Quality control and monitoring of *in situ* SSTs for satellite calibration and validation at NESDIS

Feng Xu[1,2] and Alexander Ignatov[1]

1. NOAA / NESDIS / STAR, Camp Springs, MD

2. CSU / CIRA, Fort Collins, CO

Abstract

In situ sea surface temperatures (SST) are used for calibration and validation (Cal/Val) of satellite retrievals. Quality of *in situ* SST is suboptimal and very non-uniform which requires careful Quality Control (QC). This paper presents a brief description of the QC algorithm for *in situ* SST implemented at NESDIS which includes checks for internal, mutual and external consistencies. An online in situ quality monitor, or iQuam http://www.star.nesdis.noaa.gov/sod/sst/iquam/, was set up to monitor global statistics of *in situ* SSTs with respect to a reference field (daily 0.25° resolution Reynolds SST), and serving quality controlled *in situ* SSTs to data users.

Keywords – sea surface temperature, *in situ* observation, buoys and ships, quality control

1 Introduction

In situ observations of sea surface temperature (SST) play an important role in calibration and validation (Cal/Val) of satellite retrievals. However, quality of *in situ* data is suboptimal and very non-uniform across different platforms and sensors [*Xu and Ignatov, 2010*]. Moreover, a small fraction of the data is corrupted by instrument, operator or communication errors. A few outliers, if included in Cal/Val statistics, may render its results unusable. On the other hand, rejecting abnormal but correct data can miss important climate signals [*Lorenc and Hammon, 1988*]. An effective quality control (QC) is badly needed for *in situ* SSTs to be correctly used in satellite Cal/Val.

This need has been long recognized and all Cal/Val efforts perform QC of *in situ* data. However, the practices adopted in the remote sensing community remain to be *ad hoc* and overly simplistic. For instance, outliers are often identified by merely a constant thresholding with respect to a reference (climatological or analysis) SST field [*e.g. Kilpatrick et al., 2001; Francois et al., 2001; Brisson et al., 2002*]. Some authors attempt to estimate the global thresholds from the data itself, and exclude data beyond e.g. mean ± 3 standard deviation (SD) with respect to reference SST [e.g. *O'Carroll et al., 2006; Merchant et al., 2008*]. At any rate, QC methods adopted in the remote sensing community are far superseded by more sophisticated, systematic and well developed procedures used in the meteorological community [e.g. *Slutz et al., 1985; Lorenc and*

Hammon, 1988; Woodruff et al., 1998; Worley et al., 2005; Ingleby and Huddleston, 2007].

Satellite Cal/Val is very demanding to the quality of *in situ* data and requires a flexible and scalable QC depending upon the specific Cal/Val task. This study describes implementation of such a QC for the purpose of providing QCed *in situ* SST for near real-time (NRT) Cal/Val applications. The algorithm includes duplicate removal, platform track check and SST spike check, the reference check and the cross-platform check. The latter two checks follow the Bayesian approaches proposed by *Lorenc and Hammon (1988)* and *Ingleby and Huddleston (2007)* with minor modifications. Algorithm and configurations are described in section 2. The monitoring system is described in section 3 where results are also presented. Section 4 concludes the paper and discusses future work.

2 Algorithm Development and Implementation

2.1 Principles

The basic rule of QC is to check the data for self-consistency and cross-consistency with other data. Commonly used QC checks were summarized in *Woodruff (2008)* based on the condition and the method. Those checks can be categorized into five major groups based on the physical principles they rely on.

- Prescreening – aiming to resolve data specific problems (e.g., duplicate removal, and data cleaning or re-organizing).

- Plausibility check – assuring that each individual field and relationships between different fields are realistic (e.g., field range check, geolocation check, ID versus platform type check).

- Internal consistency check – checking different measurements from the same platform for internal consistency (e.g., platform track check and SST spike check).

- Mutual consistency check – based on correlation between nearby measurements from different platforms. This check is referred to as ‘cross-platform check’ in this paper, but is also often referred to as the ‘buddy check’ [e.g. *Lorenc and Hammon, 1988*].

- External consistency check – based on consistency of individual measurements with the reference (or first-guess) SST field. It is referred to as ‘reference check’ in this paper, but is also sometimes referred to as ‘background check’ [e.g. *Lorenc and Hammon, 1988*].

2.2 Algorithm

The QC algorithm implemented in this study consists of several major checks: duplicate removal, geolocation check, platform track check, SST spike check, SST reference check and SST cross-platform check.

a) Duplicate removal

Duplicates arise from multiple receptions of the same report via different paths, or from merging different datasets. The algorithm checks the difference between any two records belonging to the same platform. Only latitude, longitude and time will be checked. Tolerances are set as the corresponding digitization precision of each field, for example, 0.01 degree for latitude and longitude.

