2016-04-27

Sampling Errors in Satellite-Derived Infrared Sea Surface Temperatures

Yang Liu
University of Miami, yliu@rsmas.miami.edu

Follow this and additional works at: https://scholarlyrepository.miami.edu/oa_dissertations

Recommended Citation
https://scholarlyrepository.miami.edu/oa_dissertations/1659

This Open access is brought to you for free and open access by the Electronic Theses and Dissertations at Scholarly Repository. It has been accepted for inclusion in Open Access Dissertations by an authorized administrator of Scholarly Repository. For more information, please contact repository.library@miami.edu.
UNIVERSITY OF MIAMI

SAMPLING ERRORS IN SATELLITE-DERIVED INFRARED SEA-SURFACE TEMPERATURES

By

Yang Liu

A DISSERTATION

Submitted to the Faculty
of the University of Miami
in partial fulfillment of the requirements for
the degree of Doctor of Philosophy

Coral Gables, Florida

May 2016
A dissertation submitted in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy

SAMPLING ERRORS IN SATELLITE-DERIVED INFRARED SEA-SURFACE
TEMPERATURES

Yang Liu

Approved:

Peter J. Minnett, Ph.D.
Professor of Ocean Sciences

Benjamin Kirtman, Ph.D.
Professor of Atmospheric Sciences

Paquita Zuidema, Ph.D.
Professor of Atmospheric Sciences

Mohamed Iskandarani, Ph.D.
Associate Professor of Ocean Sciences

Chelle Gentemann, Ph.D.
Senior Scientist
Earth and Space Research
Seattle, Washington

Guillermo Prado, Ph.D.
Dean of the Graduate School
LIU, YANG                  (Ph.D., Meteorology and Physical Oceanography)

Sampling Errors In Satellite-Derived Infrared Sea-Surface Temperatures

(May 2016)

Abstract of a dissertation at the University of Miami.

Dissertation supervised by Professor Peter J. Minnett.

No. of pages in text. (86)

Long time series of accurate Sea Surface Temperatures (SSTs) are needed to resolve subtle signals that may be indicative of a changing climate. Motivated by the stringent requirements on SST accuracy required for Climate Data Records (CDR) we quantify sampling errors in satellite SSTs. Infrared (IR) sensors, including the Moderate Resolution Imaging Spectroradiometer (MODIS), have sampling errors caused by incomplete coverage primarily due to clouds and inter-swath gaps (gaps between successive swaths). Unlike retrieval errors, the sampling errors are introduced when calculating mean values and in generating gap-free SST fields. This dissertation is focused on quantifying and parameterizing the global MODIS sampling errors.

The MODIS-sampled SST field is generated by superimposing MODIS cloud masks on top of the Multi-scale Ultrahigh Resolution (MUR) SST field for the same day. Based on the MODIS-sampled fields, sampling errors are calculated at different temporal and spatial resolutions to examine the impacts at different scales. In order to assess the robustness of the quantification using a reasonable reference field, we compare the sampling errors quantified using MUR and those generated from another very different reference SST field—HYCOM (HYbrid Coordinate Ocean Model) Global 1/12° reanalysis. Also, sampling errors are compared for variations between El Niño and La Niña events. The climatological component of the sampling errors are calculated and
assessed for its importance on sampling error estimation. Based on the error characterizations, an empirical model is proposed to parameterize the sampling errors using cloud masks and climatological or reference SST standard deviations.

Global sampling errors generated from both MUR and HYCOM reference fields are significant, more so in the high latitudes, especially the Arctic. The 30°N-30°S zonal band is found to have the smallest errors; a notable exception is the persistent negative errors found in the Tropical Instability Wave (TIW) area, where the mesoscale ocean-atmosphere interaction leads to a more frequently satellite sampling above the cold sections of the wave area. The global mean sampling error is generally positive and increases approximately exponentially with missing data fraction at a fixed averaging interval, while error variability is mainly controlled by SST variability. Areas with persistent cloud cover have large sampling errors in temporally averaged SSTs. As opposed to the fact that HYCOM and MUR SSTs are substantially different globally, geophysical patterns of the sampling errors generated from HYCOM reanalysis SSTs repeat those from MUR, giving rise to sampling error differences commonly within ±0.1 K. This result support the robustness of using a reasonable reference SST field for the quantification of MODIS sampling errors, and provide the evidence of the sampling error estimates being the consequences from missing observations and not the choice of the reference field. The negative errors in the TIW area change proportionally with the SST gradient, which is recognized of being modulated by the El Niño and La Niña events. The climatology component is demonstrated to be the dominant component in the sampling errors, especially for the errors caused by spatial averaging, therefore can be a reasonable estimates for sampling errors due to spatial averaging. Yet for the sampling errors caused
by temporal averaging, only 10% of the sampling errors can be attributed to the seasonal variation embedded in the climatology.

We propose an empirical sampling error model by incorporating the sampling error nonlinear dependence on cloud and SST variabilities. As a result, the sampling error estimates in many regions, particularly where warm sampling errors prevail, are largely improved compared to using only the climatological component.

This dissertation initiates the global characterization and parameterization of IR sampling errors due to clouds and inter-swath gaps. The results indicate the MODIS SST sampling errors can be an important or even dominant component of the error budget of mean and gap-free SST fields. Therefore, climate data generation and interpretation of satellite-derived SST CDRs and their application must be conducted with due regard to the sampling error.
Acknowledgement

First and foremost, I would like to give my big thanks to my advisor, Dr. Peter Minnett, for offering me the opportunity to study at the University of Miami, for his enduring patience, inspiring guidance, generous support, and persistent encouragement and trust, for giving me the freedom to find answers and raise more questions.

I would like to thank my other academic committee members, Dr. Paquita Zuidema, Dr. Ben Kirtmann, Dr. Mohamed Iskandarani and Dr. Chelle Gentemann, for their guidance, advice and discussions on my research.

I also want to thank Dr. Bill Johns who took me to my first research cruise. I also thank Thania, Adam, Mark, and all the crew members of the retired RV Knorr. The cruise was an unforgettable experience because of you.

I would also like to thank Honghai Zhang, who always endures listening to my frustration and complaints and returns with unconditional encouragement; Shitao Wang, who motivated me to get over difficulties and be a better me; Kieran, for all of those kitchen and lunch chats; Matt, Claire, Adam, Elizabeth, Johnna, Jian, Jie, Gustavo, Gino, Doug, Yemi for the all the help you offered in classes or in research; my gym pals, LinJie Feng and Shuwen Ma, for the happy Zumba times we share together every week.

The initial stages of this research were supported by a grant from the NASA Physical Oceanography Program (NNX11AF26G) and then by a NASA graduate fellowship to Y. Liu (NNX14AL28H). This dissertation would not been possible without these funding.

The data used in this research are public. The 1/12 deg global HYCOM+NCODA Ocean Reanalysis was funded by the U.S. Navy and the Modeling and Simulation
Coordination Office. Computer time was made available by the DoD High Performance Computing Modernization Program. The output is publicly available at http://hycom.org. AMSR data are produced by Remote Sensing Systems and were sponsored by the NASA AMSR-E Science Team and the NASA Earth Science MEaSUREs Program. Data are available at www.remss.com.

Last but not least, I would like to thank my parents for the endless love, always being supportive, and always having high expectations of me.
# TABLE OF CONTENTS

Chapter 1. Introduction ............................................................................................................... 1

Chapter 2. Sampling error quantification framework ............................................................. 6

Chapter 3. Global and regional MODIS fields ........................................................................ 12

  3.1 Sampling error annual mean characteristics .............................................................. 12

  3.2 Global sampling error dependence ............................................................................ 17

  3.3 Sampling errors seasonal and day-night variability .................................................. 19

  3.4 Cloud persistence ...................................................................................................... 22

  3.5 Regional characteristics ......................................................................................... 25

    3.5.1 Western Pacific Warm Pool ............................................................................... 26

    3.5.2 Tropical Instability Wave region ....................................................................... 27

    3.5.3 Gulf Stream and Kuroshio .............................................................................. 29

    3.5.4 Stratocumulus regions ..................................................................................... 31

  3.6 MODIS cloud mask ................................................................................................. 35

  3.7 Summary and discussion ......................................................................................... 38

Chapter 4. Sampling error sensitivity ..................................................................................... 43

  4.1 Background ............................................................................................................... 43

  4.2 Methods and data ..................................................................................................... 45

  4.3 HYCOM reanalysis vs MUR ................................................................................... 47

  4.4 Sampling error difference \( d_e \) ............................................................................... 52

  4.5 Sampling error variation with ENSO ...................................................................... 58

  4.6 Summary .................................................................................................................. 60

Chapter 5. The parameterized sampling errors ..................................................................... 63

  5.1 Background ............................................................................................................... 63

  5.2 The climatology component ..................................................................................... 65

  5.3 Sampling error parameterization ............................................................................ 67

  5.4 Results ...................................................................................................................... 69

  5.5 Summary .................................................................................................................. 74

Chapter 6. Conclusions ............................................................................................................ 75

Chapter 7. Future work ......................................................................................................... 78

References .......................................................................................................................... 80
Chapter 1. Introduction

Global Sea Surface Temperature (SST) is an essential climate variable (ECV, listed by the Global Climate Observing System) that can be used to assess climate change. In order to resolve the subtle signals that maybe indicative of a changing climate, long time-series of accurate, spatially and temporally averaged SSTs are needed. Specifically, an SST Climate Data Record (CDR) (National Research Council 2004) requires an absolute temperature uncertainty of 0.1K and stability of 0.04K per decade (Ohring et al. 2005). Such stringent requirements are intended to enable the detection of the likely regional or global signals of 0.2 K per decade. Hence, the correct quantification of errors and uncertainties in observed SSTs has become a critical need.

Among all the SST observing methods, satellites provide the most consistent global coverage. Infrared (IR) sensors in particular provide measurements of a fine resolution as well as having a long history. Therefore, for generating SST CDRs from satellite measurements (Minnett and Corlett 2012), IR measured SSTs are a potentially valuable source. The Moderate Resolution Imaging Spectroradiometer (MODIS (Esaias et al. 1998)) on board the NASA Earth Observing System satellites Terra and Aqua obtain SST retrievals in a 2330 km swath. SST is derived from MODIS measurements of top-of-atmosphere radiances in mid- and thermal-IR bands (centered at wavelengths of 3.7, 3.9, 4.0, 11 and 12 μm), at which the atmosphere is relatively transparent to the transmission of surface IR emission. The comparison with independent measurements from shipboard spectroradiometers (Minnett et al. 2001) confirms that the derived SSTs from MODIS generally have mean biases < 0.1K and scatter < 0.5 K (Minnett 2010).
However, in a general satellite data processing flow, errors from different sources are produced at each of the successive data levels (Level 0 (digitized detector output) to Level 4 (bias corrected, geo-located, gridded, and gap-free SSTs in lat/lon coordinates) and accumulate toward higher levels (Figure 1.1), as discussed by the Interim Sea Surface Temperature Science Team White Paper (ISSTST 2010). As with other satellite measurements, the MODIS SST accuracy refers to the retrieval error produced at Level 2 (derived SSTs in swath coordinates), but Level 3 (binned, gridded and averaged Level 2 values) and Level 4 fields are extensively used in climate and modeling studies, mainly because of the desirable features of being “gridded and gap-free”. Another important error
in Level 4 fields, independent of the retrieval error, is the sampling error caused by incomplete coverage of satellite measurements, and this is the focus of this dissertation.

There are two main reasons for this incomplete coverage. First, for any IR sensors such as MODIS, the presence of clouds causes gaps in the sampling, or ‘undersampling’, of SSTs. Cloudy pixels rejected by many currently used SST cloud masks constitute up to 90% of the total pixels sampled. Instead of being random, clouds around the globe form geographical patterns where some regions are prone to cloudiness while others are not. Some regions are even found with cloud-SST relations due to physical (e.g. Ramanathan and Collins (1991)) and dynamical mechanisms (e.g. Xie (2004) and Klein (1997)). Second, gaps between successive swaths of some sensors also lead to sampling errors. Sensors with narrower swaths are subject to a larger gap than others. The relatively broad swath of MODIS indicates that a scan gap of 432.8 km exists at the equator every 98.8 minutes. This gap narrows as it extends to the mid-latitudes of 32.3° poleward of which there is overlap of successive swaths. Consequently, these two factors become the fundamental issues in generating gap-free SST fields and lead to sampling errors.

In early studies, incomplete sampling issues were highlighted to ascertain the sampling errors of averaged climate data (Parker, 1984; Trenberth 1984a, b; Wigley et al. 1984). Recent sampling error studies for climatic time-space grid box averages of in-situ measured Surface Air Temperature (SAT) (Parker and Horton 2005; Shen et al. 2007) and SST (Brohan et al. 2006) are based on the quantification framework proposed by Jones et al. (1997) (referred to here as J97). In J97, the sampling error was expressed as the additional variance contributing to the grid box long-term temporal variance due to spatially incomplete sampling. Certain assumptions were made about the data statistics (e.g.,
homogeneity and stationarity) and the data spatial correlation curve. The sampling uncertainty was calculated by estimating averaged variance of stations and the inter-correlation between stations in the grid box. The SST data were mostly from ship and buoy observations and control run outputs from models. Morrissey and Greene (2009) developed a more general quantification framework by including temporally-insufficient sampling associated with ship measurements, assuming observations are randomly distributed. Kennedy et al. (2011) updated the work of J97 by applying an isotropic correlation decay function in both time and space. These previous works relied on the assumption made for the spatial or temporal inter-correlation curves within a grid box; additionally, the in-situ SSTs were used. Most recently, Hearty et al. (2014) quantified sampling biases in climatologies of atmospheric temperature and water vapor by comparing two MERRA (Modern Era Retrospective-Analysis for Research Applications; Rienecker et al. (2011)) climatologies sampled separately by the time and quality components of AIRS (Atmospheric Infrared Sounder) (Aumann et al. 2003) with a MERRA climatology sampled like a climate model, assuming that the MERRA data represents the real atmospheric state.

