NPL

fiducial reference temperature measurements

esa

Fiducial Reference Measurements for validation of Surface Temperature from Satellites (FRM4STS)

ESA Contract No. 4000113848_15I-LG

OP-90: Scientific and Technical Meeting Report: *Investigation on uncertainty budget for drifting buoy ice surface temperature observations, when compared against satellite observations*

JUNE 2018

Reference	OFE-OP-90-V1-Iss-1-Ver-2
Issue	1
Revision	2
Date of Issue	25 June 2018
Document Type	STM

Approval/Acceptance			
ESA Craig Donlon Technical Officer	Digitally signed by Device Identity 9c6f8ccd4f5eaa531db70b 0929b3d2a06f1e6010 Location: Noordwijk, The Netherlands Date: 2018.06.27 13:15:38 +02'00' *Signature*	**NPL** Andrew Brown Project Manager	Andrew Brown, NPL *Signature*

INTRODUCTION

This report contains the presentation on: "the uncertainty budget for non-recoverable ice surface temperature (IST) observations, when used to validate and calibrate satellite observations" that was given at International FRM4STS workshop held at NPL in October 2017. In addition, a summary of the key results and discussions from the workshop is included. In the end, the recommendations from the discussions are given for improving the FRM satellite calibration and validation.

The work with establishing fiducial reference measurements for ice surface temperature and agree upon protocols within the community is much less mature than for SST and LST. This means that today there is no consensus on a FRM IST data set that can be used for routine validation of the satellite IST products. Candidates for a routine IST FRM are the buoy data from the Data Buoy Cooperation Panel (DBCP) and the International Arctic Buoy Programme (IABP) that deploy iSVP buoys on the sea ice that typically report hourly temperature observations from the sea ice (see http://iabp.apl.washington.edu/index.html). However, very few studies have been performed to assess the uncertainty of these non-recoverable observations (e.g Rigor et al., 2000). Dybkjær et al., 2012 found that manual inspection of each buoy was needed in order to increase the quality of the ice drifting observations obtained from the Global Telecommunication System (GTS). In addition, they ended up discarding a vast majority of the observations. The conclusions in these papers pointed towards the need for an in-depth analysis of the uncertainties on the ice drifter observations.
The nature of the ice surface temperature observations requires that the establishment of an IST FRM to be used for satellite climate data record validation must include both assessment of the sensor performance and the representativeness effects due to spatial and temporal variability. In particular, the vertical transformation of the ice drifting observations to the skin IST is very different if the sensor is 20 cm above the sea ice or covered with 5 cm of snow, this due to the large vertical gradients within the snow. To obtain a reliable IST FRM data set to be used for climate data record and routine validation, the magnitude of all these effects need to be assessed. In this presentation we quantify the uncertainty components related to the temperature observations from the ice surface drifters. The work is connected to the work within the FRM4STS proposal: "Study of SI Traceability for non-recoverable SST and IST FRM instruments" in terms of sensor degradation and performance of the SVP buoys.

SUMMARY OF KEY FINDINGS

The results demonstrate that using iSVP as a FRM for validating satellite ice surface temperature measurements over sea ice bears several complications. The accuracy of the sensors is adequate to be used for validation. Through data analysis on iSVP buoys set out on the sea ice off Qaanaaq, western Greenland, each of the components in the uncertainty budget is assessed, when satellite and iSVP buoys observations are validated. A table has been derived for typical validation conditions, showing sampling differences to range from 0.36 °C for radiometric observations to more than 5 °C, when comparing satellite and iSVP buoys analysis. The largest effects arise from the different sampling components, where the vertical displacement of the in situ observations from the skin into the snow and ice or in the air accounts for the largest contribution.
In addition to the Qaanaaq data analysis, automatic quality control routines have been developed to filter out the observations that do not represent the skin IST. These procedures consist of 15 tests that range from to buddy checks on mean value and variability. Application of these QC tests on a test data set obtained from the GTS leads to improvements in the data

quality, but the fundamental problem still exists, namely that the IST skin temperature easily can differ several degrees from temperature observations within a few centimetres of snow in the air. These large discrepancies severely limit the usefulness of the observations for validating and improving upon the satellite IST retrieval algorithms. Several land-based snow and ice radiometric observations are available today from e.g. the ARMS sites but the snow and ice melts during summer and no all-year round radiometers exist today that can be used to monitor and validate the existing satellite IST algorithms.