For a group of "duplicates," the one with the best quality is kept. If the quality information is not available and if all the duplicates have SSTs within 0.1°C tolerance, then the first in the sequence is kept and the rest are dropped; otherwise, all of them are dropped.

In practice, reference check described in item e) below usually precedes duplicate removal as it only uses external information and so is not affected by duplicates. Quality information from reference check is then used in duplicate removal to select the record with the best quality.

b) Geolocation check

Geolocation check checks if the location of a platform is plausible. For instance, SST measurements should not be reported over land, and buoy measurements are supposed to be located in the region indicated by the area code embedded in its WMO ID number. Geolocation check may also remove those reports located too close to coastlines (< 10km), depending upon the resolution and the accuracy of the water mask employed (the UMD 1km land cover classification [*Hansen et al., 2000*] is used in this study). Near-coast *in situ* SSTs are highly variable in space and time, due to shallow waters and high dynamics and, therefore, should be avoided in satellite Cal/Val.

c) Platform track check

Platform track check verifies that consecutive locations of a platform (ID) are consistent with the respective time stamps so that the platform does not move faster than a predefined maximum moving speed. Significant errors in time and latitude/longitude will cause deviations from this expected pattern.

At first, a least-required speed is calculated assuming that the platform had traveled between the locations of any two reports through a direct linked path. Then the report with the maximum times of speed violations will be identified and excluded. The operation is iterated until convergence achieved. The maximum speed is chosen as 60 km/h for ships and 15 km/h for drifters.

For moored platforms, the procedure is simplified. Reports located far away from the majority of reports, is regarded erroneous. The maximum allowed distance is chosen as 100km to tolerate reasonable drifting and latitude/longitude error. Note that group ID (several platforms which share the same ID, e.g. call sign “SHIP” representing all anonymous ships) and single-reporter (ID with less than 3 reports in a month) are not subject to track check.

Fig. 1 gives several examples of abnormal reports identified by this check. In Fig. 1(a), one observation apparently falls out of track, due to an error caused by a swapped sign of latitude field. In this case, it would be difficult to detect such error merely by comparing to reference SST, as reference SSTs could be close for the similar latitude zones in the north and south. Another example, for drifting buoy, is given in Fig. 1(b). Such error is even more difficult to detect by comparing reference SST. Finally, Fig. 1(c) shows an example of a mooring buoy which has two observations located far off from the majority.

d) SST spike check

For a continuously reporting platform, SST spike (or step) is very likely to present along its track or time series if erroneous reports occurred, due to sensor failure or occasional maintaining operation. Spike check has the same logic as track check except that it checks the maximum SST gradient in space and time instead of travel speed, using the same algorithm. The maximum SST gradient is chosen as 0.5K/km in space and 1.0K/hour in time.

e) Reference check

Reference check (a.k.a. background check) is the major check of many QC methods which identifies most of the outliers with extreme SST values. The Bayesian-based approach by *Lorenc and Hammon, 1988* was adopted in our QC algorithm. Compared to conventional statistic-based outlier detection methods, it employs the Bayesian probability theory to better take into account factors such as the accuracy of the reference field itself, matching error between *in situ* and reference field, and the instrumental noise of *in situ* data.

In this study, Reynolds optimal interpolation (OI) global 0.25° daily analysis SST (“AVHRR only”) was selected as reference [*Reynolds et al., 2007*]. Recall that Reynolds SST is a blended product of AVHRR satellite retrievals and QCed ICOADS *in situ* SSTs and is available for a period from 1982 onward. Gridded 0.25° resolution data were bilinearly interpolated in space to each *in situ* observation. Note that Reynolds SST of previous day is used in QC of current day *in situ* data, in an attempt to preserve the independence of reference on *in situ* data. The biases and SDs of reference are estimated empirically based on local statistics, within a 1°×1°×3day running window [*Xu and Ignatov, 2010*].

f) Cross-platform check

Cross-platform check is a critical complement to reference check, as it verifies across independent platforms and may counter-balance the deficiencies resulting from possible inaccuracies in the reference field. The Bayesian cross-platform check (a.k.a. 'buddy check') is performed based on the reference check, which adjusts the *posteriori* probability of gross error by further incorporating information from nearby measurements (a.k.a. 'buddies') [*Lorenc and Hammon, 1988; Ingleby and Huddleston, 2007*].

The 'buddy check' algorithm by *Lorenc and Hammon's* (1988) was employed, with minor modifications described by *Xu and Ignatov* [2010]. In particular, the complexity of searching 'buddies' was reduced, by setting the upper limit to 300km in space and 4 day in time, to exclude those 'buddies' from cross-platform check which are too far away. In addition, space-time gridding technique was employed to accelerate the 'buddy search' process [*Xu and Ignatov, 2010*].