The sampling error of IR SSTs remains to be determined. Furthermore, IR SSTs have different sampling structures (not random) and known sources of the sampling error (i.e., clouds and inter-swath gaps). The aforementioned statistical assumptions may not be necessary nor appropriate for determining sampling error magnitudes, and they may smooth sampling error variations which in fact can give information on how the errors are generated and whether they can be reduced. These have motivated us to initiate the
sampling error quantification for satellite IR SSTs and scrutinize the characteristics of such errors to explore whether they can be reduced.

This dissertation is composed of four chapters. The first chapter lists the sampling error quantification framework that are developed in this study. The second chapter shows the MODIS sampling errors calculated using MUR as the reference and interprets the error generation processes associated with ocean-atmosphere interactions. Considering the error uncertainty that might be caused by assuming MUR as the reference field for the quantification the error variation in different years especially when climate event presents, we illustrate the sampling error sensitivity to the selection of reference field and the error magnitudes and patterns in a different year in Chapter 4. Inspired by the sampling error dependences, in Chapter 5, an empirical model is proposed to parameterize the sampling errors.
Chapter 2. Sampling error quantification framework

The global distributions of SSTs and clouds — the primary cause for missing IR surface observations — are determined geophysical process. The sampling errors caused by clouds are therefore a reflection of the undersampled geophysical SST variations, which might be connected to the corresponding cloud variations in certain SST-Cloud coupling situations. In addition, given that Level 3 SST fields are usually generated by averaging Level 2 fields into specific resolutions, the undersampling of SST variation at different scales might lead to different extents of contamination in the mean values of a target resolution. Whether the undersampling can cause a significant difference in the mean SSTs needs to be studied. Therefore, SST sampling errors should be examined for the generation processes and quantified globally, regionally, and at different scales.

We calculate the MODIS sampling errors without presuming SST spatial or temporal correlations. Instead, we assume that a reasonable Level 4 field can be the reference, or ‘true’ field to help quantify the impact of the under-sampling in IR fields, so the effects of sampling errors can be isolated and quantified. For the purpose of this work, we calculate the difference between the sampled fields and the corresponding gap-free reference fields as an ‘error’ instead of ‘uncertainty’ because of our assumption about the reference fields.

We select as the Level 4 reference field, the Multi-scale Ultrahigh Resolution (MUR) SST product (Chin et al. 2010), which is a 1km resolution daily SST analysis derived using observations from multiple sources including both satellite skin and subskin SSTs and in-situ SSTs. The satellite input fields include: 1) IR SST retrievals from polar orbiting satellite sensors AVHRR and MODIS; 2) microwave SST retrievals from AMSR-E and Windsat. The interpolated in situ data are from iQuam (in situ SST quality monitoring; Xu
and Ignatov (2014)), and are used as references for bias correction (Chin et al. 2010). MUR data can be retrieved freely from the NASA JPL PO.DAAC website: http://podaac.jpl.nasa.gov/dataset/JPL-L4UHfnd-GLOB-MUR. In order to match the MODIS cloud mask resolution, we aggregated MUR fields into 4 km daily maps.

The assumption that MUR to be a possible realization of the global SST field does not suggest that the MUR fields are error-free representations of daily SSTs. There certainly are inherent errors and uncertainties in MUR due to the blending of SSTs of different spatial scales. We build our assumption on MUR’s potential of providing realistic, if not perfectly accurate, representations of the small scale SST variability in dynamic regions (Vazquez-Cuervo et al. 2013). The consequence of the uncertainties in MUR errors are tested through the application of the methods developed here using a very different Level 4 field—HYCOM reanalysis SSTs. Details of such a sensitivity test are in Chapter 4.

The MODIS-sampled MUR is generated by eliminating the 4 km MUR pixels which are identified as cloudy or fall in an inter-swath gap. In other words, the MODIS sampling is represented by superimposing daily cloud masks on the daily MUR reference field. The MODIS retrieval errors, assessed using Level 2 data (e.g. Kilpatrick et al., 2015) do not contribute to our results. Figure 2.1 shows an illustration of the method. We superimposed the MODIS cloud mask (Figure 2.1b) to the MUR field (Figure 2.1a) to generate the sampled field (Figure 2.1c). Note that the global SST density (Figure 2.1d) of MUR and the MODIS-sampled MUR are quite different. By utilizing this approach, the undersampling impact i.e. the sampling error is represented and quantified as the difference between spatial or temporal mean fields of the MODIS-sampled-reference SSTs and the reference SSTs. We assessed the sampling errors of seasonal and 4-month means and
compared day and night fields. The merit of this approach is that we can suggest possible
deserves and impacts that can be physical and predictable, instead of just statistical, which
can further help develop solutions or predictions of the sampling error.

The cloud masks used are from the thermal IR daytime and mid-IR nighttime Level 3
fields of Terra MODIS SSTs. These data are globally gridded fields at 4 km spatial
resolution and were generated from the MODIS Collection 6 retrievals, which is the most
recent reprocessing of the MODIS SST. Global day and night cloud masks (i.e., quality
mask, referred to as cloud mask in this dissertation, with flags = 0 indicating the best quality)
were derived by considering quality flags >1 as missing data primarily due to cloud cover
and gaps between successive orbits. These Level 3 fields were generated for this study at
the University of Miami, and are available at http://oceancolor.gsfc.nasa.gov/.

Figure 2.1 Methodology. a: Reference field of MUR on 2011/04/07. b: Day mask of Terra
MODIS SST on the same day. Green colors show where quality 0 and 1 SSTs are derived.
c: The sampled MUR after superimposing b on a. d: SST probability density of a and c.
Note the difference between the two histograms shows the incompleteness of sampling.
Due to the occasional satellite instrument outages such as result from satellite maneuvers, a year-round continuously scanned global MODIS SST field is not achieved. Therefore, for each season, a 30-day period was selected as being a representative to determine the seasonal (refer to Northern Hemisphere) sampling error characteristics: winter: 20101228-20110125 (yyyymmdd); spring: 20110407-20110506; summer: 20110721-20110819; fall: 20111001-20111030. Any mean quantities calculated by averaging the four seasonal periods will be referred to here as an annual mean.

The sampling error definition and the statistical quantities used are introduced below. Let $SST^\text{ref}$ represent the reference data at a base resolution $R_0$ of 0.04° and 1 day (1d), which is the Level 4 field, MUR. Then, for a grid box centered at location $i$ and time $j$, the averaged reference field is

$$SST^\text{ref}_{i,j}(R) = \frac{1}{N_R} \sum_{n=1}^{N_R} SST^\text{ref}_0 n$$

where $N_R$ is the number of reference SSTs in the grid box. The averaged sampled field for the same grid box is expressed as

$$SST^\text{sample}_{i,j}(R) = \frac{1}{n_R} \sum_{n=1}^{n_R} SST^\text{ref}_0 n$$

where $n_R$ is the number of sampled reference SSTs; $0 < n_R \leq N_R$.

The sampling error due to cloud and inter-swath gap at grid box $i$, $j$ is given by
\[ \varepsilon_{i,j}(R) = \text{SST}_{i,j}(R) - \text{SST}^{ref}_{i,j}(R) \]  \hspace{1cm} (3)

This is the difference from the ‘true’ grid box mean and can be evaluated at any given time, location, and resolution.

In general, it can be expected that SST sampling errors are affected by two factors. One is the fraction of cloud or missing data (gap fraction) in the grid box.

\[ f_{i,j}(R) = 1.0 - \frac{n_R}{N_R} \]  \hspace{1cm} (4)

Here, \( 0 \leq f_{i,j} < 1 \). For the grid box statistics, we only use those with at least one SST value being sampled.

One important aspect of the gap fraction in this study is how long the gap persists, especially in cases of sampling errors due to monthly averaging. It is intuitive that even though the temporal gap fraction is large, the gaps might be short intervals that frequently occur during the month or clouds leading to the same gap fraction may occur in a single event. These cases may lead to different sampling errors. To examine this, we define the cloud persistence as the largest number of consecutive days during which the grid \( i, j \) was detected to be cloudy in a temporal averaging period.

The other essential factor for sampling error is the variance of the SSTs in the grid box, which is represented by the standard deviation of the reference SST in the grid box:

\[ \sigma_{i,j}^{ref}(R) = \left[ \frac{1}{N_R-1} \sum_{n=1}^{N_R} \left( \text{SST}^{ref}_{0} - \text{SST}^{ref}_{i,j} \right)^2 \right]^{\frac{1}{2}} \]  \hspace{1cm} (5)

To assess this error, statistics will be calculated for global and regional ocean areas, primarily in terms of the three quantities below: mean sampling error, root mean square error and difference of mean SST in a region of interest, \( A \), and during a time interval, \( t \). \( \text{lat}_i \) in the equations below is the latitude of each sample.
\[ \bar{\varepsilon}_{A,t}(R) = \frac{1}{\sum_{i=1}^{n_A} \sum_{t=1}^{n_t} \cos(lat_i)} \sum_{j=1}^{n_t} \sum_{i=1}^{n_A} \varepsilon_i,j(R) \cdot \cos(lat_i) \quad (6) \]

\[ \text{RMSE}_{A,t}(R) = \left[ \frac{1}{\sum_{j=1}^{n_t} \sum_{i=1}^{n_A} \cos(lat_i)} \sum_{j=1}^{n_t} \sum_{i=1}^{n_A} \varepsilon_i,j(R) \cdot \cos(lat_i) \right]^{1/2} \quad (7) \]

\[ \delta\overline{SST}_{A,t}(R) = \frac{1}{\sum_{j=1}^{n_t} \sum_{i=1}^{n_A} \cos(lat_i)} \sum_{j=1}^{n_t} \sum_{i=1}^{n_A} SST_i,j(R) \cdot \cos(lat_i) \]
\[ - \frac{1}{\sum_{j=1}^{n_t} \sum_{i=1}^{n_A} \cos(lat_i)} \sum_{j=1}^{n_t} \sum_{i=1}^{n_A} SST_{i,j}^{ref}(R) \cdot \cos(lat_i) \quad (8) \]

Regional gap fraction is defined as the missing data fraction that cannot be filled even by averaging into a low resolution of \( R \):

\[ f_{A,t}(R) = 1.0 - \frac{\sum_{i=1}^{n_A} \sum_{j=1}^{n_t} \cos(lat_j)}{\sum_{i=1}^{n_A} \sum_{j=1}^{n_t} \cos(lat_j)} \quad (9) \]

By gridding the sampled and reference SSTs into a range of temporal and spatial resolutions, the undersampling impacts are reflected in different temporal and spatial averaging situations. One way to consider the difference between the grid box statistics (\( \bar{\varepsilon}_{A,t} \) and \( \text{RMSE}_{A,t} \)) and the regional statistics (\( \delta\overline{SST}_{A,t} \)) is that the statistics calculated based on grid box \( i, j \) gives an evaluation of impacts from the “subgrid scale clouds” (clouds or gaps that have scales less than the box size defined by \( R \); also defined as \( f_{i,j} \)), while the statistics based on region \( A \) and time interval \( t \) provides additional impacts from grids that are not filled, or, from the “super-grid scale clouds” (clouds or gaps that have scales at least the box size; also defined as \( f_{A,t} \)). The impacts due to certain scale clouds and SST temporal and spatial variation are therefore revealed in the different resolution fields. Changes in global error magnitude and error patterns from one resolution to another can be attributed to the impacts from the inclusion of oceanic features and clouds on these scales.
Chapter 3. Global and regional MODIS fields

3.1 Sampling error annual mean characteristics

Annual means of the sampling error are estimated by averaging over the full time period at a certain resolution R, and are shown in Figure 3.1. Generally for these three cases, global sampling errors are prevalently within ± 0.5 K. Regions where fronts and upwelling occur have some values at about ± 1.0 K. Extreme values of ± 1.0 K to ± 6.0 K only occur in the Hudson Bay and other seasonally open water regions at high latitudes. The error pattern varies with resolution. In the [4k, mon] map, positive errors (~ 0.5 K) are evident off the west coasts of North and South America and South Africa, where maritime

Figure 3.1 a, b, and c: Global annual mean (t = 4 seasonal periods) sampling error distribution using MODIS (Terra) cloud masks in 2011. White labels on the Asian continent indicate the resolution; White color in ocean areas indicates zero error, which is denoted in the middle of the color bar; Black indicates either land or sea ice; Grey indicates the region with super-grid scale clouds, thus is excluded when generating the probability density. Note the color scale is non-linear. d: SST probability density of the maps shown. The number of grids counted is displayed in the upper right corner.
stratocumulus clouds are common. There is a large warm (~ 0.5 K) pattern in the South-Central Pacific (SCP, 175°W-135°W, 25°S-45°S) region. This is due to a significant warming event in this area in January of 2011, similar to the SCP warming of the previous year. The 2009-10 SCP warming was associated with a strong divergence of circumpolar westerlies, low wind speed (Lee et al. 2010) and stronger solar radiation at the surface (Liu et al. 2014), which are usually accompanied by gradually clearer skies and therefore lead to warm sampling errors. The cold (~ -1.0 K) zonally stripe-like patterns at the eastern equatorial Pacific and Atlantic oceans exist along the tropical instability waves (TIW, details in Section 3.5.2). The error distributions along the western boundary currents are rather diverse (details in Section 3.5.3). Similarly variable are those errors in the high latitude regions usually covered by stratus. However, the [5°, 1d] error distribution shows some differences. Tropical regions are prevalently filled with negative errors of ~ 0.1 K. Warm patterns in the west coast stratocumulus regions are variable. The positive error region in the SCP is absent. Errors along the ocean fronts and western boundary currents are more coherent, with cold errors to the warmer side of the current and warm errors to the cold side. The [5°, mon] map has the above-mentioned features with larger errors, except that the error signs in the tropics have changed to be more positive than in the [5°, 1d] case, yet the cold error in the TIW remains. Hence, a more interesting La-Niña-like error pattern becomes apparent in the tropical Pacific. The changing error sign in the tropics is reflected in the SST density in Figure 3.1d. The SST error distribution shifts toward the positive as the resolution shifts from the spatial averaging case of [5°, 1d] to the temporal averaging case of [4k, mon]. The [5°, mon] case shows even warmer errors than the [4k, mon] because of the including of monthly persistent cloudy regions that were excluded
from [4k, mon]. Xie (2004) summarized the global low-level boundary layer clouds response to SST changes that over most regions result in a negative SST-cloudiness correlation due to basin-wide dominant static stability. That more warm sampling errors are found in [4k, mon] fields supports the concept of the long-term basin-wide negative correlation, while the sampling errors in [5°, 1d] are associated with weather scale SST-cloud correspondence.

