AI for weather and climate

Bedartha Goswami

machine learning in climate science

IISER PUNE

Tutorial, CODS 2025

Deep Learning Models

~~AI~~ for weather and climate

Deep learning …

:: More than function approximation

:: Uses NNs to learn representations

:: Learnt representations help solve a task

REVIEW

doi:10.1038/nature14539

Deep learning

Yann LeCun[1,2], Yoshua Bengio[3] & Geoffrey Hinton[4,5]

Deep learning allows computational models that are composed of multiple processing layers to learn representations of data with multiple levels of abstraction. These methods have dramatically improved the state-of-the-art in speech recognition, visual object recognition, object detection and many other domains such as drug discovery and genomics. Deep learning discovers intricate structure in large data sets by using the backpropagation algorithm to indicate how a machine should change its internal parameters that are used to compute the representation in each layer from the representation in the previous layer. Deep convolutional nets have brought about breakthroughs in processing images, video, speech and audio, whereas recurrent nets have shone light on sequential data such as text and speech.

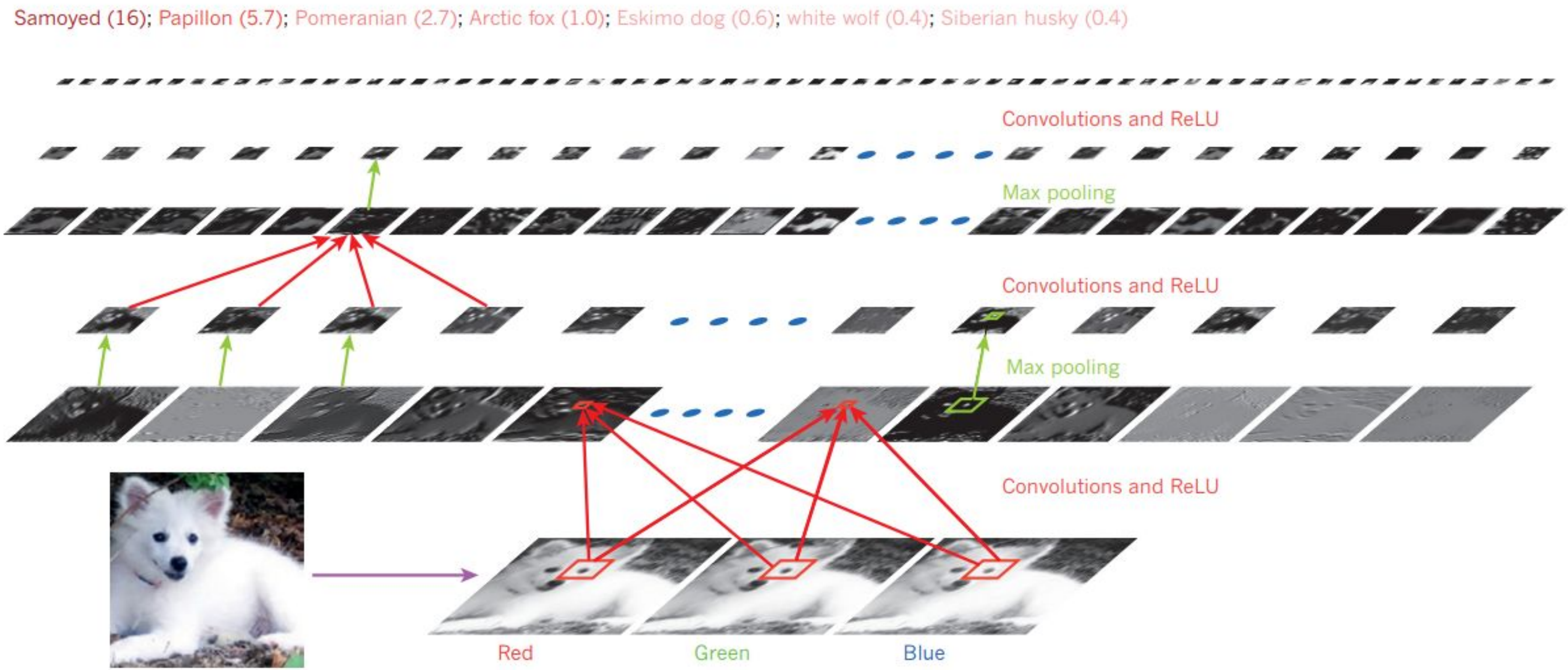


Figure 2 | Inside a convolutional network. The outputs (not the filters) of each layer (horizontally) of a typical convolutional network architecture applied to the image of a Samoyed dog (bottom left; and RGB (red, green, blue) inputs, bottom right). Each rectangular image is a feature map corresponding to the output for one of the learned features, detected at each of the image positions. Information flows bottom up, with lower-level features acting as oriented edge detectors, and a score is computed for each image class in output. ReLU, rectified linear unit.

Deep Learning Models

~~AI~~ for weather and climate

How do we cleverly set up a NN to learn useful representations?

Outline

:: Rainfall nowcasting

:: Weather forecasting

:: Downscaling (aka super-resolution)

:: Bias correction of model outputs

:: Sub-grid parameterisations

:: Climate projections

Outline

:: Rainfall nowcasting

:: Weather forecasting

:: Downscaling (aka super-resolution)

:: Bias correction of model outputs

:: Sub-grid parameterisations

:: Climate projections

Nowcasting

:: Radar images (reflectances or precipitation rates)

:: High spatial and temporal resolution (~1km, at every 15 mins to every 1 hour)

:: Training data ~ for around a few years

:: Test data ~ for around one year

:: Image is tiled and then sampled -> Many samples have no rain!

