Environmental Data Science (2026), 5: e13, 1–22
doi:10.1017/eds.2026.10041

CAMBRIDGE UNIVERSITY PRESS

METHODS PAPER

Precipitation nowcasting of satellite data using physically aligned neural networks

Antônio Catão, Leonardo Voltarelli, Melvin Poveda and Paulo Orenstein

Instituto de Matemática Pura e Aplicada, Rio de Janeiro, RJ, Brazil
Corresponding author: Leonardo Voltarelli; Email: leonardo.voltarelli@impa.br

Received: 22 October 2025 UTC; **Revised:** 09 March 2026 UTC; **Accepted:** 30 April 2026 UTC

Keywords: neural networks; physical conditioning; precipitation nowcasting; satellite data

Abstract

Accurate short-term precipitation forecasts predominantly rely on dense weather-radar networks, limiting operational value in places most exposed to climate extremes. We present TUPANN (Transferable and Universal Physics-Aligned Nowcasting Network), a satellite-only model trained on GOES-16 RRQPE. Unlike most deep learning models for nowcasting, TUPANN decomposes the forecast into physically meaningful components: a variational encoder–decoder infers motion and intensity fields from recent imagery under optical-flow supervision, a lead-time-conditioned MaxViT evolves the latent state, and a differentiable advection operator reconstructs future frames. We evaluate TUPANN on both GOES-16 and IMERG data, in up to four distinct climates (Rio de Janeiro, Manaus, Miami, La Paz) at 10–180-min lead times using the CSI and HSS metrics over 4–64 mm/h thresholds. Comparisons against optical-flow, deep learning, and hybrid baselines show that TUPANN achieves the best or second-best skill in most settings, with pronounced gains at higher thresholds. Training in multiple cities further improves performance, while cross-city experiments show modest degradation and occasional gains for rare heavy-rain regimes. The model produces smooth, interpretable motion fields aligned with numerical optical flow and runs in near real time due to the low latency of GOES-16. These results indicate that physically aligned learning can provide nowcasts that are skillful, transferable, and global.

Impact Statement

Extreme rainfall events are expected to increase in frequency and intensity, yet many high-risk regions lack the radar infrastructure behind state-of-the-art nowcasting. TUPANN lowers this barrier by using only globally available geostationary satellite imagery and a physics-aligned design that yields interpretable motion fields. It provides 10-minute refresh, transferable skills across diverse climates, and useful guidance with limited local data. This broadens access to timely storm and flood alerts, strengthens emergency management, and supports urban drainage and transport operations. Where radar and numerical prediction exist, TUPANN offers a complementary, low-latency layer. By reducing dependence on costly sensors while preserving skill and transparency, it advances more equitable, operational weather intelligence.

This research article was awarded Open Data and Open Materials badges for transparent practices. See the Data Availability Statement for details.

© The Author(s), 2026. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

1. Introduction

Extreme precipitation events are projected to become more frequent and intense under climate change, increasing the risk of floods and landslides, particularly in vulnerable regions (Marvel et al., 2023). Nowcasting—forecasting the atmosphere on time horizons up to 6 hour at high spatial resolution—is critical for early warnings and disaster management. While numerical weather prediction has improved steadily, its finite resolution and latency limit the accuracy of short-term precipitation forecasts. Radar-based nowcasting methods provide detailed observations but often require dense and well-maintained radar networks that are absent or degraded in much of the world. For example, Rio de Janeiro experiences recurrent flood-induced disasters yet lacks reliable radar coverage due to topographic blocking and limited infrastructure.

Recent advances in machine learning have shown that deep networks can outperform traditional numerical models in precipitation nowcasting when trained on high-resolution radar data. However, reliance on radar restricts their applicability to radar-rich regions, leaving large parts of South America, Africa, and Asia underserved. Furthermore, purely data-driven architectures often struggle with physical interpretability: they may produce realistic-looking precipitation maps while neglecting physically consistent motion fields, hindering forecasters' trust and operational uptake.

This paper addresses both accessibility and interpretability by leveraging geostationary satellites, which provide global coverage with near real-time latency, and by incorporating explicit physical structure into the neural network. We present TUPANN (Transferable and Universal Physics-Aligned Nowcasting Network), a model that uses only satellite-derived precipitation fields and decomposes the forecasting problem into physically motivated submodules. TUPANN comprises a variational encoder–decoder trained under optical-flow supervision to recover motion and intensity fields, a lead-time-conditioned transformer to evolve latent states, and a differentiable advection operator to reconstruct future frames. A strong physical alignment is done by explicitly penalizing the encoder-decoder output to match results from optical flow algorithms, which numerically infer motion fields. We evaluate TUPANN on data from four climate regimes—tropical rainforest (Manaus), subtropical highland (La Paz), tropical savanna with coastal influence (Rio de Janeiro), and tropical monsoon (Miami)—and report critical success indices (CSI) and Heidke skill scores (HSS) across lead times from 10–180 minutes and various precipitation thresholds. We benchmark against state-of-the-art baselines, including optical–flow methods (PySTEPS), deep learning models (Earthformer, CasCast), and hybrid approaches (NowcastNet).

Our main contributions are:

- We develop a physically aligned satellite-only nowcasting model that separates motion inference, latent dynamics, and advection. Unlike prior works that learn motion implicitly from final frame loss, our variational encoder–decoder is directly supervised by numerical optical flow, yielding smooth and interpretable motion fields.
- We leverage a lead-time-conditioned MaxViT transformer to evolve the latent representation and allow long lead-time prediction with a single network, reducing memory requirements compared with recurrent decoding.
- We perform extensive experiments on GOES-16's Rain Rate Quantitative Precipitation Estimation (RRQPE) and IMERG datasets in up to four cities with different climate regimes, comparing TUPANN with well-established and operational baselines. We demonstrate state-of-the-art CSI and HSS scores at multiple thresholds, analyze the effect of adding a generative adversarial network, and evaluate cross-city and multi-city training for transferability.
- We discuss operational considerations, including runtime and latency, and outline limitations and future directions for satellite-based nowcasting.

The remainder of this paper is structured as follows. Section 2 summarizes related work in numerical, optical-flow, deep learning, and satellite-only nowcasting. Section 3 describes the datasets and study regions.

Section 4 details the TUPANN architecture and its components. Section 5 presents our experimental design, baselines, and results. Section 6 discusses limitations and future work, and Section 7 concludes our work.

2. Related work

2.1. Numerical nowcasting methods

Precipitation nowcasting emerged in the late 1980s (Browning and Collier, 1989) and remains a fundamental tool to mitigate the impacts of extreme precipitation events (An et al., 2025). Early approaches relied primarily on Lagrangian extrapolation of radar echoes (Germann and Zawadzki, 2002), while later developments incorporated physical constraints and stochastic perturbations to enable ensemble-based probabilistic forecasts (Seed et al., 2013). Among these, PySTEPS (Pulkkinen et al., 2019) has become a widely adopted open-source Python library providing a reproducible platform for numerical nowcasting. It integrates multiple optical-flow algorithms, including Lucas–Kanade (Lucas and Kanade, 1981) and DARTS (Ruzanski et al., 2009), to estimate motion fields and applies the STEPS model (Bowler et al., 2006) for probabilistic extrapolation enhanced with downscaled numerical weather prediction input. Despite their interpretability and operational maturity, these numerical schemes typically experience a rapid decline in forecast skill with increasing lead time, and ensemble configurations such as STEPS can incur substantial computational cost.

2.2. Deep learning models

Deep learning (DL) approaches have recently achieved strong performance in precipitation nowcasting, often surpassing traditional numerical methods in both accuracy and scalability while enabling faster inference once trained. Early models include ConvLSTM (Shi et al., 2015) and PredRNN (Wang et al., 2023), which introduced convolutional and recurrent architectures to capture spatiotemporal dependencies in radar imagery. Transformer-based architectures have since extended this line of work. Earthformer (Gao et al., 2022) adapts the Transformer framework for general Earth-system forecasting through a modified Cuboid Attention mechanism that models three-dimensional spatial interactions. Similarly, Rainformer (Bai et al., 2022) extracts local and global features by combining a window-based multi-head self-attention with a gating mechanism component. These deterministic models, typically trained with pixel-wise L1 or L2 losses, target mean or median intensities and therefore tend to produce overly smooth forecasts.

To address this limitation, generative DL models have been proposed to better reproduce fine-scale variability by learning the underlying data distribution. The first class comprises Generative Adversarial Networks (GANs), including DGMR (Ravuri et al., 2021) and NowcastNet (Zhang et al., 2023), which have shown competitive performance on radar-based benchmarks. NowcastNet augments the GAN structure with an Evolution Network that estimates motion and intensity fields used to generate intermediate predictions before adversarial refinement. Despite its physics-inspired design, NowcastNet does not explicitly encode physical constraints and inherits known GAN instabilities, including mode collapse and artifact generation (Saxena and Cao, 2021).

More recently, diffusion-based generative models have emerged as stable alternatives to GANs, inspired by advances in computer vision. PreDiff (Gao et al., 2023) employs a latent diffusion model (LDM) with a knowledge-alignment mechanism that enforces domain-specific physical consistency during the sampling process. Evaluated on the SEVIR dataset (Veillette et al., 2020)—a combination of GOES-16 satellite imagery and NEXRAD radar data over the United States—PreDiff guides denoising steps toward physically plausible predictions by aligning generated intensities with those estimated from a time-series model applied to context-frame averages. CasCast (Gong et al., 2024) extends this approach through a cascaded LDM framework: it first conditions on deterministic Earthformer predictions and then refines them in latent space to generate high-resolution, small-scale structures. CasCast achieves superior results on SEVIR and other benchmarks, though it remains limited to radar data, short lead times (up to 1 hour), and lacks explicit physical regularization—sometimes producing noisy outputs and incurring high inference costs typical of diffusion models (Salimans and Ho, 2022).

2.3. Physically conditioned deep learning

Deep learning models are often regarded as black boxes, offering limited interpretability of the processes guiding their predictions. In geophysical applications, incorporating domain-specific physical knowledge can regularize training and promote physically consistent outputs. Physics-Informed Neural Networks (PINNs) (Raissi et al., 2019) exemplify this approach and have inspired numerous extensions and applications (Karniadakis et al., 2021; Kovacs et al., 2022; Wang et al., 2022; McClenny and Braga-Neto, 2023). Their core principle is to augment the loss function with a term enforcing that the network outputs satisfy a governing partial differential equation (PDE).

