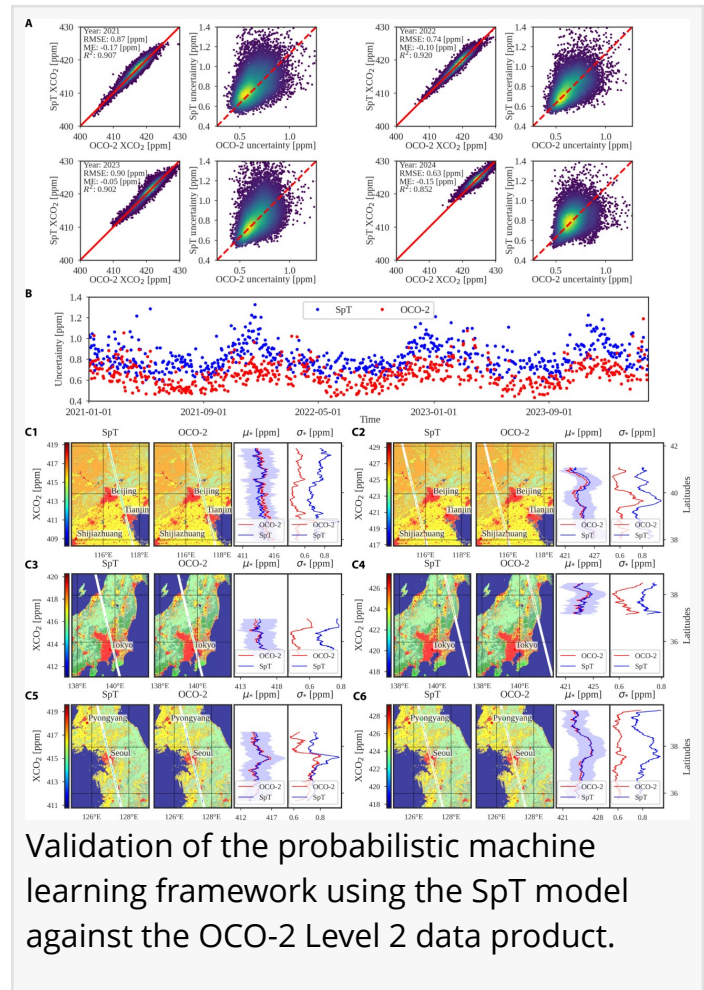# Fast AI for satellite learns to quantify uncertainty

GA, UNITED STATES, January 26, 2026 /EINPresswire.com/ -- Accurately tracking atmospheric greenhouse gases requires not only fast predictions but also reliable estimates of uncertainty. Researchers have developed a [lightweight machine learning framework](#) that enables satellite-based trace gas retrievals to quantify their own uncertainty while maintaining high computational efficiency. By combining probabilistic modeling with ensemble learning, the approach delivers rapid and trustworthy CO_2 measurements, helping bridge the gap between data-driven speed and scientific reliability.

Satellite missions generate massive volumes of atmospheric data that are essential for monitoring greenhouse gases and informing climate policy. Traditional physics-based retrieval algorithms provide reliable uncertainty estimates but are computationally intensive and struggle to scale with growing data streams. In contrast, machine learning methods offer

Validation of the probabilistic machine learning framework using the SpT model against the OCO-2 Level 2 data product.

dramatic speed improvements but typically produce only single-value predictions, lacking uncertainty information needed for scientific interpretation and decision-making. Existing probabilistic machine learning approaches often require heavy computation, complex tuning, or labeled uncertainty data, limiting their operational use. Based on these challenges, there is a strong need to develop scalable retrieval methods that combine machine learning efficiency with rigorous uncertainty quantification.

A research team from Shanghai Jiao Tong University reported this advance in Journal of Remote Sensing, published (DOI: 10.34133/remotesensing.0881) on December 26, 2025. The study introduces a probabilistic machine learning framework designed for satellite-based trace gas retrievals, with a focus on carbon dioxide monitoring. The work addresses a critical bottleneck in

current satellite data processing: how to rapidly analyze vast datasets while still providing uncertainty estimates required for climate science, data assimilation, and policy-relevant applications. The framework was validated using long-term observations from NASA's Orbiting Carbon Observatory-2 (OCO-2) mission.

The study demonstrates that uncertainty-aware machine learning can be achieved without sacrificing speed. By modifying neural networks to predict both expected values and associated uncertainty, the framework transforms standard deterministic models into probabilistic ones. When applied to satellite CO_2 retrievals, the method produced highly accurate concentration estimates while simultaneously quantifying uncertainty. Validation against OCO-2 operational products showed strong temporal and spatial consistency, with over 99% of reference values falling within the predicted uncertainty bounds. Importantly, the probabilistic model matched the accuracy of physics-based methods while operating thousands of times faster. This combination of speed, accuracy, and reliability marks a significant advance over existing machine learning approaches in atmospheric remote sensing.

The framework integrates two key innovations: likelihood-based learning and snapshot ensemble modeling. Instead of predicting a single output, the neural network simultaneously estimates the mean and variance of each retrieval, enabling direct uncertainty learning from existing satellite products. A Gaussian negative log-likelihood loss function ensures that both overconfident and underconfident predictions are penalized, encouraging well-calibrated uncertainty estimates.

To capture model-related uncertainty efficiently, the researchers employed snapshot ensembles, which extract multiple model instances from a single training run using cyclical learning rate scheduling. This avoids the heavy computational cost of training many independent models.

When tested on OCO-2 data from 2017 to 2024, the probabilistic model achieved retrieval speeds of milliseconds per satellite sounding, compared with minutes for traditional algorithms. Case studies over major cities showed that predicted uncertainty patterns closely followed those from physics-based retrievals, while remaining slightly more conservative—reflecting both measurement noise and model uncertainty.

"Uncertainty is not a luxury—it's essential for trustworthy climate data," said a member of the research team. "Our goal was to keep the speed advantages of machine learning while restoring the uncertainty information that scientists and policymakers rely on.This framework shows that we don't need complex or computationally expensive solutions to achieve reliable probabilistic predictions at scale."

The researchers adapted an existing Transformer-based neural network used for CO_2 retrievals by adding outputs for uncertainty estimation. Training data consisted of OCO-2 spectral measurements and corresponding retrieval products. The model was trained using a Gaussian likelihood loss and cyclical learning rate scheduling, with multiple snapshots collected to form an

ensemble. Performance was evaluated using independent multi-year satellite observations, statistical correlation analysis, and regional case studies over East Asia.

Beyond carbon dioxide, the framework can be applied to other satellite-based retrieval tasks, including methane monitoring and aerosol profiling. Its lightweight design makes it particularly well suited for next-generation Earth observation missions, where data volumes will continue to grow rapidly. By making uncertainty-aware machine learning practical at scale, the approach could improve climate monitoring, air quality assessment, and data-driven environmental decision-making worldwide.

DOI
10.34133/remotesensing.0881

Original Source URL
https://spj.science.org/doi/10.34133/remotesensing.0881

Lucy Wang
BioDesign Research
email us here

This press release can be viewed online at: https://www.einpresswire.com/article/886546269