iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 Issue: I Month of publication: January 2026

DOI: https://doi.org/10.22214/ijraset.2026.76920

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue I Jan 2026- Available at www.ijraset.com

AI for Weather Nowcasting: Opportunities and Challenges

Zeyad Faieq Suleiman Abuhuweidi[1], Ibrahim Khaled Qasem Farea[2]
[1]School of Computer Science, Nanjing University of Information Science and Technology
[2]School of AI, Nanjing University of Information Science and Technology

Abstract: Weather nowcasting, defined as forecasting weather phenomena on time scales from minutes to several hours, is critical for mitigating the impacts of high-impact events such as flash floods, severe convective storms, and extreme precipitation. Traditional nowcasting approaches based on radar extrapolation and convection-permitting numerical weather prediction (NWP) exhibit fundamental limitations in representing rapid storm evolution, convective initiation, and localized extremes at short lead times. Recent advances in artificial intelligence (AI) and deep learning have enabled a new generation of nowcasting systems that learn complex spatiotemporal relationships directly from high-resolution radar, satellite, lightning, and NWP data. This paper provides a comprehensive review of AI-based precipitation nowcasting, covering data sources, model architectures, and evaluation methodologies. We discuss deterministic and probabilistic approaches, including convolutional recurrent networks, encoder–decoder convolutional neural networks, transformers, generative adversarial networks, diffusion models, and emerging physics-informed and hybrid AI–NWP systems. Opportunities such as improved short-lead forecast skill, multi-sensor fusion, probabilistic decision support, and enhanced forecast equity are examined alongside key challenges related to data quality, class imbalance, generalization, interpretability, and operational deployment. Finally, we highlight current research frontiers and methodological trends, outlining open challenges and promising directions for future AI-driven nowcasting systems at the PhD level and beyond.
Keywords: Weather nowcasting; precipitation forecasting; deep learning; radar and satellite data; probabilistic forecasting; physics-informed machine learning; diffusion models; numerical weather prediction.

I. INTRODUCTION

Weather nowcasting —forecasting on time -scales from a few minutes up to roughly six hours —is critical for protecting life, infrastructure, and economic activity. Many of the most dangerous hazards, such as flash floods, severe convective storms, and microbursts, evolve on these short time -scales. Over the last half-century, a large fraction of weather-related disasters and fatalities has been linked to extreme precipitation events [1][2]. Improving short-lead-time forecasts therefore has high societal value for emergency management, aviation, road safety, agriculture, renewable energy, and urban drainage systems. However, nowcasting is particularly challenging. Hazardous convective storms can initiate, intensify, split and decay on time -scales of tens of minutes. Even modern convection-permitting numerical weather prediction (NWP) models struggle to represent the exact timing, location and intensity of localized heavy rainfall [3]. Data assimilation cycles are typically 1 h or longer, which means that valuable high-frequency observations (e.g. radar volumes every 5 min) are not fully exploited in real time [13]. For these reasons, nowcasting has been described as one of the most difficult problems in hydrometeorology [4]. Recent advances in artificial intelligence (AI) and deep learning offer a promising new path. Deep neural networks can learn complex, non-linear relationships directly from large archives of radar, satellite and NWP data, potentially overcoming some limitations of traditional extrapolation and NWP approaches [5][6]. Pioneering systems such as DeepMind's deep generative model of radar rainfall (DGMR) have already demonstrated that AI-based nowcasts can be judged more accurate and useful than existing operational products in blind evaluations by expert meteorologists [3][7]. At the same time, these models are new, sometimes opaque, and not yet universally trusted in operations. This report provides a scholarly review of AI-based nowcasting, focusing on opportunities, challenges, and directions for research at the PhD level and beyond.

II. BACKGROUND AND DEFINITIONS

The World Meteorological Organization (WMO) typically defines nowcasting as providing detailed description of current weather, plus forecasts up to about 0 –6 h, emphasizing local detail and rapid update cycles [8]. Short-range forecasts (roughly 6 –24 h) and medium-range forecasts (several days to two weeks) focus increasingly on larger-scale atmospheric dynamics and are more naturally produced by NWP on coarser grids.

In practice, operational nowcasts place particular emphasis on phenomena that evolve quickly and have large local impacts: convective cells, squall lines, mesoscale convective systems, urban thunderstorms, hail storms, and localized heavy rainfall. Short-range NWP remains indispensable beyond ~3 –6 h, but in the "nowcast window" AI models can exploit the latest high-resolution observations to add value where NWP is weakest.

A. Traditional Nowcasting Methods

Table 1 Comparison of traditional nowcasting vs AI-based nowcasting (typical characteristics and performance)

