, particularly in the Eastern and Western Pacific regions.
- It is important to have a range of forcing datasets so that model biases can be put into context compared to the model response to different forcing datasets.

4.2.3.2 Scientific analysis

4.2.3.2.1 Aims of the study

To investigate the impact that different SST and sea-ice forcing datasets have on the simulated climate of global models at different horizontal resolutions. There is particular interest in the impact on the mean state of the climate, and aspects of the variability such as tropical cyclone simulation.

We might expect that, as model resolution increases, the simulation may be able to react to small scale SST features that would otherwise be lost when the dataset is averaged onto a relatively coarse model grid.

4.2.3.2.2 Method

SST and sea-ice datasets from a variety of sources are prepared to be used for forcing an atmosphere-land climate model. The datasets used were:

1. ESA SST CCI analysis long-term product pre-release version, which is a satellite-only SST-depth analysis created by the OSTIA system from SST CCI ATSR and SST CCI AVHRR products, produced on a 1/20° grid – henceforth SST CCI analysis.

2. The ¼° Daily OI dataset (Reynolds et al, 2007; RD.76)

3. The HadISST/PCMDI monthly dataset used for AMIP-II simulations (Rayner et al, 2003, RD.74; Taylor et al, 2000, RD.366) on a 1° grid.

Each of the high resolution datasets (SST CCI analysis and Daily OI) is interpolated to the model grid in a simple fashion (since the data is on a higher spatial resolution than the models). For the HadISST/PCMDI data, a more complex interpolation is used so that artefacts are not introduced. The climate models used are at 130km and 25km resolution (as measured at mid-latitudes), which essentially compares a typical CMIP5 resolution climate model (130km), with one more typical of global weather forecasting (25km).

The models are then integrated for about 20 years, 1991-2010, using the standard protocol for such integrations such as the AMIP-II standard protocol (see http://www-pcmdi.llnl.gov/projects/amip/AMIP2EXPDSN/BCS/bcsintro.php) for CMIP3/5.

4.2.3.2.3 Results

4.2.3.2.3.1 Mean state changes

There are some notable differences in the surface temperatures over the ocean between the SST CCI analysis and the other datasets (Figure 4-19). In the Arabian Sea the SST CCI analysis is up to 0.6K cooler than other datasets in June-July-August (JJA) period. This cooler SST causes a significant reduction in the simulated Indian monsoon rainfall in the higher resolution 25km model, magnifying an existing model bias. The other main SST difference is in the Pacific, where there are slightly warmer temperatures near both the Intertropical and South Pacific Convergence Zones, in both winter and summer, of 0.4-0.6K, with a cooler SST in the N Pacific in JJA. Although relatively small, these SST patterns project onto the land surface temperatures, particularly strongly in Dec-Jan-Feb (DJF), with a warming of western N America (by 1-2K) and a cooling of eastern N America by a similar amount. This magnitude of land-surface difference is similar to (but opposite in sign from) that generated by the recent surface warming hiatus (e.g. Kosaka and Xie, 2013; RD.367).

There are also differences in surface temperature near the poles (Figure 4-19), particularly during the local winter – these differences are presumably related to different SST and sea-ice in the different forcing datasets. However, for single model integrations of 20 years, the variability at high latitudes is larger, and so it is difficult to draw strong conclusions.

The use of different SST datasets therefore enables us to better understand uncertainties in the model simulations due to forcing uncertainties – error bars can be put on different SST observational datasets, but these kinds of test can indicate how these differences translate to a model simulated state.

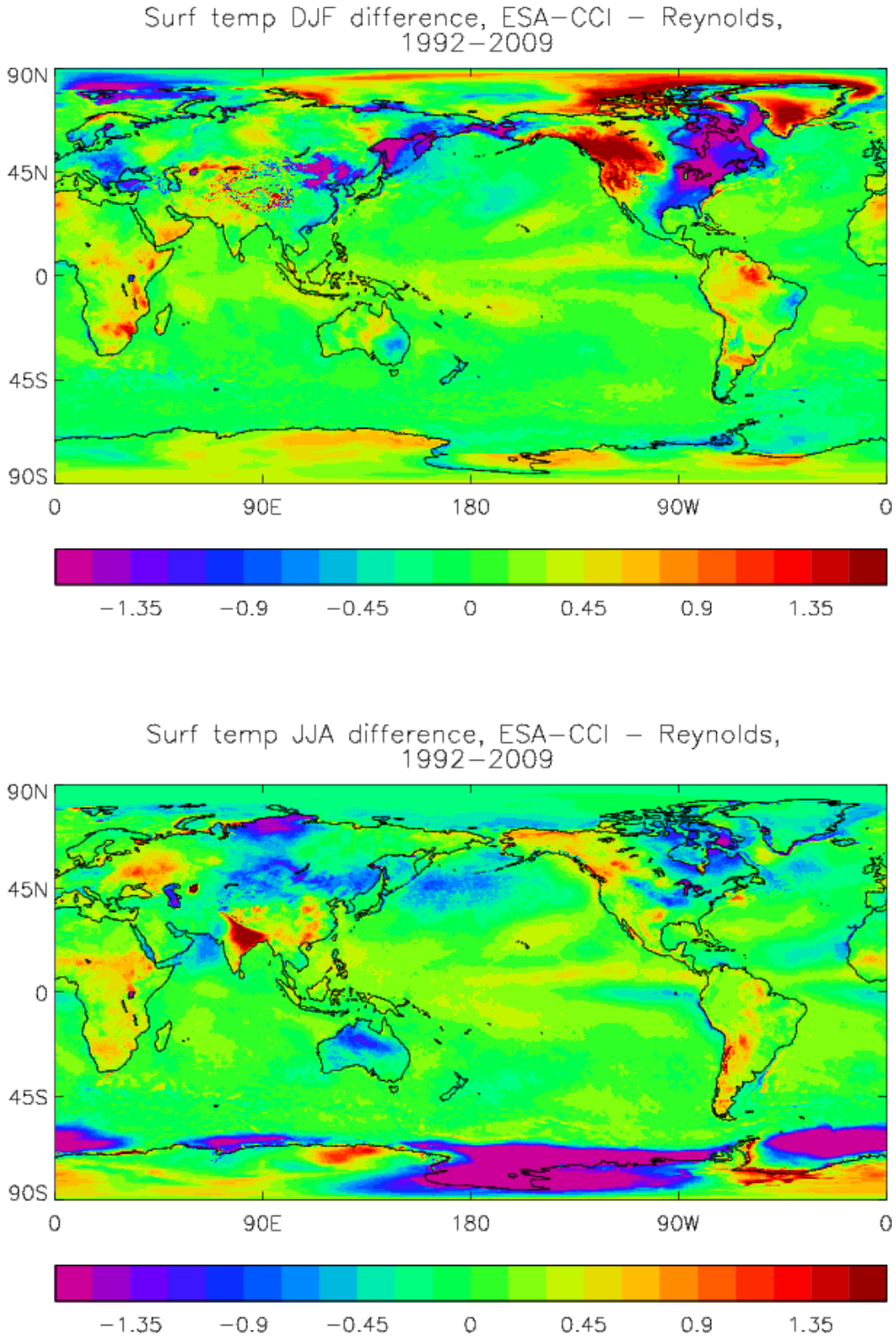


Figure 4-19. Surface temperature difference (°C) in atmospheric model simulations for 1992-2009 driven by ESA SST CCI analysis and Reynolds et al (2007) Daily OI. Top: December-February. Bottom: June-August.

4.2.3.2.3.2 Impact on aspects of variability

In addition to the mean state, aspects of climate variability can also be assessed for sensitivity to forcing dataset. In this case the performance of tropical cyclones has been studied.

In the Pacific there are quite large differences between the tropical cyclone climatologies from models using different SST forcing (Figure 4-20). For the SST CCI analysis, the tropical cyclone tracks from the 25km model in the East Pacific are shifted slightly southwards, elongated and are slightly more frequent, possibly related to the warmer tropical SSTs in this region in JJA. There is also an increase in the frequency of tropical cyclone tracks in the West Pacific, perhaps influenced by the SST increase, though this is only about 0.2K on average.

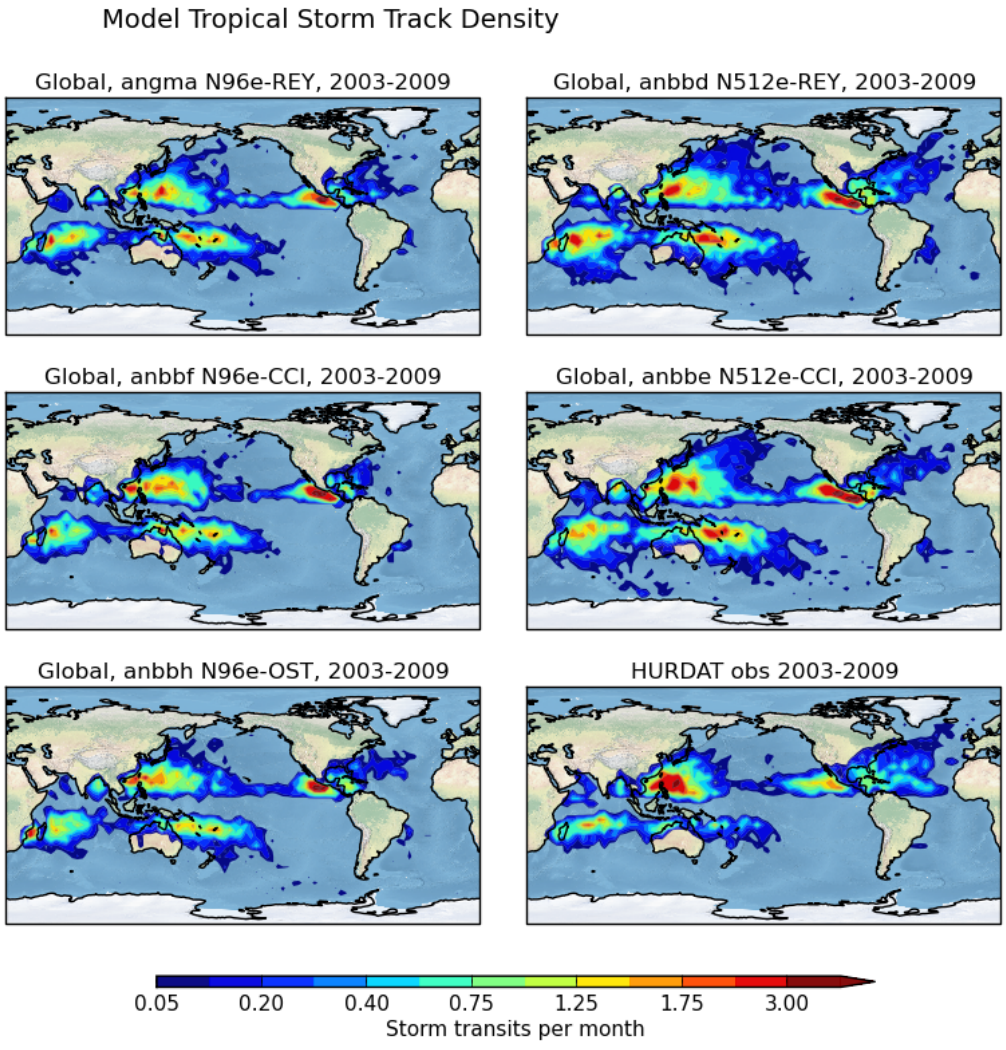


Figure 4-20. Tropical storm track density in different model simulations and observations. Left hand column: Model simulations at ~130km resolution. Right hand column: Model simulations at ~25km resolution and observations (bottom right). Atmosphere-only model simulations were driven by different SST analyses. Top: Reynolds et al Daily OI. Middle: SST CCI analysis. Bottom: OSTIA v1.0 reanalysis.

4.2.3.2.4 Conclusions

It is clear that the different observational datasets for sea surface temperature and sea-ice, which are used to force a climate model in atmosphere-land forced mode, can have a significant impact on the mean state of the simulated climate. This can give some insight into how aspects of model internal variability (weather) can be judged in comparison to the range of responses to different SST products, and hence improve our understanding of what an "acceptable" range our models should lie in for some metrics.

Future work could examine the variability of the different SST datasets, both at daily timescales, and at intra- and inter-seasonal – there is certainly evidence that some aspects of climate processes, such as the Indian monsoon, are linked to intra-seasonal SST variability.

4.2.4 DUDLEY CHELTON: COMPARISON TO OTHER HIGH RESOLUTION SST ANALYSES

Dudley Chelton and Craig Risien

Oregon State University, College of Oceanic and Atmospheric Sciences

4.2.4.1 Key messages

- The pre-release version of the ESA SST CCI dataset is a significant improvement over the pre-January 2013 Operational OSTIA SST dataset, both in terms of accuracy and spatial resolution.

- There are significant differences between monthly averages of the SST fields of the pre-release version of the ESA SST CCI dataset and the NOAA NCDC Daily OI fields (Reynolds et al, 2007; RD.76) constructed from AVHRR infrared and AMSR microwave satellite observations. Ongoing analysis is trying to determine which, if either, of these two SST products is most accurate.

4.2.4.2 Scientific analysis

4.2.4.2.1 Aims of the study

The aims of the study are to assess whether the pre-release version of the ESA SST CCI dataset is an improvement over the pre-January 2013 Operational OSTIA SST dataset. In particular, does the SST CCI product eliminate a systematic trend error of the Operational OSTIA SST fields identified from our previous analysis that consisted of a large cold bias relative to tropical buoy measurements at the April 2006 start of the OSTIA data record that decreased in magnitude by about 0.2°C by December 2007 and remained approximately constant thereafter? The constant post-December 2008 relative cold bias is likely the difference between the "foundation temperature" of the Operational OSTIA product and the mooring measurement depths of about 1m.

Another aim of the study is to assess the spatial resolution of the SST CCI product compared with the pre-January 2013 Operational OSTIA SST product and the NOAA NCDC Daily OI SST product.

A third aim is to determine whether there are significant differences between the three SST products (SST CCI, pre-January 2013 Operational OSTIA, and NOAA NCDC Daily OI SST) for scientific applications to study short-term climate variability (time scales of a month and longer).

4.2.4.2.2 Method

The aims of the study summarized above have been addressed from a variety of statistical analyses of the three SST products (the pre-release version of the ESA SST CCI, pre-January 2013 Operational OSTIA, and NOAA NCDC Daily OI SST), including:

- Comparisons of monthly averages of the three SST products with monthly averages of measurements by ~100 moorings in the TAO, PIRATA and RAMS tropical mooring arrays.
- Comparisons of monthly averages of the three SST products with each other.
- Visual inspection of maps of monthly averages of SST and SST gradient magnitude from the three products.
- Zonal wavenumber spectral analysis of monthly averages of SST from the three products. These spectra were computed separately for wintertime and summertime.
- Global animations of differences between monthly averages of SST from the three products.
- Global maps of the mean, standard deviation and RMS of differences between monthly averages of SST from the three products.
- Global empirical orthogonal function analysis of differences between monthly averages of SST from the three products.
- Global comparisons of seasonal cycles of SST from each of the three products by least-squares regression onto an annual plus a semiannual cycle at each grid location.
- Global maps of the positive and negative extrema of differences between monthly average anomaly SST from the three products, where “anomaly” is defined to be the deviation from the seasonal cycle for each month.
- Global map of the cross correlations between monthly average anomaly SST from the three products.
- Comparisons of two climate indices computed from monthly averages of the three products (El Nino and the Pacific Decadal Oscillation).

4.2.4.2.3 Results

- The systematic trend of errors of the Operational OSTIA SST product from April 2006 to December 2007 is eliminated in the pre-release version of the ESA SST CCI product.
- Based on comparisons with tropical mooring measurements, the pre-release version of the ESA SST CCI analyses are about 0.2°C warmer than the pre-January 2013 Operational OSTIA SST analyses in the tropics. They are warmer than the NOAA NCDC Daily OI SST analyses by about 0.1°C in the tropics.
- Outside of the tropics, the pre-release version of the ESA SST CCI analyses appear to be warmer than pre-January 2013 Operational OSTIA SST analyses by a smaller amount of about 0.1°C.
- The pre-release version of the ESA SST CCI product has higher spatial resolution than either the pre-January 2013 Operational OSTIA SST product or the NOAA NCDC Daily OI SST product.
- From animations of the differences between the pre-release version of the ESA SST CCI product and the NOAA NCDC Daily OI SST product, it is apparent that differences of

~0.5°C with scales of ~1000 km and persistence of 2-3 months occur randomly in space and time throughout the World Ocean.

- The differences between monthly averages of the pre-release version of the ESA SST CCI product and the NOAA NCDC Daily OI product are disturbingly large. About 37% of the grid points globally have monthly average differences with a standard deviation larger than 0.2°C and about 10% of the grid points have a standard deviation larger than 0.4°C.
- Large differences between monthly averages of the pre-release version of the ESA SST CCI product and the NOAA NCDC Daily OI product occur throughout the World Ocean. Extrema values of positive and negative differences both exceed 0.3°C nearly everywhere and exceed 0.6°C in regions of energetic SST variability (e.g., the eastward extensions of all of the western boundary currents, the equatorial cold tongues of the eastern tropical Pacific and eastern tropical Atlantic, and several other regions.
- The dominant EOF of differences between monthly averages of the ESA SST CCI product and the NOAA NCDC Daily OI product consists of large differences in the ice covered regions of the northern hemisphere, the ITCZ of both the Pacific and the Atlantic, and the southeastern tropical Pacific off the coast of Peru.
- The second EOF of differences between monthly averages of the ESA SST CCI product and the NOAA NCDC Daily OI product consists of large differences over most of the subpolar North Pacific, the tropical Atlantic north of the equator and the Arabian Sea. The latter two regions may be related to effects of Sahara dust storms while the former may be related to problems with persistent clouds over the subpolar North Pacific.
- The cross correlations between monthly averages of the pre-release version of the ESA SST CCI product and the NOAA NCDC Daily OI product exceed 0.95 at most grid points, but about 20% of the grid points of cross correlations less than 0.9 and about 4% of the grid points have cross correlations less than 0.8.
- The large-scale climate indices (El Nino and the Pacific Decadal Oscillation) computed from monthly averages of each of the three SST products do not differ significantly from each other.

4.2.4.2.4 Conclusions

- The pre-release version of the ESA SST CCI analyses is a significant improvement over the pre-January 2013 Operational OSTIA SST analyses, both in terms of accuracy and spatial resolution.
- The pre-release version of the ESA SST CCI analyses are warmer than the pre-January 2013 Operational OSTIA SST analyses. These differences are likely attributable to the reference depth of about 0.2 m (the “drifting buoy depth”) for the ESA SST CCI analyses compared with the deeper “foundation depth” for the pre-January 2013 Operational OSTIA SST analyses.
- The magnitudes of the differences between the pre-release version of the ESA SST CCI analyses and the NOAA NCDC Daily OI SST analyses are larger than expected. The reasons for these differences are not yet known. In part, they may be attributable to effects of persistent cloud cover (the North Pacific) and Sahara dust storms (the northern tropical Atlantic and the Arabian Sea) on the satellite SST retrievals. However, unacceptably large differences occur throughout the World Ocean.

- Ongoing analysis is aimed at trying to determine which, if either, of the two products (the pre-release version of ESA SST CCI product or the NOAA NCDC Daily OI product) is most accurate.

4.2.4.3 Feedback on scientific utility of the SST CCI products

Thus far, we have not used the data for scientific analysis. Our efforts have focused on the degree of similarity and differences between the various GHRSST products for potential applications to investigate short-term climate variability with time scales of a month and longer.

4.2.5 JACOB HOYER: USE IN AN ANALYSIS FOR THE ARCTIC

Jacob L. Høyer

Center for Ocean and Ice, Danish Meteorological Institute

(Note: although part of the SST CCI team, here Hoyer reports work undertaken as a trail-blazer user on a voluntary basis.)

4.2.5.1 Key messages

- Very easy to access and use the SST CCI AVHRR data sets
- Quality 5 pixels with missing values may be confusing when using the data. (Note: this problem has been addressed in the for-release SST CCI products.)
- SST CCI AVHRR L3 SST data better than Pathfinder 5.1, years 2006-2008
- SST CCI AVHRR L3 SST showed cold summer bias relative to reference data and seasonal variation
- L4 DMI-OI showed improvements for 2006-2008 using CCI data

4.2.5.2 Scientific analysis

4.2.5.2.1 Aims of the study

The aim of this study is to assess the impact using the L2P ESA SST CCI AVHRR Long-term product pre-release version (SST CCI AVHRR). The SST CCI AVHRR products have been used to produce three years of a test data set. The test data set has been compared to a reference data set, which is obtained from an Arctic level 4 climate dataset based upon Pathfinder AVHRR observations.

