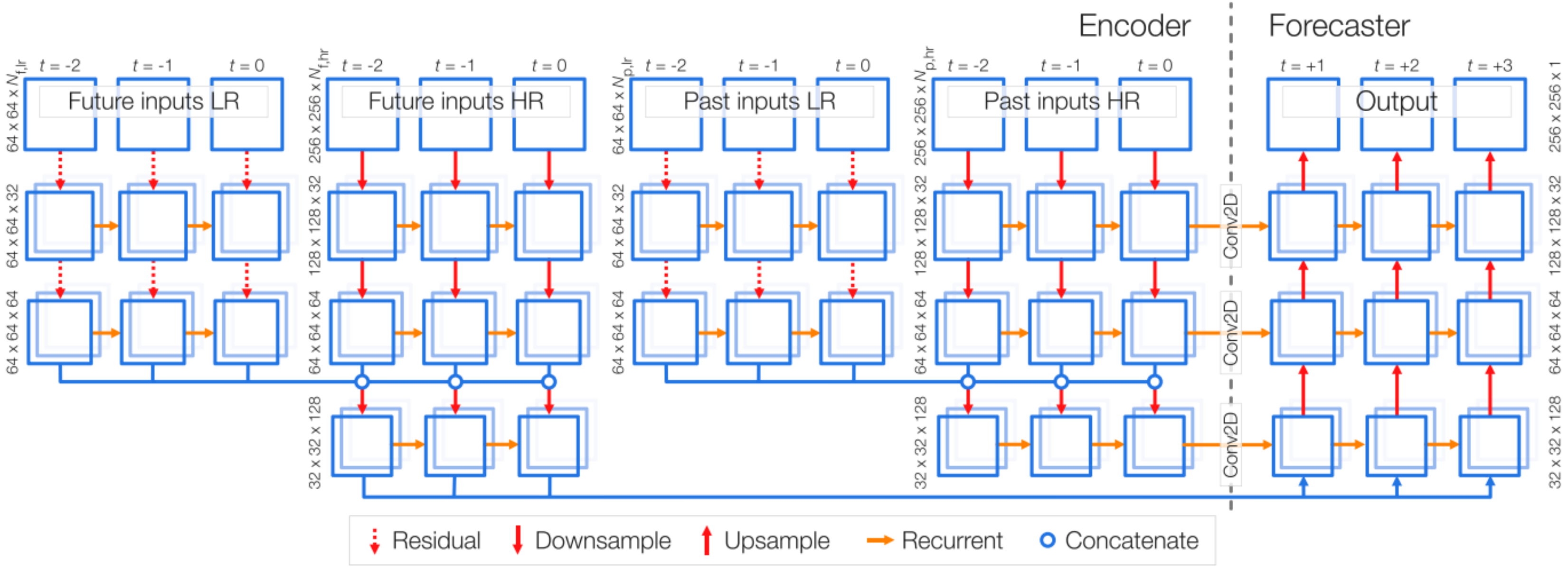
Multi-hazard nowcasting system over Romania

5th ECMWF-ESA Machine Learning Workshop

Claudiu Adam

Bologna, Italy | April 15th 2026

Model architecture – COALITION-4



Leinonen, J., Hamann, U., & Germann, U. (2022). Seamless Lightning Nowcasting with Recurrent-Convolutional Deep Learning. Artificial Intelligence for the Earth Systems, 1(4). https://doi.org/10.1175/aies-d-22-0043.1

Product importance analysis

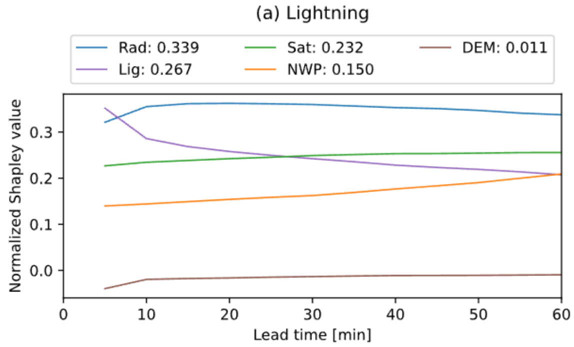
Note: The following analysis was performed on the 12-timestep output variant of the model (5-minute timesteps, up to 60 minutes), which was discarded due to operational constraints — the requirement to ingest future timesteps during inference, **making the architecture incompatible with real-time nowcasting deployment considering NMA's constraints.**

The training dataset consisted of radar, LINET, MSG SEVIRI, NWCSAF, and NWP data.

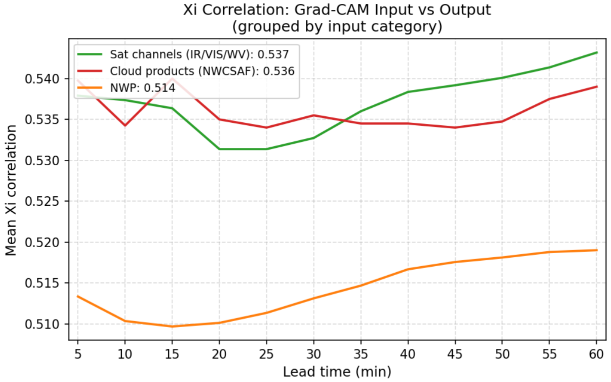
"Why Xi correlation on Grad-CAM maps instead of Shapley values?"

1) Shapley's has a hidden decision to consider, in which to represent the "absence" of a variable (replaced with channel mean, zeros, sampled from distribution etc.), and results are sensitive to this choice. Xi on Grad-CAM requires no equivalent assumption.

2) Shapley values as applied in COALITION-4 collapse feature importance to a scalar per input group, discarding all spatial information. Xi on Grad-CAM operates on full activation maps, preserving the spatial structure of where the model learns within each input — could also be used for bias correction.



