Uncertainty & Precision: Conversations about data set reprocessing and utility of the data we produce

B Guenther and Cole Rossiter
JPSS Algorithm Management Project
Stellar Solutions

CALCON, SDL, Logan, Ut 19 June 2018

www.calcon.sdl.usu.edu

Outline

- Role of Uncertainty in Metrology
- NASA HQ Guidance
- Examples of Uncertainty and Reprocessing
- New Horizons for Meteorology

Role of Uncertainty in Metrology

Metrology

- NIST Technical Note 1297
  - Measurement result complete only when accompanied by a quantitative statement of its uncertainty.
    - Type A: are evaluated by statistical methods
    - Type B: are evaluated by other means
  - Error and accuracy
  - Uncertainty
  - Recommends correction for significant, systematic measurement errors

From Note 1297

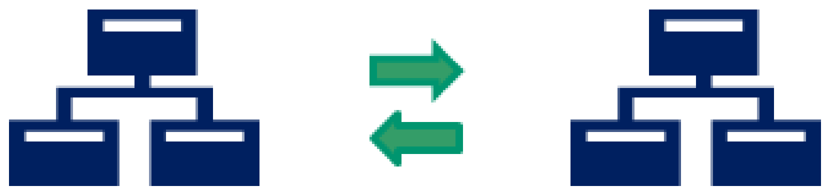
- Section 5.2 - apply a correction (or correction factor) to compensate for each recognized significant systematic effect
- Every effort has been made to identify such effects

NASA HQ Guidance

Program Office Guidance

- I got this from NASA HQ, but expect identical advice for NOAA, Interior, and others
  - Information actually predates Earth Observing System development
- Mission defines the quantities to be measured
  - Mission defines the accuracy/uncertainty of the measurement to meet mission objectives
- Metrologists will agree but sense of guidance significantly different for uncertainty of 1 part in 100 compared to 1 part in 500,000
  - Guidance on significant systematic effects

From **NESDIS' Budget & Organizational alignment**



Greater Ability to Leverage Commercial Partnerships

- Agility to strengthen partnerships with commercial entities and provide more innovative solutions to critical data sets and products

Examples of Uncertainty and Reprocessing

Relationship of Uncertainty to Reprocessing

- Reprocessing is only considered when datasets have discontinuities, changes in physics or processing codes, and improved performance or consistency
- In other words, improved uncertainty for the comprehensive product is only motivation for reprocessing
  - Includes characterization scale synchronization within dataset from multiple sensors

MODIS Uncertainty

- Uncertainty data on a *pixel-by-pixel basis*. These values (called "uncertainty indexes") may be converted to percent uncertainties associated with each value in the corresponding scaled integer SDS (MODIS L1B User guide)
  - These are 4-bit values scaled by ratio ln(estimated uncertainty/specification uncertainty)
- Product provided as 16-bit unsigned integer, and a value > 2^{15} represents failed calibration for that pixel

AIRS Uncertainty

- QC provided each FOV called “state” and applies to all channels
- QC provided each (2378) channels each scan line for calibration quality
- Estimated Noise equivalent Delta Temperature based on 250K scene, each channel, each granule
- Glint latitude and longitude each scan

from AIRS Level 1, V5 Users Guide

VIIRS Uncertainty

Flag	Description	Bits in Granule	Bits in cell
Summary SDR Quality	% good quality pixels in Granule	8	
Scan Quality Exclusion	No. scans could not be processed	5	
Moon in excluded FOV	Lunar intrusion, 1 bit/band/scan	352	
Bad Detectors (M-bands)	1 bit/detector	16	
Bad Detectors (I-bands)	1 bit/detector	32	
HAM side		16	
Imagery scan-line	Each detector/each band	128	
M-band scan line	Each detector/each band	128	
SDR Quality	Calibration quality		2
Saturated pixel			2
Missing Data	(calibration inputs)		2
Out of Range			2
Out of Range DNB			1

Materials from VIIRS SDR Users Guide
VIIRS SDR QF Descriptions

CrIS Uncertainty

- *Data Quality Indices* (21),
  - Actual numbers, including some at 32-bit float values
  - Some are at the spectrum level and some are at the individual bin level
  - Explicit listing in Appendix to presentation
- Each scan has 9 FOV (fields of view, discrete detectors) on each of 3 bands, 30 earth-view and 4 calibration FOR (fields of regard). Each band (SWIR, MWIR and LWIR) has 808, 1052 and 874 bins.
  - Spectrum level is for individual interferogram for one FOV, FOR, each band

- *Data Quality Flags* exist at Scan level, and spectrum level.