4.2.5.2.2 Method

The level 4 processing chain is indicated in the diagram below:

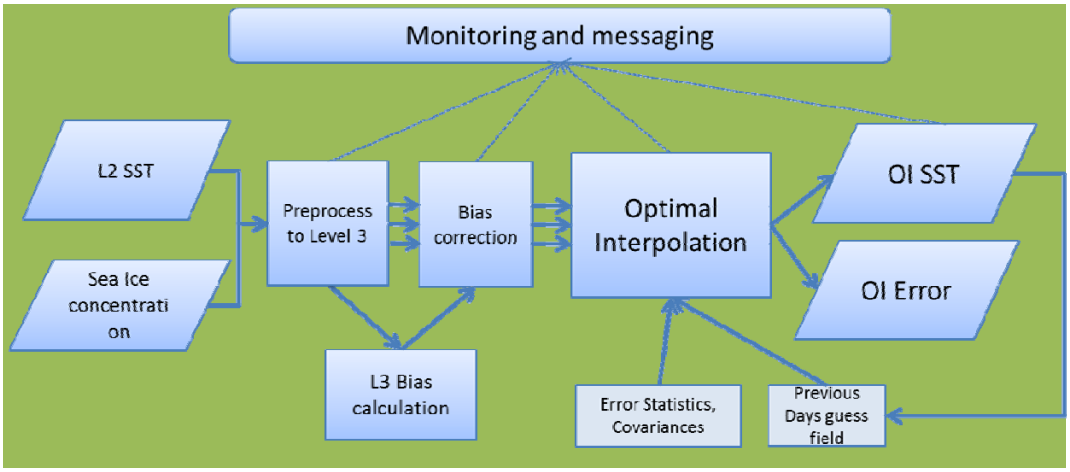


Figure 4-21. Arctic analysis processing chain.

The following inputs are used as input to the reference processing system, called DMI-OI:

(A)ATSR Reprocessing for Climate (ARC). The AATSR Reprocessing for Climate (ARC) SST level 3 dataset v1.0 and v1.1 is used, for the period 01/2006-01/2010. Data are obtained through the NERC Earth Observation Data Centre (http://www.neodc.rl.ac.uk/browse/neodc/arc). The selected file types are i) Day-time dual-view 2-channel and ii) Night-time dual-view 3-channel SST retrievals. The data series include observations from the ATSR 1 instrument on board the ERS-1 satellite, ATSR 2 on board the ERS-2, satellite and the AATSR on board ENVISAT.

Pathfinder SST: The 4 km AVHRR Pathfinder Version 5 SST Project (Pathfinder V5.1) is a reanalysis of the AVHRR data stream developed by the University of Miami's Rosenstiel School of Marine and Atmospheric Science (RSMAS) and the NOAA National Oceanographic Data Center (NODC). The L3 data series include observations from the AVHRR instruments on board NOAA 17 and 18.

In situ SST data from drifting buoys, moored buoys and ship observation are obtained from the ICOADS project and used for validation only.

Sea-ice concentration data: High resolution sea ice data from Swedish Meteorological and Hydrological Institute (SMHI) are used in DMI-OI reprocessing.

SST CCI AVHRR: The specific level 3 pre-processing has been performed to resemble the Pathfinder pre-processing as much as possible. The pre-processing steps include:

- AVHRR 17 + 18 GAC
- DMI-OI L3/L4 CCI processing similar as for Pathfinder data , using only local nighttime observations (21-07 local time).
- Only quality level 5 observations accepted
- Years 2066-2008

DMI-OI Level 4 products

Two setups have been used here:

- **Level 4 DMI-OI Reference run, with Pathfinder AVHRR observations**
- **Level 4 DMI-OI test run, with CCI SST observations**

The processing for the reference and test runs is very similar. For the test run, the Pathfinder observations have been substituted with the SST CCI AVHRR observations, whereas the ARC, and sea ice observations have been kept the same in the two runs.

4.2.5.2.3 Results

The impact of using the CCI products has been assessed by validating the level 3 and level 4 products from both the reference and test runs year 2006-2008. All validation results are based upon comparisons with drifting buoy observations not included in the analysis.

Level 3 results:

The overall results are shown in the table below, indicating smaller bias and standard deviations for the SST CCI AVHRR products when compared to reference data for 2006-2008.

Table 4-2. Overall validation of level 3 SST products compared against independent in situ observations from drifting buoys.

	Pathfinder	ESA CCI
Bias, °C	-0.21	-0.07
Stddev, °C	0.56	0.45

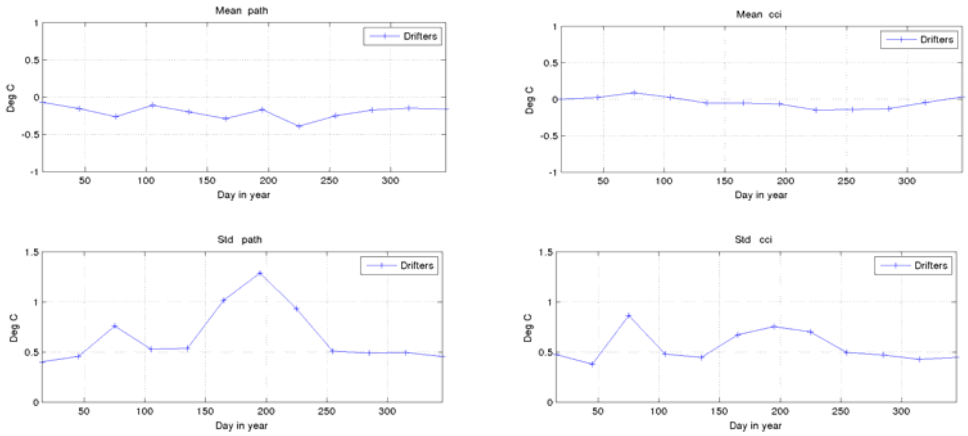


Figure 4-22. Validation statistics calculated as a function of day in the year. Left figures show the Pathfinder results and right figures show the SST CCI AVHRR results. The upper figures show bias and lower figures show standard deviations, when compared against independent drifting buoy observations.

Level 4 results:

The level 4 test data set for the years 2006-2008 has been constructed as described in Section 4.2.5.2.2.

The overall RMS differences are shown for the reference and test data sets in Figure 4-23 for each observation type (drifting buoys, moored buoys and ship observations).

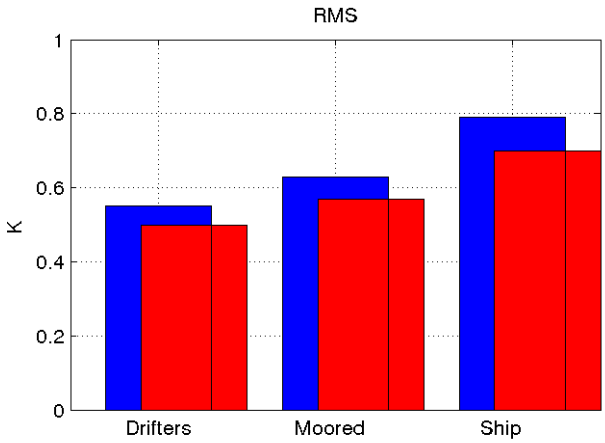


Figure 4-23. Level 4 performance for year 2006-2008. RMS differences for drifting buoys, moored buoys and ship observations. Blue indicate the reference level 4 run with Pathfinder AVHRR data and red is the test run with the SST CCI AVHRR observations.

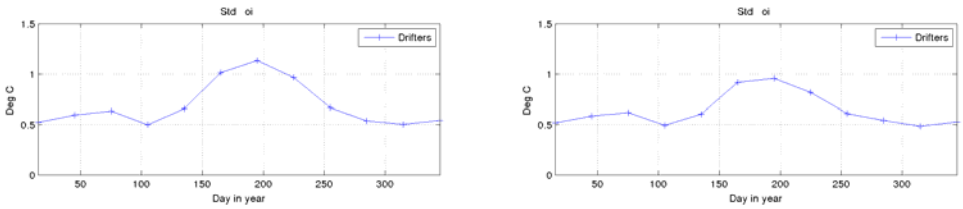


Figure 4-24. Standard deviation of the level 4 products including Pathfinder (left) and ESA CCI AVHRR (right) observations.

4.2.5.2.4 Conclusions

The L3/L4 test runs with ESA CCI data showed:

- SST CCI AVHRR data have smaller absolute biases and standard deviations relative to reference data than the Pathfinder AVHRR version 5.1
- SST CCI AVHRR showed cold summer bias relative to reference data and seasonal variation in the performance
- L4 OI showed improvements using CCI data

4.2.5.3 Feedback on scientific utility of the SST CCI products

The SST CCI AVHRR products in the Arctic region appear to be of higher quality than e.g. the Pathfinder AVHRR observations. In the Arctic, however, it is evident that are still

some issues that need to be solved. The enhanced summer time errors have also been found in other AVHRR data sets and have been reduced in the SST CCI AVHRR data set but not resolved yet.

4.2.6 SIMONE MORAK: COMPARISON TO OTHER SST DATA SETS

Simone Morak-Bozzo

Meteorology Department, University of Reading

4.2.6.1 Key messages

- Analysis product easily usable, with full coverage
- Products pre-averaged to coarser resolution, e.g. 1 degree, would be useful
- Inter-annual variability and EOF patterns are geophysically convincing at ~1 degree resolution, more so than in-situ-only aggregated data
- Will use CCI-based EOFs further in developing historical reconstruction of global SST

4.2.6.2 Scientific analysis

4.2.6.2.1 Aims of the study

We used the ESA SST CCI analysis data set for a comparison analysis, with another satellite-only (obs4MIPS derived from ARC SSTs), two in situ (ICOADS, HadSST3) and one reanalysis data set (HadISST). During the analysis we focused on comparison of the monthly mean climatology pattern, inter-annual variability and EOF patterns. The intention is to assess SST CCI-based EOFs for use in historical reconstruction of in situ SST.

4.2.6.2.2 Method

We computed the monthly mean climatologies as well as the inter-annual variability for each month over the common period 1992-2010.

4.2.6.2.3 Results

The comparison analysis of the spatial patterns of the monthly mean climatologies showed a high agreement across all data sets, despite their difference in spatial resolution.

The patterns of the inter-annual variability, were overall in less precise agreement. The differences most likely arise from the different data sources. Inter-annual variability in the gridded ICOADS dataset shows a noisy pattern except where ship-observations are very frequent. Gridded ICOADS is based solely on in situ measurements, which have been minimally filtered, and therefore measurement outliers can exaggerate apparent variability in sparsely observed areas. HadSST3, shows qualitatively similar characteristics to gridded ICOADS. ARC-obs4MIPS and SST CCI analysis, show a much smoother picture, due to greater sampling. HadISST shows a very smooth pattern because it is a historical reconstruction.

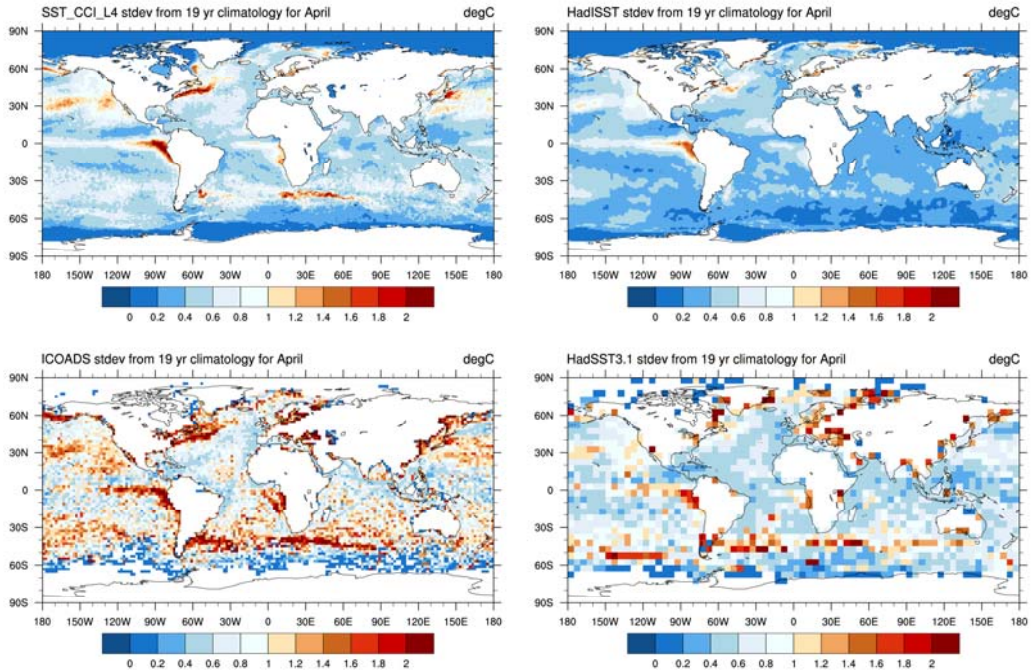


Figure 4-25. Std Dev of April-mean SSTs over same 19 yrs in four data sets: top left, SST CCI; top right, HadISST (reconstruction); lower left, gridded ICOADS (in situ only); lower right, HadSST3 (in situ only).

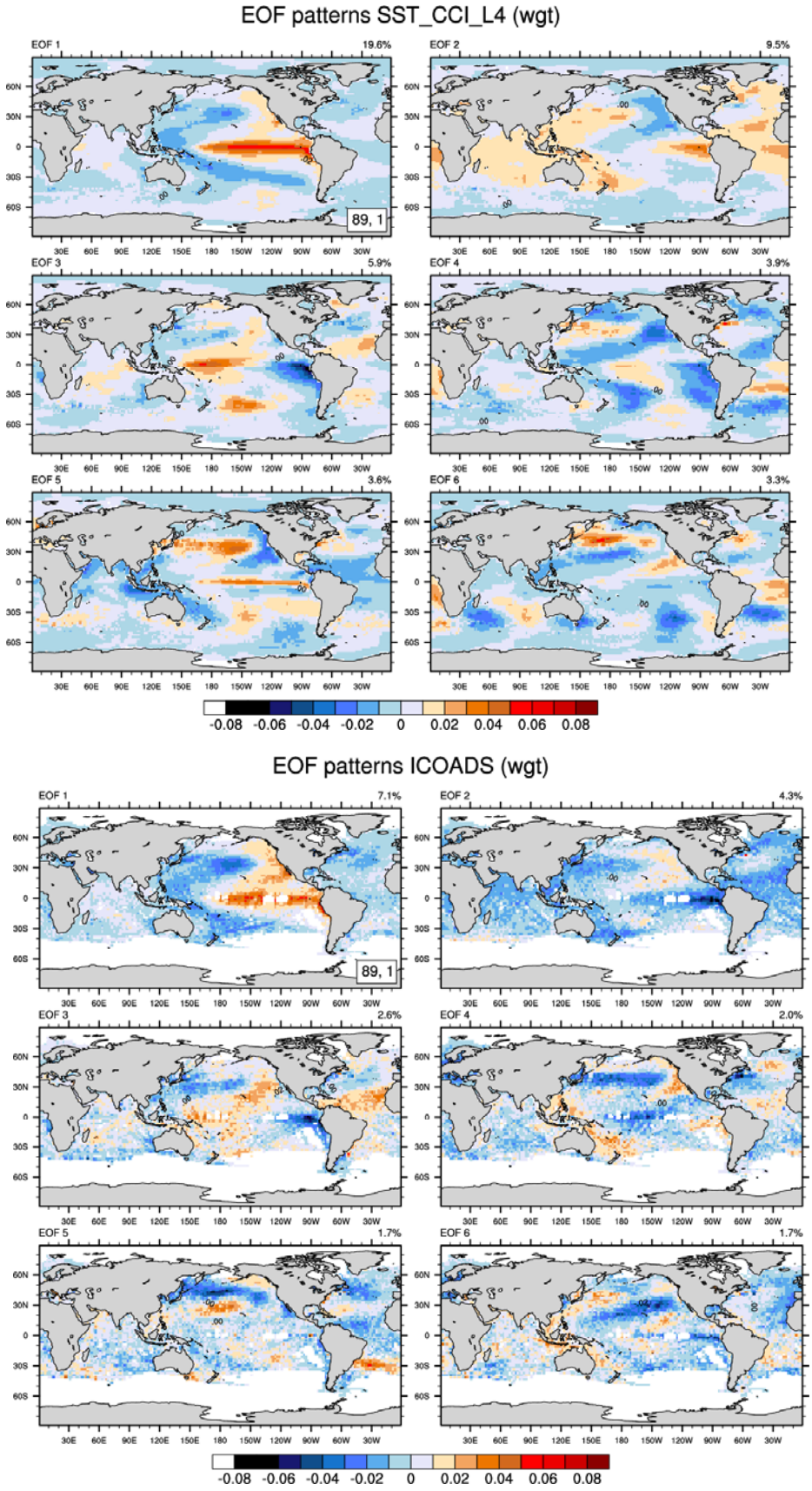


Figure 4-26. First six modes of SST anomaly variability in (top panel) SST CCI analysis and (lower panel) gridded ICOADS.

The leading EOF modes from SST CCI are identifiably the same as those derived from ICOADS, but account for about twice the fraction of total variance than in the case of ICOADS, and are more spatially coherent.

4.2.6.2.4 Conclusions

The SST CCI analysis product appears promising as a source of EOFs for historical SST reconstruction.

4.2.6.3 Feedback on scientific utility of the SST CCI products

The SST CCI analysis data set is a very valuable data set for the analysis of sea surface temperatures, due to its coverage and spatial detail. However, due to the high spatial resolution the data size is large compared to gridded in-situ only datasets, which isn't very practical for quick downloads or a preliminary analysis.

While the EOFs from the dataset look promising for use in historical reconstruction because they cleanly explain a high fraction of total variance, the time series is only 20 years long. This means, for example, that modes of variability associated with ENSO will be dominated by the major El Nino of the late 1990s. A longer time series, preferably >30 years duration, will be preferable for sampling more of the possible variability (e.g., the El Ninos of the 1980s, the SST response to two major volcanic eruptions, etc).

4.2.7 KATIE BROWN: INVESTIGATING LINKS BETWEEN SST AND HYDROLOGICAL CYCLE VARIABLES

Katie Brown

On secondment to Met Office Hadley Centre

4.2.7.1 Key messages

- The new SST CCI analysis data has few notable differences when compared to the AMIP dataset, whilst the Daily OI is generally more poorly correlated to precipitation and cloud amount. All datasets exhibit the largest variability in the equatorial tropical pacific and at the mid latitudes and show similar spatial patterns in their anomaly fields. The climatologies of the data sets are offset from each other. The Daily OI is coolest whilst the AMIP is warmest overall.

- The correlation between precipitation and sea surface temperature is mostly positive, where significant, whilst the correlation between cloud amount and sea surface temperature is both positive and negative. Areas of positive correlation, mostly across ENSO regions, are strongly linked to the relationship between precipitation and SST whereas areas of negative correlation are not and indicate regions of marine stratocumulus cloud.

4.2.7.2 Scientific analysis

4.2.7.2.1 Aims of the study

We have analysed the datasets and their relationships to each other by considering four different areas:

- How do the SST anomalies differ between datasets?

- What is the relationship between precipitation and SST and does each of the datasets represent this? What are the differences?

- How does cloud coverage correlate to sea surface temperatures in different regions? Is this the same in each of the datasets?

- How does seasonality affect the relationships between SST, precipitation and cloud amount?

4.2.7.2.2 Method

The three datasets used in this study are:

- SST CCI analysis. Spatial grid resolution of 0.05°.

- Reynolds et al (2007) Daily OI. Version used includes data from AMSR-E from June 2002 onwards. Spatial grid resolution of 0.25°.

- Atmospheric Model Intercomparison Project (AMIP) dataset. Created by joining the monthly mean HadISST1 with the OI.v2 (Hurrell et al., 2008). Spatial grid resolution of 1°.

4.2.7.2.3 Results

4.2.7.2.3.1 Sea Surface Temperature Anomalies

In order to aid comparison between the datasets each dataset was re-gridded to a resolution of 1°. Using the re-gridded data a climatology for each dataset was created enabling calculation of the anomalies. Figure 4-27 indicates the overall variation in the climatologies is approximately 0.15K across the three datasets with the AMIP dataset the warmest for the period considered (1991-2010) and Daily OI the coolest.

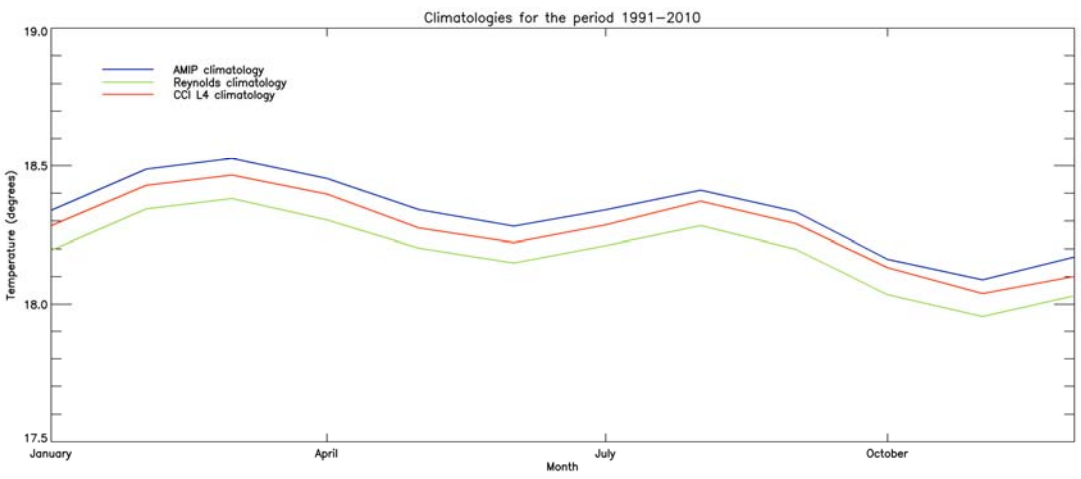


Figure 4-27: Climatology time series for each dataset. The AMIP climatology is calculated between January 1991 and December 2010 whilst the Daily OI and SST CCI analysis climatologies use data beginning September 1991.

