climate change initiative

→ SEA SURFACE TEMPERATURE

# Determining Parameters for Optimal Estimation by Exploiting Matched References

Chris Merchant, Stéphane Sauxpicart and Joanne Waller

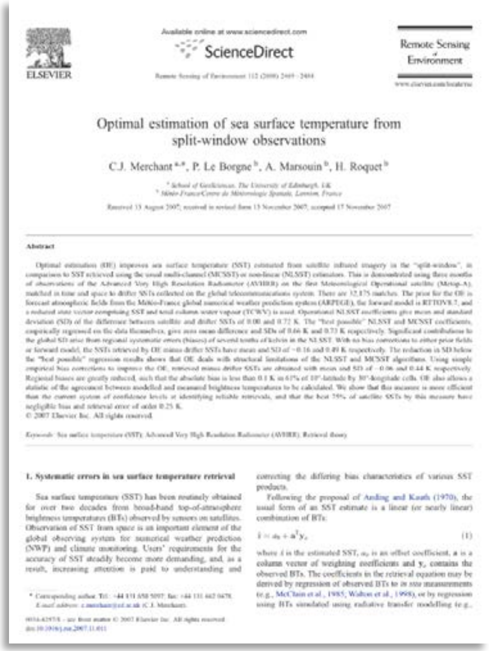
sst cci

# Introduction

- Optimal estimation is only truly optimal if underlying assumptions are met
- Typically, the assumptions are not met, and the retrieval is sub-optimal
- Talk shows how to use reference data to make the retrieval more optimal

# About Optimal Estimation (OE)

- A ‘classic’ retrieval theory, effectively applying Bayes’ theorem
- Widely applied (clouds, sea ice, aerosol, GHGs, water vapour…)
- Applied to SST since 2008
  - Experimentally at OSI-SAF
  - SST CCI for AVHRRs
- Why OE is good
  - Well-understood retrievals
  - Uncertainty and sensitivity
  - Deals with channel failures etc
  - Smooth-atmosphere low noise
- Known limitations (until now)
  - Biases
  - Covariance matrices


ELSEVIER

Available online at www.sciencedirect.com

ScienceDirect

Remote Sensing of Environment

Optimal estimation of sea surface temperature from split-window observations

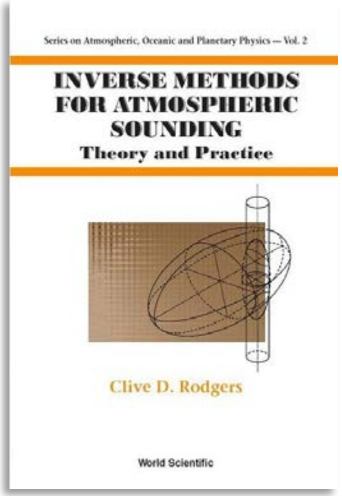
C.J. Merchant [a,*], P. Le Borgne [b], A. Marsouin [b], H. Roquet [b]

Abstract

© 2007 Elsevier Inc. All rights reserved.

1. Systematic errors in sea surface temperature retrieval

correcting the differing bias characteristics of various SST products.



# OE (MAP) formulation for SST

*a priori* knowledge

observation-simulation difference

$$\hat{\boldsymbol{z}} = \boldsymbol{z}_a + \boldsymbol{G}(\boldsymbol{y} - \boldsymbol{F})$$

retrieval

$$\boldsymbol{z}_a = \begin{bmatrix} x_a \\ w_a \end{bmatrix}$$

RTM
(here: RTTOV)

# OE (MAP) formulation for SST

error covariance of prior state

$$\hat{\boldsymbol{z}} = \boldsymbol{z}_a + \left(\boldsymbol{K}^{\mathrm{T}}\boldsymbol{S}_{\epsilon}^{-1}\boldsymbol{K} + \boldsymbol{S}_a^{-1}\right)^{-1}\boldsymbol{K}^{\mathrm{T}}\boldsymbol{S}_{\epsilon}^{-1}(\boldsymbol{y} - \boldsymbol{F})$$

$$\boldsymbol{z}_a = \begin{bmatrix} x_a \\ w_a \end{bmatrix}$$

$$\boldsymbol{K} = \left.\frac{\partial \boldsymbol{F}}{\partial \boldsymbol{z}}\right|_{\boldsymbol{z}_a}$$

error covariance of observations and simulations

RTM

# What does the covariance matrix mean?

It embodies what we know about

- uncertainty
- error correlation

$$\boldsymbol{S} = \begin{bmatrix} u_1 & 0 & 0 \\ 0 & u_2 & 0 \\ 0 & 0 & u_3 \end{bmatrix} \begin{bmatrix} 1 & r_{1,2} & r_{1,3} \\ r_{1,2} & 1 & r_{2,3} \\ r_{1,3} & r_{2,3} & 1 \end{bmatrix} \begin{bmatrix} u_1 & 0 & 0 \\ 0 & u_2 & 0 \\ 0 & 0 & u_3 \end{bmatrix}$$