Aspect	Traditional nowcasting (radar extrapolation / NWP)	AI-based nowcasting (deep learning models)
Mechanism	Extrapolates recent radar echoes using an estimated motion field (optical flow / cross-correlation) or integrates NWP equations for 1–6 h on fixed grids.[10][21][22]	Learns a spatiotemporal mapping from recent radar, satellite and/or NWP fields to future precipitation using neural networks (RNNs, CNNs, transformers, diffusion models).[3][4][6][12][13][14][18]
Strengths	Very fast and cheap; radar extrapolation can update every 5–10 min; NWP provides physically consistent multi-variable fields and synoptic-scale guidance.[10][21][22]	Captures non-linear storm growth/decay and environmental context; can blend multiple sensors; once trained, inference is very fast and scales well over large domains.[1][2][3][6][13][16][18]
Limitations	Assumes storms persist and simply move; cannot represent new convective initiation or rapid dissipation; skill decays sharply beyond ~1–2 h, especially for convection.[10][21][22]	Requires large high-quality datasets and heavy training compute; can blur fine-scale extremes; may generate unphysical fields if not constrained; generalization and interpretability are active research challenges.[4][6][7][11][27]
Typical skillful lead time	Good skill for ~0–30 (sometimes up to ~60) min at convective scales; NWP becomes more useful beyond ~2–3 h for larger-scale systems.[10][21][22]	State-of-the-art radar nowcasting shows clear added value out to ~2–3 h for convective rainfall; probabilistic large-context models extend useful precipitation skill to ~8–12 h for moderate events.[1][2][3][16][18]
Representative systems	Radar extrapolation systems such as STEPS, MAPLE and open-source PySTEPS; convection-permitting rapid-update NWP systems (e.g. HRRR, UKV).[10][21][22]	DGMR GAN nowcasting system [3]; MetNet and MetNet-2 large-context CNNs [13][18]; PredRNN family of spatiotemporal RNNs [14]; RainNet CNN baseline [25]; Rainformer hybrid transformer–CNN [28]; diffusion-based PreDiff models [27].

Before the AI era, operational nowcasting relied primarily on extrapolation of recent observations. The simplest baseline is persistence, assuming conditions remain unchanged over the forecast period. More sophisticated systems perform Lagrangian advection of radar reflectivity, estimating a motion field (often via optical flow or cross-correlation between successive radar images) and then translating echoes forward in time [9][10]. Frameworks
such as STEPS and the open-source PySTEPS library add stochastic perturbations and simple growth/decay models, which partially respect conservation and yield probabilistic nowcasts [10][11].

These advection methods are surprisingly skillful at very short leads (~0 –1 h), because they preserve observed storm structures and simply advect them. However, they cannot predict initiation of new convection, rapid upscale growth, splitting or sudden dissipation beyond what can be inferred from recent motion [12]. Skill typically drops sharply after ~1 –2 h, especially for convective precipitation.

Another strand of "traditional" nowcasting is to use high-resolution NWP in rapid-update mode. For example, the U.S. HRRR model runs hourly at ~3 km grid spacing out to ~18 h.

Such models provide physically consistent multi-variable fields and can, in principle, simulate storm growth and environmental interactions. Yet they remain computationally expensive, have spin-up issues, and often misrepresent small-scale convective structures at very short lead times [3][13]. In practice, many centers blend radar extrapolation for the first 0–2 h with NWP beyond that.

B. Data sources for nowcasting

Effective nowcasting requires dense, frequent observations:

1) Weather radar is the workhorse for precipitation nowcasting, providing volumetric reflectivity scans every 2 –10 min at horizontal resolutions of order 1 km [14]. Radar captures the evolving structure of rain and snow fields and is well suited as both input and verification for AI systems. However, radar coverage is uneven globally, and radars see precipitation but not pre-convective cloud development.
2) Geostationary satellites provide near-global coverage with scan intervals of 5 –15 min. Infrared (IR) and visible channels give information on cloud-top temperature, texture and growth, which are valuable precursors of convective initiation [15]. AI models can learn relationships between satellite features and surface rainfall, particularly when trained in regions with both radar and satellite.
3) Surface rain gauges, automatic weather stations, lightning detection networks and crowdsourced observations provide additional constraints on precipitation intensity and severe weather. Lightning in particular is a strong indicator of deep convection and has been used in multi-task deep learning models for concurrent rainfall and lightning nowcasting [17].

Environmental fields from global reanalyses or NWP (e.g. instability indices, vertical wind shear, moisture flux convergence) provide large-scale context that helps AI models distinguish environments favorable for sustained convection from those where showers quickly decay [16][18].

Table 1 Example datasets used for AI-based precipitation nowcasting

Dataset (region)	Data source(s)	Resolution (space / time)	Time span (approx.)	Notes and typical use
HKO-7 (Hong Kong)	C-band weather radar reflectivity	≈1 km grid, 5-min frames	2009–2015 (~7 years)	Classic benchmark for deep learning nowcasting; used in Shi et al.'s benchmark and TrajGRU work.[23]
Shanghai radar (China)	S-band radar reflectivity (single site)	≈0.5 km, ≈6-min frames (subset)	2014–2015	Used in the original ConvLSTM paper for 10-frame radar prediction experiments.[12][23]
Iowa Rain / CONUS subset (USA)	NEXRAD national radar mosaic	≈1 km, 5-min frames	2016–2019 (multi-year)	Used by Lebedev et al. for satellite-aided nowcasting and as a benchmark for radar extrapolation vs AI.[24]

MRMS (USA, national)	Multi-radar / multi-sensor (MRMS) mosaic	1 km grid, 2–5-min products	2017–present	High-resolution national composite widely used in stochastic nowcasting frameworks such as PySTEPS.[10]
SEVIR (USA)	Co-located radar, satellite and lightning	1 km, 5-min (radar/satellite)	2017–2019	Public benchmark for multi-modal severe-weather prediction and nowcasting; curated events for deep learning.[26]
OPERA (Europe)	Pan-European radar composite	≈2 km grid, 15-min	≈2011–2019 (operational)	Used in European nowcasting research and EUMETNET/OPERA products; often a source for regional composites.
RainNet (Germany)	National radar mosaic (DWD network)	1 km grid, 5-min frames	2015–2018	Basis of the RainNet CNN baseline for radar-based nowcasting; open dataset for method comparison.[25]
MeteoNet (France)	Radar + NWP-derived fields	1 km, 5-min radar	2016–2018 (challenge period)	Dataset released for the MeteoNet challenge by Météo-France; used to test radar+NWP fusion.

