RUSH

Rapid-Update High-Resolution precipitation nowcasting and Global AI/NWP downscaling in a single Latent-Diffusion model

April 15th, 2026

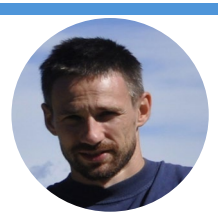
ESA-ECMWF workshop, Machine Learning for Earth System Observation and Prediction, 13-17 April 2026 | ECMWF, Bologna (IT)

Authors: Simon De Kock (1, 2), Gabriele Franch (3), Elena Tomasi (3), Giacomo Tomezzoli (3)
Co-Authors: Rishabh Wanjari (3), Matteo Angelinelli (4), Lesley De Cruz (1, 2), Marco Cristoforetti (3), Renata Pelosini (5), Thomas Gastaldo (6), Virginia Poli (6), Francesca Vittorioso

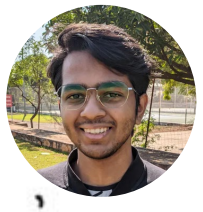
(1) Electronics and Informatics (ETRO), Vrije Universiteit Brussel, Brussels, Belgium
(2) Royal Meteorological Institute of Belgium, Brussels, Belgium
(3) Fondazione Bruno Kessler, Data Science for Industry and Physics, Trento, Italy
(4) HPC Department, Cineca, Bologna, Italy
(5) Arpa, Piemonte, Italy
(6) Arpae, Emilia-Romagna, Italy

The team

FONDAZIONE BRUNO KESSLER







VUB RMI


arpae emilia-romagna


CINECA

PIEMONTE Arpa Agenzia Regionale per la Protezione Ambientale


The goal

Downstream users need **high resolution**, **sub-hourly**, **probabilistic**, **rapidly updated** forecasts to act on fast-changing rainfall (eg. disaster response, hydrology, …)

- **High resolution**:
  1 km (same as observations)
- **Probabilistic:**
  Ensembles → uncertainty estimation
- **Rapid update**:
  forecast up to date with latest observation (every 5 min with latest radar & satellite)
- **National scale:**
  Belgium and Italy
- **24 h Lead time**
- **Sub-hourly timestep**

Nowcast	NWP/LAM	MLWP
✓	✓	✗
✓	✓	✓
✓	✗	✗
✓	✓	!
✗	✓	✓
✓	✗	✗

Our approach

Rapid update:

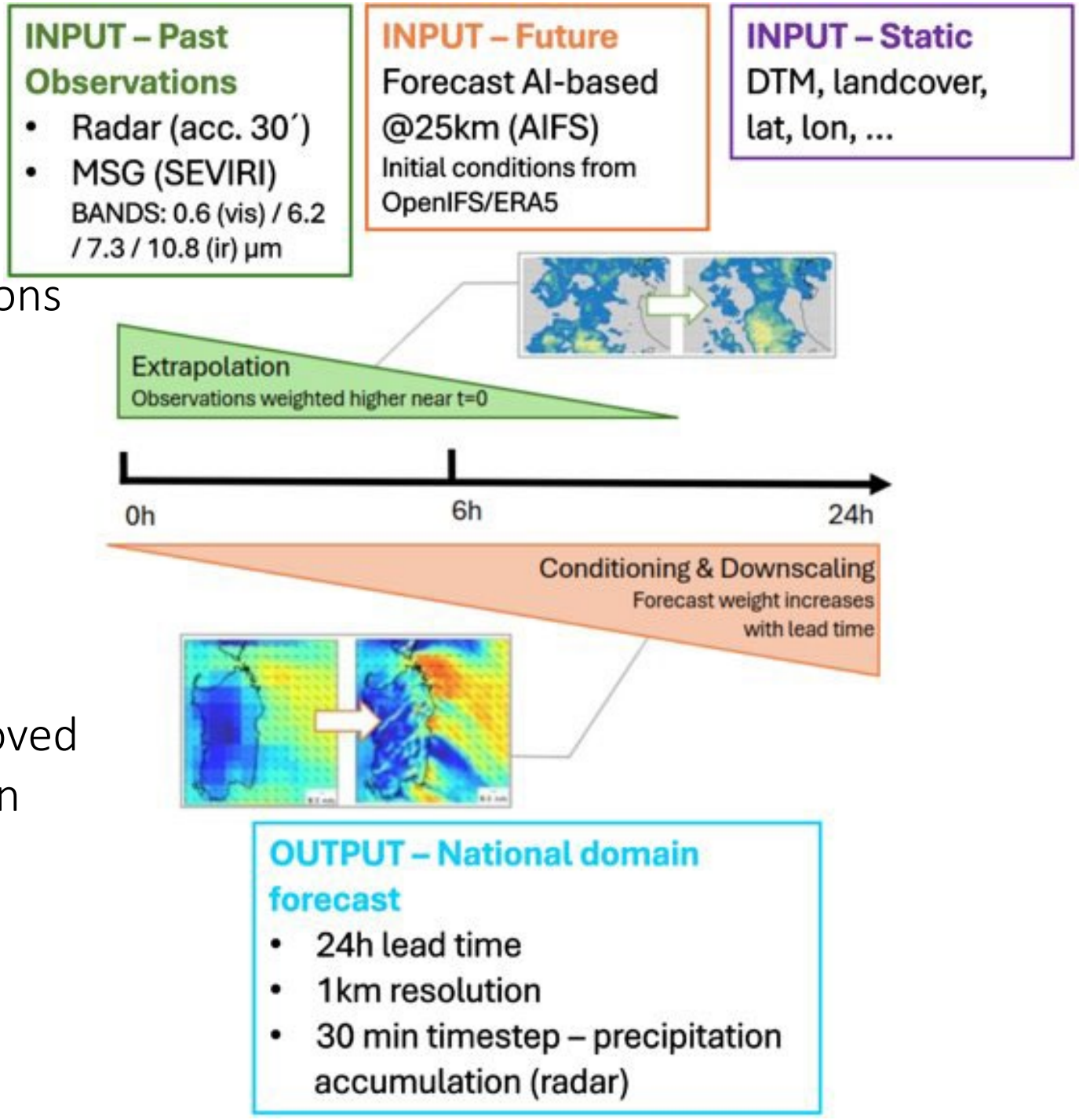
- Take as many readily available observations as possible: **Radar + Satellite**
- **Nowcast** the future states

Use predictions from N/ML WP:

- **Downscale** spatio-temporal course predictions using static information

Model choice

- **Latent diffusion** models have already proved themselves in nowcasting (**LDCast**) and in downscaling (**LDM_res**) of precipitation
- We combine both tasks in one model