# A old guesstimate for observation-simulation *S*

$$\boldsymbol{S}_{\epsilon} = \underbrace{\begin{bmatrix} 0.11^2 & 0 & 0 \\ 0 & 0.11^2 & 0 \\ 0 & 0 & 0.15^2 \end{bmatrix}}_{\text{Estimates of NEDT, origin lost}} + \underbrace{\begin{bmatrix} 0.15^2 s^2 & 0 & 0 \\ 0 & 0.15^2 s^2 & 0 \\ 0 & 0 & 0.15^2 s^2 \end{bmatrix}}_{\text{Estimates of RTTOV \textit{circa} 2003}}$$

# Optimality assumptions

- OE is only actually optimal given certain assumptions
  - All error distributions are zero mean (prior, satellite obs, RTM)
  - Two error covariance matrices (prior and S-O) are well estimated
    - at least to within a common scaling factor

- But in general
  - Satellite calibration and RTM are biased
  - Prior is biased
  - Error covariance matrices are informed guesses

- How to make the assumptions less untrue?

# Determine OE parameters

## Parameters:

$\boldsymbol{\beta}$ = obs. corr.

$\boldsymbol{\gamma}$ = prior corr.

$$\boldsymbol{S}_a$$

$$\boldsymbol{S}_\epsilon$$

# Determine OE parameters

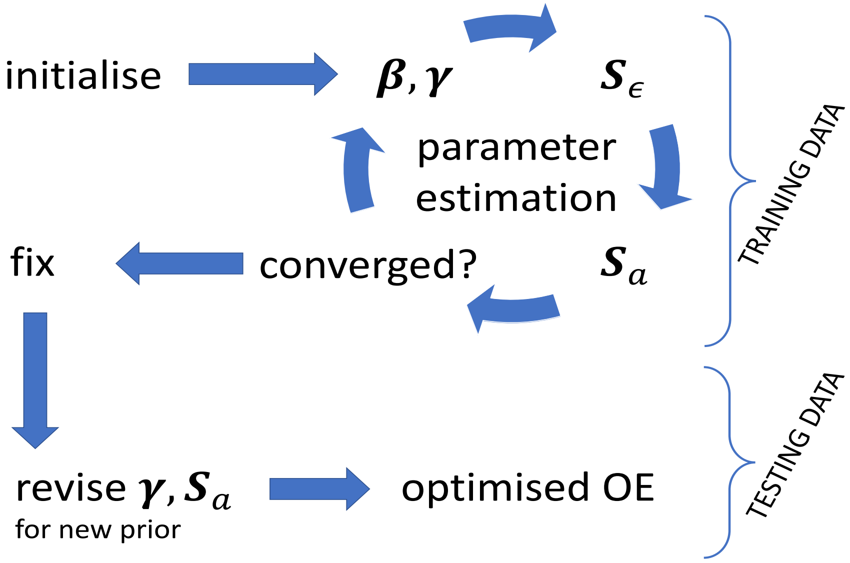
## Parameters:

$\boldsymbol{\beta}$ = obs. corr.

$\boldsymbol{\gamma}$ = prior corr.