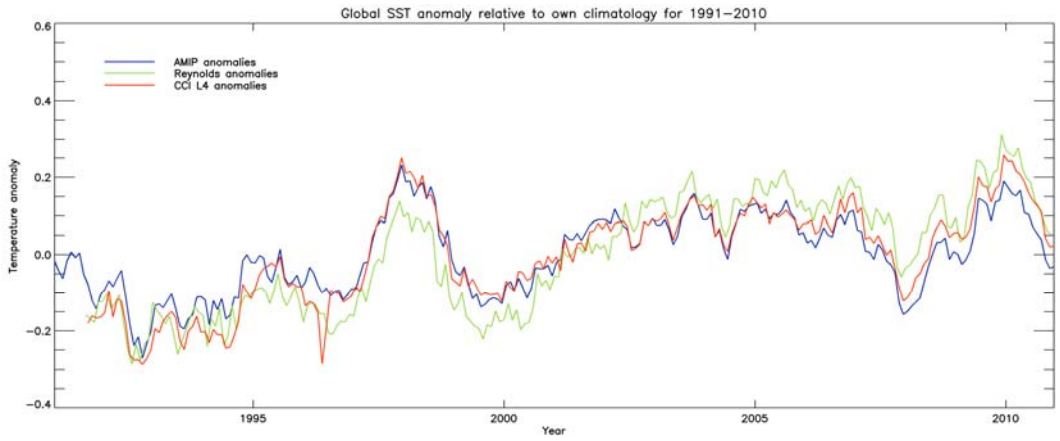


Figure 4-28: Global sea surface temperature anomaly time series (°C) for each dataset (relative to their own climatologies).

It is clear to see from Figure 4-28 that the Daily OI dataset is cooler than the other two datasets pre 2002 whilst it warms with anomalies greater than AMIP and SST CCI analysis post 2002. The ENSO event of 1997-1998 is weaker in the Daily OI dataset not reaching as higher peak in the anomalies than AMIP and SST CCI analysis. It is also noticeable that there is a distinct spike in the SST CCI analysis data in May 1996 where it is significantly cooler than both AMIP and Daily OI

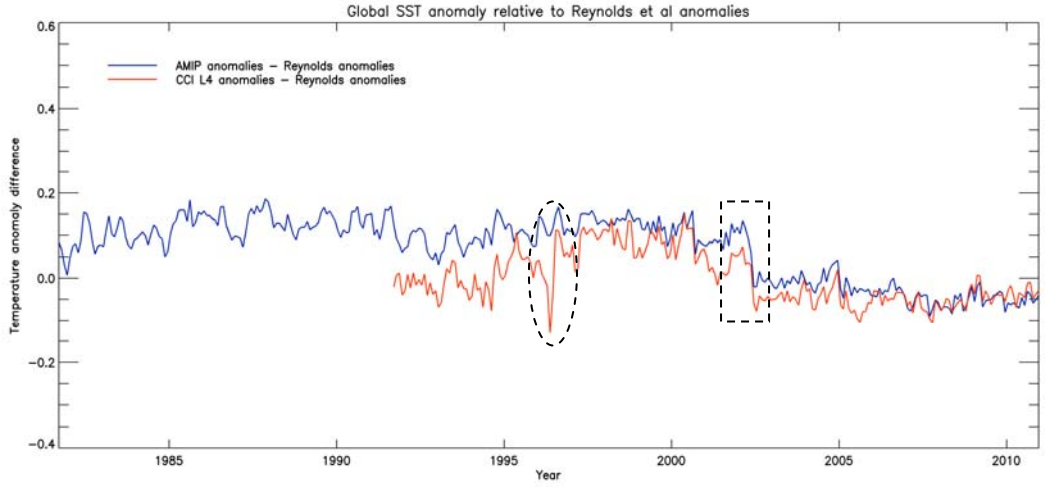


Figure 4-29: Global SST anomaly difference time series. The plot displays time series of AMIP and SST CCI analysis anomalies minus the Daily OI anomalies.

The dotted rectangle in Figure 4-29 indicates the Daily OI anomaly increasing where the AMIP and SST CCI analysis anomalies do not. This occurs in June 2002 and is due to the incorporation of AMSR data. In addition, the spike identified in the SST CCI analysis in Figure 4-28, highlighted by a dotted oval shape, can still be clearly seen and occurs in May 1996. Further investigation has identified that this is from problems with the first measurement instrument ATSR1, which were apparent towards the end of its life.

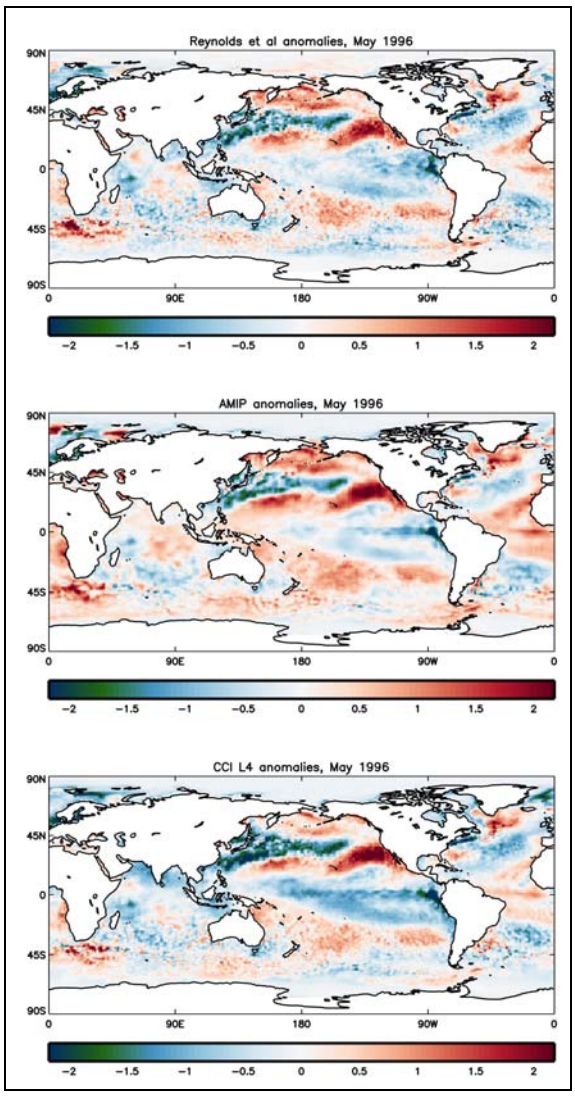


Figure 4-30: Temperature anomaly contour plot in May 1996 for each of the datasets. (a) Daily OI; (b) AMIP; (c) SST CCI analysis.

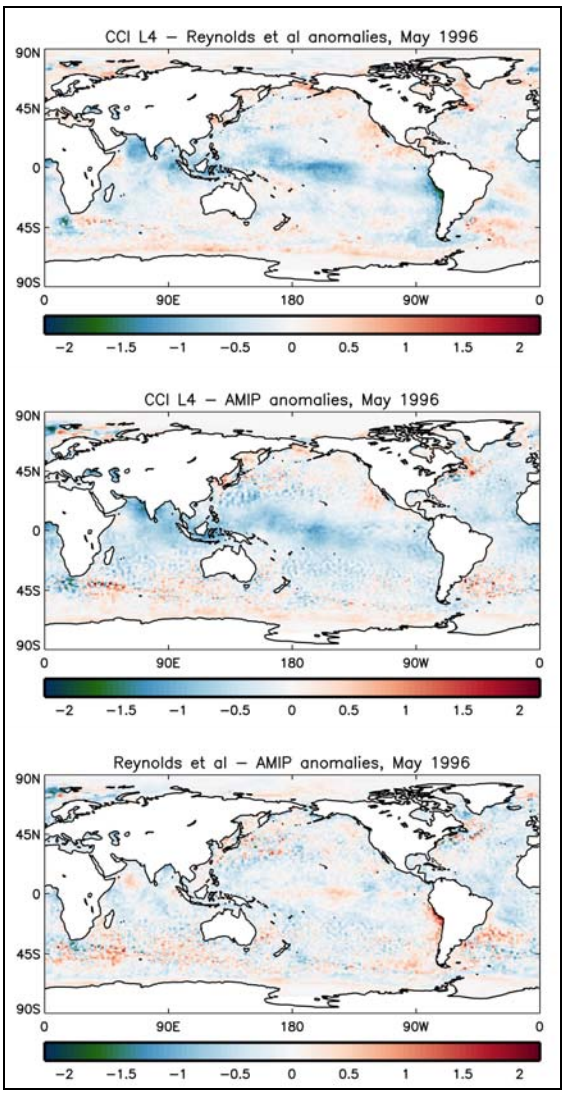


Figure 4-31: Anomaly difference contour plot for May 1996. Differences between each of the datasets calculated.

The SST CCI analysis dataset appears cooler in the tropical Pacific and Indian Ocean compared to both AMIP and Daily OI as shown in Figure 4-30 and Figure 4-31. Parts of the Atlantic Ocean are also cooler in the SST CCI analysis, notably around the Gulf Stream.

Figure 4-32 highlights the differences in the average anomaly between the two periods 1991-2002[6] and 2002-2010. Each anomaly was created using each datasets own climatology for the period 1991-2010. The boundary between the two time periods is chosen as this is when the AMSR measurements are introduced to the Daily OI.

[6] *Although the period used to calculate average anomalies for the AMIP data set is slightly different from that used for the others, this does not affect the conclusions drawn,*

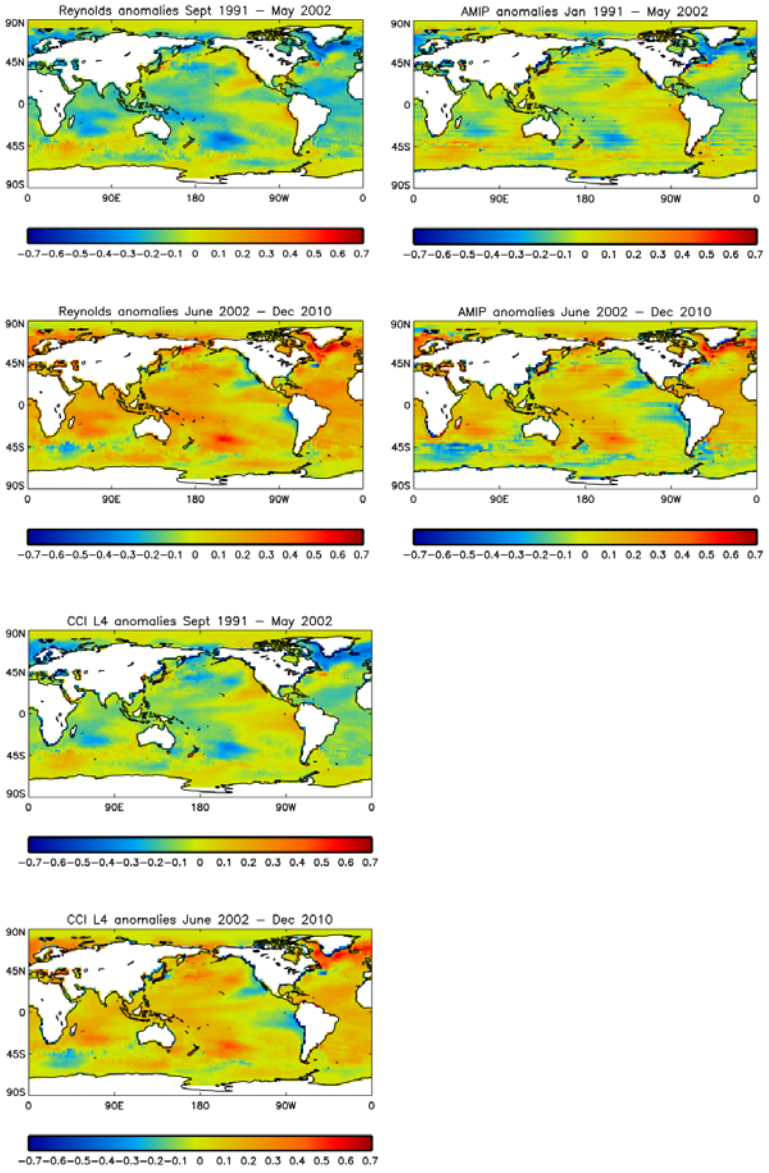


Figure 4-32: SST anomaly contour plot for each of the three data sets considered. SST anomalies are averaged in time at each grid box with one plot showing average SST from 1991-May 2002 and another June 2002 – December 2010.

The Daily OI anomalies are cooler than AMIP and SST CCI analysis for the period prior to June 2002 whilst they are notably warmer for the period which follows.

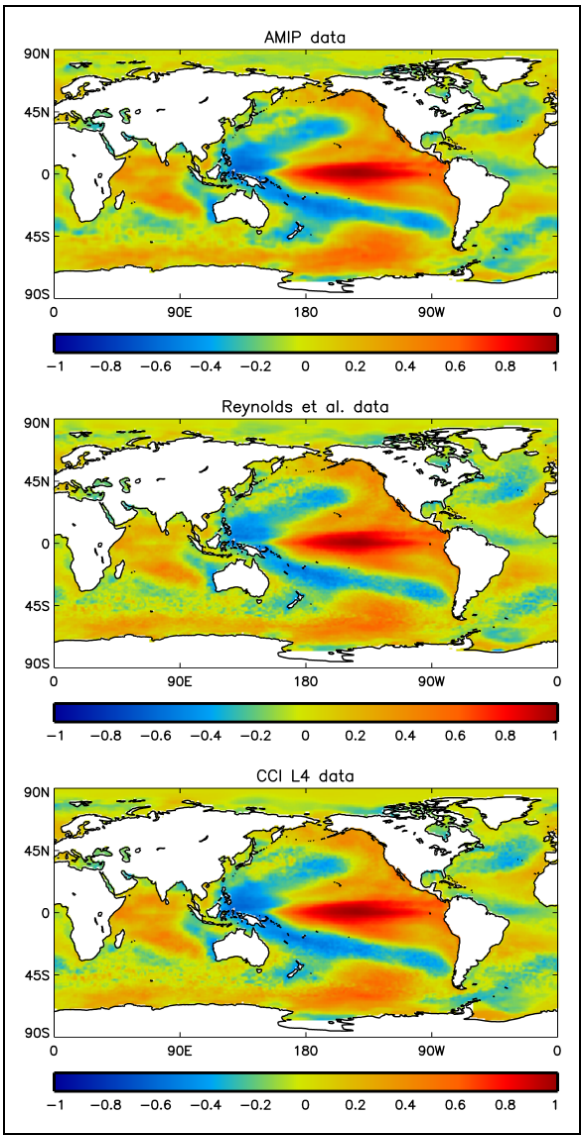


Figure 4-33: Correlation coefficient plot for 1991 – 2010. Grid box correlation coefficients are calculated between the centre of Nino region 3.4 (140W, 0S) SST anomaly and SST anomalies elsewhere.

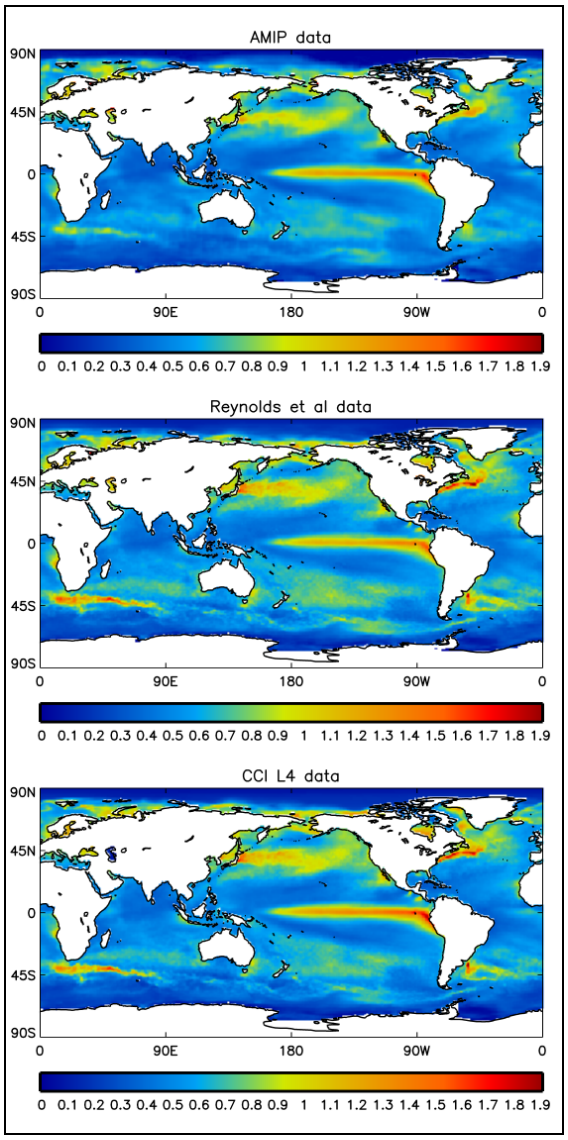


Figure 4-34: Standard deviation plot of SST anomalies. Standard deviation is calculated based on data from period 1991 – 2010.

Figure 4-33 highlights a visibly strong positive correlation across all of the ENSO regions with region 3.4. Surrounding this is an area of strong negative correlation which demonstrates the well known horseshoe shape, loosely wrapping around the ENSO regions.

The most variation in the anomalies occurs in the Equatorial Tropical Pacific Ocean as we might expect. The patterns shown in Figure 4-34 quite clearly identify the high variation in the Equatorial Tropical Pacific Ocean to be related to ENSO events. Between the three datasets the differences are only slight whilst there is reduced variability off the South East Coast of South America in the AMIP data when compared to Daily OI and SST CCI analysis. Further, the variability seen in the mid latitudes is also weaker in the AMIP data compared to the others.

4.2.7.2.3.2 Sea Surface Temperature and Precipitation

This section looks at the relationship between SST and precipitation. The GEWEX GPCP dataset is used to investigate the precipitation patterns. It contains rain gauge data, satellite geostationary and low-orbit infrared, passive microwave and sounding observations to provide a spatial complete data set on a 2.5° spatial grid resolution. The dataset begins in 1979 and runs until present. In order to assess the relationship between precipitation and SST it was necessary to re-grid all three SST datasets to a 2.5° spatial grid resolution to match the precipitation dataset.

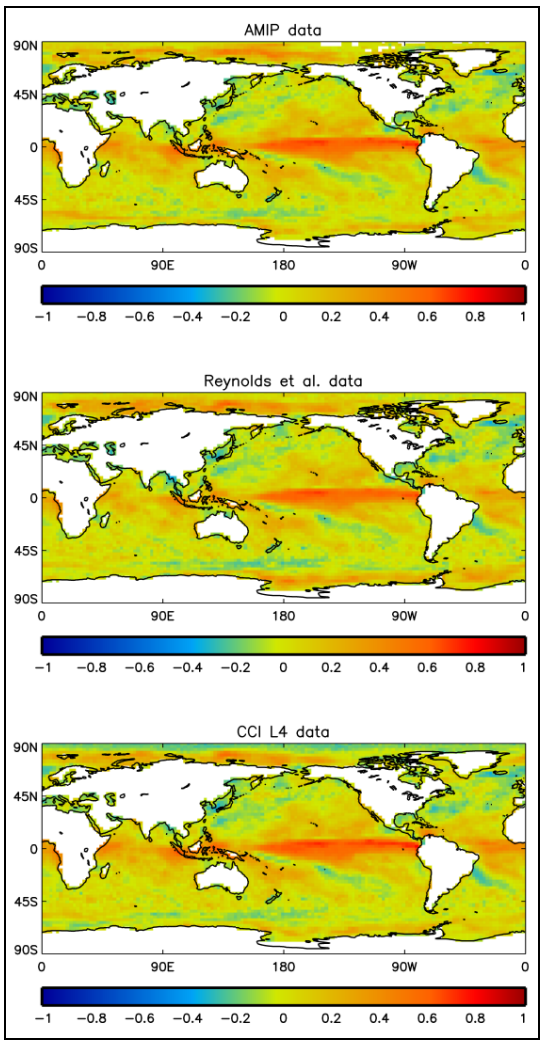


Figure 4-35: Point-by-point correlation coefficient plot between SST anomalies and precipitation anomalies. Precipitation anomalies are calculated using a precipitation climatology for the period 1991-2010.