Leinonen, J., Hamann, U., Sideris, I. v., & Germann, U. (2022). Thunderstorm nowcasting with deep learning: a multi-hazard data fusion model. https://doi.org/10.1029/2022GL101626

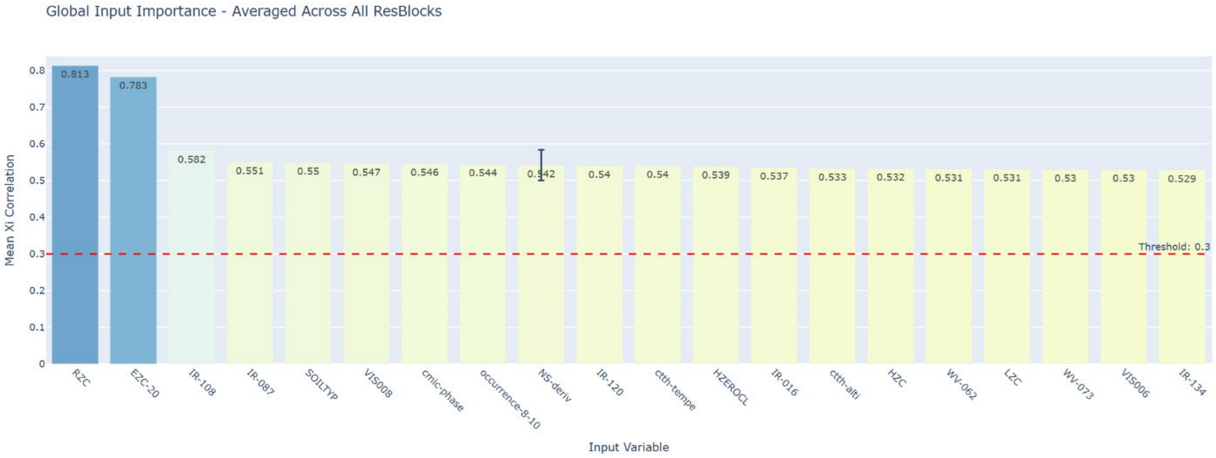


Grad-CAM & Xi methodology

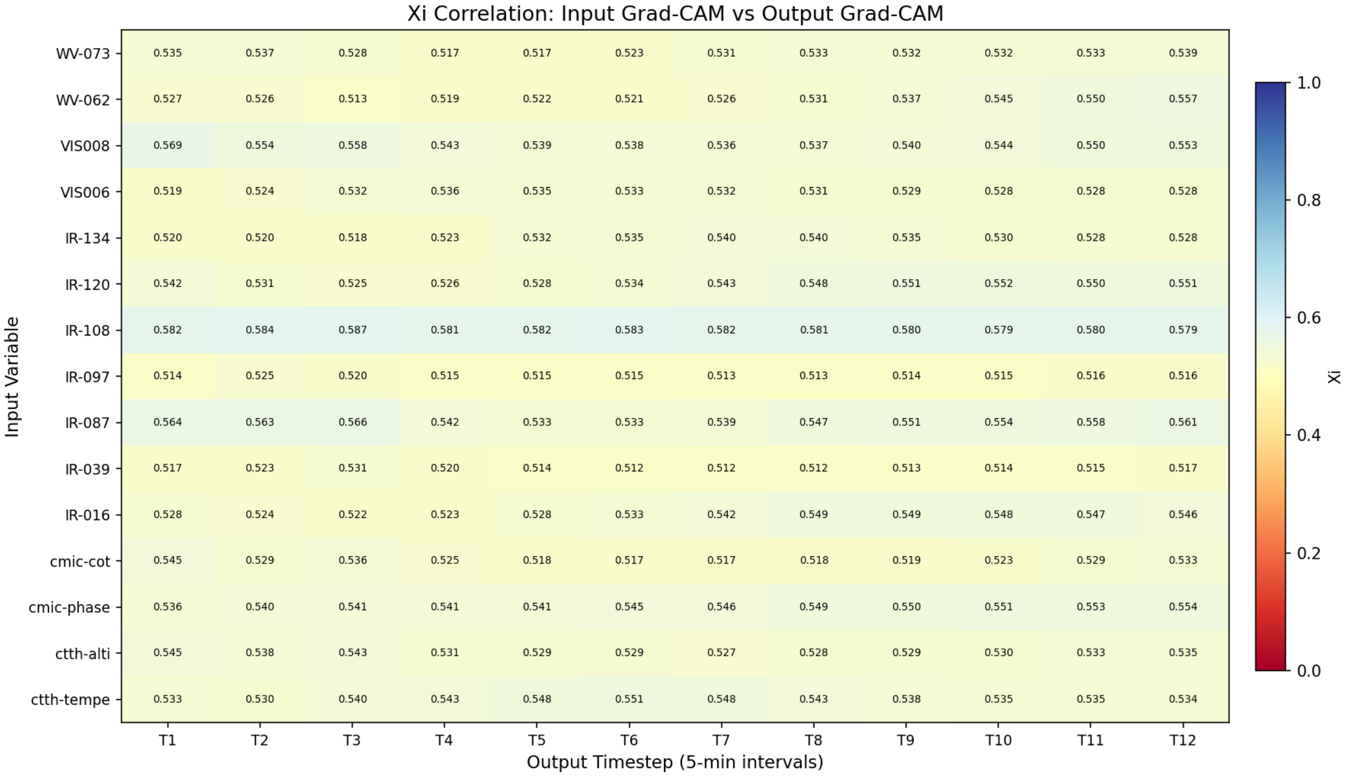
- Analysis of variable importance through activation maps generated by the Gradient-weighted Class Activation Mapping (Grad-CAM) algorithm. Analysis was performed on all 40 model products (the first 20 in the image below)
- Xi correlation for nonlinear dependence measurement (0=independent, 1=dependent)
- Combined approach for correlating input and output activation patterns
- Features below threshold ξ < 0.52 removed for increased model efficiency

