FIDUCEO : Principles for Satellite retrieval

Jonathan Mittaz
NPL/University of Reading
The FIDUCEO project

• **Fidelity and Uncertainty in Climate data records from Earth Observation**

• **Ambition**: develop a widely applicable metrology of Earth Observation (EO)

• **Motivation**: establish defensible, uncertainty-quantified evidence (CDRs) for climate and environmental change from space assets

• **Limitation of the status quo**: the FCDR uncertainty is not characterised, and cannot be propagated to the CDR
### FIDUCEO FCDRs (L1)

FCDR: fundamental climate data record (calibrated radiances) from which climate data can be derived

<table>
<thead>
<tr>
<th>DATASET</th>
<th>NATURE</th>
<th>POSSIBLE USES</th>
</tr>
</thead>
<tbody>
<tr>
<td>AVHRR FCDR</td>
<td>Harmonised infra-red radiances and best available reflectance radiances, 1982 - 2016</td>
<td>SST, LSWT, aerosol, LST, phenology, cloud properties, surface reflectance ...</td>
</tr>
<tr>
<td>HIRS FCDR</td>
<td>Harmonised infra-red radiances, 1982 - 2016</td>
<td>Atmospheric humidity, NWP re-analysis, stratospheric aerosol ...</td>
</tr>
<tr>
<td>MW Sounder FCDR</td>
<td>Harmonised microwave BTs for AMSU-B and equivalent channels, 1992 – 2016</td>
<td>Atmospheric humidity, NWP re-analysis ...</td>
</tr>
<tr>
<td>Meteosat VIS FCDR</td>
<td>Improved visible spectral response functions and radiances 1982 to 2016</td>
<td>Albedo, aerosol, NWP re-analysis, cloud, wind motion vectors,...</td>
</tr>
</tbody>
</table>
FIDUCEO CDRs (L2/L3)

CDR: climate data record, the evidence base for high-level climate information and services

<table>
<thead>
<tr>
<th>DATASET</th>
<th>NATURE</th>
<th>USE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surface Temperature CDRs</td>
<td>Ensemble SST and lake surface water temperature</td>
<td>Most of climate science ... model evaluation, re-analysis, derived/synthesis products ..</td>
</tr>
<tr>
<td>UTH CDR</td>
<td>From HIRS and MW, 1992 - 2016</td>
<td>Sensitive climate change metric, re-analysis ...</td>
</tr>
<tr>
<td>Albedo and aerosol CDRs</td>
<td>From M5 – 7 (1995 – 2006)</td>
<td>Climate forcing and change, health ...</td>
</tr>
<tr>
<td>Aerosol CDR</td>
<td>2002-2012 aerosol for Europe and Africa from AVHRR</td>
<td>Climate forcing and change, health ...</td>
</tr>
</tbody>
</table>
The FIDUCEO Approach

• At Level 1 we start with the “Traceability Tree”
  – Starts with the measurement equation

\[ R_E = a_0 + \frac{a_1 R_T - a_2 \dot{C}_T^2}{\dot{C}_T^2} C_E + a_2 C_E^2 + f(T_{Instr}) + O \]

  – Looks at each term and breaks it down into however many underlying processes are needed to get back to root process

  – Links lowest level processes to their impact and associated uncertainty on the observed Earth radiance via sensitivity coefficients

\[ u_c^2(y) = \sum_{i=1}^{n} \left( \frac{\partial f}{\partial x_i} \right)^2 u^2(x_i) + 2 \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \frac{\partial f}{\partial x_i} \frac{\partial f}{\partial x_j} u(x_i, x_j), \]
\[ a = f(R_{\text{ref}}, C_i, T_i, K_i) + O \]

\[ R_T = B(T_i) \cdot d + O \]

\[ R_E = a_0 + \frac{a_1 R_T}{C_T} - \frac{a_2 C_T^2}{C_E} + \frac{a_2 C_T^2}{C_E} + f(T_{\text{Instr}}) + O \]
The $+0$ term

- Appears in a number of places
- Intended to force investigation of extra assumptions
  - To think about and characterize known unknowns
  - AVHRR L1 examples
    - Quadratic assumption for non-linearity effect
    - Constant non-linear coefficient
    - Numerical issues (digitisation/numerical integration)
    - etc.
- Will be much more important for geophysical retrievals (Level 2+) then Level 1
AVHRR Effect Tables

• How FIDUCEO codifies different uncertainty components
  – Uncertainties caused by random effects
  – Uncertainties caused by systematic effects
  – Uncertainties due to correlated errors
    • Include correlation length scale/shape
    • Covariance for channel-to-channel case
      – Strong correlations seen in TIROS-N AVHRR

• Effects tables covers
  – Effect size, correlation type and scale, covariance information and sensitivity coefficient
Impact of time dependent biases for SST retrieval

Used Pathfinder retrieval algorithm form and estimated error in SST from instrument dependent BT errors and compared with Pathfinder V6 errors

Related to instrument temperature

Overlay matches many of the features (after adding in a shift of 0.1K).