Note that both reference check and cross-platform check produce a continuous quality indicator, which serves as the probability of gross error. Different thresholds can be applied to this probability for different requirements of data quality. In this study, a threshold of $P > 0.5$ is used to determine whether the measurement is corrupted.

2.3 Evaluation

As a demonstration of the efficacy of the proposed QC algorithm, the NCEP-GTS data (from 1991 to 2009) are quality controlled by the proposed algorithm. In Fig. 2, it shows the SD of SST anomalies "in situ minus reference" before and after QC. The reference used is Reynolds optimal interpolation global 0.25° daily analysis SST ("AVHRR only") available from 1982 onward. Gridded 0.25° resolution Reynolds SST of the same day was bilinearly interpolated in space to each *in situ* observation.

It is clearly seen that non-QCed reports show abnormally high SDs emphasizing the critical importance of QC. The effect is particularly large for drifters and tropical moorings, which are the major source of *in situ* data used in satellite Cal/Val. Note that the average level of removed reports by QC is 10-12% for ships, 5-10% for drifters, 1-3% for tropical moorings and 5-10% for coastal moorings.

More theoretical and technique details about implementation of QC algorithm will be given in future publication [*Xu and Ignatov, 2010*], as well as objective evaluation of the performance of each QC checks.

3 iQuam – the online monitoring system

In addition to operational QC system, long-term and prompt monitoring of both global and single-platform statistics is very helpful to users of *in situ* data as well as the maintainer of QC system. Besides, such prompt statistics can also be used as feedback to QC system for algorithm self-adjusting purpose.

There are currently many useful online monitoring tools existed. For example, the Data Buoy Cooperation Panel maintains a quality control relay networks and provides real-time statistics (*http://www.jcommops.org/dbcp/data/qc.html*); the UK Met Office runs an operational quality control and monitoring system for in situ and satellite measurements (*http://www.metoffice.gov.uk/research/nwp/observations/*); another QC tool is developed by Meteo France (*http://www.meteo.shom.fr/qctools*); the NCEP/NCO Quality Assessment Project (*http://www.nco.ncep.noaa.gov/pmb/qap/*); and more other related websites can be found through these websites. However, most of them were initially designed for meteorological application, not specifically for satellite Cal/Val.

At NESDIS, the QC algorithm described in previous section was implemented and running in NESDIS/STAR with input *in situ* data from NCEP Global Telecommunication System (GTS). QCed data in self-described HDF format are available online [*c.f.* www.star.nesdis.noaa.gov/sod/sst/iquam/data.html]. Currently, all processing of the previous month data are done on the 5th day of the current month. Near real time (daily or hourly) updates to the current month data file will be provided in later versions.

In addition, a monitoring system is running in parallel to calculate timely statistics of QC results and SST anomalies. Reports are presented on web to facilitated human monitoring. This system, called 'in situ SST QUAlity Monitor' (iQuam), is organized as shown in Fig. 3. Raw *in situ* data are preprocessed and re-organized in to monthly files, which are further input to rotation buffer of QC system. QC is performed with ancillary data, such as water mask and reference SSTs, and feedback information from monitoring system. QC results as well as *in situ* data are then input to monitor where statistics are calculated and reports are generated.

Such parallel monitoring and feedback mechanism is critical for a robust and accurate QC system. It provides statistics supporting a more accurate QC configuration. For example, SST variance of *in situ* vs. reference can be used to estimate the actual error level in the reference. The monitoring system also enables maintainers to interfere and adjust QC system in real time.

The iQuam web interface is shown in Fig. 4 as well as a monthly global map of four types of *in situ* measurements, i.e. ships, drifters, tropical moorings and coastal moorings. Outliers detected by QC are shown in grey color. Some other examples of iQuam monitoring reports are shown in Figs. 5-7.

In situ data with quality flags (QF) appended are aggregated in to monthly files in self-described HDF format. These data are served in the iQuam website (see Fig. 8) and can also be downloaded via FTP (ftp://www.star.nesdis.noaa.gov/pub/sod/sst/iquam/). Evolving data of the latest month will be available in later versions. More information can be found on iQuam website www.star.nesdis.noaa.gov/sod/sst/iquam.

4 Conclusion

A QC approach is implemented in NESDIS which includes operations of five categories, prescreening – duplicate removal, plausibility check – geolocation check, internal consistency check – tracking and spike check, mutual consistency check – cross-platform check and external consistency check – reference check. Two major checks are the Bayesian-based reference check and cross-platform check. The algorithm is implemented straightforwardly and the complexity is within reasonable range, e.g. ~0.5 hour per year-data for early years and ~6 hour per year-data for years after 2005 on mainstream PC platform.

The online monitoring system, iQuam, is developed and running in parallel to QC. It calculates statistics of QC results and SST anomalies and sends feedback information to QC system. Global maps, monthly statistics, time-series and individual platform statistics are generated and presented online.