Global ranking reveals radar products far exceed others

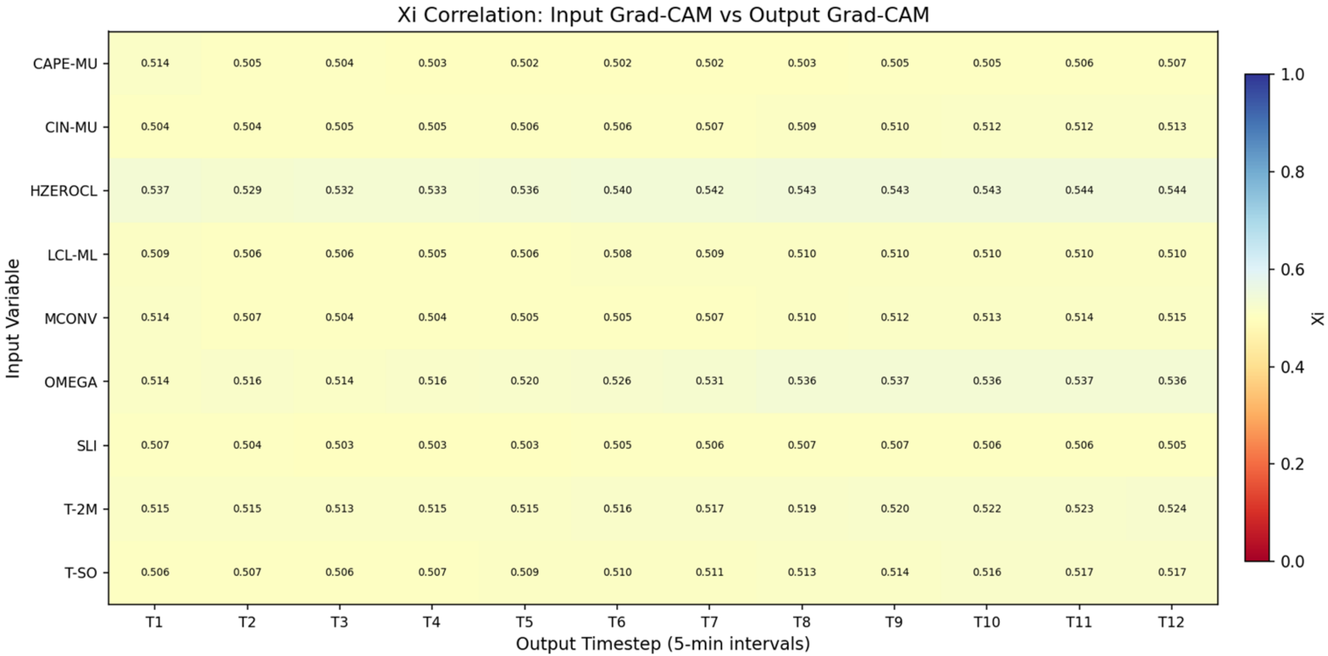
$\xi_{RZC} = 0.813$, $\xi_{EZC\text{-}20} = 0.783$



Feature importance results – MSG and NWCSAF products

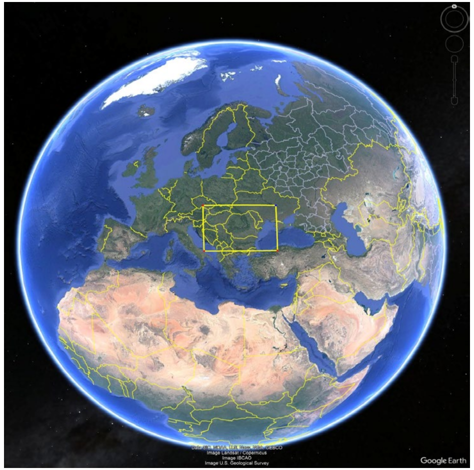


Feature importance results – ICON products

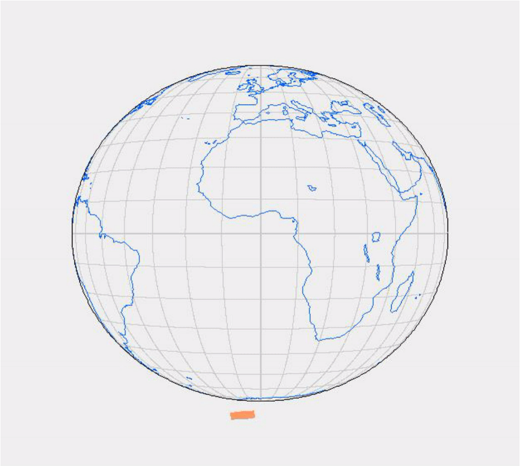


Data gathering — satellite data acquisition

- Full disk at 10-min cadence, 5 channels. Romania intersects only chunks 34–38 out of 41 total. ~90% download reduction.
- Study area lat/lon bounds are projected into the geostationary scanning frame to compute row indices. Only the chunk subset covering Romania is downloaded, with one additional chunk N/S as a safety buffer.



Study area (yellow box over Romania)



MTG FCI full disk scanning

Data gathering — data sources

“What multi-source observational streams feed the nowcasting pipeline?”