$\boldsymbol{S}_a$

$\boldsymbol{S}_\epsilon$




Anchored by matched reference data

Doesn’t use references

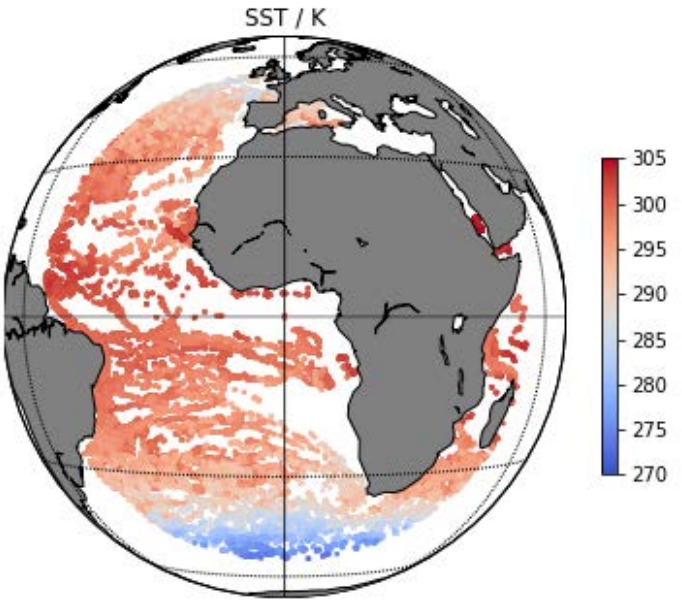
# Data used

- Matches between SEVIRI and drifting buoy observations
- Retrieve SST from 8.7, 10.8 and 12.0 um

- 2011 for training
  - RTTOV 11.2 simulations using buoy SST – 0.17 K

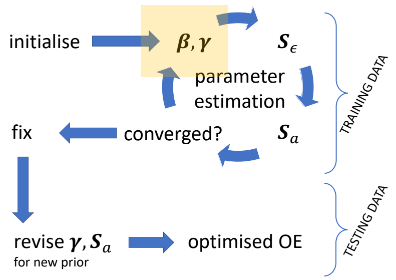
- 2012 for application
  - RTTOV 11.2 simulations using SST climatology

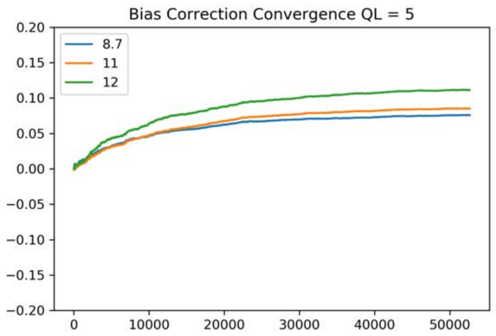
- About 90,000 matches in each year
- GHRSST QL 5 considered



# Estimate biases

- Want parameters for corrections of
  - Each channel BT (observation relative to simulation)
  - Prior total column water vapour
- Assume drifting buoy + skin effect provides unbiased SST
  - With non-zero uncertainty
  - Use this SST for RTTOV simulations
- Iterating over many matches, do a simultaneous retrieval
  - the state estimate (SST and TCWV)
  - and these parameters as well
- Variant of Kalman filtering for state and model parameter estimation



# Estimate observation error covariance

Use a “Desroziers equation”, increasingly used in data assimilation

- Do new retrievals using

$$\beta_{8.7} \quad \beta_{10.8} \quad \beta_{12.0} \quad \gamma_{w_a}$$

- From the results calculate
- $\boldsymbol{d}_r^o = \boldsymbol{y} - \boldsymbol{F}'(\hat{\boldsymbol{z}}) - \langle \boldsymbol{y} - \boldsymbol{F}'(\hat{\boldsymbol{z}}) \rangle$
- $\boldsymbol{d}_a^o = \boldsymbol{y} - \boldsymbol{F}'(\boldsymbol{z}_a) - \langle \boldsymbol{y} - \boldsymbol{F}'(\boldsymbol{z}_a) \rangle$

and use these for better estimate for the observation error covariance matrix is

*Q. J. R. Meteorol. Soc.* (2005), **131**, pp. 3385–3396 doi: 10.1256/qj.05.108

## Diagnosis of observation, background and analysis-error statistics in observation space

By G. DESROZIERS†, L. BERRE, B. CHAPNIK and P. POLI
*Météo-France, Centre National de Recherches Météorologiques, Toulouse, France*

(Received 31 May 2005; revised 31 October 2005)

SUMMARY

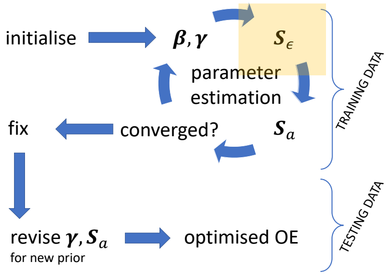
Most operational assimilation schemes rely on linear estimation theory. Under this assumption, it is shown how simple consistency diagnostics can be obtained for the covariances of observation, background and estimation errors in observation space. Those diagnostics are shown to be nearly cost-free since they only combine quantities available after the analysis, i.e. observed values and their background and analysis counterparts in observation space. A first application of such diagnostics is presented on analyses provided by the French 4D-Var assimilation. A procedure to refine background and observation-error variances is also proposed and tested in a simple toy analysis problem. The possibility to diagnose cross-correlations between observation errors is also investigated in this same simple framework. A spectral interpretation of the diagnosed covariances is finally presented, which allows us to highlight the role of the scale separation between background and observation errors.

