CUTTING-EDGE GENERATIVE AI FOR INTRA-HOUR SOLAR FORECASTING

Yann Fabel, Dominik Schnaus, Bijan Nouri, Stefan Wilbert, Niklas Blum, Luis F. Zarzalejo, Julia Kowalski, Robert Pitz-Paal

EUPVSEC 2025

22th of September 2025, Bilbao, Spain

Agenda

- Introduction
- Generative Forecasting Approach
- Ramp Event Prediction
- Conclusion & Outlook

MOTIVATION

Motivation

Why forecasting solar irradiance?

- Predict expected energy yield of PV power plants
- Anticipate local short-term fluctuations caused by cloud passages (ramp events)

Challenges by ramp events

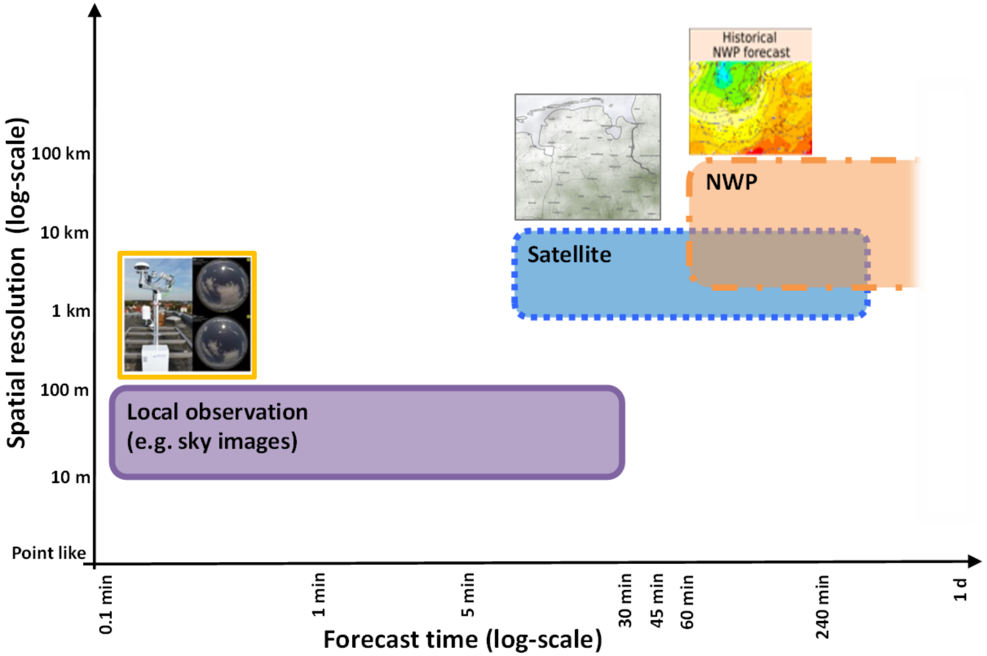
- Local power output variability
- Potential risk of grid instabilities at high solar penetration

Benefits of intra-hour forecasting

- Better situational awareness for plant and grid operators
- Reduced storage requirements
- Improved market trading strategies
- More efficient operation of CST plants