- **2 training datasets generated – one using radar, LINET, MSG SEVIRI, NWCSAF and one using radar, LINET, MTG FCI, NWCSAF**
- **5 romanian national radar mosaic products reprojected to EPSG:31700 Stereo70 at 1 km resolution: RZC (rain rate), CZC (composite reflectivity), EZC-20 (echo top at 20 dBZ), LZC (VIL), and CPCH (precipitation accumulation).**
- **Three LINET derived products at 15-min resolution: stroke density, peak current, and binary occurrence maps over Romania.**
- **Cloud top height (ctth_alti), cloud top temperature (ctth_tempe), cloud microphysics phase (cmic_phase), and cloud optical thickness (cmic_cot) derived from MSG.**
- **MSG SEVIRI (5 channels, 3 km, 15-min cadence): VIS006 (0.6 μm), IR_039 (3.9 μm), IR_108 (10.8 μm), WV_062 (6.2 μm), WV_073 (7.3 μm).**
- **MTG FCI (5 equivalent channels, 1–2 km, 10-min cadence): vis_06, ir_38, ir_105, wv_63, wv_73 — spectral equivalents of the MSG channels above. Higher spatial resolution and 10-min native cadence, synthesized to 15-min through interpolation.**
- **OBS: The NWCSAF products (ctth, cmic) are derived from MSG SEVIRI in both dataset configurations (i.e. the MTG dataset uses MSG-derived NWCSAF, not FCI-derived products).**

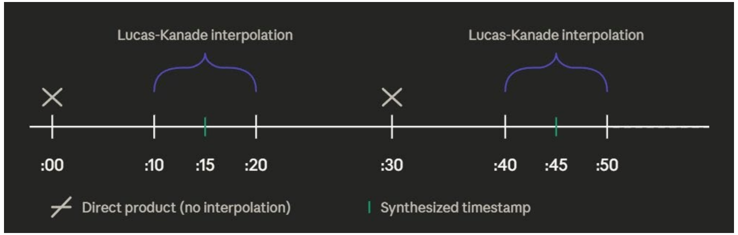
Data reprocessing — temporal alignment

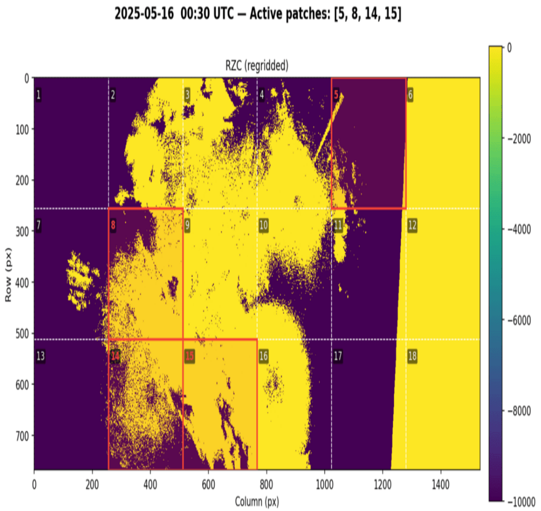
"How do we harmonize the 10-min MTG FCI cadence to the 15-min COALITION grid?"

Direct products at :00 and :30 require no interpolation. Synthesized timestamps at :15 and :45 are generated via Lucas-Kanade optical flow applied to (:10, :20) and (:40, :50) pairs using pysteps.

Error accumulation – better to keep the closest neighbor frame.

Interpolation strategy





Data preprocessing — labels & loss functions

“How are prediction targets structured for the radar and lightning branches?”

- **Rain rate (mm/h) is divided into 5 classes through fixed thresholds: $R<10$, $10≤20$, $20≤30$, $30≤40$, $R≥40$ mm/h. Each pixel becomes a one-hot vector.**

  -> Categorical Crossentropy loss

- **Lightning labels are binary (1 = lightning present, 0 = absent).**

  -> Weighted Focal loss

Data preprocessing — focal loss

$$FL(p_t) = -\alpha_t (1 - p_t)^\gamma \log(p_t)$$

pt* — *predicted probability for the correct class

αt* — *class weight, set inversely proportional to class frequency

$\alpha_1 = 1/(2f)$ — weight for lightning-present pixels ("1" pixels)

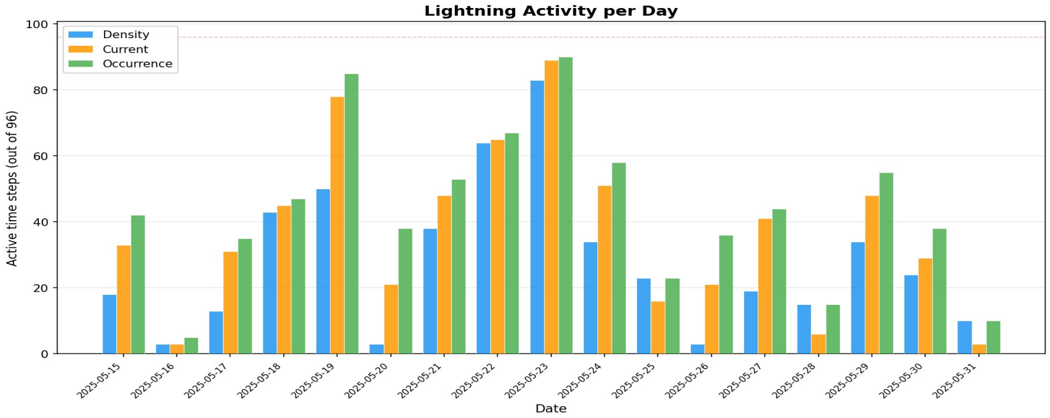
$\alpha_0 = 1/(2(1-f))$ — weight for lightning-free pixels ("0" pixels)

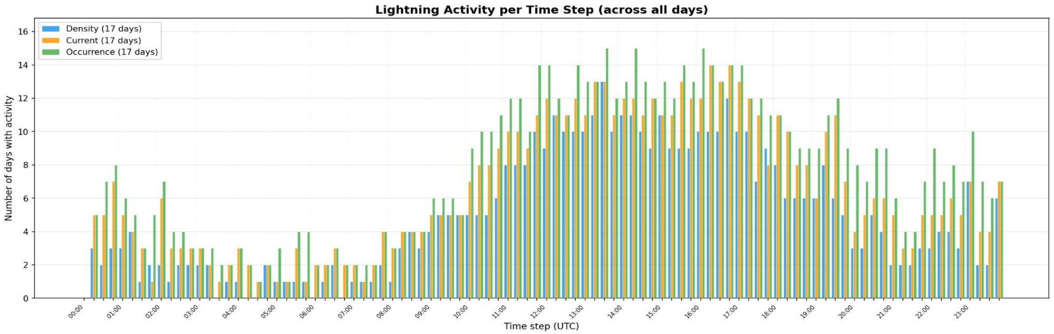
$f \approx 0.00012$ — pre-computed lightning occurrence fraction across the study period

$\gamma = 2$ — focusing parameter; used to penalize well-classified pixels (mostly 0 – no lightning activity)

Data statistics — lightning activity

“When and how does lightning occur across the study period?”