Acknowledgements

This work is conducted under the Algorithm Working Group funded by the GOES-R Program Office, NPOESS Ocean Cal/Val funded by the Integrated Program Office, and Polar Product System and Implementation, NPOESS Data Exploitation, and Ocean Remote Sensing Projects funded by NOAA/NESDIS. F. Xu also acknowledges the CIRA visiting scientist fellowship. Thanks to Scott D. Woodruff of NOAA/ESRL, Rick Lumpkin of NOAA/OML, Steven Worley of NSF/NCAR/SCD, Diane Stokes of NOAA/NCEP, Nancy L. Baker of NRL, Richard Crout of NOAA/NDBC, and our colleagues John Sapper, Prasanjit Dash and Yury Kihai for helpful discussions. The views, opinions and findings contained in this report are those of the authors and should not be construed as an official NOAA or U.S. Government position, policy, or decision.

References

Brisson, A., P. Le Borgne, and A. Marsouin, 2002: Results of One Year of Preoperational Production of Sea Surface Temperatures from GOES-8. *J. Atmos. Oceanic Technol.*, **19**, 1638–1652.

Francois, C., A. Brisson, P. Le Borgne, A. Marsouin, 2002: Definition of a radiosounding database for sea surface brightness temperature simulations - Application to sea surface temperature retrieval algorithm determination, *Remote Sensing of Environment*, 81, 2-3, 309-326.

Hansen, M., R. DeFries, J.R.G. Townshend, and R. Sohlberg, 2000: Global land cover classification at 1km resolution using a decision tree classifier, *International Journal of Remote Sensing*. 21: 1331-1365.

Ingleby, B. and M. Huddleston, 2007: Quality control of ocean temperature and salinity profiles – historical and real-time data, *J Marine Systems*, 65, 158–175.

Kilpatrick, K.A., G.P. Podesta and R. Evans, 2001: Overview of the NOAA/NASA advanced very high resolution radiometer Pathfinder algorithm for sea surface temperature and associated matchup database, *J Geophy. Res.*, 106, C5, 9179–9197.

Lorenc, A.C. and O. Hammon, 1988: Objective quality control of observations using Bayesian methods. Theory, and a practical implementation. *Q. J. R. Meteorol. Soc.* 114, 515–543.

Martin, M.J., Bell, M.J., Hines, A., 2002. Estimation of threedimensional error covariance statistics for an ocean assimilation system. Ocean Applications Tech. Note, vol. 30. Met Office, FitzRoy Road, Exeter, EX1 3PB, UK.

Merchant C.J., P. Le Borgne, A. Marsouin, H. Roquet, 2008: Optimal estimation of sea surface temperature from split-window observations, *Remote Sensing of Environment*, 112(5), 2469-2484.

O'Carroll, A.G., J.R. Eyre and R.W. Saunders, 2008: Three-way error analysis between AATSR, AMSR-E, and *in situ* sea surface temperature observations, *J. Atmos. Oceanic Technol.*, 25(7), 1197–1207.

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

Slutz, R.J., S.J. Lubker, J.D. Hiscox, S.D. Woodruff, R.L. Jenne, D.H. Joseph, P.M. Steurer and J.D. Elms, 1985: *Comprehensive Ocean-Atmosphere Data Set; Release 1*. NOAA Environmental Research Laboratories, Climate Research Program, Boulder, CO, 268 pp. (NTIS PB86-105723).

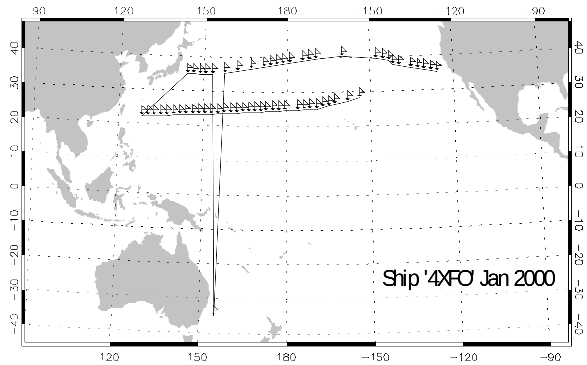
Woodruff, S.D., H.F. Diaz, J.D. Elms and S.J. Worley, 1998: COADS release 2 data and metadata enhancements for improvements of marine surface flux fields. *Phys. Chem. Earth*, 23, 517–527.

Woodruff, 2008: Marine QC scoping document, *JCOMM Data Management Coordination Group 3rd Session*, Ostend, Belgium, 26-28 March 2008.