KEYWORDS: Estimation theory Optimality criterion Parameter estimation

### 1. INTRODUCTION

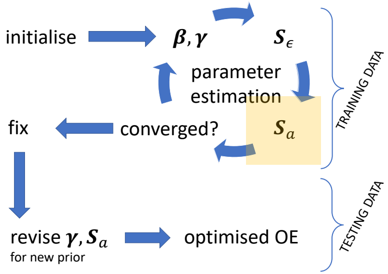
Most main operational assimilation systems are now based on the variational formalism (Lewis and Derber 1985; Courtier and Talagrand 1987, Rabier *et al.* 2000). Such a formalism allows the use of a large spectrum of observations and in particular satellite data that are not directly and linearly linked with model variables. However, those variational algorithms still rely on the theory of least-variance linear statistical estimation (Talagrand 1997). In the linear estimation theory, each set of information is given a weight proportional to the inverse of its specified error covariance. The pieces of information are classically given by observations and a background estimate of the state of the atmospheric flow. Analysis systems are then dependent on appropriate statistics for observation and background errors. Unfortunately those statistics are not perfectly known and their determination remains a major challenge in assimilation systems. One source of information on the observation and background errors is contained in the statistics of the innovations, that is the differences between observations and their background counterparts. Those statistics have for example been used by Hollingsworth and Lönnberg (1986), assuming that background errors carry cross-correlations while observation errors do not. From a slightly different point of view, Dee and da Silva (1999) have used a maximum likelihood method to estimate the information error statistics. Desroziers and Ivanov (2001) have proposed an approach based on a consistency criterion of the analysis relying on statistics of *observation-minus-analysis* differences to adapt observation-error statistics. The consistency criterion used in this method was defined by Talagrand (1999). Chapnik *et al.* (2004) investigated the properties of the algorithm and especially showed that it was equivalent to a maximum likelihood method, though less expensive to implement. Chapnik *et al.* (2006) also applied the same algorithm in an operational framework to tune observation-error variances.

This paper presents a set of diagnostics based on combinations of *observation-minus-background* (O–B), *observation-minus-analysis* (O–A) and *background-minus-analysis* (B–A) differences, which provide an additional consistency check of an analysis scheme.



# Estimate prior error covariance

Use an expression derived from a "Desroziers equation"

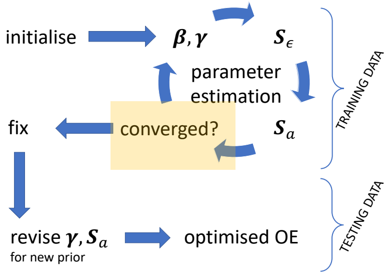



- Do new retrievals using $\beta_{8.7}$ $\beta_{10.8}$ $\beta_{12.0}$ $\gamma_{w_a}$ and $\widehat{\boldsymbol{S}}_\epsilon$
- From the results calculate
- $\boldsymbol{d}_a^r = \boldsymbol{F}'(\hat{\mathbf{z}}) - \boldsymbol{F}'(\mathbf{z}_a) - \langle \boldsymbol{F}'(\hat{\mathbf{z}}) - \boldsymbol{F}'(\mathbf{z}_a) \rangle$
- $\boldsymbol{d}_a^o = \boldsymbol{y} - \boldsymbol{F}'(\boldsymbol{z}_a) - \langle \boldsymbol{y} - \boldsymbol{F}'(\boldsymbol{z}_a) \rangle$

- and use these to get a better estimate for the prior error covariance matrix is

# Test for convergence

Use yet another adapted “Desroziers equation”




- Do new retrievals using $\beta_{8.7}$ $\beta_{10.8}$ $\beta_{12.0}$ $\gamma_{w_a}$, $\widehat{\boldsymbol{S}}_\epsilon$ and $\widehat{\boldsymbol{S}}_a$
- From the results calculate
- $\boldsymbol{d}_a^o = \boldsymbol{y} - \boldsymbol{F}'(\boldsymbol{z}_a) - \langle \boldsymbol{y} - \boldsymbol{F}'(\boldsymbol{z}_a) \rangle$