Key observations (15–31 May 2025):

Three LINET products: density, current, occurrence across 17 days.

Peak activity on 23rd of May with >80 active timesteps for all three products.

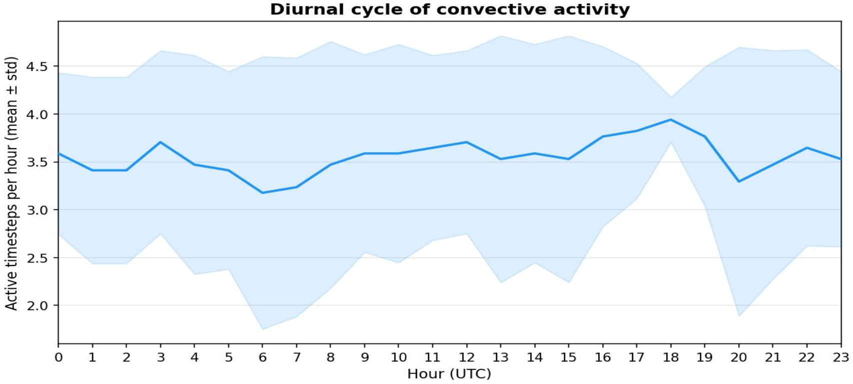
Clear diurnal signal: activity increases from 10 UTC, peaks between 15–18 UTC, then decays.

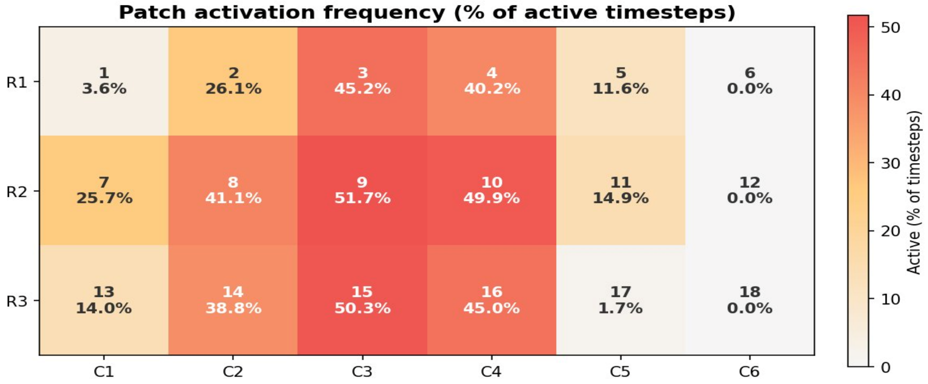
Occurrence product shows highest counts, followed by current and density.

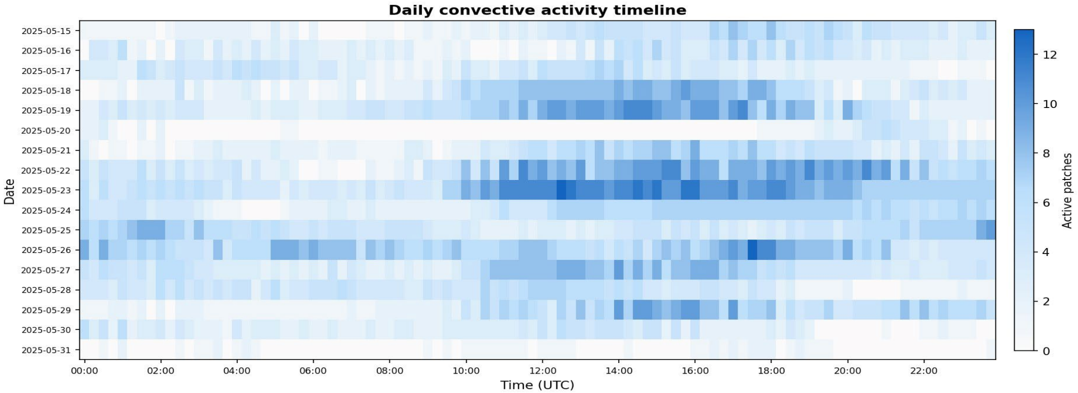
Low-activity days (16 May, 28 May, 31 May).

Data statistics — convective activity (rain rate)

“What is the spatio-temporal structure of convective events?”