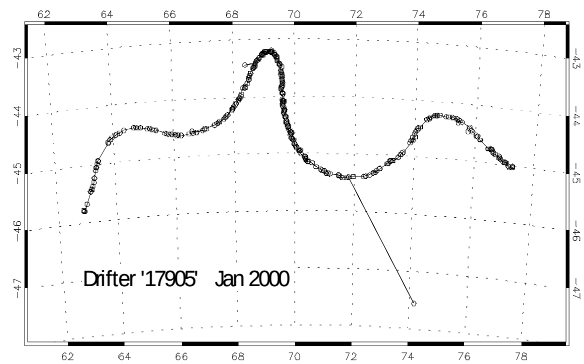
Worley S.J., S.D. Woodruff, R.W. Reynolds, S.J. Lubker and N. Lott., 2005: ICOADS release 2.1 data and products, *Int J Clim.*, 25(7), 823–842.

Xu F. and A. Ignatov, 2009: Evaluation of in situ SSTs for use in the calibration and validation of satellite retrievals, *JGR-Ocean*, submitted.

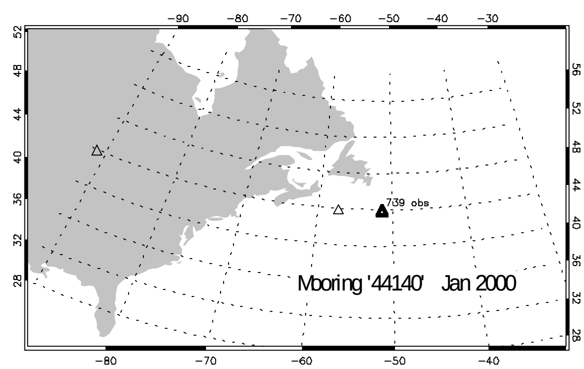
Xu F. and A. Ignatov, 2010: Quality control of in situ sea surface temperature measurements: implementation and evaluation, in preparation.



(a) latitude sign swapped

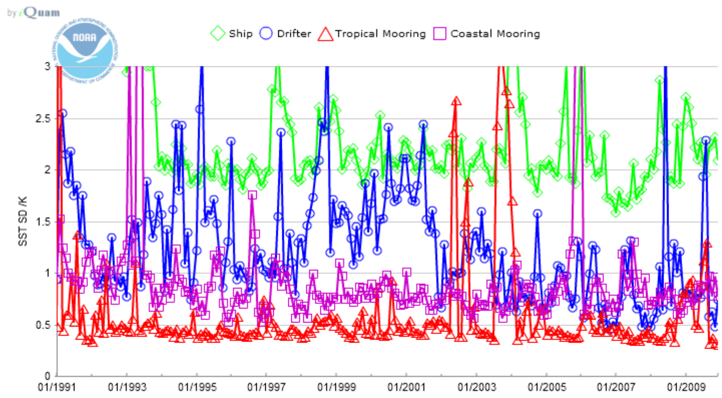


(b) latitude and longitude shifted by 1-2°

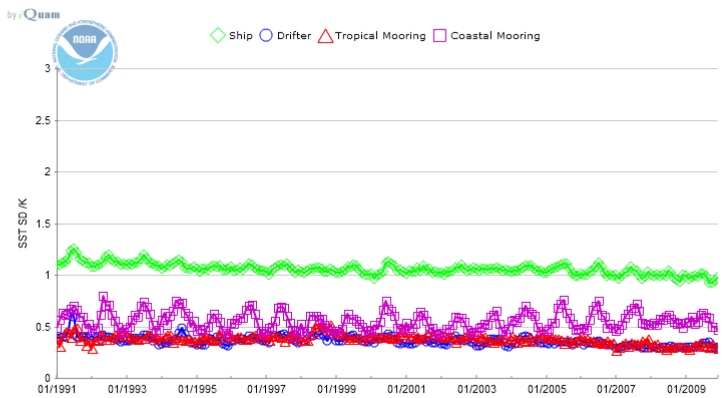


(c) located off from the moored position

Figure 1. Cases of erroneous records of (a) ship, (b) drifter and (c) mooring detected by tracking.



(a) Before QC



(b) After QC

Figure 2. Time series of SST SD (in situ - Reynolds): (a) 'No QC' vs. (b) 'QCed'.

