GENERATIVE AI FOR INTRA-HOUR DNI FORECASTS

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

SolarPACES 2025

September 2025, Almería, Spain

Agenda

- Introduction
- Generative Forecasting Approach with All Sky Images
- Ramp Event Prediction
- Conclusion & Outlook

MOTIVATION

Motivation

Why forecasting solar irradiance?

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

Challenges by ramp events

- Local heat/power output variability
- Potential risk of supply instabilities

Benefits of intra-hour forecasting

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

Requirements

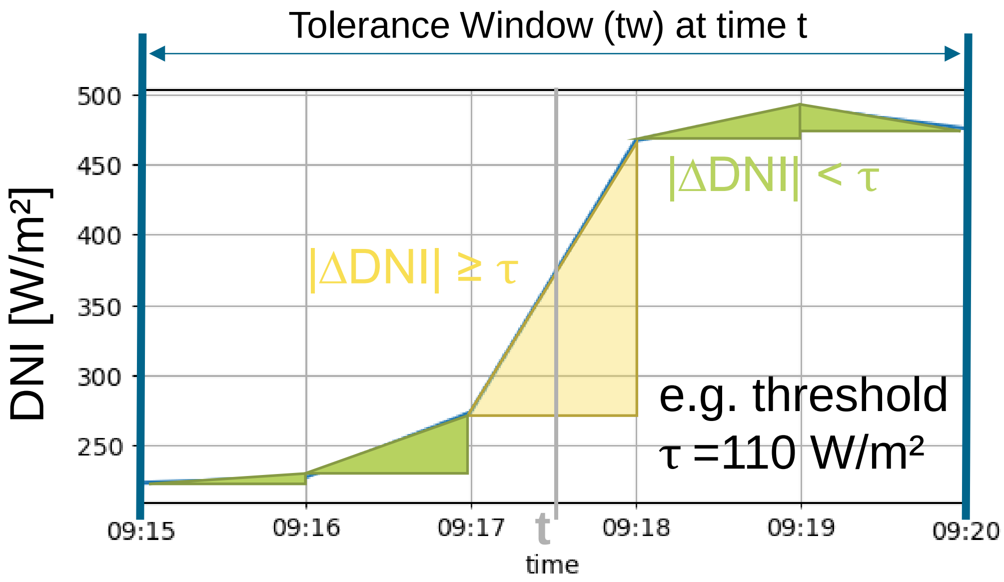
- High-resolution cloud information in space and time
  → All-Sky-Imagers

Motivation
Ramp Events

- Common forecasting metrics (e.g., RMSE, MAE, MBE) of the target quantity (e.g, DNI)
  - are a good measure to assess expected energy yield (integration of irradiance over time)
  - provide no information on variability within the forecast
- Definition ramp event and ramp metrics [1]:

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

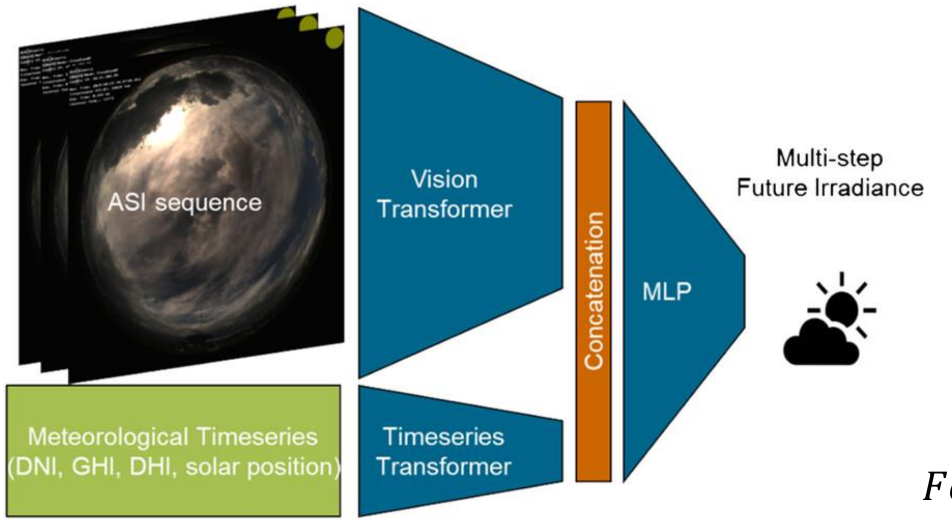
$$t:\ if\ \exists\ Ramp\ in\ [t - tw/2, t + tw/2] \Longrightarrow Ramp\ Event\ at\ t$$