Multi-sensor fusion is non-trivial because different datasets have distinct spatial/temporal resolutions, coverage, and noise characteristics. Nevertheless, combining radar's high-resolution rainfall view with satellites' broad cloud view and NWP environment has emerged as a key strategy for robust AI nowcasting [15][16][18].

III. AI AND DEEP LEARNING METHODS FOR NOWCASTING

A. Deterministic vs. probabilistic models

Early deep learning nowcasting systems typically framed the problem as supervised video prediction: given a sequence of past radar images, predict future images by minimizing a pixel-wise loss such as mean squared error (MSE). This yields a single "best-guess" deterministic forecast. A well-known drawback is that MSE encourages "averaging" over many plausible futures, leading to overly smooth, blurry forecasts, especially at longer lead times [5].

To address this, the community has increasingly shifted to probabilistic models that represent uncertainty. Instead of one deterministic forecast, a generative model produces an ensemble of possible future radar sequences consistent with the recent past. DeepMind's DGMR is a prominent example: a conditional generative adversarial network (GAN) that outputs an ensemble of high-resolution radar scenarios up to 90 min ahead [3]. The adversarial loss encourages realistic, sharp structures that better match observed convective cells [3][7]. More recently, diffusion models have emerged as a competitive alternative. Leinonen et al. (2023) applied a latent diffusion model (LDM) to precipitation nowcasting, showing improved sharpness and more reliable uncertainty compared to both GAN-based DGMR and traditional extrapolation [6][19]. Diffusion models iteratively "denoise" random noise into a forecast conditioned on past observations, naturally producing ensembles whose spread can be calibrated to match forecast uncertainty [20].

Probabilistic nowcasts can be evaluated using metrics such as the continuous ranked probability score (CRPS) [21] and reliability diagrams, and they have clear advantages for decision-making because users can base actions on forecast confidence rather than a single deterministic scenario [22].

B. Convolutional Recurrent Networks (ConvLSTMs, PredRNN)

One of the earliest deep learning approaches to nowcasting was the Convolutional LSTM (ConvLSTM) of Shi et al. (2015) [12]. ConvLSTM extends standard LSTMs by replacing fully connected operations with convolutions, allowing the network to model spatio-temporal evolution of radar images. In an encoder –forecaster configuration, ConvLSTM learns both motion and intensity changes and was shown to outperform optical-flow-based extrapolation for short-term precipitation prediction in Hong Kong [12].

Subsequent work introduced variants such as Trajectory GRU (TrajGRU), which learns dynamic connection structures that better follow advecting rain cells, and PredRNN, which adds additional memory cells to better capture long-term dependencies [14]. PredRNN and its successors (e.g. PredRNN-v2, MIM) have achieved strong performance on benchmark radar datasets and, importantly, have seen early operational adoption. For example, the China Meteorological Administration has deployed a PredRNN-type system for short-term precipitation guidance [14].

ConvLSTM-family models are relatively lightweight and well suited to GPU inference in real time. However, when trained with simple pixel-wise losses they still tend to blur small-scale, intense features at longer lead times. Hybrid training strategies (e.g. combining MSE with structural or threshold-based losses) and integration with generative objectives partially mitigate this, but many state-of-the-art systems now combine recurrent backbones with GAN or diffusion modules [1][3][6].

C. Encoder –decoder CNNs and Vision Transformers

Fully convolutional encoder –decoder networks, often in U-Net form, provide another widely used architecture for nowcasting. These models treat a stack of recent radar (and sometimes satellite) images as input channels and predict multiple future frames in a single feed-forward pass. Skip connections help preserve fine spatial detail, and 3-D convolutions over space – time can capture motion implicitly. With sufficient data, U-Net-type models have been shown to rival or surpass ConvLSTM in some settings [4].

Attention-based architectures, such as Vision Transformers (ViTs) and spatio-temporal transformers, have recently entered the field. Google's MetNet[13] and MetNet-2[18] are examples that combine convolutional backbones with attention mechanisms over large spatial contexts to produce probabilistic precipitation forecasts up to 8 –12 h ahead.

Transformer-style self-attention allows the model to link distant upstream features with local rainfall, a key advantage for longer-lead nowcasts where remote precursors become relevant.

The main trade-off is computational cost: pure transformers scale poorly with resolution and domain size. Modern systems therefore use hybrids (e.g. CNN encoders plus local –global attention blocks) and hierarchical, multi-scale tokenization to keep inference fast enough for operational use [4][18].

D. GANs and diffusion models

Generative adversarial networks (GANs) and diffusion models explicitly target the realism and uncertainty representation of nowcasts. DGMR uses a U-Net generator and a discriminator that evaluates entire radar sequences, producing sharp, high-resolution ensembles that human forecasters often prefer over extrapolation and NWP in the 0 –2 h range [3][7]. However, first-generation GAN systems tended to under-predict very rare extremes and did not guarantee calibrated probabilities [3][18].

Diffusion models, including LDM-type architectures, address some of these issues. Their iterative denoising structure produces ensembles that can be tuned for reliability and supports conditioning on additional inputs (e.g. NWP fields, satellite features) via "guided diffusion" [6][19][20]. For instance, Physical-Driven Diffusion Networks (PDDN) condition the diffusion process on fields from a limited-area NWP model (WRF), leading to improved 6-h precipitation nowcasts that outperform both pure ML and pure NWP baselines in several case studies [8].