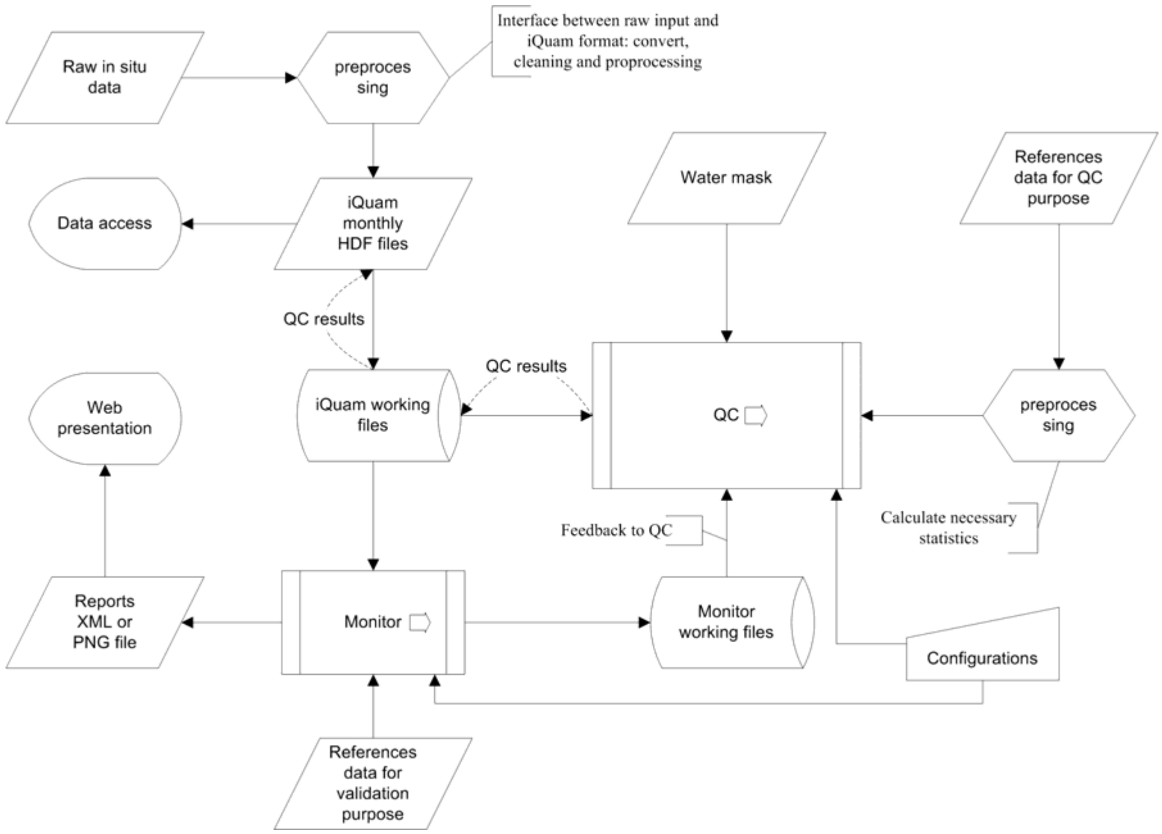


Figure 3. iQuam – *in situ* SST QC and monitoring system.

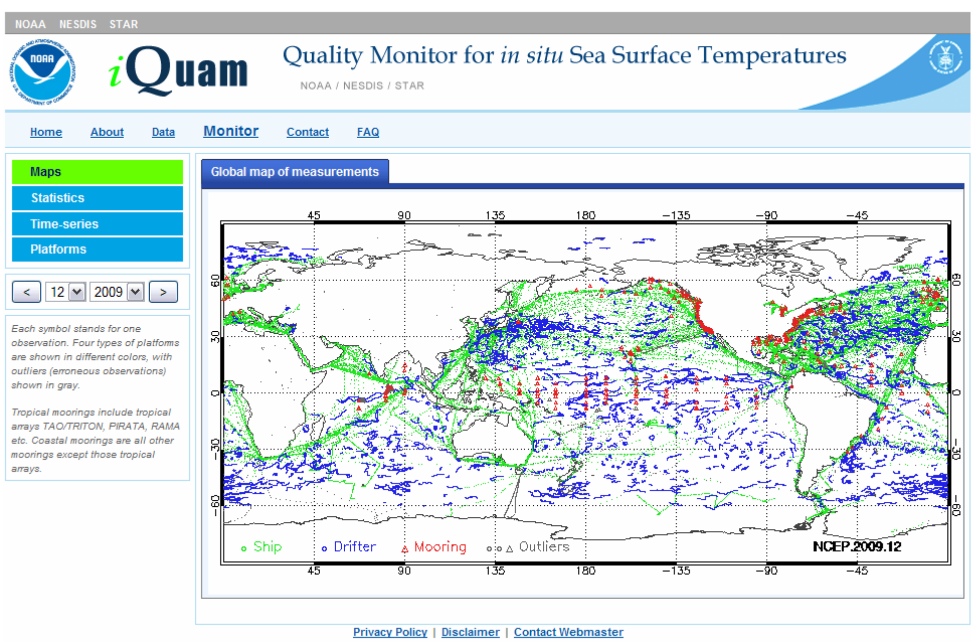


Figure 4. iQuam web interface and global map of in situ measurments.

