Validating uncertainty estimates

Andrew Sayer

GESTARII-UMBC at NASA GSFC

andrew.sayer@nasa.gov

So your retrieval has an uncertainty estimate…

- How do we know if the uncertainty estim[illegible]useful or realistic?

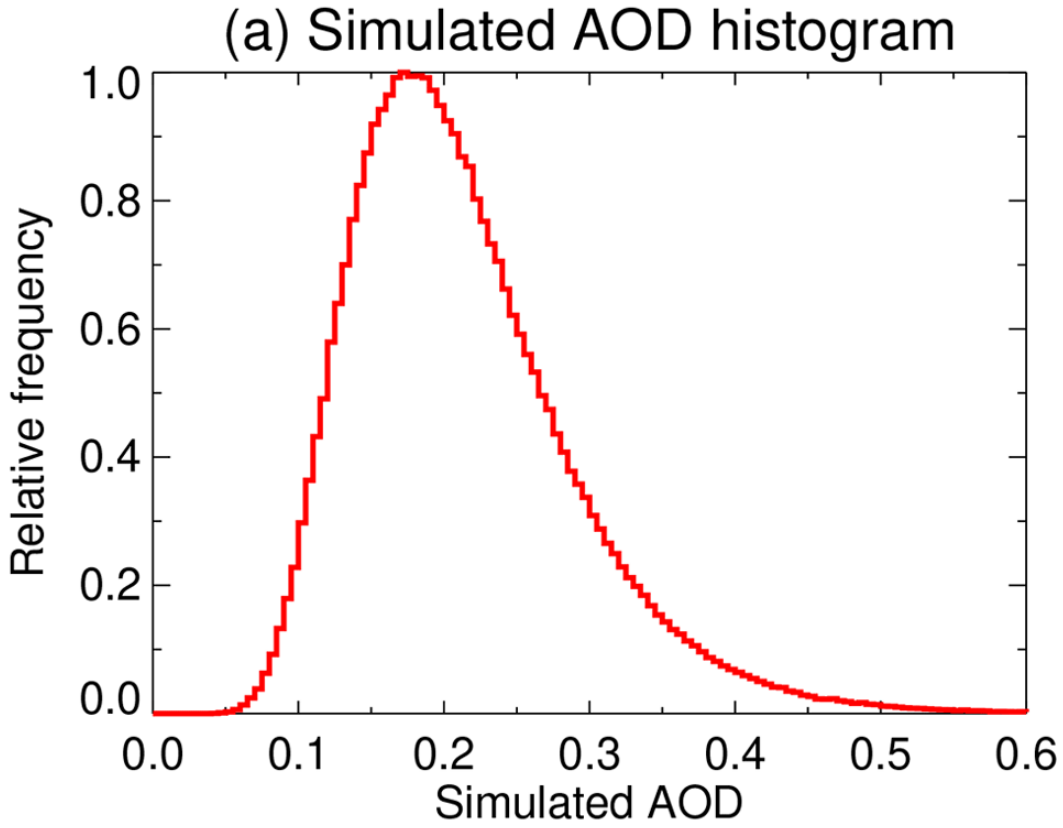
Reminder: uncertainty and error are different

- Uncertainty is an *expected level of dispersion*
  - Typically one standard deviation (1σ, ~68% of the time as good as X)
- Error is an observed quantity
  - A *draw from the distribution* of possible errors given the uncertainty level