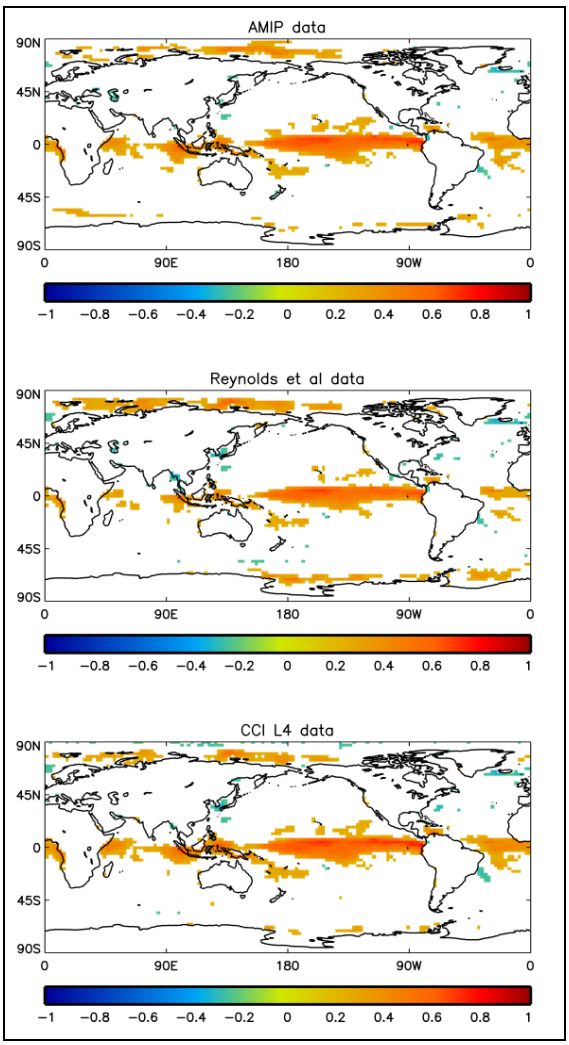


Figure 4-36: Significant correlations at the 5% level.

All three datasets demonstrate similar relationships to precipitation, as seen in Figure 4-35, with each dataset showing the same general spatial patterns. The correlation in the tropical Pacific is particularly notable as it is both large scale spreading ~80W to 170E and strongly positive. This is what one may expect as the higher SST's result in increased precipitation due to higher rates of evaporation at the surface. As a result precipitation is

mostly positively correlated to SST. However, there are some regions, such as the Southern Ocean where an area of warm current meets cooler current creating a frontal gradient and here there is a weak negative correlation to precipitation.

Figure 4-36 helps to highlight significant differences between the datasets and their correlation to precipitation. The Daily OI data is more weakly correlated than the SST CCI analysis or AMIP which is particularly noticeable in the equatorial regions. These differences are investigated further in Figure 4-37 to Figure 4-42.

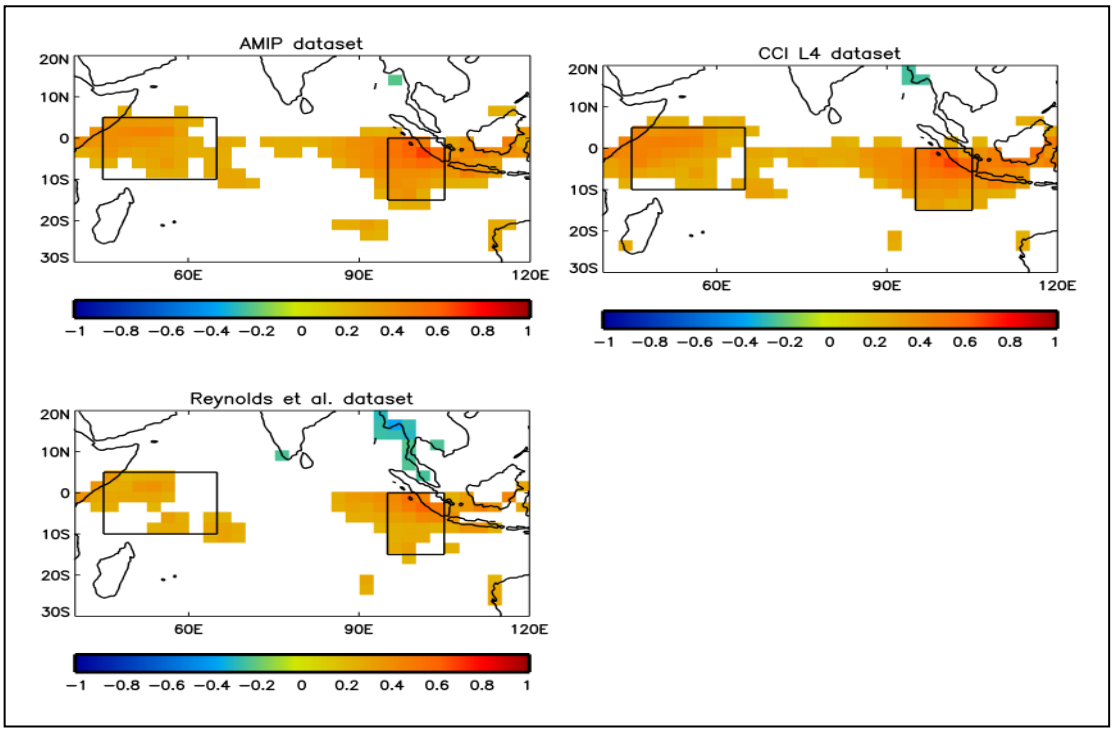


Figure 4-37: Point-by-point correlation coefficient plot showing relationship between SST anomalies and precipitation anomalies in the Indian Ocean. Plot is a zoomed in section of Figure 4-36. Areas highlighted by boxes are shown as time series plots in Figure 4-38 and Figure 4-39.

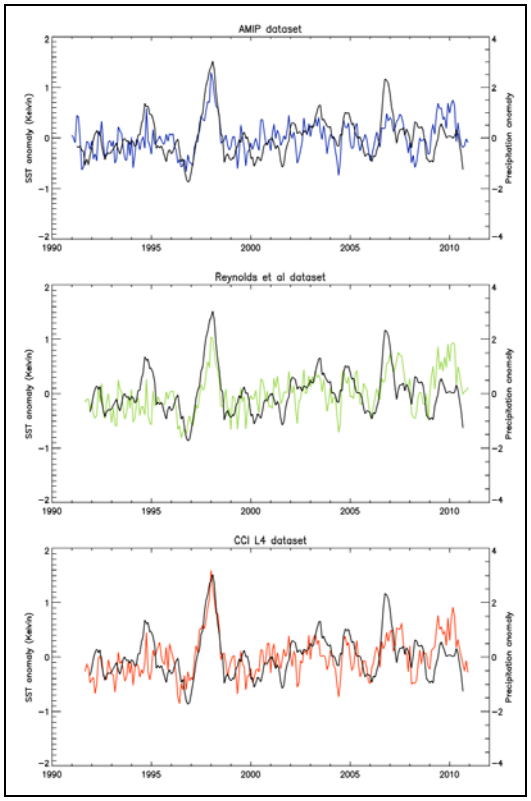


Figure 4-38: Time series of the western Indian Ocean box as seen in Figure 4-37. Precipitation anomalies are shown in black and have been plotted using a 7 month rolling average.

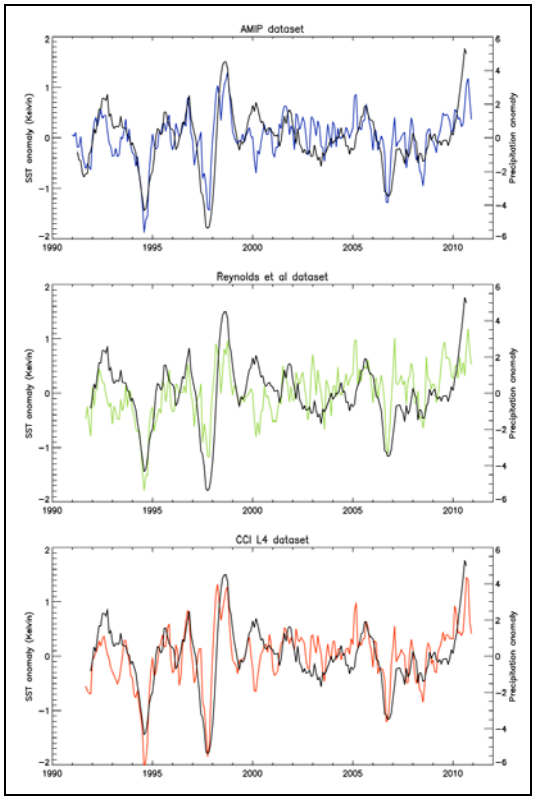


Figure 4-39: Time series of the eastern Indian Ocean box as seen in Figure 4-37. Precipitation anomalies are shown in black and have been calculated as described in Figure 4-38.

Figure 4-37 focuses on the Indian Ocean. This area has been chosen as there were clear differences between the three datasets on inspection of the correlation plot, Figure 4-36. It is clear that the Daily OI dataset is more weakly correlated with precipitation in this region when compared to both the AMIP and SST CCI analysis datasets. The SST CCI analysis dataset is most strongly correlated with precipitation.

In addition, the time series of interesting sub-regions within the area considered are shown in Figure 4-38 and Figure 4-39. Particularly noticeable are the differences between the datasets during the ENSO event of 1998. Here it is plain that the Daily OI SST anomalies are more poorly correlated with precipitation.

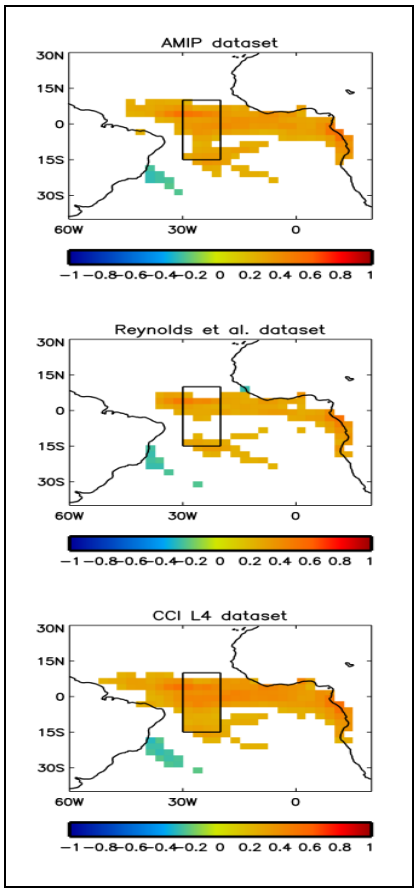


Figure 4-40: Point-by-point correlation coefficient plot showing relationship between SST anomalies and precipitation anomalies in the Atlantic Ocean. Plot is a zoomed in section of Figure 4-36. Boxed area shown as time series in Figure 4-41.

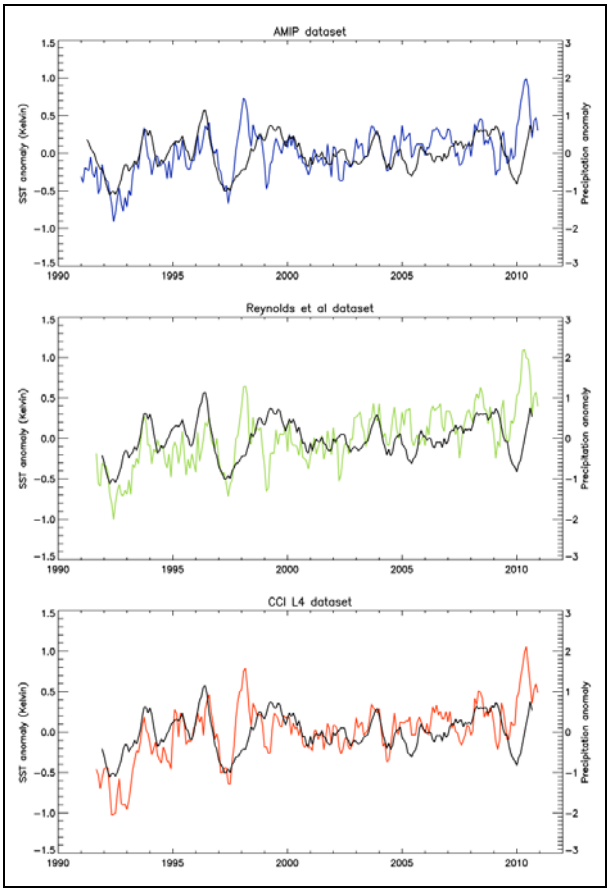


Figure 4-41: Time series of the Atlantic Ocean box as seen in Figure 4-40. Precipitation anomalies are shown in black and have been calculated as described in Figure 4-38.

Figure 4-40 corroborates what is seen in Figure 4-37; the Daily OI is more weakly correlated with precipitation than the other two datasets. Further, the SST CCI analysis anomalies remain the most strongly correlated to precipitation in terms of both stronger positive and stronger negative correlations. The correlation differences between the datasets are less obvious in the time series plots, Figure 4-41, which focus on an area where the correlation is noticeably different between datasets.

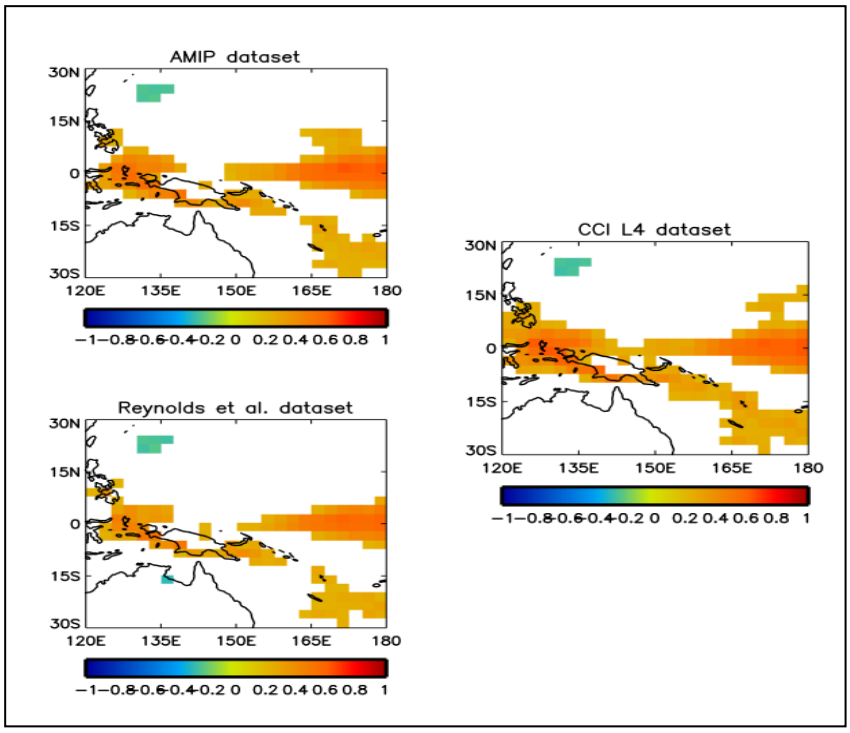


Figure 4-42: Point-by-point correlation coefficient plot showing relationship between SST anomalies and precipitation anomalies in the West Pacific Ocean. Plot is a zoomed in region from Figure 4-36.

Once again in Figure 4-42, as in Figure 4-37 and Figure 4-40 the local relationship between precipitation and SST is weakest in the Daily OI dataset. As before, the SST CCI analysis remains the most strongly correlated.

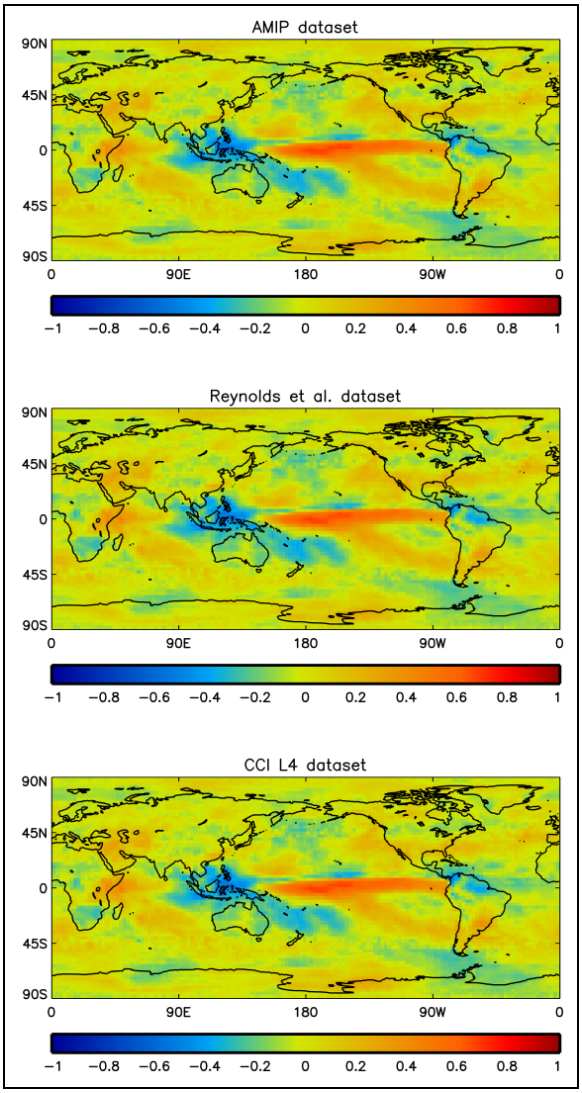


Figure 4-43: Plot displaying correlation coefficients between the centre point of Nino region 3.4 (140W, 0S) SST anomaly and Precipitation anomalies elsewhere. Correlation calculated for period 1991-2010.

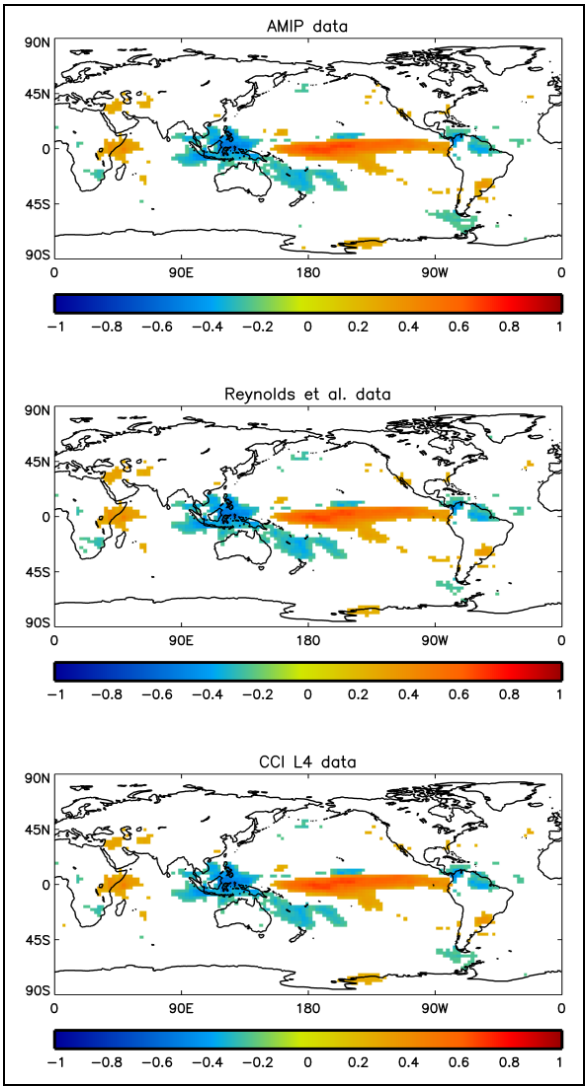


Figure 4-44: Figure 4-43 reproduced showing only significant correlations at the 5% level. See Appendix for method used in significance calculations.

There are large scale high positive correlations across the tropical equatorial Pacific as well as strong negative correlations around the Indonesian warm pool. These correlations are statistically significant at the 5% level as highlighted by Figure 4-44. In addition, Figure 4-43 shows weakly correlated regions over North America and Canada with two bands of weakly positive correlation surrounding a band of weak negative correlation.

The negative correlation around the Aleutian Islands is for the most part only observed in Figure 4-43, although there is a small region, shown in Figure 4-44, which is significant. This correlation identifies the Aleutian low, a factor associated with El Nino events.