Sensor-to-Sensor Harmonisation

- For FCDR crucial to get consistent calibration taking into account real sensor differences – we call this Harmonisation

  - Measurement equation fitted to sensor-to-sensor matchup data
  - Full uncertainty information including correlated error terms
  - Fit process using Error-in-Variable (uncertainty in both ‘X’ and ‘Y’)
    - OLS will give biased result (e.g. backup slide)
FIDUCEO at Level 2 (CDR)

• SST CDR based on ESA CCI SST methodology
  – Use FIDUCEO Level 1 AVHRR data with full uncertainty information
  – Will provide an Ensemble
  – Exact Level 2 framework still being worked on

• Stability of Level 1 data crucial (within uncertainties)
  – Need good Harmonisation

• Note depending on required timescales/spatial scales (e.g. climate related scales) different parts of uncertainty budget will become important
  – Require full uncertainty budget at Level 2 including modelling uncertainties
    • Skin-to-depth
    • Input model data
    • etc
Breakdown of (part of) the ESA CCI SST Algorithm (a start at least)

- Taken from simple OE processing code at UoR (GBCS)
- (not showing smoothed version of OE actually used)

\[ \hat{x} = x_a + S^{-1} K^T (K S^{-1} K^T + S_a^{-1})^{-1} (y - F(x_a)) + O \]

• Likely different set of correlated errors compared to Level 1
Start some early studies via Simulations

'TRUTH'

Reynolds SST

ECMWF ERA-Interim Profiles

RTTOV

GBCS Emissivity Model

AVHRR Instrument Model

'Available'

NOAA GFS SST

CSTM

Emissivity Model

NCEP Profiles

Interpolate to position

SST Algorithm

Simulated BT/Tangent linears

Retrieved SST

1000 BTs
Simulator

‘TRUTH’
- Reynolds SST
  - GBCS Emissivity Model
  - RTTOV
  - ECMWF ERA-Interim Profiles
  - AVHRR Instrument Model
  - 1000 BTs

‘Available’
- NOAA GFS SST
  - CRTM
  - Emissivity Model
  - Simulate BT/Tangent linears
  - Interpolate to position
  - Retrieved SST

SST Algorithm

Available SST

Observed BTs...
‘Truth’ Simulations

- Uses real AVHRR orbit data for locations/angles/timings
  - Only over clear sky/ocean data
- Positions – AVHRR tie points as defined by ESA CCI SST processing
  - 137 points across track
  - Of order 684 along track
  - Note no need for ‘BT Harmonisation’ in this case as SRF errors not included (yet)

SatZA distribution of simulated data
SST retrievals

• Standard OE with no bias correction
  – ‘Buoy’ data statistics (note skin-to-skin)
  – No Aerosol loading/Perfect cloud detection

Bias = -0.007757
Stdev = 0.265158
SST retrieval individual PDFs - digitisation

• For early designed sensors digitisation is important...

• Note – not a ‘single Gaussian’ in sight!
• More modern sensors won’t have this problem...
Regional biases

- Map of residuals
  - Note that this is basically model-to-model difference

Daytime

Nighttime
Statistics of PDFs

• Can then look at statistics (mean/stdev) of PDFs

Bias term related to more systematic effects and is much larger than random effect case

Standard deviation related mostly to random uncertainties and is small < 0.1K

Understanding the all terms in the retrieval is crucial including model data
How can FRM4STS help FIDUCEO?

• Concept behind FIDUCEO methodology is to understand the individual sources of error and uncertainty
  – We don’t simply ‘bias correct’
• Use validation data to make sure of uncertainty model
  – Accurate in-situ SST data would then be used as a source of information as part of validation
  – Can be used to check that uncertainty modelling hasn’t missed anything
    • Key will be uncertainty model from in-situ from FRM4STS
    • If uncertainties not consistent then would investigate current model and modify if necessary
• Eventually through Harmonisation processes (sensor-to-sensor) can lead to links to SI
  – Ideally make sure the SI traceable values cover complete range of all uncertainty effects including time evolution effects of errors/uncertainty of input data
Conclusions

• FIDUCEO is applying metrological processes to in-orbit satellite data
  – Currently finishing Level 1 analysis
    • Will be evaluating SLSTR as well
  – Moving onto Level 2 (CDR) analysis

• By breaking down problem into original sources of possible error
  – Traceability
  – Understanding of importance of different sources of error/uncertainty
  – Use validation against accurate reference to ensure uncertainty model contains all dominant sources