You *expect* 3.5
You *get* 1, 2, 3, 4, 5, or 6

Uncertainty distributions and error distributions

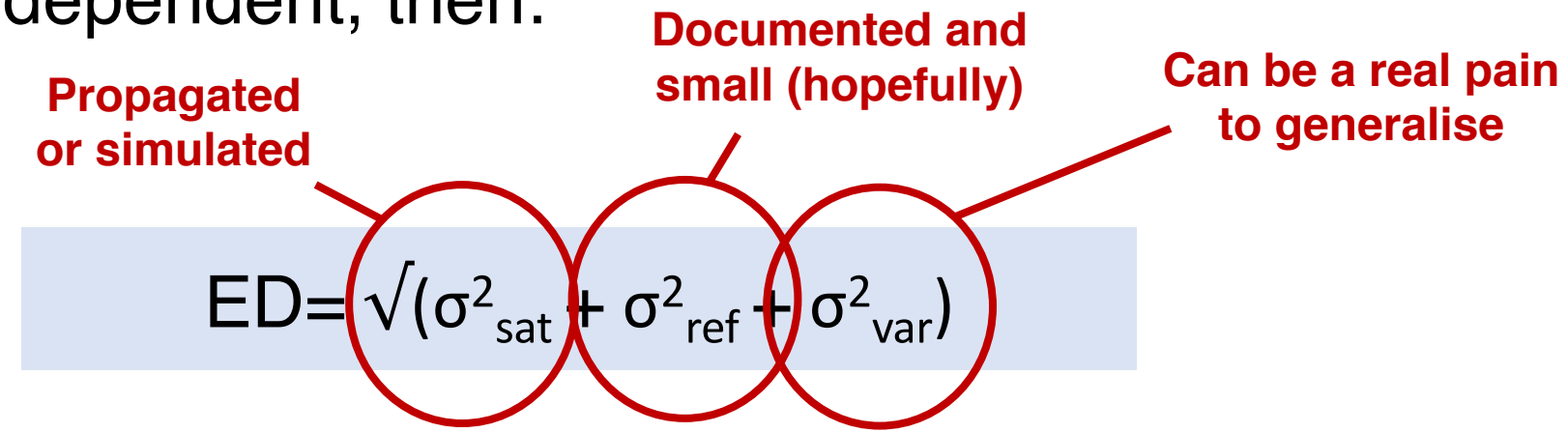
- Suppose AOD uncertainty is ±(0.05+15%). Then:



Sayer *et al.*, AMT (2020)

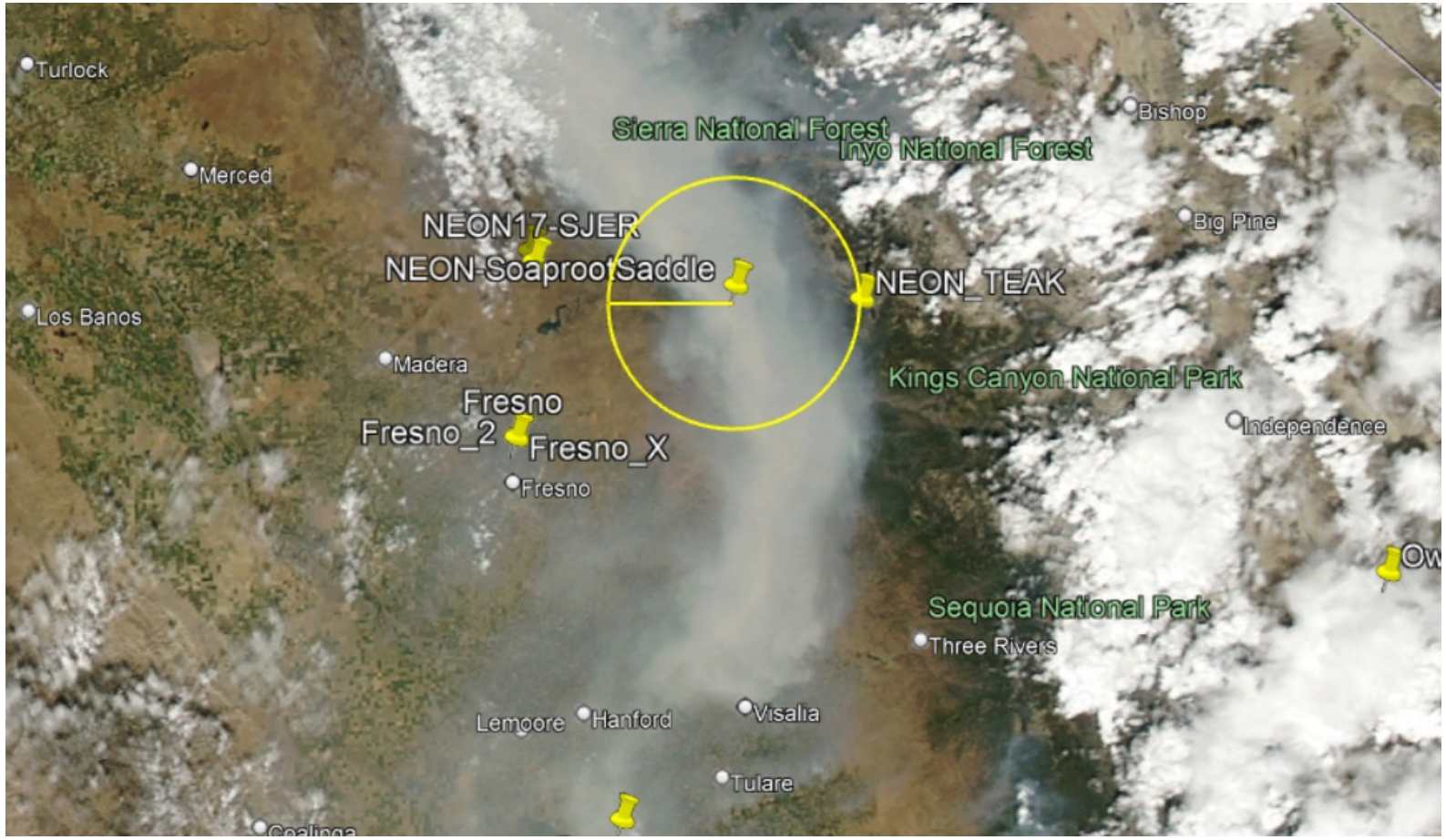
Solution: compare distributions of observed errors to expected discrepancy (ED)