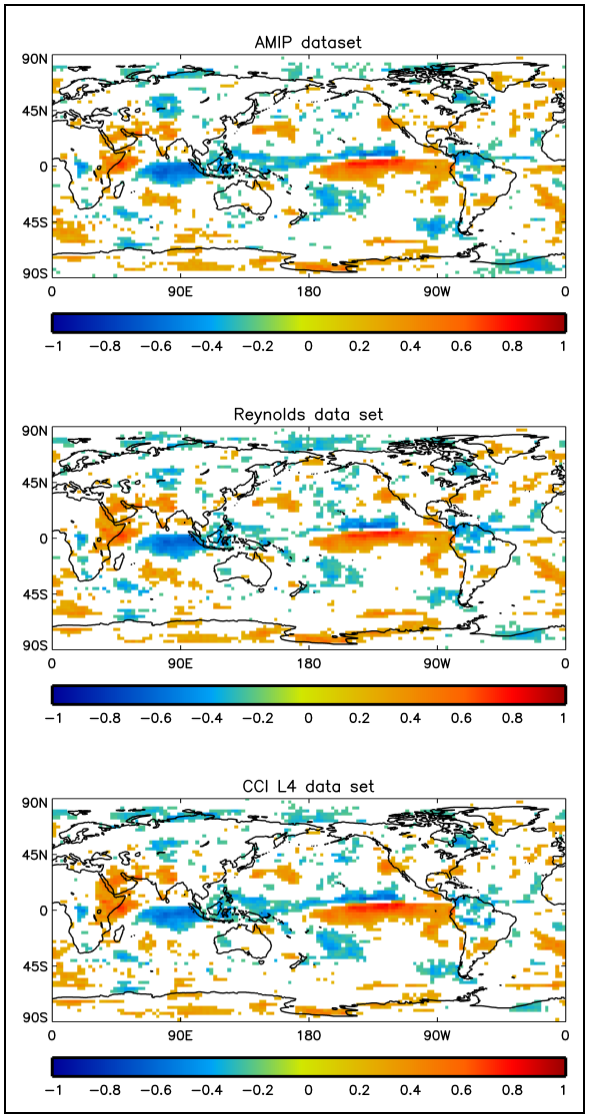


Figure 4-45: Contour plot of significant (at 5% level) correlation coefficients between Nino region 3.4 centre (140W, 0S) SST anomaly and precipitation anomalies elsewhere. Correlation calculated for period 1991 – 2010 when El Nino events were occurring.

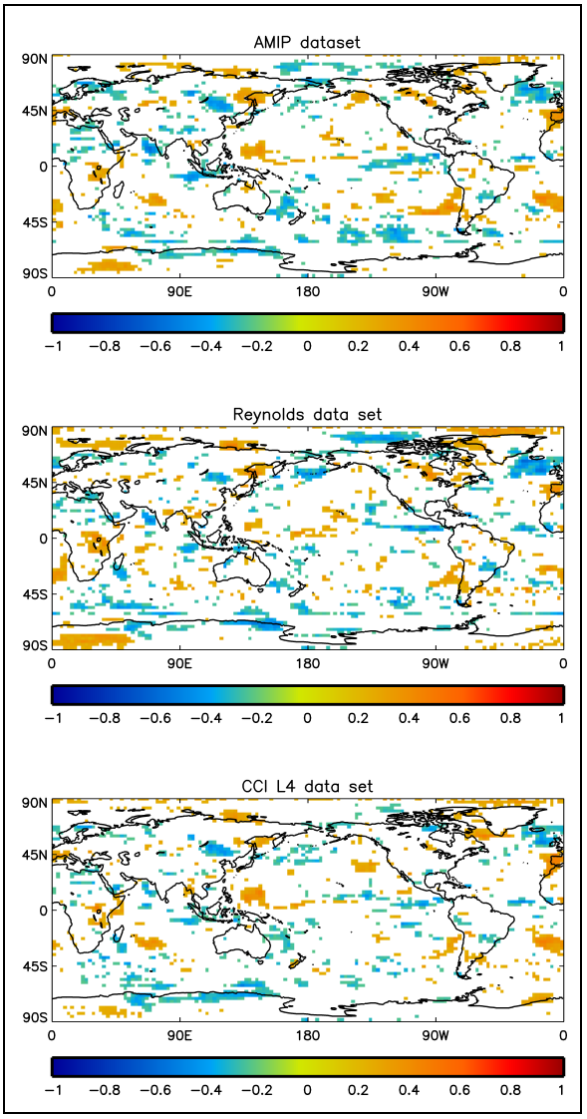


Figure 4-46: Contour plot of significant (at 5% level) correlation coefficients between Nino region 3.4 centre (140W, 0S) SST anomaly and precipitation anomalies elsewhere. Correlation calculated for period 1991 – 2010 when La Nina events were occurring.

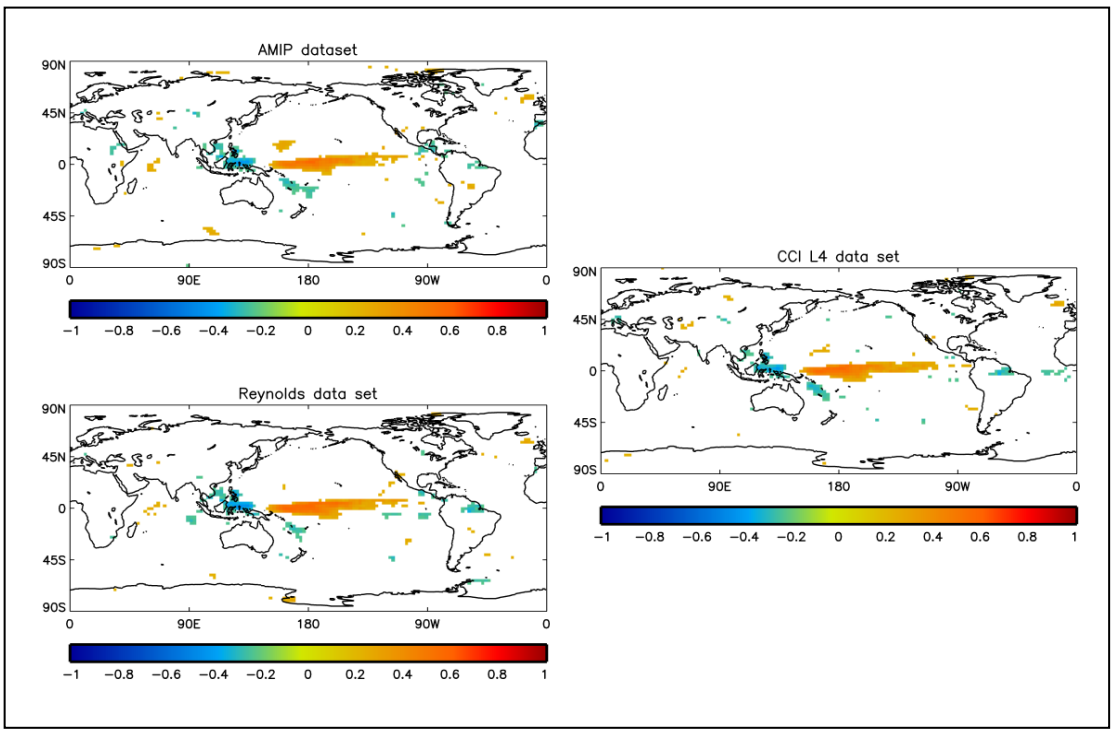


Figure 4-47: Contour plot of significant (at 5% level) correlation coefficients between Nino region 3.4 centre (140W, 0S) SST anomaly and precipitation anomalies elsewhere. Correlation calculated for period 1991 – 2010 when no ENSO events were occurring.

There are only slight differences in correlation strength and spatial pattern between the three data sets in Figure 4-45. There is a clear similarity between the El Nino correlations and Figure 4-44, particularly in the equatorial regions. The correlation between precipitation and SST Nino region 3.4 has less distinct spatial patterns during La Nina than El Nino. Figure 4-46 shows lots of small regions of positive and negative correlation making it difficult to identify differences between the data sets. Generally there are few notable differences in the spatial patterns between the datasets.

Figure 4-47 shows similar spatial patterns to those observed in Figure 4-44 although the area of negative correlation around the Indonesian warm pool is visibly smaller. Further, the correlation seen is associated to El Nino events and corresponds closely to correlations seen in the equatorial regions of Figure 4-45. This could suggest that the strength of the correlation seen in Figure 4-47 may be due, if only in part, to some of the El Nino periods possibly captured by the criterion set for the non ENSO events.

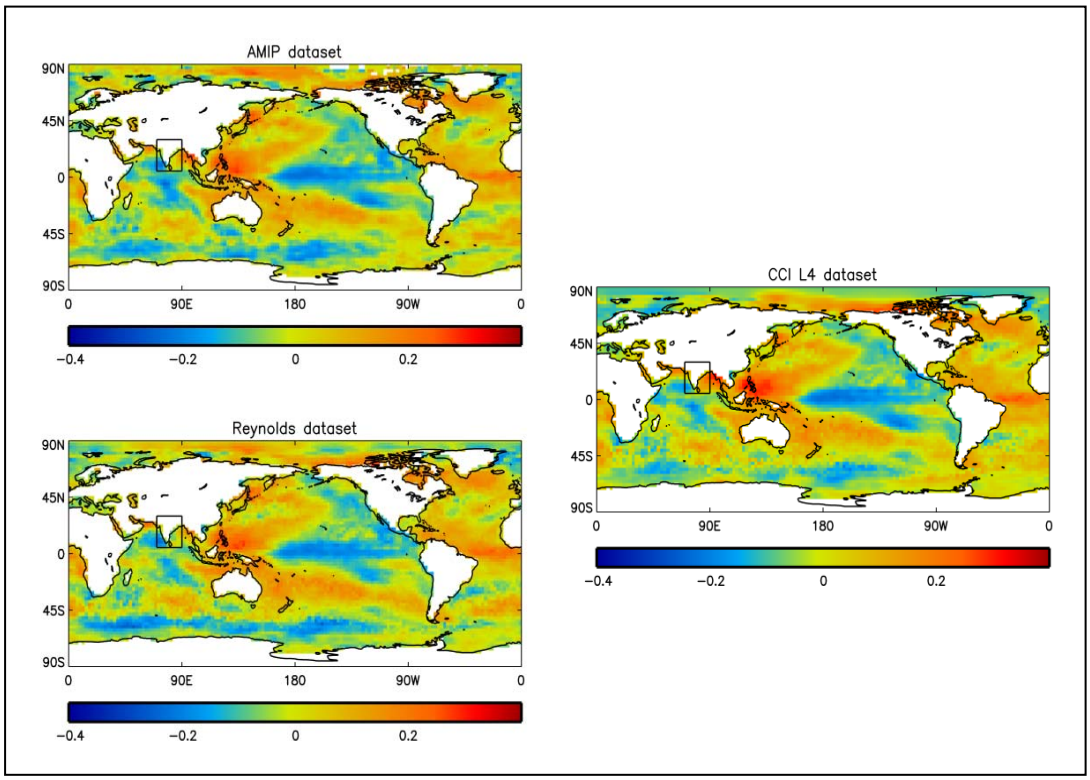


Figure 4-48: Correlation coefficient plot highlighting teleconnections between precipitation anomalies in India and the near Ocean with SST anomalies elsewhere. The box in each plot is the area over which precipitation anomalies were area averaged before correlating with SST anomalies.

Figure 4-48 shows similar spatial patterns to those seen in Figure 4-33 except that the correlation are reverse, i.e. in the tropical pacific Figure 4-33 shows a strong positive correlation whilst in Figure 4-48 we see a strong negative correlation. This implies that as precipitation increases over India, there is a decrease in SST in the tropical equatorial Pacific. A further reason for this could be because the Indian monsoon season occurs in the northern hemisphere summer whilst El Nino events tend to peak in the northern hemisphere winter and this is when the SST would be warmest.

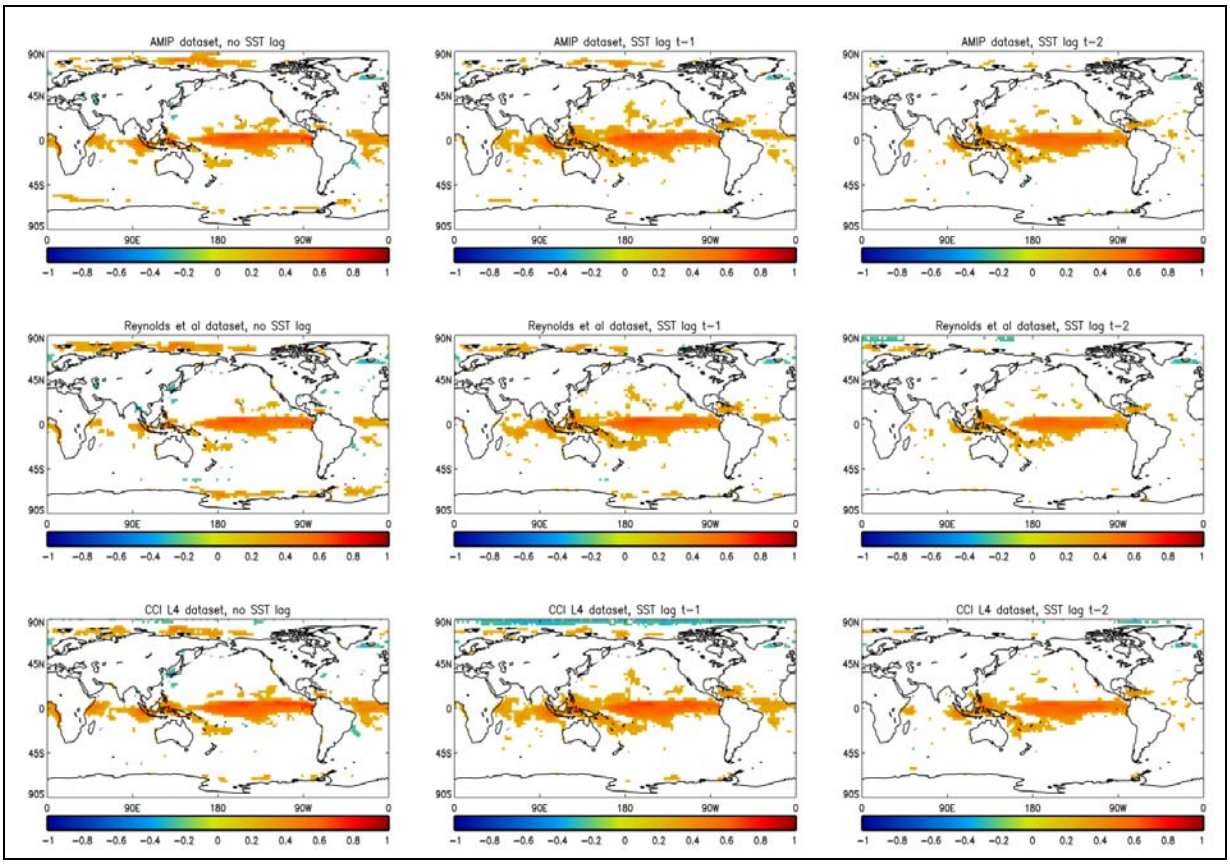


Figure 4-49: Significant (at the 5% level) point-by-point correlation coefficient plots between SST anomalies and precipitation anomalies. The first column is a repeat of Figure 4-36, the second and third columns show plots where precipitation anomalies at time t are correlated with SST anomalies at time t-1 and time t-2 respectively.

Interestingly, the correlation between SST and precipitation increases in strength between SST with no lag and a one month SST lag. This occurs most noticeably around the Indian Ocean and the Indonesian warm pool. The correlation then weakens in the two month lags plots although is still stronger in some areas than the correlation with no lag. Therefore, it is evident that the lag in the time it takes for the atmosphere to respond to the changes in SST and cause changes in the precipitation patterns is approximately one month.

The difference in correlation between the three datasets is consistent with previous findings; that is the Daily OI dataset presents the weakest correlations whilst the SST CCI analysis shows the strongest. However, the differences are all small and the spatial patterns are the same.

4.2.7.2.3.3 Sea Surface Temperature and Cloud Amount

Here we investigate the relationship between SST and cloud amount. The International Satellite Cloud Climatology Project (ISCCP, Rossow et al., 1999; RD.368) D-series data for cloud amount has been used to assess the relationship between SST's and cloud coverage. Both the SST data used and the cloud data are on a 2.5° spatial grid.

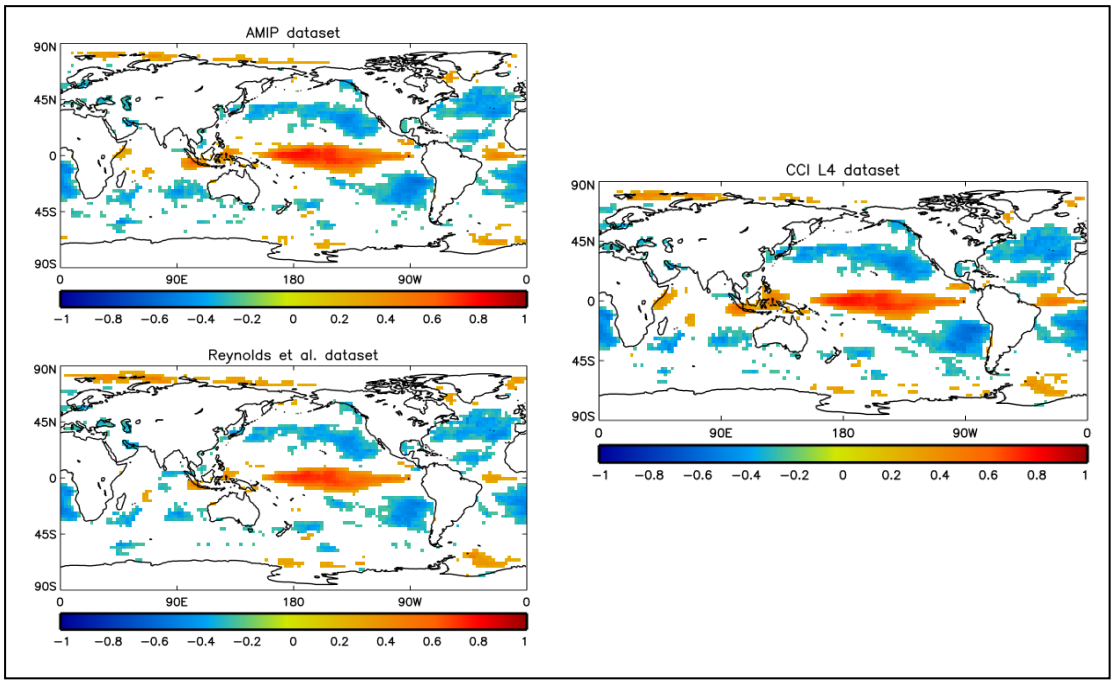


Figure 4-50: Point-by-point correlation plot of significant correlation coefficients at the 5% level. The correlation is between SST anomalies and cloud amount anomalies and is for the period 1991-2008.

As shown in Hanson (1991) the overall correlation between cloud and SST is negative. Whilst there is a strong positive correlation in the equatorial tropical pacific, there are several regions of large negative correlation which correspond to the four well-known marine stratocumulus regions.

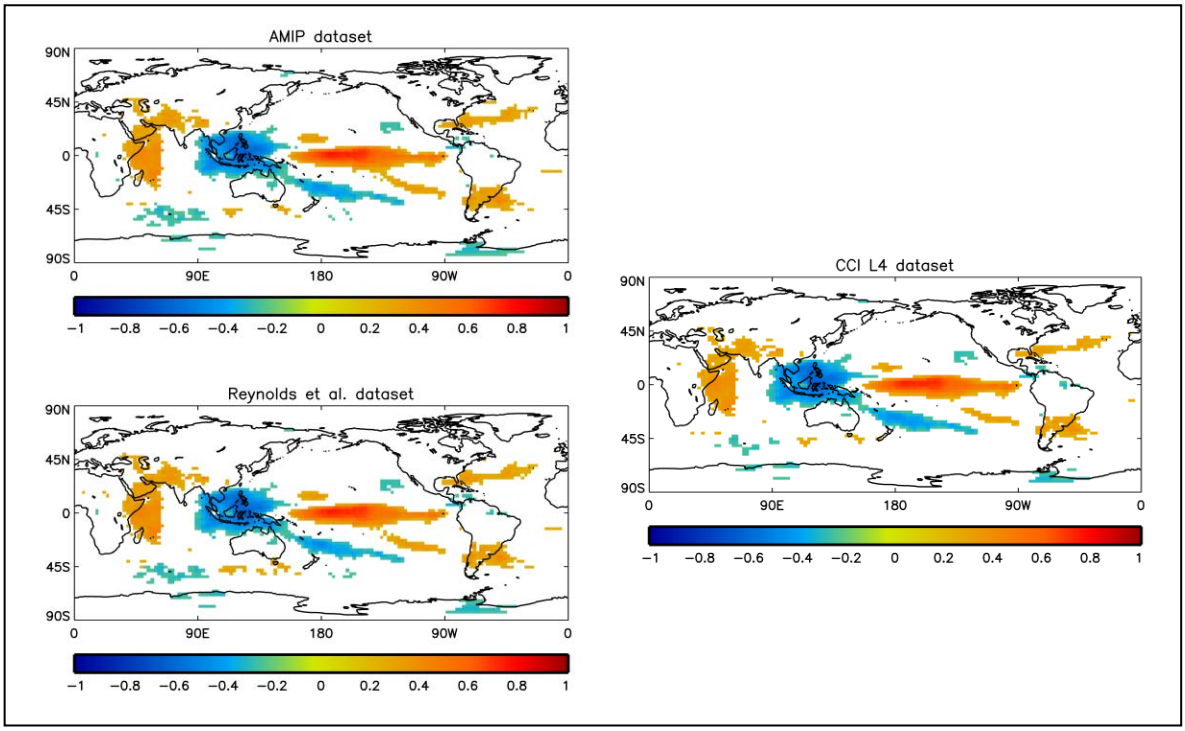


Figure 4-51: Contour plot displaying significant correlation coefficients at the 5% level between the centre point of ENSO region 3.4 (140W, 0S) SST anomaly and cloud amount anomalies elsewhere. The correlation is calculated for the period 1991-2008.