Once the whole set of parameters are converged, should have

- $\langle \widehat{\boldsymbol{S}}_\epsilon + \boldsymbol{K}\widehat{\boldsymbol{S}}_a\boldsymbol{K}^{\mathrm{T}} \rangle^{-1} \langle \boldsymbol{d}_a^o \boldsymbol{d}_a^{o\,\mathrm{T}} \rangle - \boldsymbol{I} \sim \boldsymbol{0}$

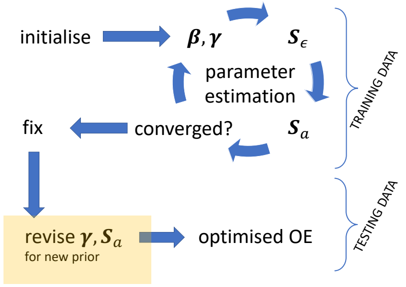
Measure convergence as element-wise sum of squares of LHS

Also check standard deviation of change in retrieved SST change

# New prior biases and uncertainty

- Estimate the correction for
  - Prior SST to be used in the application of the OE
  - SST CCI climatology
    - 1982 – 2010, cold bias relative to the year 2012
  - Correction in latitudinal bands
- Estimate the prior SST uncertainty

- Do the above using Kalman-filtering-like method as before



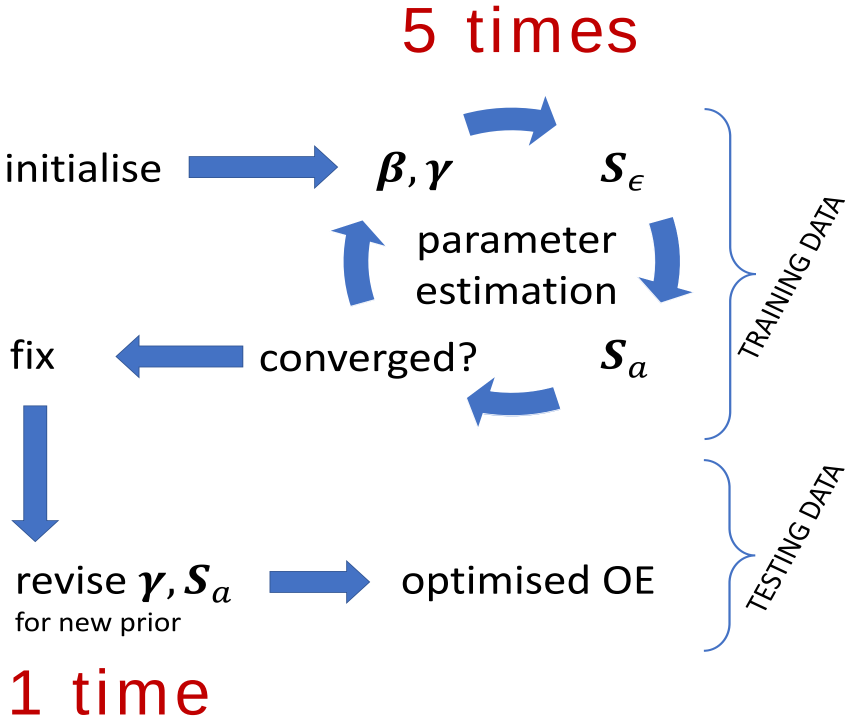
# Parameters:

$\boldsymbol{\beta}$ = obs. corr.

$\boldsymbol{\gamma}$ = prior corr.

$\boldsymbol{S}_a$

$\boldsymbol{S}_\epsilon$




Anchored by matched reference data

Doesn't use references

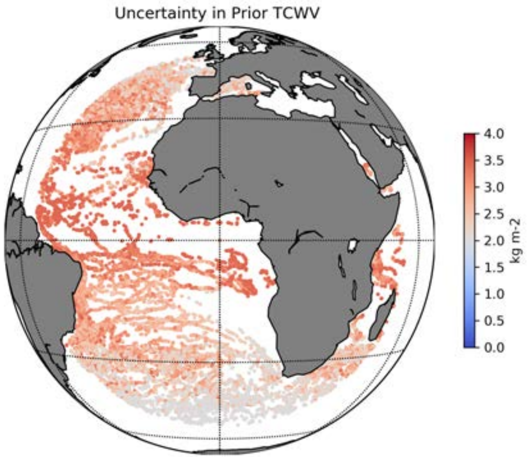
# Validation against drifting buoys using the 'test' matches