Materials extracted for the CrIS SDR User Guide. Additional help provided by Yong Chen, STAR CrIS Cal/Val Team Lead

Materials is this Session

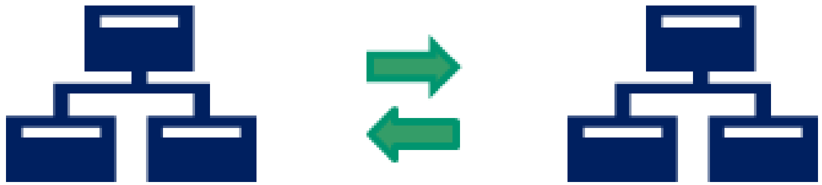
- We will have 4 talks upcoming in this Session on reprocessing
- Firstly want to turn attention to use of data in Data Assimilation and project forward to considerations of using evolving data sources in operational forecast models

New Horizons for Meteorology

New Horizons (NH)

- NASA and NOAA are Federal Agencies
  - Not able to represent NASA
  - Only participant in NOAA trying to contribute and influence planning
- Do you read Space News? Then you know about as well as I do
  - Pressures for broader industrial & academic participation
  - Acceptance of greater amounts of risk
  - Advanced innovation, injection of technological improvements

From **NESDIS' Budget & Organizational alignment**



Greater Ability to Leverage Commercial Partnerships

- Agility to strengthen partnerships with commercial entities and provide more innovative solutions to critical data sets and products

Faster, Cheaper, Better

- Usual paradigm is faster, better, cheaper
- When I started in this business the saying used to be F, B, C – pick any two
- On current large programs, paradigm frequently becomes F, B, C – pick any one
- New Horizons could be F, C, HNS (Hope for Nearly the Same performance) which is faster in schedule, cheaper in cost, and hope for nearly the same in performance

HNS

- Current facility-level calibrations and characterizations take lots of time
  - From initial system level test (Pre-Environmental Review) to sensor shipment (Pre-Ship Review) is approximately 13 months for VIIRS and CrIS and 8 months for ATMS and OMPS
  - Planet team last year talked about characterization of their constellation and not particularly individual units

How Do These Threads Knit Together

- Presume cadence of Small Sat developments will persevere
  - Strong enough to drive our 25+ yr old Conference out of our August schedule slot
- New missions include NOAA Mission Of Record (2025) will be LEO, GEO and Earth-Moon Lagrangian points such as L1 and L5
  - LEO and GEO are potential orbits for New Horizon missions for NOAA
  - Ride sharing, low cost launches, rapid construction, high technology applications are candidates

Knitting - 2

- Think of the data more as quasi-curated Crowd-Sourced datasets rather than an Expert Witness dataset we now use
- For operational forecast applications, where data latency is important attribute
- And if cadence of F, C, HNS continues, then requirement may be not if we may use those datasets but how can we use them

DISCUSSION

- What do I expect to hear in next presentations
  - High fidelity uncertainty and reprocessing are expensive endeavors
  - Sometimes they are requirements for specified missions
  - And in Data Assimilation, I expect to hear detailed uncertainty not important (maybe not used at all) but broad uncertainty metrics may be used (QF indicate which data may be excluded)
- And more in Wednesday Panel Discussion

CONCLUSIONS

- Maybe the minimum requirement for usefulness of New Horizon constellation datasets is a low bar
  - How low, how high?
- What would studies look like to demonstrate adequate dataset fidelity
- Would we need to make changes to forecasting paradigms to make NH data useful?
- How may our Metrology Community drive or contribute to this transition
- Can we influence this evolution, or do we achieve uselessness for some/many operational uses?
- These are disruptive technologies