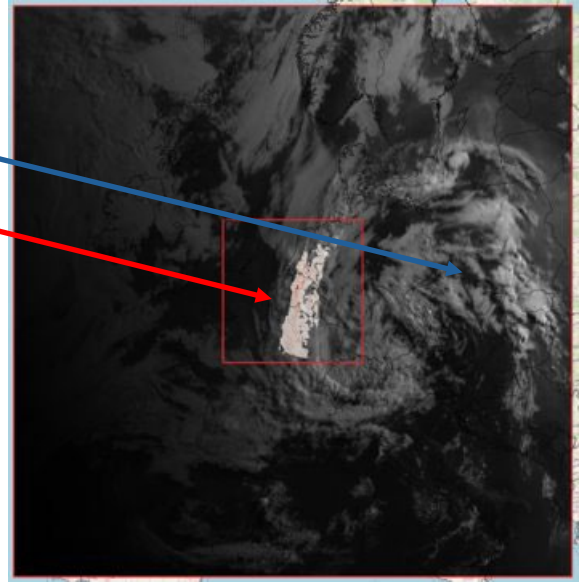


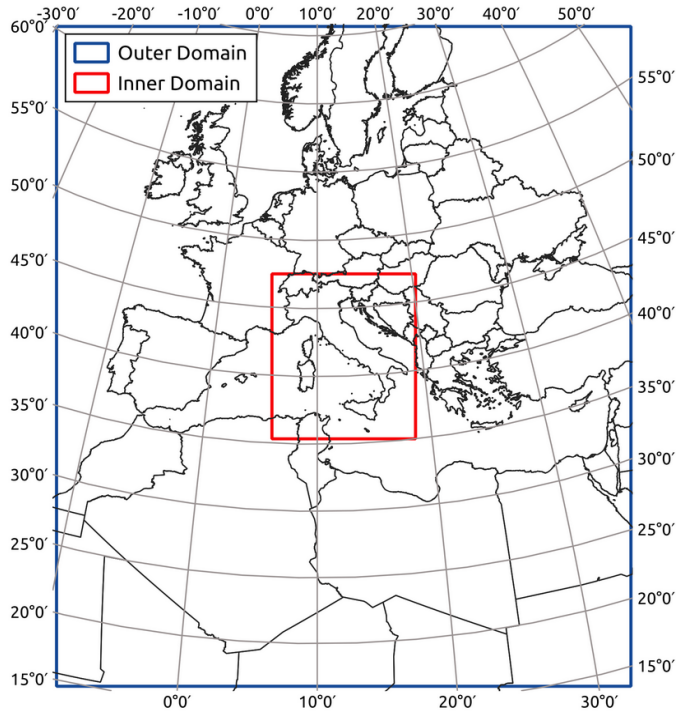
LDCast:
https://doi.org/10.48550/arXiv.2304.12891
LDM_res:
https://doi.org/10.5194/gmd-18-2051-2025

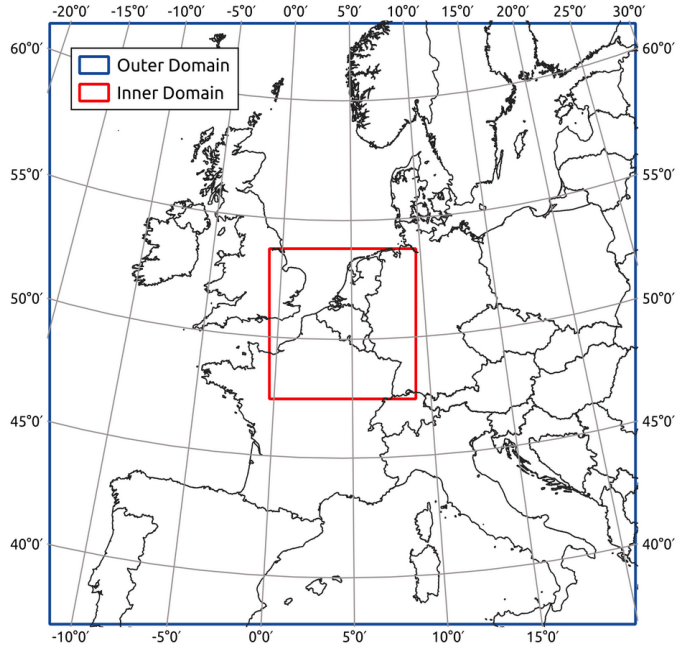
Data

Satellite

Radar







Harmonised to common grid @30 min

Radar (Input and Ground truth)

Rolling window 30 min accumulated precipitation from national radar QPE (e.g. for Belgium useful domain of 500x500@**1km**)

Satellite (Input)

Four MSG SEVIRI bands: (IR + 2 WV + VIS) every 15 min

Domain 4 times larger centred on radar domain (e.g. for Belgium 2816x2816@**4km**)

NWP fields: Determined in survey (at RMI and ItaliaMeteo) covering Thermodynamics, Dynamics and Moisture & Radiation

Static data: Use radar mask, orography and dynamic landcover information for downscaling purposes

Data Size and split

11 years archive processed → **over 55 TB raw data**

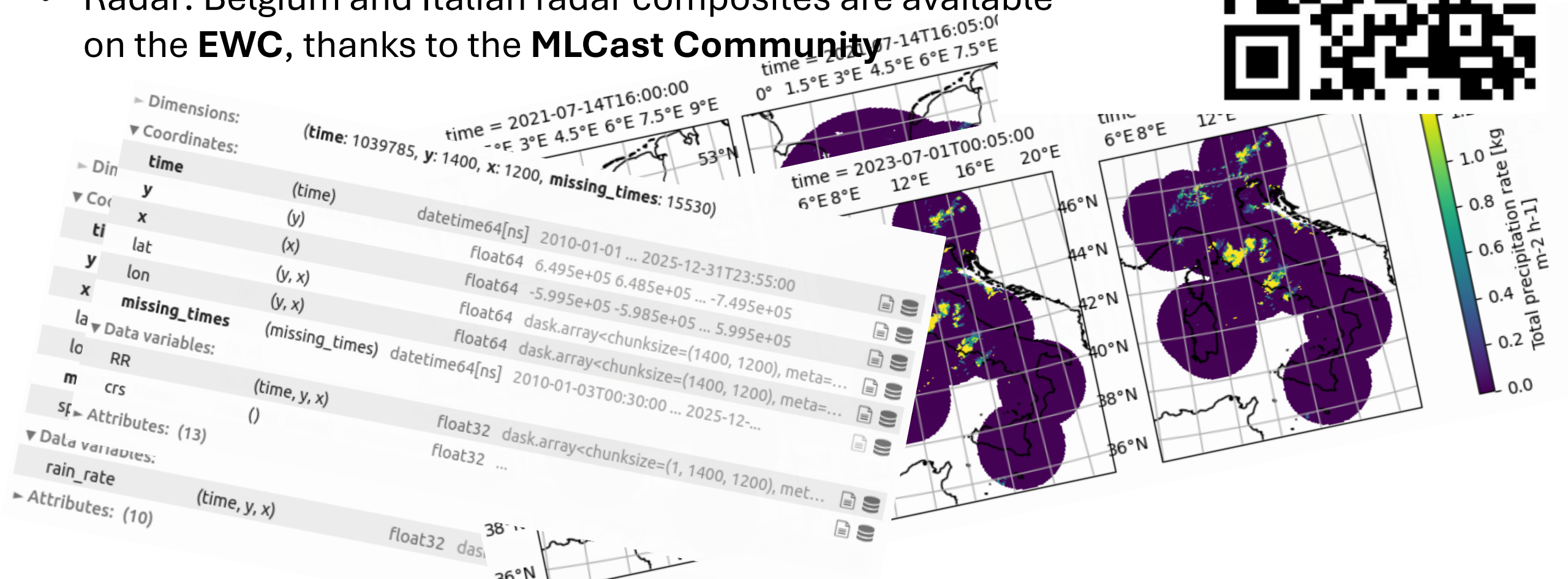
Train (80%) / Val (20%): 2014 – 2025

Test: 2024 + selected extreme cases

Data availability

All data used in this project is publicly available:

- Satellite: MSG SEVIRI
- Proxy NWP: ERA5
- Radar: Belgium and Italian radar composites are available on the **EWC**, thanks to the **MLCast Community**



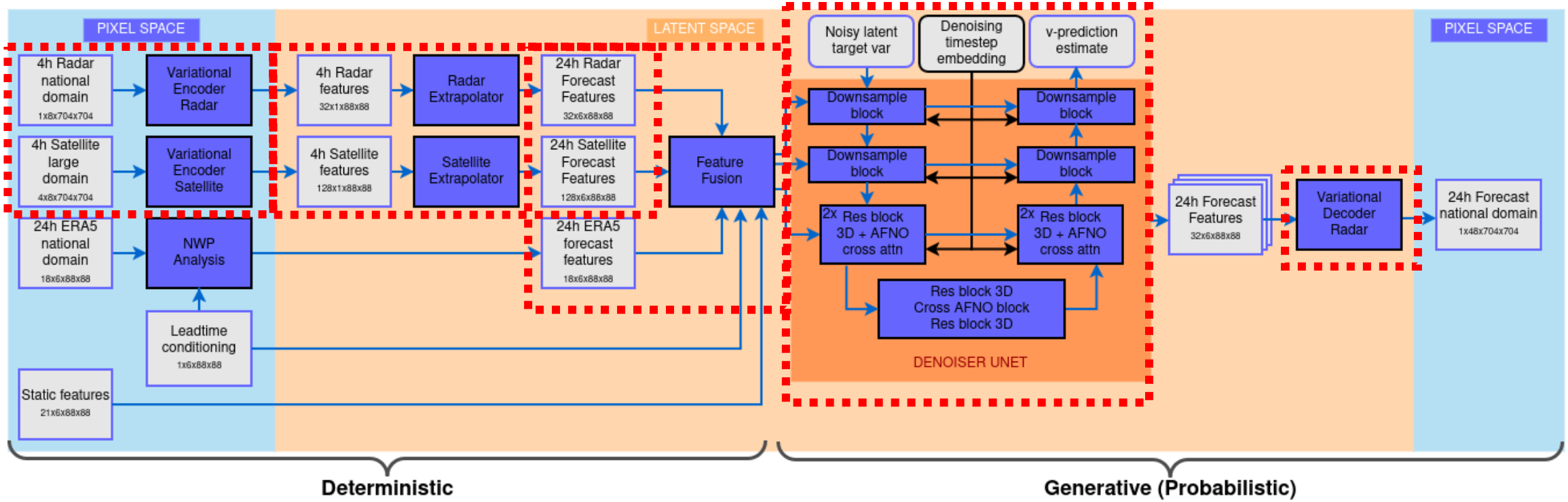
The latent diffusion model

Two processing spaces

- Observation space (pixel space): direct radar & satellite inputs.
- Latent space: compressed, efficient representation learned by VAEs.

Model architecture (4 main components)

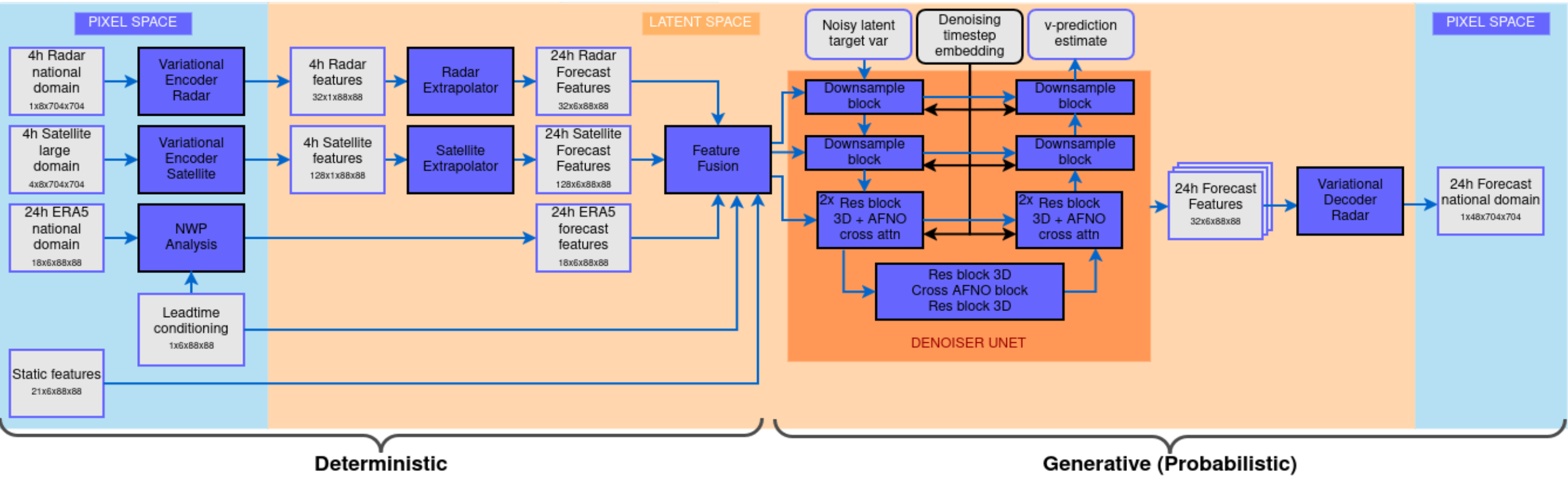
- **VAEs** – encode 4 h stacks of radar & satellite → latent representation
- **Extrapolators** – provide a first guess of future observation states.
- **Feature fusion** – combines extrapolators with NWP/MLWP and static input.
- **Denoiser (U-Net)** – conditioned on fused features, refines the forecast.