Limitations of State-of-the-Art Data-Driven Models

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

Fabel et al. [4]

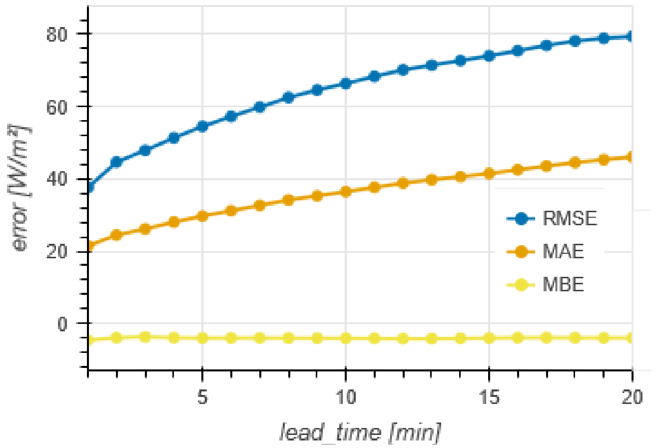


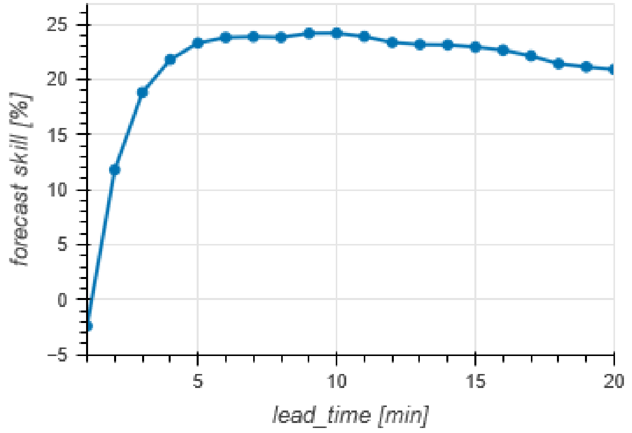
$$RMSE = \sqrt{1/n \sum_{i=1}^{n} (\hat{y}_i - y_i)^2}$$