In precipitation nowcasting, several DL models include such physical conditioning. PID-GAN (Yin et al., 2024) combines a GAN framework with a physics-informed loss derived from the moisture conservation equation, trained on radar data. FourCastNet (Kurth et al., 2023) employs an Adaptive Fourier Neural Operator architecture (Guibas et al., 2021), a physics-inspired design widely used for PDE solutions, and applies it to global-scale forecasting with ERA5 reanalysis inputs. The previously discussed PreDiff and NowcastNet also incorporate elements of physical conditioning: NowcastNet's design draws inspiration from the continuity equation without numerical constraints, whereas PreDiff explicitly penalizes deviations from numerical-model intensities. The conditioning proposed in this work builds on both ideas by enforcing the continuity equation through an explicit loss between predicted physical terms and those derived from a numerical optical-flow method, applied to a specific module of the architecture. As shown later, this formulation achieves competitive predictive skill while improving interpretability and physical plausibility.

2.4. Use of satellite data

The use of satellite observations for precipitation nowcasting remains relatively limited. Several studies have explored this direction. Shukla et al. (2025) evaluated the use of PySTEPS with satellite imagery under different optical flow methods. Lebedev et al. (2019) employed a variant of the U-Net architecture (Ronneberger et al., 2015) to predict precipitation up to 2 hours ahead over Russia using EUMETSAT data. Rahimi et al. (2024) proposed a hybrid U-Net–ConvLSTM model evaluated on IMERG and Global Forecast System (GFS) data, using GFS as ground truth but without comparison against baseline models. More recently, Park et al. (2025) introduced a two-phase neural network that first predicts future satellite imagery using a video prediction model and then performs image-to-image translation to obtain radar reflectivity. Their approach leverages the Sat2Rdr dataset, derived from the Korean GK2A geostationary satellite and ten ground-based radar stations. While effective for light precipitation, reported results are limited to low critical success index (CSI) thresholds (below 8 mm/h), leaving high-intensity events mostly unassessed. Agrawal et al. (2025) also leverage geostationary satellite mosaics, providing global skillful forecasts up to 12 hours into the future. The model uses an encoder-decoder architecture with multiple high-dimensional inputs, including numerical weather prediction (NWP), leading to an intensive hardware use of more than 500 TPU cores during training.

Other studies, including Niu et al. (2024), Zheng et al. (2024), and Andrychowicz et al. (2023), integrate both satellite and radar data as inputs, which limits their applicability in radar-sparse regions. In contrast, the present work relies exclusively on satellite imagery, enabling global scalability. Comparisons are conducted against established baselines and evaluated across a range of precipitation intensities, including high and extreme-rate events.

3. Data and study regions

We train and evaluate TUPANN using precipitation data from satellite products. The primary data source is GOES-R and IMERG, but our work can also be extended to other satellites. For each dataset, we identify rain events and split them into training, validation, and test sets as described below.

3.1. GOES-16 RRQPE

GOES-R Advanced Baseline Imager Rain Rate Quantitative Precipitation Estimation (RRQPE) (GOES-R Algorithm Working Group and GOES-R Program Office, 2018). RRQPE provides precipitation estimates over the Americas every 10 min at a spatial resolution with a latency of approximately 5 min, enabling real-time nowcasting. This product is highly correlated with rain-related bands and has been validated against ground radars and the GPM CORRA dataset (Agrawal et al., 2025), highlighting the value of predicting geostationary satellite observations. We use RRQPE from January 2020 to December 2023. Rain events are defined as contiguous periods when the precipitation rate exceeds a chosen threshold (see Section 5.1 for details); we sample events uniformly at random and allocate 70% to training, 15% to validation, and 15% to testing. Figure 1 shows the proportion of observations above various thresholds for four different cities, while Figure 2 displays the accumulated precipitation and dataset splits.

3.2. IMERG

To test generalization across data sources, we also use the Integrated Multi-satellitE Retrievals for GPM (IMERG) Final Run product (Huffman et al., 2014). IMERG provides precipitation estimates every 30 min at $_{a}$ resolution and is widely used in remote sensing research. Its latency is about 3.5 months (the Early Run version has 4 hour latency), which precludes real-time use but offers an independent validation dataset. We extract IMERG data from January 2020 to December 2023, split it using the same event-based procedure, and consider only the Rio de Janeiro region. Figure 3 compares the proportion of heavy-rain observations and accumulated precipitation for IMERG.

3.3. Study regions

To evaluate model performance across different climates, we select four 512 km × 512 km subregions of the GOES-16 domain centered on Rio de Janeiro (Brazil), La Paz (Bolivia), Manaus (Brazil), and Miami (USA). These regions span coastal, high-altitude, rainforest, and subtropical environments. For IMERG,

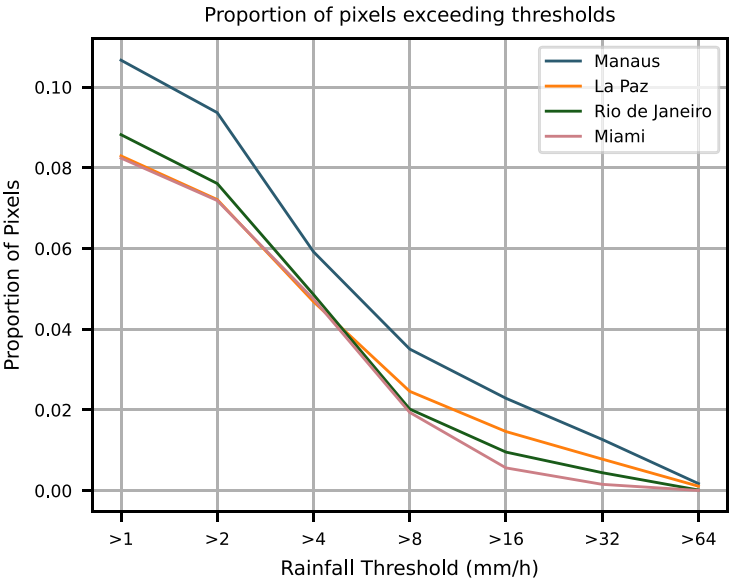


Figure 1. *Proportion of observations above different precipitation thresholds in the GOES-16 RRQPE dataset for each of the four study regions. Manaus exhibits the highest frequency of heavy rainfall across thresholds, while Rio de Janeiro and La Paz show intermediate levels.*

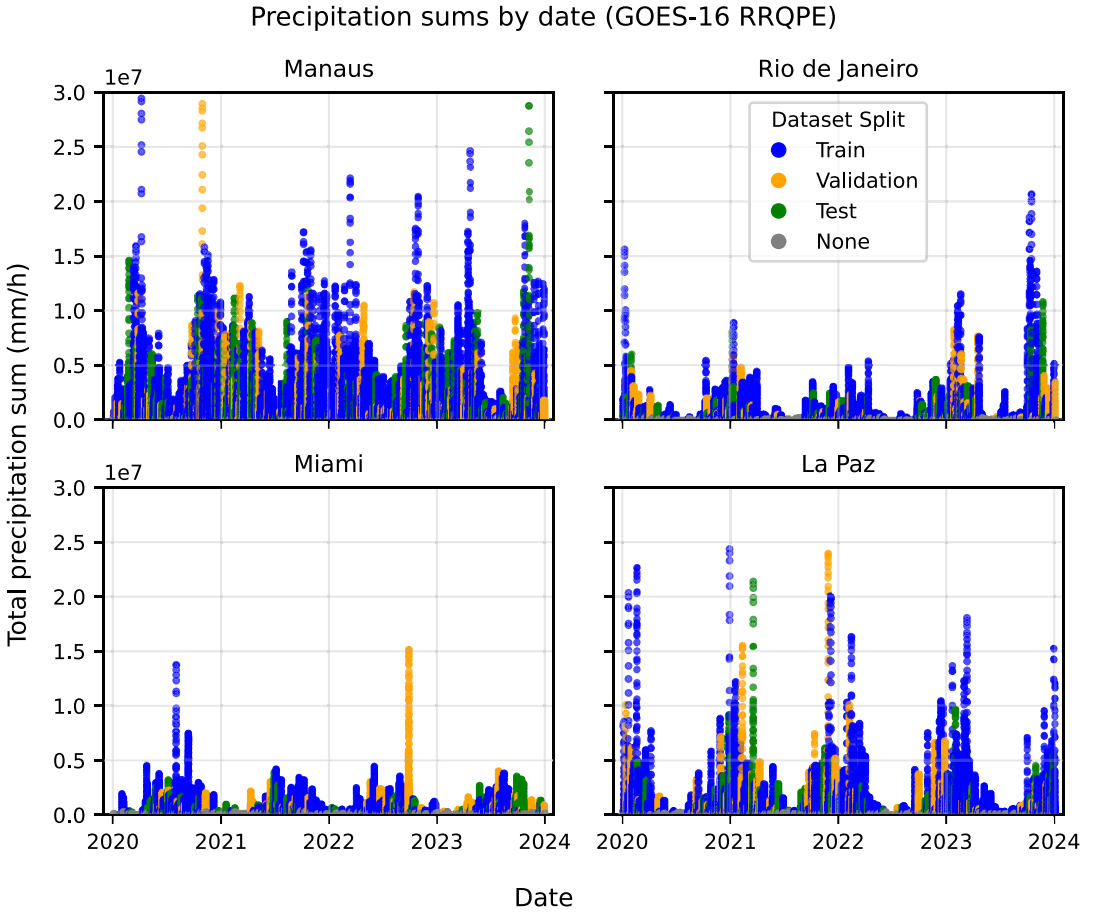
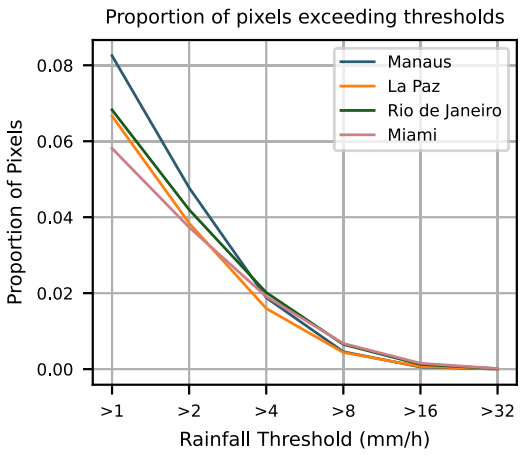