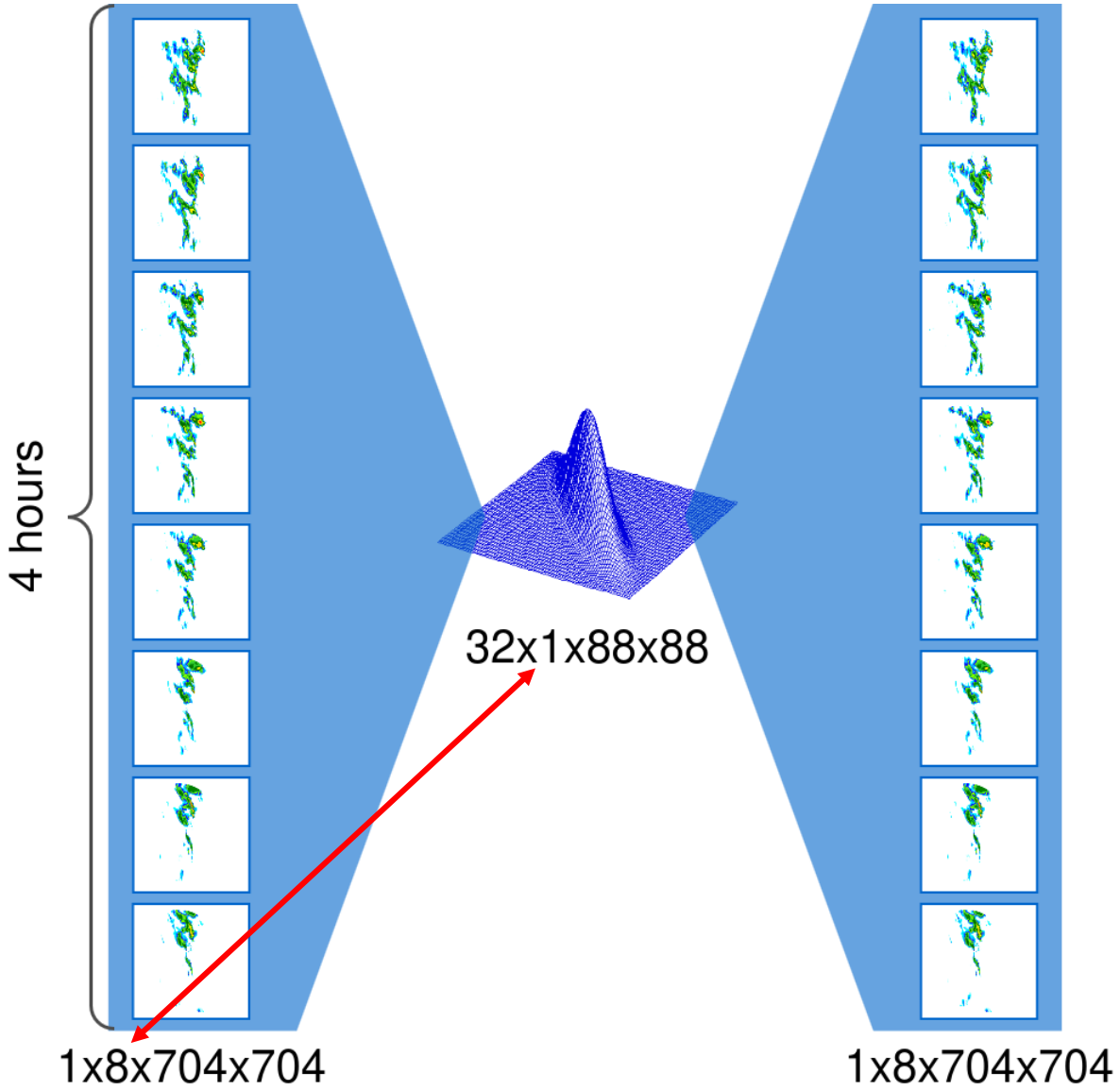
RUSH – Experiments

Many tasks, split up in different experiments:

- **VAEs:** Proper evaluation of VAEs
- **Nowcasting mode:** Evaluate the performance of the extrapolators
- **Downscaling mode:** Effect of low spatio-temporal resolution of **ERA5** on the model
- **Full RUSH:** Combining all modalities



Variational autoencoders - Experiments



Crucial: high-quality latent representation is essential, since all downstream forecasting relies on it.

Challenge:

Compression to single latent 'timestep'

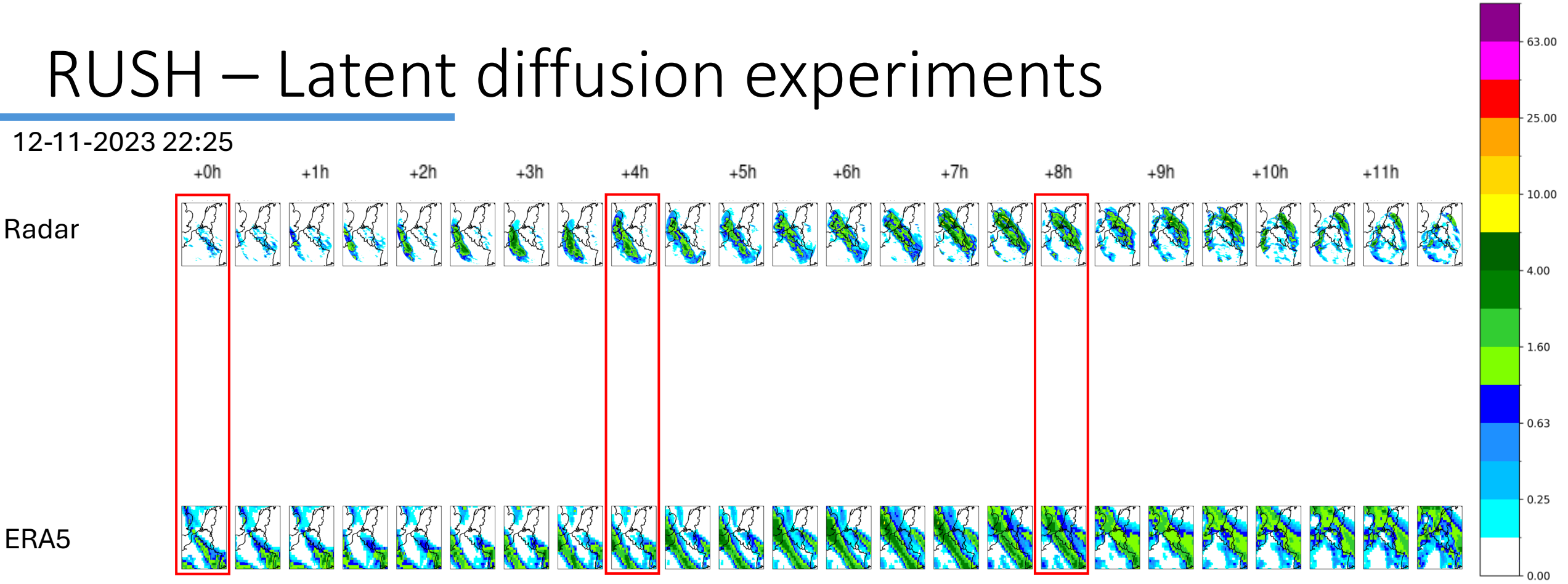
- Maintain "good" compression
  - Preserve temporal dynamics
  - Preserve spatial reconstruction
  - Semantically meaningful representation

Radar images come with added challenges (compared to satellite):