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

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

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





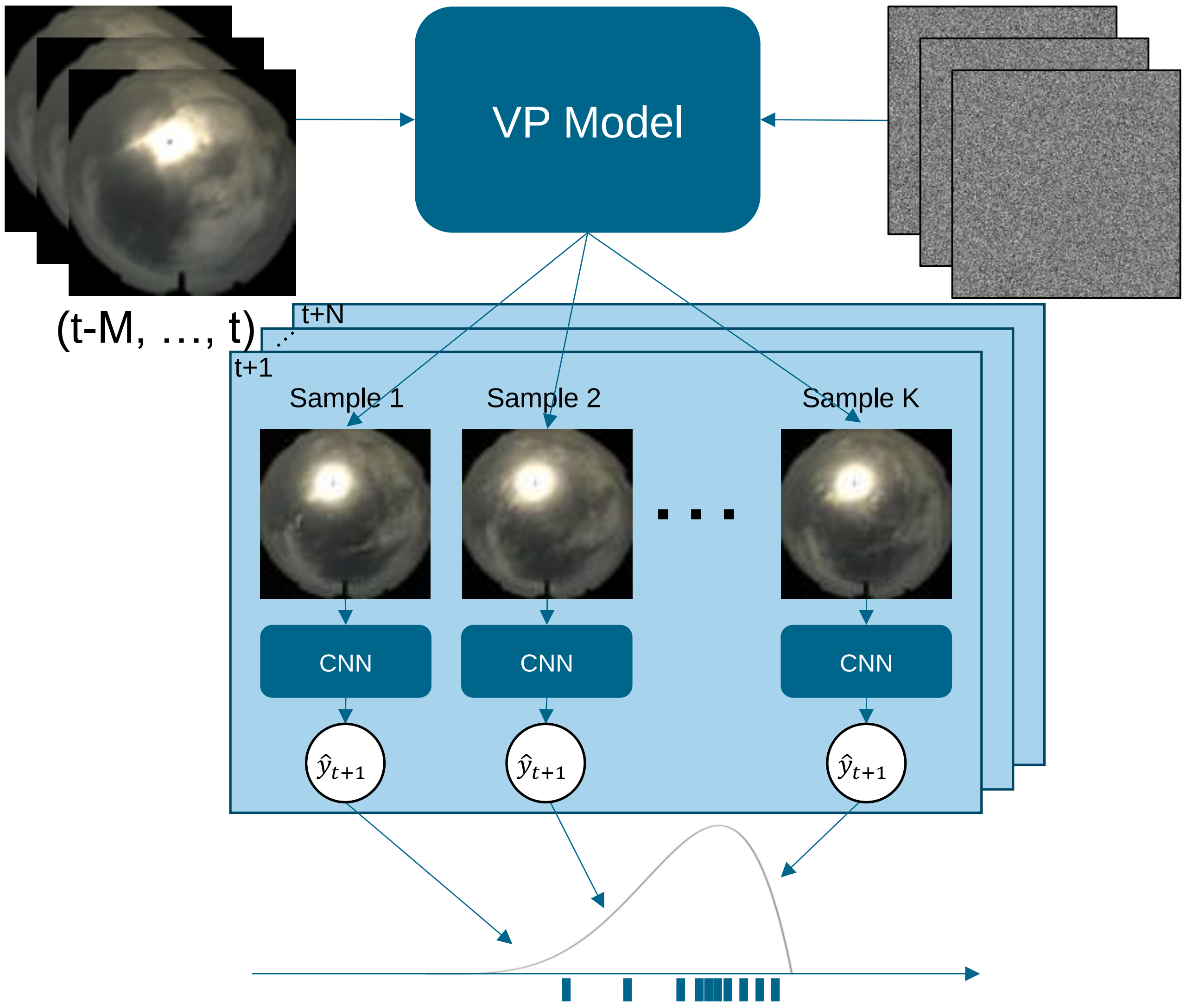
- Strong performance on standard error metrics, but ramp events undetected due to smoothed forecast curves
- How can we circumvent smoothing of the forecast curve?

GENERATIVE FORECASTING APPROACH

Generative Forecasting