The zonal and annual means of the sampling error are shown in Figure 3.2. The mean error generally increases with a decrease of resolution or increase of latitude in both spatial and temporal averaging cases. This is because the Rossby radius and eddy scales decrease with increasing latitude (Hoyer et al. 2012), and the lower the resolution, the more scales of variability become involved and affect the sampling errors. The maximum appears at high latitudes for both hemispheres. The secondary peak exists in the northern hemisphere mid latitudes, where SST variability is influenced by western boundary current systems and along storm tracks. The increase of mean error with latitudes does not apply to the spatial averaging in the tropics, where zonal mean cold errors exist with the maximum (~
-0.15K at 5°) close to the equator for all spatial averaging cases, and decrease away from the equator. This is associated with the significant cold error along the TIWs shown in Figure 3.1.

In order to assess the sampling error at all resolutions in different latitude bands, the regional and annual statistics of the RMSE, the mean sampling error, the difference between the regional mean SST, and the regional gap fraction are generated for 5 zonal

Figure 3.3 Variation of the 4 quantities (columns) in the resolution domain (x: spatial; y: temporal) are evaluated at different latitude bands (rows). First column: RMSE (contour interval = 0.1K); Second column: the zonal-band and 4-month averaged sampling error $\varepsilon_{A,t}$ (contour interval =0.05K); Third column: The difference between the sampled annual mean SST and the annual mean MUR SST of the zonal band (contour interval = 0.2K); Fourth column: The regional gap fraction of the field due to super-grid scale clouds/gaps. Definitions are given in Eqs. (6)-(9).
bands (Figure 3.3). By examining the evolution of RMSE in the resolution domain, one finds the distribution follows a saddle-like pattern with extremely large values falling into either the spatial averaging or temporal averaging case (the bottom-right and upper left corners) and smaller values appear along the diagonal indicating more symmetrically spatiotemporal averaging. The difference of mean SST ($\delta \bar{SST}_{A,t}$) is mostly positive and is maximum at the highest resolution, primarily because of the sparsest data coverage for cold SSTs. This suggests large scale data loss can cause significant high mean SST sampling errors. Going to lower resolutions, the missing data become filled by averaging and the $\delta \bar{SST}_{A,t}$ grows smaller and closer to $\bar{e}_{A,t}$. This illustrates that eliminating the significant positive biases in the regional or global mean SSTs often comes at the cost of coarser resolution. The regional gap fraction decreases monotonically when going to the lower resolution or lower latitudes, which is associated with the more super-grid scale clouds in the high latitudes. In the 30°S-30°N region, the RMSEs and mean sampling errors are the smallest among the latitude zones. The mean errors are within ±0.05 K, and the RMSEs are within 0.3 K except in the [5°, 1d] case. The sampled mean SST does not differ greatly from the true mean as shown in the $\delta \bar{SST}_{A,t}$. Even though this is the region affected by both clouds and inter-swath gaps, gap fraction in this region remains the lowest when compared to others. These indicate the relative reliability of sampling within 30°S-30°N for most gridded SSTs. Going to the higher latitudes, the RMSE and $\delta \bar{SST}_{A,t}$ distributions exhibit more complexity. RMSE values are larger in the northern hemisphere than in the southern. The $\delta \bar{SST}_{A,t}$ is remarkably large (> 2.5 K) in the 30°N-50°N zone.
3.2 Global sampling error dependence

The results above show sampling errors vary substantially over the globe. As one might expect, natural SST variation and gap fraction are the two reasons. However the extent of such expected dependence as yet remains to be quantified. So, for the global sampling error,
shown as representations (Figure 3.4 a and b). In these cases, the sampling error magnitude increases substantially as SST standard deviation or gap fraction increases. Note that the portion where sampling errors are within 0.1K is limited to: monthly standard deviation of <0.3K or 5° spatial standard deviation of <0.7K; monthly gap fraction of <10% or 5° spatial gap fraction of <4%. Such restrictions can be relaxed when assessing cases averaged more symmetrically, or single averaging cases at higher resolutions (Figure 3.5). Figure 3.4, c and d, show the mean sampling error is positive and increases approximately exponentially with gap fraction. On the contrary, averaging over standard deviation intervals (not shown) does not indicate any dominant dependence.

The spread of the sampling error depends on SST variability as well as the gap fraction. The RMSE shows stronger dependence on the reference SST variability (Figure 3.5 c and

![Figure 3.5](image)

Figure 3.5 Dependence of the RMSE on gap fraction (a. temporal averaging of: [4k, 3d], [4k, 1w], [4k, 2w] and [4k, mon]; b. spatial averaging of: [12k, 1d], [0.25°, 1d], [0.5°, 1d], [1°,1d], [2.5°, 1d], and [5°, 1d].) and reference standard deviation (c. temporal and d. spatial). Resolutions are represented by the size of the circles. The smaller the circle is, the higher the resolution.
d) than on the sampling gap fraction (Figure 3.5 a and b). Note that the RMSE varies approximately linearly with the reference standard deviation. Increasing spatial resolution reduces the RMSE dependence on spatial gap fraction, while different temporal resolutions barely change the RMSE dependence on temporal gap fraction. Nonetheless, the change of resolution makes minimal difference to the near-linear dependence of RMSE on reference SST variability. This demonstrates that as long as the gap-free SST variability is known, the RMSE generated by the current cloud masks can be predicted regardless of the resolution selected.

3.3 Sampling errors seasonal and day-night variability

Mid- (30°-50°) and high- (50°-70°) latitudes are found to have large sampling error variability (Figure 3.3). The sampling error in these regions is further decomposed by seasons and day-night sampled fields. Monthly error variation in the resolution domain does not necessarily follow the patterns of annual mean statistics shown in Figure 3.3. For example, spatiotemporal averaging might cause the RMSE in some seasons to exceed that in the single dimension averaging case. Monthly mean errors no longer necessarily increase monotonically with decreases of spatial or temporal resolution. In this section, we continue to present [4k, mon] and [5°, 1d] monthly statistics (Figure 3.7) as examples, and compare the seasonal variation of sampling error statistics among different zonal bands, and between time (day or night cloud masks).

Evidently, RMSEs and mean errors (the first and second rows in Figure 3.7) of the Northern Hemisphere (NH) show more in-phase fluctuations regardless of latitude, resolution, and day-night sampled fields. The maximum RMSE and mean error generally exist in the summer month of the NH mid- and high-latitudes, due to the larger SST
variability and higher amounts of cloud, probably stratus. One exception is the spatial averaging case in the NH mid-latitude band, where the largest RMSE occurs in the spring month, possibly due to a Pacific storm. The high-latitude summer peak is due to the increased SST variability brought by open water around the ice edge (Figure 3.6). Also in this area, the mean sampling error is negative in the winter month, primarily due to the large amount of negative sampling errors of -0.5 - -1.0 K that occur in the Gulf Stream and Kuroshio in this season. On the other hand, RMSEs of the Southern Hemisphere (SH) exhibit small variation across the 4 months. The time of the weak RMSE peaks

Figure 3.6 The seasonal patterns of gap fraction (first column), sampling error in monthly averaged ssts (second column), and the standard deviation (third column) calculated for that month. Shown here are the statistics of day time. Grey indicates the region with super-grid scale clouds.
differs from one to another. The mean sampling error of the SH high latitudes at [5°, 1d] shows a large variability among the seasons. In aspects of the regional mean temperature ($\delta$SST$_{A,t}$) and the super-grid scale clouds ($f_{A,t}$), seasonal variations of these two quantities are generally in phase (Figure 3.7 the third and fourth rows). This indicates again the significance of large scale cloud coverage on causing high mean sampling errors in SST. Seasonal variations of the $\delta$SST$_{A,t}$ and $f_{A,t}$ in the NH generally peak in summer except for

Figure 3.7 Seasonal and day-night variation. Columns represent Southern Hemisphere and Northern Hemisphere. The seasons in the x-axes refer to those in each hemisphere. Mid- and high-latitude bands are differentiated by color: Orange: mid-latitude; blue: high-latitude. Resolution is represented by line style: Solid:[4k, mon]; Dot: [5°, 1d]. Circled and starred lines represent daytime and night time cloud masks respectively.
the temporal averaged case in the high-latitude band. In the SH, seasonal variations of the \( \delta \overline{\text{SST}}_{A,t} \) and \( f_{A,t} \) present a large spread across different resolutions and sampling time.

Day and night differences of the 4 statistical quantities in the NH are small, compared with the SH. Note only the NH high-latitude band shows considerable differences in the \( \delta \overline{\text{SST}}_{A,t} \) and \( f_{A,t} \) between day and night. On the contrary, notable inconsistencies in the aspect of the seasonal peak of the SH RMSE and mean sampling error exists between day and night field. Especially, the daytime \( \delta \overline{\text{SST}}_{A,t} \) and \( f_{A,t} \) generally peak in the winter month, while the nighttime values peak in the summer month. The most significant day-night difference remains in the SH high latitude band, and thus blurs the seasonality. Since the same daily MUR field is used for both this difference cannot result from SST diurnal variability and the diurnal differences are therefore caused by diurnal cloud variability, and possibly indicate an imperfection in the cloud masks. This will be discussed in Section 3.6.

### 3.4 Cloud persistence

The maximum cloud persistence at each grid box around the globe is extracted from both day and night cloud masks (Figure 3.8). Long persistence is found in the high latitudes. Also, regions known to have prevailing stratus and stratocumulus clouds, for example, the north Pacific and regions off-shore of California, Peru, and Namibia, are shown to have long cloud persistence. In addition, the Indian-Pacific warm pool and the ITCZ, which generally are expected to have long-time presence of convective systems, and do indeed display long cloud persistence in Figure. 3.8. From the previous figures, large sampling errors correspond to those regions with long cloud persistence. Seasonally, the ice-edge-region of the Arctic Ocean shows high variability. The summer and fall months are when the persistence can be less than 25 days in this region, while in the winter and spring, some
areas are covered by ice, the remaining areas have very long persistence of almost 30 days. It is also notable that Pacific mid-latitudes have a persistence peak in summer. This area is known to have pronounced stratus clouds in summer.

Figure 3.8 The maximum cloud persistence at each grid point. The boreal season and time of sampling (D: day; N: night) are denoted on the Asian continent.
However, large differences exist between day and night cloud persistence around Antarctica. The difference is more evident in the austral autumn and winter (boreal spring and summer in Figure 3.8), when some areas are determined to be consistently cloudy in the daytime for almost the whole month, but cloud-free at night for most of the days. This is consistent with the diurnal variation in the statistics of sampling error (Figure 3.7 $\delta\overline{SST}_{A,t}$) due to super-grid scale clouds (Figure 3.7 $f_{A,t}$) in this region. Again, we believe this is related to imperfections in the cloud masks, which will be discussed in Section 3.6.

A larger sampling error is expected in areas with longer cloud persistence. But, a large gap fraction does not necessarily dictate long persistence. Therefore, for the sampling error in temporal averaged SSTs, even though gap fraction is suggested as a source of error (Figure 3.4 a), the dominant cause is shown to be the cloud persistence (Figure 3.9). The mean absolute sampling errors are calculated at each gap fraction and persistence for the 4 months. For the same gap fraction, the sampling error tends to increase as the persistence increases.
3.5 Regional characteristics

The magnitudes and dominant causes of the sampling errors vary globally. In this section, we characterize and interpret the sampling error in 4 important types of region (Figure 3.10). Firstly, the Western Pacific Warm Pool (WPWP: 140°E-160°W, 10°S-15°N) is known to have the highest SSTs, consistently ≥28°C. In this convective region, prevalent clouds include precipitating mid- to high-level stratiform clouds that heat the upper troposphere and cool the lower troposphere, and cirrus clouds close to the tropopause; SSTs have relatively low variability due to small horizontal temperature gradients. Secondly, the Tropical Instability Wave (TIW: 160°W-80°W, 5°S-10°N) region is selected because the marked SST gradients are associated with distinct forcing of the atmosphere above. Thirdly, the two important NH western boundary currents

Figure 3.10 Geographic locations of the regions: Western Pacific Warm Pool (WPWP: 140°E-160°W, 10°S-15°N); Tropical Instability Wave (TIW: 160°W-80°W, 5°S-10°N); Kuroshio Current (KS: 120°E-150°E, 15°N-50°N); Gulf Stream Current (GS: 80°W-45°W, 25°N-55°N); Peruvian Coast (PERU: 95°W-75°W, 20°S-10°S), and Namibian Coast (NAMI: 5°W-15°E, 20°S-10°S).
regions are defined as 120°E-150°E and 15°N-50°N for the Kuroshio (KS), and 80°W-45°W and 25°N-55°N for the Gulf Stream (GS). SST variability in these two regions is greater than in many other areas, and with strong and more complex air-sea interactions present. Lastly, we analyze the two most pronounced marine stratocumulus regions, where satellite sampling of SST is rather scarce: the off-shore coastal regions of Peru (90°W-80°W, 20°S-10°S), and of Namibia (0°W-10°E, 20°S-10°S).

3.5.1 Western Pacific Warm Pool

![Figure 3.11 a. The dependence of mean sampling error on gap fraction in the WPWP. The single temporal and single spatial cases are shown. b. The WPWP $\delta SST_{A,t}$ variation in the resolution domain.](image)

Overall, sampling errors in the WPWP are relatively small, as the SST variability is small (standard deviation $< 1K$, calculated from one month of daily SST fields). Although the sampling error magnitude is not as large as in the high latitudes, there is a difference of sign between the means of the temporal and spatial sampling errors (suggested in Figure. 3.1 and shown in Figure 3.11 a). Also, the mean sampled SST of this region is lower than
the mean of the reference SST field (Figure 3.11 b), which is opposite to the global sampled mean (Figure 3.3 $\delta \text{SST}_{A,t}$). The preponderance of negative sampling errors on daily spatial averaged SSTs might indicate the significance of local convergence to the SST-cloud relationship (Xie 2004), or the spatial correspondence between deep convection and warmer SSTs indicative of the local regulation of SSTs by cirrus clouds being part of the thermodynamical constraint (Ramanathan and Collins 1991). For non-weather scales (monthly, as considered here), the radiative cooling effect of clouds on SST plays a role (Ramanathan and Collins 1991; Ramanathan et al. 1995), so that the long-term mean sampling error tends to be positive. From Figure 3.11 b, we suggest that the sampling error in mean SST of WPWP region can be within 0.05K, if the SST fields to be averaged are at a temporal resolution of 1-week to 1-month, or at 3-day if the spatial resolution is at or lower than 0.5°.