- Very sparse data
- Complex spatial variability
- Lower spatial redundancy

RUSH – Latent diffusion experiments

12-11-2023 22:25

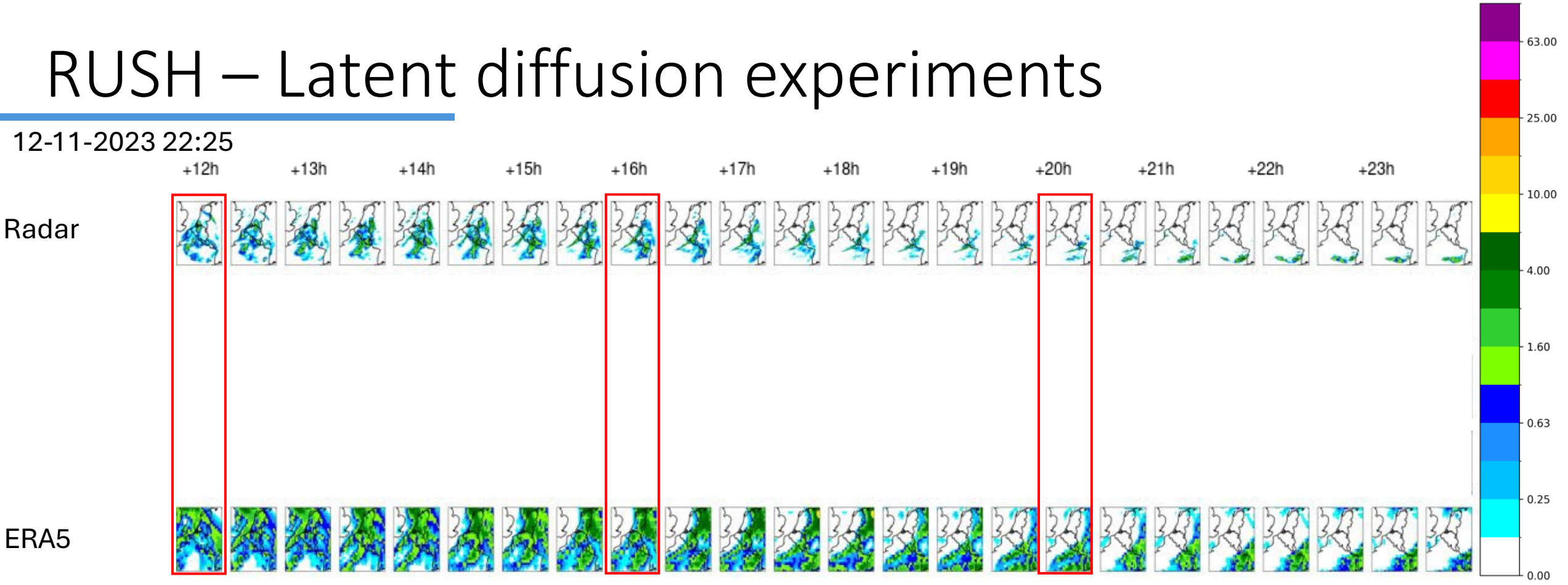


Notes:

- Looking at a "deterministic" instance of a probabilistic model
- Used ERA5 as a proxy for NWP to remove component of model predictability

RUSH – Latent diffusion experiments

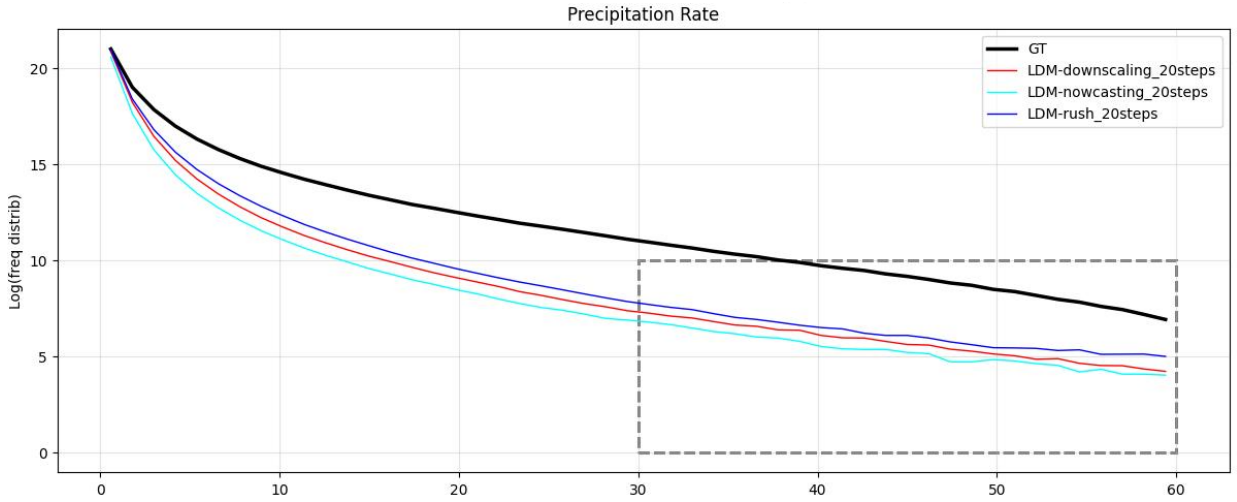
12-11-2023 22:25

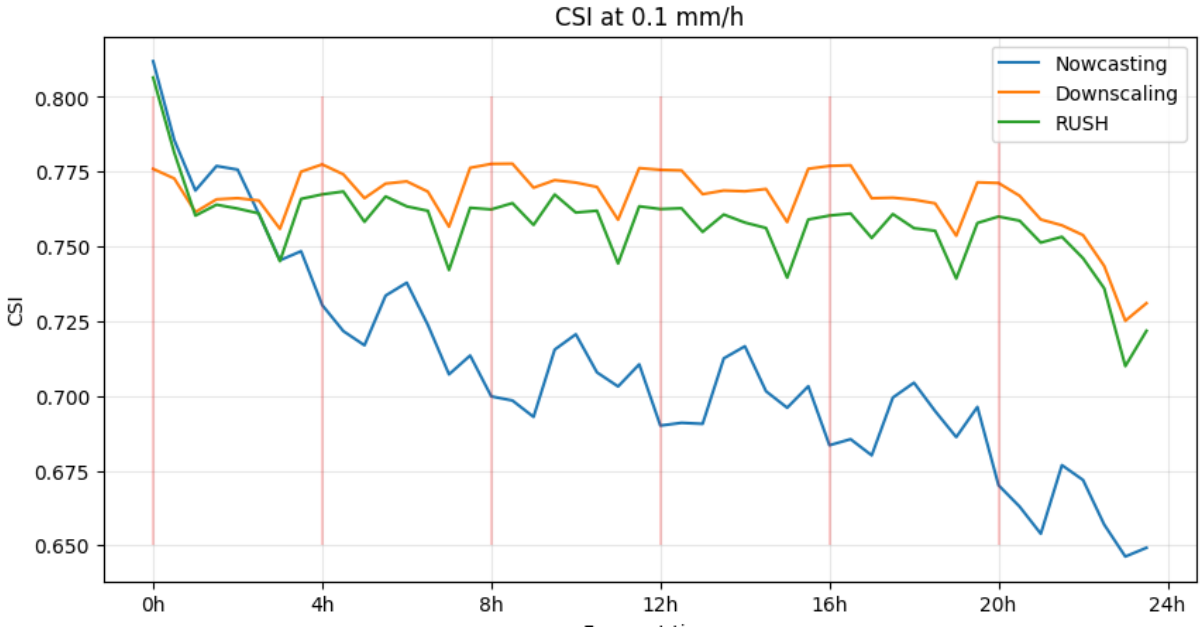


Conclusions:

- We are underpredicting intensities
- Downscaling is spatially correct
- Nowcasting is doing well in the first 2 hours
- RUSH is doing a combination of Nowcasting and Downscaling