Model Architecture

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

- **Classifier:**
  - Given individual predicted future frames, a second model (e.g. CNN) is used to derive target quantity
  - E.g., classifier predicts directly **ramp events** based on predicted sky images
  - Trained separately on real images



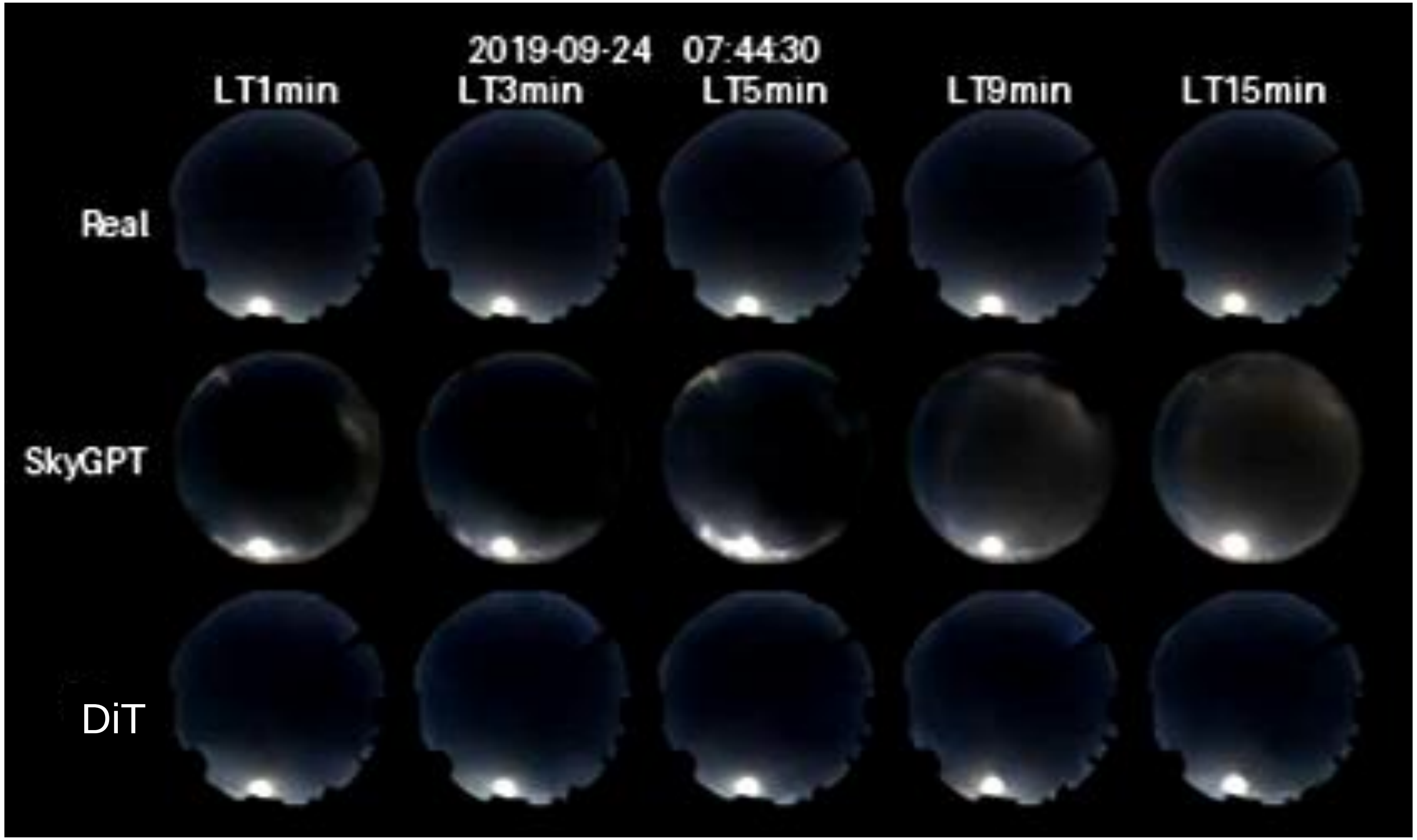
Video Prediction Models
Model options / Train and Test setup