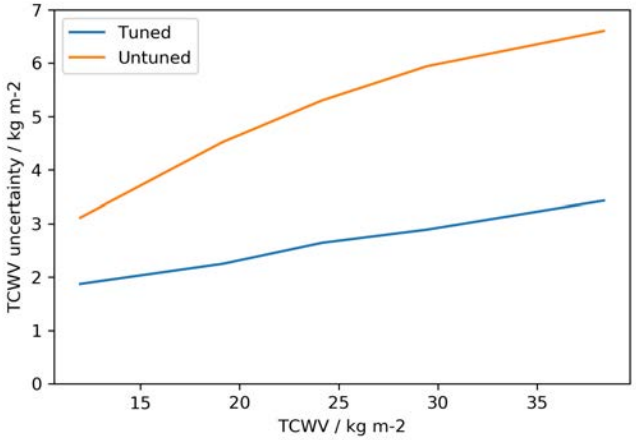
| Retrieval | Mean diff. / K | SD. diff. / K | RSD. diff. / K | $\partial\hat{x}/\partial x_{true}$ | $SD\left(\frac{\hat{x}-x_b}{\sqrt{S_{\hat{x}}+u_{x_b}^2}}\right)$ |
|---|---|---|---|---|---|
| Clim. (prior) | -0.19 | 0.79 | 0.74 | 0% | - |
| Operational | 0.02 | 0.45 | 0.39 | ? | - |
| Untuned OE | -0.05 | 0.44 | 0.38 | 71% | 0.77 |
| Tuned OE | 0.00 | 0.43 | 0.37 | 79% | 1.04 |