So in summary, it was concluded from the work that:

- This is the first time a systematic study has been performed to assess the different components in the satellite vs in situ inter-comparison budget.
- The 4 iSVP bouys deployed within few metres showed:
    - Up to 14 degrees C differences among themselves
    - More than 20 degrees away from Tskin
- The sampling effects using iSVP observations can be much larger than potential algorithm uncertainties
    - Sampling effects can be up to 5 degrees
    - The TIR FRM observations of skin IST is significantly better than other types of IST observations for satellite IST monitoring and validation
- Automatic QC procedures can improve the in situ observations, but cannot completely remove the sampling effects.

WAY FORWARD

At the workshop, it was discussed that the lack of FRM radiometric observations over sea ice is limiting the routine monitoring of operational products and the development within the IST algorithms. The present day validation of these retrieval algorithms is usually performed using temperature observations from the sea ice or from fixed AWS stations reporting T2m. The discussion ended with a clear recommendation to the stake holders:

It is recommended that all-year FRM radiometric observations of IST are being performed over a homogeneous permanent ice area, such as Summit, Greenland or Dome C on Antarctica for the routine calibration and validation of satellite IST observations.

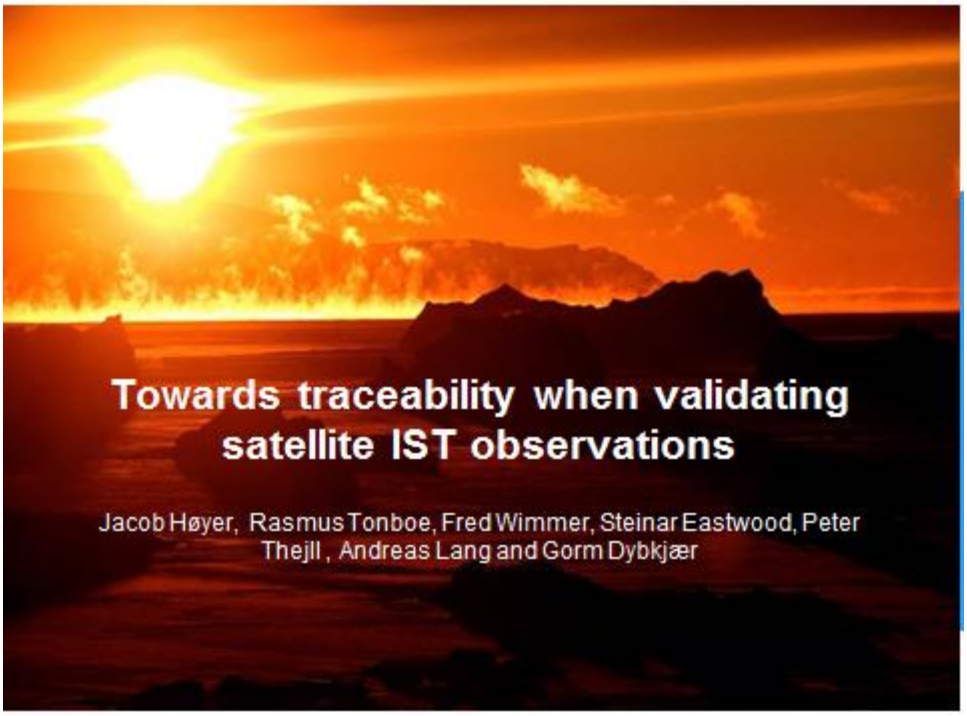
PRESENTATION FROM THE FRM4STS INTERNATIONAL WORKSHOP

The slides from the presentation '*Towards traceability when validating satellite ice surface temperature observations'* are reproduced on the following pages.