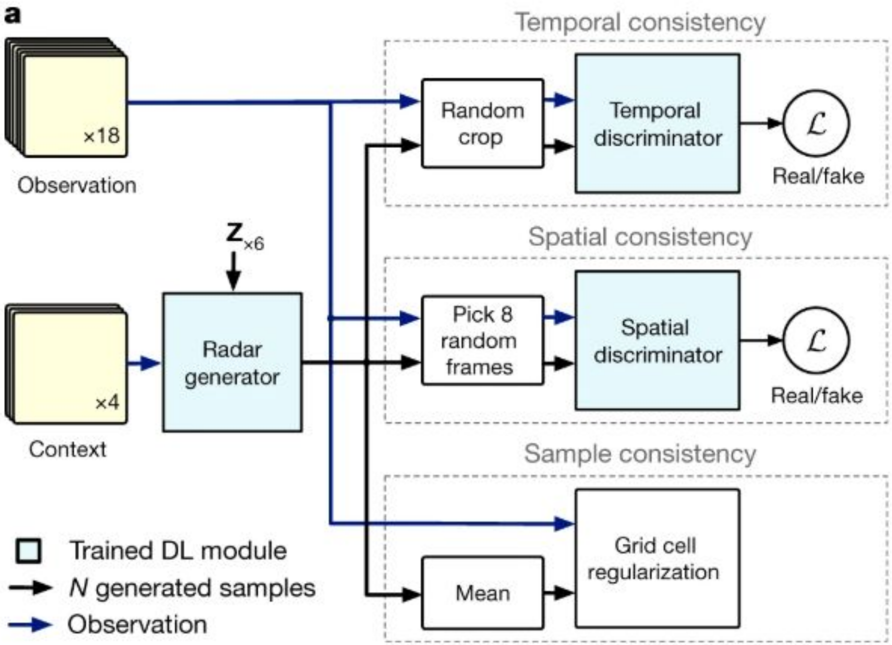
DGMR
Ravuri et al., *Nature*, 2019

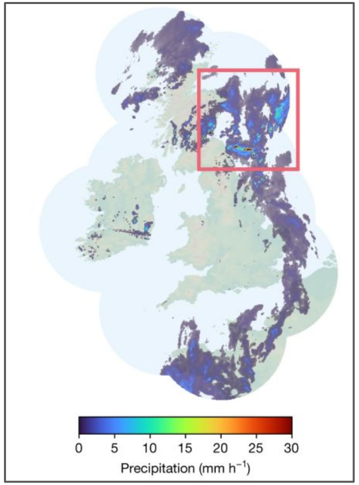
:: GAN architecture

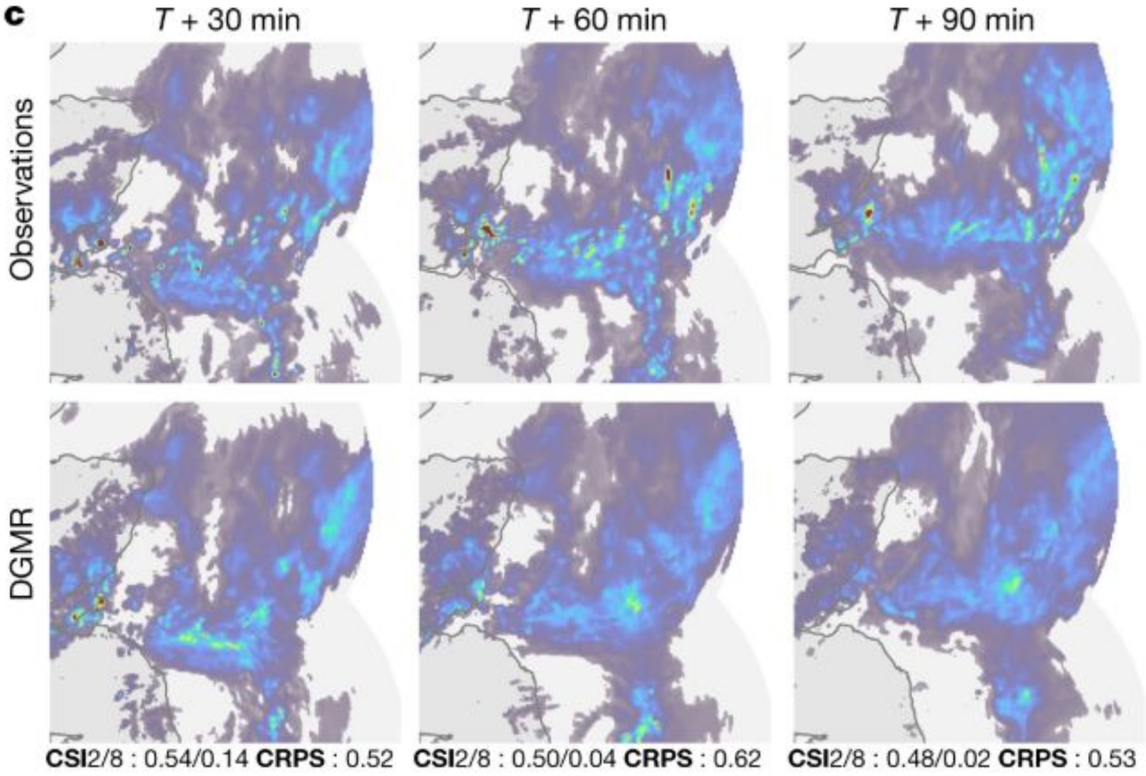
:: Discriminate space & time separately

:: Radar data -> importance sampling

:: Purely data-driven




0
5
10
15
20
25
30
Precipitation (mm h⁻¹)


c
T + 30 min
T + 60 min
T + 90 min
Observations
DGMR
CSI2/8 : 0.54/0.14 CRPS : 0.52
CSI2/8 : 0.50/0.04 CRPS : 0.62
CSI2/8 : 0.48/0.02 CRPS : 0.53

