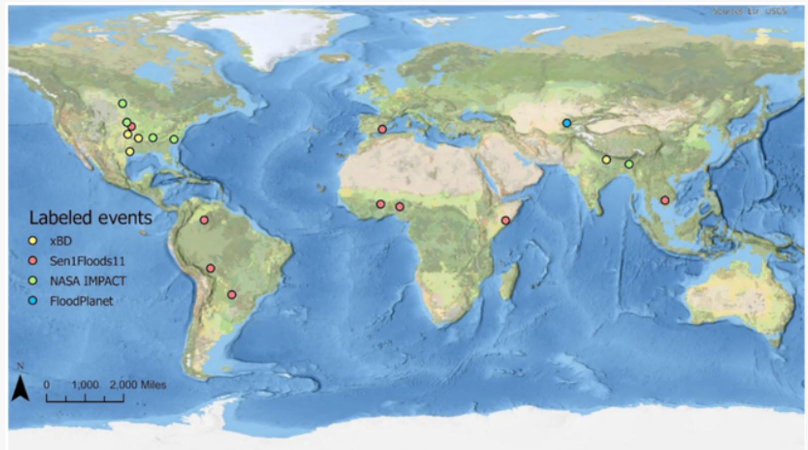
EIN PRESSWIRE

FloodPlanet dataset enhances global inundation monitoring

FAYETTEVILLE, GA, UNITED STATES, July 24, 2025 /EINPresswire.com/ -- A new high-resolution flood dataset, FloodPlanet, is enhancing satellite-based flood monitoring through more accurate training of deep learning models. By manually labeling inundation data from commercial satellites and aligning them with public sensor imagery, researchers improved flood detection accuracy by up to 15.6%.



Locations of the 19 FloodPlanet events.

<NWSLUG />Flooding affects more people globally than any other environmental hazard, yet accurate monitoring remains a challenge. Public satellite sensors often suffer from limited spatial resolution, revisit frequency, or cloud cover interference. While commercial satellites offer sharper imagery, their cost and restricted access hinder broad usage. Meanwhile, deep learning requires large, high-quality datasets to train models effectively. Most existing datasets use coarser labels that limit model accuracy. As interest grows in geo-foundation models and climate-adaptive infrastructure, the need for precise, accessible inundation data becomes pressing. Due to these challenges, a comprehensive high-resolution dataset is needed to advance satellite-based flood detection and machine learning capabilities.

Researchers from the University of Arizona and collaborators from institutions including NASA and Columbia University have developed FloodPlanet, a new global flood dataset released (DOI: 10.34133/remotesensing.0575) on May 15, 2025, in the Journal of Remote Sensing. The dataset addresses key limitations in satellite-based inundation monitoring by providing manual annotations derived from 3-meter resolution commercial imagery. Designed to improve flood detection by public sensors like Sentinel-1 and Sentinel-2, this effort enhances data quality for training deep learning models and supports the development of more robust disaster response systems.

The study found that models trained on FloodPlanet labels significantly outperformed those trained on lower-resolution public datasets. For example, using FloodPlanet data improved intersection-over-union (IoU) scores by up to 15.6% when evaluating flood detection via Sentinel-1 imagery. When tested on the same flood events, FloodPlanet-trained models consistently delivered more precise flood extent mapping, particularly in diverse ecoregions and complex terrain. The dataset enabled public sensors to achieve near-commercial accuracy levels, providing a cost-effective way to boost model performance without requiring continuous access to expensive satellite data. This innovation addresses a critical gap in the field: how to bridge high-resolution data advantages with publicly available resources.

FloodPlanet contains 366 manually labeled image chips from 19 global flood events between 2017 and 2020. Labels were created using 3-meter resolution PlanetScope imagery and co-aligned with Sentinel-1 and Sentinel-2 data. The team evaluated model performance through a leave-one-region-out cross-validation method, training a UNet-based deep learning model on public and commercial sensors. Models trained on FloodPlanet labels showed clear gains across all performance metrics. For instance, Sentinel-1 models improved IoU scores from 0.52 (NASA IMPACT) to 0.601, while Sentinel-2 models saw a jump from 0.571 (S1F11) to 0.624. PlanetScope models achieved a mean IoU of 0.691, outperforming both public sensors. Additionally, spatial analysis showed better results in vegetated and coastal regions, with lower performance in arid zones due to spectral confusion. The research also found that integrating even limited commercial data into model training can dramatically enhance performance, helping public-sector agencies and global researchers improve flood mapping at scale.

“Our goal was to make high-resolution flood data more accessible and impactful,” said co-author Dr. Zhijie Zhang. “Even without real-time commercial imagery, training public satellite models with FloodPlanet labels bridges the performance gap. It’s a scalable solution for global flood monitoring, particularly in vulnerable regions where timely, accurate information is vital for disaster response.”

The research team curated FloodPlanet by selecting diverse flood events across continents and ecoregions. Each event was represented by 1,024×1,024-pixel chips manually labeled using NASA’s ImageLabeler software, combining true- and false-color composites to identify water. Models were trained using a UNet architecture in PyTorch, with PlanetScope, Sentinel-1, and Sentinel-2 imagery resampled to match spatial resolutions. Performance was assessed through precision, recall, F1-score, and IoU, using cross-validation to ensure generalizability across unseen flood events.

FloodPlanet sets a new standard for training flood detection models with high-quality data. Its open-access format allows researchers worldwide to develop more accurate flood prediction systems, especially in regions underserved by commercial satellite access. The dataset could inform early warning systems, emergency response planning, and climate adaptation strategies. As foundation models for Earth observation evolve, integrating FloodPlanet may further unlock insights into hydrological extremes and accelerate the development of AI-driven environmental monitoring tools.

References
DOI
10.34133/remotesensing.0575

Original Source URL
https://spj.science.org/doi/10.34133/remotesensing.0575

Funding Information
This work was funded by the NASA CSDA Program (award number 80NSSC21K1163).

Lucy Wang
BioDesign Research
email us here