The correlation between ENSO region 3.4 and cloud cover is strong in the equatorial tropical pacific suggesting that ENSO events have a large impact on cloud cover in the region. The spatial patterns in Figure 4-51 are similar to those seen in Figure 4-44 the precipitation/SST region 3.4 correlation plot, particularly in equatorial regions. The negative correlation over the Indonesian warm pool is also seen in Figure 4-44, although in the total cloud amount correlation plot the area of negative correlation spreads further south east across the Southern Ocean. Again it is not surprising that these two correlations go hand in hand as where there is reduced cloud there is likely to be reduced rain.

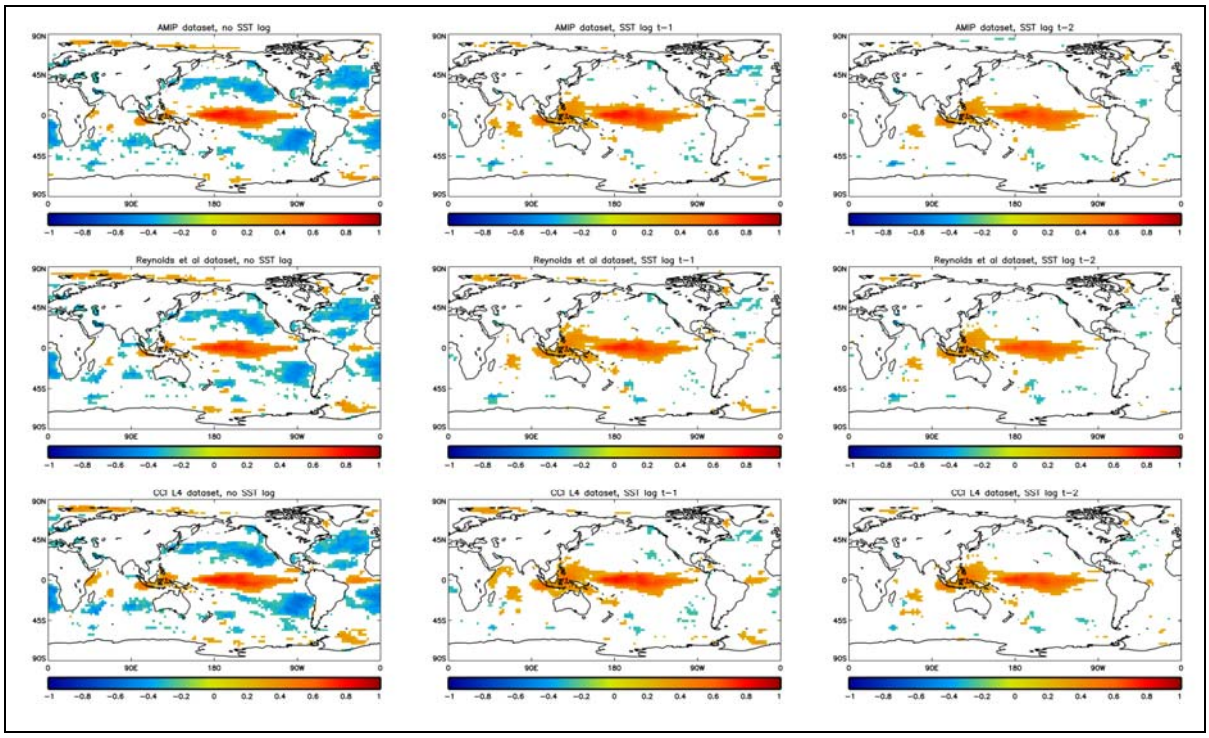


Figure 4-52: Series of contour plots showing significant point-by-point correlations at the 5% level between SST anomalies and cloud amount anomalies. The first column is a repeat of Figure 4-36, the second and third columns show plots where precipitation anomalies at time t are correlated with SST anomalies at time t-1 and time t-2 respectively.

Exploring the link between cloud amount and lagged SST in Figure 4-56, we see the correlation between SST and cloud amount increases in strength where positive correlations are observed. The largest increase in correlation strength can be seen noticeably across the Indonesian warm pool. The correlation then weakens as cloud amount is correlated to SST with a two month lag.

Conversely, where negative correlations are seen, in the marine stratocumulus regions, there is a distinct reduction in strength of the correlation and it is no longer significant when related to SST the month before. This would suggest that the atmospheric response to changes in SST in these regions are more instant than in equatorial regions where it appears to take approximately one month to respond. The differences between the datasets are small and all three show the same general spatial patterns.

4.2.7.2.3.4 Sea Surface Temperature, Precipitation, Cloud Amount and the Seasons

This section investigates the seasonality of SST and the relationship it has to precipitation and cloud coverage. The seasons are May through September and November through March. This aims to capture both summer and winter seasons for each of the northern and southern hemispheres.

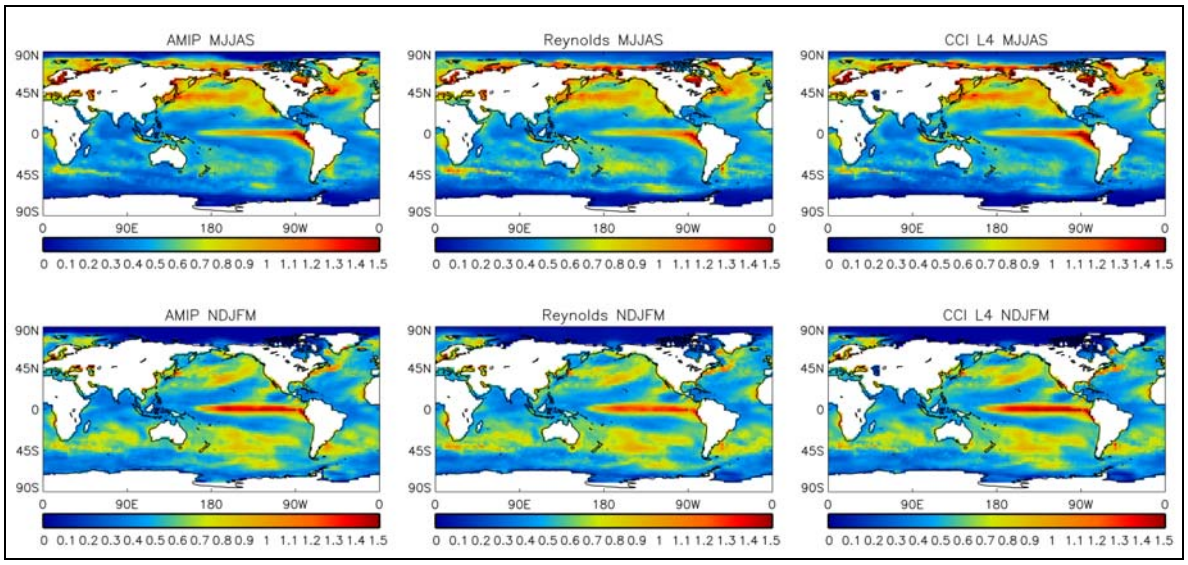


Figure 4-53: Contour plots displaying the seasonal standard deviation of SST for each dataset.

The largest variability in SST is seen in the northern hemisphere summer. Differences between data sets can be seen in the mid-latitude Southern Hemisphere in MJJAS in particular.

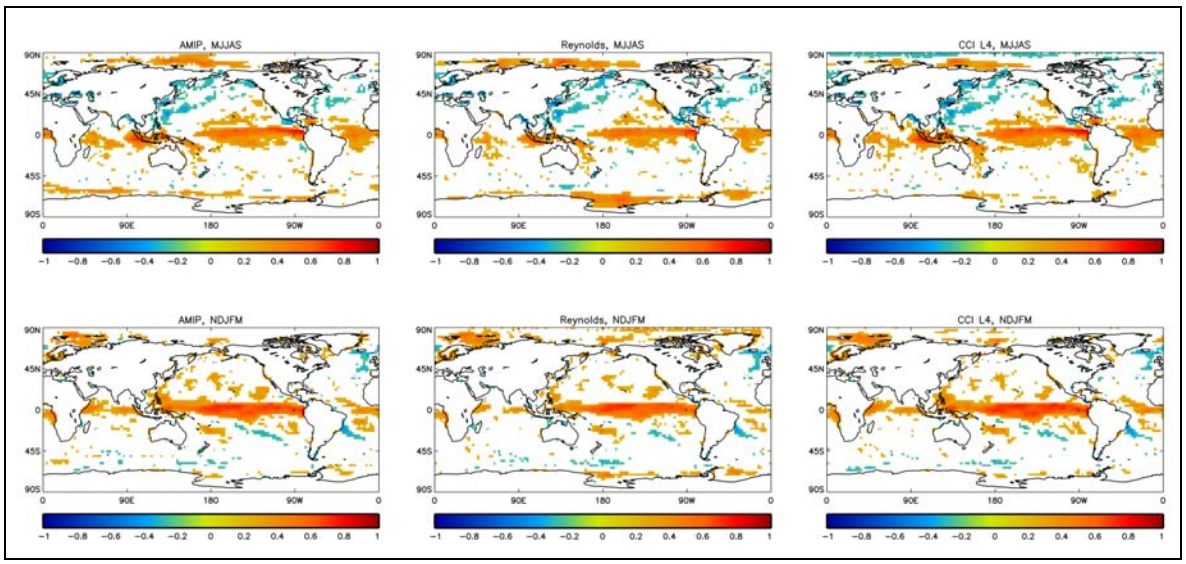


Figure 4-54: Correlation plot showing correlations between SST and precipitation for the summer and winter seasons which are significant at the 5% level.

Figure 4-54 shows areas of negative correlation in the northern hemisphere mid-latitudes which are not seen in Figure 4-36, the correlation plot encompassing all twelve months. This may be an artefact of other processes such as the way atmospheric circulation drives precipitation. Nevertheless, overall the relationship appears to remain consistent throughout the year between SST and precipitation with a dominant region of positive correlation across the tropical pacific associated with ENSO.

The three datasets are generally quite similar and display similar spatial patterns, however there are some differences to note. In MJJAS there is a region of positive correlation on the Eastern edge of Australia. This correlation appears more positive in the SST CCI analysis plot compared to AMIP whilst the Daily OI is weaker still. Further, during MJJAS the Daily OI data displays a large area of positive correlation around the Ross Sea which is not seen in either of the other two datasets. This may be due to trends in the Daily OI, not seen in either AMIP or SST CCI analysis, as a result of the artificial jump seen in 2002 when microwave observations are added to the data.

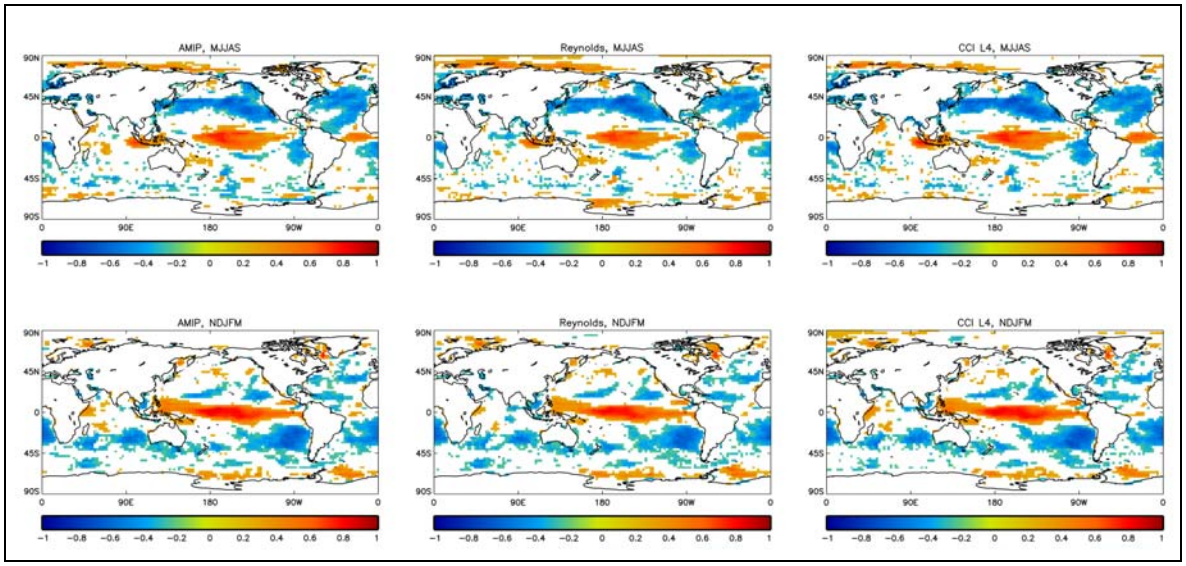


Figure 4-55: Correlation plot showing correlations between SST and cloud amount for the summer and winter seasons which are significant at the 5% level. The correlation coefficients are calculated for the period 1991-2008.

Figure 4-55 highlights the marine stratocumulus regions; these can be seen as a negative correlation in the summer months for each hemisphere. The clouds develop as the SST cools and boundary layer stability increases. There is an area of negative correlation signifying low level cloud off the South coast of Australia which is confined to the northern hemisphere winter. In addition, there is negative correlation in the marine stratocumulus region off the west coast of Africa in the MJJAS plots. This begins to occur in September/October/November and has been captured by the northern hemisphere summer plots.

4.2.7.2.4 Conclusions

The Daily OI data set is the coolest of the three data sets between 1991 and 2002. However when microwave observations are added to the dataset in June 2002 there is a jump in the anomalies which make the dataset warmest overall. The AMIP dataset is warmest with a difference of approximately 0.15K between it and the Daily OI.

SST variability is greatest in the mid-latitudes and in equatorial regions, for example, in the Tropical Pacific and in the Atlantic around the Gulf Stream. All three datasets show similar SST standard deviation spatial patterns although AMIP is slightly less variable with a lower standard deviation than SST CCI analysis and Daily OI in areas. Disparities between the datasets are generally small.

The Daily OI dataset is more weakly correlated to precipitation in the point-by-point correlation plot than the other two datasets. The SST doesn't reach the precipitation anomaly peaks and trough extremes as well as the AMIP and SST CCI analysis data do, seen in Figure 4-38, Figure 4-39 and Figure 4-41. The teleconnections with Nino region 3.4 are similar between all datasets, the Daily OI does not appear noticeably weaker correlated here. Correlation patterns are most pronounced in the correlation plot with El Nino periods only (Figure 4-45) whilst there is little distinct pattern in the same plot during La Nina's (Figure 4-46). There are only slight differences in correlation strength and spatial pattern between the three datasets.

The overall point-by-point correlation between cloud amount and SST is negative, although there is a strong positive correlation in the equatorial tropical pacific. The relationship between SST in ENSO region 3.4 and cloud amount is strongly positive across the equatorial tropical pacific as well as a large region of strong negative correlation in the Indonesian warm pool area. Differences between the three datasets are small in these plots particularly in Figure 4-51 which displays the relationship between ENSO regions 3.4 and SST. In the point-by-point correlation plot, Figure 4-50, the Daily OI data is most weakly correlated and SST CCI analysis is strongest, although is very similar in strength to AMIP.

Precipitation and SST are most strongly correlated between SST at time t = t – 1 and precipitation at time t = t. This is also the case for cloud amount in regions where the correlation is positive. Conversely, in marine stratocumulus regions the correlation ceases to be significant in the lag plots. The correlation differences in the lag plots are consistent with previous findings, that is, Daily OI is the most poorly correlated and SST CCI analysis is the most strongly correlated.

4.2.8 ROBERT MCEWAN: DIAGNOSIS OF OCEAN MODEL ERRORS

Name of user Robert McEwan

Institution of user *Met Office Hadley Centre*

4.2.8.1 Scientific analysis

Aim: Determine the 4km Persian Gulf Model (PGM4) SST errors that may be significant to the European Regional Seas Ecosystem Model (ERSEM).

Method: Comparison of SST CCI analysis seasonal averages to model surface level SST seasonal averages.

Results: The model has a warm bias across the Persian Gulf during spring summer and autumn. During winter the model has a cool bias in coastal regions and a warm bias in central regions.

Conclusions: The model displays a general warm bias. This bias is small and persists for most of the year except in coastal regions during winter. This may have an effect on seasonal mixing and therefore impact ERSEM.

4.2.8.2 Feedback on scientific utility of the SST CCI products

The SST CCI analysis product was useful as there are limited in situ data for the region.

4.2.9 ALEXEY KAPLAN: COMPARISON TO OTHER SST ANALYSES

Alexey Kaplan

LDEO of Columbia University

4.2.9.1 Key messages

- Initial evaluation of the pre-release long-term SST CCI satellite-only SST-depth analysis created by the OSTIA system from SST CCI ATSR and SST CCI AVHRR products for the period Sep 1991 to Dec 2010 (hereafter OSTIA-CCI) showed that it is warmer than the OSTIA version currently in use (hereafter OSTIA-MyOcean) by about 0.15C on average.
- This warm shift makes OSTIA-CCI[7] on average more consistent with in situ and microwave SSTs than OSTIA-MyOcean is, suggesting that the OSTIA-CCI might be an improvement on the OSTIA-MyOcean, and by extension, that reprocessed SST CCI L2P data sets are improvements on their earlier versions. (Although more detailed evaluations are necessary).
- The SST CCI data file system seems well-organized and well-documented (once these documents are located; suggest to put them into the docs/ ftp subdirectory).
- The only technical difficulty in using the SST CCI data, which was felt by this user, was the slow data transfer rate of the NEODC ftp site.

4.2.9.2 Scientific analysis

4.2.9.2.1 Aims of the study

To see what effect consistently-produced (and in situ-independent) SST CCI L2P products can have on the L4 analysis products, which are using them as inputs into the L4 analysis procedure.

4.2.9.2.2 Method

The pre-release long-term version of the satellite-only SST-depth analysis created by the OSTIA system from SST CCI ATSR and SST CCI AVHRR products for the period Sep 1991 to Dec 2010 (hereinafter called OSTIA-CCI) was obtained from the NEODC ftp site and compared with the “traditional” OSTIA, as obtained from the myocean.eu.org. (OSTIA reanalysis was extended by operational OSTIA through 2010), hereinafter called OSTIA-MyOcean. To clarify the differences, comparisons with the ICOADS Release 2.5 1x1 degree enhanced MSG monthly summaries for SST were made as well as with some other data sets of interest.

4.2.9.2.3 Results

Mean differences between the two OSTIA products for the entire long-term SST CCI product period (Sep 1991- Dec 2010) and standard deviations of monthly mean differences around their long-term means are shown below in Figures 4-55 and 4-56, respectively. Standard deviations exceed 0.5C only in the Gulf Stream and Kuroshio areas and in some of eddy-rich areas. Incidentally, these are the same areas where mean differences are small or negative. Everywhere else mean differences are positive (the

[7] *This is Kaplan’s nomenclature in this section for the SST CCI analysis product.*

SST CCI product is warmer than the MyOcean one), and they exceed 0.3C in the ITCZ, SPCZ, and other cloudy areas near equator. This behaviour not only occurs in long-term means but persists most of the time, as can be seen in Y-T diagram for zonal means of monthly differences (Figure 4-57). Sharp temporal changes in zonal mean differences (beginning of 1995, early 1996, middle of 2002, beginning of 2008) are likely of non-climatic origin and are due to discontinuities in the input data coverage for both products and in the analysis procedure of the MyOcean product.