QC Statistics

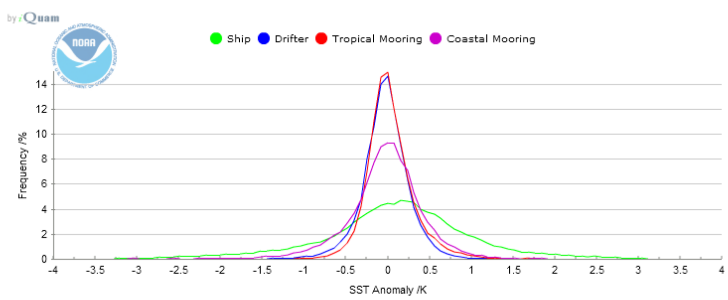
Platform	N_Obs	N_QC	DR	TC	SC	RC	XC
Ship	78,661	66,072	164	7,409	150	11,626	12,087
Drifter	1,048,270	939,558	84,745	3,197	870	15,257	23,878
Tropical Mooring	33,566	32,994	212	140	17	314	351
Coastal Mooring	187,319	174,938	6	1,354	426	6,313	12,180

(a)

SST Statistics

Platform	BIAS	SD	SKEW	KURT	MED	RSD	N_Mtchp
Ship	0.14	0.94	-0.32	1.77	0.17	0.73	60,687
Drifter	0.02	0.29	-0.29	4.43	0.02	0.23	937,136
Tropical Mooring	0.07	0.29	0.79	3.39	0.04	0.22	32,947
Coastal Mooring	0	0.5	-1.24	7.32	0.04	0.35	148,818

(b)



(c)

Figure 5. Monthly statistics stratified by platform types: (a) statistics of QC results (N_obs: total number of observations; N_QC: total number of observations passed QC; DR/TC/SC/RC/XC: number of measurements detected by each QC checks); (b) statistics of SST anomalies (Bias, SD, Skewness, Kurtosis, Median, Robust SD, Number of matchups of in situ vs. reference); (c) histogram of SST anomalies.

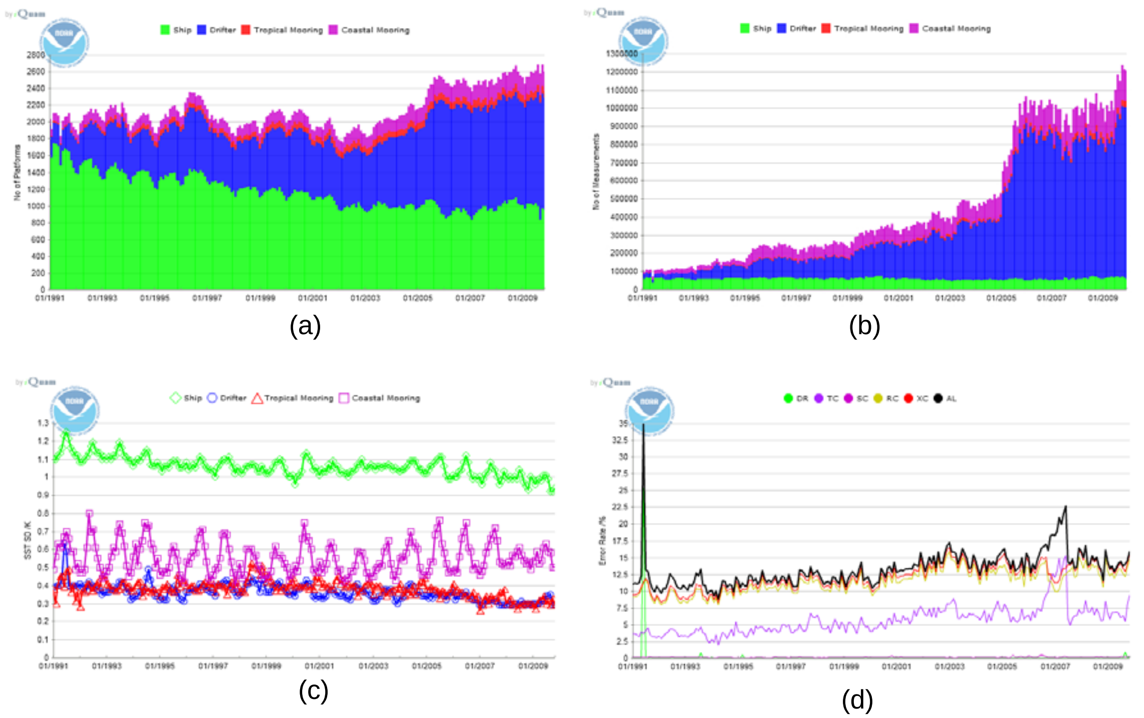
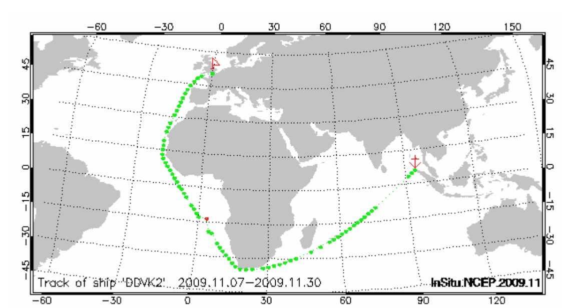


Figure 6. Time-series of monthly statistics stratified by platform types: (a) number of platforms; (b) number of observations; (c) mean biases of SST anomalies after QC; (d) error rates (percentage of detected erroneous measurements) of each QC checks for ships.

