Experimental approaches to uncertainty: Communicating the unquantifiable

Adam Povey and Don Grainger

Thanks to Thomas Holzer-Popp, Greg McGarragh, Kerstin Stebel

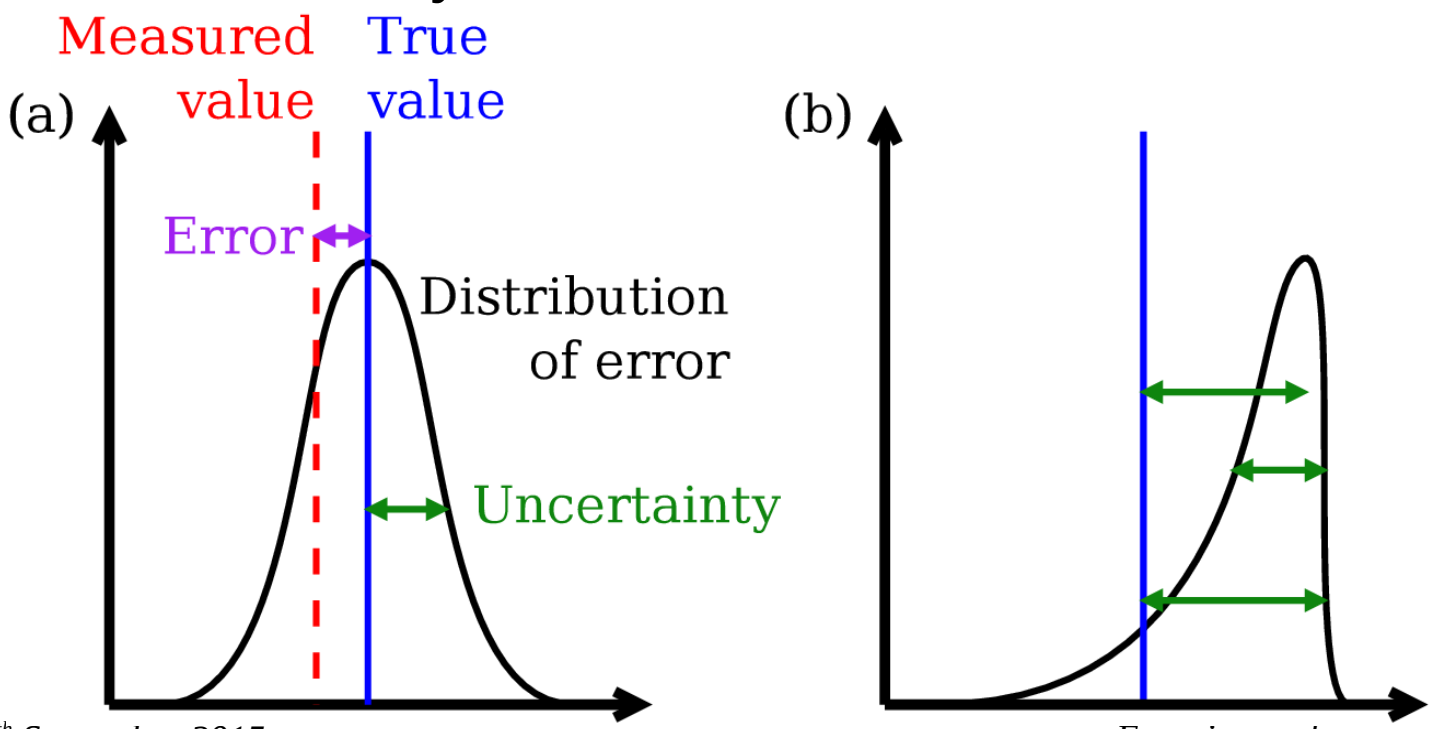
UNIVERSITY OF OXFORD

National Centre for Earth Observation
NATURAL ENVIRONMENT RESEARCH COUNCIL

Representing uncertainty

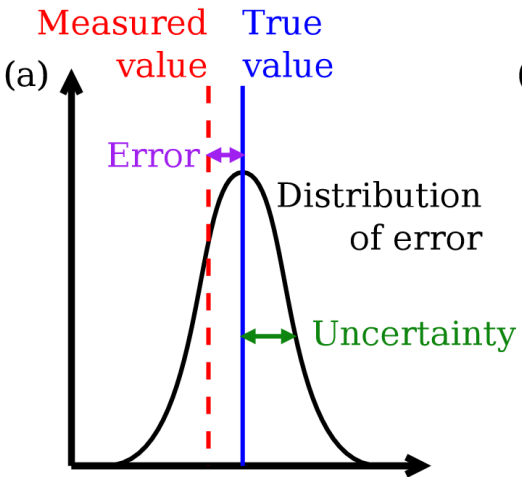
a) Typical understanding of uncertainty as the width of the distribution of error around the measurand. This is generally accurate for random errors.