E. Graph neural networks and geometric deep learning

Most nowcasting architectures operate on regular latitude –longitude or Cartesian grids. Graph neural networks (GNNs) provide a way to represent precipitation fields and observation networks on irregular meshes or learned adjacency graphs. Zhao et al. (2023) proposed a geometric deep learning framework in which each grid cell is treated as a node, and the model learns an adjacency matrix that captures dynamic, flow-dependent relationships rather than fixed local neighborhoods [7]. Temporal graph convolutions then propagate information along these learned edges, improving representation of complex advection and deformation patterns.

GNNs also offer a natural way to fuse heterogeneous data sources such as radar pixels, rain gauges, and lightning sensors into a single graph. While still an emerging area, geometric approaches align well with recent global forecasting models like GraphCast [19], and similar ideas are likely to be applied at nowcasting scales.

F. Physics-informed and Hybrid Models

A major research trend is to embed physical knowledge into AI architectures, creating hybrid physics-AI systems. Pure data-driven models can produce unphysical outputs (e.g. non-conservative rain fields, unrealistic storm growth), which raises concerns for high-impact applications. Physics-informed designs aim to incorporate constraints such as mass conservation, approximate continuity equations, or known advection operators directly into the network or loss function.

NowcastNet, a physics-conditioned generative model, is a leading example [1]. It embeds differentiable operators inspired by the precipitation continuity equation into a GAN-style framework, encouraging realistic advection and growth of storms while still allowing flexible learning of non-linear processes. Evaluations over the U.S. and China show that NowcastNet outperforms both DGMR and high-resolution NWP for extreme precipitation thresholds, with much higher critical success index (CSI) for very heavy rain [1][2].

Hybrid diffusion models such as PDDN go a step further by directly conditioning on NWP fields [8]. In these systems, AI and NWP complement each other: NWP provides dynamically consistent large-scale context, while AI learns fine-scale structures and corrects systematic NWP errors. This hybridization improves robustness, extends useful lead times to 3 –6 h, and can enhance interpretability because some components have explicit physical meaning.

Table 2 Representative deep learning models for radar nowcasting (selected examples, approaches, and reported highlights)

Model (year)	Type / approach	Key idea(s)	Example dataset & performance (approximate)	References
ConvLSTM (2015)	ConvLSTM encoder–decoder RNN	Adds convolutions inside LSTM gates so the hidden state is a feature map, well-suited to radar image sequences.	Demonstrated improved MSE/CSI vs optical-flow extrapolation on Shanghai radar 10-frame nowcasting tasks.	[12]
TrajGRU (2017)	Trajectory-aware GRU (recurrent)	Learns location-variant recurrent connections that move with flow, enabling better representation of advection.	On HKO-7 radar, reduces MSE and improves CSI relative to vanilla ConvLSTM at 1-h lead times.[23]	[23]
PredRNN (2017/2022)	Spatiotemporal LSTM (RNN)	Uses dual memories and a “gradient highway” to mitigate blurring over long sequences in video / radar prediction.	Strong results on generic video (e.g. Moving-MNIST) and improved structural similarity for radar nowcasts.	[14]
RainNet (2020)	U-Net-style CNN	Fully convolutional encoder–decoder taking a short radar history and predicting future frames in one pass.	On German 5-min radar, outperforms persistence at 30–90 min; CSI $\gtrsim 0.5$ for light rain at 1-h lead.[25]	[25]
MetNet (2020)	Large-context CNN with attention (probabilistic)	Aggregates a very large spatial context and produces	Over CONUS, provides skillful 8-h precipitation probabilities vs	[13]

		calibrated probability distributions for rain rates.	operational baselines.	
DGMR (2021)	GAN (deep generative radar model, DeepMind)	Uses a generative adversarial network to produce ensembles of realistic radar futures, tuned for extremes.	For 90-min UK radar, ~89% of 50 Met Office forecasters preferred DGMR over the operational system.	[3]
Rainformer (2022)	Hybrid transformer + CNN	Swin-style transformer blocks capture long-range dependencies; CNN layers refine local convective structure.	On Chinese radar, improves CSI for moderate rain and better preserves small convective cells vs ConvLSTM.	[28]
MetNet-2 / large-context CNNs	Deep CNN with larger context + NWP conditioning	Extends MetNet with larger spatial context and NWP inputs to 12-h horizons; produces probabilistic “cubes”.	Shows improved precipitation skill compared with high-res NWP (e.g. HRRR) across 0–12 h leads.	[13][18]
NowcastNet family (2023–2024)	Hybrid physical–ML nowcasting	Embeds explicit advection / warping inside a neural architecture and focuses on extremes and interpretability.	Demonstrates skilful extreme-precipitation nowcasts and hybrid physics–AI gains over NWP and pure ML.	[1][2]
PreDiff / diffusion models (2023+)	Latent diffusion generative models	Models full predictive distributions via iterative denoising in a learned latent space; sharp, realistic radar fields.	On SEVIR and related datasets, produces sharper, more realistic rain structures than GAN baselines while giving calibrated uncertainty.	[6][27]

IV. OPPORTUNITIES

AI-based nowcasting creates several major opportunities:

1) Improved short-lead forecast skill: Deep learning models consistently outperform optical-flow extrapolation and often outperform NWP for 0 –2 h convective precipitation [2][3][13]. Gains are particularly large for localized extremes, where additional 30–60 min of reliable lead time can translate directly into saved lives and reduced damage.

2) Multisensor fusion and global coverage: AI architectures naturally ingest heterogeneous inputs, enabling quantitative fusion of radar, satellite, lightning, gauges and NWP environment [4][5][15][16]. Satellite-driven AI systems can provide useful nowcasts in radar-sparse regions, significantly improving forecast equity worldwide [5].
3) Probabilistic, user-oriented products: Generative models and ensembles provide full probability distributions rather than single deterministic forecasts [3][6][19]. This supports risk-based decision-making and more transparent communication of uncertainty.
4) Application-specific nowcasts: AI models can be tailored to aviation, hydrology, renewable energy, urban flooding and other sectors by optimizing for sector-specific targets (e.g. probability of runway lightning, probability of exceedance of a flood threshold) [16][17].