### 3.5.2 Tropical Instability Wave region

As seen from previous figures, the TIW region is the exception in the tropics regarding the sign of the sampling error. This region has significant horizontal SST gradients to which the atmosphere boundary layer strongly responds. The vertical mixing mechanism in the atmosphere (Hayes et al. 1989; Wallace et al. 1989) dominates the mesoscale ocean-atmosphere coupling in this area. To the north of the TIW, where the SST gradients in this cusp-shaped front are stronger, the surface wind converges slightly northeast of the warm SST anomalies, which results in the formation of low level clouds because of the convergence of surface moisture (Chelton et al. 2001; Chelton et al. 2004; Chelton and Xie 2010; Hashizume et al. 2001; Hashizume et al. 2002; Mason et al. 1996; Small et al. 2008; Xie 2004). The gap fraction shown in Figure 3.12 b can be deemed as cloud fraction, except
that the inter-swath gaps may generate striped patterns that are more obvious in regions where cloud occurrence is low. Two phenomena can be observed. First, the significant warm sampling errors are found in the offshore region of Ecuador and Peru, where the SSTs are the coolest, due to upwelling and the cloud fraction is large. Second, in the downstream region further away from the shore, significant cold sampling errors can be found along the larger SST gradient, northern side of the TIW. The warm sampling errors can be attributed to the modulation of SST on the lower troposphere static stability. The decrease of SST can stabilize the frequent atmospheric temperature inversions, which in

Figure 3.12 The 2011 October mean SST (a), gap fraction (b) and the resulting sampling error (c) in the TIW region. SST contours (1 K interval) are shown in b. and c. for comparison.
most cases favor greater occurrence of low-level stratus (Klein and Hartmann 1993). This usually can be found in a stratocumulus deck above ocean upwelling. Note that fog forms frequently in such upwelling regions and is treated as cloud in the satellite SST retrieval process, so the warm errors could also be caused by the presence of fog. The large negative sampling errors along the north side of the TIW indicate that there are more colder-than-average SST measurements, which support the robust SST-cloud coupling suggested by the atmospheric vertical mixing mechanism. Figure 3.12 shows the monthly mean SST, which does not perfectly reveal the instantaneous coupling of warm SSTs beneath low-level clouds. The dynamic activity (e.g. meridional shifts and westward wave propagation) in a grid box, among the fixed grids defined for any resolution in this study, can lead to temporal changes of both clouds and SST variations. Although not shown in Figure 3.12, it has been seen in satellite images that clouds propagate westward along with the SST warm anomalies in the TIW (Chang 1970; Legeckis 1977). Thus, mean sampling errors of temporal averaged SSTs in this region are negative. Similarly, negative sampling errors can be found along the strong SST gradient in the spatial averaged fields as well. Simply put, when the skies are clear and SSTs can be measured from satellite in the infrared, the SSTs are lower than average, as the warmer-than-average SSTs tend to be associated with cloudy skies when SST cannot be derived from infrared satellite measurements.

3.5.3 Gulf Stream and Kuroshio

Sampling errors of the spatially averaged SSTs vary substantially in the GS and KS regions and among seasons (Figure 3.13). According to the aforementioned linear relation between SST standard deviation and RMSE for the globe (Section 3.2), errors here seem to be largely determined by the SST variability. Nonetheless, several well-known processes
First, and also in concert with the TIW, atmospheric response to SSTs in these two ocean frontal regions involves the ubiquitous vertical momentum mixing mechanisms and the resultant wind convergence as well as low level cloud forming above the downwind warm water (Chelton and Xie 2010; Kilpatrick et al. 2013). Here direct pressure adjustment is also found (Lindzen and Nigam 1987), in which a good correspondence exists between wind convergence and the sea level pressure Laplacian (Minobe et al. 2008). Low level clouds show a sharp transition across the GS front due to the secondary circulation induced by pressure adjustment (Liu et al. 2014). Li et al. (2004) reported a cloud line forming right above the GS axis, where the pressure-induced wind curl is the maximum. Similar adjustments were found in the KS region (Kawai et al. 2014; Liu et al. 2013; Tokinaga et al. 2009). Both mechanisms suggest that the SST-cloud relationship, if it exists, should be positively correlated and thus lead to negative sampling errors. Second, seasonal variations of the direct pressure adjustment have been shown to strongly affect the cloud distribution in the GS area: in winter, wind convergence is strongest right over the GS between Cape...

Figure 3.13 Monthly mean sampling errors of 5° spatially averaged daytime sampled SSTs in: the Kuroshio region (first row) and the Gulf Stream region (second row). The 4 columns show different seasons.
Hatteras and the Grand Banks, and is accompanied by enhanced precipitation and more midlevel clouds (Minobe et al. 2010). Indeed, in Figure 3.13, the GS region shows strong negative sampling errors over the northern sector in winter, yet the dominant negative errors are missing in the other seasons. Regardless of the mechanisms, the pathways of these two western boundary currents can affect synoptic variability. The GS front is more intense and confined to a narrow meridional band of about 100 km, while the KS front appears to be more diffusely distributed. This geographical feature influences atmospheric variables, such as latent and sensible heat release and horizontal divergence (Joyce et al. 2009). We found similar features in the sampling error patterns with large errors in the KS spread wider in the meridional direction (Figure 3.13). It is also widely known that frequent fog and haze form over the slope waters. Fog can be identified as clouds in IR and thus affect the sampling errors. To summarize, the strong variability in the sampling error results from the concurrent coupled air-sea processes and the error cannot be attributed to any single process.

3.5.4 Stratocumulus regions

![Box-whisker plots of sampling errors at different cloud persistence. Peruvian (left) and Namibian (right) coasts are shown.](image)

The Peruvian and Namibian coasts are recognized as being covered by persistent stratocumulus clouds. As seen in Figure 3.8, the maximum monthly persistence can be ~
25 days in these areas. In Figure 3.14, we show that generally sampling errors become warmer as the cloud persistence becomes longer in both regions. In other words, the longer the sampled cloudiness, the colder the sampled mean SST for that grid box. This relationship is seen in each season, although not evidently in the Peruvian coast winter (Figure 3.15). Causes for this relationship could be several. Klein and Hartmann (1993)
found that the modulation of SST on the lower troposphere static stability is important for the low-level stratus cloud cover in these two regions. The decrease of SST can stabilize the frequent temperature inversions, which in most cases favor frequent cloud occurrence. Later Klein (1997) pointed out that such a modulation also exists at synoptic time scales. The prevalence of positive sampling error here can possibly be due to such a modulation, which gives a negative correlation between SST and cloud fraction.

Besides, if the reduced solar radiation at the sea surface due to persistent cloudiness decreases the SST in these regions, the variation in the amplitude of the SST could be magnified by a positive feedback and warmer sampling errors can be found at longer cloud persistence. The amplification of the SST annual cycle by cloud “shading” effects was found in oceanic observations off the Peruvian coast (Takahashi et al. 2005). Kubar et al. (2012) found that starting from the time scale of ~15 days the low-topped cloud fraction (LCF) and SST show negative correlation comparable to annual cycle statistics; where the primary LCF variance is explained by the annual cycle, the maximum LCF leads the minimum SST by 15-30 days. The domains where this was found include the two stratocumulus regions here. We believe that the time scale of 15-30 days indicates the dominant warm sampling errors commonly starting from 20-day cloud persistence and the likely internal positive feedback. But more research is needed to support this mechanism. Note that in the Peruvian coastal area a positive feedback exists between negatively correlated low-level cloud amount and SST on a decadal time scale (Clement et al. 2009). Our findings here suggest that any long term relationship between SSTs that include IR measurements and clouds in these regions could be potentially obscured, especially during periods of cold SSTs, because of the cloud induced sampling errors in IR measurements.
Although during the studied 4 months, warm sampling errors seem to dominate the two regions, the sign of local sampling errors are in fact associated with the correlations between SSTs and cloud occurrence. In the winter month, sampling errors of the Peruvian stratocumulus region are found slightly decrease with the increase of cloud persistence after 22 days (Figure 3.15). To further investigate whether this is an exception as oppose to the prevalent low-level cloud and SST negative correlation found in June-July-August (JJA) by Klein and Hartmann (1993), we generate the time series of SST-CLW (Cloud Liquid Water) correlations by using AMSR-E measurements from 2002-2011. Figure 3.16 shows a strong seasonal signature of SST-CLW correlations: positive correlations centered at the JJA and negative ones centered at the DJF (December-January-February). The positive correlation in JJA of the Namibian region is pronounced and stronger than those of the Peruvian, while the negative correlations in DJF of the Peruvian region is stronger than the Namibian. These results indicate the possibility of a sampling error seasonal cycle with more cold errors in JJA and warm errors in DJF. The decrease of error with cloud persistence in the JJA Peruvian region falls into the season when there are positive correlations between SST and CLW, and could be a consequence. However, using CLW
to deduce characteristics of the sampling errors might be problematic since the MODIS cloud mask cannot represent CLW measurements but are more close to the cloud occurrence. High CLW content could indicate low cloud occurrence if there are more thick clouds (usually puffy cumulus) and the total latent heat flux into the atmosphere remains the same. The interpretation of the SST-CLW correlations and the relationships to the sampling errors need further investigation. That said, the stronger dependence on the cloud persistence than the gap fraction (in temporal averaging can be deemed as the cloud occurrence) found in sampling errors could suggest new mechanisms since this has not been analyzed before.

3.6 MODIS cloud mask

One noticeable feature in our results is the day and night inconsistency of sampling error statistics, especially in the Southern Ocean. The reference field (MUR) is the same for both day and night sampling; the only source of the different day and night statistics is the different day and night cloud masks of MODIS. Therefore, unless the diurnal variability in cloud cover is substantial, the diurnal difference in gap fraction should be small. However, note that the MODIS SST cloud masks not only flag pixels identified as clouds, but also flag pixels with poor quality level due to radiometric errors. For the cloud flags, the day-night difference can be related to the use of a visible band for cloud detection during the daytime but which cannot be used at night; for other low quality level flags, the day-night difference is related to the different thresholds used in the quality tests. Consequently, even if there were no diurnal changes of clouds in reality, differences between day and night cloud masks could be introduced simply through the methods of cloud identification.
In fact, the day-night gap fraction difference shows meridional and seasonal variations (Figure 3.17). The winter hemisphere shows higher daytime gap fraction than the summer hemisphere, while in spring and fall, mid and high latitudes show higher daytime gap fraction than the tropics. The daytime gap fraction in high latitudes is about 30% higher than the nighttime, especially in the Southern Ocean. However, the tropical daytime gap fraction is lower than the night-time in all four seasons. Interestingly, over the North and South Pacific Ocean gyre centers, where it is mostly “clear”, the daytime gap fraction is higher than at the nighttime. For sampling errors, the day-night difference
does not show analogous patterns to the mask difference. Instead, large differences mostly occur in the cloudy regions. There can be multiple causes of the day-night cloud mask meridional and seasonal differences, but the identification of the causes needs further investigation into the cloud screening algorithm.

The sampling error day-night differences show different patterns not only by seasons but also by resolutions. In the [4k, mon] case (second column in Figure 3.17), the sampling error pattern shows large day-night differences in regions with frequent cloudiness. The error difference in the “clear” regions is essentially negligible. On the other hand, the error difference in the [5°, 1d] case shows more correspondence with the gap fraction difference. Particularly in the Southern Ocean, where clouds are often multilayer, warmer daytime sampling errors coexist with higher daytime gap fraction and vice versa.

Mesoscale SST anomalies associated with oceanic eddies modify the atmospheric boundary layer, including in the Southern Ocean: Frenger et al. (2013) found a positive correlation between SST anomalies and cloud fraction in this area; Chelton (2013) commented that the surface expression of eddies imprinted on the surface wind, low-level clouds, and precipitation are omnipresent. Such robust dynamical links can be manifested in the sampling error, but only if the MODIS cloud mask represents realistic cloud fraction and if low level cloud dominates. However, it seems that both conditions apparently do not hold in the Southern Ocean. Taken together, an SST-cloud correlation is less likely reflected in the high latitude sampling errors because of the enhanced uncertainty in the cloud mask, and this uncertainty is most likely generated by intrinsic cloud screening algorithm instead of physical causes.
3.7 Summary and discussion

We analyzed the effects of daily MODIS SST cloud masks of 4 months to characterize the resulting sampling errors. Note that because the MODIS swath width is insufficient to provide overlap of successive swaths equatorward of 32.3° latitude, the resulting systematic gaps in the SST fields are included in this analysis with the missing data that result from the presence of clouds. The annual mean sampling error magnitude might fluctuate around the values we calculated for the 4 sample months. Similarly, the seasonal variation extracted from the 4 months might be less representative of the error seasonality, which connects to complete seasonal cycles of clouds and SSTs. Sampling errors also occur in generating Level 3 fields (4 km as used here) from Level 2 swaths (1 km for MODIS), but we anticipate these errors to be small, as the error growth is very flat from 4 km to 0.25° (Figure 3.18). Therefore sampling errors in 4km fields are neglected in this work. However, the results shown here are indicative of how large are the error magnitudes, compared with the CDR target accuracies of 0.1 K and decadal stability of 0.04 K (Ohring et al., 2005).

