EARTH SYSTEM SCIENCE INTERDISCIPLINARY CENTER

Decomposing Uncertainty in ML-Based Precipitation Type Retrievals

Aleatoric and epistemic uncertainty for GMI convective / stratiform classification

Veljko Petković[1], Marko Orescanin[2], Barret Manfre[2], Jackson Tan[3], Pierre Kirstetter[4], Daniel Watters[4], Malarvizhi Arulraj[1], Lei Ji[1], Huan Meng[5], Ralph Ferraro[1]

[1] Earth System Science Interdisciplinary Center, University of Maryland, College Park, MD, USA
[2] Naval Postgraduate School, Monterey, CA, USA
[3] UMBC, NASA GSFC
[4] University of Oklahoma
[5] NOAA/NESDIS/STAR, College Park, MD, USA

IPWG-12; Kraków, Poland; 8 July 2026

MOTIVATION

Why uncertainty matters

ML-based precipitation retrievals are moving into operational use. Transparent uncertainty is what makes that move trustworthy.

Assess algorithm maturity

Quantified confidence shows how much a retrieval can be trusted.

Enable comparability

A shared confidence measure lets users compare retrievals on equal terms.

Build user confidence

Users can see, per pixel, where an estimate is reliable.

The gap: Most ML models give point estimates, with no meaningful confidence attached.

CONCEPT

Total uncertainty splits into two distinct parts

Aleatoric

INHERENT DATA AMBIGUITY

- Irreducible; the measurement itself is ambiguous
- Different precipitation types, similar microwave signatures
- More training data will not reduce it

“the observation cannot tell them apart”

Epistemic

MODEL KNOWLEDGE GAP

- Reducible; reflects what the model has not learned
- Large where training data is sparse or out of distribution
- More representative training data reduces it

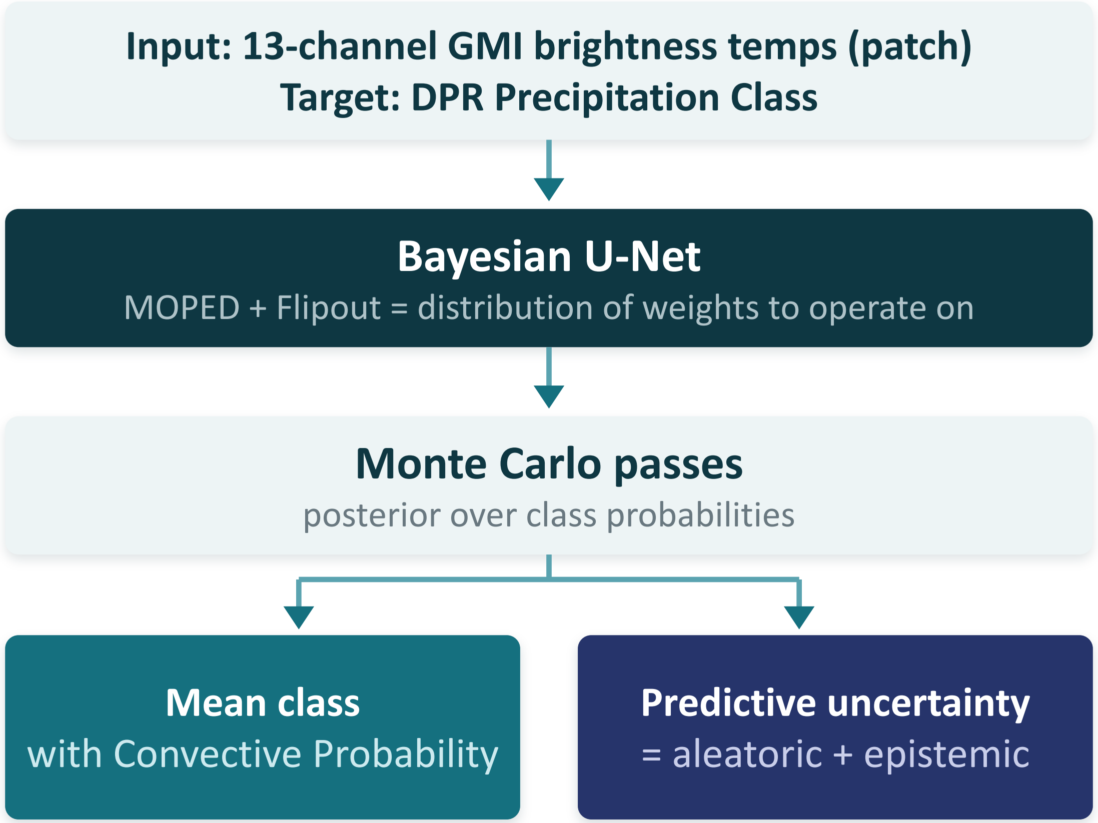
“the model has not seen enough like this”