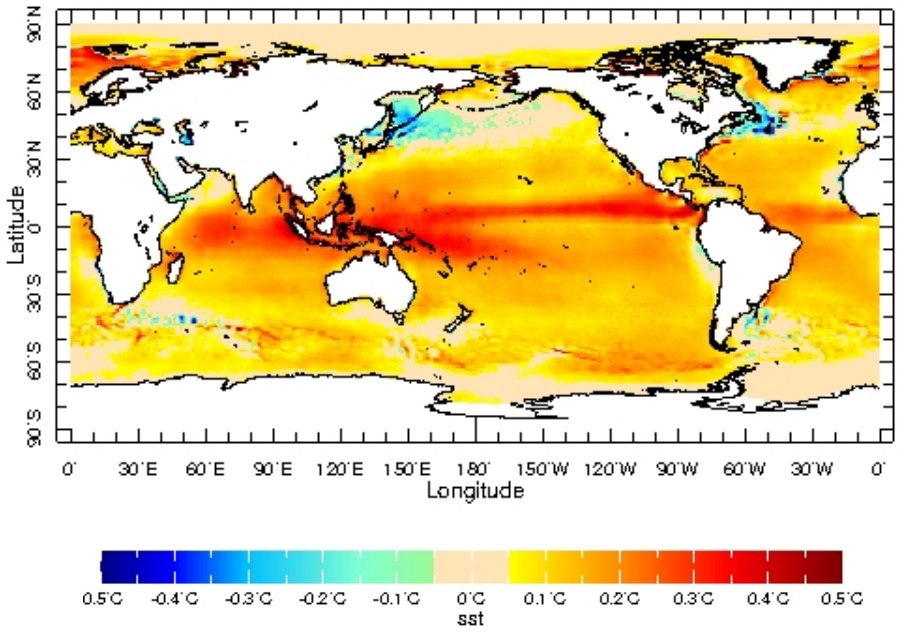


Figure 4-55: Mean difference <OSTIA-CCI minus OSTIA-Myocean> (Sep 1991 to Dec 2010 period)

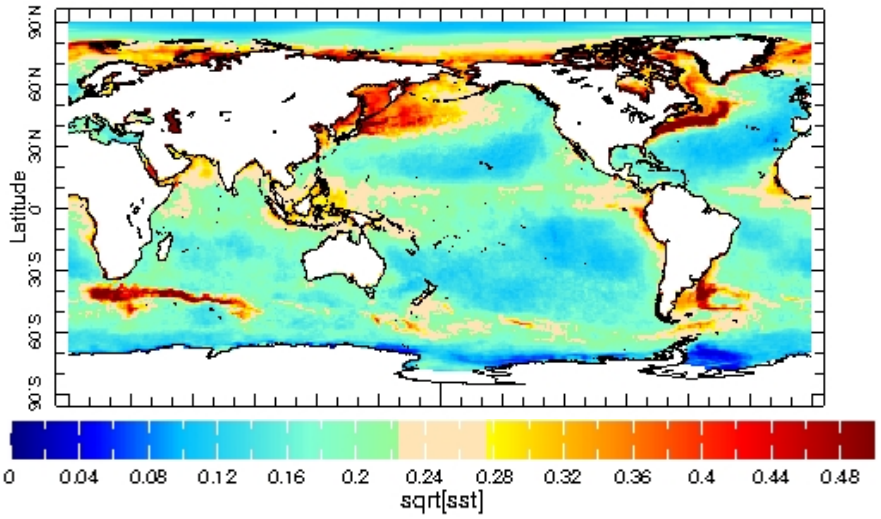


Figure 4-56: Standard deviation of (OSTIA-CCI minus OSTIA-MyOcean) for Sep 1991 - Dec 2010 period

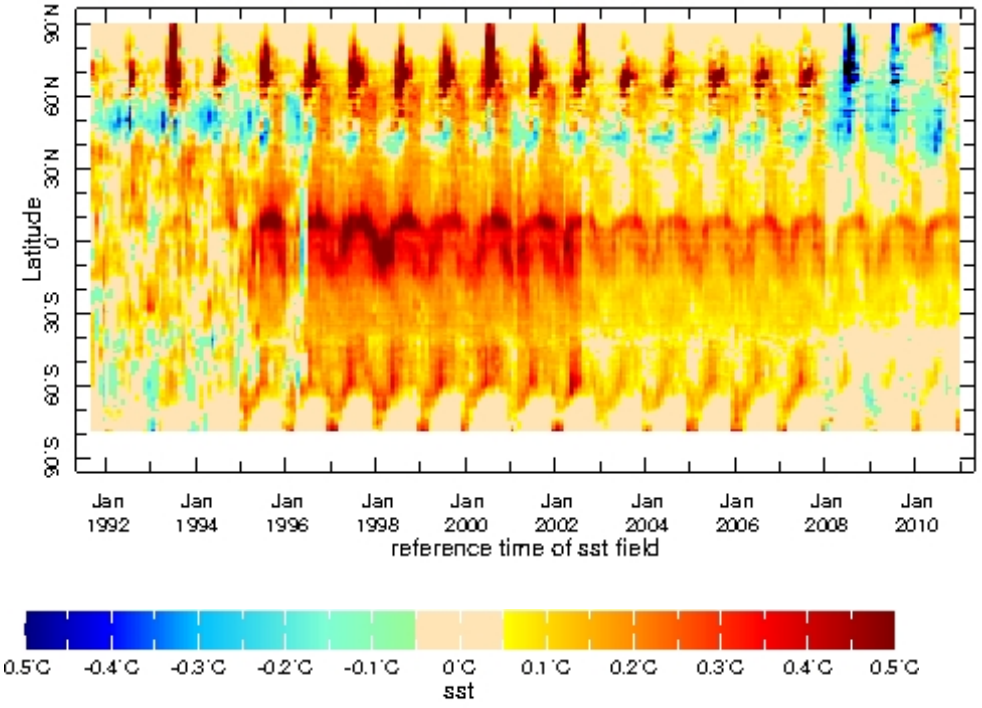


Figure 4-57: Zonal mean monthly difference <OSTIA-CCI minus OSTIA-MyOcean> (Sep 1991 to Dec 2010 period)

It is easy to deduce from the mean difference values in Figures 4-55 and 4-57 that the SST CCI product has higher global mean SST than the MyOcean one. Indeed, as Figure 4-58 shows, it is so for each individual year as well (1992-2010). In the period 1997-2001 OSTIA-CCI is warmer than OSTIA-MyOcean by more than 0.2C in annual global means (60S-60N).

Annual mean difference of 0.2C in global mean is not a small value and needs to be explained away by, e.g., tracing it to differences between the input data sets used by the two OSTIA products. However, as Figure 4-58 shows, this shift towards higher values brings the OSTIA analysis product closer, on average, to the products that use in situ (ICOADS and HadISST1) and microwave data types (NCDCirmw), which were not used in OSTIA-CCI (even though they were used in OSTIA-MyOcean). This improved consistency with independent data is a very encouraging sign for the SST CCI products, although more detailed evaluation is necessary.

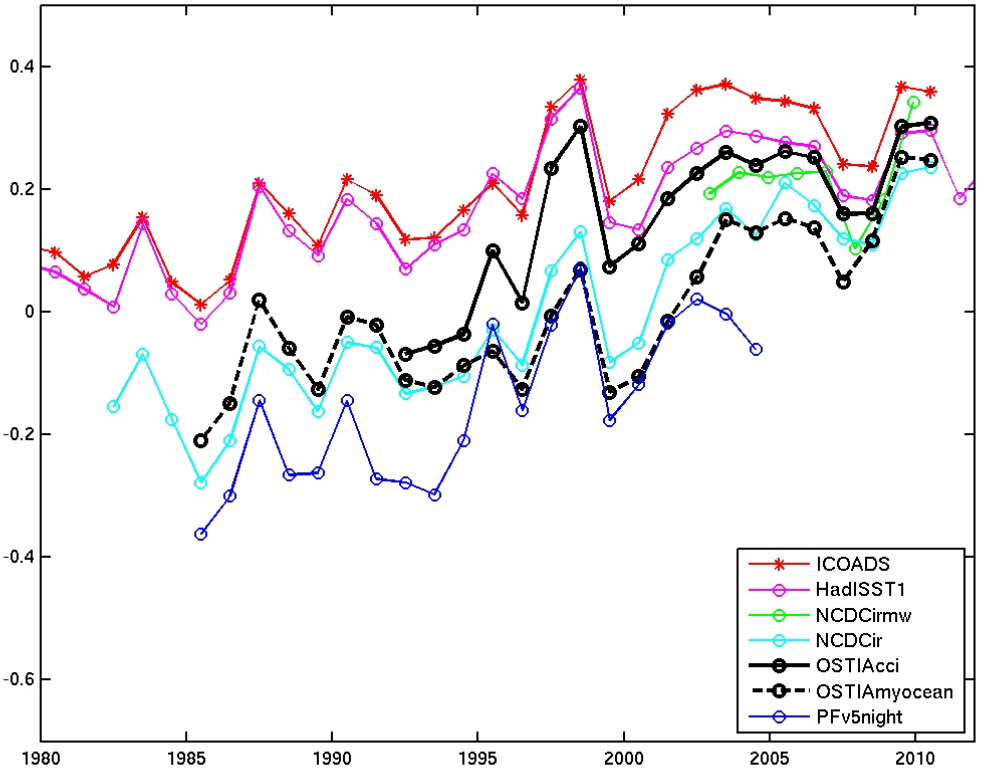


Figure 4-58: Global (60S-60N) means for annually averaged (calendar year) anomalies w.r.t. NCEP OI.v2 climatology for different SST products. In addition to OSTIA-CCI and OSTIA-MyOcean shown are ICOADS Release 2.5 1x1 Enhanced MSG monthly summaries, HadISST1, NCDC 0.25deg Daily OI versions based on the IR only and on both IR and MW data, and Pathfinder Night SST, v5.0.

4.2.9.2.4 Conclusions

OSTIA-CCI is warmer on average than OSTIA-MyOcean, their global annual means (60S-60N) exceeding 0.2C in some years (1997-2001). This warmer shift actually brings the OSTIA fields closer to independent data (in situ, microwave SST), at least on average, which is an encouraging sign.

4.2.9.3 Feedback on scientific utility of the SST CCI products

Only the OSTIA-CCI has been explored in any detail so far. But the structure of the SST CCI data base, namely providing the L4 analysis product based on consistently

reprocessed L2P inputs, as well as L3U and L3C compilations of these inputs[8], is exactly what is needed for developing scientific understanding of properties of L4 analysis products and explaining the differences between different L4 analysis products. So SST CCI products are going to be very useful.

[8] L3U and L3C compilations of the L2P SST CCI products are not, in fact, provided by the SST CCI project thus far. SST CCI AVHRR is provided in L2P format and SST CCI ATSR as L3U format. Nonetheless, this user comments that it would be beneficial to provide L2P, L3U and L3C versions of all products and this will be fed into the SST CCI Phase II User Requirements analysis.

5. COMPARISON TO OTHER CCI ECVS: OCEAN COLOUR

5.1 Introduction

It is important to explore the feedback between chlorophyll (from phytoplankton) and temperature, as this impacts:

- the ocean mixed layer depth and fluxes of heat;
- the carbon cycle. Although, in some places, the carbon cycle can be physically driven, in others it is biologically driven. Here, the carbon cycle is influenced by the amount of plankton present; carbon sequestration is achieved through the life cycle of plankton with larger shells. Chlorophyll retrievals can be used to explore where the carbon cycle is biologically driven; and
- plankton succession, which is important for studies of food security, via the follow on from plankton to higher organisms. Tuna fisheries are particularly interested in phytoplankton to understand the migration patterns of tuna.

SST and ocean colour (OC) then are intimately connected. Frontal regions in the global oceans tend to be associated with higher biological activity, leading to areas of high chlorophyll (from phytoplankton). With the CCI products for both SST and OC we have the benefit of information on high spatial resolution: 4km in both cases. At this resolution, it is possible to identify the surface signal of ocean fronts from both SST and OC data, since an ocean front is a boundary between two distinct water masses with different temperature (and other) properties. This study briefly examines the detection of ocean fronts in the SST and OC CCI products with a view to exploring (i) to what extent they provide a coherent picture of contemporaneous thermal and biological frontal activity and (ii) how that picture relates to that provided by the MODIS instrument, from whose measurements both SST and OC can be retrieved simultaneously. Using new data sets to help us to confirm or determine the details of relationships between phytoplankton and temperature is valuable. However, those new data sets need to be checked for representation of basic features first before more complicated analyses (such as the exploration of lag relationships between chl-a and SST fronts or studies involving different plankton functional types) can be undertaken or justified.

Comparisons between chlorophyll-a (chl-a) and SST fronts on a global scale are not appropriate as a means to assess consistency between SST and ocean colour products due to the differing physical and biological forcing mechanisms that govern their evolution. Temperature is one of several factors that contribute to making any one location a suitable environment for plankton growth. Nonetheless, it is an interesting coupling to explore, in general, as it sheds further light on the processes involved and may better enable us to monitor such processes in future. There are a number of regional cases where it is appropriate to explore the consistency between fronts seen in SST and chl-a retrievals.

The Single Image Edge Detection (SIED) algorithm [Cayula and Cornillon, 1992, RD.346] provides a fast and objective means of detecting ocean fronts in satellite-derived fields of SST and chl-a. Whilst Cayula and Cornillon [1992, RD.346] applied the SIED algorithm to AVHRR SST fields with a resolution of 1-km, later studies have successfully used SIED on daily SST and chl-a fields at a range of grid resolutions between 1-25 km [Stegmann and Ullman, 2004, RD.347; Castelao et al., 2006, RD.348; Belkin and Cornillon, 2005, RD.349; Belkin and Cornillon, 2009, RD.350; Wall et al., 2008, RD.351; Obenour et al., 2010, RD.352; Kahru et al., 2012, RD.353]. The performance of the SIED algorithm has been assessed through a comparison with other automated front detection methods, with

SIED detecting fewer false fronts and detecting more true fronts than other methods (Ullman and Cornillon, 2000, RD.247).

5.2 Data

Two sets of SST and chlorophyll-a fields were used in this study for each day during 2003.

The precursor set was data from the MODIS instrument on the Aqua satellite (hereafter MODIS-A): these provide estimates of both SST and chl-a from coincident measurements by the same instrument. [MODIS-A 4km data were downloaded from http://oceandata.sci.gsfc.nasa.gov/MODISA/L3SMI/]

Compared to these precursor data were fields of SST and chl-a from the SST and OC CCI products. SST was taken from the pre-release version of the SST CCI AVHRR product and chl-a was taken from v0.95 of the OC CCI L3 4km product, which is a combination of data from MODIS, SeaWIFS and MERIS [Storm and Boettcher, OC CCI Product User Guide, RD.354]. Prior to use, all SST CCI AVHRR retrievals for each day (from all available instruments) were gridded onto a 4km latitude by 4km longitude grid.

Neither the SST nor OC CCI products provide a choice of quality flags. No flags are given in the OC CCI product and only data with quality flag 5 (best) is given in the SST CCI products. MODIS-A products provide data with a range of quality flags from 0 (best) to 2 (poorest). The impact on the results of the use either of MODIS-A data with only quality flag 0 or with any of quality flag 0-2 was determined.

5.3 Method

The SIED algorithm was designed to operate using a Digital Number (DN) representation of SST, which has been filtered for cloudy pixels and had a 3x3 median filter applied [Cayula and Cornillon, 1992, RD.346]. The SIED algorithm requires both the 'true' geophysical field (i.e. as °C for SST or mg/m^3 for Chl-a) and the DN representation of this field as inputs. What now follows is a basic description of the SIED algorithm.

Having obtained cloud-free, median-filtered fields (true geophysical values and the DN representation), automated front detection with the SIED algorithm follows a series of steps:

1. Create a square window (i.e. 32×32 pixels), which slides across the DN field with a step size equal to half its window size (i.e. 16 pixels).

2. Run a histogram analysis on each window, to test whether the data within a window displays a bimodal distribution. Whilst the presence of a unimodal distribution implies the presence of one water mass, a bimodal distribution suggests the presence of two water masses that are separated by a front. Find the threshold value (i.e. DN value) that best separates the two water masses.

3. If the data within the window displays a sufficiently bimodal distribution, test for the spatial cohesion of the two water masses. This tests whether they form two distinct 'patches' within the window (a front) or if they form a 'checker-board' pattern (noise).

4. If the window has two water masses, which are also spatially distinct, the algorithm flags this window as having a front. The position of the frontal pixels within this window are marked using the threshold value previously identified that best separates the two populations.

5. Once all of the windows of data have been processed within the field, the algorithm outputs a field with grid values set to the number of times a front was identified at this grid location. An edge-thinning algorithm is run over this field to reduce the width of the fronts to being 1-pixel wide.

6. Finally, a contour following algorithm is run to extend the fronts at either end, provided that the underlying gradient is sufficiently high and continues in the direction of the front. This extends fronts where the presence of a strong front may have been missed by the histogram and cohesion step.

7. The output of the SIED algorithm is a binary-field indicating which pixels in the input field are located at ocean fronts (0 = no front, 1 = front).

Consideration needs to be given to what window size is used to search within for ocean fronts, and the tolerance threshold on how bimodal a distribution needs to be in order to identify a window as containing a front. The original implementation of the SIED algorithm used a window size of 32×32 pixels and a tolerance threshold of 0.76 [Cayula and Cornillon, 1992]. Window sizes between 16-32 pixels are suggested: any smaller and there will likely be too few data to examine in the histogram step, any larger and the likelihood of multiple fronts within the window increases the probability of the algorithm missing some fronts.

The tolerance threshold should be set at a value between 0.72-0.78. Increasing this value provides more certainty on the presence of two water masses, whilst relaxing this threshold increases the detection rate of less-distinct fronts. Here, a value of 0.76 was used.

5.4 Discussion of results

Fronts off the southern Mexican coast line were examined in detail and are summarised in this section.

Figure 5-1 compares fronts detected in MODIS-A and SST CCI fields for 16th January 2003. The swath of missing data in the MODIS-A field arises from the gap between overpasses and is a consequence of the use of data from a single MODIS instrument in this study. The combination of different sensors in the SST CCI SST CCI AVHRR product allows more of the cloud-free ocean to be observed. Clearly, the two fields are observing the same system of fronts and many coincident detections can be seen.

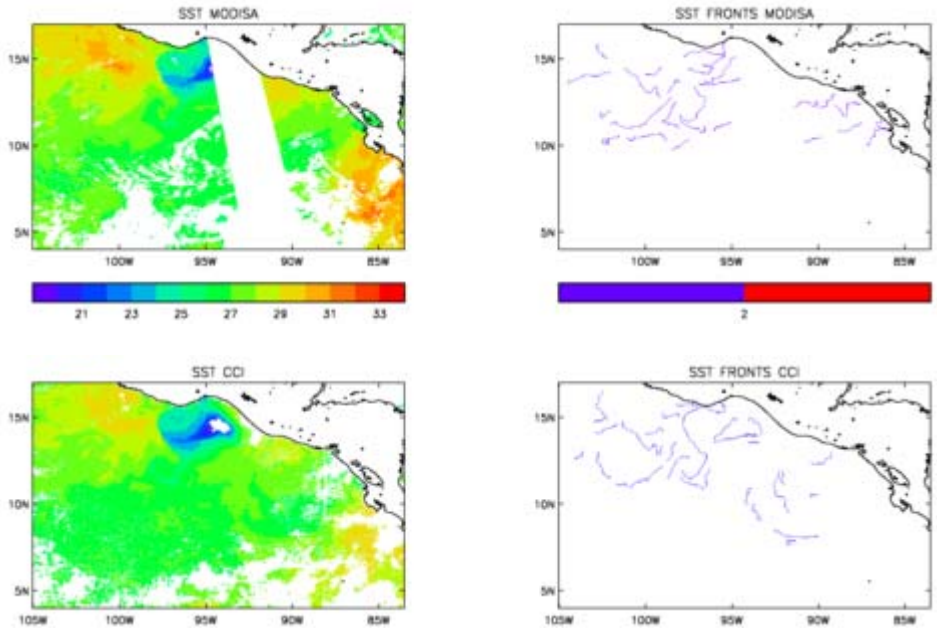


Figure 5-1. Analysis of fronts detected in SST fields off the southern Mexican coast on 16th January 2003. Top: MODIS-A. Bottom: SST CCI SST CCI AVHRR averaged onto 4km grid. Left: SST (°C). Right: frontal positions.

The same is true of Figure 5-2, where frontal detections from chl-a fields in MODIS-A and OC CCI products are shown. Again, the combination of different instruments in the OC CCI merged product allows more of the cloud-free ocean to be observed and the portions of fronts that are detected in both products have the same location and the same shape.

Considering the correspondence between the SST and chl-a fields (cf Figure 5-1 and 5-2) we see, as expected, a high chl-a concentration where an area of cold SST is seen around [95W, 14N]. However, it appears that the SST CCI SST CCI AVHRR product has flagged the coldest part of this water here as cloud, where the OC CCI product clearly shows an area of very high chl-a concentration, indicating that perhaps the SST should be cold here. This is highlighting a deficiency in the cloud-clearing scheme used for the SST CCI SST CCI AVHRR product and is an example of how expected consistency between different variables could be used to highlight potential problems in one or other retrieval. In contrast, the area of missing data around [84W, 8N] seen in the MODIS-A, SST CCI SST CCI AVHRR and OC CCI L3 products is robustly detected in all cases and so might be definitely concluded to be cloudy.

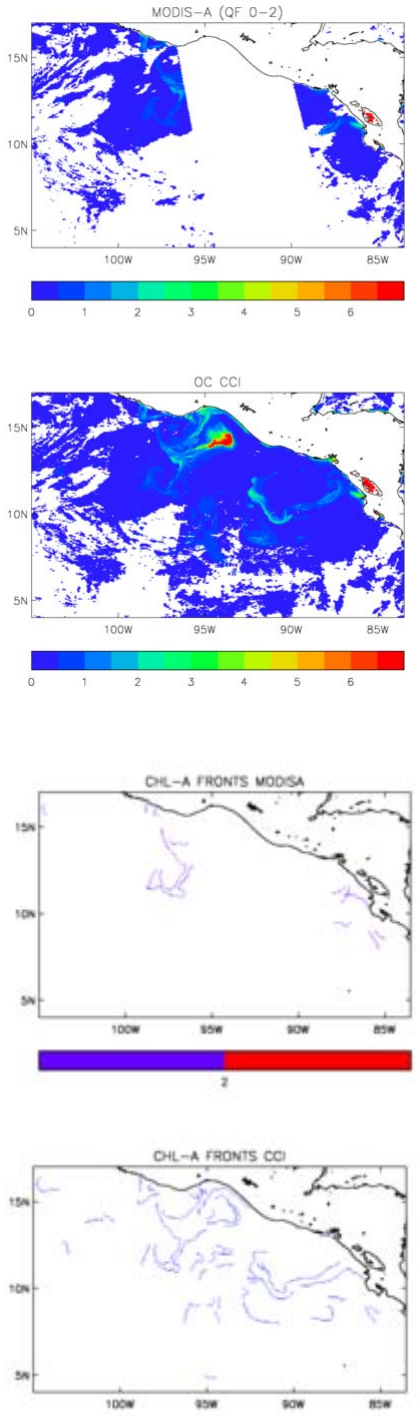


Figure 5-2. Analysis of fronts detected in chl-a fields off the southern Mexican coast on 16th January 2003. Top: MODIS-A chl-a concentration (mg/m^3). Second: OC CCI L3 v0.95 chl-a concentration (mg/m^3). Third: MODIS-A frontal positions. Bottom: OC CCI L3 frontal positions.