Because the same core architectures can be re-trained on different targets, this opens a large space of targeted decision-support products.

Synergy with physics-informed learning: Hybrid physics-AI models such as NowcastNet and PDDN show that coupling ML with NWP and physical constraints can extend useful lead times to 3 –6 h while retaining dynamical consistency [1][2][8]. This suggests a pathway toward seamless integration of nowcasting and short-range forecasting.

V. CHALLENGES

Despite rapid progress, significant challenges must be addressed before AI nowcasting can be fully relied upon in operations.

A. *Physics-informed and hybrid models*

AI models inherit all the imperfections of their training data. Radar fields can contain ground clutter, bright -band artifacts, beam blockage and calibration drifts; satellite rainfall estimates are noisy and biased; gauge networks are sparse and uneven [10][14]. If these issues are not carefully handled, models may learn spurious patterns (e.g. always predicting rain near a radar range edge) rather than true meteorology.

Global coverage is highly uneven. Large parts of Africa, South America and the oceans lack dense radar networks [5]. Models trained in data -rich mid -latitude regions may not generalize to tropical regimes or sparsely observed areas. Satellite -based AI helps, but satellite retrievals themselves are uncertain and climate -dependent.

Non-stationarity is another concern: climate change is altering the frequency and intensity of extreme events, and observing systems evolve in time. A model trained on historic data may become sub -optimal as the underlying distribution shifts. Continual learning, transfer learning and routine re -training are promising but operationally non -trivial.

B. *Models Class Imbalance and Extremes*

Precipitation fields are extremely imbalanced: most pixels at most times have zero or light rain, while the high -impact extremes occupy a tiny fraction of space –time [4][11]. Standard losses like MSE or MAE focus on minimizing average error and under - emphasize rare heavy rainfall. As a result, naïve deep learning nowcasts often smooth out or miss intense convective cores.

To address this, researchers use weighted losses, focal losses, threshold -oriented losses (e.g. Maximizing CSI for specific rain thresholds), and specialized distributions such as Tweedie or compound Poisson losses tailored to zero -inflated, heavy -tailed rainfall [11]. Data -level techniques such as oversampling rainy cases or augmenting extreme events are also applied.

Nevertheless, accurately predicting the location and timing of extremes remains very difficult, both because they are rare in training data and because they are inherently less predictable.

C. *Generalization and transferability*

Many AI nowcasting models are trained on a specific region, season or radar network. Domain shifts —different climate regimes, orography, observing systems or microphysical characteristics —can markedly degrade performance. A model trained on U.S. Great Plains convection may struggle with tropical cyclones, monsoon convection or winter stratiform precipitation.

Possible strategies include:

- Global or multi -regional training with regional conditioning [5][13]. Transfer learning: pretraining on large global datasets then fine -tuning on local data [15][19].
- Self -supervised learning on massive unlabelled archives to learn general weather representations [15].

Developing models that are both globally applicable and locally adaptable is key research frontier.

D. Interpretability and trust

Operational meteorologists must understand, to some degree, why a model is predicting a hazardous event in order to trust it and diagnose failures. Traditional extrapolation is conceptually simple, and NWP provides physically interpretable fields (e.g. CAPE, shear, convergence). In contrast, most deep networks are black boxes with millions of parameters.

Explainable AI tools —saliency maps, feature attribution, analog -based retrievals —can help illuminate which input regions or features drive a given forecast [4]. Physics -informed architectures also provide partial interpretability because some components correspond to recognizable physical operators [1][8]. Nonetheless, building trust will require extensive validation, careful documentation of failure modes, and training for end -users.

E. Evaluation, Metrics and Benchmarks

Precipitation nowcasting is difficult to evaluate because small spatial displacement errors can lead to large pointwise errors (the "double penalty" problem). Simple metrics such as MSE and correlation are therefore inadequate for high -resolution precipitation fields. The community increasingly relies on threshold -based scores (CSI, POD, FAR), scale -aware metrics (fractions skill score, SAL), and probabilistic scores (CRPS, Brier score, reliability diagrams) to fully characterize performance [4][21][22].

A major challenge is lack of standardization. Different studies use different datasets, thresholds, scales and metrics, hindering fair comparison. Recent surveys call for open benchmarks and shared testbeds for AI nowcasting [4]. Libraries like PySTEPS already provide baseline extrapolation methods and verification tools [10], but consistent, widely accepted benchmark datasets are still emerging.

F. Evaluation, Metrics and Benchmarks

Deploying AI nowcasting in real time raises practical and ethical questions:

1) Latency and cost: Some advanced transformer or diffusion models are computationally intensive. Operational centers must ensure that inference can keep pace with observation updates using available hardware [6][18].

2) Robustness: Models must behave sensibly in edge cases (e.g. missing radar tiles, sensor glitches). Fallback strategies and sanity checks are needed to avoid implausible outputs entering warning systems.

3) Human –AI interaction: Forecasters need guidance on when and how to use AI nowcasts, and how to integrate them with NWP and conceptual models [7].

4) Fairness and equity: Systems should perform adequately across all regions and populations, not just data -rich urban centers [5].

Transparency and accountability: For high -impact decisions, agencies must understand model limitations, version changes, and responsibilities if AI guidance contributes to missed events or false alarms.