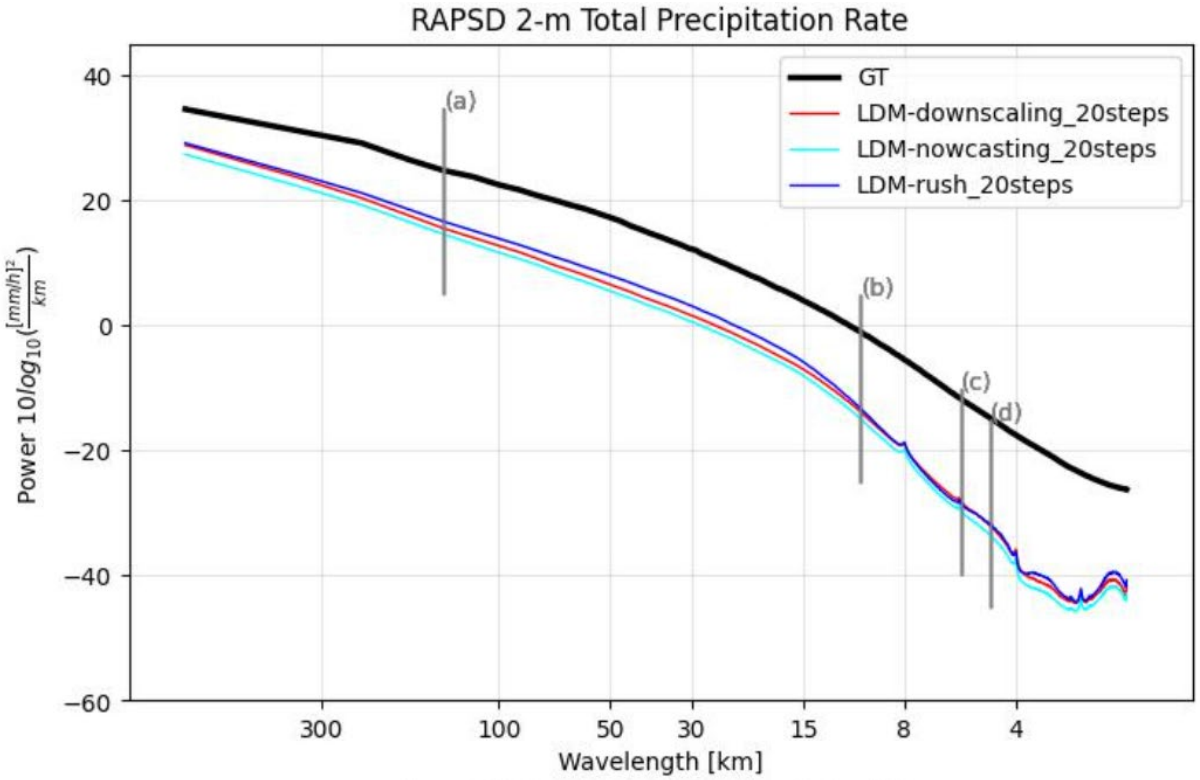
RUSH – Latent diffusion experiments





Conclusions (500+ predictions):

- We are underpredicting intensities
  Likely due to overfitting
- Predictions are spatially correct



Conclusion and outlook

What's done

- All data is prepared
- All modalities have been trained at least once
- VAE architecture is fixed

Next steps / To do

- Fix intensity mismatch (likely due to overfitting)
- Train on larger subset of the data/Shrink the model
- Do full evaluation: compare against existing nowcasting, NWP and MLWP baselines.

Thank you for your attention

Simon De Kock

PhD Student – AI4WCM group - Electronics and Informatics (ETRO)

