A deep learning approach to evaluate individual predictors for extreme precipitation in Greece

Authors: Vasileios Vatellis, Stelios Karozis, Iraklis A. Klampanos, Antonis Troumpoukis and Antonis Gkanios

Overview

- Motivation
- Problem statment
- Related work
- Data
- Methodology
- Results
- Future work

Motivation

- Global warming
- Rise of extremes (floods, flashfloods, rainstorms)
- Forecasting of local hazard/impacts
- Enabling proactive adaptation.

Image: https://edition.cnn.com/2023/09/07/europe/greece-floods-storm-rescue-climate-intl

Problem Statement

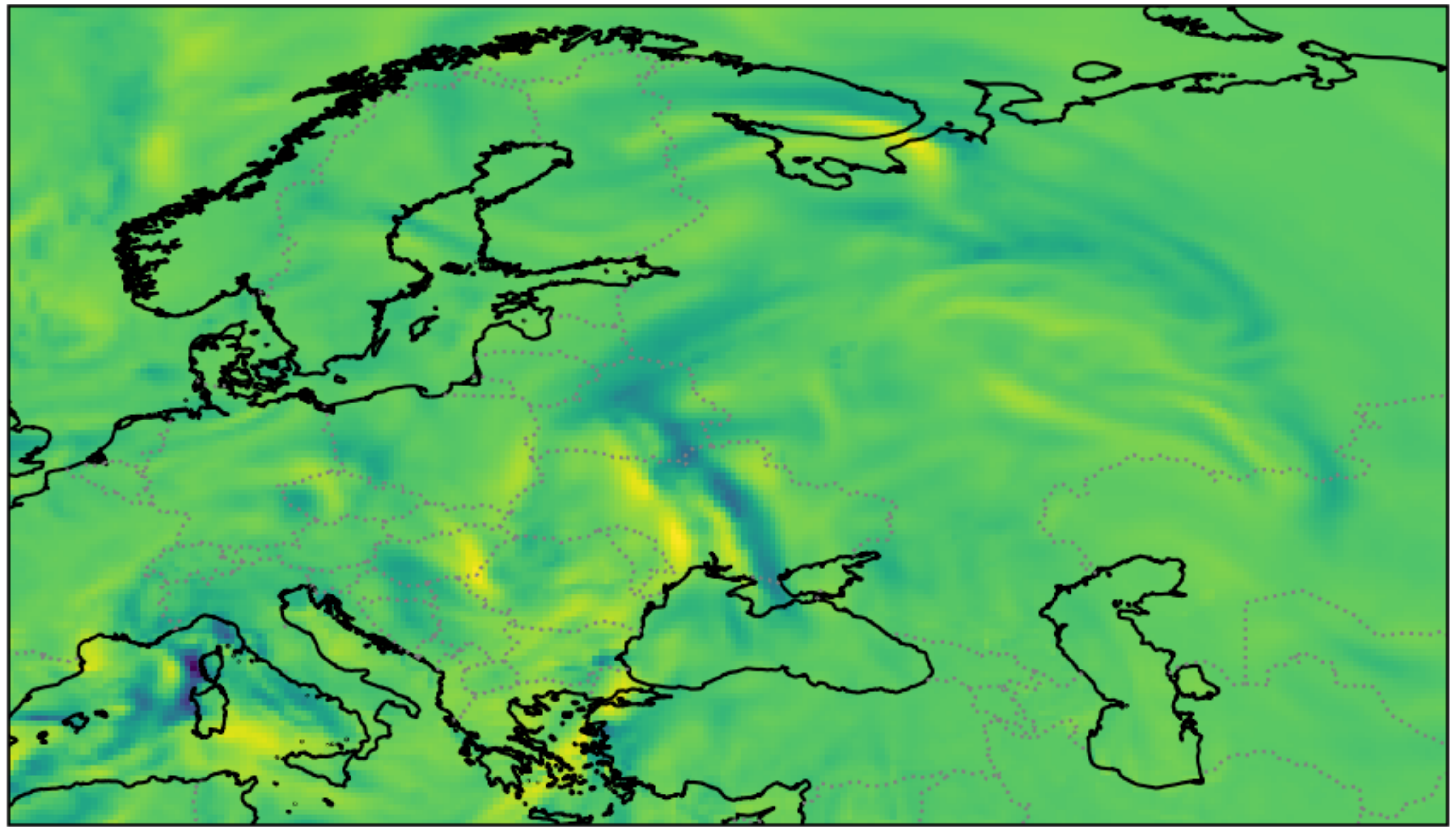
Can AI forecast extreme precipitation?

- ML/DL can forecast “normal” precipitation
- Extreme precipitation events are rare and high-impact
- **which predictors are the most informative for extreme precipitation?**

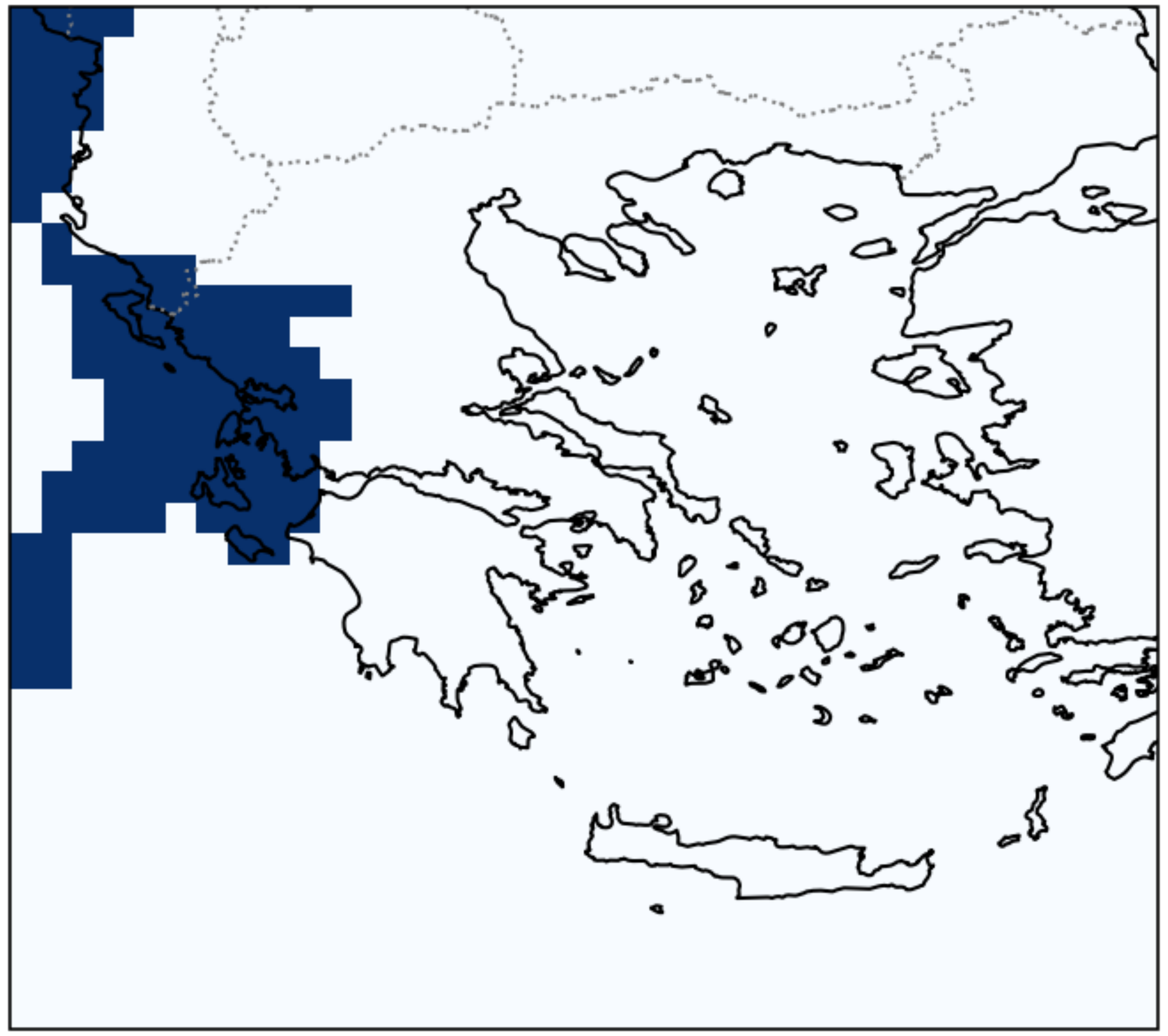
Related Work

- Foundation models for Earth Observation forecast: GraphCast, GenCast, Pangu-Weather
  Leverage GNNs, diffusion models, transformers
  Achieve **high accuracy and speed**.
- Specialized precipitation nowcasting
  Example: **SmaAt-UNet** (lightweight attention U-Net)
- Dominant drivers
  In United States for monthly extreme precipitation

Data



- ERA5: Atmospheric reanalysis data for climate
  - Input: Europe region (34°–72° N, 25° W–65° E)
  - Output: Greece (34°–42° N, 19°–28° E)



- Binary mask of extreme
  - Threshold 0.0015 m

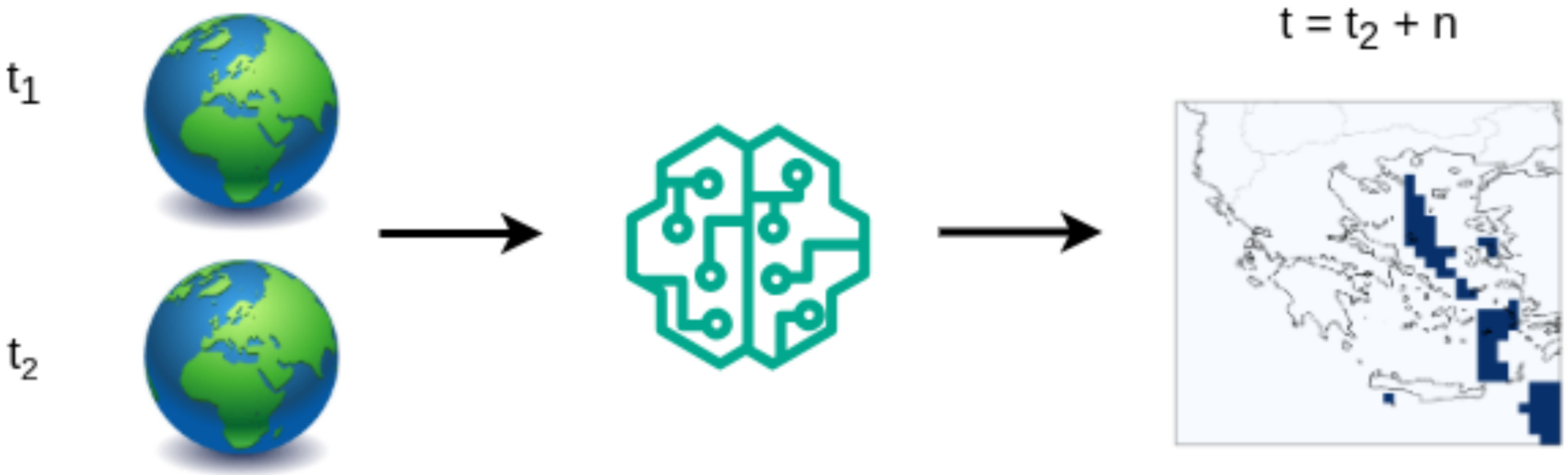
Intelligent Data-Intensive Systems

Total precipitation : total amount of water accumulated over a particular time period which depends on the data extracted