REFERENCES

Dybkjær, G., Tonboe, R., & Høyer, J. L. (2012). Arctic surface temperatures from Metop AVHRR compared to in situ ocean and land data. Ocean Science, 8(6), 959-970.

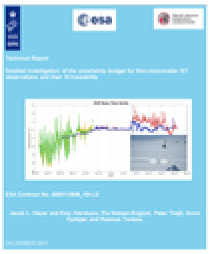
Rigor, I. G., Colony, R. L., & Martin, S. (2000). Variations in surface air temperature observations in the Arctic, 1979–97. Journal of Climate, 13(5), 896-914.


Towards traceability when validating satellite IST observations
Jacob Høyer, Rasmus Tonboe, Fred Wimmer, Steinar Eastwood, Peter Thejll, Andreas Lang and Gorm Dybkjær

fiducial reference temperature measurements

Motivation

NPL National Physical Laboratory



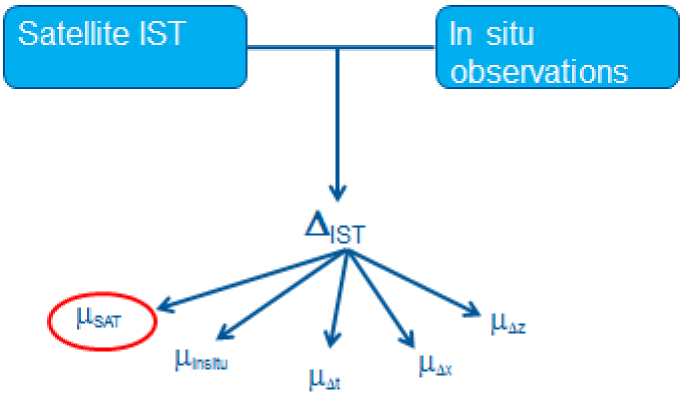
- Little use of Satellite IST despite large potential
  - Limited IST validation studies
    - Lack of reference data
    - Lack of understanding the uncertainty budgets
- Need for a reference data set for validation and calibration of satellite IST
- Here: Investigate the implications of using:
  - iSVP buoys
  - Radiometers observations
  - T2m air observations
- Developed automatic quality control procedures
- Purchased two iSVP buoys to put out in Qaanaq Western Greenland

fiducial reference temperature measurements

DMI Danish Meteorological Institute

Comparing satellite and in situ observations

NPL National Physical Laboratory



fiducial reference temperature measurements

DMI Danish Meteorological Institute

fiducial reference temperature measurements

iSVP buoys

NPL National Physical Laboratory

- Two iSVP buoys deployed in January 2017, at AWS site
- Wrong software, reporting -5°C !
- Two new deployed in April, 2017
- New buoys recovered in June
- Old buoys left on ice -> ocean

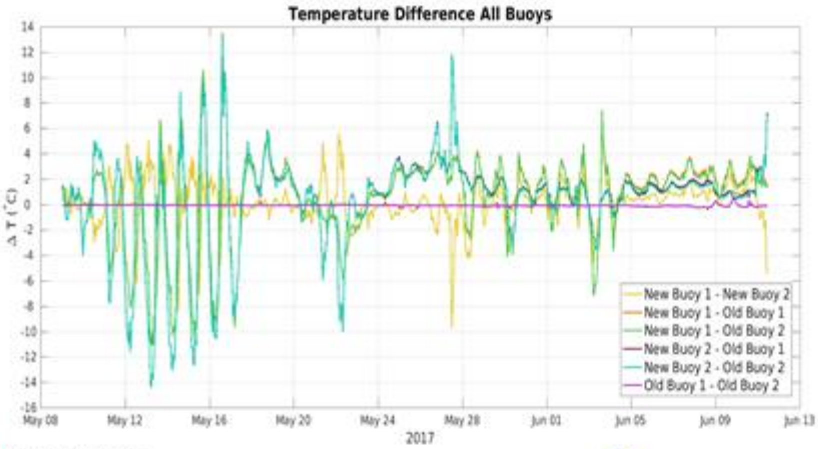
fiducial reference temperature measurements