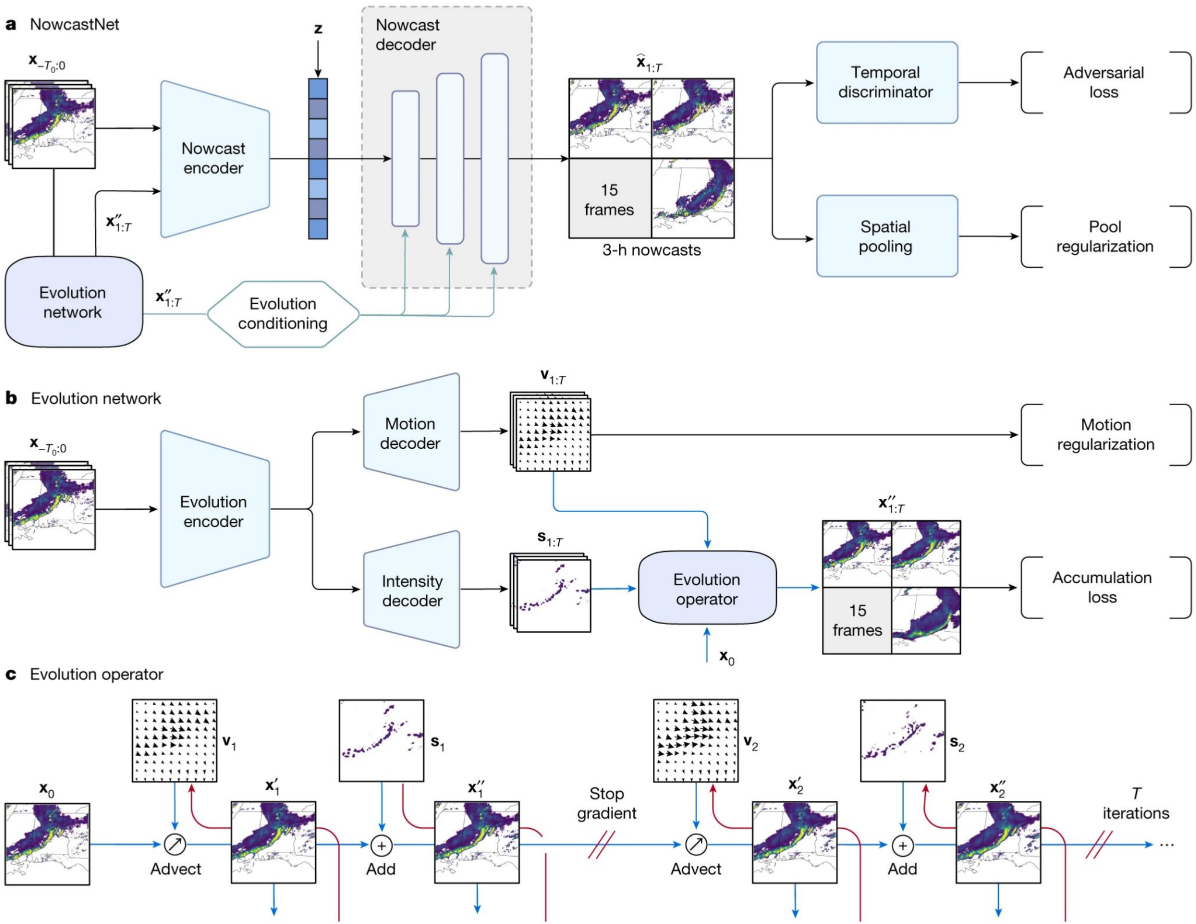
NowCastNet
Zhang et al., *Nature*, 2023

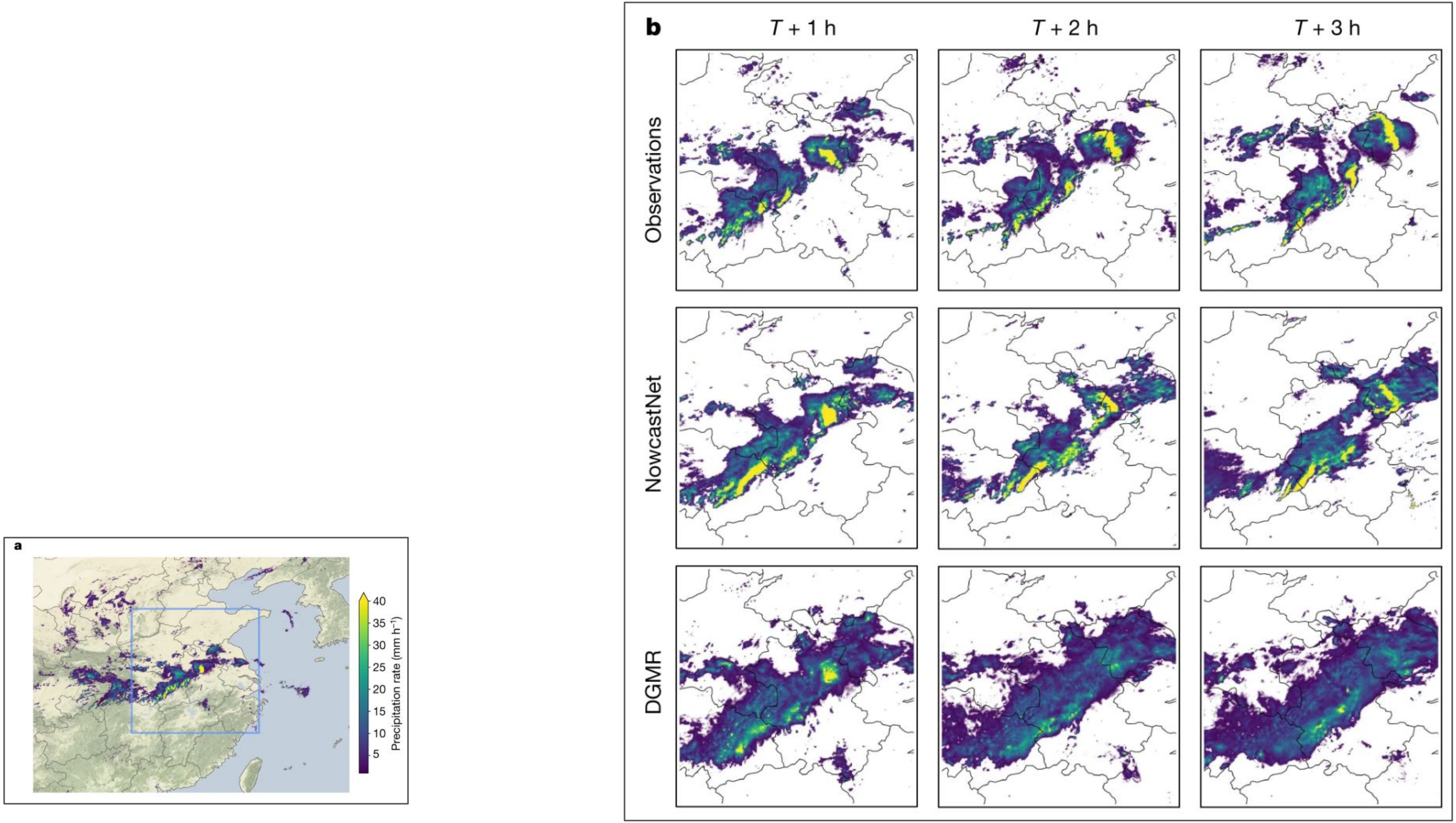
:: Hybrid model

:: GAN architecture

:: Conditioned on advective scheme

:: Captures extremes well




a
40
35
30
25
20
15
10
5
Precipitation rate (mm h⁻¹)
b
T + 1 h
T + 2 h
T + 3 h
Observations
NowcastNet
DGMR

Outline

:: Rainfall nowcasting

:: **Weather forecasting**

:: Downscaling (aka super-resolution)