- Two different generative models were tested
  - SkyGPT [5]: Adaptation of the VideoGPT [6] model combined with PhyCells [7]
  - DiT: Adaptation of the diffusion-based transformer model [8]
- Training
  - Both models were trained on selected camera data from CIEMAT‘s PSA
  - 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

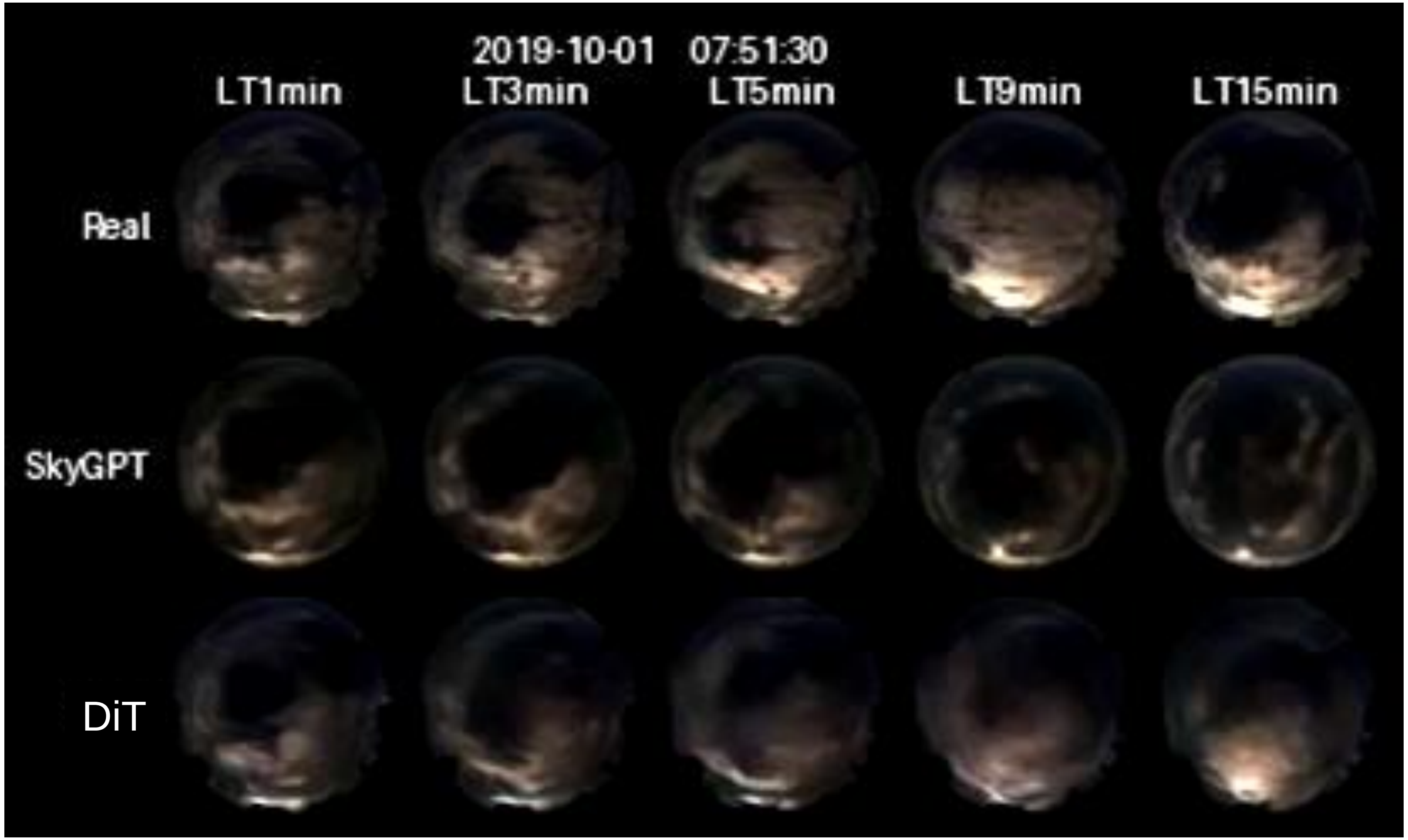
	SkyGPT	DiT
Image res.	64x64	128x128
Temporal res.	2min	1min
Forecast horizon	15min	30min
Variational Auto Encoder	VQ-VAE [6]	Pretrained VAE [9]