DMI Danish Meteorological Institute

Relation between Buoy observations

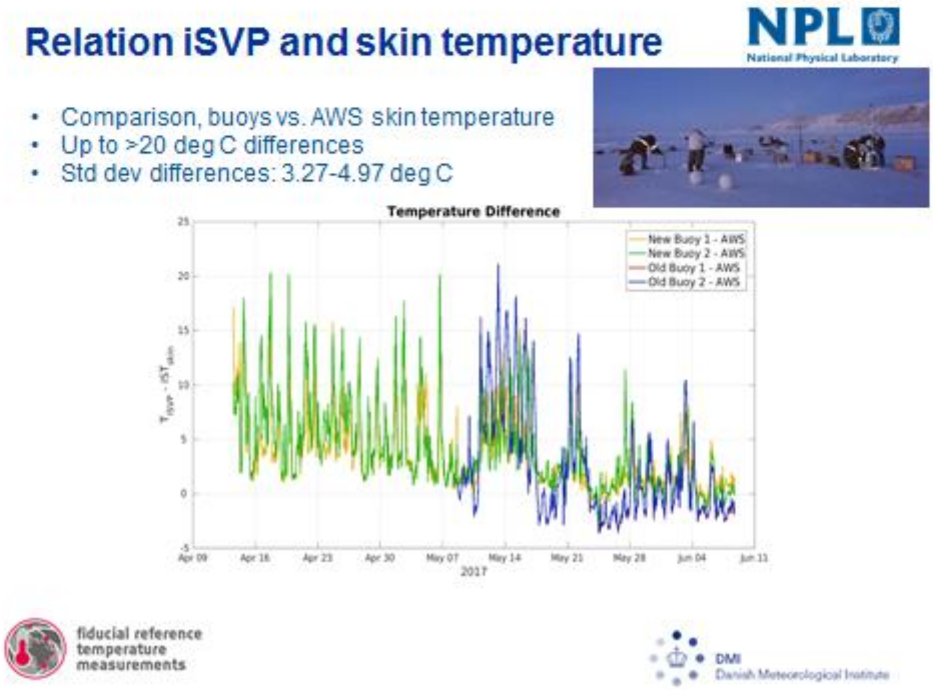
NPL National Physical Laboratory

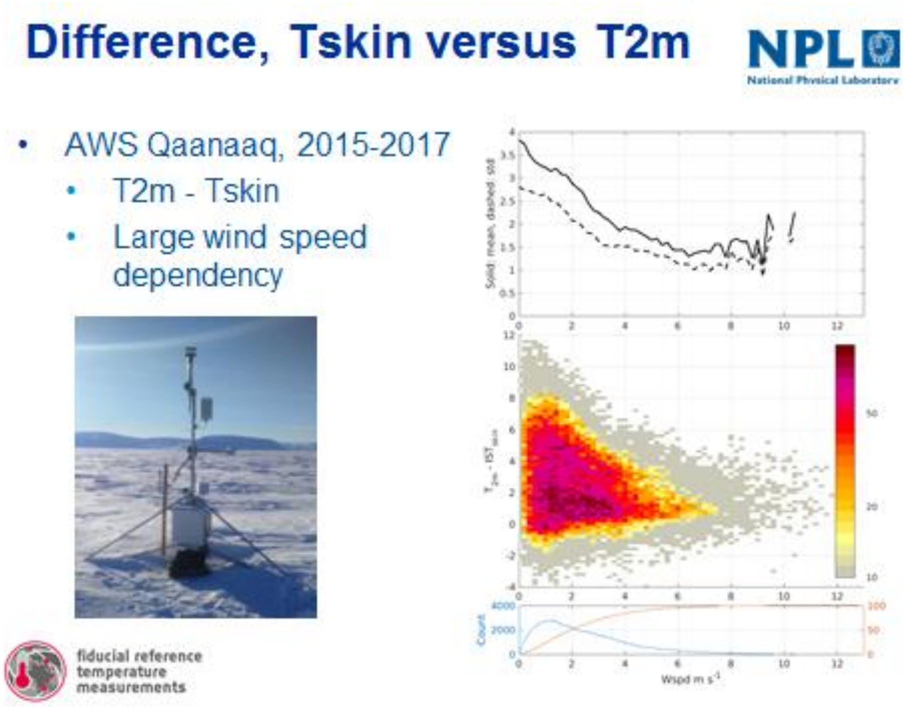
- Pairwise differences can be large, > 14 degrees C
- Std of differences: 0.07-3.77