:: Bias correction of model outputs

:: Sub-grid parameterisations

:: Climate projections

Weather forecasting

:: Predict future weather states based on current state

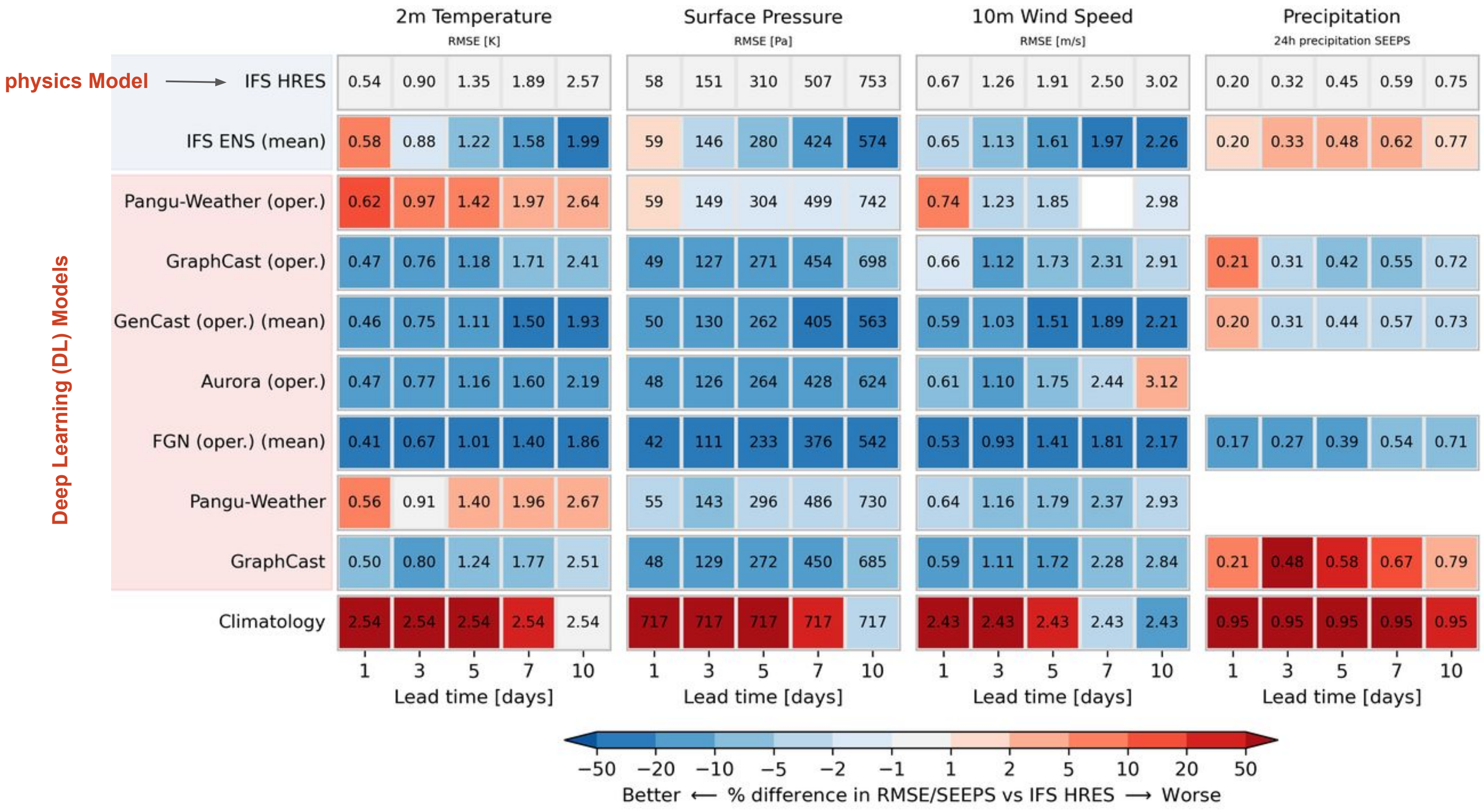
:: Weather state defined by multiple variables stacked together (surface variables + atmospheric)

:: Typically, two time steps in the past are used to predict the next time step -> iterative rollout

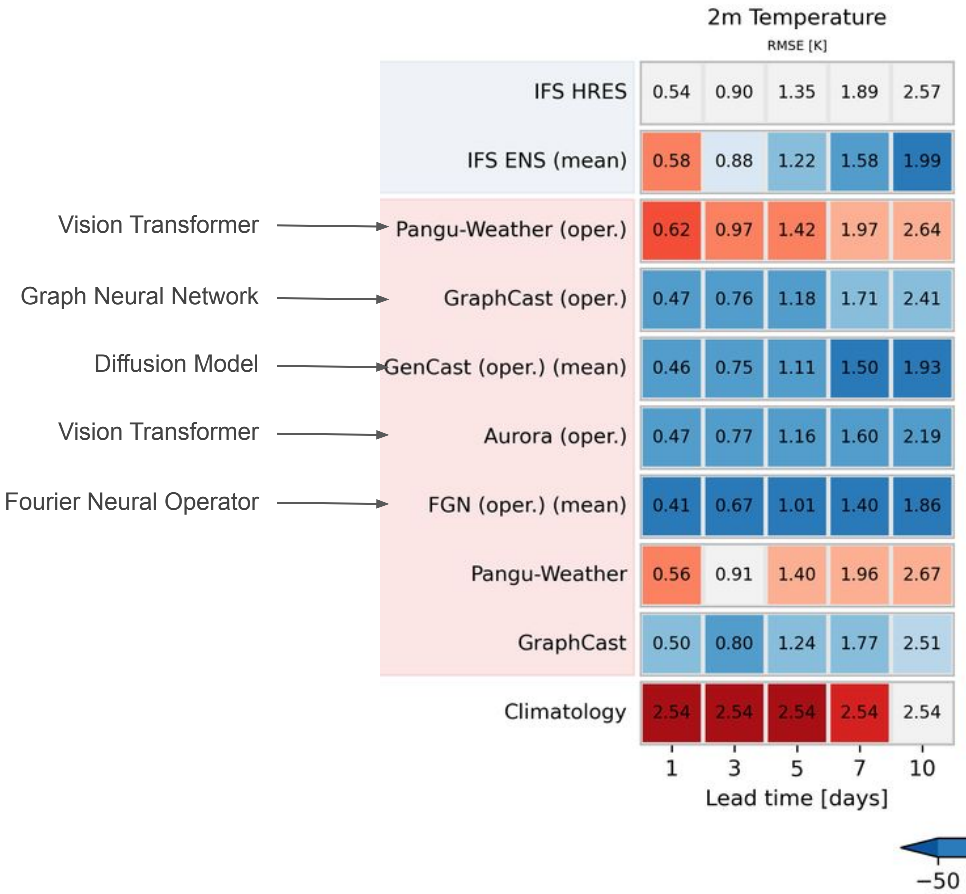
:: Training data ~ 40 years at 6h or 1h resolution, at 0.25 degrees globally

:: Test data ~ similar resolution but for 3-4 years

:: Popular benchmark - WeatherBench(2)



https://sites.research.google/gr/weatherbench/


2m Temperature
RMSE [K]
IFS HRES
0.54 0.90 1.35 1.89 2.57
IFS ENS (mean)
0.58 0.88 1.22 1.58 1.99
Vision Transformer
Pangu-Weather (oper.)
0.62 0.97 1.42 1.97 2.64
Graph Neural Network
GraphCast (oper.)
0.47 0.76 1.18 1.71 2.41
Diffusion Model
GenCast (oper.) (mean)
0.46 0.75 1.11 1.50 1.93
Vision Transformer
Aurora (oper.)
0.47 0.77 1.16 1.60 2.19
Fourier Neural Operator
FGN (oper.) (mean)
0.41 0.67 1.01 1.40 1.86
Pangu-Weather
0.56 0.91 1.40 1.96 2.67
GraphCast
0.50 0.80 1.24 1.77 2.51
Climatology
2.54 2.54 2.54 2.54 2.54
1 3 5 7 10
Lead time [days]
−50

ARCHITECTURE DOESN’T MATTER

2m Temperature

RMSE [K]

Architecture	Model	1	3	5	7	10
	IFS HRES	0.54	0.90	1.35	1.89	2.57
	IFS ENS (mean)	0.58	0.88	1.22	1.58	1.99
Vision Transformer	Pangu-Weather (oper.)	0.62	0.97	1.42	1.97	2.64
Graph Neural Network	GraphCast (oper.)	0.47	0.76	1.18	1.71	2.41
Diffusion Model	GenCast (oper.) (mean)	0.46	0.75	1.11	1.50	1.93
Vision Transformer	Aurora (oper.)	0.47	0.77	1.16	1.60	2.19
Fourier Neural Operator	FGN (oper.) (mean)	0.41	0.67	1.01	1.40	1.86
	Pangu-Weather	0.56	0.91	1.40	1.96	2.67
	GraphCast	0.50	0.80	1.24	1.77	2.51
	Climatology	2.54	2.54	2.54	2.54	2.54