VI. METHODOLOGICAL TRENDS AND RESEARCH FRONTIERS

AI for nowcasting is evolving rapidly. Key research directions include:

1) Foundation models and self -supervised learning: Large, generic "weather foundation models" trained on heterogeneous global data (e.g. ClimaX, FourCastNet, GraphCast, Pangu -Weather) demonstrate that a single architecture can support many tasks after fine-tuning [15][19][20]. Extending such models down to nowcasting scales is an active area.

2) Multimodal and multi -sensor fusion: Architectures that jointly ingest radar, satellite, lightning, gauges and NWP fields using cross -attention, multi -branch encoders or graph structures aim to exploit complementary strengths of each dataset [4][5][16][17].

3) Physics -informed deep learning: Embedding conservation laws, symmetry constraints, differentiable advection operators or NWP -like modules into AI architectures promises better generalization and physical consistency [1][2][6][8].

4) Uncertainty quantification and extremes: Generative models, Bayesian techniques and tailored loss functions are being explored to deliver well -calibrated probabilities and improved representation of extreme events [6][11][19].

5) Continual and transfer learning: Methods that allow models to update incrementally as new data arrive, while avoiding catastrophic forgetting, are crucial for adapting to changing climates and observing systems.

6) Benchmarks and open science: Community datasets, leaderboards and open -source reference implementations are increasingly recognized as vital infrastructure for progress [4][10][19].

VII. AI VS. TRADITIONAL NOWCASTING

It is useful to compare AI nowcasting qualitatively against two traditional approaches: optical -flow extrapolation and convection -permitting NWP.

1) Forecast skill (0 –2 h): Deep learning systems generally outperform extrapolation and often outperform NWP for high -resolution convective precipitation, especially on structural and threshold -based metrics [2][3][13].
2) Forecast skill (>3 –6 h): Purely data -driven nowcasts typically lose skill beyond a few hours as chaos and large -scale dynamics dominate. NWP, with full physical equations and data assimilation, remains superior at these lead times [18][19][20].
3) Physical consistency: NWP enforces conservation and dynamical balance by design. Optical -flow extrapolation preserves observed reflectivity but does not simulate physics. AI models may violate physical constraints unless specifically regularized, though hybrid systems mitigate this [1][8].
4) Computational cost: Once trained, AI models are very fast at inference, often much cheaper than running a full NWP cycle, particularly at very high resolution [3][5][13]. Extrapolation is cheapest but less capable; NWP is the most expensive.
5) Interpretability: Extrapolation and NWP provide clear, physically grounded reasoning pathways. AI models are currently less interpretable, though physics -aware designs and XAI tools help.
6) Global equity: Satellite -driven AI can provide high -frequency, higher -resolution guidance in data-sparse regions where local NWP and radar networks are limited [5]. Traditional radar -based extrapolation is not possible where radars do not exist.

In practice, the most promising operational paradigm is hybrid: use extrapolation for the first tens of minutes, AI nowcasts for 0 –2 (or 0 –3) h, and gradually transition to NWP beyond that, possibly with AI assisting in blending and post -processing. Human forecasters remain central in synthesizing guidance and managing warnings.

VIII. DISCUSSION

In practice, the most promising operational paradigm is hybrid: use extrapolation for the first tens of minutes, AI nowcasts for 0 –2 (or 0 –3) h, and gradually transition to NWP beyond that, possibly with AI assisting in blending and post -processing. Human forecasters remain central in synthesizing guidance and managing warnings. The emergence of AI nowcasting has generated both excitement and hype. High -profile results from DGMR, MetNet, NowcastNet and others demonstrate real, substantial gains in short -range precipitation skill [1][2][3][5][13]. At the same time, operational uptake is still in early stages. Most national weather services are cautiously experimenting with AI guidance alongside established methods rather than replacing them.

A. *Key Themes in the Current Discourse Include*

1) Practical impact vs. hype: While AI has produced impressive case studies, widespread operational transformation will depend on sustained validation, robust engineering, and user training.
2) Standardization and reproducibility: The community needs shared benchmarks, clear baselines and open code/data where possible to compare methods fairly and avoid redundant effort [4][10].
3) Interdisciplinary collaboration: The most successful systems have arisen from close collaboration between meteorologists, hydrologists and ML researchers [1][2][8][16][17].
4) Responsible deployment: Issues of equity, transparency, and accountability must be addressed explicitly when AI nowcasts influence public warnings and disaster risk management [5].

IX. CONCLUSION AND FUTURE DIRECTIONS

In practice, the most promising operational paradigm is hybrid: use extrapolation for the first tens of minutes, AI nowcasts for 0 –2 (or 0 –3) h, and gradually transition to NWP beyond that, possibly with AI assisting in blending and post -processing. Human forecasters remain central in synthesizing guidance and managing warnings. The emergence of AI nowcasting has generated both excitement and hype. High -profile results from DGMR, MetNet, NowcastNet and others demonstrate real, substantial gains in short -range precipitation skill [1][2][3][5][13]. At the same time, operational uptake is still in early stages. Most national weather services are cautiously experimenting with AI guidance alongside established methods rather than replacing them. AI-based nowcasting has rapidly progressed from proof -of-concept ConvLSTMs to sophisticated hybrid diffusion and transformer systems. These models leverage large archives of radar, satellite and NWP data to produce detailed, frequently updated, and increasingly probabilistic precipitation forecasts. Evidence from multiple studies indicates that, for 0 –2 h convective rainfall, AI systems can deliver higher skill than both optical -flow extrapolation and high -resolution NWP in many settings [2][3][13]. Physics -informed and hybrid architectures extend useful lead times and improve robustness for extremes [1][2][6][8].