Requirements

- High-resolution cloud information in space and time
  → All-Sky-Imagers

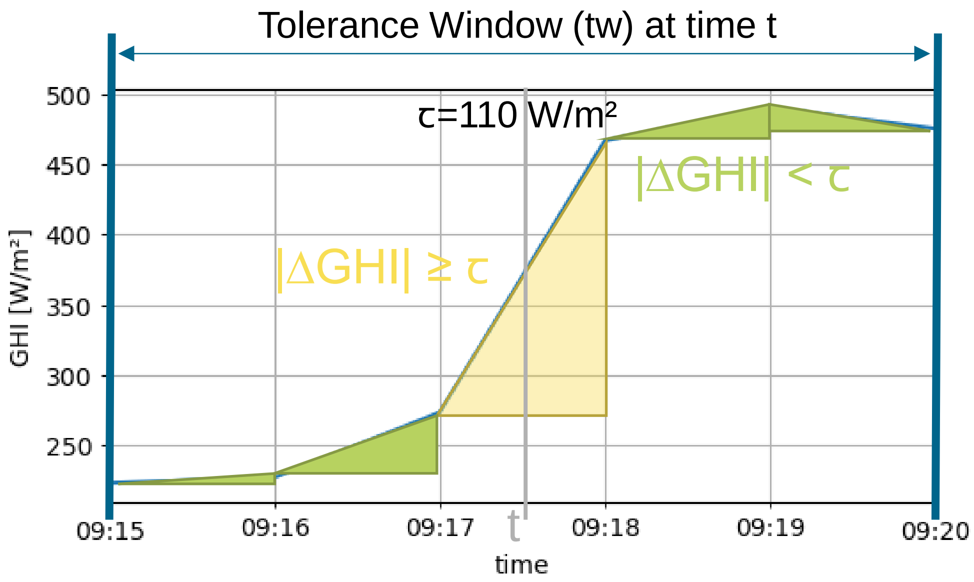


Motivation
Ramp Events

- Typically no evaluation of predicting the variability of solar irradiance
- Common forecasting metrics (e.g., RMSE, MAE, MBE) represent an average error of the target quantity (e.g., GHI)
  - Good measure to assess expected energy yield (integration of irradiance over time)
  - No information on variability within the forecast
- Definition Ramp Event [1]:

$$\frac{|\Delta GHI|}{\Delta t} > \tau \Longrightarrow Ramp$$

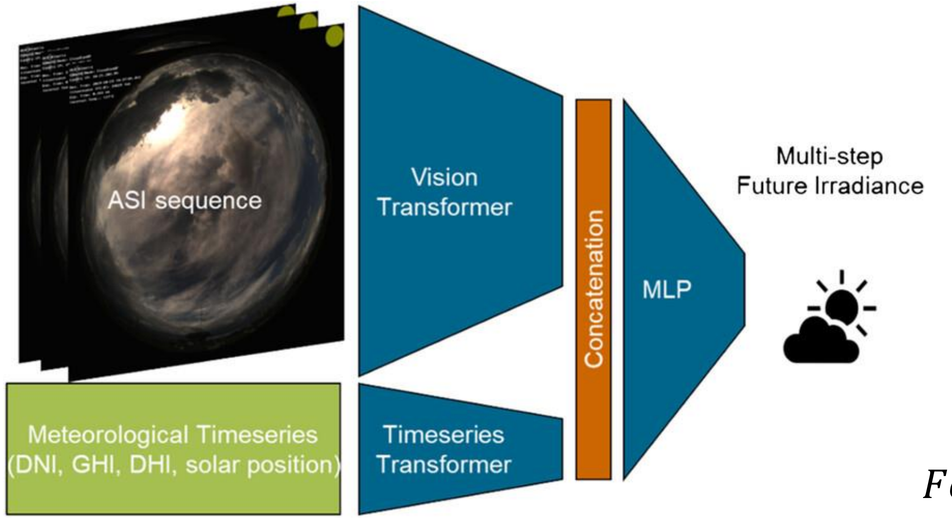
$$t:\ if\ \exists\ Ramp\ in\ [t - tw/2, t + tw/2] \Longrightarrow Ramp\ Event\ at\ t$$



Limitations of State-of-the-Art Models

- State-of-the-Art direct data-driven models are often optimized on RMSE [2, 3, 4]

Fabel et al. [4]



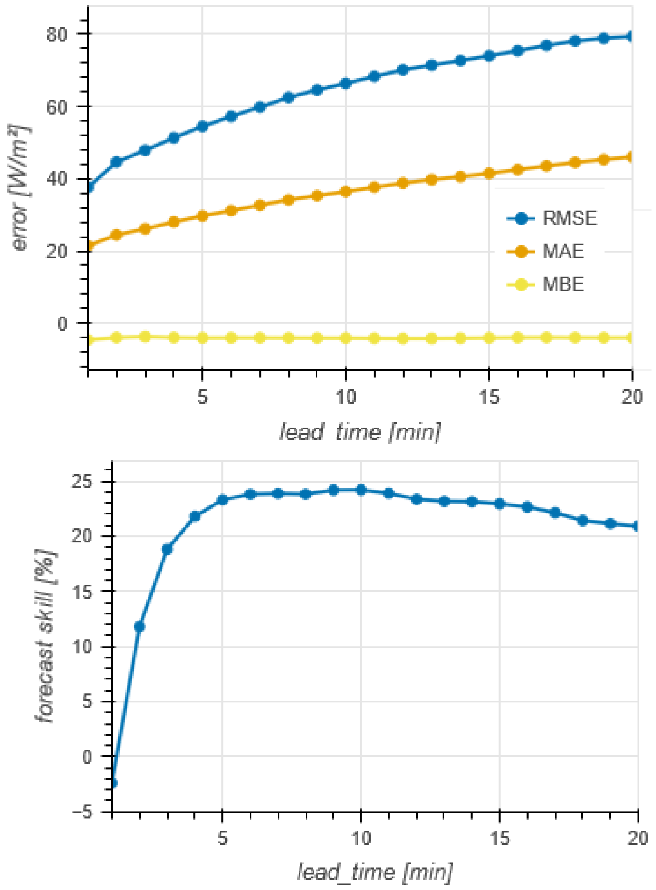
$$RMSE = \sqrt{1/n \sum_{i=1}^{n} (\hat{y}_i - y_i)^2}$$