Lead time [days]

−50

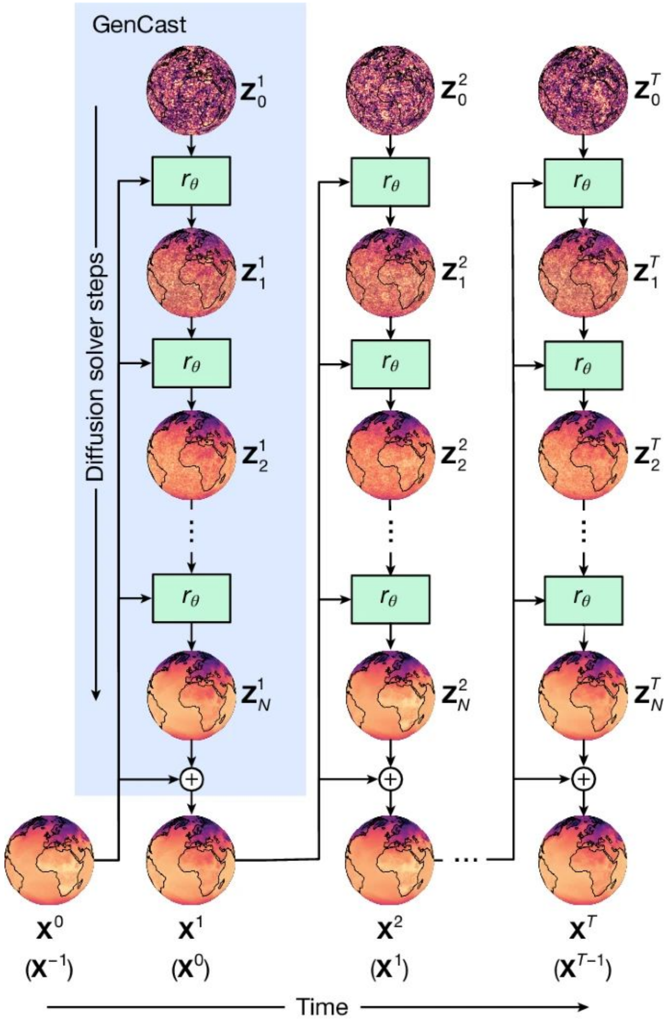
GenCast
Price et al., *Nature*, 2025

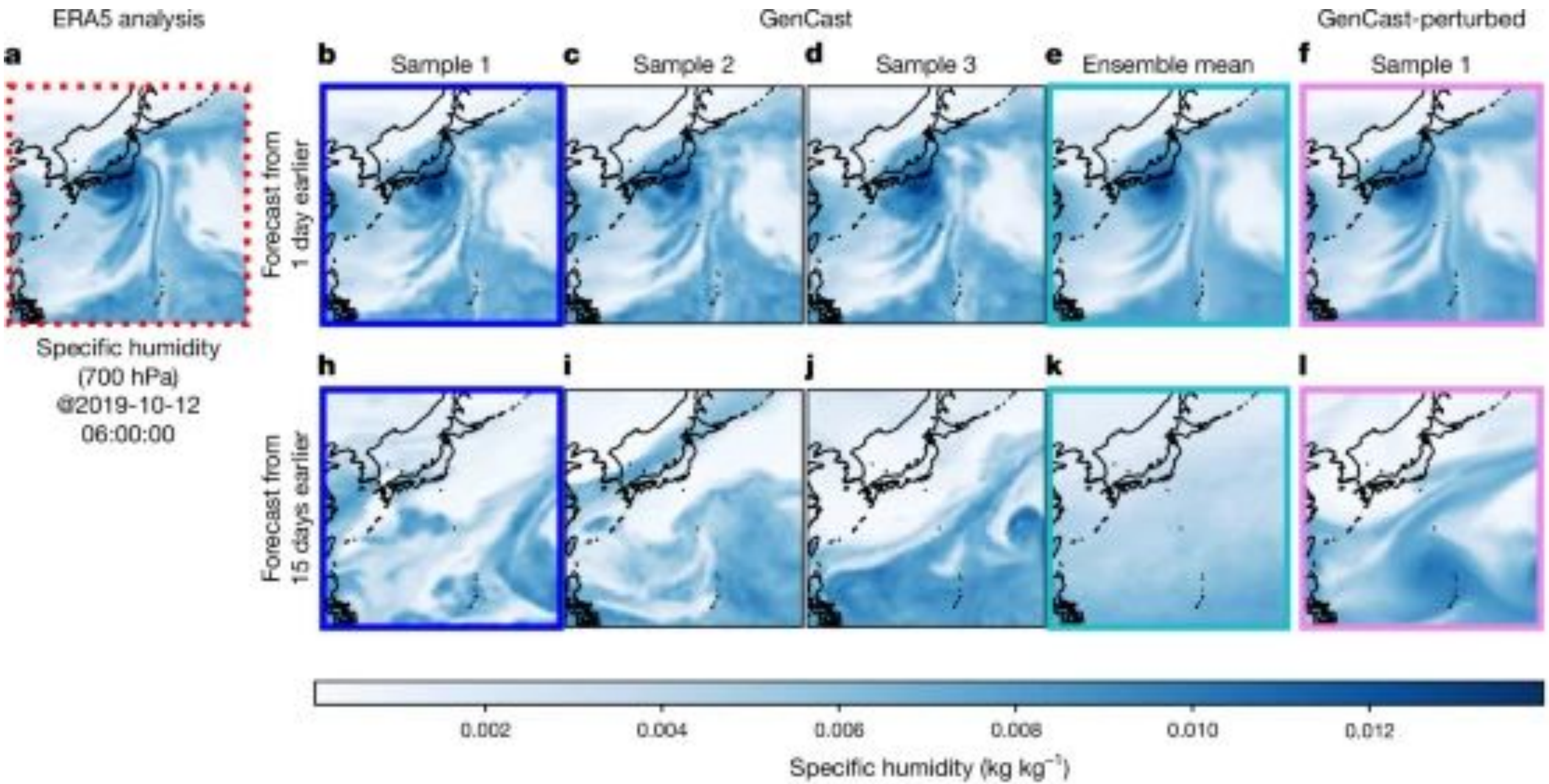
:: Diffusion based model

:: Denoiser based on GraphCast (GNN)

:: Predictions are well-calibrated

:: Predictions are skilful





ERA5 analysis
GenCast
GenCast-perturbed
a
b
Sample 1
c
Sample 2
d
Sample 3
e
Ensemble mean
f
Sample 1
Forecast from 1 day earlier
Specific humidity (700 hPa) @2019-10-12 06:00:00
h
i
j
k
l
Forecast from 15 days earlier
0.002
0.004
0.006
0.008
0.010
0.012
Specific humidity (kg kg⁻¹)