Appendix A

CrIS Data Quality Indicators and CrIS Data Quality Flags

CrIS DQI -1

Data Quality Indicators (dimension: scan=4, sweep direction=2, FOV = 9, FOR = 30, band =3,
channel bin = (717, 869, and 637 for FSR, 717, 437, and 163 for NSR).
Also notice that spectrum level dimension is in user grid, not in sensor grid:

```
DATASET "DS_SpectralStability" {
  DATATYPE  H5T_IEEE_F64LE
  DATASPACE  SIMPLE { ( 4, 2, 9, 3 ) / ( H5S_UNLIMITED, H5S_UNLIMITED, H5S_UNLIMITED, H5S_UNLIMITED ) }
}
DATASET "DS_Symmetry" {
  DATATYPE  H5T_IEEE_F64LE
  DATASPACE  SIMPLE { ( 4, 9, 3 ) / ( H5S_UNLIMITED, H5S_UNLIMITED, H5S_UNLIMITED ) }
}
DATASET "DS_WindowSize" {
  DATATYPE  H5T_STD_U16LE
  DATASPACE  SIMPLE { ( 4, 2, 9, 3 ) / ( H5S_UNLIMITED, H5S_UNLIMITED, H5S_UNLIMITED, H5S_UNLIMITED ) }
}
DATASET "ES_ImaginaryLW" {
  DATATYPE  H5T_IEEE_F32LE
  DATASPACE  SIMPLE { ( 4, 30, 9, 717 ) / ( H5S_UNLIMITED, H5S_UNLIMITED, H5S_UNLIMITED, H5S_UNLIMITED ) }
}
DATASET "ES_ImaginaryMW" {
  DATATYPE  H5T_IEEE_F32LE
  DATASPACE  SIMPLE { ( 4, 30, 9, 869 ) / ( H5S_UNLIMITED, H5S_UNLIMITED, H5S_UNLIMITED, H5S_UNLIMITED ) }
}
DATASET "ES_ImaginarySW" {
  DATATYPE  H5T_IEEE_F32LE
  DATASPACE  SIMPLE { ( 4, 30, 9, 637 ) / ( H5S_UNLIMITED, H5S_UNLIMITED, H5S_UNLIMITED, H5S_UNLIMITED ) }
}
DATASET "ES_NEdNLW" {
  DATATYPE  H5T_IEEE_F32LE
  DATASPACE  SIMPLE { ( 4, 30, 9, 717 ) / ( H5S_UNLIMITED, H5S_UNLIMITED, H5S_UNLIMITED, H5S_UNLIMITED ) }
}
```

Please excuse font selection for the Appendix, materials provided for completeness and not intended for projections

CrIS DQI-2

```
DATASET "ES_NEdNMW" {
  DATATYPE  H5T_IEEE_F32LE
  DATASPACE  SIMPLE { ( 4, 30, 9, 869 ) / ( H5S_UNLIMITED, H5S_UNLIMITED, H5S_UNLIMITED, H5S_UNLIMITED ) }
}
DATASET "ES_NEdNSW" {
  DATATYPE  H5T_IEEE_F32LE
  DATASPACE  SIMPLE { ( 4, 30, 9, 637 ) / ( H5S_UNLIMITED, H5S_UNLIMITED, H5S_UNLIMITED, H5S_UNLIMITED ) }
}
DATASET "ES_RDRImpulseNoise" {
  DATATYPE  H5T_STD_U8LE
  DATASPACE  SIMPLE { ( 4, 30, 9, 3 ) / ( H5S_UNLIMITED, H5S_UNLIMITED, H5S_UNLIMITED, H5S_UNLIMITED ) }
}
DATASET "ES_ZPDAmplitude" {

  DATATYPE  H5T_STD_I16LE
  DATASPACE  SIMPLE { ( 4, 30, 9, 3 ) / ( H5S_UNLIMITED, H5S_UNLIMITED, H5S_UNLIMITED, H5S_UNLIMITED ) }
}
DATASET "ES_ZPDFringeCount" {
  DATATYPE  H5T_STD_U16LE
  DATASPACE  SIMPLE { ( 4, 30, 9, 3 ) / ( H5S_UNLIMITED, H5S_UNLIMITED, H5S_UNLIMITED, H5S_UNLIMITED ) }
}
DATASET "ICT_SpectralStability" {
  DATATYPE  H5T_IEEE_F64LE
  DATASPACE  SIMPLE { ( 4, 2, 9, 3 ) / ( H5S_UNLIMITED, H5S_UNLIMITED, H5S_UNLIMITED, H5S_UNLIMITED ) }
}
DATASET "ICT_TemperatureConsistency" {
  DATATYPE  H5T_IEEE_F32LE
  DATASPACE  SIMPLE { ( 4 ) / ( H5S_UNLIMITED ) }
}
```