fiducial reference temperature measurements

DMI Danish Meteorological Institute

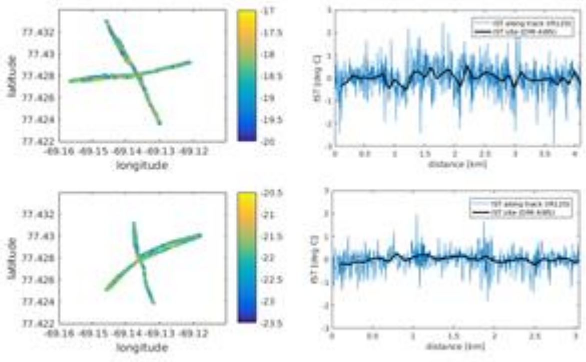

Relation iSVP and skin temperature
NPL
National Physical Laboratory
Comparison, buoys vs. AWS skin temperature
Up to >20 deg C differences
Std dev differences: 3.27-4.97 deg C
Temperature Difference
New Buoy 1 - AWS
New Buoy 2 - AWS
Old Buoy 1 - AWS
Old Buoy 2 - AWS
Apr 09
Apr 16
Apr 23
Apr 30
May 07
May 14
May 21
May 28
Jun 04
Jun 11
2017
fiducial reference temperature measurements
DMI
Danish Meteorological Institute


Difference, Tskin versus T2m
NPL
National Physical Laboratory
AWS Qaanaaq, 2015-2017
T2m - Tskin
Large wind speed dependency
Solid: mean, dashed: std
Count
Wspd m s-1
fiducial reference temperature measurements

fiducial reference temperature measurements

Spatial Variability

NPL National Physical Laboratory



	N(obs)	Distance	Duration	Stdv (σ)	Bias to AWS	Spatial stdv
Part 1 (Apr-02)	718	4.08 km	00:59:45	0.69 °C	-0.01 °C	0.25 °C
Part 2 (Apr-03)	709	3.04 km	00:59:00	0.42 °C	0.50 °C	0.12 °C

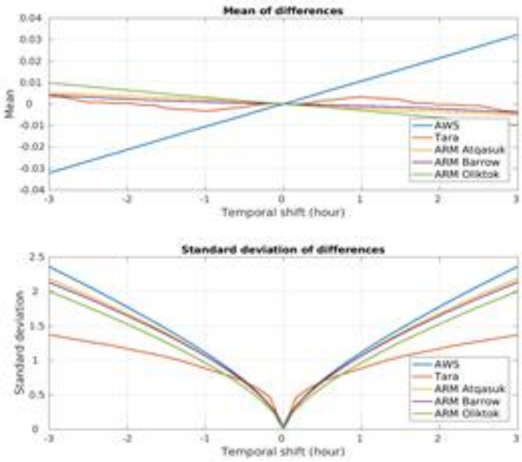
fiducial reference temperature measurements

DMI Danish Meteorological Institute 8

Temporal difference

NPL National Physical Laboratory

- Large hourly variability, compared to SST
- Three years of AWS data, + Tara + ARMS data



fiducial reference temperature measurements

DMI Danish Meteorological Institute

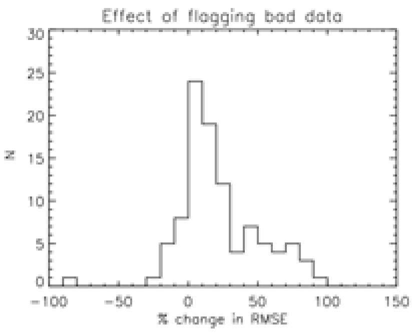
Matchup contributions

- Variability from contributions not related to satellite retrieval errors
- Differences from 0.4-5.1°C
- iSVP buoys can contain large errors