Outline

:: Rainfall nowcasting

:: Weather forecasting

:: Downscaling (aka super-resolution)

:: Bias correction of model outputs

:: Sub-grid parameterisations

:: Climate projections

Downscaling (i.e. super-resolution)

:: Based on a low resolution input (space or time), provide high resolution output

:: Useful for regional, localised forecasts

:: Useful for generating fine grained climate projections

:: Training data usually involves a hi-res regional weather product

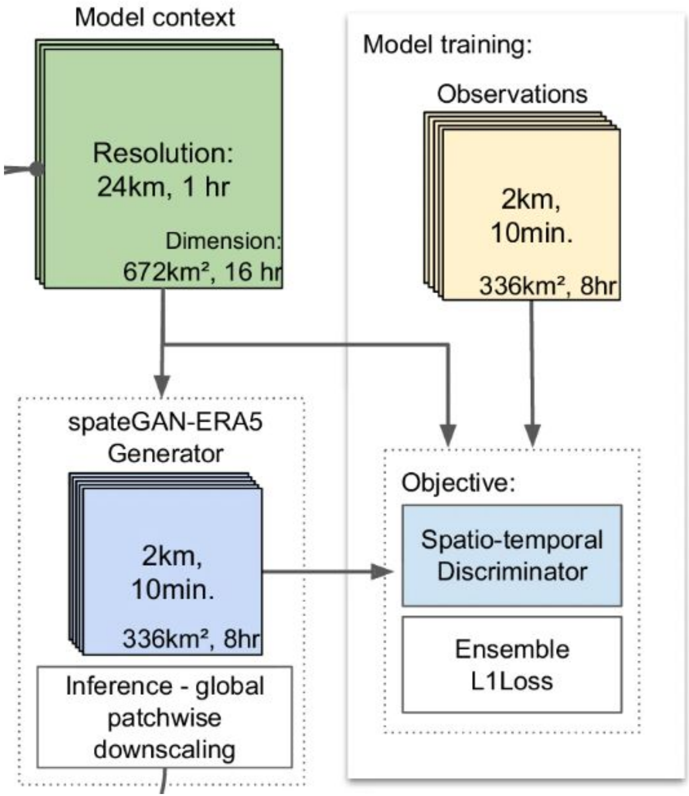
spateGAN-ERA5
Glawion et al., *npj Atmos. Sci*, 2025