b) The 'uncertainty' of a systematic error can be more complex as the distribution of error is asymmetric with non-zero mean.



Classifying sources of error

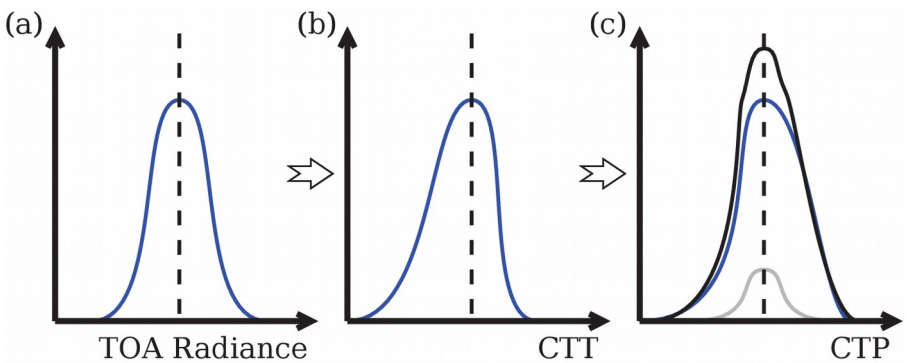
- Intrinsic sources of error
  - Measurement
    - Statistical variation in measurand or detector
    - e.g. dark current, radiometric calibration
  - Parameter
    - Uncertainty in auxiliary information used
    - e.g. spectroscopic data, meteorological profiles
  - Both also known as “parametric errors”



- Generally well-represented by traditional techniques for calculating uncertainty.