Total predictive uncertainty = aleatoric + epistemic. The decomposition gives distinct, complementary diagnostics.

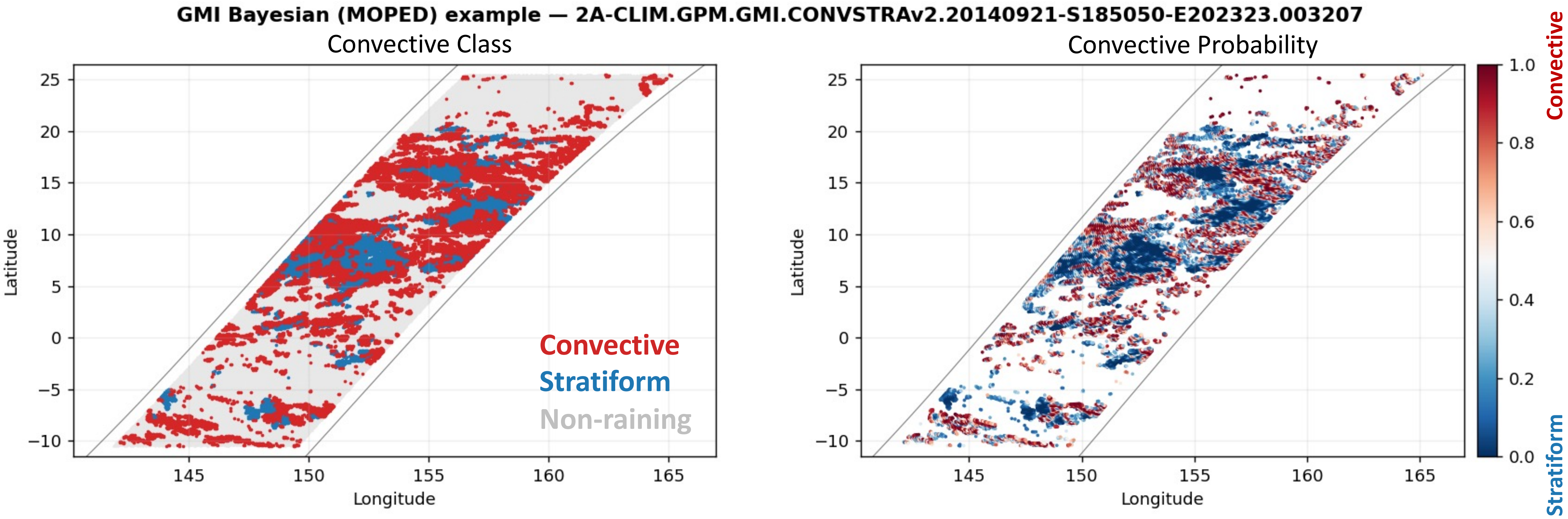
METHOD

A Bayesian Deep Learning U-Net

- Bayesian U-Net (ResNet50) maps 13-channel GMI patches to the DPR per-pixel class (stratiform/convective)
- Bayesian with MOPED priors (+ Flipout), giving a Monte Carlo posterior over class probabilities
- Total predictive uncertainty is split into aleatoric and epistemic parts
- Estimates are physically interpretable, spatially coherent, and informative for retrieval improvement

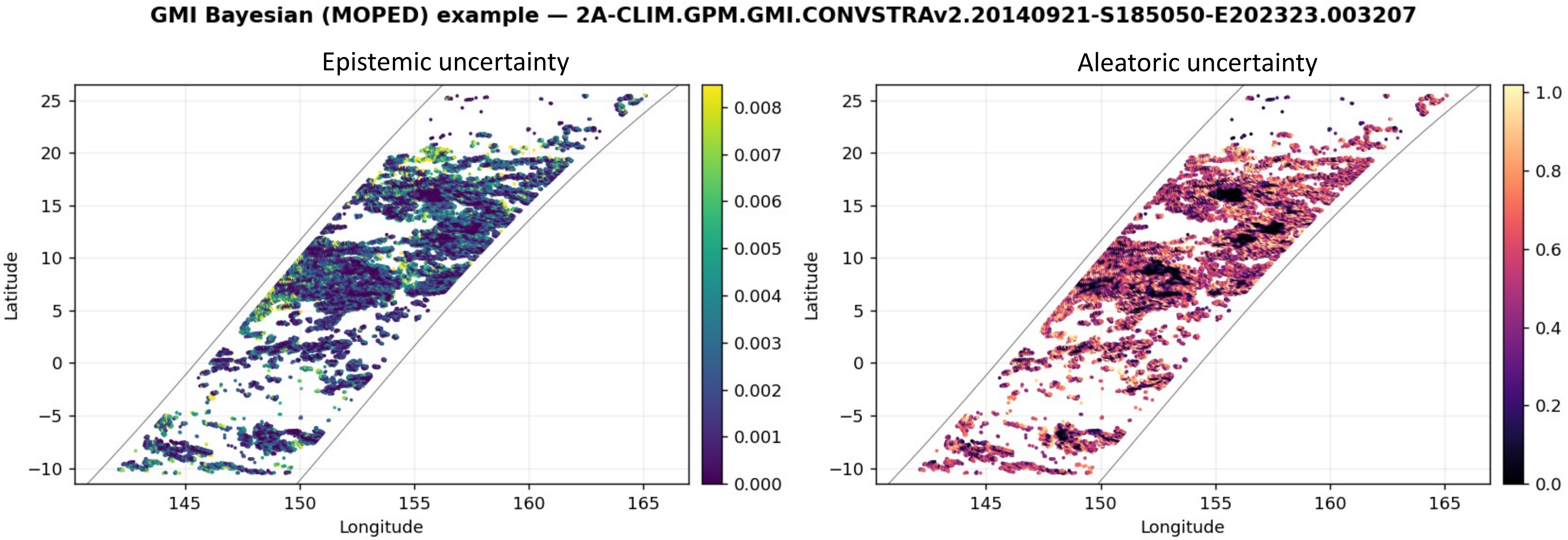


Every classification comes with per-pixel uncertainty



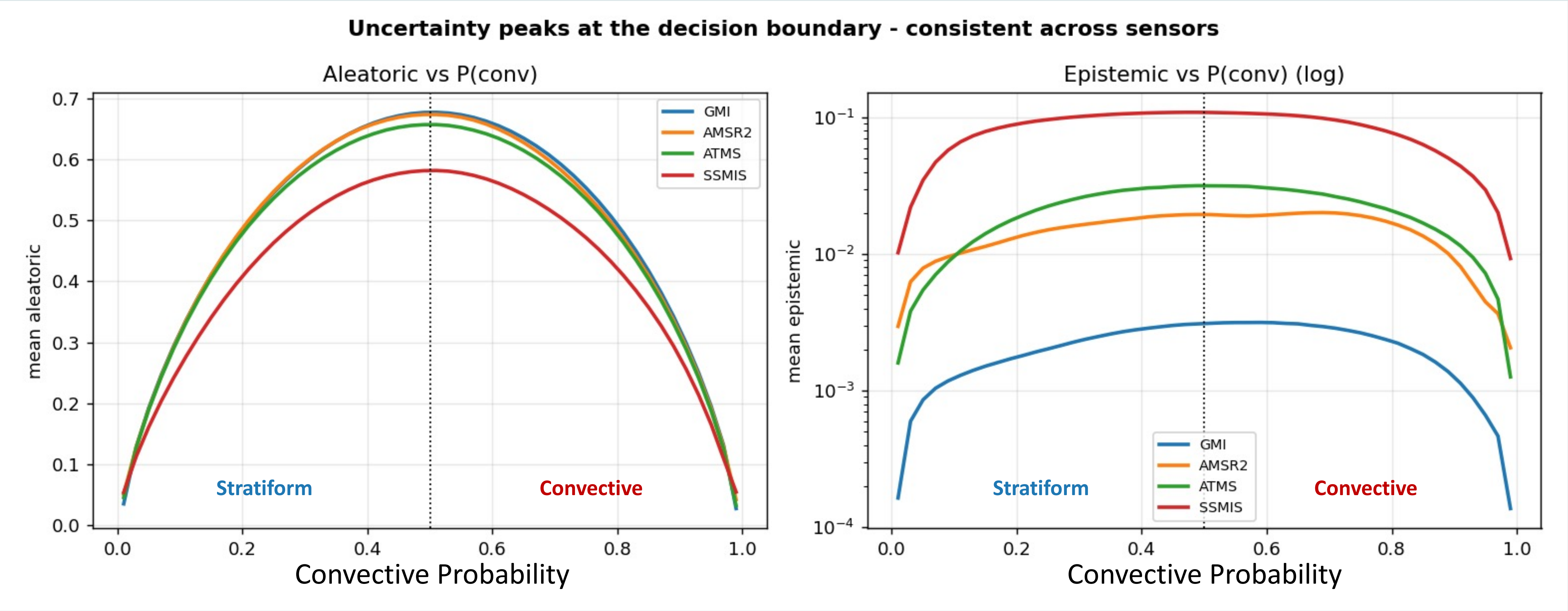
Per-pixel class and probability, P(convective), for a portion of a single GMI orbit (tropical West Pacific).

Every classification comes with per-pixel uncertainty



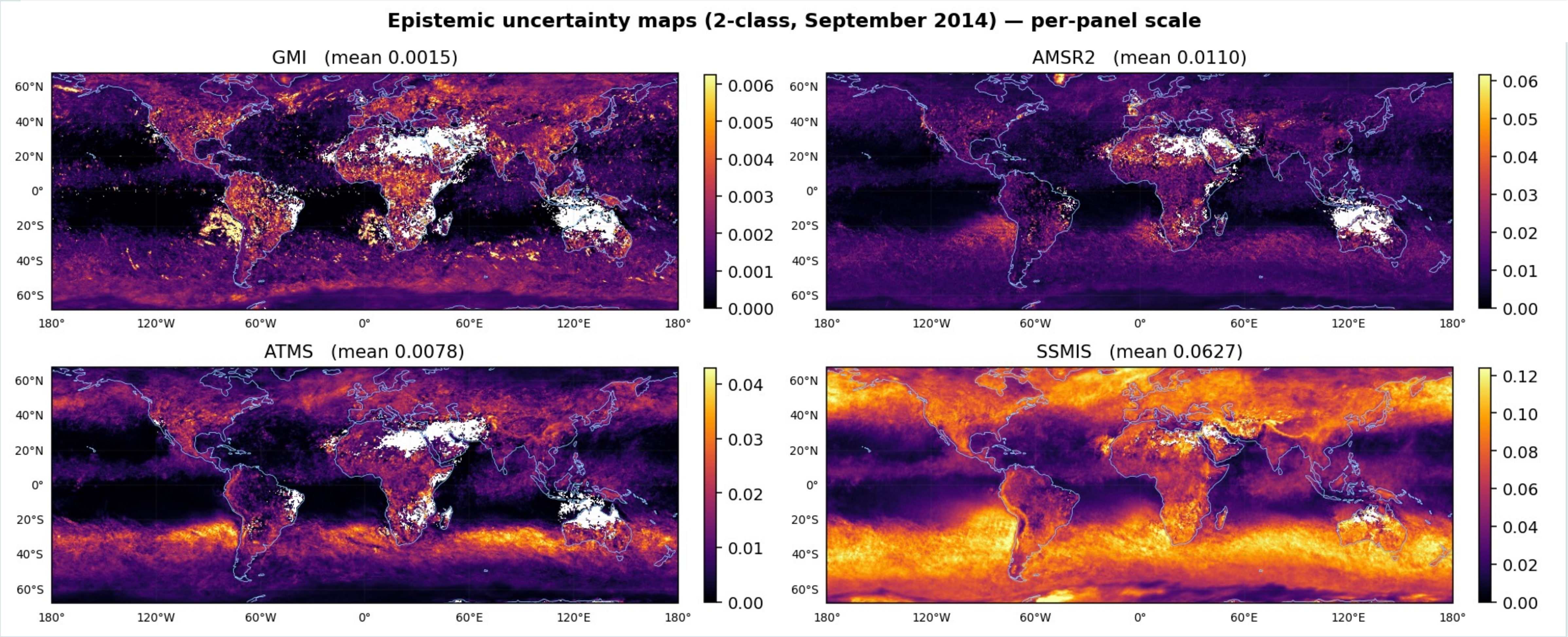
The two uncertainty components for a single GMI orbit (tropical West Pacific).

Aleatoric uncertainty dominates by two orders of magnitude



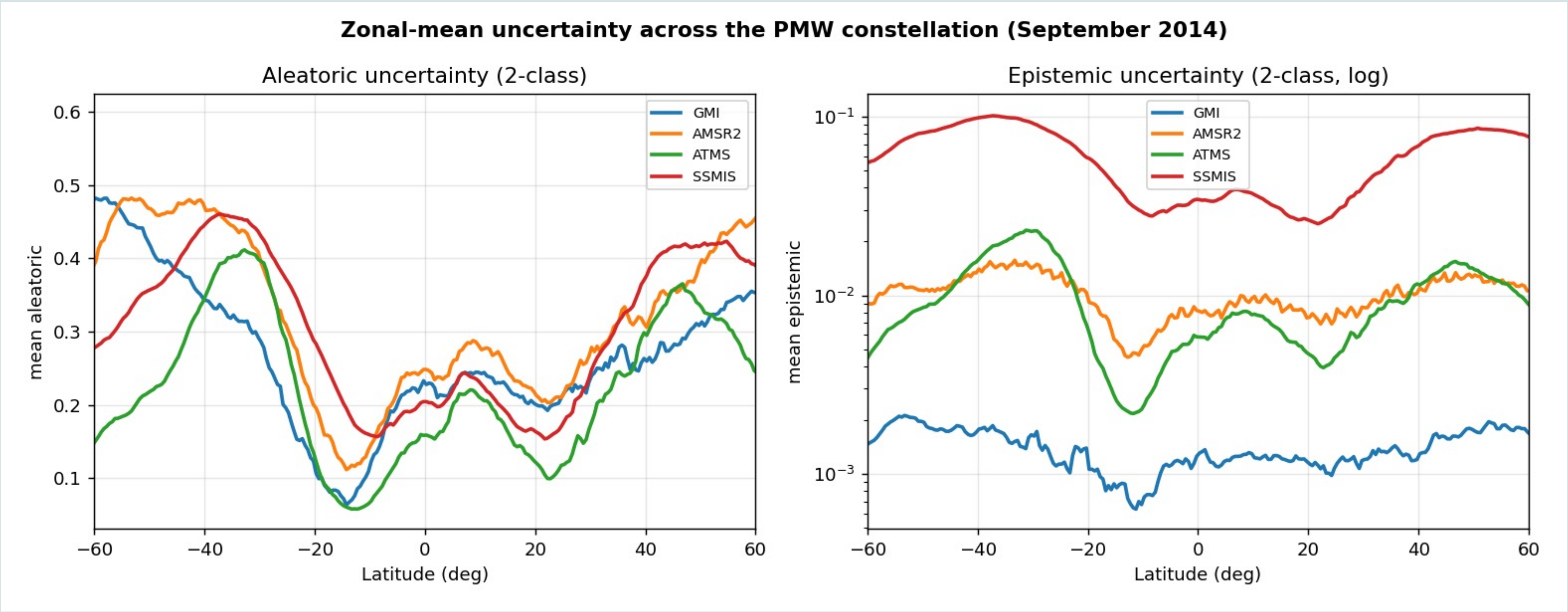
Aleatoric (left) and epistemic (right, log scale) versus Convective Probability; aleatoric peaks at the P = 0.5 (decision threshold). The limit is the ambiguity of passive microwave signatures, not insufficient training.

Uncertainty is regime dependent (flagging high-latitude gaps)



Per-pixel epistemic uncertainty, September 2014. Largest at high latitudes (mixed-phase processes, frozen surfaces, light precipitation), all hard for passive microwave; values rise poleward beyond the ±67° training domain.

The findings hold across the wider PMW constellation



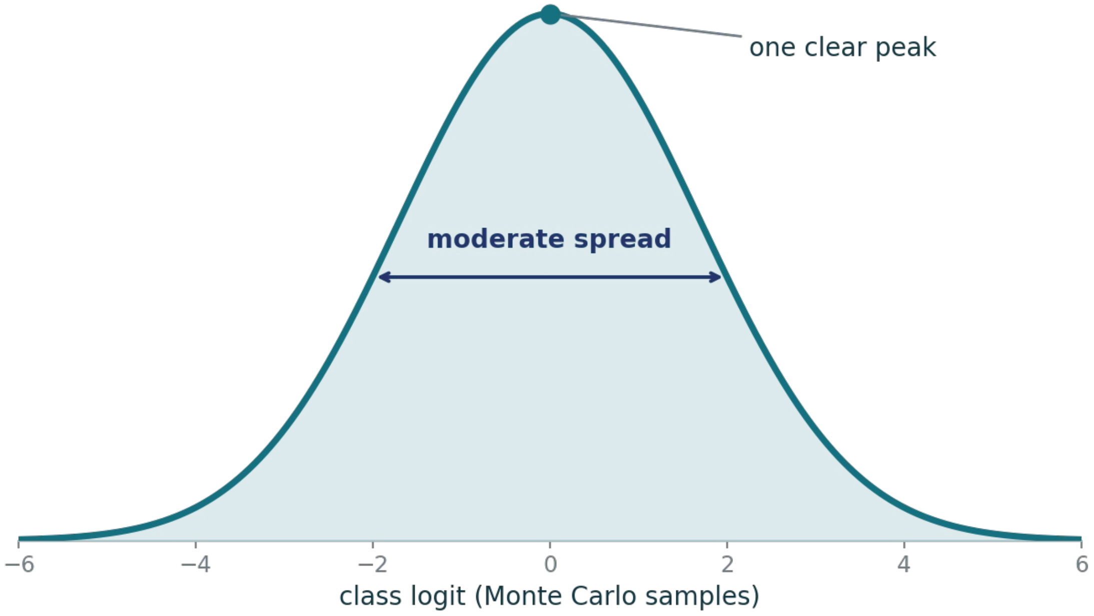
Zonal-mean aleatoric and epistemic uncertainty for GMI, AMSR2, ATMS, and SSMIS (September 2014)

Posteriors are unimodal, pointing to model's stability

Consistently unimodal

with moderate logit-space spread

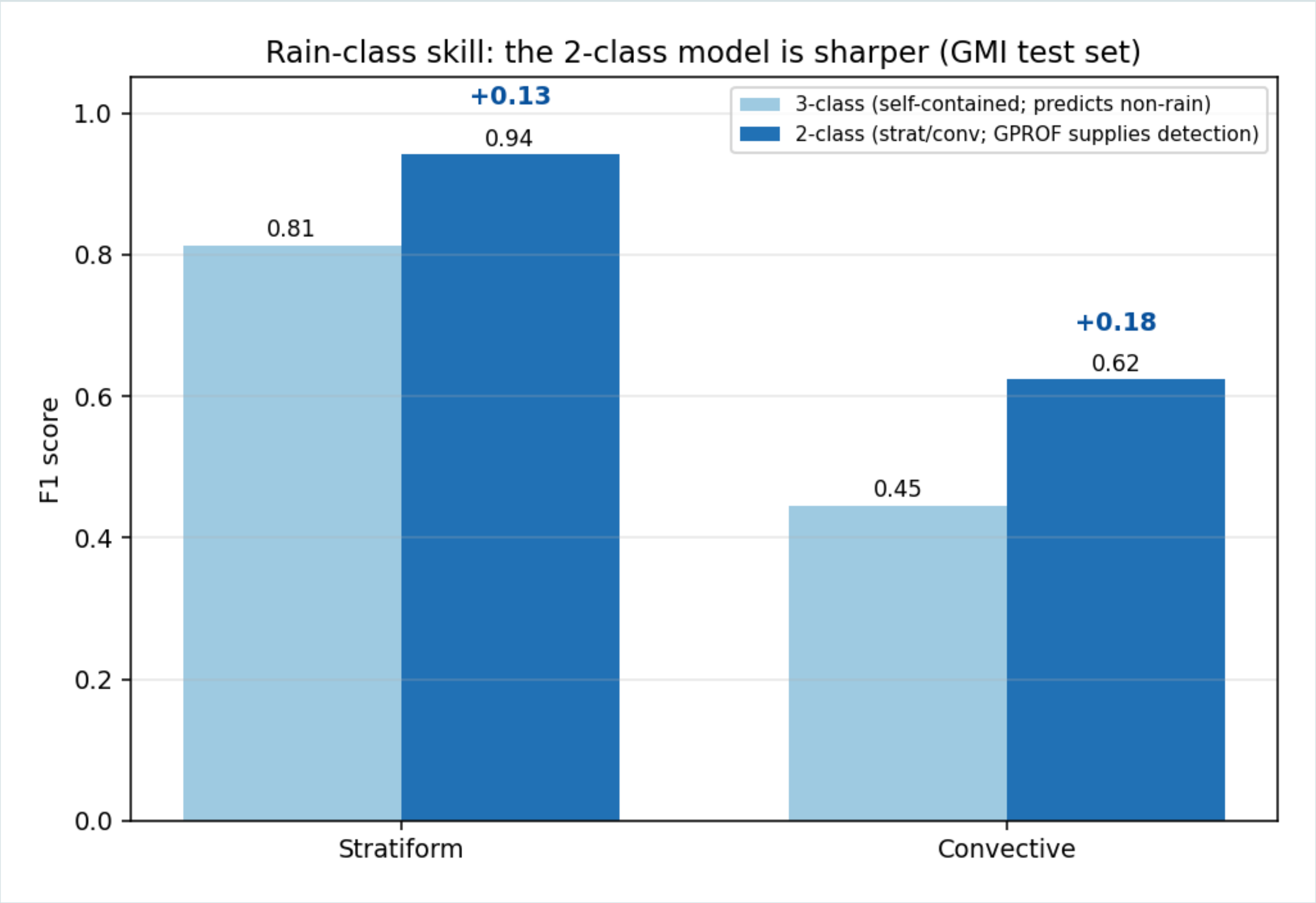
- A single, well-defined peak means the model is not flip-flopping between classes.
- Stable, predictable behaviour, suitable for operational use.



Illustrative; per-pixel posterior over the class logit

ANALYSIS

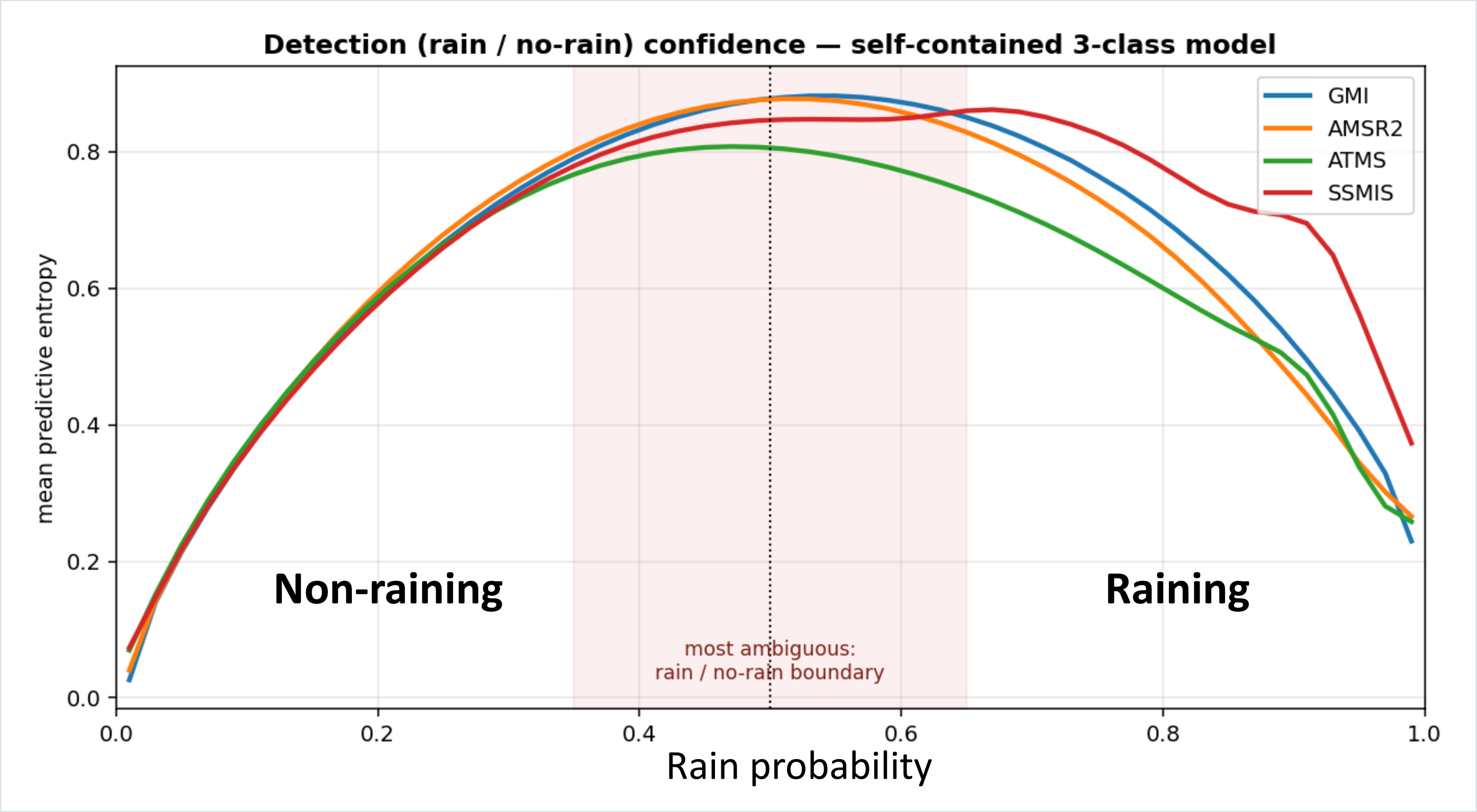
Label scheme is a design choice the uncertainty informs



Rain-class skill (GMI example): A 2-class scheme (uses external rain/no-rain mask) is more accurate on rain classes than a self-contained 3-class scheme.

ANALYSIS

The model flags its own rain/no-rain ambiguity



Self-contained 3-class model; predictive entropy peaks at P(rain) ≈ 0.5, which is a built-in confidence in whether it is raining at all.

IMPLICATIONS

What the decomposition tells us to do next

Aleatoric dominates: seek richer information

Add information content (more channels, active plus passive sensors), rather than more training data.

Epistemic locates the gaps

It pinpoints where targeted training-data augmentation would be most impactful, such as high-latitude regimes.

Stable posteriors signal readiness

Unimodal, moderate-spread posteriors support operational deployment.

Summary and next steps

- Decomposed uncertainty from a Bayesian U-Net is physically interpretable and spatially coherent
- Aleatoric error dominates by roughly two orders of magnitude; the limit is signal ambiguity, not training volume
- Epistemic error flags regime gaps (high latitudes); posteriors are unimodal, indicating operational stability

Next: Richer information content where aleatoric dominates; targeted augmentation where epistemic is high.

Veljko Petković ESSIC, University of Maryland Veljko.Petkovic@umd.edu

PMW-based retrieval

University of Maryland; Naval Postgraduate School

- Develop Passive Microwave retrieval of precipitation type
- Applicable to any GPM constellation member
- **Attach uncertainties to the product** (this talk)

IR-based retrieval (poster 6.15)

University of Oklahoma; GSFC NASA

- Develop InfraRed retrieval of precipitation type
- Applicable to all IR geostationary sensor
- Attach uncertainties to the product

Creating a Global High-Resolution Convective/Stratiform Type Product
University of Maryland BC and GSFC NASA
Jackson Tan (talk 8.9 on Thursday)