- Observed error = (truth – retrieval)
- Expected discrepancy contains contributions from:
  - Retrieval uncertainty σ_{sat}
  - Reference uncertainty σ_{ref}
  - Spatiotemporal variability inherent in the comparison σ_{var}
- If assumed independent, then:

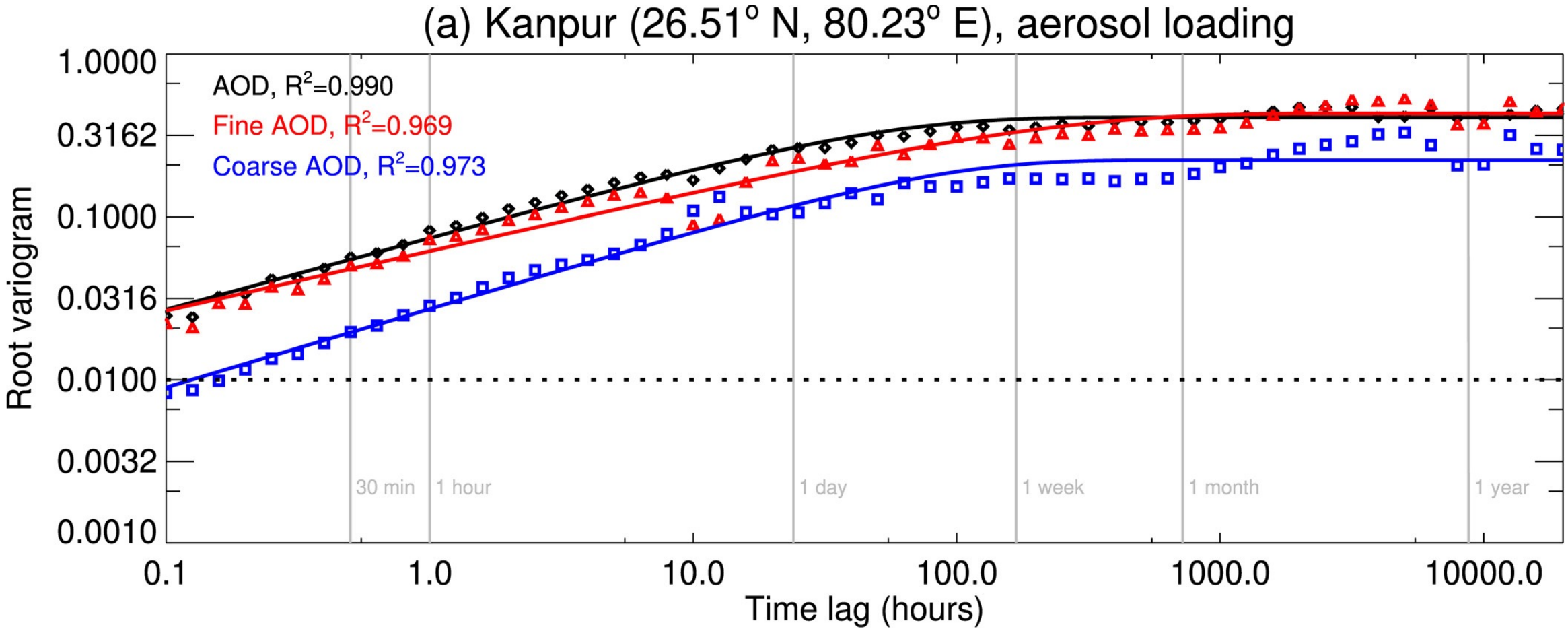


Spatial and temporal variations can both be important

- Look at *spatial statistics* of satellite data in the area, *temporal statistics* of the ground data
- *Generalise* based on a model
- Use *robust statistics* (e.g., median not mean)
- *Shrink* area aggregated or comparison window