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



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

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

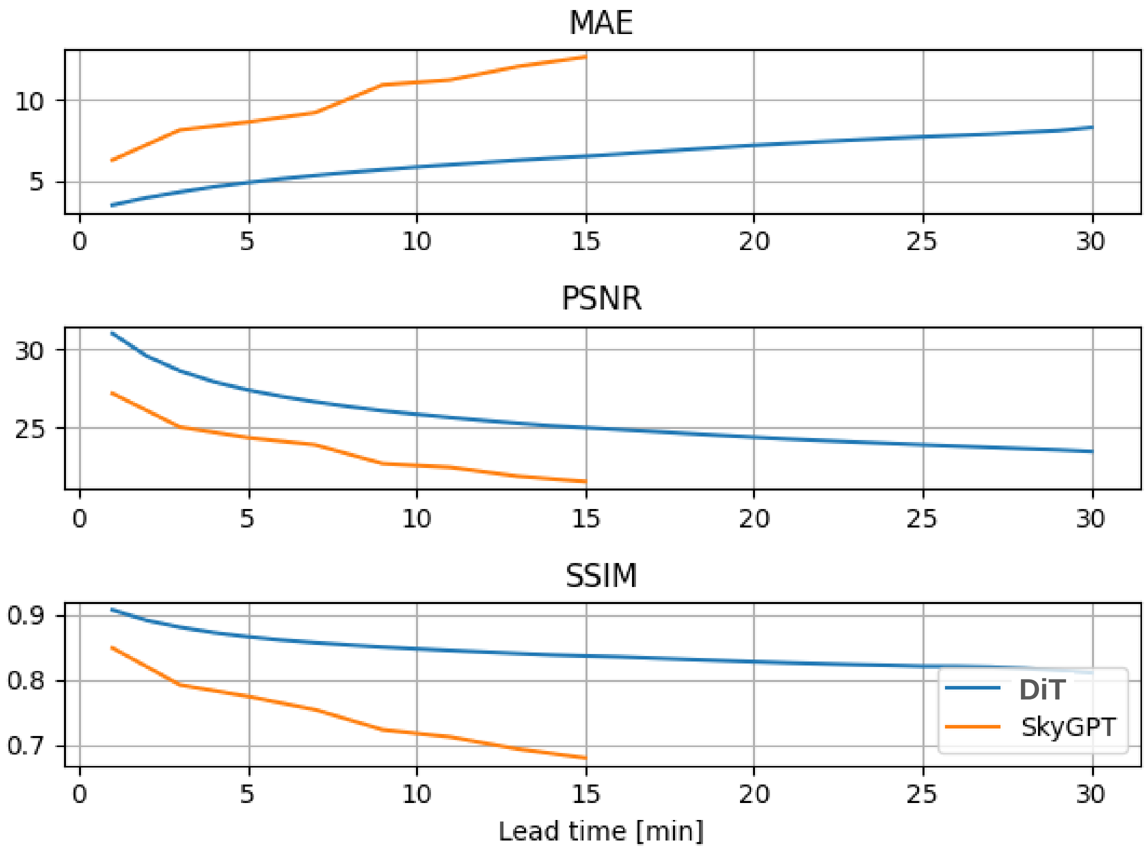


→ Strong deviations in terms of cloud coverage for larger lead times

Video Prediction Evaluation
Quantitative Evaluation of Images

- Evaluation of predicted sky image frames
  - Image pixel value 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

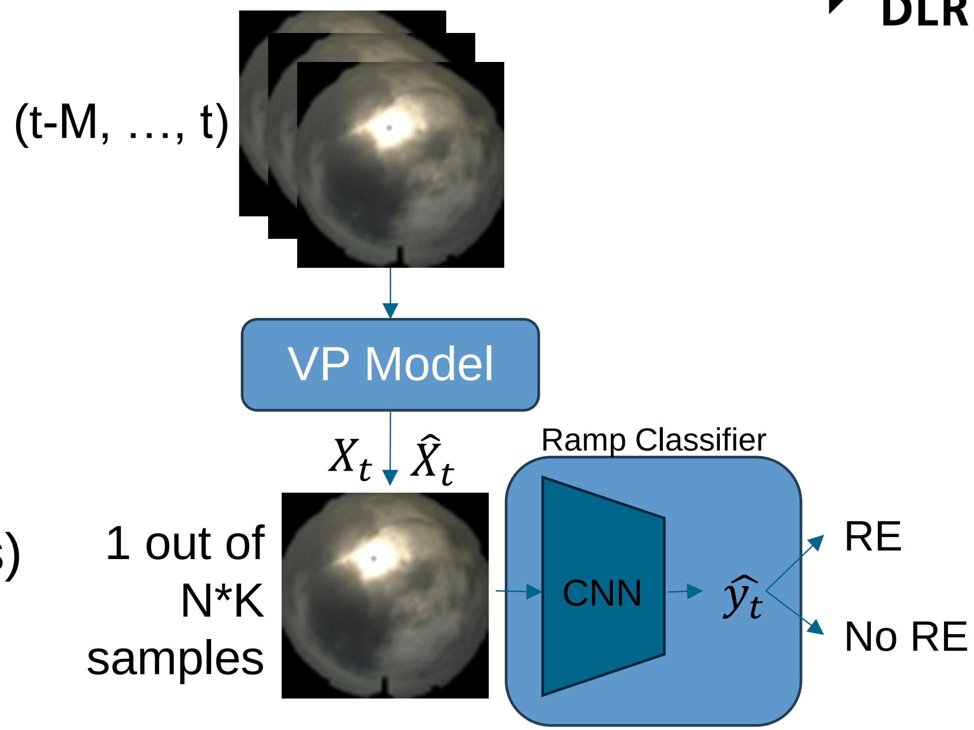
RAMP EVENT PREDICTION

Ramp Event Prediction

- <u>Ramp Classifier</u>
  - CNN predicting probability of ramp event based on corresponding image ($\hat{y}_t = \mathrm{P}(\mathrm{RE}|\mathrm{X_t})$)
  - Ground truth labels based on GHI measurements

$$\mathrm{y_t} = \begin{cases} 1 \;\; if \; \exists \; ramp \; in \; [t-3, t+3] \\ 0 \;\; otherwise \end{cases}$$