Vrije Universiteit Brussel, Brussels, Belgium

simon.de.kock@vub.be

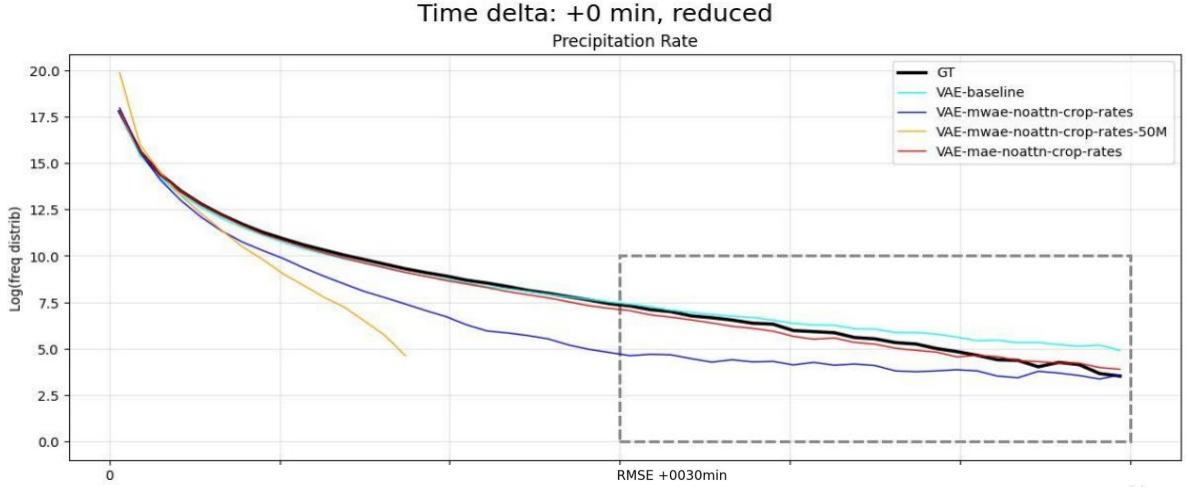
Backup slides

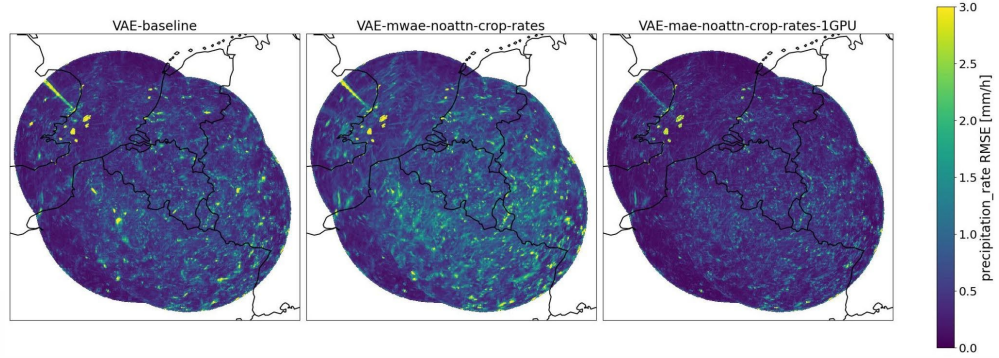
Results - VAE radar

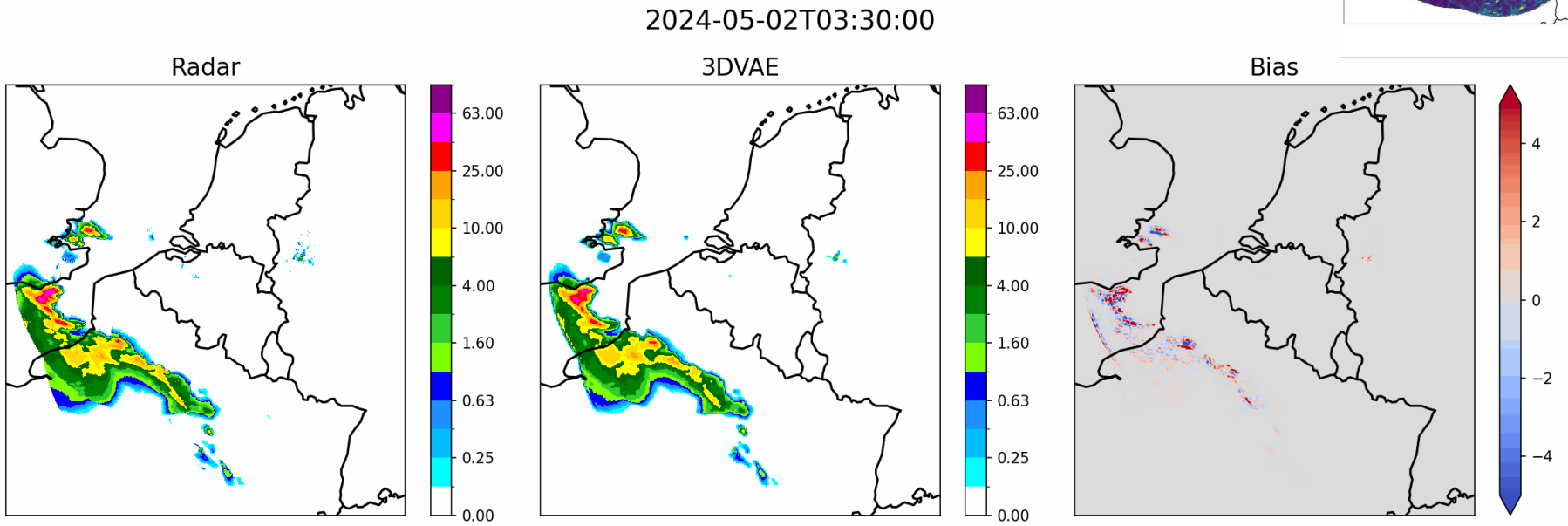
Compare different model configurations:

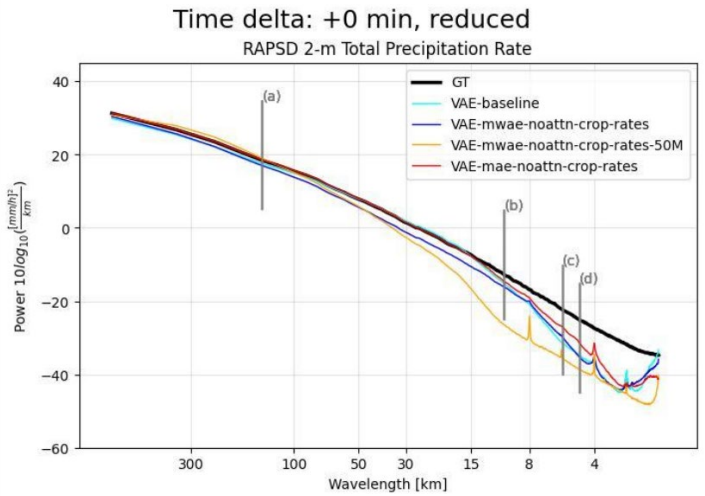
- Reconstruction (RMSE & continuity)
- Frequency distribution and power spectra