We selected one Level 4 field, MUR, as the reference, assuming it to be a reasonable, not necessarily a perfectly accurate, realization of the daily SST fields. For example, MUR has unresolved SST variability on scales < 25 km in regions where extensive clouds are present, due to only microwave SSTs being used. To assess the error in the estimate of the mean that is caused by unresolved variability in microwave SSTs, we selected the grid boxes at [0.25°,1d] resolution that are identified with gap fraction f=0, and use these to represent the SST fields that would be given by on the microwave (AMSR-E) SST retrieval grid, i.e. the SST variability in these grid cells is taken as the
sub-cloud SST variability missing from MUR in cases where the cloud extent totally covers a 25 km grid cell. The standard error of the estimate of the mean of sampled \([4k,1d]\) sub-grids in the \([0.25º,1d]\) areas is then considered to be the error in the SST that is caused by unresolved variability in the microwave SST retrieval grid cell. We show the global standard error cumulative histogram of the 4 months (Figure 3.18). These are day time statistics, but those from night time is very similar. In all cases, more than two-thirds of the error is less than 0.02K. We anticipate the error due to unresolved variability to be very small and thus can be neglected. Being aware that the input of MODIS data in generating MUR field might cause cloud mask coupling with the underlying SSTs, we conducted two tests shifting the cloud masks by ± 1 day, in order to decouple MUR and MODIS. The two tests yield very similar sampling error patterns, magnitudes and statistics (not shown). There might also be differences in the sampling error if other Level

Figure 3.18 The cumulative histogram of estimated errors in MUR that are caused by unresolved SST variability underneath clouds. SE is the standard error of the estimate of the mean.
4 reference fields were to be used. We will compare the difference of sampling error referenced to different Level 4 fields in future work.

Although also affected by the inter-swath gaps, the 30°S-30°N zone shows small sampling errors primarily due to the low SST variability in this region. Here the absence of the anticipated striped pattern in the monthly averaged SST fields indicates a relatively adequate sampling of MODIS for monthly SSTs. This is also a reason for the small sampling error found in this region.

For mesoscale applications, which require SSTs at spatial resolutions of higher than 1°, our analysis is able to reveal possible sampling error impacts on the data quality and we have described possible physical and dynamical causes. An interesting question is, can we rely on an SST-cloud correlation to anticipate or even correct the sampling error? The previously reported SST-cloud relationships involve several mechanisms at various scales in the lower atmospheric boundary layer. Our analysis implies that over some regions where the vertical system is strongly coupled, for example the TIW region, sampling errors can be explicitly anticipated. However, in most cases, an SST-cloud relationship does not hold linearly or may not be manifested, because the sub-seasonal cloud variability might be more related to surface meteorological variables than to the SST. In fact, many factors in microphysics more directly modulate the cloud formation. Therefore, more analysis is needed before the SST and cloud attributions to the sampling error of many regions can be made in a quantitative fashion.

Cloud detection in IR imagery is difficult and has its own issues. As mentioned in Section 3.6, there appears to be some imperfections in the MODIS cloud masks. Although improving cloud masks is a subject beyond the sampling errors discussed here,
we believe that the further understanding of sampling errors in the high latitudes is obstructed by this additional cloud mask uncertainty. Nonetheless the type of analysis presented here can illuminate some of the possible causes of failures in the cloud screening algorithms.

Our results show warm sampling errors in the monthly averaged SSTs dominating in most regions around the globe. Without analyzing a longer period, it is unclear whether the long-term climate signal in the current IR SST climatologies is also affected. In a related study on sampling errors in climatology, Hearty et al. (2014) found that sampling biases in climatologies of air temperature and water vapor measured by AIRS can be up to 2K cold and >30% dry over mid-latitude storm tracks and deep convective regions, and >20% wet over stratus regions. They also found seasonal and diurnal variations in the sampling biases and mentioned that clouds might be the main cause for these since the bias pattern resembles the cloud distribution. Here we argue that over regions with possible SST-cloud feedbacks and longtime cloud persistence, a climate trend can even be biased by the possibly weakened data stability. Although the small SST sampling error in the tropics implies good data quality, the La-Niña-like sampling error pattern of striking negative errors in the cold tongue region potentially depresses the quality of IR SSTs for ENSO studies.

However, there is one component of the sampling error that can be quantified and separated from the overall sampling error: the sampling error caused by the known seasonal climatological cycle of SSTs. We expect for regions with long cloud persistence, the sampling error can be reduced by eliminating the climatological seasonal variation in the data.
Using microwave SSTs can reduce the sampling error caused by clouds, but there are sampling errors caused by precipitation. For regions like the GS, where precipitation tends to line up with ocean fronts (Chelton and Xie 2010; Liu et al. 2014; Minobe et al. 2010; Small et al. 2003), sampling errors can also be expected.
Chapter 4. Sampling error sensitivity

4.1 Background

Clouds and inter-swath gaps are the primary reasons for incomplete coverage of satellite infrared (IR) measurements of the Earth’s surface, and yield sampling errors in averaged IR Sea-surface Temperature (SST) fields. In the last chapter (and in Liu & Minnett. 2016) we found that the MODIS monthly SST sampling error referenced to MUR SSTs (Multi-scale Ultrahigh Resolution (Chin et al. 2010), see details in Section 4.2) is up to O(1 K), which far exceeds the error threshold needed for climate research.

Considering that the magnitudes of the quantified sampling errors are substantial in many cases, the sensitivity of such sampling errors is therefore critical to support their significance. Since the MODIS sampling error was initially calculated based on the reference MUR SST fields, whether the different SST variability embedded in a different SST reference field causes different sampling error patterns remains to be studied. The international Group for High Resolution SST (GHRSST: https://www.ghrsst.org/) was set up to help coordinate efforts to improve the accuracy of satellite-derived SST fields at all processing levels to standardize data format and facilitate the exchange of different SST fields to the research and operational communities (Donlon et al, 2007). With the growing number of Level 4 SST fields that blend observations, often including simulations, SST structure differences exist among the different data products, especially in many dynamic and rarely observed regions. SST differences among sixteen daily Level 4 fields are monitored and discrepancies are revealed in L4-SQUAM (SST Quality Monitor (Dash et al. 2012): http://www.star.nesdis.noaa.gov/sod/sst/squam/L4/). For example, compared with the GHRSST multi-product ensemble (GMPE, (Martin et al.
MUR frequently shows lower estimates in the Southeast Asian Maritime Continent region, Falkland Islands (Islas Malvinas), and the Pacific and Atlantic eastern equatorial upwelling areas, while higher estimates are found in the Northern Hemisphere high latitudes. Such non-negligible differences between L4 fields may constitute the source of uncertainty to the previously quantified sampling errors and are examined here.

In addition to the potential sampling error sensitivity on different reference fields, another important question is whether the error magnitudes and patterns change significantly in different years. Sampling errors may have interannual variability due to ocean-atmosphere interactions associated with climate events such as ENSO (El Niño–Southern Oscillation). It is recognized that the eastern equatorial Pacific TIW activity can be influenced by ENSO (Yu & Liu, 2003; An, 2008; Kug, Ham, Jin, & Kang, 2010): stronger (weaker) activity due to the increased (decreased) the eastern equatorial Pacific meridional SST gradient during La Niña (El Niño). Sampling errors found in Chapter 3 are quantified using the data of year 2011, which was during the 2010-2011 moderate La Niña event. How the negative TIW sampling errors may evolve with ENSO requires an error quantification using data from an El Niño event. This can yield an assessment of the sampling error interannual changes.

In this chapter, we compare the sampling errors using two very different reference SST fields, and explain the prevalently small error differences and the few exceptions when the sampling errors could depend on the reference field selection. As an explorative test toward sampling errors interannual variability, we quantified the errors in the El Niño year of 2009, as oppose to the previously studied La Niña year of 2010-2011. The sampling error changes on different inputs of reference SST field and on ENSO years are
studied as the sensitivity that indicates systematic uncertainty and expectable interannual variability respectively in this study.

4.2 Methods and data

This chapter continues the appreciation of the sampling error quantification framework described in Chapter 1. For the sensitivity test for TIW negative errors during ENSO events, we use the data of 20091001-20091030 to represent the El Niño TIW variations, and compare with the month of 20111001-20111030.

The HYCOM (HYbrid Coordinate Ocean Model) Global 1/12° reanalysis SSTs at 00Z reference field is added to calculate the sampling errors and compare the results with those generated from MUR. The former reference data is generated from the Navy Coupled Ocean Data Assimilation system (NCODA), which uses HYCOM model forecast as a first guess and assimilates both satellite and in situ SSTs. The latest distributed operational version GOFS3.0 (Global Ocean Forecast System 3.0) assimilates SST observations from the 5-day hindcast up to the nowcast time (Metzger et al. 2008; Metzger et al. 2010a). After the initial NCODA analysis, the first 6-hour of the 24-hour HYCOM run is incrementally updated by the NCODA analysis. Then the HYCOM hindcast cycle and NCODA analysis repeat themselves daily for 5 days until the nowcast time (Metzger and Smedstad, 2009). The assimilation system applies the 3DVAR algorithm with the first guess at appropriate time (FGAT) method, which reduces the mean analysis error that occurs when the observation-hindcast comparison is not at the same time (Cummings and Smedstad, 2013). Hourly HYCOM forecasts are employed in the FGAT to maintain the diurnal cycle in the model (Metzger et al. 2010b). The analysis SST root-mean-square-error (RMSE) assessed by comparing with temperatures
measuring from drifting buoys between 45ºS - 45ºN is reported as ~0.3K (Metzger et al. 2008). The HYCOM-NCODA SST data used here are the average temperature of the ocean top 1 meter depth layer. We refer to these temperature data as HYCOM SSTs hereafter.

On the other hand, MUR is a 1km resolution daily analysis of foundation temperatures (without diurnal warming) derived using observations from nighttime-only satellite skin and subskin SSTs and in-situ SSTs (Chin et al. 2010). Unlike other Level 4 fields such as OISSTs (Reynolds 2007), MUR does not use any climatology field to remove “outliers”, these could indicate extreme events or bad retrievals. The multi-resolution variational analysis (MRVA) is used as the interpolation method and is based on wavelet decomposition (Chin et al. 1998). The GHRSST Level 2 SST data (L2P) from multiple satellite radiometers within the -5-day and +12-hour window of the analysis time are used for the interpolation. Each L2P has its unique sampling pattern and resolution determined by the satellite instrument that took the measurements. With a 25-fold resolution difference (25 km microwave (MW) retrievals vs. 1 km IR retrievals) in the input L2P, MUR grid points have only large scale features from the MW if there were no IR observations at the location during the > 5-day time window. However, such an unresolved SST variability due to this missing IR observation issue was discussed in Chapter 3. The consequence was calculated as the standard error in the 25 km grid mean, and is found negligible. In fact, this study compares MUR and HYCOM SSTs starting from a resolution of 25km to assess the difference in the resultant sampling errors caused by any difference in the reference fields.
Since the only sources of differences in the estimates of the sampling error come from
the different input reference fields, the reference difference (Eq. 10) is studied for its
contribution to the error difference (Eq. 11). Global \( d_{\text{SST}} \) is presented in Section 4.3.

\[
d_{\text{SST}}(R) = \frac{1}{N_R} \sum_{n=1}^{N_R} S_{\text{SST}_0}^{HYCOM} - \frac{1}{N_R} \sum_{n=1}^{N_R} S_{\text{SST}_0}^{MUR} \quad (10)
\]

\[
d_{\varepsilon}(R) = \varepsilon_{HYCOM}(R) - \varepsilon_{MUR}(R) \quad (11)
\]

We are interested in not only the magnitudes of the reference SSTs, but also the field
patterns that represents the ocean features. The SST power spectrum of the two level 4
fields help indicate how well they can represent reality. The comparisons of 2D spatial
power spectra reveals the spatial scales at which the two references are consistent with
regard to spatial variability, caused by gradients, planetary waves, and eddies, while the
power spectrum of the SST time series can be used to identify the differences in the
temporal scales. Details of the reference field comparisons are presented in Section 4.3.

Further characterization of the \( d_{\varepsilon} \) sensitivity is studied by using different statistical
parameters such as Level 4 SST standard deviation, \( \sigma \), of the averaged grid cell, gap
fraction \( f \), cloud persistence \( p \), cloud spatial scale factor \( c \). The cloud spatial scale factor
is quantified as the percentage of the maximum connected cloud area, similar to the
definition for the cloud persistence in the time domain. The cloud persistence and cloud
scale factor indicate the highest temporal and spatial sampling frequency of MODIS
available for the averaged grid cell.

4.3 HYCOM reanalysis vs MUR

HYCOM and MUR SST fields have their own peculiarities. HYCOM SST
assimilated IR and MW SSTs from AVHRR, AATSR, and AMSR-E, and in situ SSTs
from ships, drifters and buoys; MUR is basically a satellite analysis of night time only IR and MW SSTs from AVHRR, MODIS, AMSR-E, and WindSat, using quality-controlled in situ SSTs from iQuam (in situ SST quality monitoring (Xu and Ignatov, 2014)) for bias corrections. The SSTs in HYCOM are the 00Z grid values representing the mean temperatures of the top 1 m of the ocean. As a result, a global pattern of diurnal warming with the afternoon maximum heating at around 150ºW could exist in this field. On the contrary, MUR represents the foundation temperature that is intended to be without diurnal effects. Therefore, we expect the 00Z HYCOM data be warmer than MUR values in many areas, especially in the Eastern Pacific Ocean.
The differences between the global monthly mean SSTs represented in the two fields are up to several degrees at high latitudes and are not negligible in many regions (Figure 4.1). In some regions, HYCOM SSTs are warmer, while others are cooler. The maximum diurnal warming of HYCOM expected at around 150°W is not evident in the SST difference. Due to the incorrect assignment of daytime SSTs in the nighttime data files, binned night-time SSTs could be contaminated by strong diurnal warming from daytime data especially in the summer time of the Arctic and around Antarctica. This might lead to MUR SSTs close to the HYCOM in those affected regions. But as shown in Figure 4.1, HYCOM is mostly colder in the high latitudes and exceptionally warmer in the Arctic summer. Since the clear-sky measurements in the Arctic and Antarctica are rather few, the warming contamination from daytime SSTs used in MUR might not cause significant difference from HYCOM in the monthly mean. Therefore, the differences in the cloudy sky SSTs between the two is expected to be substantial. Hence, there are
causes other than the diurnal effect that contribute to the large SST difference between the two reference fields.