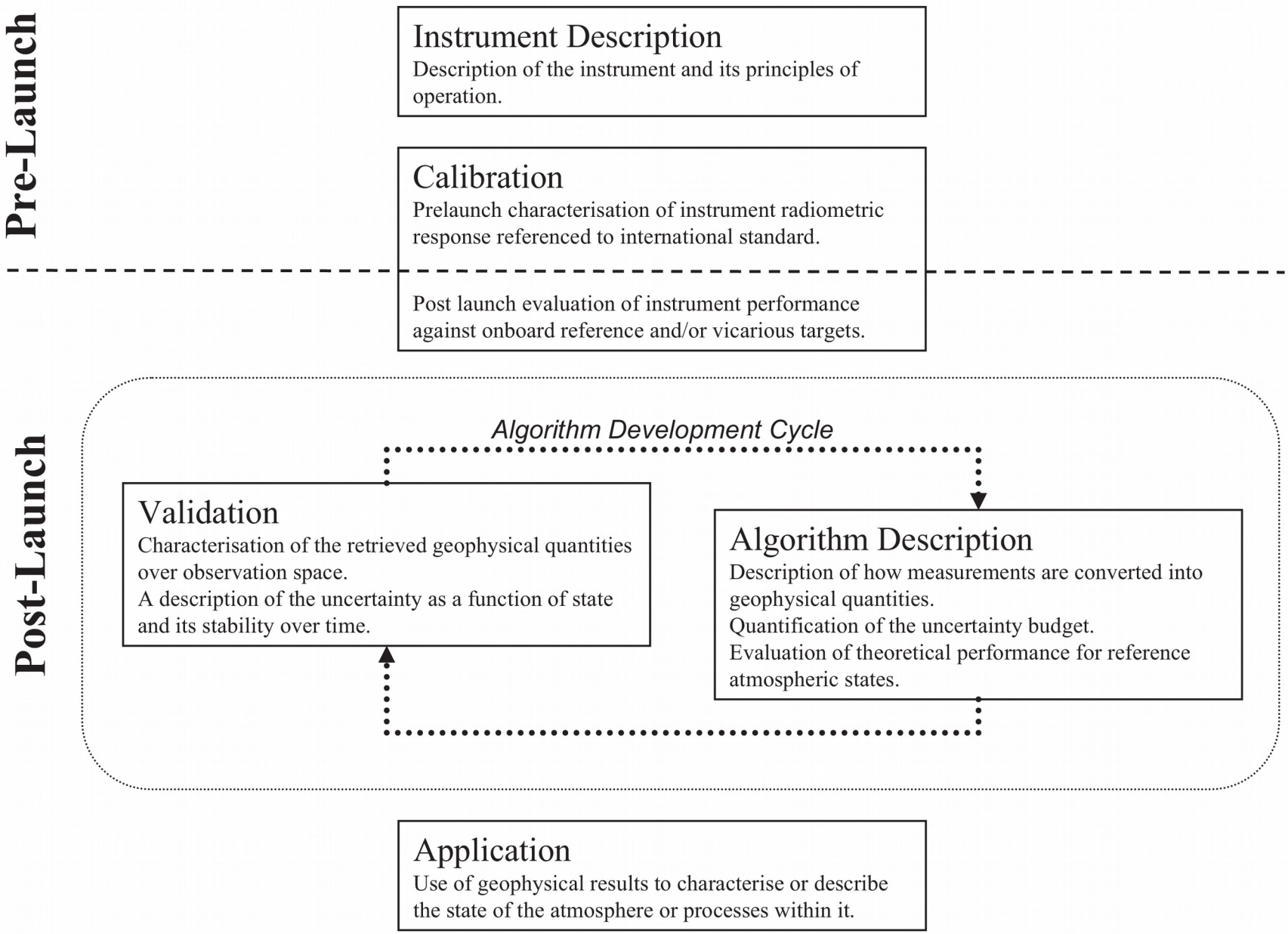
Classifying sources of error

- Structural errors, resulting from choices made in the measurement and analysis systems
  - Approximation
    - Simplifications and approximations in calculations
    - e.g. using a LUT or plane-parallel atmosphere
  - System
    - Physically meaningful assumptions
    - e.g. choice of aerosol optical model
  - Resolution
    - Finite sampling of a constantly varying system
    - e.g. fair-weather bias, MODIS “bow-tie effect”



- Potentially non-linear and circumstantial.
  Thus, the source of error affects how it needs to be reported to users.

Communicating with scientific community



Communicating with users

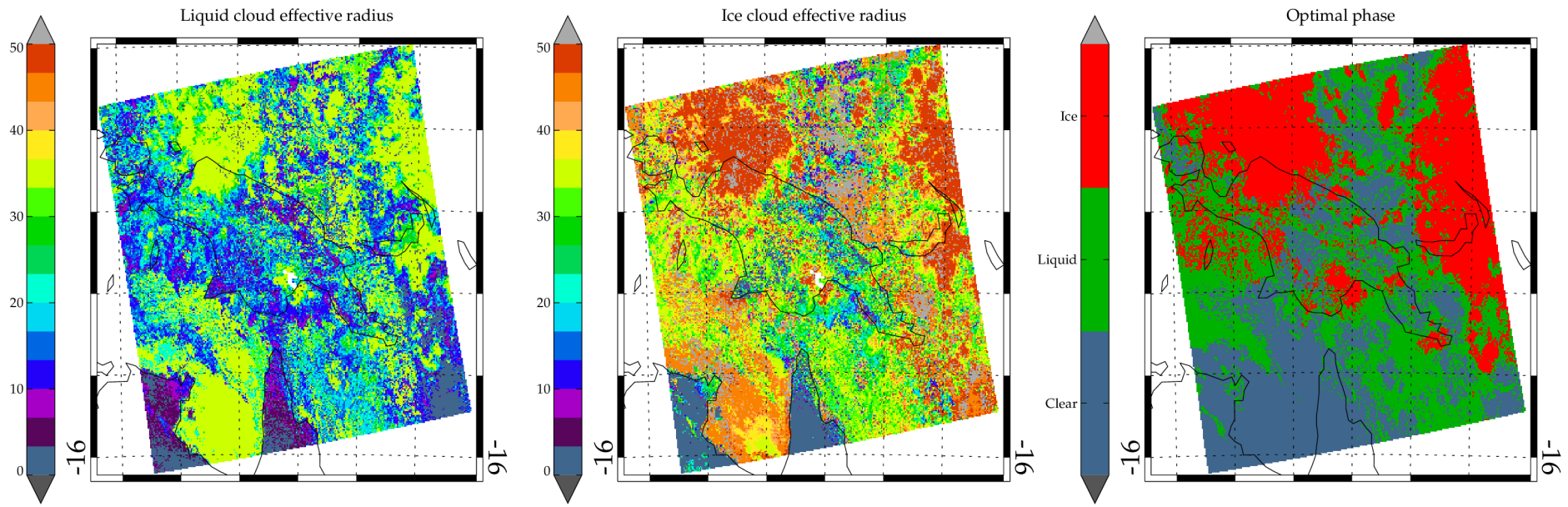
- Important to clarify the difference between what is measured and what the user wants to know.
  - For example, satellites only sample one time of day but users may need observations at other times.
    - For SST, in situ observations indicate the diurnal cycle is somewhat predictable so empirical corrections may be useful.
    - For cloud, data indicates the diurnal cycle is highly circumstantial so a single observation is not representative of an entire day.
- A quantity can be adjusted to meet a user's needs but
  - The transformation will introduce additional uncertainty and
  - The new quantity may not meet other user's needs.
- We prefer to report only the measured value and provide secondary support (i.e. communication) to produce transformations as needed.

Communicating with users

- Many satellite products require significant assumptions to constrain their analysis.
  - Some simplify the problem, introducing approximation errors
    - e.g. neglecting multiple scattering
  - Others are physical constraints, introducing system errors
    - e.g. classifying land type into finite, idealised categories
  - In either case, it isn't obvious how to quantify the uncertainty

Ensemble techniques

- Ensemble techniques can better communicate uncertainties resulting from such errors.
  - A “multi-model” ensemble of analyses with differing assumptions and approximations.
  - A “multi-run” ensemble of analyses with different constraints.



System maturity

CORE-CLIMAX System Maturity Matrix

Maturity	SOFTWARE READINESS	METADATA	USER DOCUMENTATION	UNCERTAINTY CHARACTERISATION	PUBLIC ACCESS, _x000D_FEEDBACK, UPDATE	USAGE
1	Conceptual development	None	Limited scientific description of the methodology available from PI	None	Restricted availability from PI	None
2	Research grade code	Research grade	Comprehensive scientific description of the methodology, report on limited validation, and limited product user guide available from PI; paper on methodology is sumitted for peer-review	Standard uncertainty nomenclature is idenitified or defined; limited validation done; limited information on uncertainty available	Data avaliable from PI, feedback through scientific exchange, irregular updates by PI	Research: Benefits for applications identified_x000D_DSS: Potential benefits identified
3	Research code with partially applied standards; code contains header and comments, and a README file; PI affirms portability, numerical reproducibility and no security problems	Standards defined or identified; sufficient to use and understand the data and extract discovery metadata	Score 2 + paper on methodology published; comprehensive validation report available from PI and a paper on validation is submitted; comprehensive user guide is available from PI; Limited description of operations cocept available from PI	Score 2 + standard nomenclature applied; validation extended to full product data coverage, comprehensive information on uncertainty available; methods for automated monitoring defined	Data and documentation publically available from PI, feedback through scientifc exchange, irregular updates by PI	Research: Benefits for applications demonstrated._x000D_DSS: Use occuring and benefits emerging
4	Score 3 + draft software installation/user manual available; 3rd party affirms portability and numerical reproducibility; passes data providers security review	Score 3 + standards systematically applied; meets international standards for the data set; enhanced discovery metadata; limited location level metadata	Score 3 + comprehensive scientific description available from data provider; report on inter comparison available from PI; paper on validation published; user guide available from data provider; comprehensive description of operations concept available from PI	Score 3 + procedures to establish SI traceability are defined; (inter)comparison against corresponding CDRs (other methods, models, etc); quantitative estimates of uncertainty provided within the product characterising more or less uncertain data points; automated monitoring partially implemented	Data record and documentation available from data provider and under data provider's version control; Data provider establishes feedback mechanism; regular updates by PI	Score 3 + _x000D_Research: Citations on product usage in occurring_x000D_DSS: societal and economical benefits discussed
5	Score 4 + operational code following standards, actions to achieve full compliance are defined; software installation/user manual complete; 3rd party installs the code operationally	Score 4+ fully compliant with standards; complete discovery metadata; complete location level metadata	Score 4 + comprehensive scientific description maintained by data provider; report on data assessment results exists; user guide is regularly updated with updates on product and validation; description on practical implementation is available from data provider	Score 4 + SI traceability partly established; data provider participated in one inter-national data assessment; comprehensive validation of the quantitative uncertainty estimates; automated quality monitoring fully implemented (all production levels)	Score 4 + soure code archived by Data Provider; feedback mechanism and international data quality assessment are considered in periodic data record updates by Data Provider	Score 4+_x000D_Research: product becomes reference for certain applications_x000D_DSS: Societal and economic benefits are demonstrated
6	Score 5 + fully compliant with standards; Turnkey System	Score 5 + regularly updated	Score 5 + journal papers on product updates are and more comprehensive validation and validation of quantitative uncertainty estimates are published; operations concept regularly updated	Score 5 + SI traceability established; data provider participated in multiple inter-national data assessment and incorporating feedbacks into the product development cycle; temporal and spatial error covariance quantified; Automated monitoring in place with results fed back to other accessible information, e.g. meta data or documentation	Score 5 + source code available to the public and capability for continuous data provisions established (ICDR)	Score 5 + _x000D_Research: Product and its applications becomes references in multiple research field_x000D_DSS: Influence on decision and policy making demonstrated

- Designed for judging the suitability of an algorithm for operational use and it's uptake by the community.
- Does not measure the science value, i.e. if the product extracts all the information available or addresses the needs of users.
- We are concerned that using these to publicly evaluate products could mistake data with a short/long heritage as in/accurate.

Summary

- Uncertainty must represent the non-linear and circumstantial nature of errors often dominant in satellite products.
- Communication of uncertainty to users should be tailored to
  - The data available to calculate and validate the uncertainty,
  - The underlying cause of the error, and
  - A broad range of users' needs.
- As the distribution of these errors is not always well understood, data producers must engage in a dialogue with data users to work towards useful estimates.
  - This can include ensemble techniques, quality assurance, qualitative descriptions.
  - This is complimented, not replaced, by validation activities.

Questions and comments?

Atmos. Meas. Tech. Discuss., 8, 8509–8562, 2015
www.atmos-meas-tech-discuss.net/8/8509/2015/
doi:10.5194/amtd-8-8509-2015

Atmospheric Measurement Techniques Discussions
Open Access

Known and unknown unknowns: ~~the application of ensemble techniques to~~ uncertainty estimation in satellite remote sensing ~~data~~

A. C. Povey and R. G. Grainger

National Centre for Earth Observation, University of Oxford, Clarendon Laboratory, Parks Road, Oxford OX1 3PU, UK

Received: 23 June 2015 – Accepted: 20 July 2015 – Published: 10 August 2015

Correspondence to: A. C. Povey (adam.povey@physics.ox.ac.uk)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Uncertainty

- Uncertainty is a vital component of data as it provides
  - a means of efficiently and consistently communicating the strengths and limitations of data to users, and
  - a metric with which to compare different data sets.
- Uncertainty should be
  - Universal: calculated in a manner applicable to all data;
  - Internally consistent: derived from the same information as the data and not draw on external validation data;
  - Transferable: of use to users for various purposes.

Ensemble Techniques

- Consider estimating the volume of a bucket, knowing only its mass.
- Shape is assumed but different assumptions produce different errors.
- The magnitude of error due to that assumption depends on the underlying state and other parameters of the retrieval.