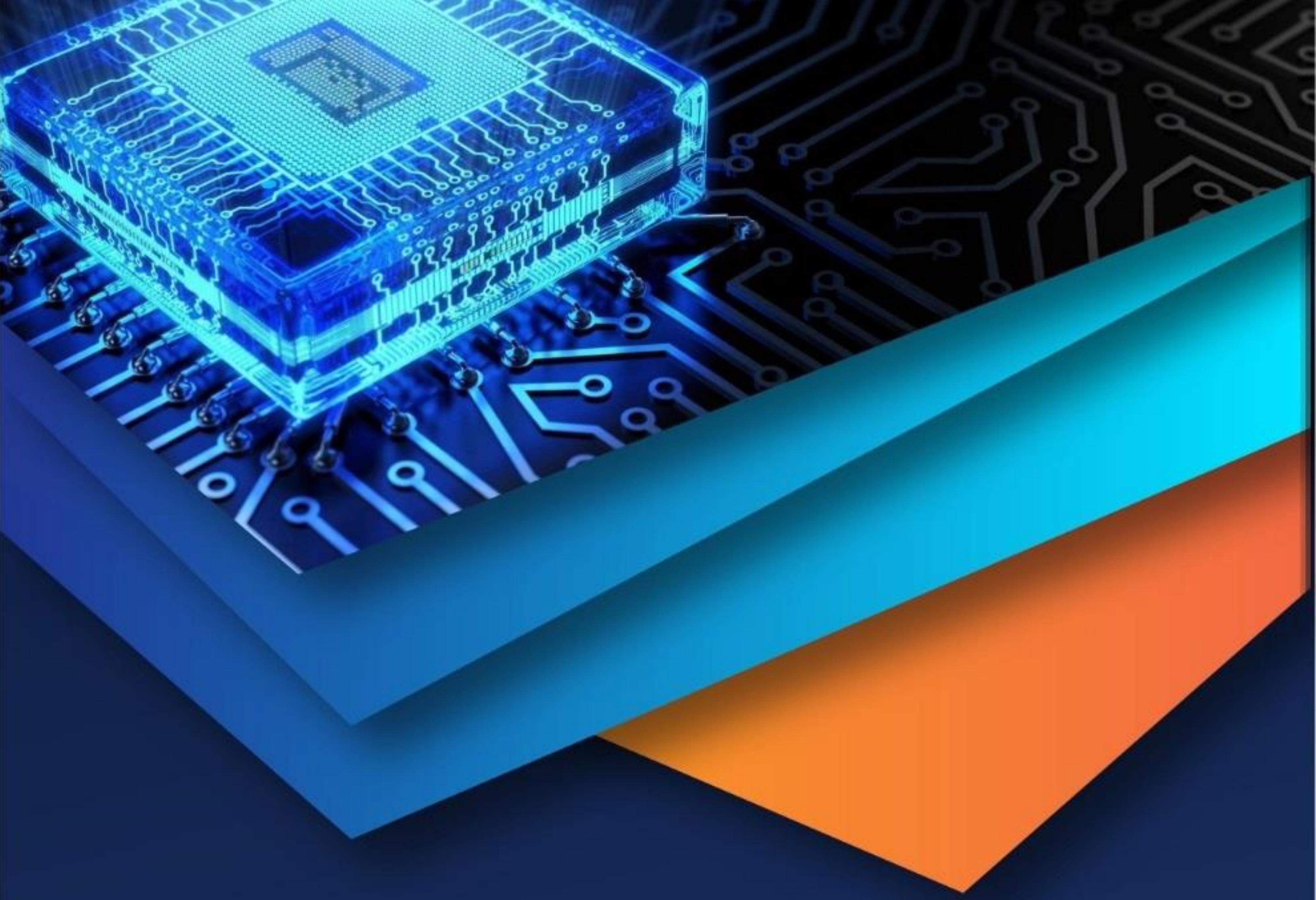
At the same time, AI is not a magic bullet. Data limitations, non -stationarity, class imbalance, generalization issues, interpretability challenges and operational constraints remain active research and engineering problems. To fully realize the promise of AI nowcasting, the community should:

1) Invest in high -quality, multi -sensor datasets and open benchmarks.
2) Embrace hybrid physics -AI designs that respect known dynamics.
3) Develop standardized verification practices, especially for probabilistic products and extremes.
4) Focus on interpretability, user training, and human –AI teaming in forecast offices.
5) Prioritize equity, ensuring that advances benefit data -sparse, vulnerable regions as well as data-rich ones [5].
6) Design governance frameworks that clarify responsibilities and manage model updates transparently.

For PhD -level research, open questions include optimal multi -sensor fusion strategies, foundation -model approaches for high -resolution nowcasting, principled physics -ML coupling, robust uncertainty quantification, continual learning under climate change, and human -centric design of decision -support products. Progress on these fronts could yield AI nowcasting systems that are not only more accurate, but also more trustworthy, interpretable and impactful.