As an example of the SST spectrum comparison result, the spatial spectrum characteristics of HYCOM and MUR SSTs of the equatorial eastern Pacific TIW region are shown in Figure 4.2. A sub-region of the predefined TIW away from shore is selected to avoid the coastline. At the first glance, the two fields do represent the main features of this cusp-shaped front: wavelength of about 1000 km and the stronger SST gradient to the north of the front (Chelton et al. 2000). However, MUR shows a relatively more coherent and smoother structure of the wave than HYCOM. Both SST spectra show one peak centered at the lowest zonal frequency, which indicates that the strongest edge/gradient in the analyzed domain is along the zonal direction. The secondary peak energy in both spectra is located at -9º in the longitude, which approximates the common wavelength of the TIW found here. By comparing the energy of the two peaks in the case, we found that HYCOM exhibits a weaker TIW wave structure, but with a stronger mean meridional SST gradient of the front. The stronger mean meridional gradient can also be seen from the wintertime mean reference SST difference (Figure 4.1), where the cold upwelling seems to be strengthened in HYCOM SSTs.

One month periods might not be sufficient to cover a complete period of the TIW. The way we study the TIW temporal representation generated by the HYCOM and MUR is to select and compare the most variable (standard deviation $\sigma > 1$ K) cells in each month. Figure 4.3 shows the cell locations with the most variability in the two reference fields are very different: HYCOM shows more temporally variable grids especially around the north-east corner of the domain, while MUR shows more mildly variable
SSTs along and to the west/offshore end of the TIW region. For those highly variable cells in either HYCOM or MUR fields, the spectrum difference primarily located at the 30-day and 15-day cycles. The largest spectrum difference is at the 30-day periodicity, where HYCOM shows ~25% lower energy. This energy component difference is primarily from where HYCOM shows the lower temporal standard deviation than MUR (Figure 4.3 blue solid line in the insert plot). In the western (further offshore) area of the TIW domain, HYCOM commonly shows less variability than MUR. This is probably caused by the less coherent structure of the TIW represented by HYCOM (Figure 4.2).

As seen from the comparisons, there are many aspects to the HYCOM-MUR SST differences, of which the TIW SST spectral differences are one. Thus, the sensitivity of the sampling error to such different fields becomes critical in determining the robustness of the method used to quantify the sampling errors.

Figure 4.3 Upper: Boreal winter month SST standard deviations and the difference between the two. Black contours indicate where either HYCOM or MUR shows 1 K SST standard deviation in the month. Lower panel: Temporal PSD and the PSD difference (expressed as the percentage of the SST variance explained by each frequency component) of the contoured area. In the HYC-MUR PSD plot, the zoom-in plot shows the averaged difference of variance.
4.4 Sampling error difference $d\varepsilon$

$$d\varepsilon(R) = -dSST(R) + \frac{1}{n_R} \sum_{n=1}^{n_R} SST_0^{HYCOM} - \frac{1}{n_R} \sum_{n=1}^{n_R} SST_0^{MUR}$$  \hspace{1cm} (12)$$

In fact, the sampling errors calculated from the two references are very similar (Figure 4.4). Geophysical patterns of the sampling errors generated from HYCOM reanalysis SSTs repeat those from MUR SSTs in both sign and magnitude and in both the
temporal and spatial averaging cases (Figure 4.4). Globally, no obvious correlated patterns can be seen from the $d\varepsilon$ and $dSST$ seasonal maps. This supports our approach to the quantification framework of sampling error estimation using a reference field as a possible realization of SSTs, regardless of the inherent errors of the reference. Sampling errors quantified are indeed primarily consequences of the missing observations due to clouds and inter-swath gaps of MODIS, and not of the choice of the reference fields.

We further compared the sampling errors generated by using the two reference fields. Sampling errors derived from HYCOM reference fields generally show smaller magnitudes than those from MUR in both spatial and temporal cases (Figure 4.5). This can be related to the assimilated HYCOM model output, which typically has low frequency variations and a memory of previous states. The spatial sampling errors derived from the two reference fields correlate better than the temporal errors. In the 4 seasonal months considered, the spatial sampling error correlation is close to 1.0. Temporal sampling errors of the summer month have the most cases of the extreme
differences. For monthly Level 3 SST field generated from MODIS, 25%-28% of the field is subject to >0.1 K uncertainty due to the temporal sampling errors, while >0.1 K spatial (5º) sampling error uncertainties can affect 4%-7% of the global SSTs (Figure 4.6). The uncertain temporal errors are mostly located in regions of large sampling errors. Uncertainties from resolutions of higher than monthly or 5º are not shown and have less contributions to monthly Level 3 SST fields.

![HYCOM-MUR sampling error difference on a monthly field](image)

Figure 4.6 HYCOM-MUR sampling error difference on a monthly field. Upper row: [25, mon] i.e. temporal averaging; lower row: [5º, 1d] i.e. spatial averaging. Differences within ±0.1 are in white and are 72%-75% and 93%-96% of the total ocean area for the temporal and spatial cases respectively. Each plot shows a boreal season denoted on the Asian Continent.

When sampling an SST signal, a linear approximation for SST variation between successive samples is feasible only if the gap is within an appropriate finite interval. When the gap between samples is larger than the dominant variation scale, the linear approximation inherent in the sampling tends to fail and may yield large sampling errors in the mean SST. According to the Nyquist sampling theorem, to adequately capture the variation of the signal the sampling frequency has to be at least twice the highest
frequency contained in the signal, otherwise sampling aliasing can occur resulting in misleading measurements and likely sampling errors. If the MODIS sampling is adequate to capture the dominant SST variation represented by both reference fields, the sampling error should be negligible. Similarly, if the MODIS sampling frequency is considerably higher than the highest frequency of the reference difference, $\delta \varepsilon$ should be negligible. In other words, if the reference difference commonly exists at the mean SST value (zero frequency), $\delta \varepsilon$ should also be negligible. This is the usual case as seen from Figure 4.5 and Figure 4.6. Therefore, in the perspective of sampling and signal frequencies, error uncertainties ($\delta \varepsilon$ not negligible) can be found when reference differences are located at high frequencies that are more than half of the sampling frequency of MODIS. As seen from the comparisons of spatial and temporal spectra of HYCOM and MUR SSTs in the dynamic TIW region (Section 4.3), the representations differ at the highest frequency of zonal spatial scales of $9^\circ$ and temporal scales of 15 days, where significant SST variation also exists. For the resolutions considered, $\delta \varepsilon$ tends to be larger in the temporal averaging as the highest frequency of the difference here often occurs at shorter than 15 days and the cloud persistence can be longer than 7 days, while in the spatial averaging $\delta \varepsilon$ is
smaller due to the highest frequency of the reference difference that can contribute to the mean (except for small eddies) is usually at a substantially larger spatial scale than the averaging grid size. Hence, $d\epsilon$ shows more significant dependence on the cloud persistence in the monthly averaging, than on the cloud spatial scale factor in the $[5^\circ,1d]$ averaging (Figure 4.7).

Another aspect that also affects the temporal $d\epsilon$ is the time-series correlation between HYCOM and MUR SSTs in each seasonal month. Since HYCOM and MUR fields yield many inconsistencies, the monthly correlation coefficients are very diverse globally. Results of the $d\epsilon$ varying with monthly correlation coefficients and cloud persistence (Figure 4.8) indicate that, even when the two signals are completely different with trivial correlations, as long as the same cloud mask is used the quantification is valid for a major portion of the sampling errors. Temporal sampling error uncertainty of $\pm 0.3$ K due to the selection of reference appears when the representation of reality by the references are contradictory and cloud persistence is up to 15 days. The summer month uncertainty is higher than the others because of the considerable overlap between large MUR sampling errors (Figure 4.8 second row) and $d\epsilon$ (Figure 4.8 third row) at the least correlation between MUR and HYCOM SSTs but the longest cloud persistence. As a result, the $d\epsilon$ and dSST are in significant negative correlations of $< -0.75$ in this overlapping domain, indicating the direct dependence of sampling errors on the selection of reference as mentioned in Eq. (12).

There are some exceptions found in the boreal summer month, although there is no obvious linearly related geographic patterns between $d\epsilon$ and dSST. We carefully examine the $d\epsilon$ and dSST correlation coefficients by applying different restrictions of many
parameters including cloud persistence, HYCOM or MUR SST, gap fraction, HYCOM or MUR standard deviation, and the difference in the reference standard deviation. The evident dependence is found in the summer month and usually in the area with cloud persistence of >22 days, and the regression slope generally increases with the cloud persistence.

Figure 4.8 The comparisons of different variables against cloud persistence (y-axis) and HYCOM-MUR correlation (x-axis) in the studied 4 seasons (columns). First row: joint histograms of cloud persistence $p$ and HYCOM-MUR correlation. Second row: the mean absolute MUR sampling errors due to temporal averaging. Third row: temporal sampling error difference magnitudes. Fourth row: the correlations between $d\varepsilon$ and $dSST$. In the 2nd and 3rd row, the two contour lines are 0.1K and 0.3K separately and denoted with cumulated percentage of the total $d\varepsilon$. 
persistence (Figure 4.9). These negative correlations explain that only 1.7% summer month sampling errors are directly affected by the selection of the reference. In other months the correlations are not as significant and consist of 0.4% (Winter), 0.1% (Spring), and 0.0% (Fall) of the total $d\varepsilon$ area. The sampling aliasing is the case when considering long cloud persistence. In the studied summer month, when the SST temporal variation is stronger (more high frequency variation) and cloud persistence is longer (lower sampling frequency) than other months, the sampling aliasing correspondingly becomes more severe such that the resulting sampling error will be more influenced by the reference field, which may represent completely different time series as seen from correlation coefficients Figure 4.9, resulting $d\varepsilon$ being more correlated with $dSST$.

Figure 4.9 Linear regressions between $d\varepsilon$ and $dSST$ of temporal averaging in the boreal summer month under different selection of cloud persistence (left). Notice that the correlated grid cells are marked with circles in the global map (right), in the corresponding color showing cloud persistence and consisting 1.7% of the total global sampling error area.

4.5 Sampling error variation with ENSO

Because the TIW sampling errors are strongly dependent on the SST gradient, primarily meridional, in the equatorial eastern Pacific, here we focus on the spatial
sampling errors in the TIW and calculate the sensitivity of the error change with the SST anomalies. MUR SSTs of October 2009 and 2011 are used for the comparisons. The

Figure 4.10 Spatial sampling errors of October 2009 and 2011, and the corresponding error and SST differences (2009-2011).

Nino3.4 index (SST anomalies in 5S-5N, 120W-170W) estimated by Climate Prediction Center is 0.94K for October 2009 and -0.97K for October 2011.

Global spatial sampling errors and MUR SST differences of October 2009 and 2011 are shown in Figure 4.10. The TIW negative errors of October 2009 is as obvious as those of 2011, but show smaller magnitudes in several grids. Gulf Stream and Kuroshio sampling errors of 2009 are colder, and can be associated with the negative 2009-2011 SST differences shown in these two regions. As shown in Figure 4.11, the meridional gradient of TIW SSTs in 2009 is indeed weaker than 2011, and this could be due to the El Niño event. The sampling errors also show smaller magnitudes in 2009, although there
are many outliers indicating the opposite. Those are the larger 2009 errors at the offshore end of the TIW. Roughly 80% weakening of the negative sampling errors can be related to the 70% decrease of the positive meridional SST gradient.

4.6 Summary

Extending the work of sampling error quantification, we have explored the sampling error sensitivity on the choice of reference field and the error variation when data from an El Niño (La Niña year data used in Chapter 3) year are used. Comparisons conducted in both time and frequency dimensions show that HYCOM and MUR represent two different SST fields. Reasons for HYCOM-MUR difference remain unclear due to the complex generation processes of the two. Despite of the marked reference SST differences, the corresponding sampling errors are very similar. Geographic patterns of previously identified large sampling errors using MUR remain in the HYCOM sampling error field. As for generating a Level 3 monthly SST field of spatial resolutions of at least

Figure 4.11 2009 (y-axis) -2011 (x-axis) TIW SST gradient and spatial sampling errors comparisons.
5°, uncertainties caused by the selection of the reference are generally within ±0.1K, in contrast to the substantial differences in the monthly SST reference fields. These results indicate the utilization of Level 4 reference SSTs on sampling error quantification is robust over most of the global oceans and the sampling error is primarily attributable to the MODIS cloud mask.

Compared with the spatial sampling errors, temporal sampling errors are subject to relatively larger uncertainty in a MODIS monthly Level 3 field. Efforts have been made to elucidate the generation for such uncertainties. The uncertainties related to the selection of reference can be linked to the cloud persistence, yet the sign of the sampling error difference ($d\varepsilon$) cannot be easily determined. We found that only in the boreal summer month when cloud persistence exceeds 20 days, the resulting sampling errors can be weakly correlated with the reference $dSST$, which indeed indicates the largest inconsistency and least positive or strong negative correlation between the two references in the boreal summer month. The differences in the reference temporal SST spectra become more important for the averaging scales of 30 days, due to the cloud persistence commonly exceeds 7 days and the HYCOM-MUR inconsistencies exist at the biweekly frequency.

Therefore, previously quantified sampling errors in regions with very long cloud persistence may have risks of being uncertain, but not necessarily so, because this can also depend on the availability of measurements from systems other than satellite IR sensors. So far Level 4 SSTs commonly use composite observations from a large variety of sources. As long as the time series generated by interpolation agrees in a large extent owing to considerable amount of observations being ingested, correlations between
different Level 4 fields cannot be significantly negative, which is the precondition for significant sampling error uncertainty.

The negative sampling error magnitudes in the Pacific TIW region indeed change with ENSO events, due to the variability of the meridional gradient becoming weaker (stronger) during El Niño (La Niña) events, although they are less directly affected because of the additional variability from clouds. Generally in the TIW region, the changes in the sampling error magnitudes are proportional to the changes in the SST gradient. Owing to the significant negative SST anomalies in the Northern hemisphere mid-latitudes, sampling errors correspondingly exhibit colder than 2011. Therefore the characterization of sampling errors in the Northern hemisphere mid-latitudes needs further study by using more observations, in order to contribute the sampling error parameterization.
Chapter 5. The parameterized sampling errors

5.1 Background

In reality, when assessing sampling errors in satellite-derived IR SST fields we do not have an appropriate reference field at the various temporal and spatial averaging intervals. However, we do have access to a number of relevant variables that can be used with the results of Chapter 3 and presented here to estimate the sampling errors. The next question to ask is, can the sampling errors be predicted, for example in terms of the local SST difference from a reference, gap fraction, cloud persistence (the number of consecutive days during which a location is detected to be cloudy), or season and region? That is, can additional, readily-available information be used to predict the sampling errors, and be used to reduce them? As the structure of the error characteristics becomes better known and the uncertainty due to the choice of the reference field barely undermine the error pattern and magnitudes, the further step is approached to parameterize such sampling errors. We study this by first investigating the importance of the climatology on estimating the error, then generating a preliminary empirical model to reproduce the sampling errors.

It is widely known that the primary component in any time series of Level 4 SST fields is the annual cycle. Therefore, a component of the sampling error that is caused by the annual cycle can be explicitly quantified by sampling a seasonal climatology, assume that the sampling errors in the IR inputs are eliminated during the climatology generation. Whether this error component is the dominant in the averaged SST fields will be studied here. We calculate the climatology component of the sampling error by using the 0.25º OISST daily climatology (Banzon et al. 2014) as the third reference field. The result of
this component can direct the parameterization of the sampling error. Later in this chapter, a preliminary empirical model is suggested by utilizing the error dependence on SST variability, gap fraction, and cloud persistence and the sensitivity to the climatology to estimate and predict the MODIS SST sampling errors. The results indicate promising usage of the model in the sampling error estimation.

Figure 5.1 Sampling errors generated using OISST climatology. Upper: temporal averaging of [0.25, mon]; lower: spatial averaging of [5, 1d]. Boreal seasons are denoted on the Asian continent.

Figure 5.2 MUR sampling errors (x-axis) explained by the OISST climatology (y-axis). Upper: temporal averaging of [0.25, mon]; lower: spatial averaging of [5, 1d].
5.2 The climatology component

Since the most pronounced temporal SST signal is the seasonal cycles, the sampling errors in temporal averaged fields might be reduced if the averaging is applied to SST anomaly fields, in which the SSTs are differences to a long term climatology, rather than the SSTs themselves. We use the OISST daily climatology (Banzon et al. 2014) to represent the seasonal signal, but first we calculate the sampling errors by superimposing MODIS masks, similarly to Eq. (3). The mid- and high- latitudes again stand out with the relatively large sampling errors (Figure 5.1). For the spatial averaging case, the sampling errors found here are similar to those using both HYCOM/ and MUR SST reference fields. In fact, > 80% of the MUR spatial sampling errors can be attributed to the climatological variation, while for the temporal sampling errors <30% is due to the seasonal cycle for monthly averages (Figure 5.2). In the [5, 1d] case, the negative

Figure 5.3 Global seasonal sampling errors quantified using MUR SSTs (red) and MUR SST anomalies (blue).
sampling errors in the TIW regions are significant across the seasons. So are the sampling
errors in the Gulf Stream, Kuroshio, Northern hemisphere mid-latitude storm track
regions, and the polar fronts of both hemispheres. The high correlations found between
sampling errors generated from OISST and MUR spatial averaging indicate that for the
averaging scales considered in this study, the spatial sampling errors can be related to the
seasonal migration of the major ocean fronts and storm tracks, and thus the

<table>
<thead>
<tr>
<th>Regression results</th>
<th>$\varepsilon_{ctim}$</th>
<th>$\varepsilon_m (\sigma_{ctim})$</th>
<th>$\varepsilon_m (\sigma_{MUR})$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>R</td>
<td>Rmse</td>
<td>Slope</td>
</tr>
<tr>
<td>Global</td>
<td>0.40</td>
<td>0.09</td>
<td>0.21</td>
</tr>
<tr>
<td></td>
<td>0.94</td>
<td>0.22</td>
<td>0.86</td>
</tr>
<tr>
<td>60ºN - 80ºN</td>
<td>0.21</td>
<td>0.07</td>
<td>0.10</td>
</tr>
<tr>
<td></td>
<td>0.94</td>
<td>0.19</td>
<td>0.88</td>
</tr>
<tr>
<td>30ºN - 60ºN</td>
<td>0.49</td>
<td>0.13</td>
<td>0.32</td>
</tr>
<tr>
<td></td>
<td>0.94</td>
<td>0.30</td>
<td>0.91</td>
</tr>
<tr>
<td>60ºS - 30ºS</td>
<td>0.36</td>
<td>0.09</td>
<td>0.16</td>
</tr>
<tr>
<td></td>
<td>0.95</td>
<td>0.25</td>
<td>0.85</td>
</tr>
<tr>
<td>80ºS - 60ºS</td>
<td>0.33</td>
<td>0.15</td>
<td>0.22</td>
</tr>
<tr>
<td></td>
<td>0.89</td>
<td>0.30</td>
<td>0.98</td>
</tr>
<tr>
<td>Topical Instability Wave</td>
<td>0.22</td>
<td>0.04</td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td>0.77</td>
<td>0.16</td>
<td>0.56</td>
</tr>
<tr>
<td>Gulf Stream</td>
<td>0.36</td>
<td>0.14</td>
<td>0.22</td>
</tr>
<tr>
<td></td>
<td>0.96</td>
<td>0.37</td>
<td>1.00</td>
</tr>
<tr>
<td>Kuroshio</td>
<td>0.63</td>
<td>0.20</td>
<td>0.42</td>
</tr>
<tr>
<td></td>
<td>0.95</td>
<td>0.30</td>
<td>0.86</td>
</tr>
<tr>
<td>Namibian Coast</td>
<td>0.45</td>
<td>0.08</td>
<td>0.28</td>
</tr>
<tr>
<td></td>
<td>0.85</td>
<td>0.16</td>
<td>0.81</td>
</tr>
<tr>
<td>Peruvian Coast</td>
<td>0.36</td>
<td>0.14</td>
<td>0.28</td>
</tr>
<tr>
<td></td>
<td>0.86</td>
<td>0.12</td>
<td>0.70</td>
</tr>
</tbody>
</table>

Table 5.1 Sampling error parameterization (Northern Hemisphere (NH) Winter). Top number of each entry shows the regression result (correlation R, RMSE, and fitting slope) from temporal averaging of [0.25º, mon] and bottom number shows the result from spatial averaging of [5º, 1d]. “N” indicates the model fails to calculate the optimal fitting.
aforementioned comparisons of using MUR and HYCOM as references show a large
degree of similarity in the spatial averaging. Therefore, the climatology component can
be a reasonable estimate for the spatial sampling errors. Indeed as shown in Figure 5.3,
the sampling errors quantified using MUR SST anomalies with respect to the OISST
climatology as the reference are smaller than when SSTs are used in terms of both mean
error and RMSE, especially in the spatial averaging that both mean error and RMSE are
substantially (>90% mean ERR and >50% RMSE) reduced after the seasonal signals are
removed. However, the RMSE and the mean error barely change magnitudes after the
removal of the climatology.

5.3 Sampling error parameterization

Even though we have shown the estimates of sampling errors are essentially
independent of the choice of the reference L4 SST field, there may be situations where a
reliable reference field might not be readily available to calculate the sampling error. An
example might be the use of assimilation of a reduced resolution SST field into
atmosphere or ocean forecast models running in real time. In such cases a simple
parametrization of the sampling error in terms of easily accessible variables could be
useful. According to the results presented in Chapter 3, sampling errors grow with the
gap fraction, cloud persistence, and the reference SST standard deviation. Combining
these with the considerable error sensitivity on the seasonal cycle, we can build a
regression model based on the previously learned error characteristics and the
climatology component. As a preliminary exploration, we assume the error function can
take the form:

$$\varepsilon_m = \alpha_0 \varepsilon_{clim} + (\alpha_1 f^{a_2} + \alpha_3 \rho^{a_4}) \sigma + \alpha_5$$

(13)
where $\alpha_0, \alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5$, are coefficients found by a Levenberg-Marquardt algorithm (Marquardt, D.W. 1963) to reach a non-linear least squares fit; $\alpha_3 = 0$ in spatial sampling error estimates, $\hat{p}$ is the normalized cloud persistence ($0 < \hat{p} < 1$), and $\sigma$ is the standard deviation of the SSTs in the averaging grid cell from MUR ($\sigma_{MUR}$). We also test the model using SST standard deviation of the climatology ($\sigma_{clim}$). When $\sigma_{clim}$ is applied, the model computes the sampling error estimates $\varepsilon_m'$ without any inputs from a Level 4 reference field and thus is predictive. Based on the prior knowledge learned from error statistics, the initial guess of coefficients to start searching for the local optimal fitting are: 1.0, 1.0, 2.0, 1.0, 2.0, -0.10, yet can be changed to any reasonable values.

<table>
<thead>
<tr>
<th>Regression results</th>
<th>$\varepsilon_{clim}$</th>
<th>$\varepsilon_m (\sigma_{clim})$</th>
<th>$\varepsilon_m (\sigma_{MUR})$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>R</td>
<td>Rmse</td>
<td>Slope</td>
</tr>
<tr>
<td>Global</td>
<td>0.44</td>
<td>0.09</td>
<td>0.26</td>
</tr>
<tr>
<td></td>
<td>0.94</td>
<td>0.21</td>
<td>0.88</td>
</tr>
<tr>
<td>60ºN - 80ºN</td>
<td>0.30</td>
<td>0.09</td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td>0.97</td>
<td>0.19</td>
<td>0.89</td>
</tr>
<tr>
<td>30ºN - 60ºN</td>
<td>0.47</td>
<td>0.13</td>
<td>0.30</td>
</tr>
<tr>
<td></td>
<td>0.96</td>
<td>0.29</td>
<td>0.90</td>
</tr>
<tr>
<td>60ºS - 30ºS</td>
<td>0.38</td>
<td>0.10</td>
<td>0.22</td>
</tr>
<tr>
<td></td>
<td>0.95</td>
<td>0.26</td>
<td>0.88</td>
</tr>
<tr>
<td>80ºS - 60ºS</td>
<td>0.16</td>
<td>0.16</td>
<td>0.16</td>
</tr>
<tr>
<td></td>
<td>0.92</td>
<td>0.26</td>
<td>1.02</td>
</tr>
<tr>
<td>Topical Instability Wave</td>
<td>0.28</td>
<td>0.13</td>
<td>0.28</td>
</tr>
<tr>
<td></td>
<td>0.77</td>
<td>0.16</td>
<td>0.77</td>
</tr>
<tr>
<td>Gulf Stream</td>
<td>0.55</td>
<td>0.13</td>
<td>0.28</td>
</tr>
<tr>
<td></td>
<td>0.97</td>
<td>0.39</td>
<td>0.93</td>
</tr>
<tr>
<td>Kuroshio</td>
<td>0.44</td>
<td>0.18</td>
<td>0.42</td>
</tr>
<tr>
<td></td>
<td>0.95</td>
<td>0.31</td>
<td>0.89</td>
</tr>
<tr>
<td>Namibian Coast</td>
<td>0.45</td>
<td>0.06</td>
<td>0.20</td>
</tr>
<tr>
<td></td>
<td>0.87</td>
<td>0.19</td>
<td>0.92</td>
</tr>
<tr>
<td>Peruvian Coast</td>
<td>0.33</td>
<td>0.10</td>
<td>0.20</td>
</tr>
<tr>
<td></td>
<td>0.93</td>
<td>0.16</td>
<td>0.73</td>
</tr>
</tbody>
</table>

Table 5.2 Similar to Table 4.1, the results of the spring month are listed.
Notice that for some regions the sampling errors are better characterized than others due to the limited prior knowledge learnt from the small portion of data, we test the model on the global sampling errors as well as errors in different regions and seasons, and focus on exploring the contribution from the additional learned error dependence on cloud and SST properties to the final estimation, i.e. seeking further improvements if only the climatology component were used. Further development of this empirical model will benefit from more statistically robust characterization of the sampling errors using more data.

5.4 Results

| Regression results | NH Summer | | | | | | | |
|--------------------|----------|----------------|----------------|----------------|----------------|----------------|
|                     | R | $\varepsilon_{c_{tim}}$ | Rmse | Slope | R | $\varepsilon_m \left( \sigma_{c_{tim}} \right)$ | Rmse | Slope | R | $\varepsilon_m \left( \sigma_{MUR} \right)$ | Rmse | Slope |
| Global | 0.29 | 0.12 | 0.11 | 0.36 | 0.08 | 0.12 | 0.44 | 0.11 | 0.21 |
| | 0.91 | 0.24 | 0.81 | 0.91 | 0.25 | 0.82 | 0.91 | 0.25 | 0.82 |
| 60°N - 80°N | 0.31 | 0.15 | 0.07 | 0.38 | 0.21 | 0.13 | 0.66 | 0.35 | 0.44 |
| | 0.86 | 0.50 | 0.78 | 0.86 | 0.47 | 0.74 | 0.87 | 0.46 | 0.76 |
| 30°N - 60°N | 0.30 | 0.26 | 0.17 | 0.32 | 0.12 | 0.10 | 0.35 | 0.13 | 0.13 |
| | 0.89 | 0.36 | 0.76 | 0.90 | 0.38 | 0.80 | 0.90 | 0.38 | 0.81 |
| 60°S - 30°S | 0.22 | 0.05 | 0.06 | 0.23 | 0.03 | 0.05 | 0.22 | 0.03 | 0.05 |
| | 0.93 | 0.26 | 0.88 | 0.93 | 0.26 | 0.87 | 0.93 | 0.26 | 0.87 |
| 80°S - 60°S | 0.10 | 0.04 | -0.02 | 0.25 | 0.03 | 0.05 | 0.45 | 0.07 | 0.20 |
| | 0.88 | 0.27 | 1.13 | 0.88 | 0.19 | 0.78 | 0.89 | 0.18 | 0.78 |
| Topical Instability Wave | 0.15 | 0.03 | 0.03 | 0.18 | 0.02 | 0.03 | N | N | N |
| | 0.90 | 0.16 | 0.77 | 0.90 | 0.16 | 0.81 | 0.90 | 0.16 | 0.82 |
| Gulf Stream | 0.24 | 0.22 | 0.22 | 0.36 | 0.07 | 0.14 | 0.33 | 0.07 | 0.13 |
| | 0.95 | 0.31 | 0.89 | 0.95 | 0.31 | 0.91 | 0.95 | 0.31 | 0.91 |
| Kuroshio | 0.29 | 0.21 | 0.18 | 0.30 | 0.08 | 0.08 | N | N | N |
| | 0.90 | 0.30 | 0.83 | 0.91 | 0.29 | 0.82 | 0.91 | 0.29 | 0.82 |
| Namibian Coast | 0.50 | 0.16 | 0.24 | 0.60 | 0.15 | 0.37 | 0.66 | 0.16 | 0.44 |
| | 0.93 | 0.21 | 0.86 | 0.93 | 0.21 | 0.87 | 0.93 | 0.21 | 0.87 |
| Peruvian Coast | 0.46 | 0.13 | 0.27 | 0.53 | 0.09 | 0.29 | 0.54 | 0.09 | 0.30 |
| | 0.92 | 0.14 | 0.73 | 0.92 | 0.17 | 0.84 | 0.92 | 0.17 | 0.84 |

Table 5.3 Similar to Table 4.1, the results of the summer month are listed.
Our results (Table. 4.1 – Table. 4.4) show that the global temporal sampling error estimation is improved, especially in the boreal or austral summer season, when the sampling error is the largest. This is because the temporal error estimates in the high latitudes summer of both hemispheres are largely improved (Figure. 5.4). Although the $\varepsilon_m$ estimates fail at negative sampling errors, for the positive sampling errors, which are primarily existed in the mid- and high- latitudes, the $\varepsilon_m$ shows overall improvements of the correlation with MUR error from 0.31 to 0.66 in the 60ºN-80ºN region, and from 0.33 to 0.54 in the 80ºS-60ºS.

<table>
<thead>
<tr>
<th>Regression results</th>
<th>$\varepsilon_{clim}$</th>
<th>$\varepsilon_m$ ($\sigma_{clim}$)</th>
<th>$\varepsilon_m$ ($\sigma_{MUR}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>R</td>
<td>Rmse</td>
<td>Slope</td>
</tr>
<tr>
<td>Global</td>
<td>0.48</td>
<td>0.12</td>
<td>0.33</td>
</tr>
<tr>
<td>60ºN - 80ºN</td>
<td>0.51</td>
<td>0.30</td>
<td>0.47</td>
</tr>
<tr>
<td>30ºN - 60ºN</td>
<td>0.55</td>
<td>0.19</td>
<td>0.52</td>
</tr>
<tr>
<td>60ºS - 30ºS</td>
<td>0.28</td>
<td>0.06</td>
<td>0.09</td>
</tr>
<tr>
<td>80ºS - 60ºS</td>
<td>0.04</td>
<td>0.06</td>
<td>0.01</td>
</tr>
<tr>
<td>Topical Instability Wave</td>
<td>0.17</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>Gulf Stream</td>
<td>0.42</td>
<td>0.22</td>
<td>0.36</td>
</tr>
<tr>
<td>Kuroshio</td>
<td>0.61</td>
<td>0.15</td>
<td>0.67</td>
</tr>
<tr>
<td>Namibian Coast</td>
<td>0.23</td>
<td>0.07</td>
<td>0.07</td>
</tr>
<tr>
<td>Peruvian Coast</td>
<td>0.57</td>
<td>0.04</td>
<td>0.11</td>
</tr>
</tbody>
</table>

Table 5.4 Similar to Table 4.1, the results of the fall month are listed.

Sampling error estimates in the Peruvian and Namibian regions are improved, especially in the in October (Figure 5.5), when the stratocumulus clouds are most developed (Klein & Hartmann, 1993; Adebiyi et al., 2015) and the negative SST-Cloud
correlation (positive sampling errors) found in Chapter 3 becomes the most pronounced.

Again, the model (Eq. 13) fails to capture the negative sampling errors. However, for those significant warm errors in these two regions that are commonly found due to the atmosphere lower boundary layer modulation, this model can generate very close estimates of the sampling errors.

Spatial sampling errors estimated by using the OISST climatology component are accurate within ±0.3K according to the RMSE. Sampling errors in spatially averaged SST anomalies are therefore largely reduced. The contribution from our model is to further decrease the RMSE of the spatial error fitting. Many regions are found with RMSE

<table>
<thead>
<tr>
<th>Regression results</th>
<th>ε_{clim}</th>
<th>ε_{m} (σ_{clim})</th>
<th>ε_{m} (σ_{MUR})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global</td>
<td>0.42</td>
<td>0.18</td>
<td>0.34</td>
</tr>
<tr>
<td>60°N - 80°N</td>
<td>0.93</td>
<td>0.31</td>
<td>0.84</td>
</tr>
<tr>
<td>30°N - 60°N</td>
<td>0.54</td>
<td>0.42</td>
<td>0.59</td>
</tr>
<tr>
<td>60°S - 30°S</td>
<td>0.93</td>
<td>0.32</td>
<td>0.87</td>
</tr>
<tr>
<td>80°S - 60°S</td>
<td>0.52</td>
<td>0.33</td>
<td>0.54</td>
</tr>
<tr>
<td>Topical Instability Wave</td>
<td>0.23</td>
<td>0.04</td>
<td>0.04</td>
</tr>
<tr>
<td>Gulf Stream</td>
<td>0.94</td>
<td>0.28</td>
<td>0.86</td>
</tr>
<tr>
<td>Kuroshio</td>
<td>0.54</td>
<td>0.30</td>
<td>0.39</td>
</tr>
<tr>
<td>Namibian Coast</td>
<td>0.96</td>
<td>0.31</td>
<td>0.85</td>
</tr>
<tr>
<td>Peruvian Coast</td>
<td>0.58</td>
<td>0.24</td>
<td>0.55</td>
</tr>
<tr>
<td></td>
<td>0.93</td>
<td>0.27</td>
<td>0.85</td>
</tr>
<tr>
<td></td>
<td>0.21</td>
<td>0.11</td>
<td>0.10</td>
</tr>
<tr>
<td></td>
<td>0.94</td>
<td>0.22</td>
<td>0.98</td>
</tr>
<tr>
<td></td>
<td>0.35</td>
<td>0.03</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>0.82</td>
<td>0.17</td>
<td>0.66</td>
</tr>
</tbody>
</table>

Table 5.5 Similar to Table 4.1, parameter verification results using data of October 2009 are shown.
decrease of a small percentage. As a matter of fact, the 80°S-60°S zone is where there is a large decrease in the RMSE: 20%, 20%, 33%, and 22% respectively in the boreal winter to fall seasons. Other regions do not display such an obvious improvement, or even occasionally show an increase in the RMSE. The largest increase of RMSE is found in the 60°N-80°N zone in winter. Therefore, without training the coefficients using more observations, the improvements made by our model for spatial sampling errors are not substantial. Although global spatial sampling errors are warm especially in the high latitudes, sometimes they can be cold, such as in the Gulf Stream and Kuroshio winter, when the direct pressure adjustment from the atmosphere become very strong (see Section 3.5.3). The TIW sampling errors are nearly always negative due to the strong vertical coupling between SST and clouds. Our model does adequately capture the strong negative errors and thus shows least improvement.

Figure 5.4 The high latitude (upper row: 60°N-80°N; lower row: 80°S-60°S) summer month error estimates using $\epsilon_{\text{clim}}$ (first column), $\epsilon_m'$ (second column), and $\epsilon_m$ (third column). R denotes the correlation coefficients; N is the total number of grids at the resolution [0.25°, mon]; RMSE and the fitting line are denoted at the bottom of each plot.
In order to assess the model and coefficients using independent data under somewhat different circumstances, we estimate the sampling errors in the 2009 October month using the proposed error function and the coefficients found for October 2011. Our results are listed in Table. 4.5. Again, the global error estimates are improved. Temporal sampling error estimates in the high latitudes, Peruvian and Namibian stratocumulus regions are improved for higher correlations and lower RMSE (except RMSE found in

Figure 5.5 Similar to Figure 5.4. The Peruvian (upper row) and Namibian (lower row) stratocumulus region sampling error estimates for October 2011 are shown.

the Peruvian region). Northern hemisphere 30º-60º latitudes show ~50% decrease in RMSE, but the correlation is not improved. This can also be seen from the statistics of Gulf Stream and Kuroshio regions, and can be attributable to the cold SST anomalies (Figure 4.10) in the two regions in 2009 leading to colder sampling errors that are not well represented by the model.
5.5 Summary

We propose an empirical model to parameterize the sampling errors using the so far learned error characteristics. The model includes the climatology as well as the sampling error dependence on information from the MODIS cloud mask and the SST variability. For sampling errors due to temporal averaging, this model can largely improve the error estimation from by only using the climatology component of the sampling error, especially in regions where warm sampling errors are prevalent, such as the polar regions and the maritime stratocumulus decks. For sampling errors due to spatial averaging, the climatological component provides estimates of accuracy within 0.3K. This result is supported by the evidence that sampling errors in SST anomalies are largely reduced after the climatological signals are removed. Adding the additional sampling error dependence to the climatology component can reduce the estimation spread in some regions but also increase in others. However, for negative sampling errors that cannot be trained in this model with positive sampling errors in the mean, the model indeed produces the wrong sign. This flaw indicates that an additional term to describe the negative sampling error causes and variations might be missing in the model, or, additional classifications of sampling errors are needed before the optimal regression can be reached.
Chapter 6. Conclusions

Motivated by the demanding accuracy of SSTs required for the SST CDR generation, in this dissertation the sampling errors of MODIS SSTs are quantified and parameterized. In Chapter 3, by using the developed sampling error quantification framework (Chapter 2), we found that the MODIS monthly sampling error, using MUR SSTs as reference fields, is up to $O(1 \text{ K})$, which far exceeds the error threshold needed for climate research and monitoring. Although the high latitudes are measured most frequently by polar-orbiting satellites, the largest sampling error ($> 5 \text{ K}$) is found in the Arctic, which is believed to be the most vulnerable to climate change (Serreze and Barry, 2011). The 30°N-30°S zonal band, exclusive of the TIW region, is found to have the smallest sampling errors. Although the error magnitudes are very diverse globally, essentially the sampling error distribution is decided by both the gap fraction and the SST variability. Our results show that the sampling error increases approximately exponentially with the gap fraction at a fixed averaging interval, while the RMSE correlates approximately linearly with the SST variability. By comparing sampling errors at 34 different spatiotemporal averaging scales, we found that the global mean error increases monotonically with the averaging scales considered here, yet the RMSE can be decreased by averaging over a more equivalent spatiotemporal scale. Based on this, we suggest that gridding pixel measurements into a resolution with relatively equivalent sub-resolution variability in the spatial and temporal domains generates lower error uncertainty. We also investigated the seasonal and diurnal changes of the sampling error. Not surprisingly, seasonal error changes are related to the seasonality of SSTs and clouds. Diurnal error changes are relatively small except for those in the high latitudes. We defined a new quantity—the cloud persistence—for error source
attribution in regions of long duration cloudiness. Imperfections in the cloud masks are seen in the unrealistic diurnal changes in the cloud persistence and can explain the apparent errors that are not attributable to either physical or dynamical causes, especially those around Antarctica. For the regions where strong ocean-atmosphere interactions exist, our results demonstrate remarkable geophysical error patterns, which can be explained by previously described mechanisms.

We compared the sampling errors generated using two very different reference SST fields—MUR and HYCOM. The results suggest that the quantification of MODIS sampling errors using the current method appears to be an appropriate approach to reflect the impacts of missing observations primarily caused by the presence of clouds. Patterns and magnitudes of the sampling errors revealed are intrinsic, regardless of substantial difference in the used reference fields. The sampling errors found here can be expected in MODIS Level 3 fields, and may affect Level 4 fields depending on how gaps in the IR SSTs are filled.

We also quantified sampling errors using OISST climatology as the reference field. Sampling errors due to spatial averaging generally can be estimated by the climatology component with accuracy of within 0.3K. This is because that the large spatial sampling errors can be related to the seasonal migration of the major ocean fronts and storm tracks, while the temporal sampling errors are primarily associated with SST intra-monthly variations.

By proposing a nonlinear empirical model built on the identified characteristics of the sampling errors, we improved the sampling error estimates especially in the high latitudes and stratocumulus regions, where warm sampling errors dominate in the studied period.
We expect that the application of this sampling error estimation model can be further explored for the benefits in Level 3 and Level 4 SST product generation and use.

This research is the first attempt to quantify and parameterize the sampling error of IR SSTs. The sampling error found is substantial and can be the primary component in the Level 3 and Level 4 SST error budget. Regional sampling errors even exhibit geophysical patterns, which indicate potential risks of misinterpretation in many applications. Therefore, we suggest that, the SST CDRs generated from IR SSTs should include the sampling error in the final error budget. Climate applications of IR SST fields should be conducted with due regard to the sampling errors.
Chapter 7. Future work

Knowing that the seasonal characteristics of IR SST sampling errors found here are derived from 4 month data in 2011, we suggest future tests being made using longer time series (with satellite data outages carefully handled) for a more sufficient characterization of the sampling error seasonal variation. This is important as the additional evidence from Microwave SST and CLW correlations show pronounced seasonal cycle in Peruvian and Namibian regions, which could imprint on the local sampling errors. Examining sampling errors in IR SST long time series can also foster additional assessments on the SST stability affected by the sampling pattern of each sensor other than systematic drift and consecutive sensor discrepancies that have been studied (Merchant et.al. 2008), and thus can benefit the generation of SST CDRs.

The improvements of sampling error estimation shown by applying our sampling error empirical model to the global sampling errors are generally small, with some exceptions, due to the large diversity of the sampling errors. For the model local modifications towards better descriptions in negative sampling errors and sampling error variations in climate events (e.g. ENSO), more characterization of the sampling error dependence and variations, and optimal classification of the sampling errors should be exploited. More exploration can also be made to the application of the parameterized SST sampling errors in data assimilation and numerical weather predictions.

As a further step towards the final error budget in SST Level 4 fields, the propagation of the sampling errors need to be studied by revisiting many Level 4 interpolation algorithms, which could either introduce more errors or reduce the final error.
Last but not least, artifacts in the current MODIS cloud masks are found in this research and can be a new problem to be studied in the future. The characterization of sampling error in the high latitudes can also be benefited from the elucidation of the cloud mask artifacts.
References


Kilpatrick, T., Schneider, N., & Qiu, B. (2013). Boundary layer convergence induced by strong winds across a midlatitude sst front. *Journal of Climate, 27*, 1698-1718