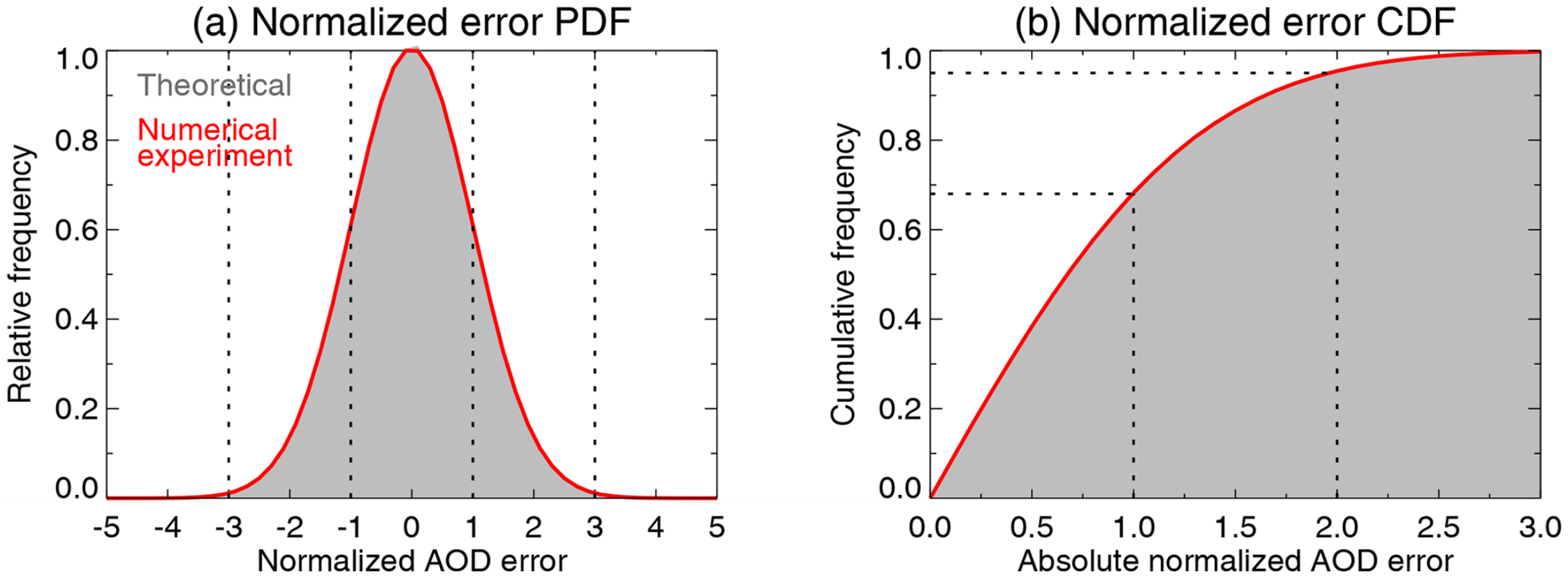


Spatial and temporal variations can both be important



Sayer, ESS (2020)

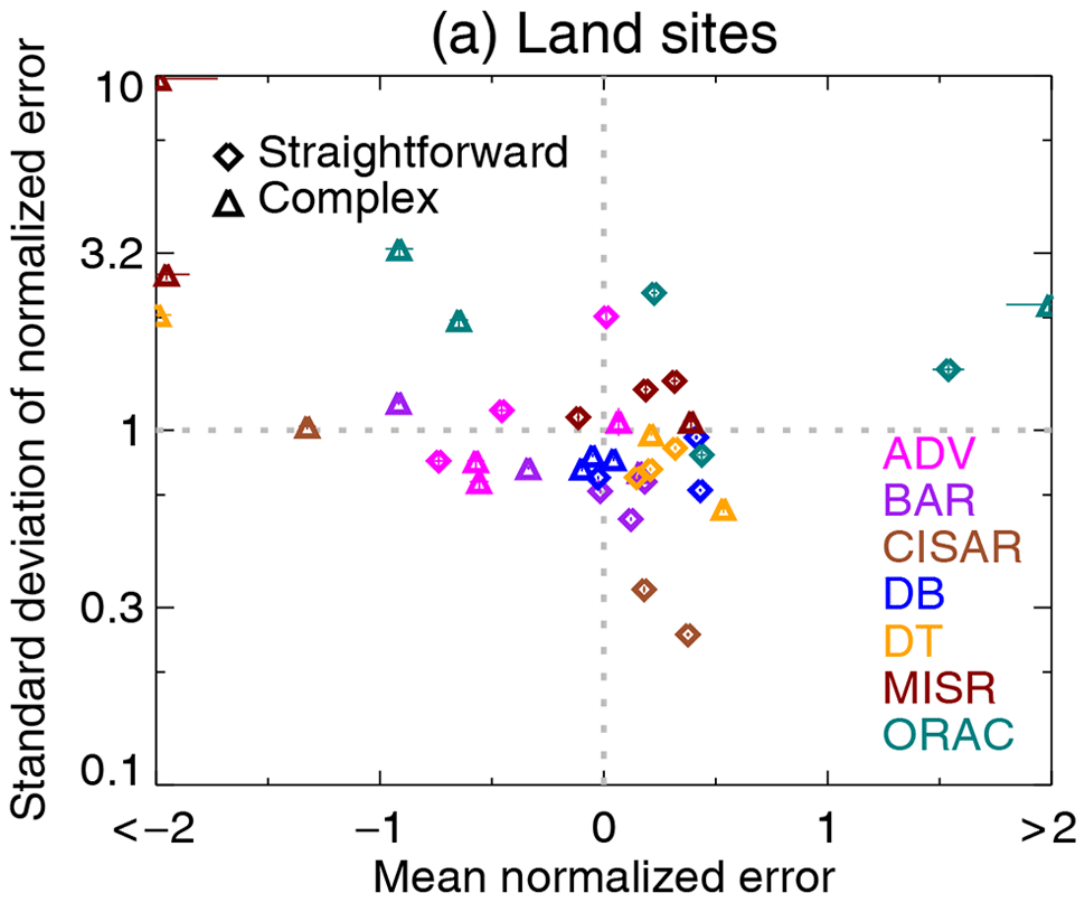
Look at distributions of normalised error