REFERENCES

[1] Zhang, Y., Long, M., Chen, K. et al. Skilful nowcasting of extreme precipitation with NowcastNet. Nature 619, 526–532 (2023). https://doi.org/10.1038/s41586-023-06184-4.
[2] Das, P., Posch, A., Barber, N. et al. Hybrid physics-AI outperforms numerical weather prediction for extreme precipitation nowcasting. npj Clim Atmos Sci 7, 282 (2024). https://doi.org/10.1038/s41612-024-00834-8.
[3] Ravuri, S., Lenc, K., Willson, M. et al. Skilful precipitation nowcasting using deep generative models of radar. Nature 597, 672–677 (2021). https://doi.org/10.1038/s41586-021-03854-z
[4] S. An et al., “Deep learning for precipitation nowcasting: A survey from the perspective of time series forecasting,” arXiv preprint arXiv:2406.04867, 2024.
[5] S. Agrawal et al., “An operational deep learning system for satellite -based high -resolution global nowcasting,” arXiv preprint arXiv:2510.13050, 2025.
[6] Leinonen, Jussi & Hamann, Ulrich & Nerini, Daniele & Germann, Urs & Franch, Gabriele. (2023). Latent diffusion models for generative precipitation nowcasting with accurate uncertainty quantification. 10.48550/arXiv.2304.12891.
[7] S. Zhao, S. Saha, Z. Xiong, N. Boers and X. X. Zhu, "Exploring Geometric Deep Learning for Precipitation Nowcasting," IGARSS 2023 - 2023 IEEE International Geoscience and Remote Sensing Symposium, Pasadena, CA, USA, 2023, pp. 3760-3763, doi: 10.1109/IGARSS52108.2023.10282387.
[8] Wang, R., Fung, J. C. H., & Lau, A. K. H. (2024). Skillful precipitation nowcasting using physical-driven diffusion networks. Geophysical Research Letters, 51, e2024GL110832. https://doi.org/10.1029/2024GL110832.
[9] L. Bouchard, “DeepMind uses AI to predict more accurate weather forecasts,” Louis Bouchard AI Blog, Oct. 2021. [Online]. Available: https://www.louisbouchard.ai/deepmind-rain-nowcasting/.
[10] S. Pulkkinen et al., “Pysteps: an open-source Python library for probabilistic precipitation nowcasting (v1.0),” Geosci. Model Dev., vol. 12, pp. 4185–4219, 2019. doi: 10.5194/gmd-12-4185-2019
[11] Hunt, Kieran. (2025). Stop using root-mean-square error as a precipitation target! 10.48550/arXiv.2509.08369.
[12] X. Shi et al., “Convolutional LSTM network: A machine learning approach for precipitation nowcasting,” in Proc. NIPS, 2015, pp. 802 –810.
[13] B. Sønderby et al., “MetNet: A neural weather model for precipitation forecasting,” arXiv preprint arXiv:2003.12140, 2020.
[14] Y. Wang et al., "PredRNN: A Recurrent Neural Network for Spatiotemporal Predictive Learning," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 45, no. 2, pp. 2208-2225, 1 Feb. 2023, doi: 10.1109/TPAMI.2022.3165153.
[15] Nguyen, Tung & Brandstetter, Johannes & Kapoor, Ashish & Gupta, Jayesh & Grover, Aditya. (2023). ClimaX: A foundation model for weather and climate. 10.48550/arXiv.2301.10343.
[16] L. Espeholt et al., “Deep learning for twelve hour precipitation forecasts,” Nat. Commun., vol. 13, art. no. 5145, 2022, doi: 10.1038/s41467-022-32483-x.
[17] Li, Y., Liu, Y., Sun, R. et al. Convective Storm VIL and Lightning Nowcasting Using Satellite and Weather Radar Measurements Based on Multi-Task Learning Models. Adv. Atmos. Sci. 40, 887–899 (2023). https://doi.org/10.1007/s00376-022-2082-6
[18] L. Espeholt et al., “Skillful Twelve Hour Precipitation Forecasts using Large Context Neural Networks,” arXiv preprint arXiv:2111.07470, 2021.
[19] L. Lam et al., “GraphCast: Learning skillful medium -range global weather forecasting,” arXiv preprint arXiv:2212.12794, 2022.
[20] H. Bi et al., “Pangu -Weather: A 3D high -resolution model for fast and accurate global weather forecast,” arXiv preprint arXiv:2211.02556, 2022.
[21] K. A. Browning and C. G. Collier, “Nowcasting of precipitation systems,” Reviews of Geophysics, 27(3), 345–370, 1989.
[22] U. Germann and I. Zawadzki, “Scale-dependence of the predictability of precipitation from continental radar images. Part I: Description of the methodology,” Monthly Weather Review, 130(12), 2859–2880, 2002.
[23] X. Shi et al., “Deep learning for precipitation nowcasting: A benchmark and a new model,” in Advances in Neural Information Processing Systems 30 (NeurIPS), 2017, pp. 5617–5627.
[24] V. Lebedev et al., “Precipitation nowcasting with satellite imagery,” in Proc. 25th ACM SIGKDD Int. Conf. on Knowledge Discovery & Data Mining (KDD), 2019, pp. 2680–2688.
[25] G. Ayzel, T. Scheffer, and M. Heistermann, “RainNet v1.0: A convolutional neural network for radar-based precipitation nowcasting,” Geoscientific Model Development, 13(6), 2631–2644, 2020.
[26] M. S. Veillette et al., “SEVIR: A Storm Event Imagery Dataset for Deep Learning Applications in Radar and Satellite Meteorology,” in NeurIPS Datasets and Benchmarks Track, 2020.
[27] Z. Gao et al., “PreDiff: Precipitation Nowcasting with Latent Diffusion Models,” in Advances in Neural Information Processing Systems (NeurIPS), 2023.
[28] X. Bai et al., “Rainformer: Features extraction balanced network for radar-based precipitation nowcasting,” IEEE Geoscience and Remote Sensing Letters, 19, 1–5, 2022



doi crossref

10.22214/IJRASET

INDEX COPERNICUS

45.98

ISRA JIF

IMPACT FACTOR: 7.129

SJIFactor Scientific Journal Impact Factor TOGETHER WE REACH THE GOAL

IMPACT FACTOR: 7.429

ISI

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY