BlockGPT: Spatio-Temporal Modelling of Rainfall via Frame-Level Autoregression

Cristian Meo[1,2*], Varun Sarathchandran[2,*], Avijit Majhi[2,4,*], Shao-Hsuan Hung[2], Carlo Saccardi[2], Ruben Imhoff[3], Roberto Deidda[4], Remko Uijlenhoet[2], Justin Dauwels[2]

*Equal Contribution; [1]LatentWorlds AI, NL; [2]Delft University of Technology, NL; [3]Deltares, NL; [4]University of Cagliari, IT.

Background

1. Climate-related disasters are becoming more frequent and costly [1]
2. 90% of major disasters in the last 20 years have been weather-related [2].
3. Short-term, high-resolution, accurate weather prediction has thus become essential [3]

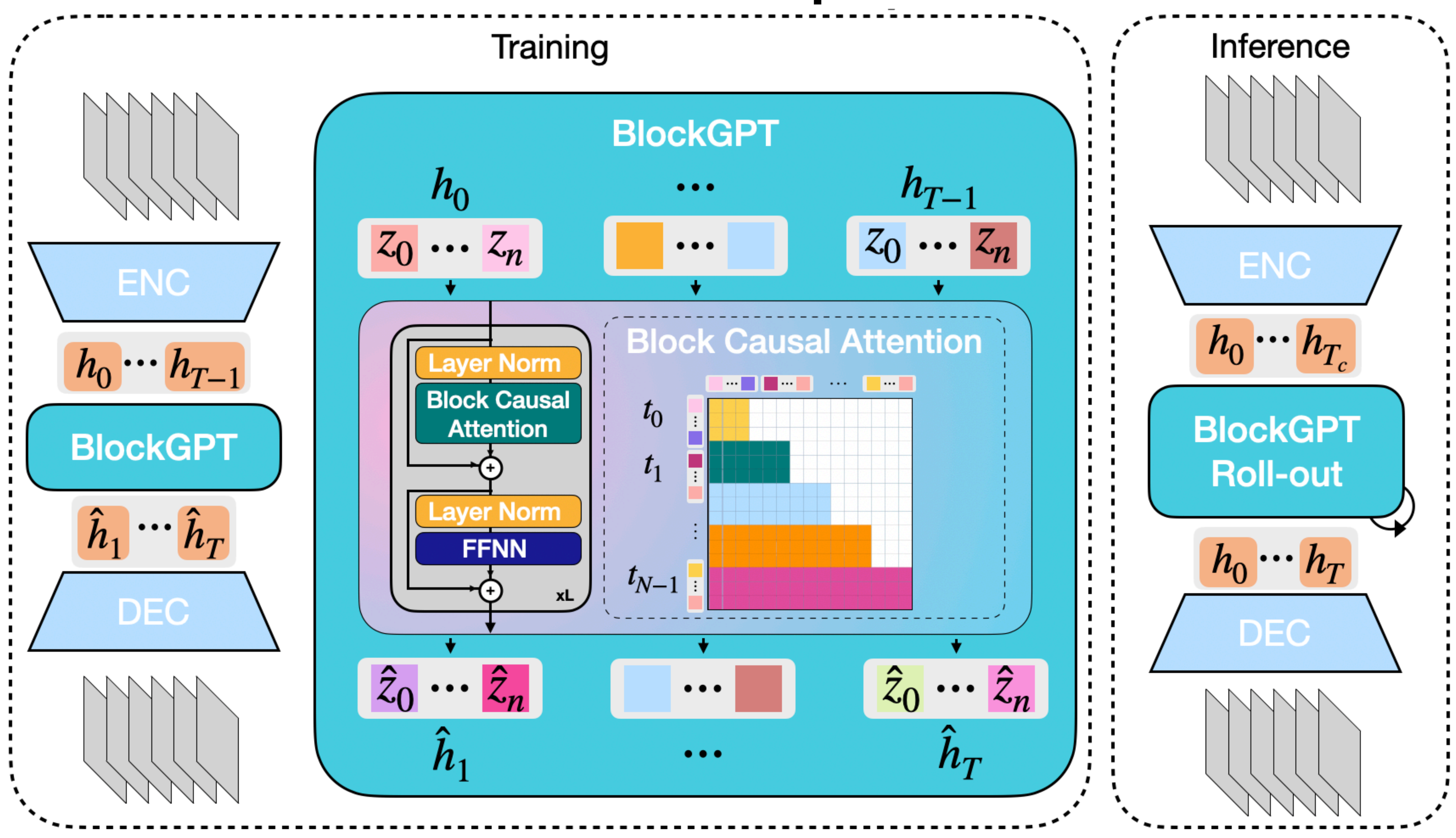
Motivations

1. ML-based nowcasting has relied on transformer-based [4,5] and diffusion-based techniques [6].
2. Diffusion pipelines are computationally heavy, and nowcastingGPT (SOTA on KNMI) imposes sub-optimal inductive biases and is also slow.

Contributions

1. We propose BlockGPT, which, unlike NowcastingGPT, treats precipitation fields as two-dimensional entities.
2. We use block-attention to impose more optimal inductive biases. This also leads to faster training and inference times (up to 31x faster inference).
3. Compared to SOTA (NowcastingGPT and Diffcast), we achieve better accuracy and event localization on both KNMI and SEVIR data.

BlockGPT Pipeline



Methodology

Given $\mathcal{X}_{\text{context}} = \{\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_{T_c}\}$, $\mathbf{X}_t \in \mathbb{R}^{H \times W}$, predict $\mathcal{X}_{\text{target}} = \{\mathbf{X}_{T_c+1}, \ldots, \mathbf{X}_T\}$.

Learn $f_{compress} : \mathcal{X} \mapsto \mathcal{T}$ **and** $f_{predict} : \mathcal{T}_{context} \mapsto \mathcal{T}_{target}$

NowcastingGPT flattens 2D fields,

$\mathbf{z} = \text{flatten}(\mathcal{T})$, and models $p(\mathbf{z})$ as

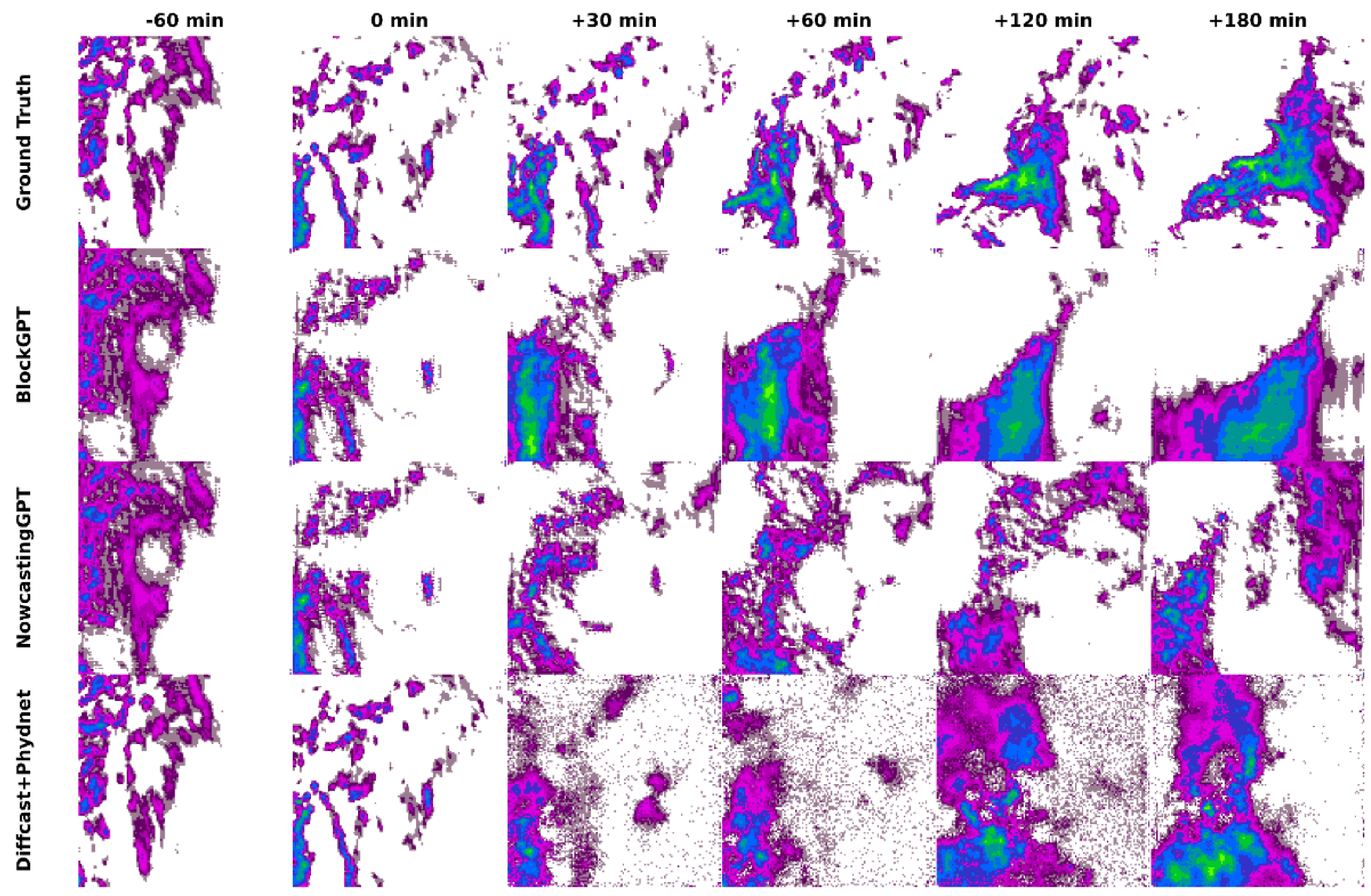
$$p(\mathbf{z}) = \prod_{i=1}^{T \cdot H'W'} p\left(z^{(i)} \mid z^{(1)}, z^{(2)}, \ldots, z^{(i-1)}\right).$$

This treats an video like a 1D structure.

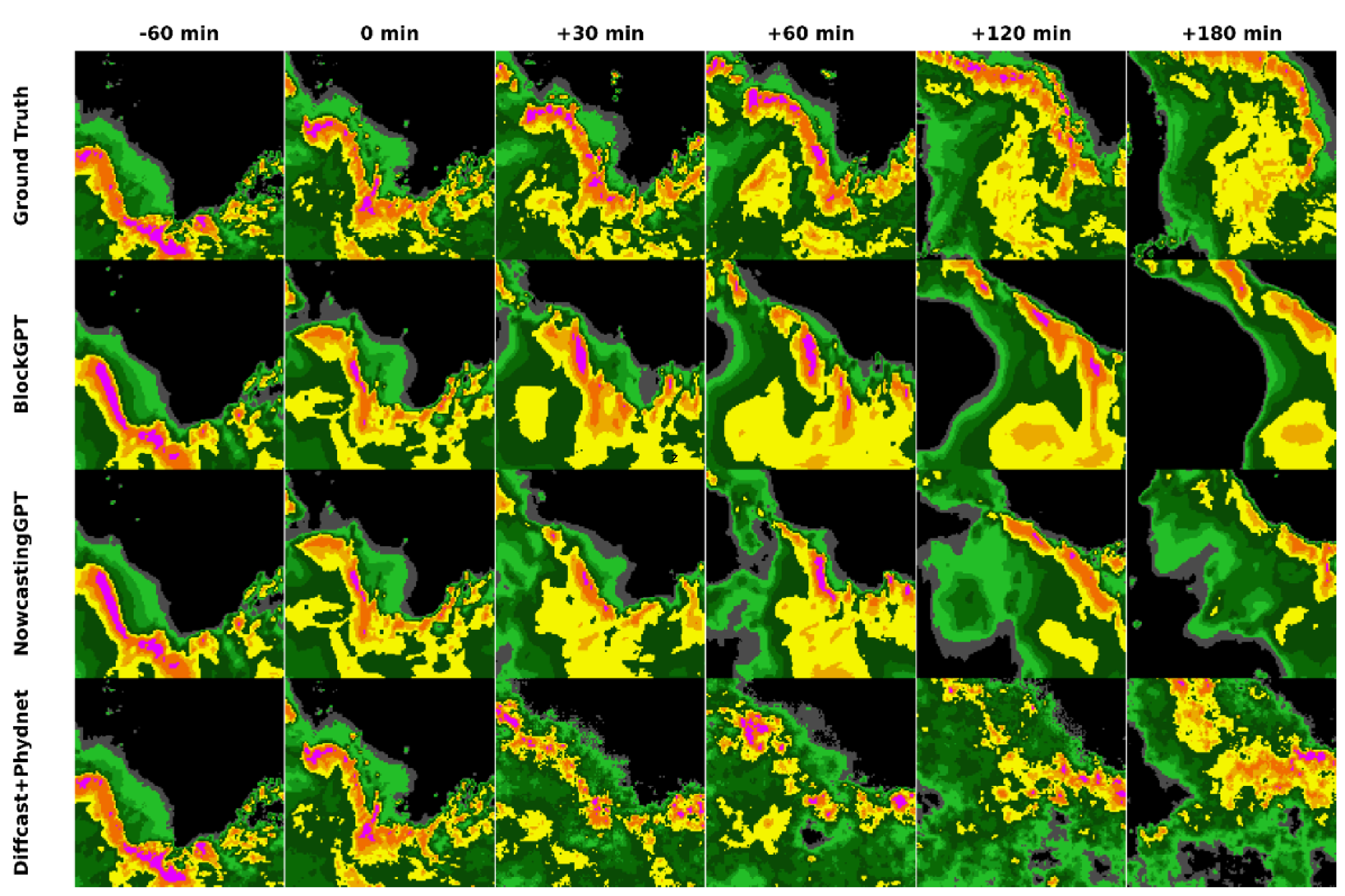
BlockGPT instead directly models $\mathcal{T} = \{\mathbf{T}_1, \mathbf{T}_2, \ldots, \mathbf{T}_T\}$ **as**

$$p(\mathcal{T}) = \prod_{t=1}^{T} p\left(\mathbf{T}_t \mid \mathbf{T}_1, \ldots, \mathbf{T}_{t-1}\right)$$

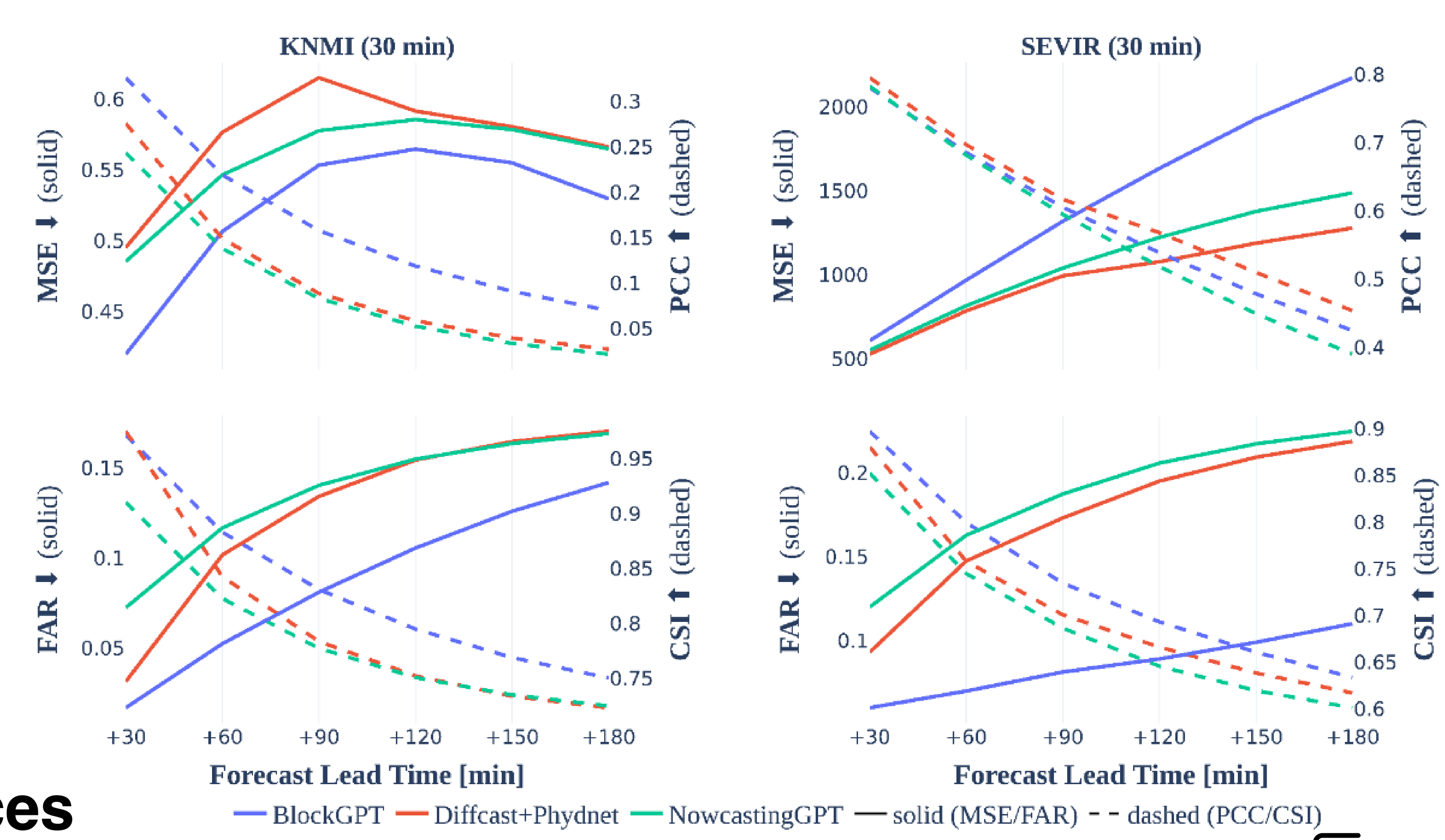
Qualitative Results (KNMI)