$$MAE = 1/n \sum_{i=1}^{n} |\hat{y}_i - y_i|$$

$$MBE = 1/n \sum_{i=1}^{n} \hat{y}_i - y_i$$

$$Forecast\ Skill = 1 - \frac{RMSE_{model}}{RMSE_{persistence}}$$

GHI Forecast



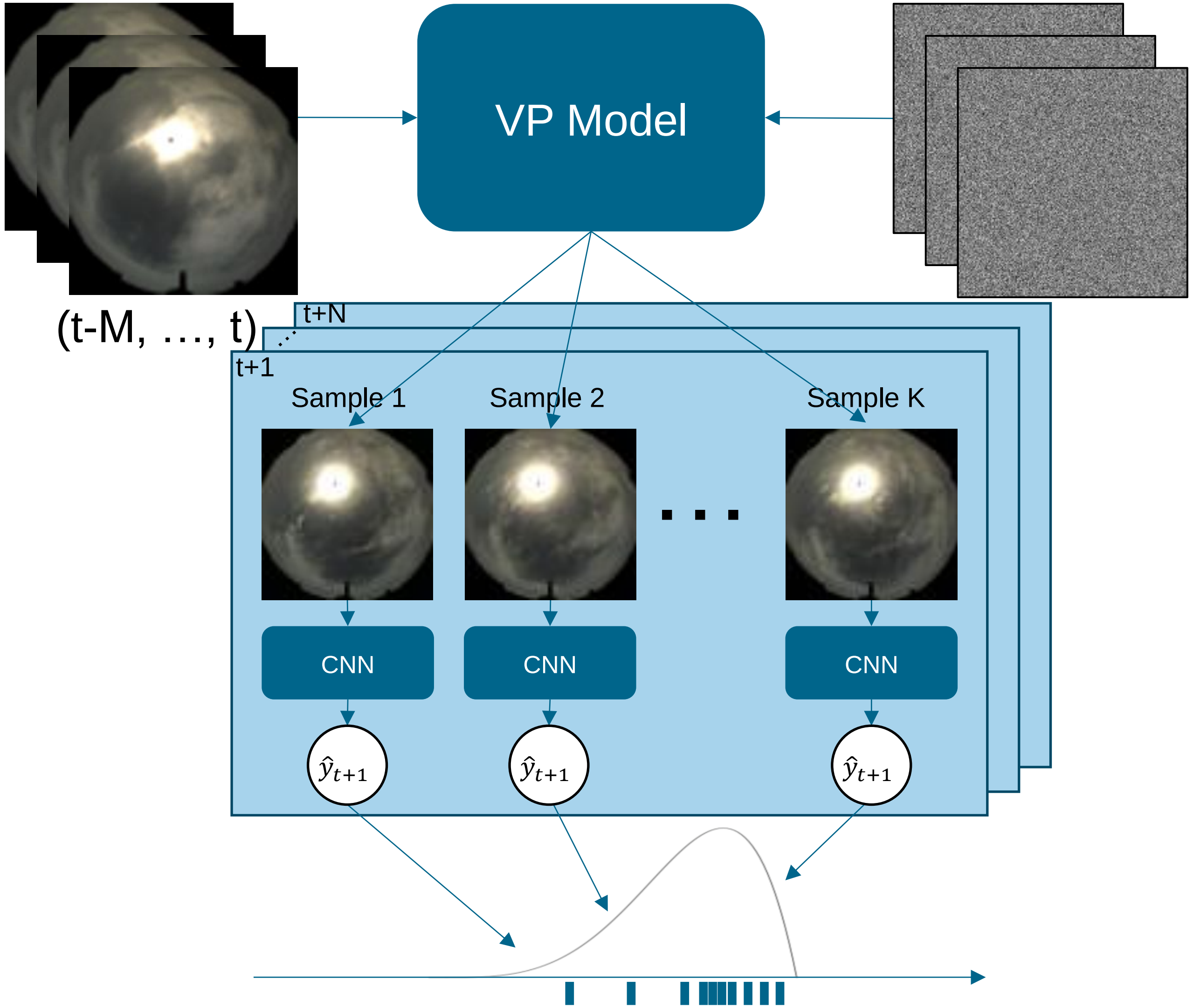
- Strong performance on standard error metrics, but ramp events remain undetected due to overly smoothed forecast curves (see later slides)
- How can we circumvent smoothing of the forecast curve?

GENERATIVE FORECASTING APPROACH

Generative Forecasting
Model Architecture

- **Video Prediction (VP)**:
  - Given a sequence of M past images the next N images are predicted
  - Multiple future scenarios can be generated from the same input sequence by sampling from Gaussian noise
    → Measure for uncertainty

- **Regressor/Classifier:**
  - Given individual predicted future frames a second model (e.g. CNN) is used to derive the desired target quantity
  - Trained separately on real images
  - E.g., a model predicts ramp events or GHI corresponding to the predicted sky image



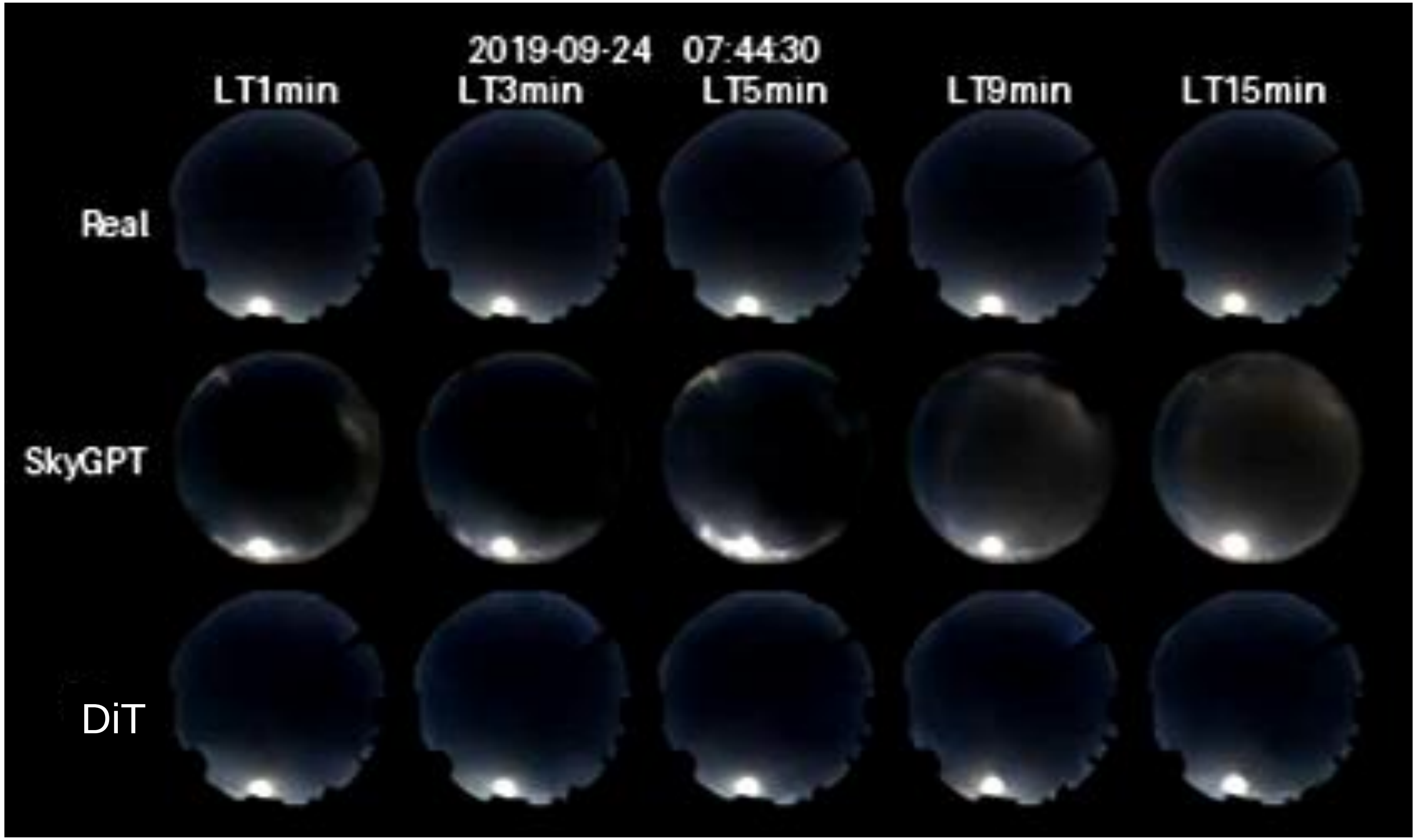
Video Prediction Models
Train and Test setup

- Two different generative models were tested
  - SkyGPT [5]: Adaptation of the VideoGPT [6] model combined with PhyCells [7]
  - Ours: Adaptation of the DiT model [8] (diffusion-based transformer)
- Training
  - Both models were trained on selected camera data from CIEMAT‘s PSA (Almería, Spain)
  - SkyGPT model trained with the same hyperparameter configuration as in the original publication
- Testing
  - Evaluation on separate benchmark dataset defined in All-Sky Imager-based forecasting study [9]
  - 28 selected days from a single camera at PSA representing diverse sky conditions
  - 4 image samples per model were generated for each lead time

Image taken at CIEMAT‘s PSA

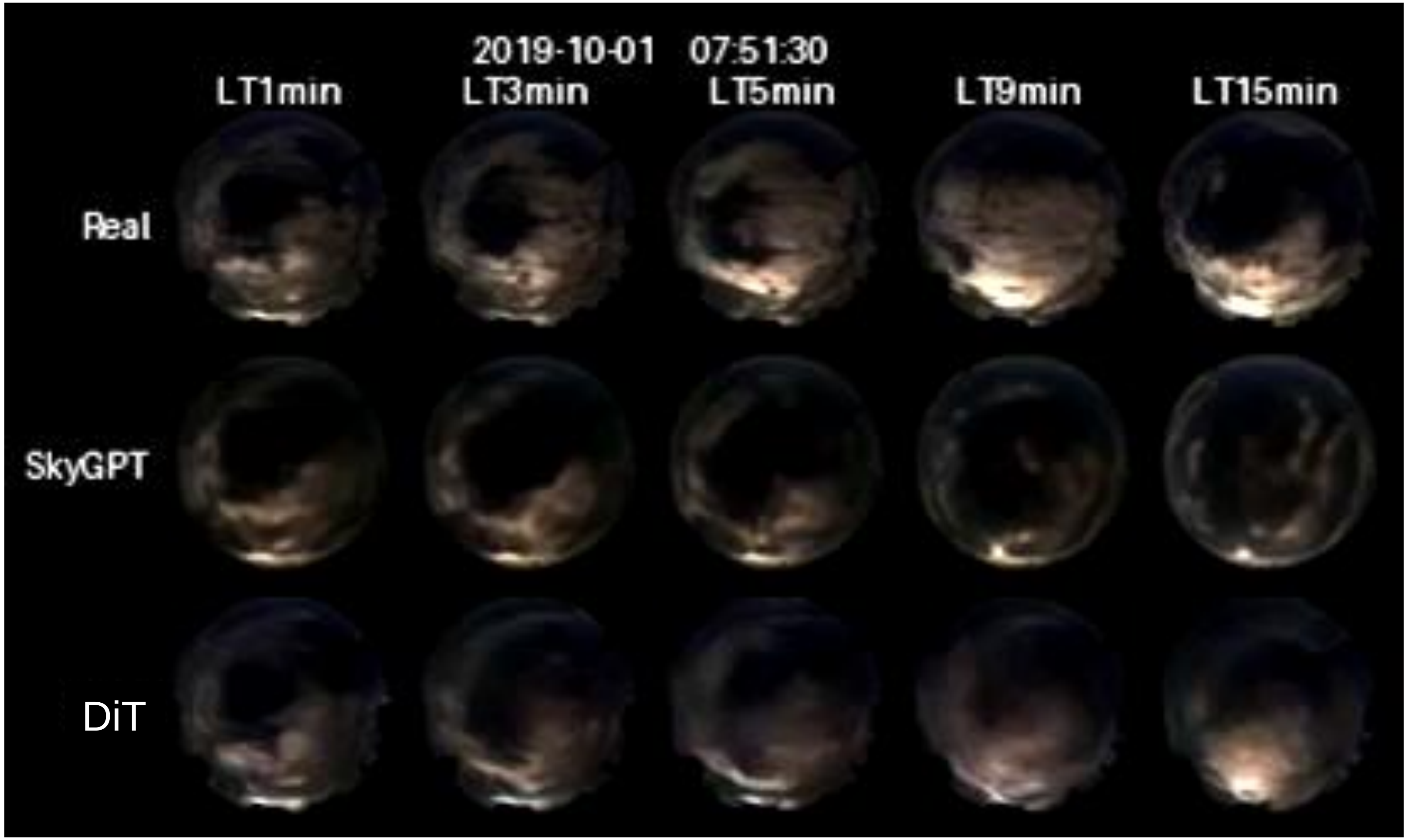
	SkyGPT	DiT
Image res	64x64	128x128
Temporal res.	2min	1min
Forecast horizon	15min	30min
Auto-Encoder	VQ-VAE [6]	Pretrained VAE [9]

Video Prediction Evaluation
Examplary Results – Single Sample, Selected Lead Times



→ A lot of „halluzinations“ even for clear sky conditions

Video Prediction Evaluation
Examplary Results – Single Sample, Selected Lead Times

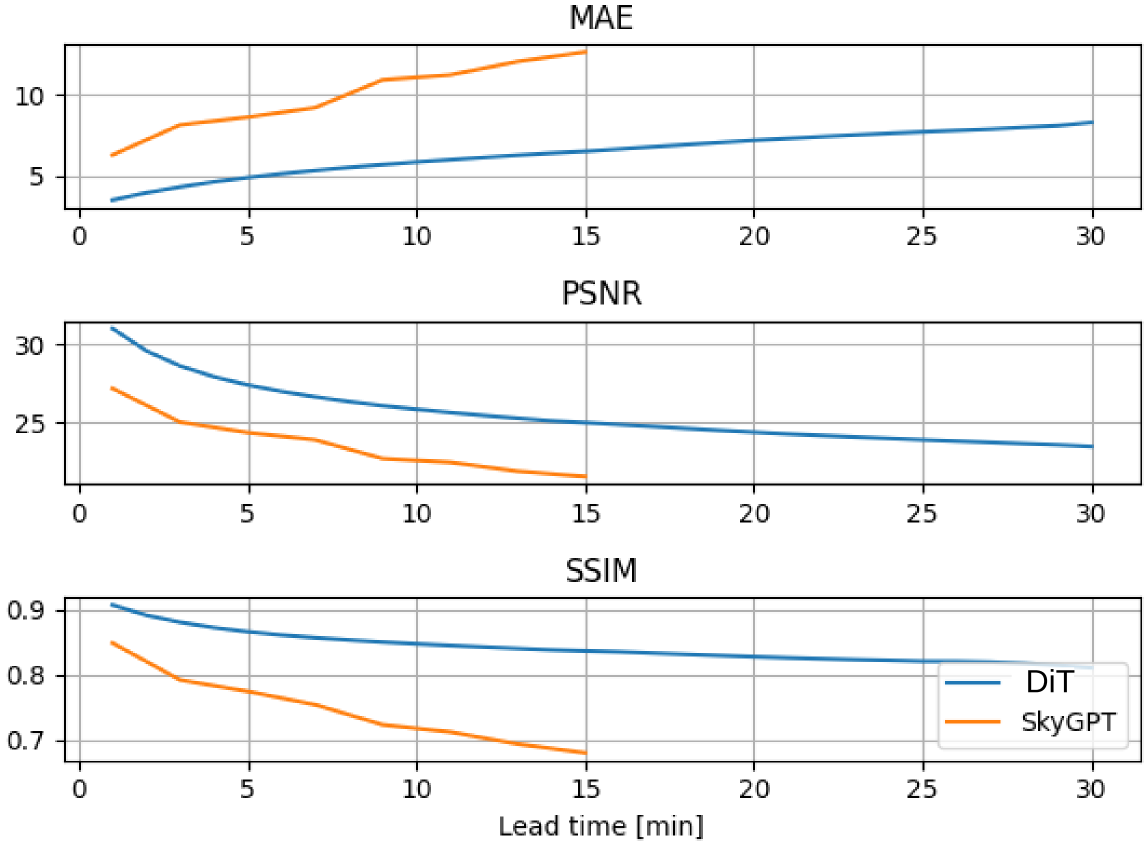


→ Strong deviations in terms of cloud coverage for larger lead times

Video Prediction Evaluation
Quantitative Results

- Evaluation of predicted sky image frames
  - Image data range: [0, 255]
  - Lead-time specific calculation averaged over all generated future scenarios

- Image-wise pixel metrics
  - Mean Absolute Error (MAE)↓:
    - Average error per pixel
  - Peak Signal-to-Noise Ratio (PSNR)↑:
    - Ratio of maximum possible signal to error in decibels (measure of fidelity)
  - Structural Similarity Index (SSIM)↑:
    - Measure for perceptual similarity (capturing sharpness and structure)



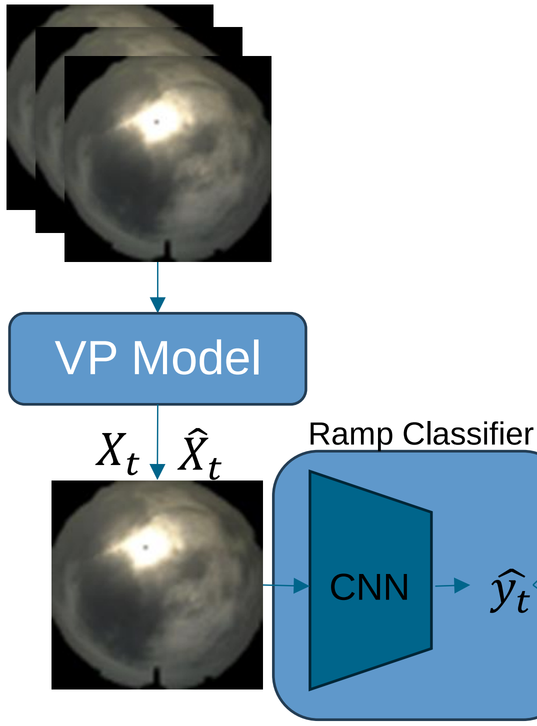
→ Better performance of DiT model in terms of image quality

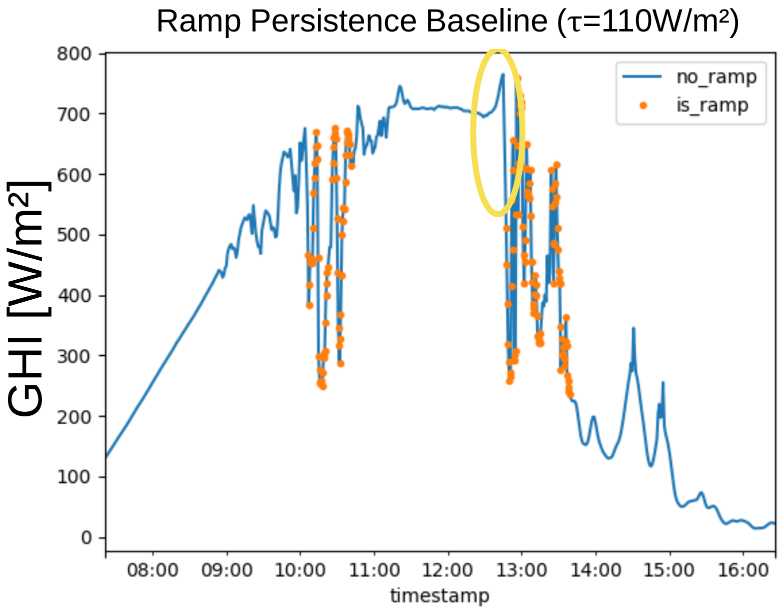
RAMP EVENT PREDICTION

Ramp Event Prediction

- <u>Ramp Classifier</u>
  - CNN predicts likelihood of ramp event (RE) based on corresponding image
  - Ground truth $y_t = \begin{cases} 1 & if\ \exists\ ramp\ in\ [t-3, t+3] \\ 0 & otherwise \end{cases}$
  - **Trained on real sky images**
  - **Evaluated on synthetic images** from generative model to obtain ramp event prediction

- <u>Baseline model</u>: Ramp Persistence
  - If a ramp was observed in the measured irradiance curve in the last T=30min a ramp is expected in the next T minutes too
  - Independent of sky images





Ramp Event Prediction Evaluation

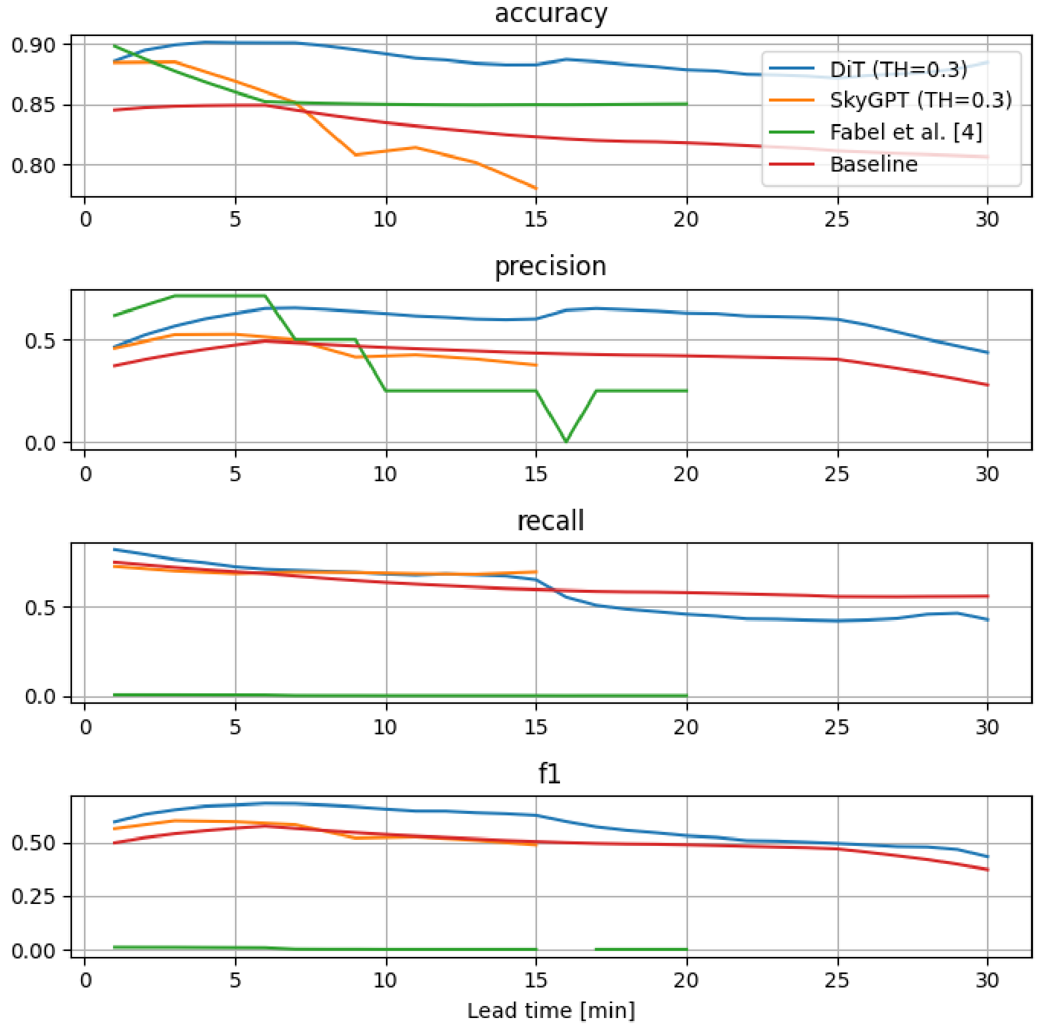
- Video prediction models generate K=4 samples (images) for all lead times for each forecast
- Predicted RE from average probability of ramp classifier over all samples
- Observed RE by measured ramp within tolerance window (tw=10min)
- Low classifier threshold (TH=0.3) chosen to prioritize recall
- Evaluation of classification metrics over lead times

$$accuracy = \frac{TP + TN}{TP + FN + FP + TN}$$

$$precision = \frac{TP}{TP + FP}$$

$$\text{f1} = 2 \times \frac{precision \times recall}{precision + recall}$$

$$recall = \frac{TP}{TP + FN}$$



Conclusion

- Summary:
  - Low RMSE does not guarantee realistic representation of irradiance variability (e.g., ramp events)
  - Generative, image-based modeling of cloud dynamics offers a promising alternative to capture short-term variability
  - Current video prediction models for ASI still struggle at longer horizons (inconsistencies and physically unrealistic cloud scenes)
  - But generated images remain useful for detecting ramp events
- Outlook
  - Enhance ASI-based video prediction
    - Focus training on highly variable cloud conditions
    - Leverage advances in generative video modeling (e.g., noise warping, motion conditioning)
  - Combine multiple perspectives to learn better cloud representations

References

1. Nouri et al. 2024, **Ramp Rate Metric Suitable for Solar Forecasting**, DOI: 10.1002/solr.202400468
2. Sun et al., **Solar PV output prediction from video streams using convolutional neural networks**, DOI: 10.1039/c7ee03420b
3. Paletta et al. 2021, **Benchmarking of deep learning irradiance forecasting models from sky images - An in-depth analysis**, DOI: 10.1016/j.solener.2021.05.056
4. Fabel et al. 2023, **Combining deep learning and physical models: a benchmark study on all-sky imagerbased solar nowcasting systems**
5. Nie et al. 2024, **SkyGPT: Probabilistic ultra-short-term solar forecasting using synthetic sky images from physics-constrained VideoGPT,** DOI: 10.1016/j.adapen.2024.100172
6. Yan et al. 2021, **VideoGPT: Video Generation using VQ-VAE and Transformers**, DOI: 10.48550/ARXIV.2104.10157
7. LeGuen et al. 2020, **Disentangling Physical Dynamics From Unknown Factors for Unsupervised Video Prediction**, DOI: 10.1109/cvpr42600.2020.01149
8. Pebbles et al. 2022, **Scalable Diffusion Models with Transformers**, DOI: 10.48550/ARXIV.2212.09748
9. Blattmann et. al 2023, **Stable video diffusion: Scaling latent video diffusion models to large datasets**, DOI: 10.48550/ARXIV.2311.15127

AuSeSol AI (Grant number 67KI21007A)

Gefördert durch:

Bundesministerium
für Umwelt, Klimaschutz, Naturschutz
und nukleare Sicherheit

aufgrund eines Beschlusses
des Deutschen Bundestages

THANK YOU FOR YOUR ATTENTION!
QUESTIONS? YANN.FABEL@DLR.DE