Looking in more detail at the coincident fronts detected in the SST and chl-a fields for 16th January 2003 (Figure 5-3), we see, in this case, that more coincident fronts have been detected in the CCI data than in the MODIS-A data. However, this is not always the case.

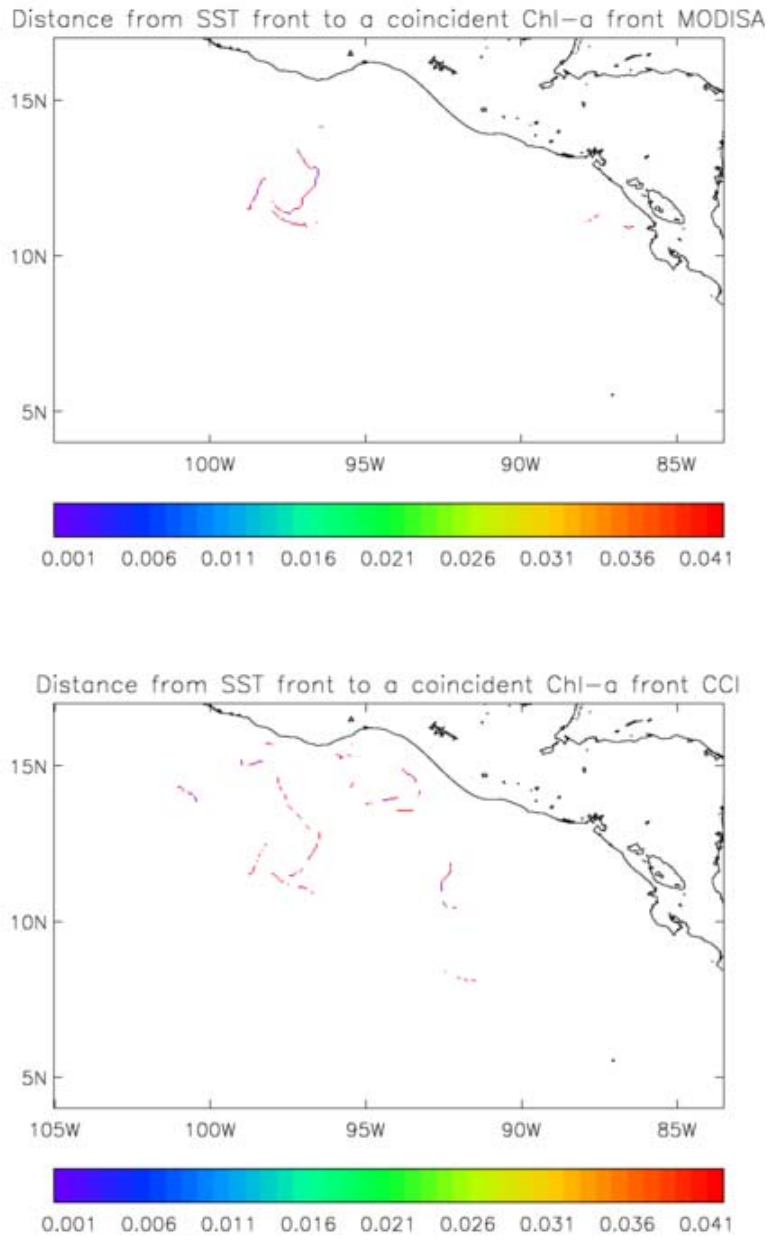


Figure 5-3. Distance between coincident SST and chl-a fronts (degrees). Top: MODIS-A (QF 0-2 used) Bottom: SST and OC CCI (see text for details).

While it is useful and interesting to examine the similarities and differences between the representation of fronts in these products on a single day, our observations don't necessarily hold for other days. So, we explore some summary statistics for the year 2003 as a whole in Figures 5-4 and 5-5.

Figure 5-4 gives the probability for each SST grid box that it contains a frontal pixel. In general, the frontal probabilities are higher in the MODIS-A data set, but the general geographical spread of the distribution is similar. Higher probabilities might arise here from random preferential sampling of the days where fronts occurred.

We explore the impact of the use of the MODIS-A quality flag here. When all data are used (QF = 0-2, bottom right panel), the pattern of frontal probability appears more diffuse than that of the SST CCI SST CCI AVHRR product. When only best quality data are used (QF=0, top left panel), the pattern is slightly more organised in places, but otherwise very similar.

The OC CCI chl-a frontal probabilities (Figure 5-5) are clearly centred around the same geographical locations as those from the SST products. Chl-a frontal probabilities appear generally higher, which may reflect the occurrence of fronts in the chl-a data arising from different processes than those determining the SST fronts.

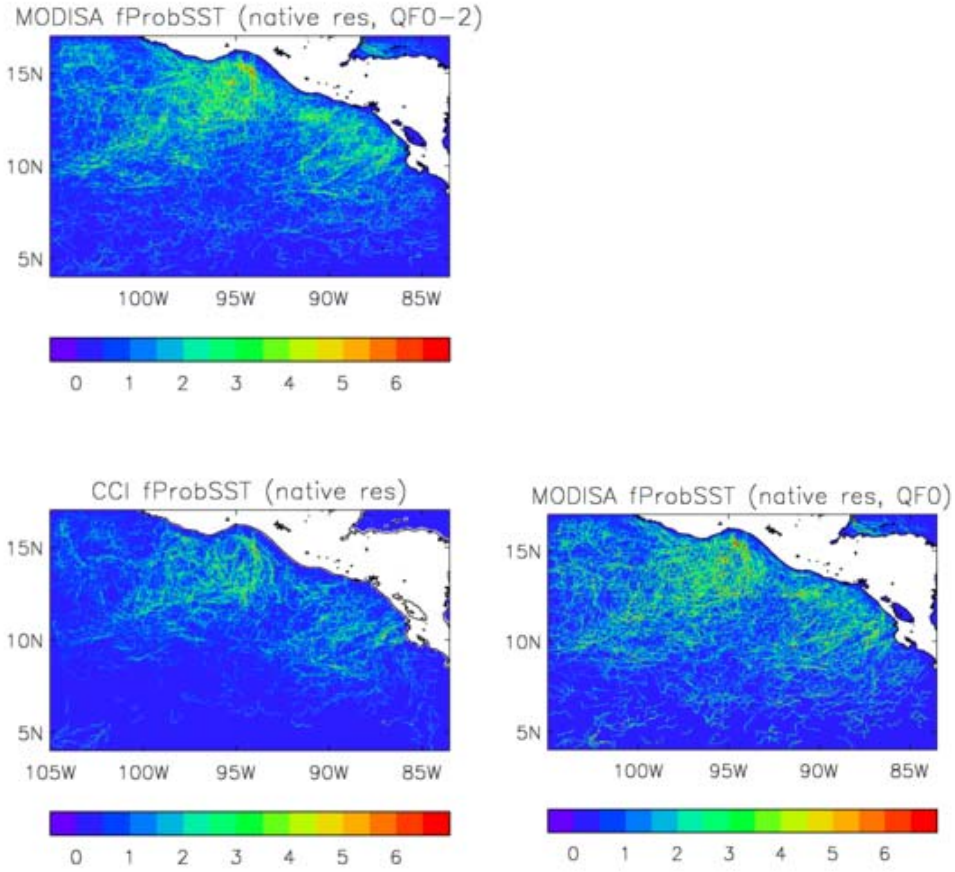


Figure 5-4. SST frontal probabilities in each 4km grid box. Top left: MODIS-A (QF0-2). Bottom left: SST CCI SST CCI AVHRR product. Bottom right: MODIS-A (QF=0 only).

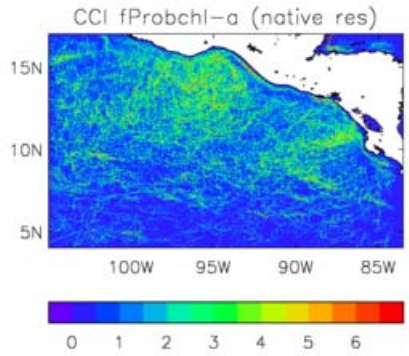


Figure 5-5. OC CCI Chl-a frontal probabilities in each 4km grid box.

We explore the frequency of coincident SST and chl-a fronts (on the same day and in the same 0.25° area) over the whole of 2003 in Figure 5-6. The frequency of detecting coincident fronts is much larger in the CCI products than in the MODIS-A product, although this is likely to be strongly affected by the relative data availability. Indeed, the use of only the best MODIS-A data reduces the frequency (but not the pattern) significantly. This robust pattern indicates that there is useful information on fronts to be found in the poorer quality MODIS-A data. It is clear from Figure 5-6 that coincident SST and chl-a fronts are a common feature of this location and that they are strongly organised in certain locations, although the exact locations do appear to vary from year to year (by comparison to the results from 2008, not shown).

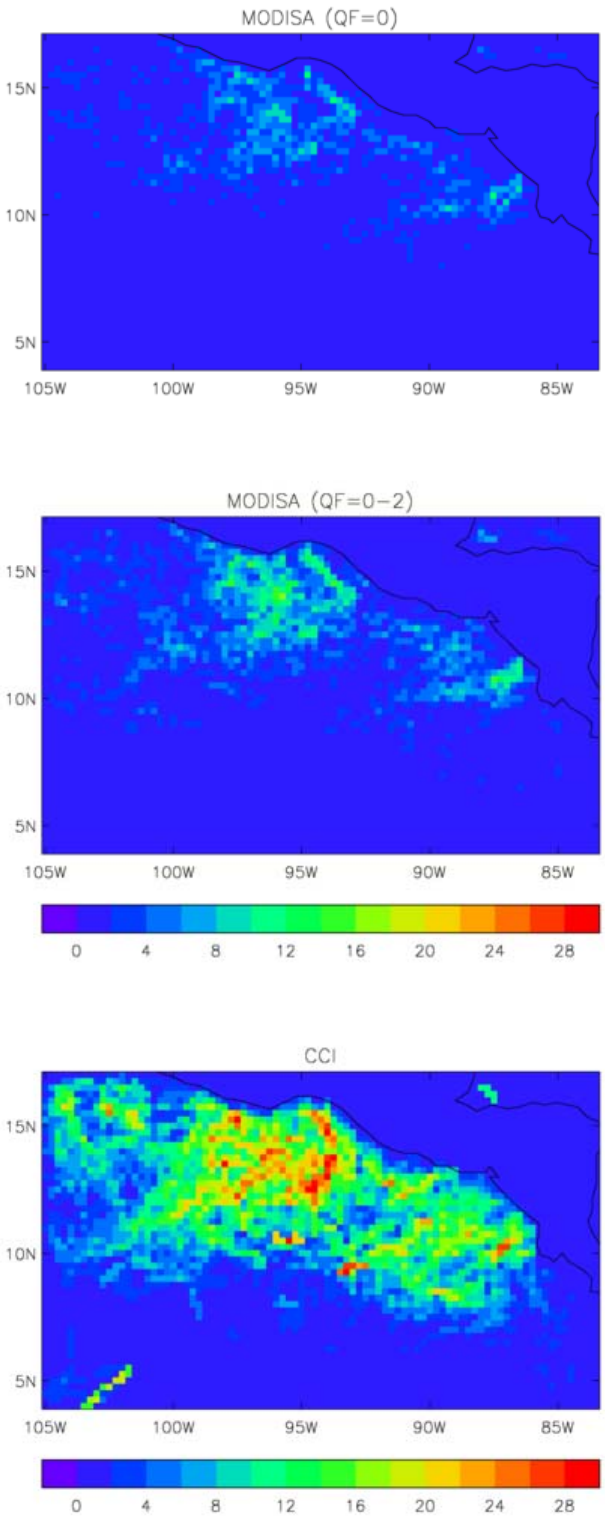


Figure 5-6. Frequency of occurrence (days) of at least one coincident SST and chl-a front within a 0.25° area over 2003. Top: MODIS-A (QF=0). Middle: MODIS-A (QF=0-2). Bottom: CCI.

5.5 Summary of key points

- SST and OC CCI products contain information on SST and chl-a concentration off the south coast of Mexico that appears to be largely consistent on daily and 4km scales.

- SST and OC CCI products, when analysed together, provide a credible source of information for the analysis of coincident SST and chl-a fronts.

- Combining sensors in the OC and SST CCI products provides sufficient coverage of the area south of Mexico to perform daily analyses of frontal positions and to produce summary statistics over a year (2003).

- There may be merit in providing lower quality data alongside best-quality data, as it appears to contain useful information for the analysis of ocean fronts.

- Coincident SST and chl-a fronts are a common feature of the ocean south of Mexico and are strongly organised in certain locations, although the exact locations appear to vary from year to year.

- Coordinated analysis of strongly connected variables can highlight potential deficiencies in aspects of the processing of either variable.

6. FURTHER ISSUES AND RECOMMENDATIONS REPORTED BY THE TRAIL BLAZER USERS

6.1 Positive feedback on ease of use of the products

- "The product has been easy to use as the output files are CF compliant making them easy to read into to IDL and python."
- In the context of producing a NW Shelf Reanalysis, the processing of the SST CCI products was more straight-forward than the corresponding Pathfinder v5.2 products used in the early years of the reanalysis.
- "The datasets were very easy to use and the documentation was fine for my purposes."
- "The products are very easy to access and use, since they are in netcdf format and follow the GHRSST GDS specification."
- "The instructions on how to download the data, as well as the documentation are understandable and relatively easy to follow."
- In the context of evaluation of a Persian Gulf model, the product required sub-sampling, re-gridding and time averaging, although this was not difficult.
- The documentation (PUG) is very well written and organized.
- "Once the data is brought in, its use is very smooth. All data sets seems well-organized, their structure well-described, and so far I haven't seen any incorrectly filled fields, parameters etc."
- "I want to give special praise to consistency with GHRSST GDS-2.0 standard, and within this standard a provision that time should be a record dimension. I've processed all files using NCO utilities (to prepare them for the use with our in-house visualization/manipulation program Ingrid), and thanks to this provision all files could be easily concatenated along the time dimension for further processing."

6.2 Reported Issues

- Data download speeds were reported to be an issue by more than one user:
    - The ESA SST CCI analysis fields were downloaded from ftp.ceda.ac.uk. The typical download speed was about 100kb/sec, which means that it took one computer using Filezilla (https://filezilla-project.org/) more than 6 days to download the 20-year data set. This download time can perhaps be reduced by compressing the NetCDF files and making the data set available via, for example, JPL PO.DAAC. [This user was in the U.S.]
    - "The only real difficulty was a speed of data transfer from the NEODC website: 150K/s, 200K/s at best. So it took a few days of concerted efforts to bring in the data. In the end I was running 5 or 6 wget jobs simultaneously to get a decent transfer speed, carrying each year of each data set by a different wget command." [This user was in the U.S.]

- "When trying to access the NetCDF files, using Matlab 2009a and functions (http://www.marine.csiro.au/sw/matlab-netcdf.html) that worked for both the Operational and Reanalysis OSTIA SST data products, we encountered Java Heap Space errors. This seems odd given that the fields for all three products have the same dimensions (3600 rows x 7200 cols) and the CCI NetCDF metadata states the following, "summary = "OSTIA L4 product from the ESA SST CCI project, produced using OSTIA reanalysis system v2.0". We were able to resolve these issues by using Matlab2010a, increasing the Java Heap Memory, and using a different NetCDF tool box (http://sourceforge.net/apps/trac/njtbx)"

- SST values with quality flag of 5 were assigned a fill value. This caused the DMI-OI processing system to break down in the initiation of the runs. This was reported to the project team. [Note. This has been rectified in the release version of the products.]

- It is an oversight, however minor, not to put any documentation into the docs sub-directory of the data ftp site.

6.3 Recommendations

- "It would be useful to work with the team to work out we could make use of the uncertainty estimates in the product improve evaluation of climate models."

- "It would also be useful to understand the discrepancy in the SST during the 1997-1998 El Niño between the CCI dataset and the Reynolds and OSTIA v1 climatologies." [Note: we think the SST CCI analysis is giving a better representation of this event (see earlier sections).]

- For future applications where the diurnal SST cycle in the assimilation of SST observations may explicitly be accounted for, e.g. in NW Shelf Seas reanalysis, it would be useful to also have the observations supplied without the diurnal correction applied.

- Future reanalyses would also benefit from a similar reprocessing of earlier observations to reduce dependency on the Pathfinder products and ensure a more consistent use of observations over the reanalysis period.

- In the context of high resolution modelling, the possible impact of the diurnal cycle of SST on the simulated variability in the climate model is of interest. When coupled atmosphere-ocean simulations are performed, the models are coupled every 3 hours in order to have some aspect of the diurnal cycle. It would be interesting to be able to do something similar with the forced atmosphere model. The model diurnal cycle over land is known to be poor, mainly due to the convective parameterisation – with only one value per day, it is difficult to investigate the ocean in the same way. This may be important for the Madden-Julian Oscillation, and possibly for Indian monsoon active/break cycles.

- No quality 5 observations with fill values should be in the data set

- A list of known gaps in the data could be useful when constructing a climate data set.

- SST CCI observations back from 1982 would be valuable.

- In the context of producing an Arctic analysis, i.e. in persistently cloudy regions of the high latitudes, it would be very valuable to have an SST CCI product based upon Passive Microwave observations.

- "My recommendation for the improvement of the ESA SST CCI analysis data set, would be the option to download a lower resolution version of the data set, or maybe even the option to only download the data of subregions. It would be also very helpful if both day and night (ascending/descending) values would be available in the analysis."

- Could the product be supplied regionally? Manipulating global fields is time consuming.

- "Just keep going! I've been very surprised by how bold and ambitious this project is, particularly in terms of its product level span (from L2 to L4 pretty much simultaneously) and its attention and effort w.r.t. uncertainty estimates. While I have not evaluated the latter yet, and to my knowledge, the kind of uncertainty estimates separated by spatial scale have not been well-established in the community, the authors seem to be moving ahead swiftly and boldly, and unless any weaknesses are discovered in these novel details in further evaluation, we should be only encouraging their momentum."

APPENDIX A HOW THIS WORK ADDRESSES THE RELEVANT TECHNICAL REQUIREMENTS

Table A-1. Summary table of relevant Technical Requirements and how they are addressed in this document.

Technical Requirement number	**Relevant parts of the description of the Technical Requirement**	**Section(s) of this document in which the Technical Requirement is addressed.**
[SST-TR-36]	2. Feedback from the user community on the utility and performance of SST test products **shall** be actively sought. 3. User conclusions and recommendations based on SST test products **shall** be analysed and translated into new requirements for further algorithm tuning or development steps.	Feedback from the user community on the utility and performance of SST CCI products has been actively sought and is documented in Section 4 and Section 6. This feedback will be analysed further for the purpose of the provision of updated User Requirements in Phase 2.
[SST-TR-37]	The consistency of ECV SST prototype products and other reference data products contributing to the overall SST ECV **shall** be tested working with the user community and SST ECV Science Team.	Comparison of SST CCI products to other SST climate data sets is documented in Section 3.1. Further comments on this from users can be found in Section 4.
[SST-TR-38]	The impact of CCI SST ECV prototype products on end user applications **shall** be tested working with the user community and SST ECV Science Team.	The impact of SST CCI products on end user applications has been assessed by the SST CCI Climate Research Group and a set of trail blazer users. This is documented in Section 4.

[SST-TR-39]	Working with other CCI Science Teams, SST ECV Science Team, and CMUG, the consistency of CCI SST prototype products with other ESA CCI ECV satellite data products (e.g. sea ice, ocean heat content estimates) **shall** be tested.	Consistency of SST CCI and OC CCI products in the representation of ocean fronts is assessed in Section 5.
[SST-TR-45]	The inter-comparison of spatial maps and regional/global mean time series of accepted and demonstrated realizations of climate SST products (e.g., [RD-17]) **shall** form part of the validation activity.	Spatial maps and regional/global mean time series of SST CCI products and other SST climate data sets are presented and discussed in Section 3.1.
[SST-TR-46]	The long term stability of each SST FCDR within this project **shall** be established and reported in a submitted peer reviewed journal article.	Section 3.2 assesses the long term stability of each SST FCDR produced by the SST CCI.
[SST-TR-47]	Trend analysis describing in quantitative terms the long-term behaviour of the time series of CCI SST ECV satellite products on global and regional spatial scales **shall** be performed and reported in a submitted peer reviewed journal article.	Trend analysis describing in quantitative terms the long-term behaviour of SST CCI products on global and regional spatial scales is presented in Section 3.1.
[SST-TR-48]	The long term consistency with established SST time series [*e.g.*, RD-17, RD-25, RD-40] **shall** be investigated.	Long term consistency of the SST CCI products with established SST time series is documented in Section 3.1.

This Page Is Intentionally Blank