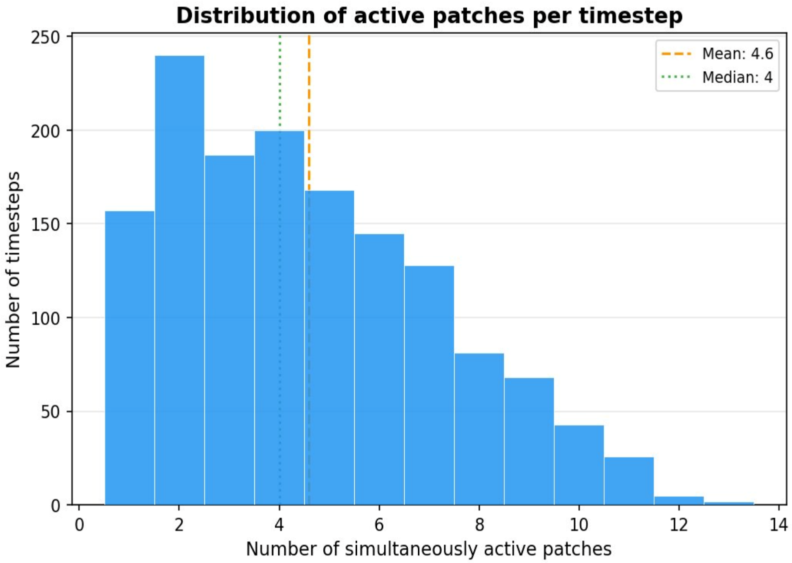


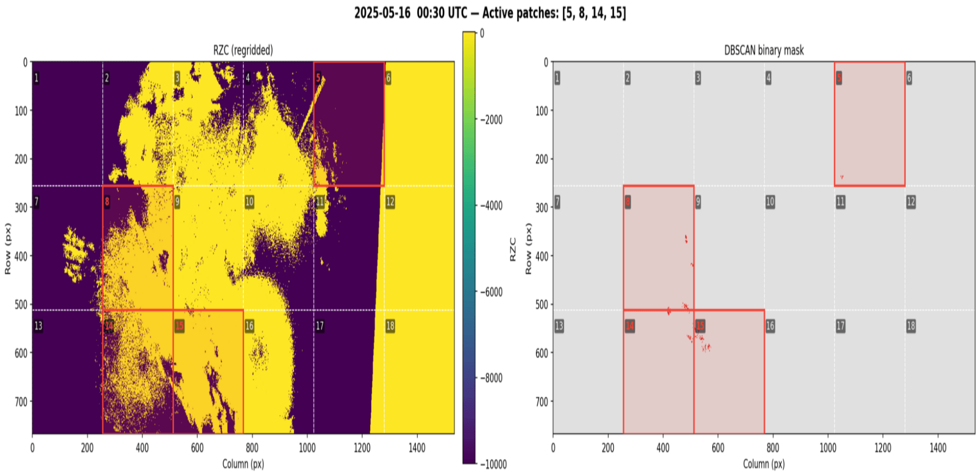




Data statistics — patch extraction and distribution

“How many concurrent convective patches exist and how are they distributed?”



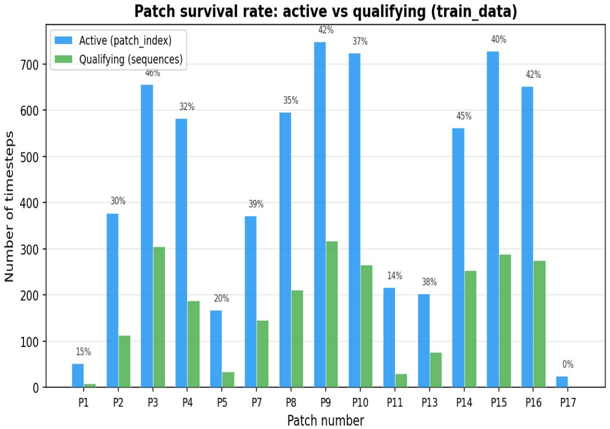


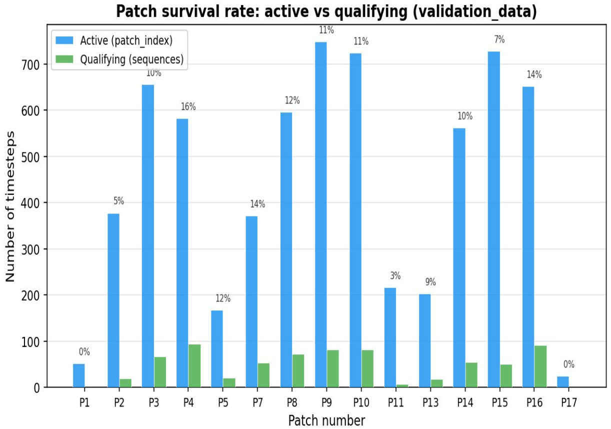
1) DBSCAN clustering applied on rain rate identifies active regions per timestep based on a threshold of >10 mm/h.

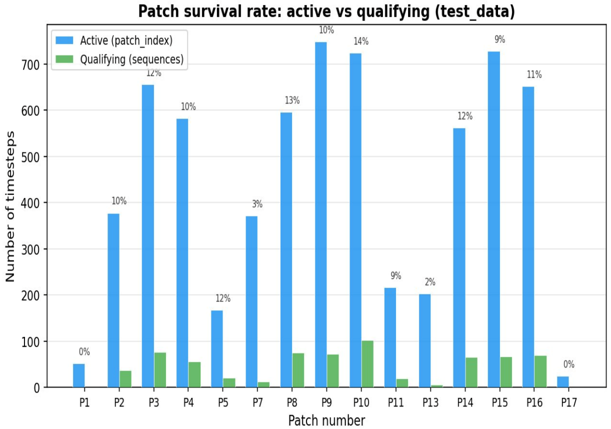
2) After DBSCAN, a binary mask is created from selected active patches and only grid cells that have at least 1 pixel different than 0 are further selected (red boxes 256 x 256 pixels - streamline the extraction of neighboring patches).

Data statistics — patch statistics

“What fraction of active patches produce valid training sequences?”