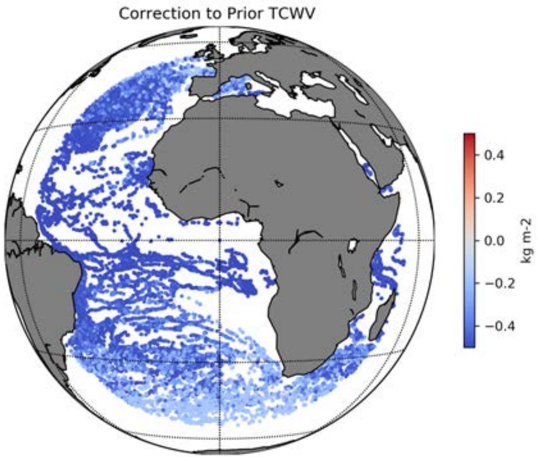
Retrieval sensitivity – ideally 100%

Scatter / uncertainty estimate – ideally 1

# Insight into ECMWF NWP TCWV

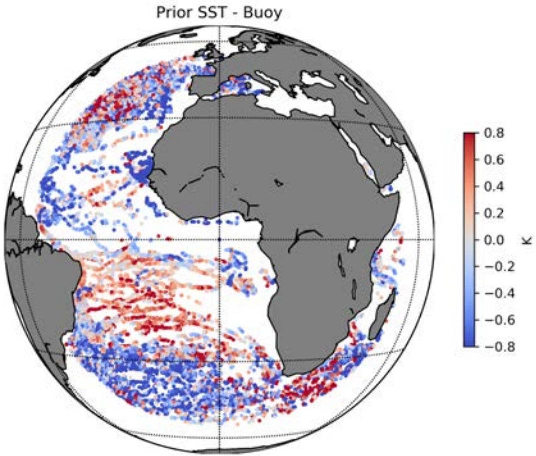


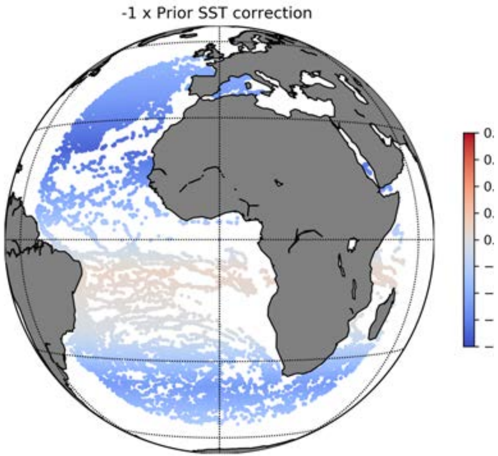



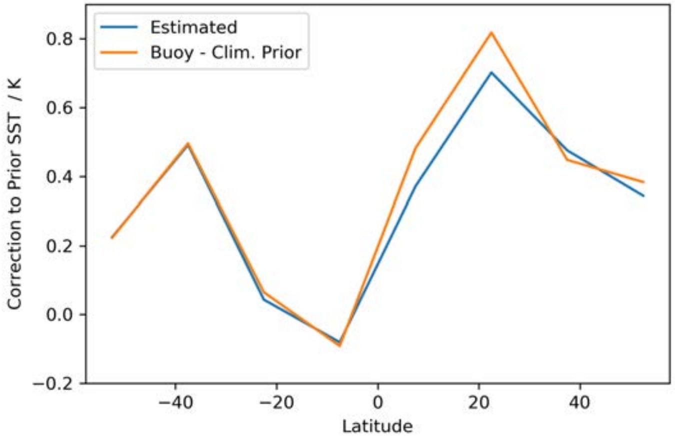



# Good estimation of prior SST parameters

- Estimate of prior SST uncertainty from difference to drifters: 0.76 K
- Estimate of same from Desroziers method (independent of drifters): 0.71 K
- Also obtained estimates for prior SST biases/corrections

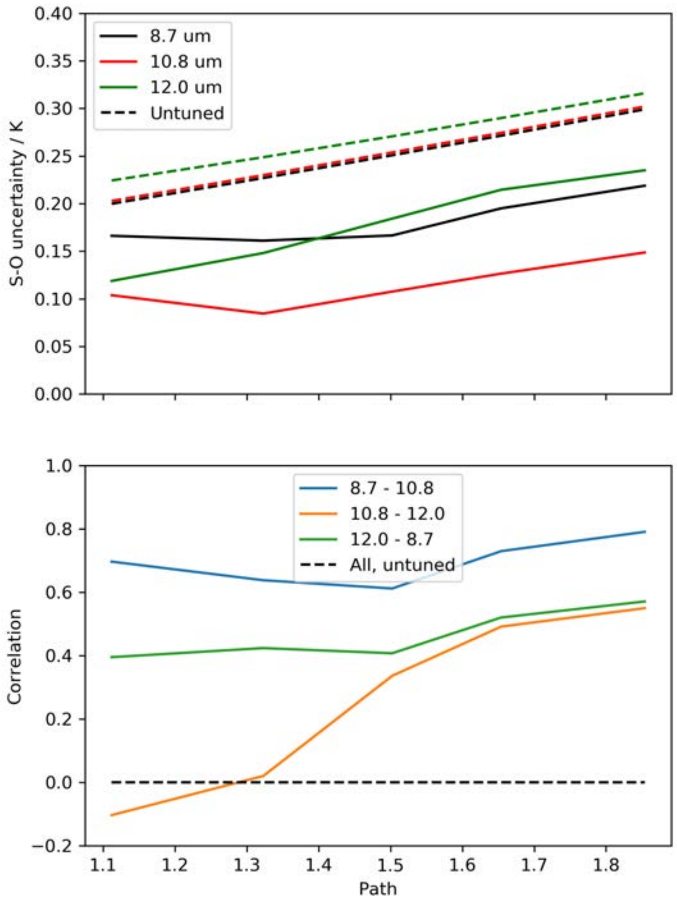


# Insights into noise & RTTOV errors

- Observation error covariance matrix parameterized in terms of path ($\sec\theta$)
- Given that NEDT is independent of path, infer that RTTOV simulation uncertainty varies with
  - path (greater for longer paths)
  - TCWV (near-nadir is also highest TCWV)
- RTTOV 8.7 simulations not as good as other channels
  - Note contrast in nadir results cf. 11 & 12
- RTTOV error correlations are high near the limb
  - >=0.5 for all channel pairs
  - Important for retrieval!



# Review

- Bias parameters and error covariance parameters are needed for OE
- Matches to in situ reference data can inform estimates for these parameters
  - Iterative "retrieval" of biases and use of "Desroziers" covariance estimates
- The "tuned" OE improved the bias, SD, RSD, sensitivity and SST uncertainty
- Solid estimates of bias and uncertainty of NWP total column water vapour
  - Uncertainty increases with TCWV
  - It is half what we previously assumed
- Can obtain an estimate of prior SST biases even where there are no in situ
- RTTOV simulation uncertainty for SST channels
  - varies with channel and TCWV
  - is $<0.1$ K at nadir and ~0.15 K at ~60º (8.7 channel is not this good)
- RTTOV simulation errors become strongly correlated ($r \geq 0.6$) high zenith angle