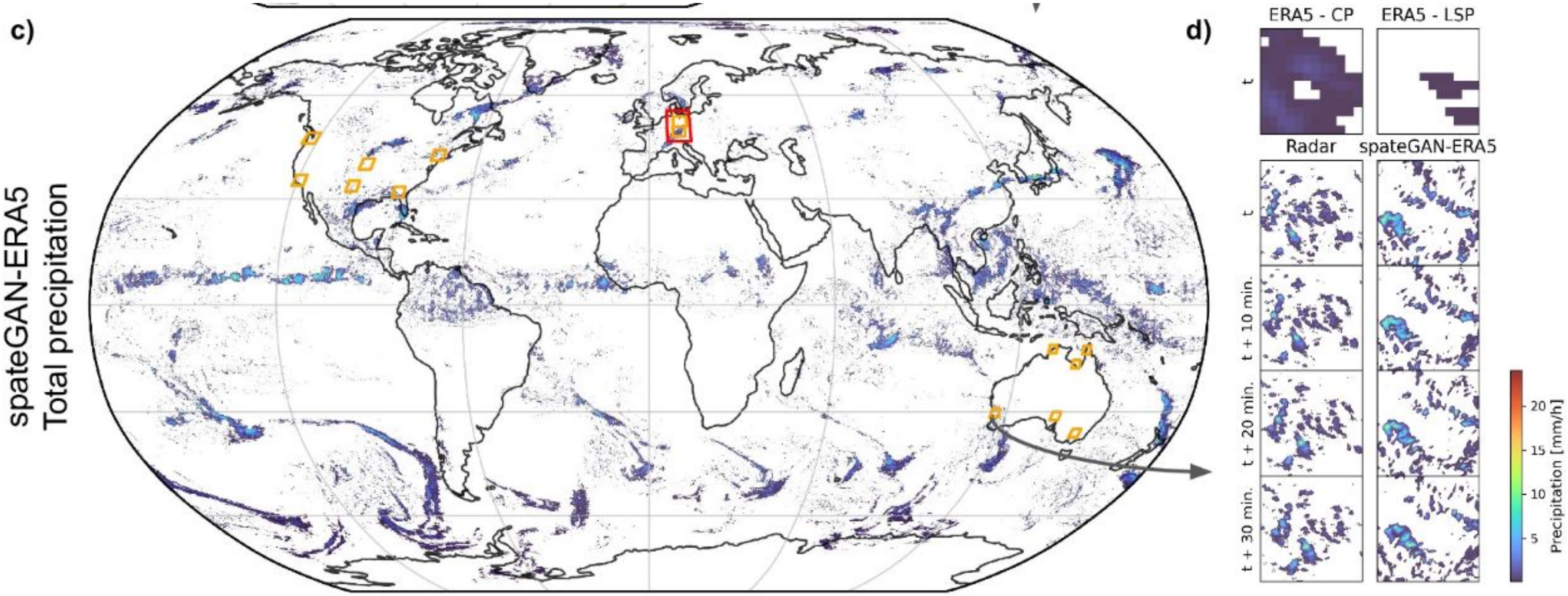
:: GAN-like loss

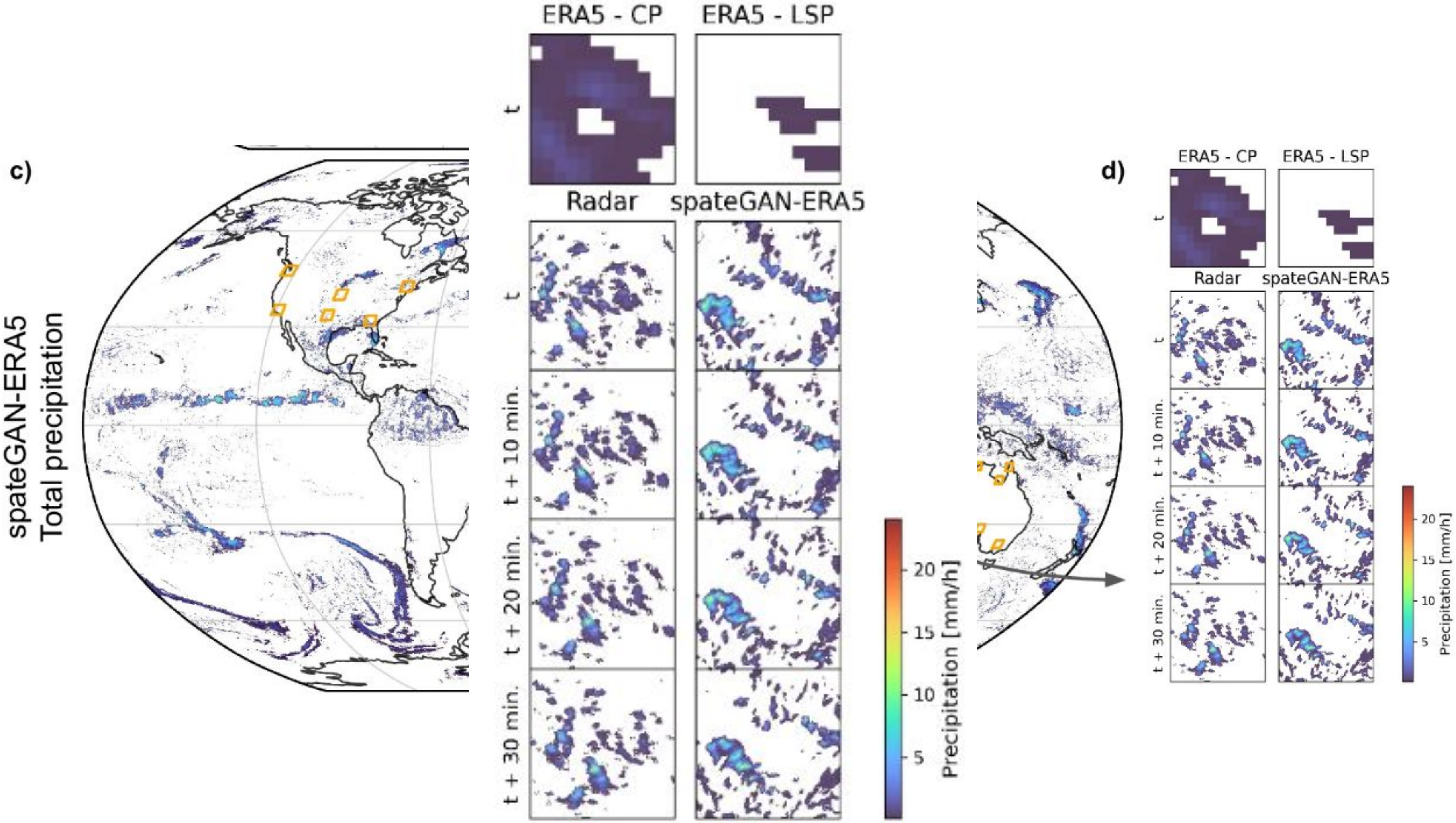
:: Trained from radar product over Germany

:: Tested over radar products elsewhere

:: Variables: Convective & large scale precipitation




c)
spateGAN-ERA5
Total precipitation
d)
ERA5 - CP
ERA5 - LSP
Radar
spateGAN-ERA5
t
t + 10 min.
t + 20 min.
t + 30 min.
Precipitation [mm/h]
5
10
15
20


c)
spateGAN-ERA5
Total precipitation
ERA5 - CP
ERA5 - LSP
t
Radar
spateGAN-ERA5
t
t + 10 min.
t + 20 min.
t + 30 min.
Precipitation [mm/h]
5
10
15
20
d)
ERA5 - CP
ERA5 - LSP
t
Radar
spateGAN-ERA5
t
t + 10 min.
t + 20 min.
t + 30 min.
Precipitation [mm/h]
5
10
15
20

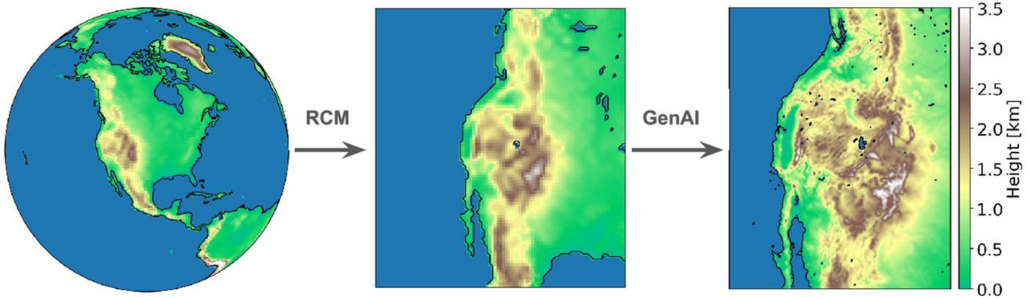
R2-D2
Lopez-Gomez et al., *PNAS*, 2025

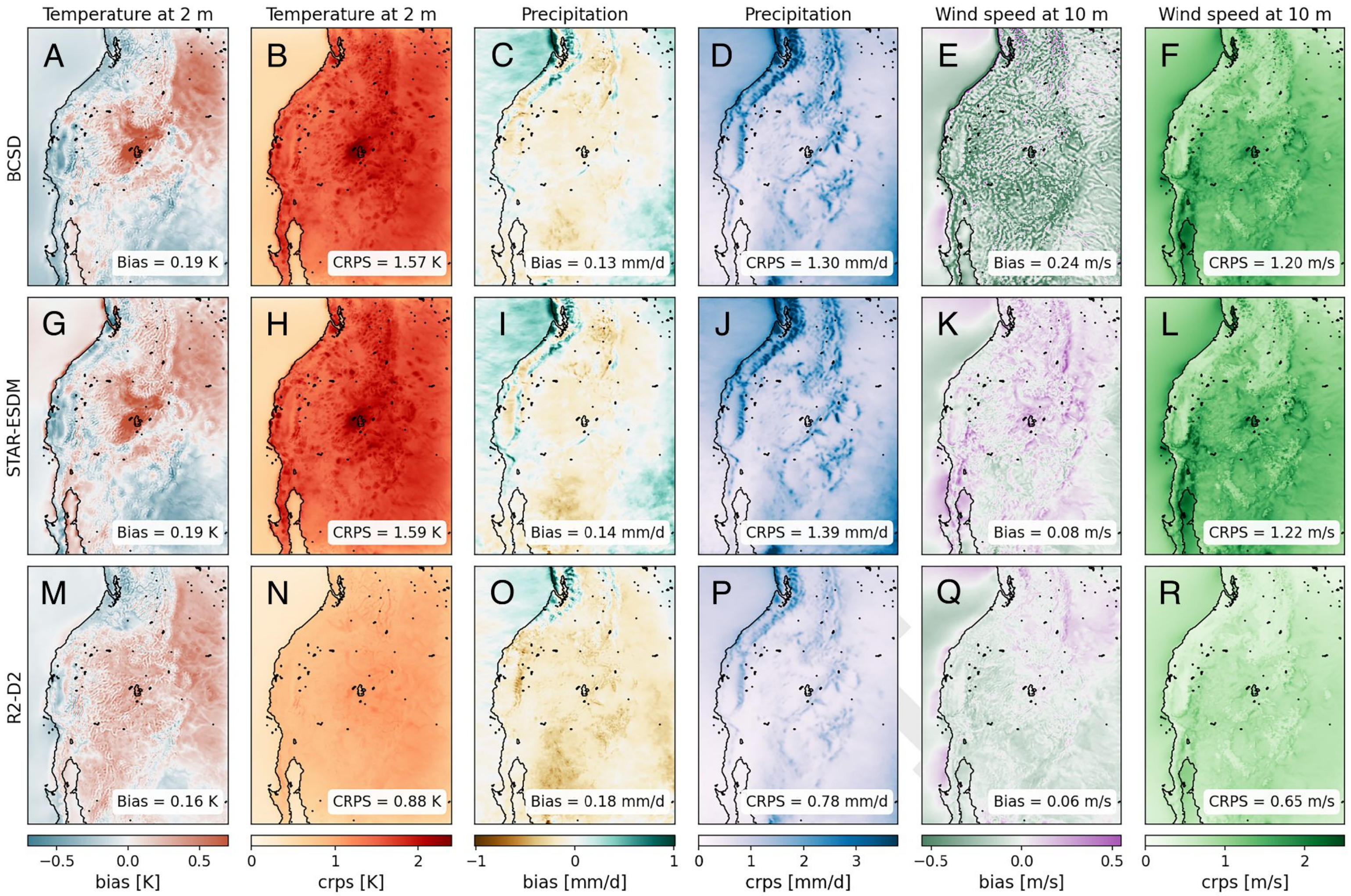
:: Uses diffusion on top of a regional model