For all frames **individually**

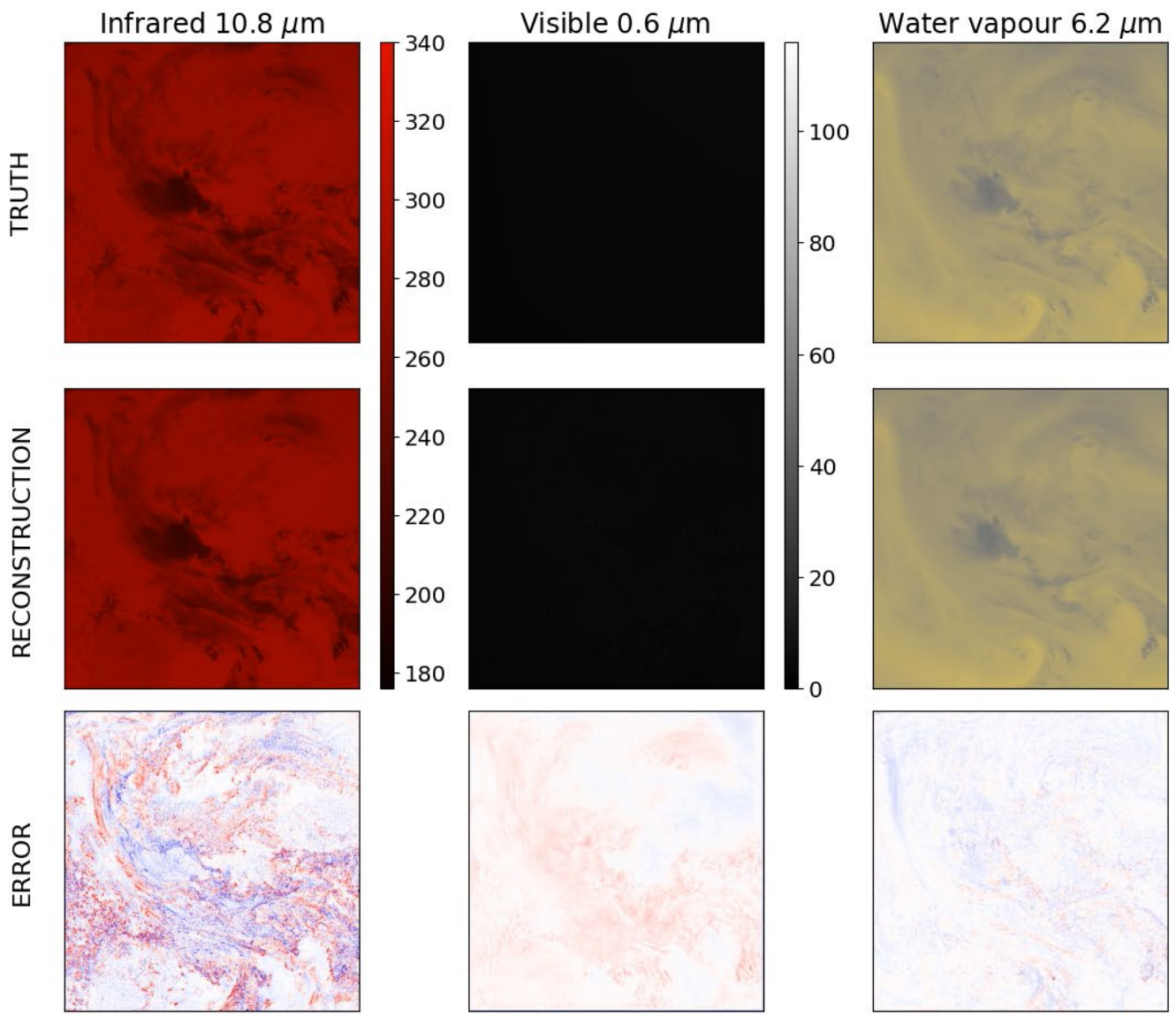


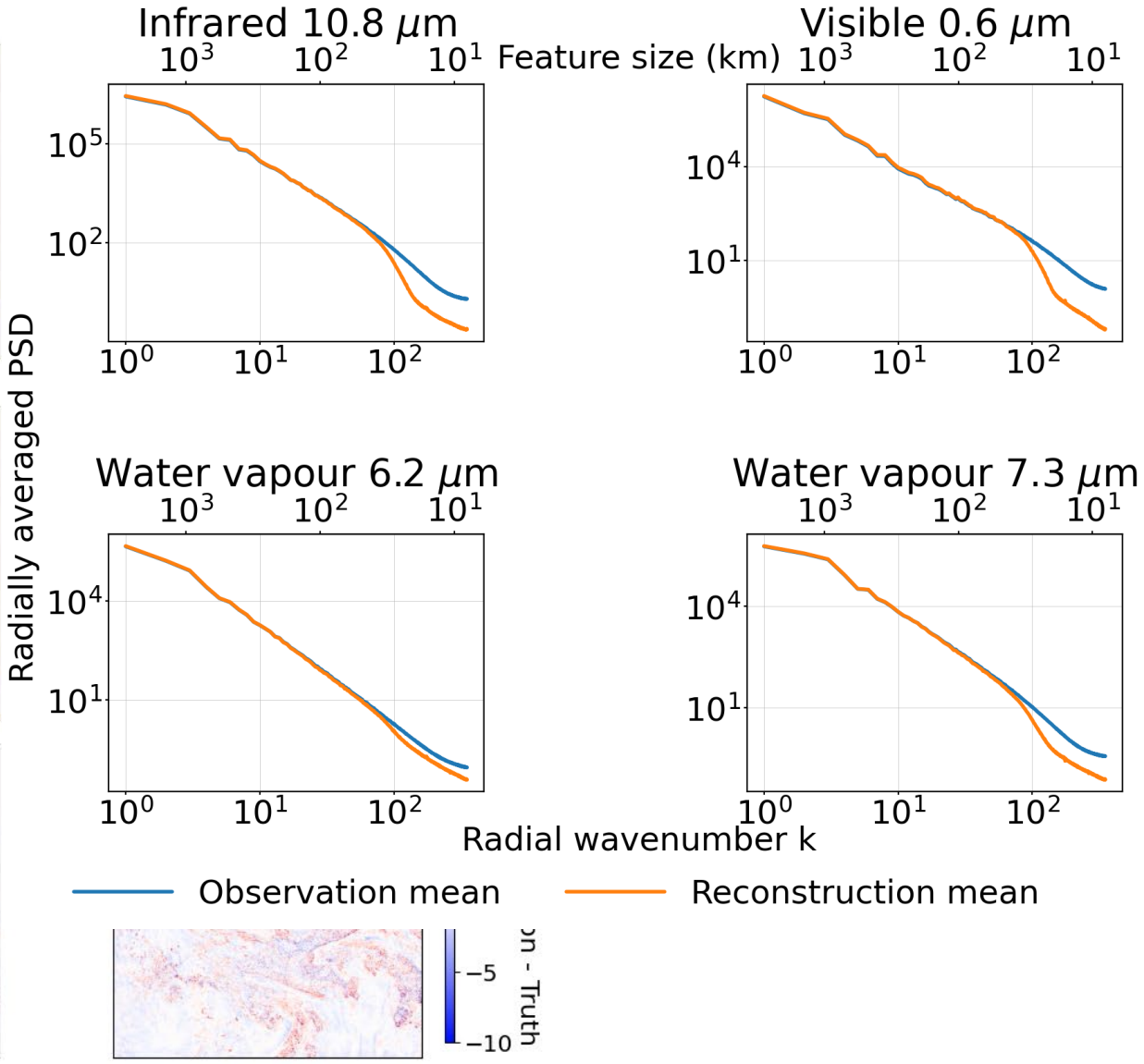




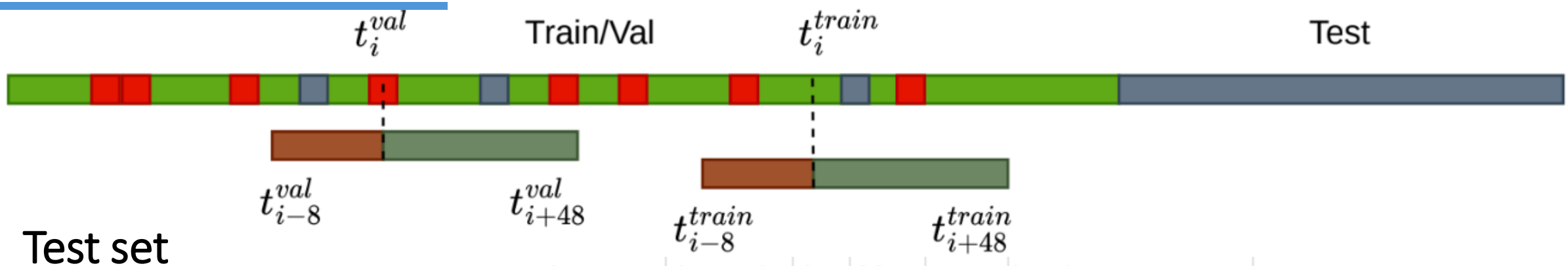


Results - VAE satellite





Train/validation/test



Test set

Large variety of:

- Length
- Years
- Seasons
- Locations
- Intensities

Full 2024 for general evaluation

date_start	date_end	days	CC	season	location	notes
2017-09-10	2017-09-10	1	IT	SON	Toscana	Alluvione a Livorno
2018-08-20	2018-08-20	1	IT	JJA	Calabria	Piena improvvisa del torrente Raganello
2018-09-04	2018-09-05	2	BE	SON	Antwerp	
2018-10-28	2018-10-29	2	IT	SON	Trentino	
2019-11-23	2019-11-24	2	IT	SON	Liguria	
2020-02-09	2020-02-10	2	BE	DJF	Antwerp	
2020-08-17	2020-08-17	1	BE	JJA	Antwerp	
2020-10-02	2020-10-03	2	IT	SON	Piemonte e Trentino	
2020-12-06	2020-12-06	1	IT	DJF	Friuli e Calabria	
2021-01-26	2021-01-29	4	BE	DJF		
2021-07-12	2021-07-15	4	BE	JJA	Vesdre catchment	400Y retrun period event
2021-07-23	2021-07-27	5	BE	JJA	Dinant	
2022-04-06	2022-04-07	2	BE	MAM		
2022-09-15	2022-09-15	1	IT	SON	Marche	Alluvione nelle Marche

Static data & NWP variables

NWP variables:

Pressure levels:
- Geopotential: z_500, z_700, z_850
- Specific humidity: q_500, q_700, q_850
- Temperature: t_500, t_850
- Vertical velocity: w_700
- U wind component: u_850, u_925
- V wind component: v_850, v_925

Surface / column variables:
- Convective precipitation: cp
- Surface solar radiation downwards: ssrd
- Total precipitation: tp
- Mean sea level pressure: msl
- Total column water: tcw

Static variables (Copernicus Global Dynamic Land Cover) :

- Radar mask
- Digital terrain model
- Closed forest, evergreen needle leaf
- Closed forest, deciduous broad leaf
- Closed forest, mixed
- Closed forest, unknown
- Open forest, evergreen needle leaf
- Open forest, deciduous broad leaf
- Open forest, mixed
- Open forest, unknown
- Shrubs
- Open sea
- Herbaceous vegetation
- Cultivated and managed vegetation / agriculture
- Urban / built up
- Bare / sparse vegetation
- Snow and ice
- Permanent water bodies
- Herbaceous wetland
- Projected x-coordinate
- Projected y-coordinate

Physical consistency

12-11-2023 20:25