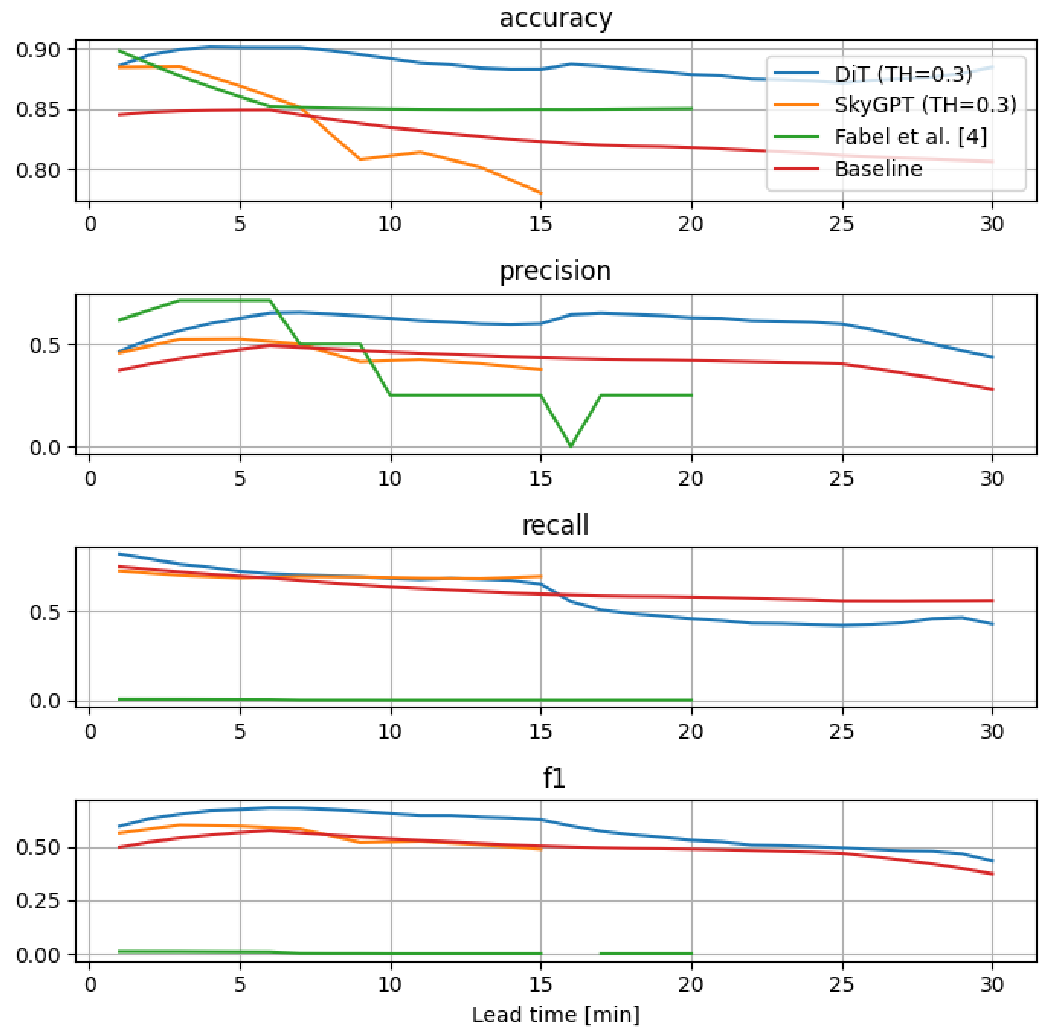
  - **Trained on real sky images**
  - **Evaluated on synthetic images** from generative model to obtain ramp event prediction
- Video prediction models generate K=4 samples (images) for all lead times for each forecast -> 4 probabilities
- Ramp is predicted if average probability > threshold
  - Low classifier threshold TH=0.3 chosen to prioritize recall



- <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

- Reference ramp data from pyranometer with tolerance window (tw=10min)
- Evaluation of classification metrics over lead times



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

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

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

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

Conclusion

- Low RMSE of forecasts does not guarantee realistic representation of irradiance variability & ramp events
- Generative, image-based modeling is promising alternative to capture short-term variability
  - Current video prediction models for ASI still struggle for higher lead times (inconsistencies & physically unrealistic cloud scenes)
  - Generated images remain useful for detecting ramp events also for higher lead times
- Outlook
  - Enhance ASI-based video prediction
    - Focus training on highly variable cloud conditions
  - 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

Acknowledgements

Funded by the Federal Ministry for the Environment, Climate Action, Nature Conservation and Nuclear Safety (BMUKN) on the basis of a resolution of the German Bundestag (Ausesol AI, grant number 67KI21007A).

THANK YOU FOR YOUR ATTENTION!
QUESTIONS?

STEFAN.WILBERT@DLR.DE
YANN.FABEL@DLR.DE