:: Leverages both physics and deep learning

:: Training data - WUS-D3 (based on WRF)

:: Diffusion model learns differences between 45 km and 9 km datasets





Temperature at 2 m
Temperature at 2 m
Precipitation
Precipitation
Wind speed at 10 m
Wind speed at 10 m
BCSD
STAR-ESDM
R2-D2
A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
Bias = 0.19 K
CRPS = 1.57 K
Bias = 0.13 mm/d
CRPS = 1.30 mm/d
Bias = 0.24 m/s
CRPS = 1.20 m/s
Bias = 0.19 K
CRPS = 1.59 K
Bias = 0.14 mm/d
CRPS = 1.39 mm/d
Bias = 0.08 m/s
CRPS = 1.22 m/s
Bias = 0.16 K
CRPS = 0.88 K
Bias = 0.18 mm/d
CRPS = 0.78 mm/d
Bias = 0.06 m/s
CRPS = 0.65 m/s
−0.5
0.0
0.5
bias [K]
0
1
2
crps [K]
−1
0
1
bias [mm/d]
0
1
2
3
crps [mm/d]
−0.5
0.0
0.5
bias [m/s]
0
1
2
crps [m/s]

Outline

:: Rainfall nowcasting

:: Weather forecasting

:: Downscaling (aka super-resolution)

:: Bias correction of model outputs

:: Sub-grid parameterisations

:: Climate projections

Sub-grid parameterisation of global circulation models

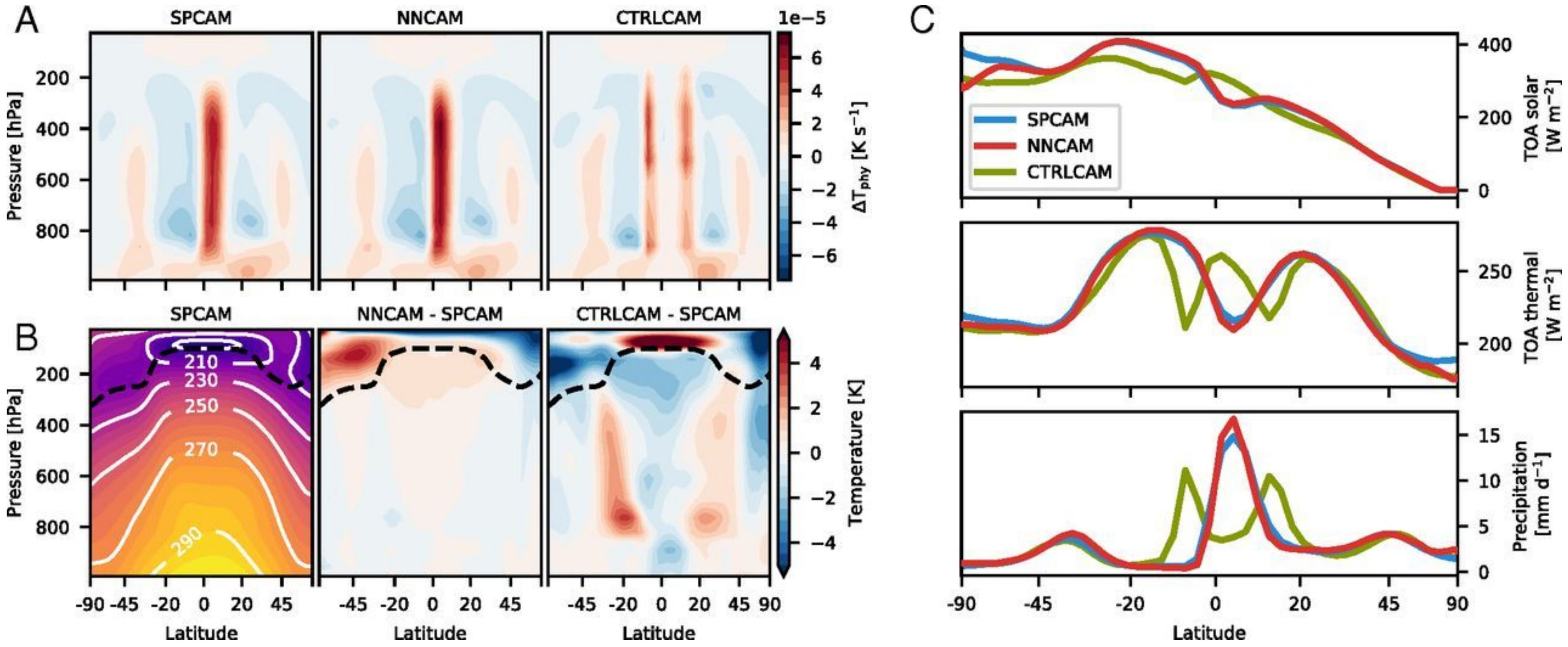
:: Unresolved processes below the grid scale influence the dynamics on the grid

:: Train an emulator on a high-resolution model (for only a few years of data)

:: Use the emulator to feed in the sub-grid information to a coarser climate model

:: Speeds up climate model runs

Rasp et al., *PNAS*, 2018



Outline

:: Rainfall nowcasting

:: Weather forecasting

:: Downscaling (aka super-resolution)

:: Bias correction of model outputs

:: Sub-grid parameterisations

:: Climate projections

How do we judge a model?

:: MAE, RMSE, … -> pixel wise error

:: CRPS -> error in modeling distributions

:: Power spectra -> error in modeling frequencies

:: Anomaly correlations -> error in spatial patterns

:: Conservation laws, physics constraints, …

:: etc …

Take Home Message

:: Deep learning can contribute in multiple ways

:: Training objective more important than architecture

:: Architecture often influenced by compute constraints

:: How we evaluate models matter

:: Including physics requires creative innovations

:: Success depends on quality/amount of training data