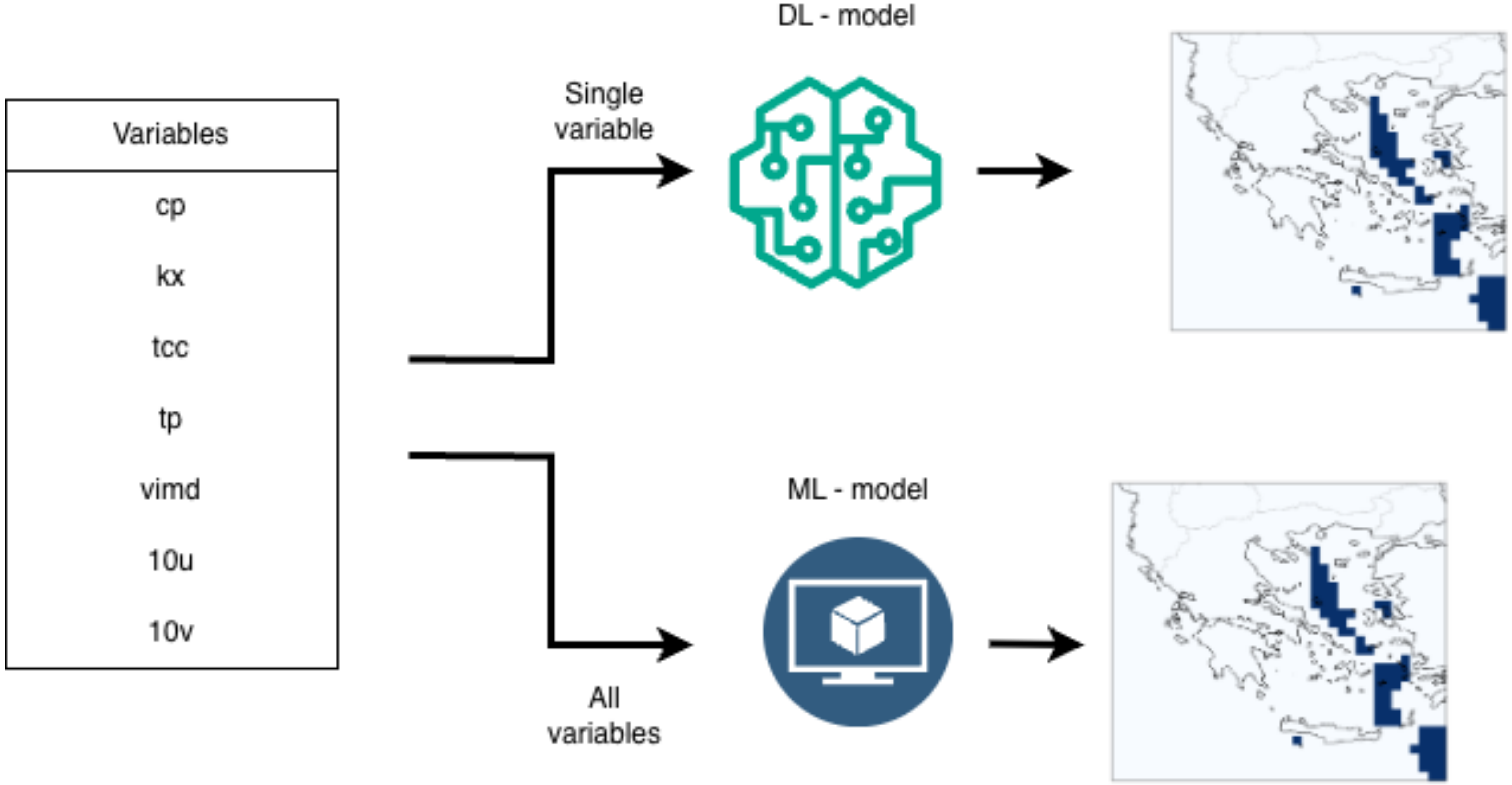
Data

7 Predictors (ERA5 single-level variables):

- Wind : 10 m u-wind (u10), 10 m v-wind (v10)
- Moisture: vertically integrated moisture divergence (vimd)
- Cloud coverage: total cloud cover (tcp)
- Stability measure (flag the posibility of a thunderstorm): K-index (kx)
- Precipitation: total precipitation (tp)



Methodology:



Models

Backbone:

- **SmaAt-UNet** (4) → compact encoder–decoder

Spatio-temporal modeling:

- **ConvLSTM bottleneck** → processes two consecutive time-step feature maps
- Captures short-term temporal dependencies

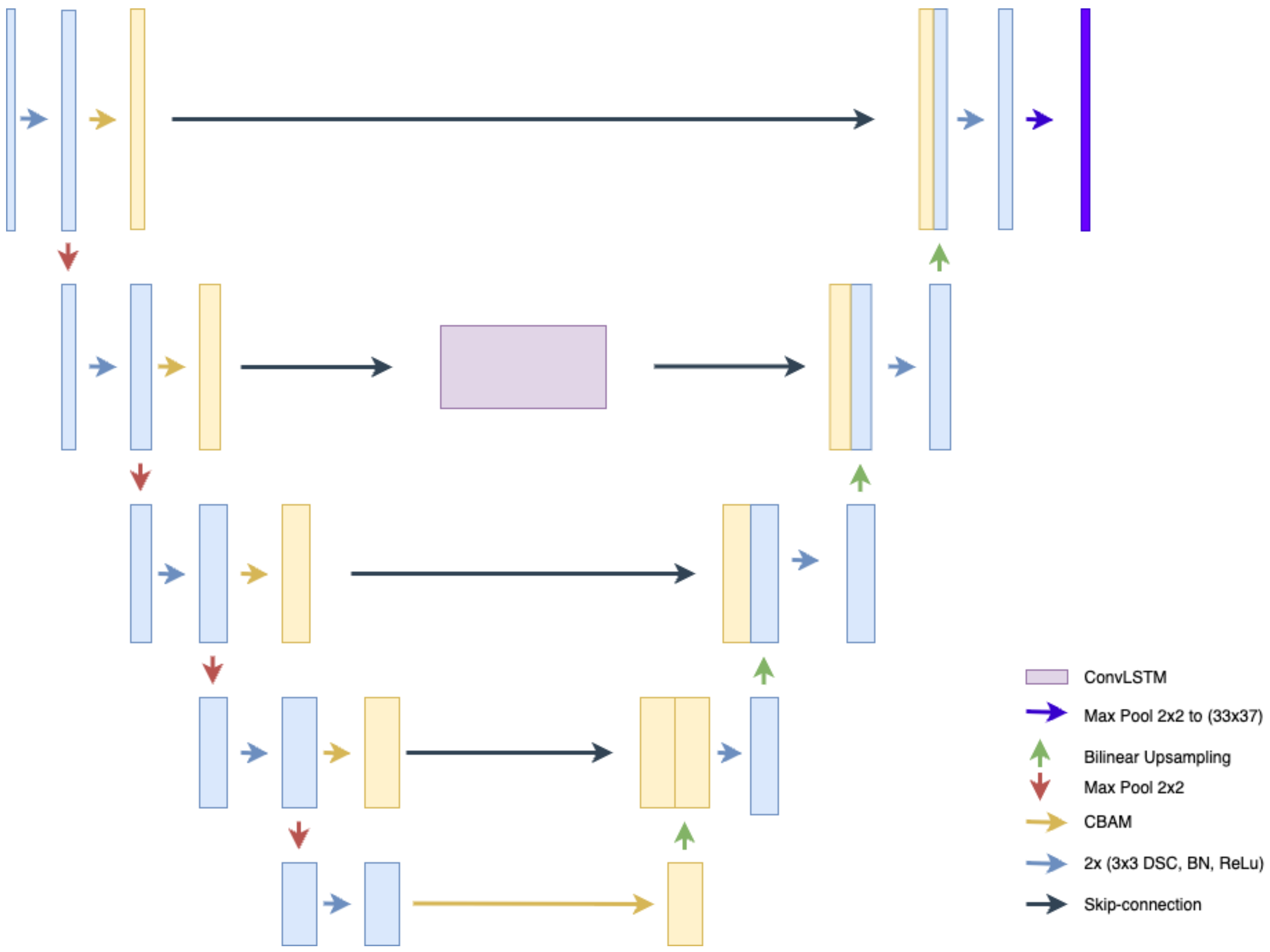
Output layer:

- Final **1×1 convolution**
- 2D adaptive average pooling → matches the 33 × 37 Greek grid

Loss function:

- **Focal loss** (6) applied to handle class imbalance (rare extreme events vs. many non-extremes)

$$\mathcal{L}_{FL}(p_t) = -\alpha(1-p_t)^{\gamma}\log p_t$$



Results: XGBoost Classifier

Top predictors across all horizons (2, 4, 6 days):

- **Total precipitation (tp)** → most important split feature
- **K-index (kx)** → consistently second

Key insight:

- Stable importance rankings across horizons → predictors carry robust signals.
- Confirms that tp and kx dominate, with vimd still relevant.

forecast/Variable	cp	kx	tcc	tp	vimd	u10	v10
2d	44.6	132	62.3	235.7	49.2	38.1	49.2
4d	58.5	124.8	78.1	191.9	32.9	42.3	58.7
6d	49.8	117.6	64.4	186.3	47.9	38.6	72.6

Results: Deep Learning Models

Performance trends

- Precision higher than recall → extremes remain hard to detect.
- All metrics degrade as forecast horizon increases (2 → 6 days).

Top predictors

- **Total precipitation (tp)** and **K-index (kx)** consistently rank among the top three.
- Wind components contribute less predictive power.

Predictor	2-days	4-days	6-days
tp	0.2027	0.1003	0.0943
kx	0.2228	0.2120	0.2120

(Values extracted from Tables 3–5) for precision

Conclusion

Which predictors are the most informative for extreme precipitation?

Evaluation of **seven ERA5 single-level predictors** for extreme precipitation in Greece.

Consistent findings across methods:

- Deep Learning → **tp, kx** strongest predictors.
- XGBoost → **tp, kx** top-ranked, with **vimd** also important.

Performance decreases with longer lead times (2 → 6 days), highlighting the challenge of rare-event forecasting.

Focusing on **interpretable and efficient predictors** can improve reliability of AI-driven extreme precipitation forecasts.

Future Work

Expand predictor set

- Add pressure-level variables (geopotential height, humidity, multi-level winds).
- Evaluate their added value for extreme precipitation classification.

Explore variable interactions

- Test a wider array of predictor combinations.
- Use fixed-architecture framework to identify the most informative subsets.

Benchmark against NWP models

- Compare data-driven extreme-event classifier with **numerical weather prediction (NWP)** outputs.
- Assess relative skill and complementarity.

Questions

Use the QR-code and ask your question

SCAN ME

DEMOKRITOS

Thank you very much for your attention!

ID-IS
Intelligent Data-Intensive Systems

References

(1) R. Lam, A. Sanchez-Gonzalez, M. Willson, P. Wirnsberger, M. Fortunato, F. Alet, S. Ravuri, T. Ewalds, Z. Eaton-Rosen, W. Hu, et al., Learning skillful medium-range global weather forecasting, Science 382 (2023) 1416–1421.

(2) I. Price, A. Sanchez-Gonzalez, F. Alet, T. R. Andersson, A. El-Kadi, D. Masters, T. Ewalds, J. Stott, S. Mohamed, P. Battaglia, R. Lam, M. Willson, Gencast: Diffusion-based ensemble forecasting for medium-range weather, arXiv preprint arXiv:2312.15796 (2023).

(3) K. Bi, L. Xie, H. Zhang, X. Chen, X. Gu, Q. Tian, Pangu-weather: A 3d high-resolution model for fast and accurate global weather forecast, 2022. URL: https://arxiv.org/abs/2211.02556. arXiv:2211.02556.

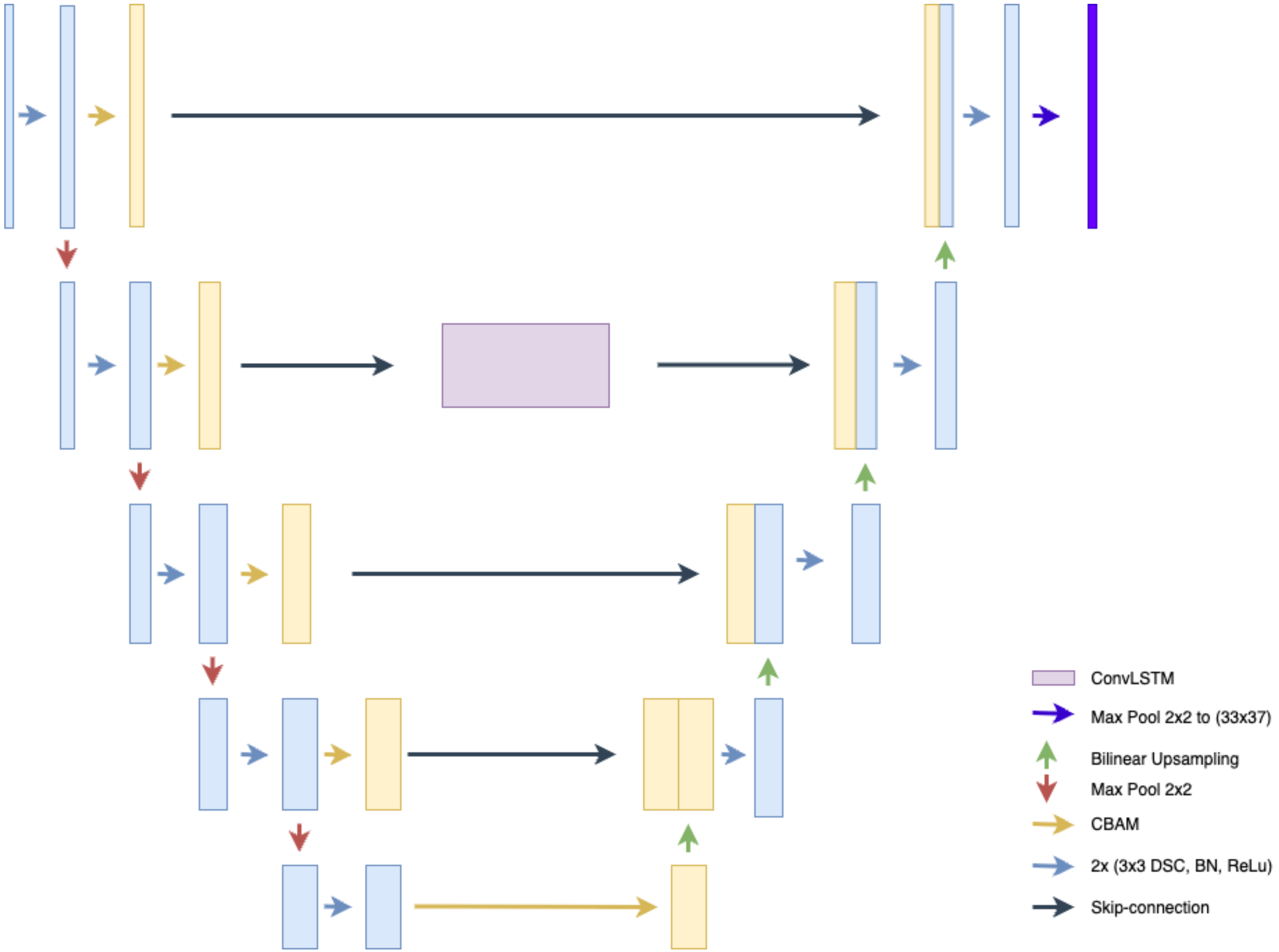
(4) K. Trebing, T. Stanczyk, S. Mehrkanoon, Smaat-unet: Precipitation nowcasting using a small attention-unet architecture, 2021. URL: https://arxiv.org/abs/2007.04417. arXiv:2007.04417.

(5) X. Lin, J. Fan, Z. J. Hou, J. Wang, Machine learning of key variables impacting extreme precipitation in various regions of the contiguous united states, Journal of Advances in Modeling Earth Systems 15 (2023) e2022MS003334.

(6) T.-Y. Lin, P. Goyal, R. Girshick, K. He, P. Dollár, Focal loss for dense object detection, 2018. URL: https://arxiv.org/abs/1708.02002. arXiv:1708.02002.

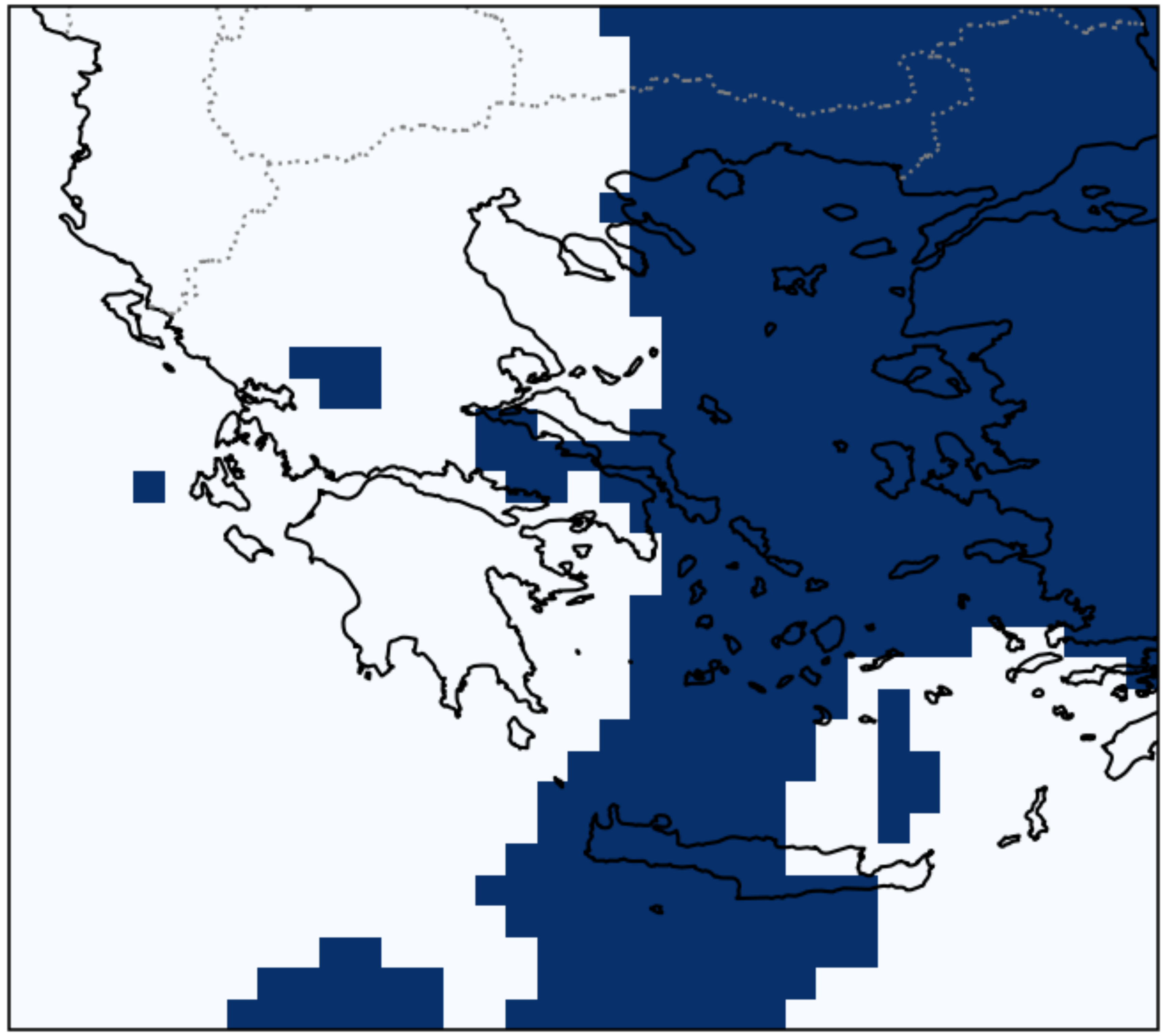
DEMOKRITOS

ID-IS
Intelligent Data-Intensive Systems


ConvLSTM
Max Pool 2x2 to (33x37)
Bilinear Upsampling
Max Pool 2x2
CBAM
2x (3x3 DSC, BN, ReLu)
Skip-connection

Motivation

- In recent years, machine-learning (ML) and deep-learning (DL) approaches have exponentially improved.
- Deep learning is becoming a powerful tool in climate science, supporting advances for example in the identification of atmospheric circulation patterns, weather forecasting, and extreme event classification.
- In some cases achieving performance comparable to traditional numerical weather prediction (NWP) in skill
- However, the complexity of atmospheric processes, particularly those that drive rare and high-impact precipitation events, still creates major challenges in data driven AI methods.



DEMOKRITOS

ID-IS
Intelligent Data-Intensive Systems

Problem Statement

- Extreme precipitation events are rare, high-impact, and difficult to forecast with deep learning and data driven methods.

- Most ML/DL studies for data driven forecasting rely on either **very large sets of variables** or a few **commonly used predictors**, without systematic testing.

- It remains unclear which atmospheric predictors from reanalysis data carry the strongest signal for extreme precipitation events in data driven AI methods.

- A **systematic evaluation** of predictors' skill is needed to improve the understanding of data-driven forecasting of extremes.

Related Work

DEMOKRITOS

Predicting precipitation has become a core task within the new generation of forecasting models, in some cases, as part of forecasting the state of the atmosphere and in some cases as the only target.

Data driven weather forecast models (GraphCast (1), GenCast (2), Pangu-Weather (3))

- Leverage GNNs, diffusion models, transformers for **global medium-range forecasts**.
- Achieve **high accuracy and speed**, trained on decades of ERA5 data.

Specialized precipitation nowcasting

- Example: **SmaAt-UNet** (lightweight attention U-Net) (4)
- Delivers comparable skill to larger networks with fewer parameters.

ID-IS
Intelligent Data-Intensive Systems

Related Work

An ensemble of machine-learning methods has also been applied to identify the dominant drivers of extreme- events intensity and frequency on a regional scale

Feature importance studies

- Ensembles (Random Forest, XGBoost, ANN) used to find **dominant drivers** of extremes (e.g., latent heat flux, humidity, soil moisture) (5).

DEMOKRITOS

ID-IS

Intelligent Data-Intensive Systems