Δx (km)	Δt (min)	Δz (m)	$\mu_{in\ situ}$ (°C)	$\mu_{\Delta x}$ (°C)	$\mu_{\Delta t}$ (°C)	$\mu_{\Delta z}$ (°C)	$\sqrt{\mu^2_{in\ situ} + \mu^2_{\Delta x} + \mu^2_{\Delta t} + \mu^2_{\Delta z}}$ (°C)
1.0	10	IST_{skin}	0.2	0.12-0.25	0.34	0	0.41-0.47
1.0	30	IST_{skin}	0.2	0.12-0.25	0.71	0	0.75-0.78
1.0	60	IST_{skin}	0.2	0.12-0.25	1.11	0	1.13-1.16
1.0	10	T_{2m}	0.05	0.12-0.25	0.34	1.45 - 2.38	1.49-2.42
1.0	30	T_{2m}	0.05	0.12-0.25	0.71	1.45 - 2.38	1.62-2.50
1.0	60	T_{2m}	0.05	0.12-0.25	1.11	1.45 - 2.38	1.83-2.64
1.0	10	T_{buoy}	0.05	0.12-0.25	0.34	3.27 - 4.95	3.29-4.97
1.0	30	T_{buoy}	0.05	0.12-0.25	0.71	3.27 - 4.95	3.35-5.01
1.0	60	T_{buoy}	0.05	0.12-0.25	1.11	3.27 - 4.95	3.46-5.08

QC control

- Automatic procedures have been developed to QC iSVP obs
- Including 16 checks
- Improvement in 30-90 % in test data set when compared to buddies



Flag #	Name	Descripti…
1	Gross Error	The temperature is outside of the interval (−80, 20)
2	Spike Test Short	The absolute temperature difference from the median temperature of a 1 day rolling window is greater than 10 degrees
3	Spike Test Long	The absolute temperature difference from the median temperature of a 3 day rolling window is greater than 20 degrees
4	Buddy Check	The absolute difference from the median of a '500 km x 500 km x 1 day' bin, to which the temperature value belongs, is greater than 20 degrees
5	Neighbouring bins check	The rolling variance (using a 1 day time window) is greater than twice the mean variance of measurements from neighbouring stations (i.e those in the same '500 km x 500 km x 1 day' bin).
6	Age Check	The data-point is greater than 1 year from start date of file
7	Sea Ice Concentration test	The sea ice concentration is less than 30%
8	Temperature variability check	The series standard deviation in a 1 day window is less than 0.1 C
9	Speed test	The speed is greater than 0.5 m/s
10	Position Sanity	The absolute latitude is greater than 50°, or the longitude is 0 while the latitude is 90°
11	Duplicates	There is another value with the same timestamp
12	Buddy checks could not be applied	Tests 4 and 5 inapplicable due to no neighbours.
13	Not used	Can be used for a new test, in the future.
14	Gappiness	The interval between successive points is greater than 2.5 times the median interval
15	Close to land	The location of the measurement is less than 15 km from land
16	Very close to land	The location of the measurement is less than 5 km from land

Notes: Each number refers to a test above. 2: 'rolling window' (here and elsewhere) refers to 1-day long windows where each window is offset from next window by 1 day. 4: The box contains information on data inside the box - not near neighbours in adjacent boxes. 9: The speed is calculated as distance travelled divided by time spent regardless of [illegible] of time interval. 11: All duplicate values are flagged.

CONCLUSIONS

NPL National Physical Laboratory

- First time a systematic study has been performed to assess the different components in the satellite vs in situ inter-comparison budget.

- 4 iSVP bouys deployed within few metres
  - Up to 14 degrees C differences among themselves
  - More than 20 degrees away from Tskin
- Sampling effects much larger than potential algorithm uncertainties
  - Sampling effects can be up to 5 degrees
  - Radiometric observations significantly better than other types of IST observations

- Recommendation:
  - *To establish an all-year reference radiometer site for validation and calibration of satellite IST observations*

Data for this phase of the project is available here: IST Observations