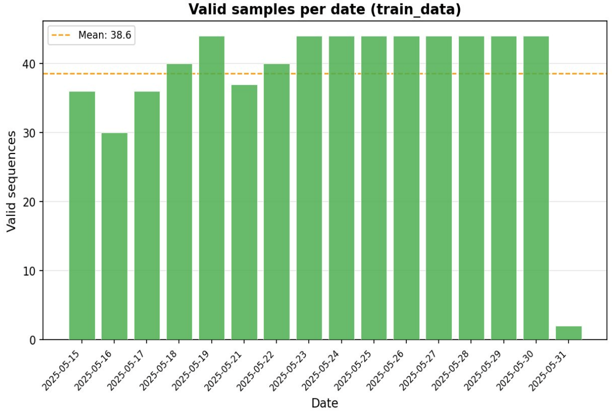
A patch is “qualifying” (survives) if it is active for at least 6 consecutive 15-minute timesteps centered on the reference time:

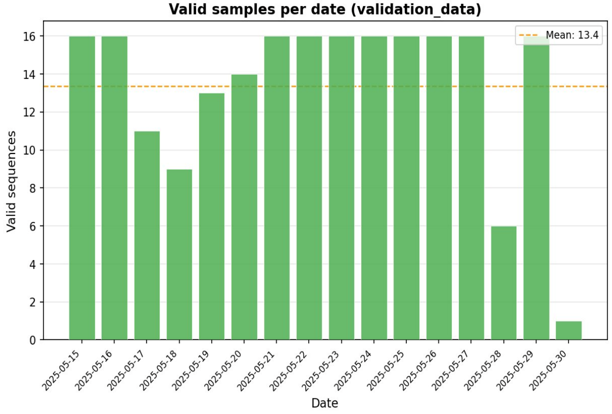
- ***2 past steps (T−30 min, T−15 min)***
- ***current step (T)***
- ***3 future steps (T+15, T+30, T+45 min)***

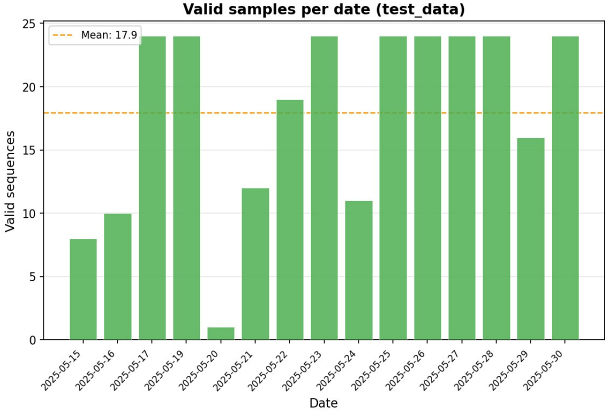
$$\text{survival rate} = \frac{\text{qualifying timesteps}}{\text{active timesteps}} \times 100$$

Data statistics — dataset split

“How are training samples distributed across the study period?”



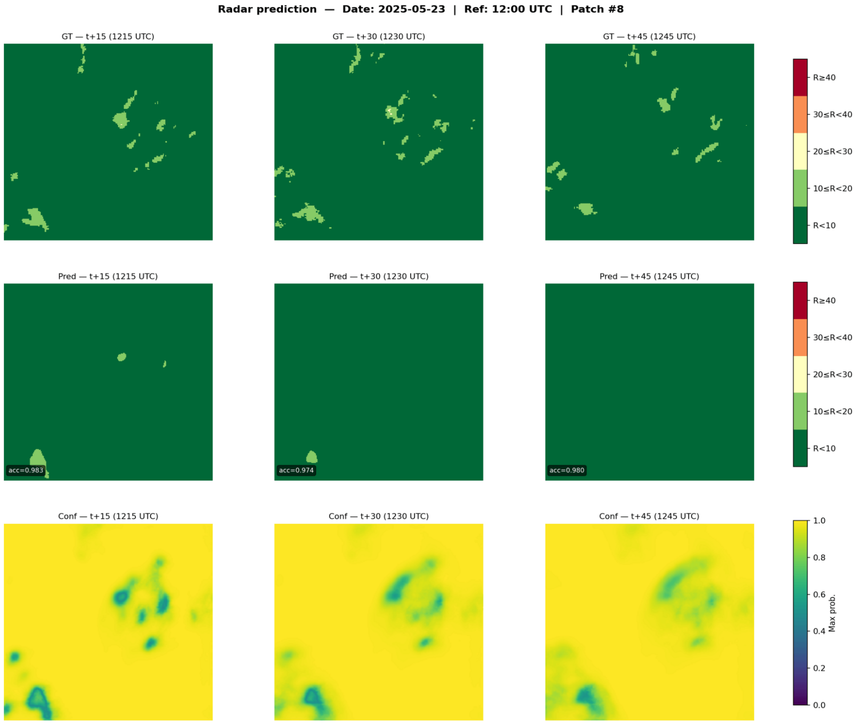


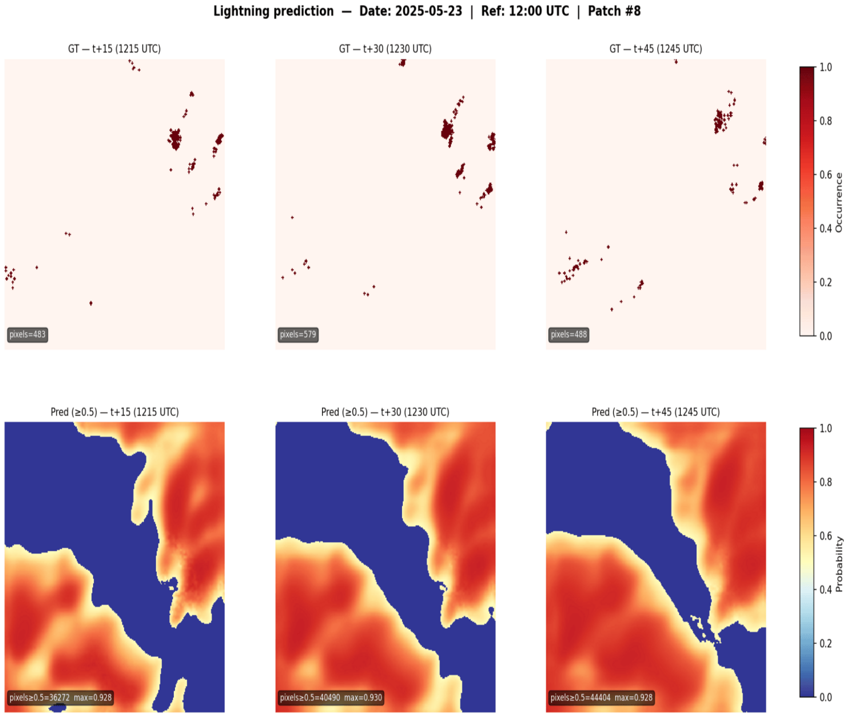


Samples within each day are divided into contiguous temporal blocks: first 10% of each 6-hour period for test, next 10% for validation, remaining 80% for training.