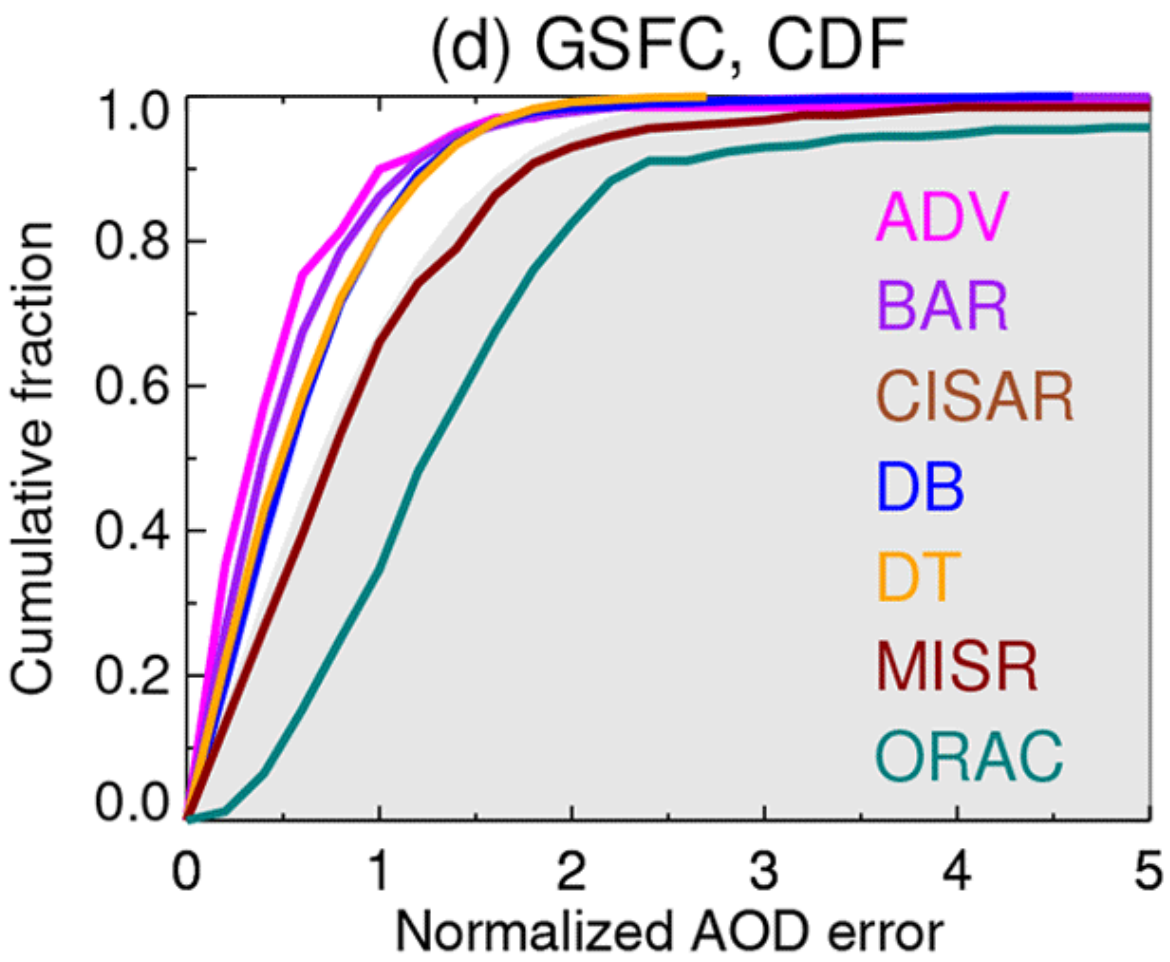


- Normalise as error/ED
- If unbiased and appropriate magnitude, distributions should follow N(0,1)

Sayer *et al.*, AMT (2020)

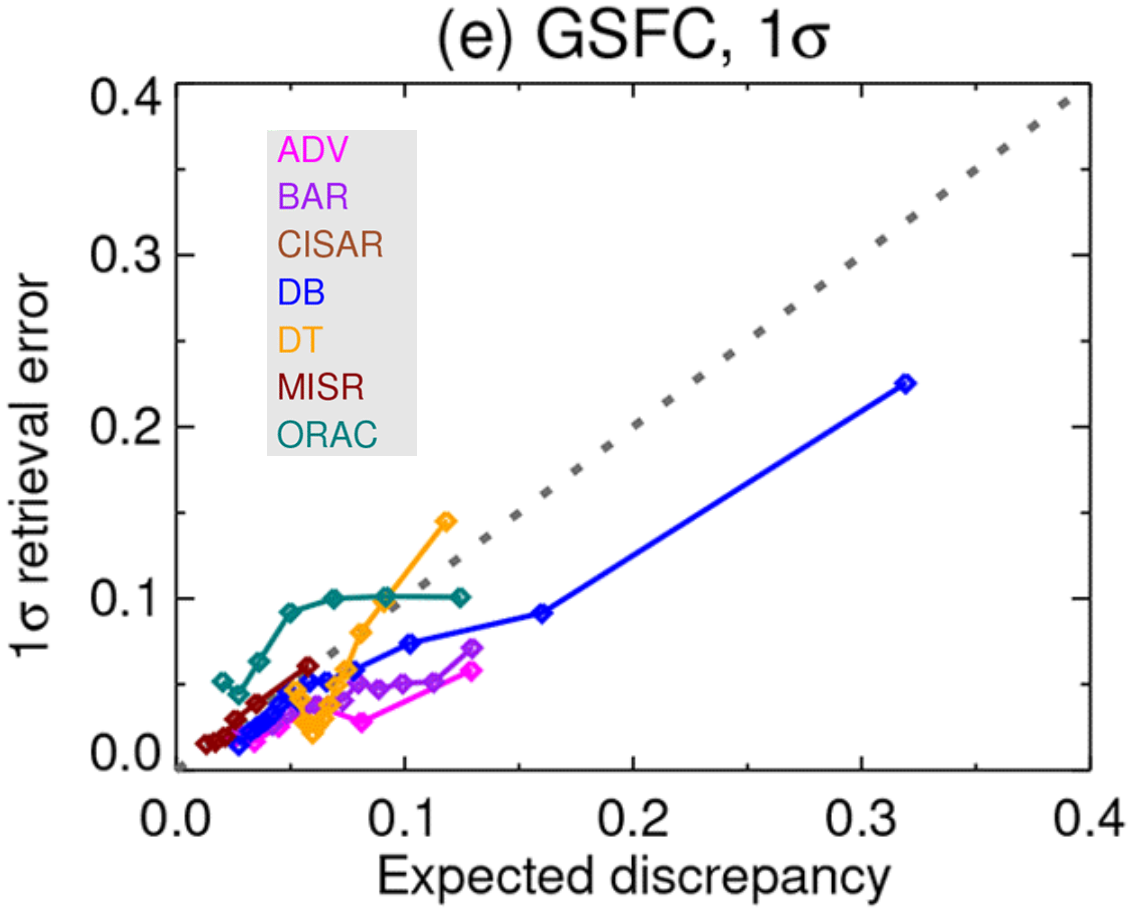
Examples on real aerosol data



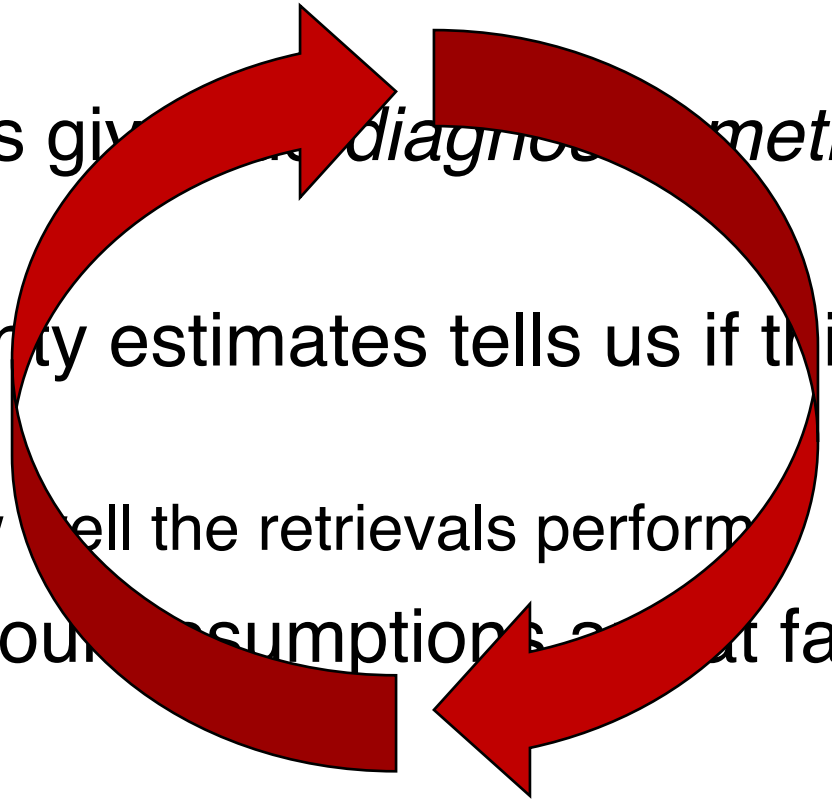


Can we tell low-uncertainty situations from high ones?

- Stratify by ED
- Look at 1σ point (68th percentile) of absolute error
  - Additional percentiles can tell you more about *how Gaussian* the error distribution is for a given ED
- Similar to *forecast calibration*, *skill scores*, etc



Closing the loop



- Validating retrievals gi[illegible] *diagnostic metrics of quality* to report uncertainty
- Validating uncertainty estimates tells us if things are *working as we expect*
  - Regardless of how well the retrievals perform
- Figure out where your assumptions are at fault, then *refine* them!

References and resources

- Merchant, C. J., Paul, F., Popp, T., Ablain, M., Bontemps, S., Defourny, P., Hollmann, R., Lavergne, T., Laeng, A., de Leeuw, G., Mittaz, J., Poulsen, C., Povey, A. C., Reuter, M., Sathyendranath, S., Sandven, S., Sofieva, V. F., and Wagner, W. (2017), Uncertainty information in climate data records from Earth observation, *Earth Syst. Sci. Data*, 9, 511–527, https://doi.org/10.5194/essd-9-511-2017
- Sayer, A. M. (2020), How long is too long? Variogram analysis of AERONET data to aid aerosol validation and intercomparison studies, *Earth and Space Science*, 7, e2020EA001290. https://doi.org/10.1029/2020EA001290
- Sayer, A. M., Govaerts, Y., Kolmonen, P., Lipponen, A., Luffarelli, M., Mielonen, T., Patadia, F., Popp, T., Povey, A. C., Stebel, K., and Witek, M. L. (2020), A review and framework for the evaluation of pixel-level uncertainty estimates in satellite aerosol remote sensing, *Atmos. Meas. Tech.*, 13, 373–404, https://doi.org/10.5194/amt-13-373-2020
- McKinna, L. I. W., Cetinić, I., & Werdell, P. J. (2021), Development and validation of an empirical ocean color algorithm with uncertainties: A case study with the particulate backscattering coefficient, *Journal of Geophysical Research: Oceans*, 126, e2021JC017231, https://doi.org/10.1029/2021JC017231