CrIS DQI - 3

```
DATASET "ICT_TemperatureStability" {
      DATATYPE  H5T_IEEE_F32LE
      DATASPACE  SIMPLE { ( 4, 2 ) / ( H5S_UNLIMITED, H5S_UNLIMITED ) }
    }
    DATASET "ICT_WindowSize" {
      DATATYPE  H5T_STD_U16LE
      DATASPACE  SIMPLE { ( 4, 2, 9, 3 ) / ( H5S_UNLIMITED, H5S_UNLIMITED, H5S_UNLIMITED, H5S_UNLIMITED ) }
    }
    DATASET "MeasuredLaserWavelength" {
      DATATYPE  H5T_IEEE_F64LE
      DATASPACE  SIMPLE { ( 4 ) / ( H5S_UNLIMITED ) }
    }
    DATASET "MonitoredLaserWavelength" {
      DATATYPE  H5T_IEEE_F64LE
      DATASPACE  SIMPLE { ( 4 ) / ( H5S_UNLIMITED ) }
    }
    DATASET "NumberOfValidPRTTemps" {
      DATATYPE  H5T_STD_U8LE
      DATASPACE  SIMPLE { ( 4, 2 ) / ( H5S_UNLIMITED, H5S_UNLIMITED ) }
    }
    DATASET "ResamplingLaserWavelength" {
      DATATYPE  H5T_IEEE_F64LE
      DATASPACE  SIMPLE { ( 4 ) / ( H5S_UNLIMITED ) }
    }
    DATASET "SDRFringeCount" {
      DATATYPE  H5T_STD_U16LE
      DATASPACE  SIMPLE { ( 4, 30, 9, 3 ) / ( H5S_UNLIMITED, H5S_UNLIMITED, H5S_UNLIMITED, H5S_UNLIMITED ) }
    }
```

CrIS DQF

Data Quality Flag, there is no spectrum level flag, each band has only one entry (band = 3 to indicate which band):

```
DATASET "QF1_SCAN_CRISSDR" {
   DATATYPE  H5T_STD_U8LE
   DATASPACE  SIMPLE { ( 4 ) / ( H5S_UNLIMITED ) }
}
DATASET "QF2_CRISSDR" {
   DATATYPE  H5T_STD_U8LE
   DATASPACE  SIMPLE { ( 4, 9, 3 ) / ( H5S_UNLIMITED, H5S_UNLIMITED, H5S_UNLIMITED ) }
}
DATASET "QF3_CRISSDR" {
   DATATYPE  H5T_STD_U8LE
   DATASPACE  SIMPLE { ( 4, 30, 9, 3 ) / ( H5S_UNLIMITED, H5S_UNLIMITED, H5S_UNLIMITED,
H5S_UNLIMITED ) }
}
DATASET "QF4_CRISSDR" {
   DATATYPE  H5T_STD_U8LE
   DATASPACE  SIMPLE { ( 4, 30, 9, 3 ) / ( H5S_UNLIMITED, H5S_UNLIMITED, H5S_UNLIMITED,
H5S_UNLIMITED ) }
}
```