# Further & *Possible* work

- In SST CCI
  - Slave Metop A AVHRR to AATSR & SLSTR (harmonization)
  - OE for 1980s AVHRRs to drifters (in situ patchy)
  - Switch to ERA-5 (needs a new prior error covariance)
  - Better constrain Bayesian cloud detection
- *Address whole-SST constellation consistency/biases systematically*
  - *Exploit metrological in situ (radiometers and new buoys)*
  - *Inter-satellite bias correction (Melbourne GHRSST)*
    - *E.g., skin SST of SLSTR as unifying reference*
    - *Or all-pairs approach (consistent parameters for all sensors)*
    - *Note: correct both SST bias and BT bias*
- *SLSTR A & B nadir-view two-channel wide-swath SST by OE*
- *Extend to a more comprehensive state vector*
  - *known that {SST, TCWV} is limiting*
  - *really want at least {SST, WV_EOF1, WV_EOF2, T_EOF1, DD}*
  - *but we had no method to estimate $S_a$ for such a state vector*
- *Should be applicable for OE beyond SST*

*Q. J. R. Meteorol. Soc.* (2006), **132**, pp. 1205–1223 doi: 10.1256/qj.05.143

## Retrievals of sea surface temperature from infrared imagery: origin and form of systematic errors

By C. J. MERCHANT[1]*, L. A. HORROCKS[2], J. R. EYRE[3] and A. G. O'CARROLL[3]
[1] *School of GeoSciences, The University of Edinburgh, UK*
[2] *Department of Environment Food and Rural Affairs, London, UK*
[3] *Met Office, Exeter, UK*

(Received 8 July 2005, revised 17 November 2005)

SUMMARY

We show that retrievals of sea surface temperature from satellite infrared imagery are prone to two forms of systematic error: prior error (familiar from the theory of atmospheric sounding) and error arising from nonlinearity. These errors have different complex geographical variations, related to the differing geographical distributions of the main geophysical variables that determine clear-sky brightness-temperatures over the oceans. We show that such errors arise as an intrinsic consequence of the form of the retrieval (rather than as a consequence of sub-optimally specified retrieval coefficients, as is often assumed) and that the pattern of observed errors can be simulated in detail using radiative-transfer modelling. The prior error has the linear form familiar from atmospheric sounding. A quadratic equation for nonlinearity error is derived, and it is verified that the nonlinearity error exhibits predominantly quadratic behaviour in this case.

KEYWORDS: Advanced Very High Resolution Radiometer (AVHRR) Along-track Scanning Radiometer (ATSR)

### 1. SCOPE OF PAPER

Sea surface temperature (SST) has been routinely obtained for over two decades from broad-band infra-red radiances observed by sensors on satellites. Generally, following the suggestion of McMillin (1975), radiances are expressed as brightness temperatures (BTs) and the SST estimate is a linear (or nearly linear) combination of these:

$$\hat{x} = a_0 + \mathbf{a}^{\mathrm{T}}\mathbf{y}^{o}, \quad (1)$$

where $\hat{x}$ is the estimated SST, $a_0$ is an offset coefficient, $\mathbf{a}$ is a column vector of weighting coefficients and $\mathbf{y}^o$ contains the observed BTs.

The coefficients in the retrieval equation may be derived by regression of observed BTs to *in situ* measurements (e.g., McClain *et al.* 1985; Walton *et al.* 1998), or by regression using BTs simulated using radiative-transfer modelling (e.g., Llewellyn-Jones *et al.* 1984; Závody *et al.* 1995). In nearly linear formulations, the coefficients are weak functions of a prior or first-guess SST. This retrieval method is convenient and simple to implement.

Linear retrieval using regression-based coefficients also appears to be conceptually straightforward. The purpose of this paper is to show that this appearance is deceptive. Linear retrieval leads to systematic errors in SST retrievals that have complex spatial and temporal characteristics. In the particular case illustrated in this paper, these errors are generally smaller than 0.3 K. While such errors may in the past have been neglected, our context is determining SST for climate applications: systematic errors of this magnitude are significant, and need to be thoroughly understood. Such systematic errors are also unlikely to remain acceptable for numerical weather prediction in the future.

We proceed as follows. Firstly, we describe empirically the nature of the systematic errors found in a particular satellite SST data set, both in observation and in simulation. A global mean bias is found in the observations, the magnitude of which is readily

* Corresponding author: School of GeoSciences, The University of Edinburgh, UK. e-mail: c.merchant@ed.ac.uk