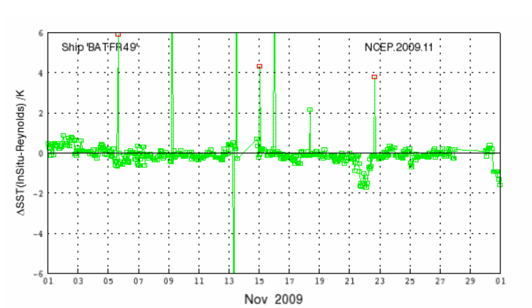
Ship | Drifter | Tropical Mooring | Coastal Mooring

ID	NOBS	N_QC	Rate	XC	RC	DR	TC	SC	MEAN	STDV
MASKST	4,951	4,130	16.6	728	684	164	415	0	0.17	0.99
LDWR	626	626	0	0	0	0	0	0	0.21	0.2
LGWS	592	592	0	0	0	0	0	0	-0.64	0.4
LF4H	609	591	3	16	16	0	0	15	0.28	0.27
DBBT	648	579	10.6	69	69	0	69	0	0.43	0.25
DFPC	578	555	4	23	17	0	0	0	0.8	0.69
DBBU	573	555	3.1	18	18	0	18	0	0.13	0.39
BATFR54	567	539	4.9	28	28	0	28	0	-0.15	0.35
VROB	639	519	18.8	120	118	0	118	0	0.04	0.44
BATFR19	588	499	15.1	88	88	0	86	3	0.28	0.52
BATFR49	483	476	1.4	7	7	0	0	6	-0.04	0.37
BATFR04	475	471	0.8	4	4	0	0	4	0.33	0.35

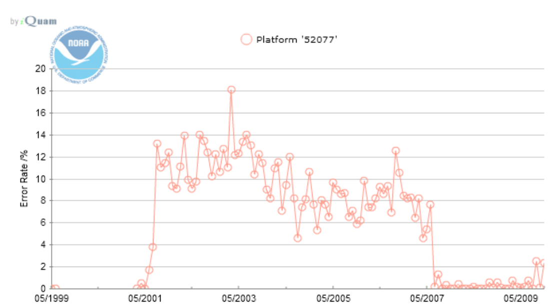
(a)



(b)



(c)



(d)

Figure 7. Individual platform statistics: (a) list of platforms and their statistics of QC results and SST anomalies; (b) monthly track map of individual platform; (b) monthly time-series of individual platform SST anomalies; (c) error rate history of individual platform.

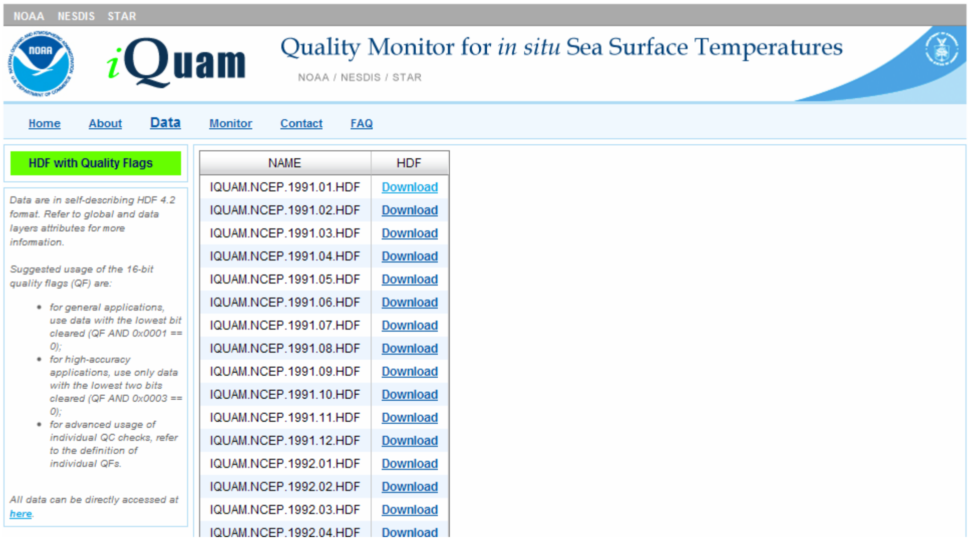


Figure 8. iQuam web page for downloading in situ data with quality information.