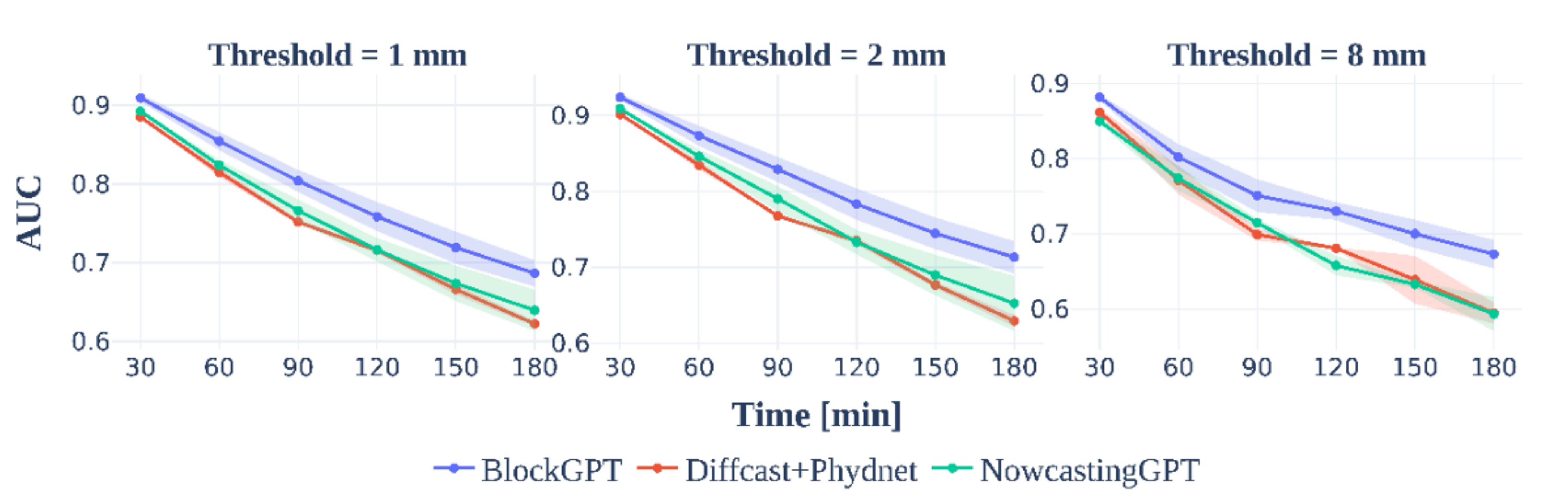
Qualitative Results (SEVIR)



Quantitative Results (Both)



Catchment Study (KNMI)



References

[1] Economic losses from climate-related extremes in Europe, EEA, 2024.
[2] The human cost of disasters: 2000–2019, CRED & UNDRR, 2020.
[3] Seamless prediction of the Earth system: From minutes to months, J. Côté et al., 2015.
[4] Earthformer: Space–time transformers for Earth system forecasting, Z. Gao et al., 2023.
[5] Extreme precipitation nowcasting using transformer-based generative models, C. Meo et al., 2024.
[6] DiffCast: A unified diffusion framework for precipitation nowcasting, D. Yu et al., 2024.