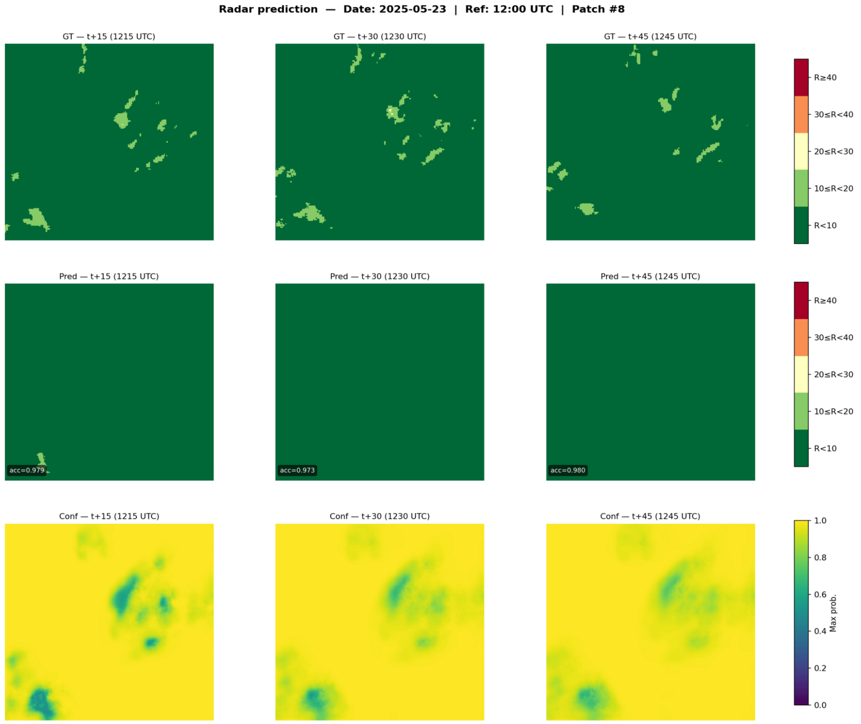
This preserves temporal ordering within each block and prevents data leakage from correlated consecutive timesteps.

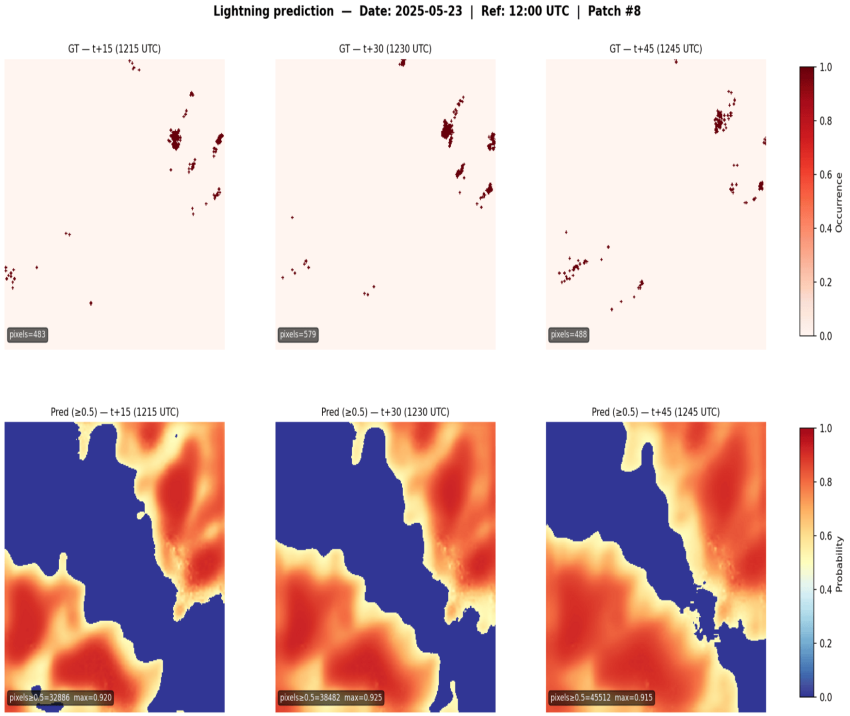
Qualitative and quantitative results — Lightning predictions and rain rate prediction using MSG SEVIRI data





Qualitative and quantitative results — Lightning predictions and rain rate prediction using MTG FCI data





Qualitative and quantitative results — MSG vs MTG

“What are the observed advantages and current drawbacks of MTG over MSG?”

MTG advantages

- Higher spatial resolution
- Enhanced multi-spectral information for convective signatures
- Lightning predictions show slightly more spatially coherent probability fields

Current drawbacks

- Shorter operational data availability
- Larger file sizes require chunk-filtering strategy for efficient data ingestion
- Optical flow interpolation introduces synthetic timestamps (errors)
- Radar predictions dominated by the R<10 mm/h class — model struggles to predict higher-intensity classes in both MSG and MTG
- Lightning prediction overestimates spatial extent (high false positive area)

Future work

- **Increase the dataset size** with sample diversity to improve generalization.
- Investigate attention-based options derived from **Transformer self-attention** to capture longer spatiotemporal dependencies.
- Apply **optical flow-based guidance** (similar to the NWP inputs from the original COALITION-4 work) to model predictions as a post-processing step to increase nowcasting skill scores.
- **Deploy the trained model within NMA's operational framework**, conducting real-time evaluation during the 2026 convective season.