Figure 2. *Accumulated precipitation in GOES-16 RRQPE from January 2020 to December 2023 over each study region. Shaded areas denote training, validation, and test splits. Seasonal variability differs markedly between regions, with pronounced dry and wet seasons in La Paz and Rio de Janeiro.*

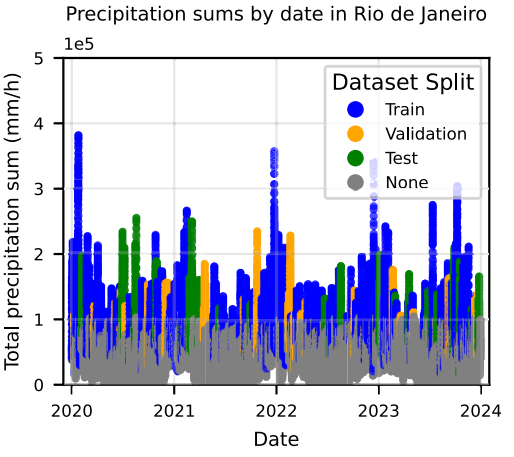
we consider only a 2560 km × 2560 km area surrounding Rio de Janeiro. The dominant precipitation processes include orographic and mesoscale convection in the Rio (Luiz-Silva and Oscar-Júnior, 2022), high-altitude convective storms in La Paz (Garreaud, 2001), monsoon-driven convection in Manaus (Oliveira et al., 2016) and sea-breeze thunderstorms in Miami (Burpee, 1979). Figure 2 illustrates the seasonal cycle and dataset splits for GOES-16 across regions.

4. Methods

The proposed model, TUPANN, forecasts precipitation fields from a sequence of past satellite images $X_{-T:0} \in \mathbb{R}^{(T+1)\times n\times n}$, where n denotes the spatial resolution (i.e., number of pixels per dimension) and $T+1$ is the number of past observations. It produces predicted fields $\widehat{X}_{1:T_f} \in \mathbb{R}^{T_f\times n\times n}$, where T_f is the forecast horizon. The architecture comprises two modules: (i) a variational encoder–decoder (VED) that learns a latent representation $L_k, k = 1,\ldots,T_f$ of the evolution of the precipitation fields under optical flow supervision; (ii) a visual transformer (MaxViT) that evolves the latent representation L_k so that the resulting application of a differentiable advection operator (warp) on the decoded motion and intensity fields, $\widehat{s}_{k-1\to k}$ and $\widehat{v}_{k-1\to k}$, closely match the ground truth frame. See Figure 4. The training procedure is sequential: the VED module is initially trained to infer the first set of motion and intensity fields, then its



(a) Proportion of observations above different precipitation thresholds in the IMERG dataset for each study region. Unlike RRQPE, Manaus only dominates the lowest thresholds.



(b) Estimated precipitation from IMERG (2020–2023) over Rio de Janeiro. The data splits are color-coded; low-level precipitation events are excluded from training and evaluation.

Figure 3. *Statistics of IMERG data, highlighting differences with respect to the RRQPE dataset.*

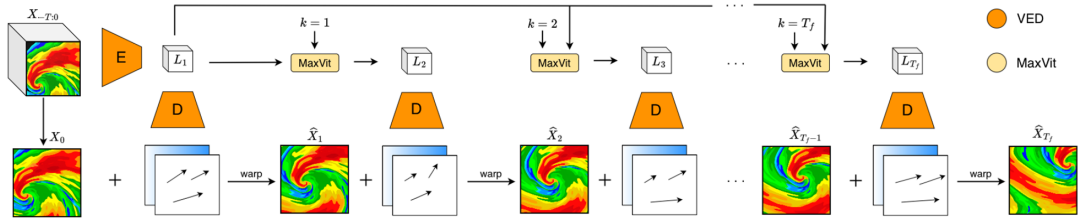


Figure 4. *TUPANN architecture. The VED and MaxViT modules displayed are learned; motion fields and the final predictions are extrapolated through a warp function.*

weights are fixed and used for the training of the MaxVit module. Details on the VED and MaxViT training are discussed in Sections 4.1 and 4.2, respectively.

Even though MaxViT offers linear complexity in the image size, the choice of such encoder–decoder architecture is guided by the idea that, apart from compressing the images, the VED is also responsible for learning the dynamics of a single-step evolution. MaxViT, on the other hand, learns to extrapolate the dynamics to further lead times.

4.1. Variational encoder–decoder

We use a variational encoder–decoder to learn an efficient representation of the precipitation evolution. Instead of reconstructing the input images as in classical variational autoencoders, our VED outputs motion fields $\widehat{v}_{0\to1} \in \mathbb{R}^{2\times n\times n}$ and intensity corrections $\widehat{s}_{0\to1} \in \mathbb{R}^{n\times n}$ given the past sequence $X_{-T:0}$. To enforce physically plausible motion, we compute the ground truth motion fields $v_{0\to1}$ applying an optical flow algorithm to $X_{-\tilde{T}+2:1}$, where $\tilde{T}$ is the context length provided to the optical flow algorithm. After that, the ground truth intensity correction $s_{0\to1}$ is obtained by subtracting the advected frame $\tilde{X}_1$, obtained using $v_{0\to1}$, from the true frame X_1 (see Figure 5). Thus, the estimated fields $\widehat{v}_{0\to1}$ and $\widehat{s}_{0\to1}$ can be used to extrapolate the last observed frame X_0 to an estimate $\widehat{X}_1$ of the next frame via an advection operator (see Section 4.2.2).

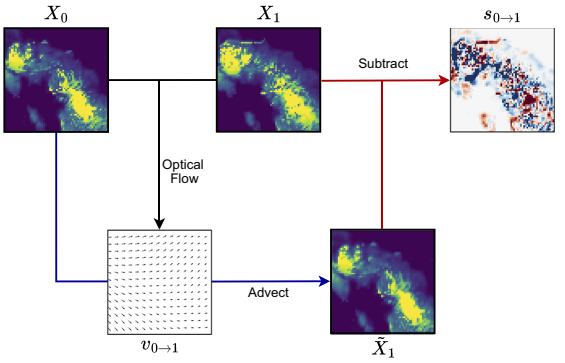


Figure 5. *Ground truth motion fields are obtained using an optical flow method from a pair of past and future images. The past image is advected to obtain an intermediate one, $\tilde{X}_1$. Finally, ground truth intensity is the subtraction of $\tilde{X}_1$ from the future image.*

4.1.1. Target loss

To supervise the predicted $\widehat{v}_{0\to1}$, ℓ_1 and cosine-similarity losses are used with respect to $v_{0\to1}$. The intensity discrepancies between $s_{0\to1}$ and $\widehat{s}_{0\to1}$ are penalized via ℓ_1 loss. Finally, a Kullback–Leibler divergence term is added to ensure the regularity of the learned latent space. The VED loss is

$$\begin{aligned}\text{Loss}_{\text{VED}}(\widehat{s}_{0\to1},\widehat{v}_{0\to1},s_{0\to1},v_{0\to1}) := \lambda_{\text{int}}\ell_1(s_{0\to1},\widehat{s}_{0\to1}) + \lambda_{\text{motion}}\ell_1(v_{0\to1},\widehat{v}_{0\to1}) \\ +\lambda_{\cos}\text{CosSimilarity}(v_{0\to1},\widehat{v}_{0\to1}) + \lambda_{\text{KL}}KL(p,\widehat{p}_\theta).\end{aligned} \tag{1}$$

Here, $\widehat{p}_\theta$ is the latent distribution inferred by the encoder, p is a standard normal distribution and $\lambda_{\text{int}},\lambda_{\text{motion}},\lambda_{\cos},\lambda_{\text{KL}}$ are hyperparameters tuned on the validation set. This loss encourages accurate motion fields, intensity corrections, and latent regularity.

4.1.2. Optical flow

An optical flow algorithm is able to infer motion fields between two images, and is thus essential to obtain the derived ground-truth motion and intensity fields (e.g., $v_{0\to1},s_{0\to1}$) used in (1). We consider two options: Lucas–Kanade (LK), which solves a local least-squares problem under the assumption of small displacements, and DARTS, a spectral method tailored to radar imagery that solves the optical-flow equation in Fourier space. For TUPANN, we have taken the choice of optical flow method as a hyperparameter; below, we use DARTS for GOES-16 and LK for IMERG results.

4.2. MaxViT

Given the latent representation L_1 from the VED, a visual transformer evolves the latent state forward in time. We adopt MaxViT (Tu et al., 2022), which combines local and grid attention to efficiently capture global context while avoiding quadratic attention cost.

4.2.1. Lead time conditioning

To predict the latent state at the lead time k, we condition the transformer on k via one-hot encoding and linear embedding, yielding $L_k = \text{VT}(L_1,k)$, $k=2,\ldots,T_f$, where where $\text{VT}(\cdot)$ represents the MaxViT model. Unlike recurrent decoding, this conditioning enables the same transformer to produce all lead times while reducing memory overhead (Andrychowicz et al., 2023). Applying the VED decoder to L_k yields motion and intensity fields $(\widehat{v}_{k-1\to k},\widehat{s}_{k-1\to k}) = D(L_k)$, where $D(\cdot)$ is the decoder module of the VED. Thus, all the necessary elements to predict the sequence recursively are obtained.

4.2.2. Warp function

Following NowcastNet (Zhang et al., 2023), we implement a fixed, differentiable advection operator that can reconstruct future precipitation frames using the predicted motion and intensity fields. Thus, given a frame $\widehat{X}_{k-1}$, motion field $\widehat{v}_{k-1\to k}$ and intensity field $\widehat{s}_{k-1\to k}$, the extrapolated frame $\widehat{X}_k$ is

$$\widehat{X}_k = \mathrm{warp}\left(\widehat{v}_{k-1\to k}, \widehat{s}_{k-1\to k}, \widehat{X}_{k-1}\right). \tag{2}$$

Roughly, following Zhang et al. (2023), the warp function takes the image $\widehat{X}_{k-1}$ and applies the motion vector field $\widehat{v}_{k-1\to k}$ pointwise; finally, the intensity field $\widehat{s}_{k-1\to k}$ is added to the result.

4.2.3. Target loss

We compute the target loss in the original image space. We assume that $\widehat{X}_{k-1} = X_{k-1}$ in equation (2) to avoid a costly recursive loss and calculate the ℓ_1 loss between the warped frame $\widehat{X}_k$ and the observed frame X_k. Thus,

$$\mathrm{Loss}_{\mathrm{MaxViT}} = \ell_1(\mathrm{warp}(\widehat{v}_{k-1\to k}, \widehat{s}_{k-1\to k}, X_{k-1}), X_k). \tag{3}$$

The gradients of this loss will flow through the VED decoder module and the MaxViT transformer. The VED is pre-trained separately; thus, optimizing this loss only affects the MaxViT modules.

5. Experiments and results

We compare TUPANN against four nowcasting benchmarks using several evaluation metrics and across regions with different climates. We also run several ablation experiments.

5.1. Rain events dataset selection

The datasets used to train and evaluate our models comprise a subsample of rainy windows drawn from either the GOES-16 RRQPE or the IMERG products. We define a rainy window as follows. For each 10-min timestamp t from 2020-01-01 00:00 UTC to 2023-12-31 23:50 UTC, we (i) form a symmetric 60-min window $[t-30\,\mathrm{min}, t+30\,\mathrm{min}]$; (ii) compute the spatiotemporal precipitation accumulation over that window (10-min steps over all grid points); and (iii) if the accumulation exceeds a threshold τ, label the larger window $[t-4\,\mathrm{hours}, t+4\,\mathrm{hours}]$ as a rainy window. Finally, we merge rainy windows that intersect. The threshold $\tau = 120{,}000$ was chosen empirically to balance excluding near-dry periods against obtaining a dataset large enough for effective learning. Using a symmetric ± 4-hour window ensures that events include both onset and dissipation phases (from no precipitation to mild or heavy precipitation and back).

5.2. Evaluation framework

We evaluate TUPANN and baselines using the critical success index (CSI) and Heidke skill score (HSS). Both metrics depend on a choice of threshold t, so that pixel values above the threshold are assigned as positive and, otherwise, negative. CSI and HSS are given by

$$\mathrm{CSI}_t = \frac{\mathrm{TP}_t}{\mathrm{TP}_t + \mathrm{FN}_t + \mathrm{FP}_t}, \qquad \mathrm{HSS}_t = \frac{2(\mathrm{TP}_t \times \mathrm{TN}_t - \mathrm{FN}_t \times \mathrm{FP}_t)}{(\mathrm{TP}_t + \mathrm{FN}_t)(\mathrm{TN}_t + \mathrm{FN}_t) - (\mathrm{TP}_t + \mathrm{FP}_t)(\mathrm{TN}_t + \mathrm{FP}_t)},$$

where TP stands for true positive, TN for true negatives, FN for false negatives and FP for false positives. Thus, CSI quantifies the overlap between forecasted and observed precipitation, ignoring true negatives (i.e., disregarding correct predictions of no precipitation) while HSS compares the forecast against random chance.

We report CSI and HSS at thresholds of 4, 8, 16, 32 and 64 mm/h and compute both pixel-wise scores (POOL1) and max-pooled scores over 4×4 blocks (POOL4). For aggregated metrics, we denote CSI–M and HSS–M as the mean across all thresholds. CSI values are reported in Section 5.4, and HSS ones are included in the Supplementary Materials.

5.3. Baseline models and tuning

In our experiments, TUPANN is compared with four baselines from different nowcasting paradigms:

- PySTEPS (LK) and PySTEPS (DARTS) (Pulkkinen et al., 2019): optical-flow baselines that estimate a motion field (LK: local Lucas–Kanade; DARTS: DFT-based spectral) from recent frames and then semi-Lagrangianly advect the precipitation field forward. As purely physical extrapolation methods, they are strong for very short lead times;
- Earthformer (Gao et al., 2022): a space–time Transformer for Earth-system data that uses Cuboid Attention (local block attention with global tokens) in a hierarchical encoder–decoder to predict future frames;
- NowcastNet (Zhang et al., 2023): a hybrid model combining a U-Net–based learnable semi-Lagrangian advection (Evolution Network) with a physics-conditioned generative network trained with a temporal discriminator to inject high-resolution convective structure;
- CasCast (Gong et al., 2024): a cascaded scheme that first uses a deterministic predictor (e.g., Earthformer) to capture mesoscale evolution, then conditions a latent-space diffusion transformer on that coarse forecast to generate small-scale features and improve extreme-precipitation skill.

For TUPANN and Earthformer, we tune learning rate, dropout rate, and loss weights on the validation set by maximizing mean CSI in the city of Rio de Janeiro, and use these for the other cities. The hyperparameters for NowcastNet and CasCast are those proposed in their original papers. The Supplementary Material to this article collects the final choice of hyperparameters for Earthformer and TUPANN. We trained each model for at most 20–300 epochs depending on the model, and chose the configuration that resulted in the highest CSI (averaged over thresholds). The optimizer for all models is Adam. After selecting the best values, we retrain on the combined training and validation data and evaluate on a held-out test set. Training uses a single NVIDIA A100 GPU, and inference typically takes under 2 seconds per forecast (for all 18 lead times).

5.4. GOES-16 results

Table 1 presents CSI scores across the four study regions and thresholds. TUPANN consistently ranks first or second for each metric. In Rio de Janeiro, it achieves the highest CSI at all thresholds; NowcastNet is the closest competitor, followed by Earthformer. In Miami, TUPANN again dominates most metrics, but now CasCast performs well. In Manaus and La Paz, Earthformer and NowcastNet obtain slightly better CSI for low thresholds, but TUPANN leads for higher thresholds and is therefore better at forecasting extreme rainfall. The 64 mm/h CSI values are small across models, reflecting the rarity of such intense events, yet TUPANN's scores remain the highest. Overall, the table highlights TUPANN's performance across different climate regimes and precipitation thresholds. Similar results are true for HSS (see Supplementary Material), although by that metric, Earthformer is much more competitive.

The graphs in Figure 6 show mean CSI (averaged across thresholds) versus lead time. TUPANN maintains the highest or second-highest skill across all lead times; the advantage over NowcastNet grows for early lead times, reflecting the benefit of explicit motion supervision and the efficiency of lead-time conditioning (see also Figure 7).

Beyond aggregated metrics, a large-scale illustration of TUPANN's prediction for a rain event in Manaus, compared with NowcastNet, CasCast, and Earthformer, is included in the Supplementary Material. Generative models such as NowcastNet and CasCast produce detailed textures but may introduce artifacts, whereas TUPANN and Earthformer yield smoother predictions. Despite the blurred appearance, TUPANN captures the timing and location of heavy rain more accurately, leading to higher CSI values.

Table 1. *Aggregated CSI metrics for GOES-16 data across cities*

	CSI-M↑		CSI_4↑		CSI_8↑		CSI_{16}↑		CSI_{32}↑		CSI_{64}↑	
Model	POOL1	POOL4	POOL1	POOL4	POOL1	POOL4	POOL1	POOL4	POOL1	POOL4	POOL1	POOL4
	Rio de Janeiro											
Earthformer	0.237	0.222	0.326	0.320	0.287	0.236	0.326	0.312	0.238	0.237	0.009	0.006
NowcastNet	0.244	0.269	0.313	0.374	0.282	**0.293**	0.318	0.325	0.247	0.278	0.059	0.074
PySTEPS (LK)	0.165	0.169	0.242	0.262	0.212	0.195	0.226	0.226	0.142	0.156	0.005	0.008
PySTEPS (DARTS)	0.166	0.166	0.231	0.243	0.216	0.191	0.229	0.228	0.140	0.152	0.013	0.015
CasCast	0.170	0.187	0.308	0.343	0.205	0.249	0.164	0.162	0.159	0.156	0.016	0.027
TUPANN (ours)	**0.259**	**0.277**	**0.330**	**0.384**	**0.289**	0.289	**0.330**	**0.336**	**0.274**	**0.287**	**0.072**	**0.090**
	Miami											
Earthformer	0.141	0.126	**0.274**	0.270	0.180	0.160	0.154	0.122	0.097	0.078	0.000	0.000
NowcastNet	0.137	0.160	0.248	0.299	0.170	0.207	0.128	0.139	0.097	0.106	0.040	0.047
PySTEPS (LK)	0.113	0.116	0.188	0.202	0.133	0.136	0.120	0.111	0.079	0.079	0.045	0.053
PySTEPS (DARTS)	0.112	0.113	0.189	0.202	0.135	0.138	0.118	0.107	0.073	0.071	0.044	0.046
CasCast	0.146	0.170	0.258	0.298	0.188	**0.229**	0.144	0.167	0.117	0.135	0.020	0.028
TUPANN (ours)	**0.169**	**0.187**	0.267	**0.312**	**0.189**	0.211	**0.177**	**0.177**	**0.135**	**0.141**	**0.079**	**0.094**
	Manaus											
Earthformer	0.276	0.256	**0.355**	0.341	**0.323**	0.297	**0.316**	0.292	0.265	0.245	0.124	0.104
NowcastNet	0.253	0.278	0.323	0.366	0.296	0.324	0.283	0.303	0.233	0.258	0.130	0.137
PySTEPS (LK)	0.200	0.196	0.258	0.266	0.237	0.233	0.218	0.212	0.160	0.156	0.125	0.112
PySTEPS (DARTS)	0.197	0.194	0.259	0.268	0.239	0.235	0.219	0.213	0.158	0.154	0.109	0.099
CasCast	0.265	0.286	0.344	**0.377**	0.303	**0.333**	0.295	0.307	0.260	0.269	0.126	0.141
TUPANN (ours)	**0.290**	**0.293**	0.339	0.367	0.316	0.321	0.315	**0.312**	**0.278**	**0.274**	**0.200**	**0.193**
	La Paz											
Earthformer	0.303	0.270	**0.337**	0.312	**0.329**	0.281	**0.359**	0.319	0.323	0.291	0.167	0.146
NowcastNet	0.291	0.301	0.330	**0.376**	0.303	0.321	0.321	0.315	0.300	0.296	0.202	0.197
PySTEPS (LK)	0.212	0.208	0.248	0.250	0.243	0.230	0.247	0.237	0.197	0.195	0.126	0.131
PySTEPS (DARTS)	0.225	0.218	0.263	0.262	0.262	0.244	0.264	0.250	0.206	0.201	0.127	0.132
CasCast	0.228	0.235	0.309	0.337	0.251	0.270	0.245	0.237	0.232	0.222	0.101	0.111
TUPANN (ours)	**0.314**	**0.317**	0.336	0.363	0.327	**0.323**	0.350	**0.340**	**0.327**	**0.322**	**0.232**	**0.239**

Note: Bold values denote the best, underlined values the second best. TUPANN obtains state-of-the-art performance at most thresholds and regions, particularly for high rain-rate events.

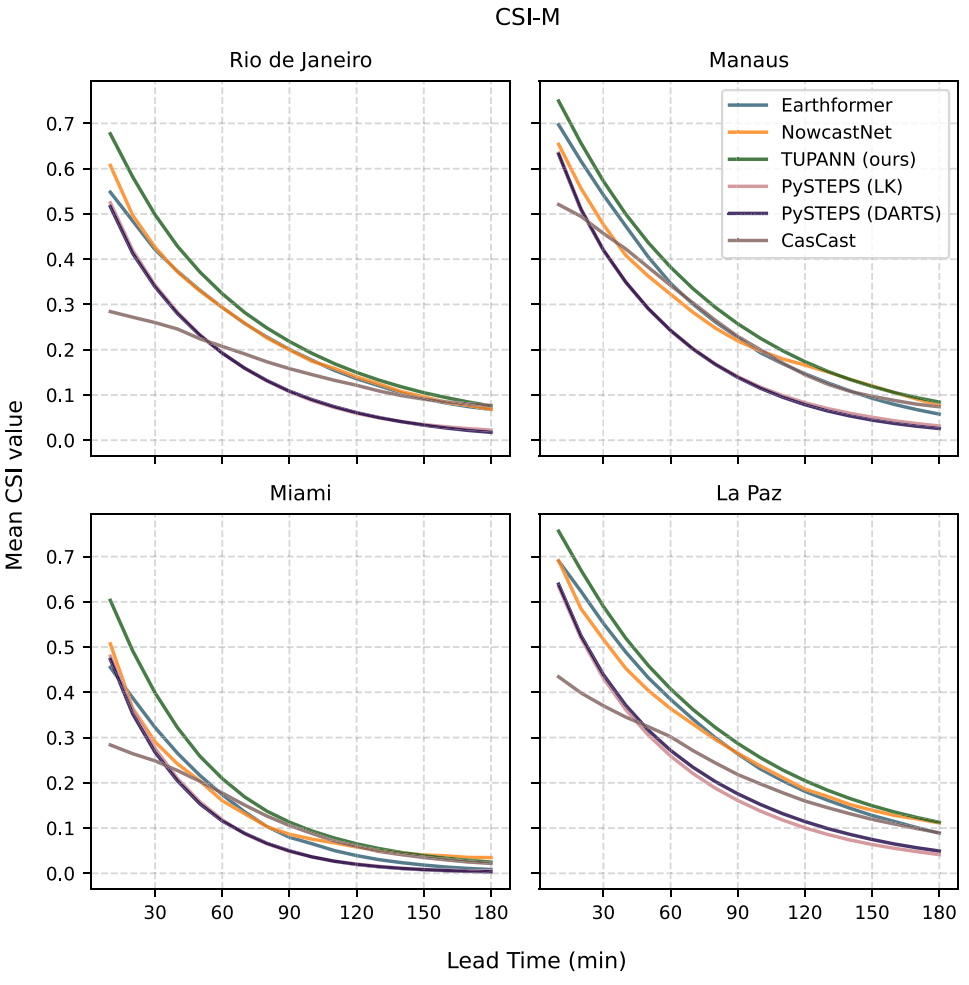


Figure 6. *Mean CSI (CSI–M) versus lead time for the four study regions using GOES-16 data. TUPANN consistently outperforms baselines across lead times.*

5.5. Ablation results

We will study how different experiments affect TUPANN: investigating its motion fields against another physics-DL hybrid model, adding a GAN head, evaluating TUPANN's cross-city generalization, and training the model jointly on all cities.

5.5.1. Interpretability and motion fields

TUPANN's interpretability stems from its explicitly learned motion fields. Figure 7 compares motion fields predicted by TUPANN and NowcastNet (which relies on an Evolution Network submodule for its motion fields). The TUPANN fields are smooth and closely resemble the numerical optical flow computed by DARTS, whereas the baselines' fields exhibit unrealistic patterns. This underscores the benefit of supervising motion fields directly.

5.5.2. GAN-TUPANN

Generative adversarial networks can improve visual realism at the cost of evaluation metrics. To study its impact on TUPANN predictions, we evaluate the variant GAN-TUPANN, which adds a GAN head to TUPANN outputs.

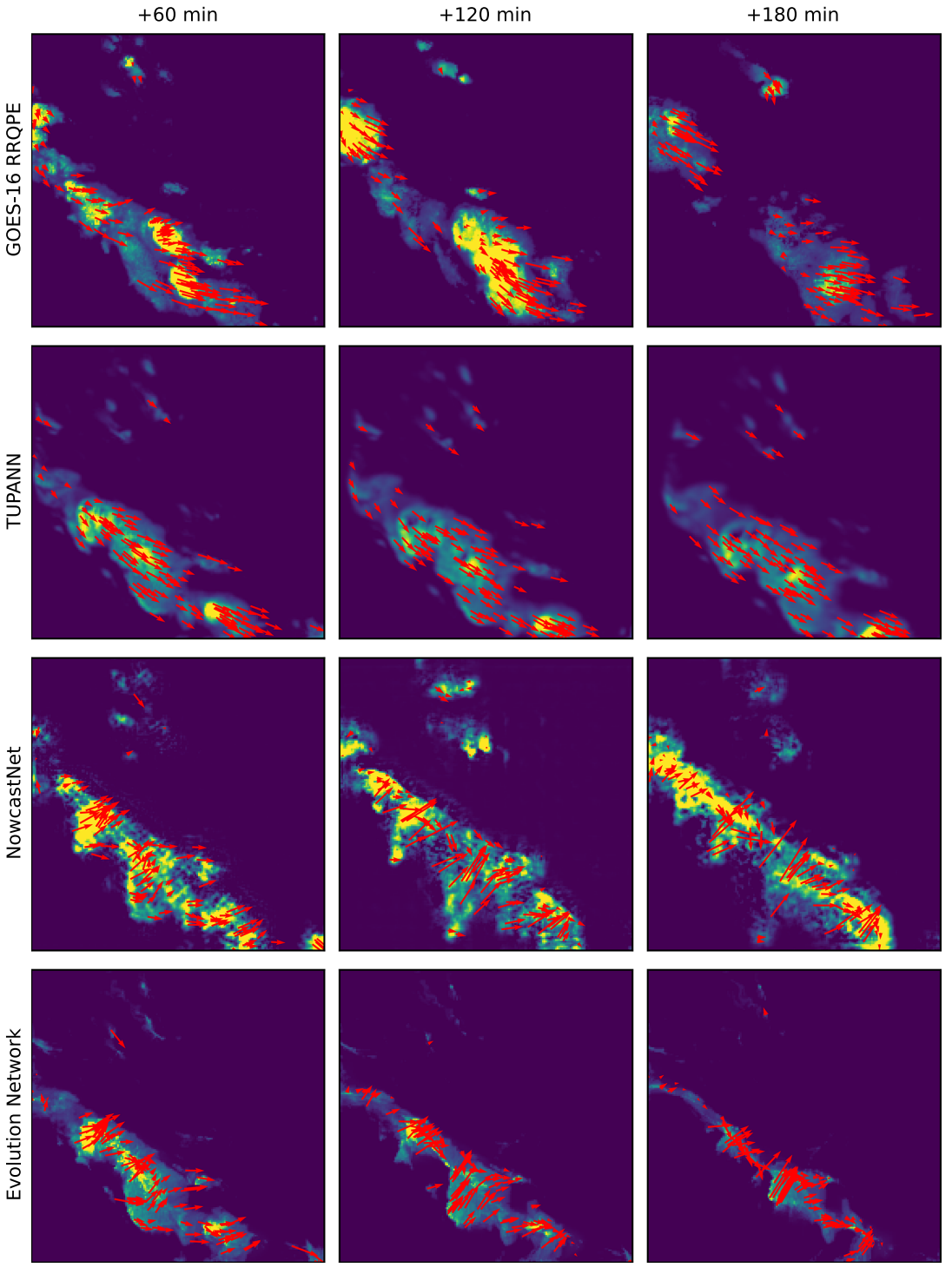


Figure 7. *Comparison of motion fields. Top row: GOES-16 RRQPE observations (ground truth) overlaid by ground-truth DARTS motion fields for future frames. Subsequent rows: learned motion fields from TUPANN, NowcastNet, and Evolution Network. In all frames, red arrows represent the motion fields. TUPANN yields smoother fields that align with physical intuition.*

Figure 8 shows that GAN-TUPANN produces significantly sharper images. Still, Table 2 shows this does not always lead to improvements in CSI scores. While in Rio de Janeiro, GAN-TUPANN increases low-threshold CSI, in Miami, the gains are either non-existent or negative. Given the computational overhead and mixed impact on metrics, there is no clear advantage in including this module unless visual fidelity is paramount.

5.5.3. Cross-city generalization

To assess generalization we train TUPANN on one city and evaluate on others. Table 3 compares TUPANN trained on Rio de Janeiro (TUPANN–Rio) against models trained separately on each city. In Manaus and La Paz, TUPANN–Rio yields lower CSI at all thresholds, as expected due to climate differences. Surprisingly, in Miami, the Rio-trained model performs comparably or better at high

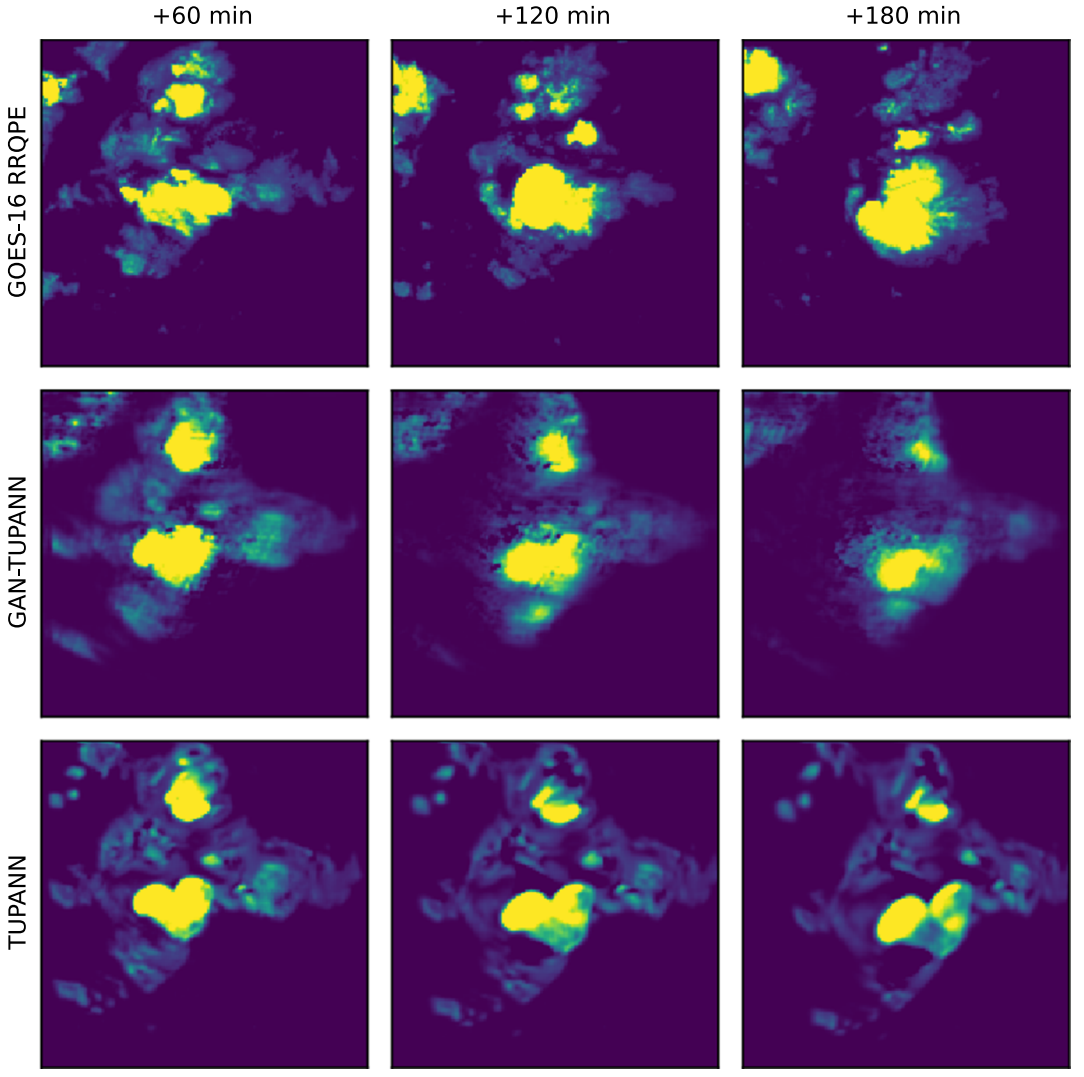


Figure 8. *Visual comparison of TUPANN and GAN-TUPANN for a rain event in Rio de Janeiro, starting at 2023-03-11 03:00 UTC. The first row shows the GOES-16 RRQPE observations (ground truth). GAN-TUPANN reduces blur, but yields mixed changes in CSI (Table 2).*

Table 2. *CSI metrics comparing TUPANN and its GAN variant (GAN-TUPANN) on GOES-16 data*

	CSI-M↑		CSI_4↑		CSI_8↑		CSI_{16}↑		CSI_{32}↑		CSI_{64}↑	
Model	POOL1	POOL4	POOL1	POOL4	POOL1	POOL4	POOL1	POOL4	POOL1	POOL4	POOL1	POOL4
	Rio de Janeiro											
NowcastNet	0.244	0.269	0.313	0.374	0.282	0.293	0.318	0.325	0.247	0.278	0.059	0.074
TUPANN	0.259	0.277	0.330	0.384	0.289	0.289	0.330	0.336	**0.274**	**0.287**	**0.072**	**0.090**
GAN-TUPANN	**0.265**	**0.290**	**0.350**	**0.413**	**0.296**	**0.319**	**0.342**	**0.348**	0.270	0.286	0.070	0.086
	Miami											
NowcastNet	0.137	0.160	0.248	0.299	0.170	0.207	0.128	0.139	0.097	0.106	0.040	0.047
TUPANN	**0.169**	**0.187**	**0.267**	**0.312**	**0.189**	**0.211**	**0.177**	**0.177**	**0.135**	**0.141**	**0.079**	**0.094**
GAN-TUPANN	0.152	0.174	0.252	0.300	0.180	**0.211**	0.160	0.164	0.116	0.128	0.052	0.066
	Manaus											
NowcastNet	0.253	0.278	0.323	0.366	0.296	**0.324**	0.283	0.303	0.233	0.258	0.130	0.137
TUPANN	**0.290**	**0.293**	**0.339**	**0.367**	0.316	0.321	**0.315**	0.312	**0.278**	**0.274**	**0.200**	**0.193**
GAN-TUPANN	0.274	0.285	0.331	0.359	**0.318**	**0.329**	0.310	**0.313**	0.258	0.267	0.156	0.156
	La Paz											
NowcastNet	0.291	0.301	0.330	**0.376**	0.303	0.321	0.321	0.315	0.300	0.296	0.202	0.197
TUPANN	**0.314**	**0.317**	0.336	0.363	0.327	0.323	**0.350**	**0.340**	**0.327**	**0.322**	**0.232**	**0.239**
GAN-TUPANN	0.306	0.312	**0.336**	0.367	**0.327**	**0.333**	0.344	0.336	0.308	0.305	0.214	0.218

Note: Bold values denote the best, underlined values the second best. GAN-TUPANN improves low-threshold performance in Rio de Janeiro but degrades or yields marginal improvements in other cities.

Table 3. *Cross-city CSI comparison between TUPANN trained on each city and TUPANN trained on Rio de Janeiro (TUPANN–Rio)*

	CSI-M↑		CSI_4↑		CSI_8↑		CSI_{16}↑		CSI_{32}↑		CSI_{64}↑	
Model	POOL1	POOL4	POOL1	POOL4	POOL1	POOL4	POOL1	POOL4	POOL1	POOL4	POOL1	POOL4
	Miami											
TUPANN	**0.169**	**0.187**	**0.267**	**0.312**	**0.189**	**0.211**	0.177	0.177	0.135	0.141	0.079	0.094
TUPANN–Rio	0.166	0.185	0.248	0.289	0.182	0.199	**0.178**	**0.182**	**0.138**	**0.152**	**0.085**	**0.103**
	Manaus											
TUPANN	**0.290**	**0.293**	**0.339**	**0.367**	**0.316**	**0.321**	**0.315**	**0.312**	**0.278**	**0.274**	**0.200**	**0.193**
TUPANN–Rio	0.236	0.249	0.283	0.319	0.257	0.276	0.254	0.264	0.214	0.221	0.172	0.167
	La Paz											
TUPANN	**0.314**	**0.317**	**0.336**	**0.363**	**0.327**	**0.323**	**0.350**	**0.340**	**0.327**	**0.322**	**0.232**	**0.239**
TUPANN–Rio	0.279	0.288	0.305	0.339	0.288	0.295	0.306	0.304	0.279	0.279	0.216	0.221

Note: Bold values denote the best, underlined values the second best. Cross-city performance declines in Manaus and La Paz but remains competitive; in Miami TUPANN–Rio improves high-threshold scores.

thresholds, possibly because heavy-rain events are rare in Miami (see Figure 1) and the Rio model may be biased towards such events. Overall, cross-city degradation is modest (at most 20%) and TUPANN–Rio still outperforms or matches baselines trained on the target city, highlighting the transferability of the architecture.

5.5.4. Multi-city training

Rather than training a model for each city, training on multiple cities can further improve skill. Table 4 compares TUPANN trained separately on each city with a multi-city model trained jointly on all regions (including Toronto, see Supplementary Material). The multi-city model (TUPANN–Multicity) yields higher CSI across most thresholds and regions. Access to diverse climates helps the network learn more generalizable features, especially for extreme rainfall events.

5.6. IMERG results

Table 5 compares TUPANN to baselines on the IMERG dataset for Rio de Janeiro. Without pooling (POOL1), TUPANN achieves the best CSI across all thresholds. With pooling (POOL4), generative models (NowcastNet, GAN-TUPANN) slightly outperform TUPANN at low thresholds, but TUPANN remains competitive and leads at higher thresholds. Figure 9 plots CSI–M versus lead time, showing TUPANN's superior performance at most lead times and small gaps only at 150 minutes. These results demonstrate that TUPANN generalizes to coarser spatial resolution and longer latency datasets, where its physics-aligned architecture can become slightly less beneficial.

5.7. Operational deployment

Downloading and processing GOES-16 data and making predictions can be done in less than 3 minutes. Due to its low latency, this enables an efficient near real-time 10 to 180 minutes nowcast with 10 minutes temporal resolution. The whole pipeline, including the predictions generated by TUPANN, can be run on a machine with relatively modest requirements: 16GB of memory, a GeForce RTX 3080 GPU. Predictions can also include the derived motion fields to help meteorologists interpret the results. This system is currently operational and has been deployed to aid Rio de Janeiro's Instituto Estadual do Ambiente (INEA) agency to prepare for floods and other extreme precipitation disasters. Due to TUPANN's reliance on widely available geostationary satellite data, we expect that this solution can be broadened to other regions as well.

6. Limitations and future work

Our study has important limitations. First, although TUPANN may use globally available satellite products, this work relies on GOES-16 coverage for real-time nowcasting. Still, our results show that the model is able to adapt to different locations. It is straightforward to retrain TUPANN to other regions such as Europe and Africa (through the EUMETSAT Meteosat series satellites products, such as P-IN-SEVIRI-PMW (H60B) (EUMETSAT H SAF, 2022)) and Eastern Asia and Western Pacific (through the Himawari satellites, from the Japanese Meteorological Agency, and GK-2A, from the Korea Meteorological Administration). Second, the optical-flow supervision requires additional computation and may not perfectly capture complex convection dynamics; using the proposed scheme with additional covariates could yield further improvements under a more detailed physical modeling. Third, CSI and HSS values decline at longer lead times (>2 h), reflecting inherent predictability limits; integrating uncertainty quantification and ensemble approaches may better characterize forecast skill. Finally, GAN-based enhancements improve visual realism but degrade or inconsistently affect skill metrics; stabilizing adversarial training and assessing perceptual quality remain open challenges.

Table 4. *Comparison between single-city TUPANN and a multi-city version (TUPANN–Multicity) trained on all regions simultaneously*

	CSI-M↑		CSI_4↑		CSI_8↑		CSI_{16}↑		CSI_{32}↑		CSI_{64}↑	
Model	POOL1	POOL4	POOL1	POOL4	POOL1	POOL4	POOL1	POOL4	POOL1	POOL4	POOL1	POOL4
	Rio de Janeiro											
TUPANN	0.259	0.277	0.330	0.384	0.289	0.289	0.330	0.336	0.274	0.287	**0.072**	**0.090**
TUPANN–Multicity	**0.271**	**0.286**	**0.337**	**0.390**	**0.303**	**0.297**	**0.353**	**0.353**	**0.293**	**0.304**	0.071	0.086
	Miami											
TUPANN	0.169	0.187	0.267	0.312	0.189	**0.211**	0.177	0.177	0.135	0.141	0.079	0.094
TUPANN–Multicity	**0.178**	**0.190**	**0.273**	**0.313**	**0.196**	**0.211**	**0.188**	**0.182**	**0.151**	**0.151**	**0.082**	**0.095**
	Manaus											
TUPANN	0.290	0.293	0.339	0.367	**0.316**	0.321	**0.315**	0.312	**0.278**	0.274	0.200	0.193
TUPANN–Multicity	**0.291**	**0.296**	**0.340**	**0.369**	**0.316**	**0.324**	0.314	**0.314**	**0.278**	**0.275**	**0.208**	**0.198**
	La Paz											
TUPANN	0.314	0.317	0.336	**0.363**	0.327	0.323	0.350	0.340	0.327	0.322	0.232	0.239
TUPANN–Multicity	**0.324**	**0.325**	**0.339**	**0.363**	**0.334**	**0.327**	**0.359**	**0.348**	**0.335**	**0.331**	**0.251**	**0.254**

Note: Bold denotes the best, underlined the second best. The multi-city model improves performance in most cases, demonstrating the benefit of pooled training.

Table 5. *Aggregated CSI metrics for Rio de Janeiro using IMERG data*

	CSI-M↑		CSI_4↑		CSI_8↑		CSI_{16}↑		CSI_{32}↑		CSI_{64}↑	
Model	POOL1	POOL4	POOL1	POOL4	POOL1	POOL4	POOL1	POOL4	POOL1	POOL4	POOL1	POOL4
Earthformer	0.153	0.141	0.361	0.343	0.270	0.243	0.130	0.114	0.006	0.006	0.000	0.000
NowcastNet	**0.209**	0.271	0.387	0.453	0.322	0.384	0.229	0.304	**0.103**	**0.201**	0.001	**0.013**
PySTEPS (LK)	0.114	0.099	0.300	0.269	0.191	0.164	0.075	0.061	0.002	0.003	0.000	0.000
PySTEPS (DARTS)	0.107	0.096	0.287	0.262	0.179	0.157	0.069	0.057	0.002	0.003	0.000	0.000
CasCast	0.180	0.229	0.362	0.420	0.288	0.339	0.188	0.248	0.061	0.139	0.000	0.000
TUPANN (ours)	**0.218**	0.248	**0.414**	0.454	**0.344**	0.379	**0.241**	0.280	0.087	0.123	**0.005**	0.006
GAN-TUPANN	0.210	**0.274**	0.391	**0.461**	0.327	**0.393**	0.234	**0.313**	0.095	0.199	0.002	0.004

Note: Bold denotes the best, underlined the second best. Without pooling TUPANN is clearly superior; with pooling, generative baselines perform slightly better at low thresholds, but TUPANN remains competitive overall.

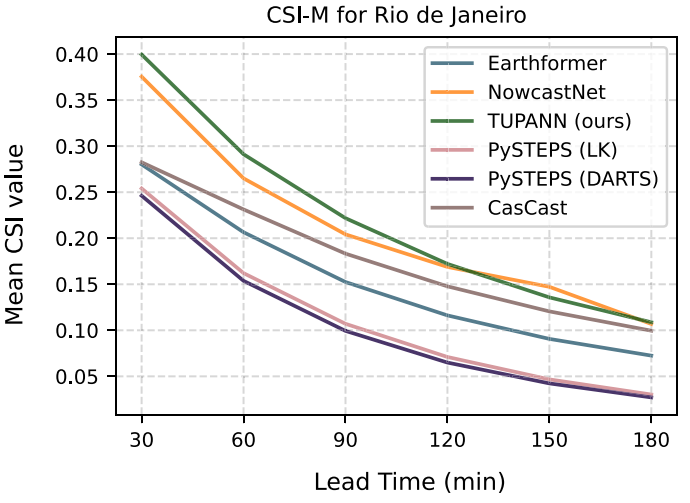


Figure 9. *Mean CSI versus lead time for IMERG data in Rio de Janeiro. TUPANN outperforms baselines at most lead times; NowcastNet overtakes slightly at 150 min but lags at shorter lead times.*

Future work includes coupling TUPANN with probabilistic post-processing to provide calibrated uncertainty estimates, extending the model to include additional inputs such as cloud imagery or microwave precipitation retrievals, and investigating transfer learning across continents.

7. Conclusions

We have presented TUPANN, a physically aligned neural network for precipitation nowcasting using satellite imagery. TUPANN's modular design—combining a variational encoder–decoder supervised by optical flow, and a transformer capable of evolving the latent representation according to physical constraints—yields interpretable motion fields and competitive forecast skill. Extensive experiments on GOES-16 and IMERG data across four climates show that TUPANN matches or surpasses state-of-the-art baselines, particularly at high precipitation thresholds. Training in multiple cities improves performance, and cross-city evaluations reveal modest degradation, highlighting the model's transferability. With its low latency and reliance on globally available satellites, TUPANN supports equitable access to short-term rainfall forecasts and provides a foundation for operational applications in radar-sparse regions.

Open peer review. To view the open peer review materials for this article, please visit http://doi.org/10.1017/eds.2026.10039.

Supplementary material. The supplementary material for this article can be found at http://doi.org/10.1017/eds.2026.10041.

Acknowledgements. The authors thank Adriana Monteiro, João Vitor Romano, Lucas Nissenbaum, and Thiago Ramos, as well as the help and support from Google, in particular, Shreya Agrawal, Boris Babenko, and Samier Merchant.

Author contribution. Conceptualization: A.C., M.P., L.V., P.O.; Data curation and visualization: A.C.; Methodology: A.C., M.P., L.V.; Writing—original draft: A.C., M.P., L.V.; Writing—review and editing: A.C., M.P., L.V., P.O. All authors approved the final manuscript.

Competing interests. The authors declare no competing interests.

Data availability statement. GOES-16 RRQPE data were obtained from the NOAA GOES-R Product Distribution and Access (PDA) system. IMERG Final Run data were obtained from the NASA GPM data portal. Processed datasets, event splits, and TUPANN code are available at Zenodo under the DOI 10.5281/zenodo.20349104 (Catão et al., 2026).

Ethics statement. The research meets all ethical guidelines, including adherence to the legal requirements of Brazil and the United States.

Funding statement. A.C. was supported by a "FAPERJ Nota 10" grant (SEI-260003/004731/2025) from Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ). M.P. and L.V. were supported by scholarships from Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES). P.O. was supported by grant SEI-260003/001545/2022 from FAPERJ.

References

Agrawal S, **Hassen MA**, **Brempong EA**, **Babenko B**, **Zyda F**, **Graham O**, **Li D**, **Merchant S**, **Potes SH**, **Russell T**, **Cheresnick D**, **Kakkirala AP**, **Rasp S**, **Hassidim A**, **Matias Y**, **Kalchbrenner N**, **Gupta P**, **Hickey J and Bell A** (2025) An operational deep learning system for satellite-based high-resolution global nowcasting. arXiv:2510.13050.

An S, **Oh T-J**, **Sohn E and Kim D** (2025) Deep learning for precipitation nowcasting: A survey from the perspective of time series forecasting. *Expert Systems with Applications 268*, 126301.

Andrychowicz M, **Espeholt L**, **Li D**, **Merchant S**, **Merose A**, **Zyda F**, **Agrawal S and Kalchbrenner N** (2023) Deep learning for day forecasts from sparse observations. arXiv:2306.06079.

Bai C, **Sun F**, **Zhang J**, **Song Y and Chen S** (2022) Rainformer: Features extraction balanced network for radar-based precipitation nowcasting. *IEEE Geoscience and Remote Sensing Letters 19*, 1–5.

Bowler NE, **Pierce CE and Seed AW** (2006) STEPS: A probabilistic precipitation forecasting scheme which merges an extrapolation nowcast with downscaled nwp. *Quarterly Journal of the Royal Meteorological Society 132*(620), 2127–2155.

Browning KA and Collier CG (1989) Nowcasting of precipitation systems. *Reviews of Geophysics 27*(3), 345–370.

Burpee RW (1979) Peninsula-scale convergence in the South Florida Sea breeze. *Monthly Weather Review 107*(7), 852–865.

Catão A, **Poveda Quimiz MH**, **Voltarelli Regini de Andrade L and Orenstein P** (2026) acataos/tupann: TUPANN v1.0.0. Zenodo. Retrieved May 22, 2026, from https://zenodo.org/records/20349104.

EUMETSAT H SAF (2022) Product validation report (PVR) for products P-IN-FCI (H40B). Retrieved 22 October, 2025, from https://hsaf.meteoam.it/Products/Detail?prod=H60B.

Gao Z, **Shi X**, **Wang H**, **Zhu Y**, **Wang Y**, **Li M and Yeung D-Y** (2022) Earthformer: Exploring space-time transformers for earth system forecasting. In *Proceedings of the 36th International Conference on Neural Information Processing Systems, NIPS '22*. Red Hook, NY: Curran Associates Inc.

Gao Z, **Shi X**, **Han B**, **Wang H**, **Jin X**, **Maddix D**, **Zhu Y**, **Li M and Wang Y** (2023) PreDiff: Precipitation nowcasting with latent diffusion models. In *Proceedings of the 37th International Conference on Neural Information Processing Systems, NIPS '23*. Red Hook, NY: Curran Associates Inc.

Garreaud RD (2001) Interannual rainfall variability over the south american altiplano and related circulation anomalies. *Journal of Climate 14*(12), 2779–2789.

Germann U and Zawadzki I (2002) Scale-dependence of the predictability of precipitation from continental radar images. Part i: Description of the methodology. *Monthly Weather Review 130*(12), 2859–2873.

GOES-R Algorithm Working Group and **GOES-R Program Office** (2018) NOAA GOES-R series advanced baseline imager (ABI) level 2 rainfall rate quantitative precipitation estimation (RRQPE). Retrieved March 2, 2025, from https://www.ncei.noaa.gov/access/metadata/landing-page/bin/iso?id=gov.noaa.ncdc:C01517.

Gong J, **Bai L**, **Ye P**, **Xu W**, **Liu N**, **Dai J**, **Yang X and Ouyang W** (2024) Cascast: Skillful high-resolution precipitation nowcasting via cascaded modelling. In *Proceedings of the 41st International Conference on Machine Learning*, ICML'24. JMLR.org.

Guibas J, **Mardani M**, **Li Z-Y**, **Tao A**, **Anandkumar A and Catanzaro B** (2021) Adaptive Fourier neural operators: Efficient token mixers for transformers. *ArXiv*, abs/2111.13587.

Huffman G, **Bolvin D**, **Braithwaite D**, **Hsu K**, **Joyce R and Xie P** (2014) Integrated multi-satellite retrievals for GPM (IMERG), version 4.4. Retrieved June 15, 2026, from http://arthurhou.pps.eosdis.nasa.gov/gpmdata/.

Karniadakis GE, **Kevrekidis IG**, **Lu L**, **Perdikaris P**, **Wang S and Yang L** (2021) Physics-informed machine learning. *Nature Reviews Physics 3*(6), 422–440.

Marvel K, **Su W**, **Delgado R**, **Aarons S**, **Chatterjee A**, **Garcia M**, **Hausfather Z**, **Hayhoe K**, **Hence D**, **Jewett E**, **Robel A**, **Singh D**, **Tripati A and Vose R** (2023). *Chapter 2 : Climate Trends. Fifth National Climate Assessment.*

Kovacs A, **Exl L**, **Kornell A**, **Fischbacher J**, **Hovorka M**, **Gusenbauer M**, **Breth L**, **Oezelt H**, **Yano M**, **Sakuma N**, **Kinoshita A**, **Shoji T**, **Kato A and Schrefl T** (2022) Conditional physics informed neural networks. *Communications in Nonlinear Science and Numerical Simulation 104*(106041), 106041.

Kurth T, **Subramanian S**, **Harrington P**, **Pathak J**, **Mardani M**, **Hall D**, **Miele A**, **Kashinath K and Anandkumar A** (2023) Fourcastnet: Accelerating global high-resolution weather forecasting using adaptive Fourier neural operators. In *Proceedings of the Platform for Advanced Scientific Computing Conference, PASC '23*. New York, NY: Association for Computing Machinery.

Lebedev V, **Ivashkin V**, **Rudenko I**, **Ganshin A**, **Molchanov A**, **Ovcharenko S**, **Grokhovetskiy R**, **Bushmarinov I and Solomentsev D** (2019) Precipitation nowcasting with satellite imagery. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, KDD '19*. New York, NY: Association for Computing Machinery, pp. 2680–2688.

Lucas BD and Kanade T (1981) An iterative image registration technique with an application to stereo vision. In *Proceedings of the 7th International Joint Conference on Artificial Intelligence - Volume 2, IJCAI'81*. San Francisco, CA: Morgan Kaufmann Publishers Inc, pp. 674–679.

Luiz-Silva W and Oscar-Júnior AC (2022) Climate extremes related with rainfall in the state of Rio de Janeiro, Brazil: A review of climatological characteristics and recorded trends. *Natural Hazards 114*(1), 713–732. https://doi.org/10.1007/s11069-022-05409-5.

McClenny LD and Braga-Neto UM (2023) Self-adaptive physics-informed neural networks. *Journal of Computational Physics 474*(111722), 111722.

Niu D, **Li Y**, **Wang H**, **Zang Z**, **Jiang M**, **Chen X and Huang Q** (2024) FsrGAN: A satellite and radar-based fusion prediction network for precipitation nowcasting. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 17*, 7002–7013.

Oliveira R, **Maggioni V**, **Vila D**, **Morales C** (2016) Characteristics and diurnal cycle of GPM rainfall estimates over the central amazon region (around Manaus). *Remote Sensing 8*(7), 544.

Park Y-J, **Kim D**, **Seo M**, **Jeon H-G and Choi Y** (2025). Data-driven precipitation nowcasting using satellite imagery. *Proceedings of the AAAI Conference on Artificial Intelligence*, *39*(27), 28284–28292.

Pulkkinen S, **Nerini D**, **Pérez Hortal AA**, **Velasco-Forero C**, **Seed A**, **Germann U and Foresti L** (2019) PySTEPS: An open-source Python library for probabilistic precipitation nowcasting (v1.0). *Geoscientific Model Development 12*(10), 4185–4219.

Rahimi R, **Ravirathinam P**, **Ebtehaj A**, **Behrangi A**, **Tan J and Kumar V** (2024) Global precipitation nowcasting of integrated multi-satellite retrievals for GPM: A U-Net convolutional LSTM architecture. *Journal of Hydrometeorology 25*(6), 947–963.

Raissi M, **Perdikaris P and Karniadakis GE** (2019) Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics 378*, 686–707.

Ravuri S, **Lenc K**, **Willson M**, **Kangin D**, **Lam R**, **Mirowski P**, **Fitzsimons M**, **Athanassiadou M**, **Kashem S**, **Madge S**, **Prudden R**, **Mandhane A**, **Clark A**, **Brock A**, **Simonyan K**, **Hadsell R**, **Robinson N**, **Clancy E**, **Arribas A and Mohamed S** (2021) Skilful precipitation nowcasting using deep generative models of radar. *Nature 597*(7878), 672–677. https://doi.org/10.1038/s41586-021-03854-z.

Ronneberger O, **Fischer P and Brox T** (2015) U-Net: Convolutional networks for biomedical image segmentation. In Navab N, Hornegger J, Wells WM and Frangi AF (eds), *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. Cham: Springer International Publishing, pp. 234–241.

Ruzanski E, **Wang Y and Chandrasekar V** (2009) Development of a real-time dynamic and adaptive nowcasting system. *In 25th Conference on International Interactive Information and Processing Systems (IIPS) for Meteorology, Oceanography,and Hydrology*, volume *2*. 89th American Meteorological Society Annual Meeting, American Meteorological Society.

Salimans T and Ho J (2022) Progressive distillation for fast sampling of diffusion models. In *International Conference on Learning Representations*.

Saxena D and Cao J (2021) Generative adversarial networks (GANs): Challenges, solutions, and future directions. *ACM Computing Surveys 54*(3), 1–42.

Seed AW, **Pierce CE and Norman K** (2013) Formulation and evaluation of a scale decomposition-based stochastic precipitation nowcast scheme. *Water Resources Research 49*(10), 6624–6641.

Shi X, **Chen Z**, **Wang H**, **Yeung D-Y**, **Wong W-k and Woo W-c** (2015) Convolutional LSTM network: A machine learning approach for precipitation nowcasting. In *Proceedings of the 29th International Conference on Neural Information Processing Systems - Volume 1, NIPS'15*. Cambridge, MA: MIT Press, pp. 802–810.

Shukla BP, **Vyas J**, **Chhari A**, **Shah S**, **Panda SK and Varma AK** (2025) Satellite based rainfall nowcasting using geospatial techniques. *Meteorology and Atmospheric Physics 137*(1), 3.

Tu Z, **Talebi H**, **Zhang H**, **Yang F**, **Milanfar P**, **Bovik A and Li Y** (2022) MaxViT: Multi-axis vision transformer. In Avidan S, Brostow G, Cissé M, Farinella G, Hassner T (Eds.), *Computer Vision – ECCV 2022*. Cham, Switzerland: Springer, pp. 459–479.

Veillette MS, **Samsi S and Mattioli CJ** (2020) SEVIR: A storm event imagery dataset for deep learning applications in radar and satellite meteorology. In *Proceedings of the 34th International Conference on Neural Information Processing Systems, NIPS '20*. Red Hook, NY: Curran Associates Inc.

Wang S, **Yu X and Perdikaris P** (2022) When and why PINNs fail to train: A neural tangent kernel perspective. *Journal of Computational Physics 449*(110768), 110768.

Wang Y, **Wu H**, **Zhang J**, **Gao Z**, **Wang J**, **Yu PS and Long M** (2023) PredRNN: A recurrent neural network for spatiotemporal predictive learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence 45*(2), 2208–2225. https://doi.org/10.1109/TPAMI.2022.3165153.

Yin J, **Meo C**, **Roy A**, **Cher Z**, **Lică M**, **Wang Y**, **Imhoff R**, **Uijlenhoet R and Dauwels J** (2024). Precipitation nowcasting using physics informed discriminator generative models. *In 2024 32nd European Signal Processing Conference(EUSIPCO)*, 967–971. IEEE.

Zhang Y, **Long M**, **Chen K**, **Xing L**, **Jin R**, **Jordan MI and Wang J** (2023) Skilful nowcasting of extreme precipitation with Nowcastnet. *Nature 619*(7970), 526–532. https://doi.org/10.1038/s41586-023-06184-4.

Zheng K, **He L**, **Ruan H**, **Yang S**, **Zhang J**, **Luo C**, **Tang S**, **Zhang J**, **Tian Y and Cheng J** (2024) A cross-modal spatiotemporal joint predictive network for rainfall nowcasting. *IEEE Transactions on Geoscience and Remote Sensing 62*, 1–23.

Cite this article: Catão A, Voltarelli L, Poveda M and Orenstein P (2026). Precipitation nowcasting of satellite data using physically aligned neural networks. *Environmental Data Science*, 5: e13. doi:10.1017/eds.2026.10041