DEEP LEARNING FOR SHORT-TERM TEMPERATURE FORECASTS WITH VIDEO PREDICTION METHODS

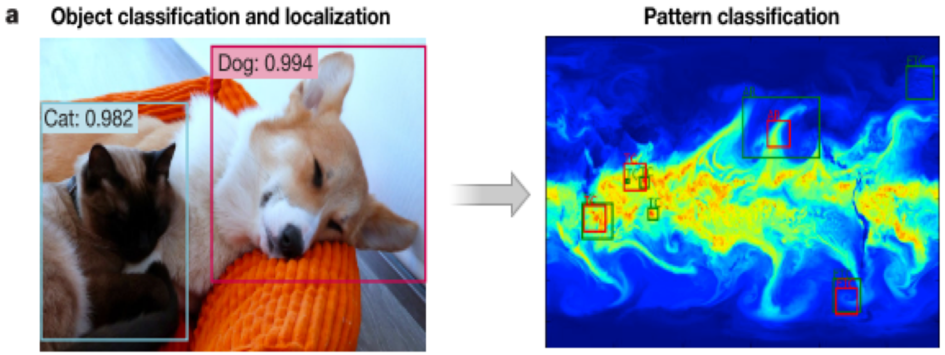
MACHINE LEARNING FOR EARTH SYSTEM MODELLING AND ANALYTICS WORKSHOP 2021-05-03

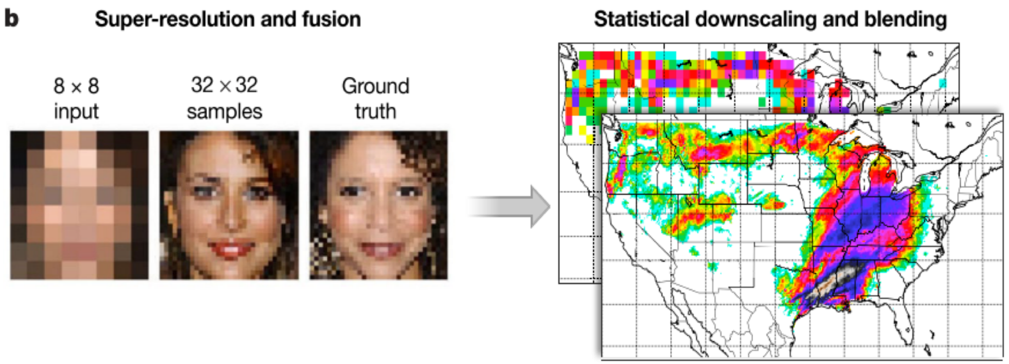
BING GONG, MICHAEL LANGGUTH, AMIRPASHA MOZAFARRI, YAN JI, SCARLET STADTLER, KARIM MACHE, MARTIN SCHULTZ

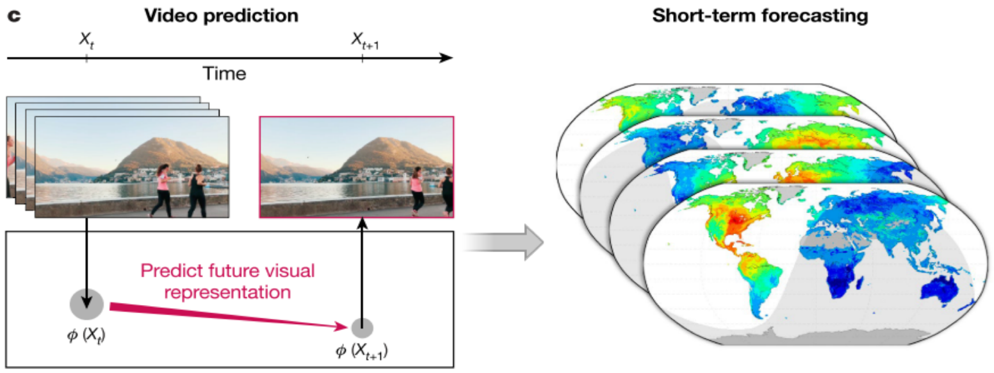
OUTLINE

- Motivation
- Experiments design
- Results
- Deep learning workflow toolkit
- Conclusions

MOTIVATION







- ➢ Video prediction → **New data-driven approach for weather forecasting**

Reichstein, Markus et al. 2019. “Deep Learning and Process Understanding for Data-Driven Earth System Science.” *Nature* 566(7743): 195–204. http://dx.doi.org/10.1038/s41586-019-0912-1.

MOTIVATION

- Video prediction → **New data-driven approach for weather forecasting**

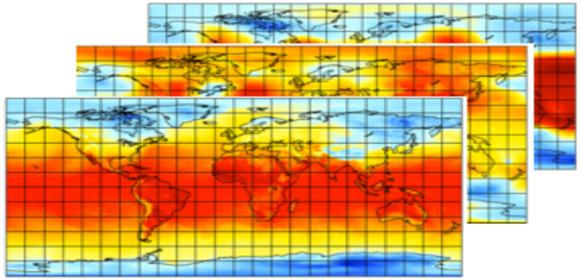
- Big data in weather forecasting and advanced Deep Learning architecture
  → **High-Performance Computing (HPC) & Parallelisation**

- Reproducibility → **End-to-end workflow toolkit**

EXPERIMENT SETTING

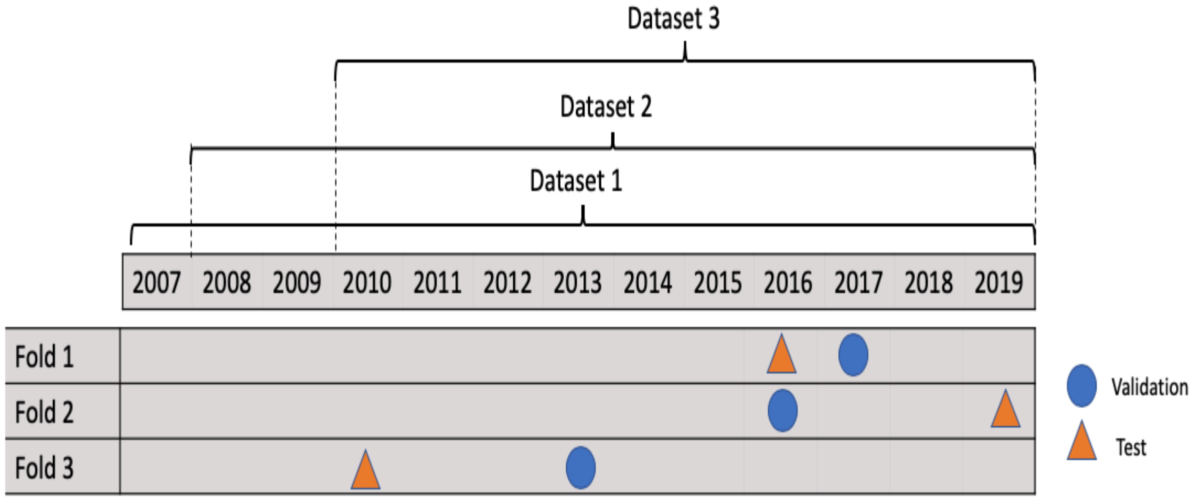
Dataset preparation

ECMWF

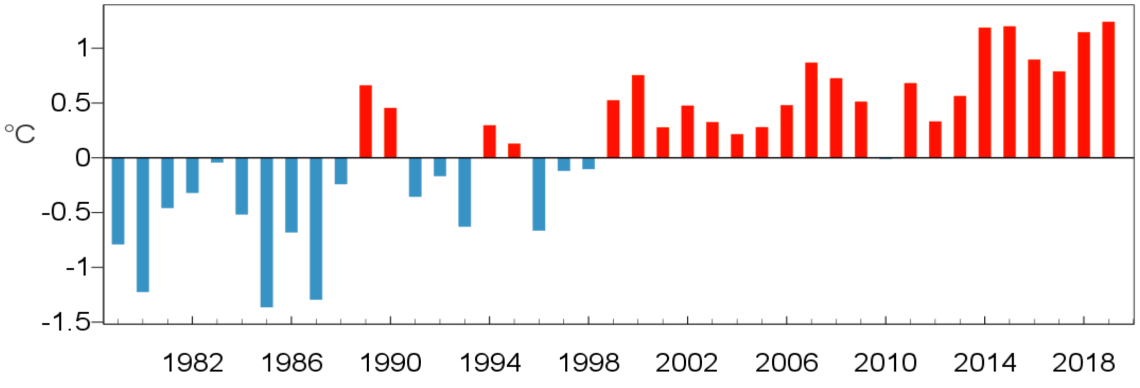


ERA5 reanalysis data

- **Grid points/Pixel**: 601 * 1200
- **Time resolution:** Hourly
- **Inputs:** 12 hours
- **Outputs:** 12 hours
- **Region:** Europe (56 * 92 grids)
- **Input variables:** 2m Temperature, 850 hPa temperature, cloud cover



Europe annual temperature anomalies 1979-2019



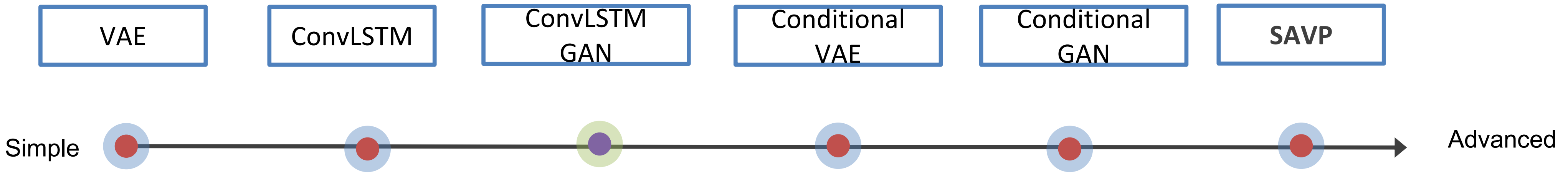
Data source: ERA5 Reference period: 1981-2010

Copernicus Climate Change Service
European State of the Climate | 2019

https://climate.copernicus.eu/ESOTC/2019/european-temperature

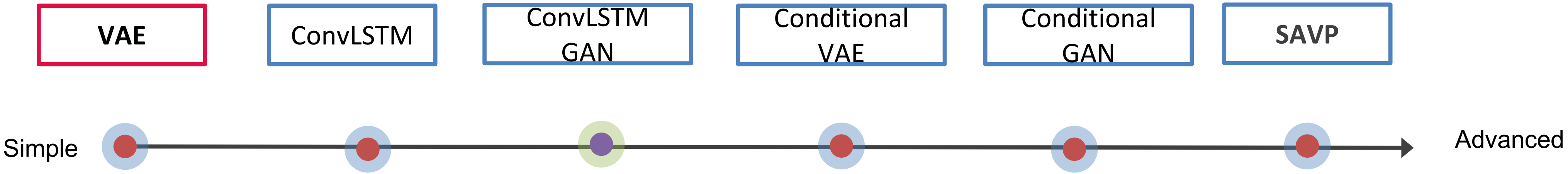
EXPERIMENT SETTING

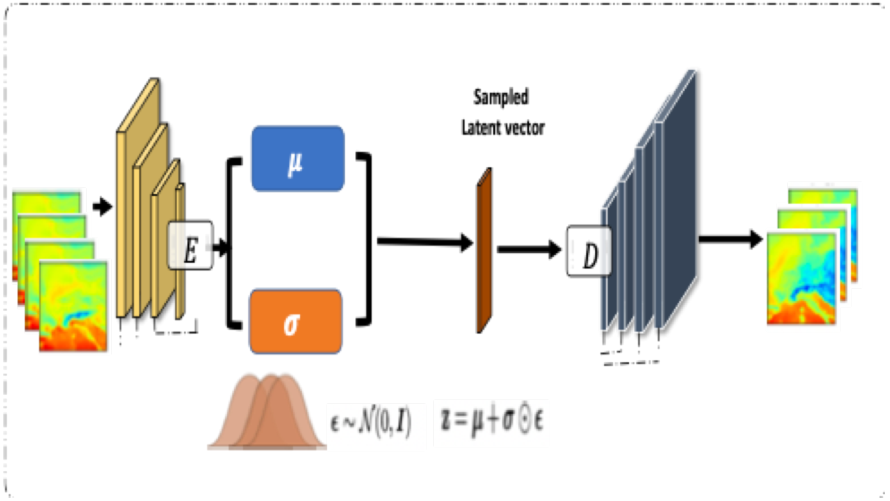
Models



EXPERIMENT SETTING

Models





Variational auto-encoder (VAE)

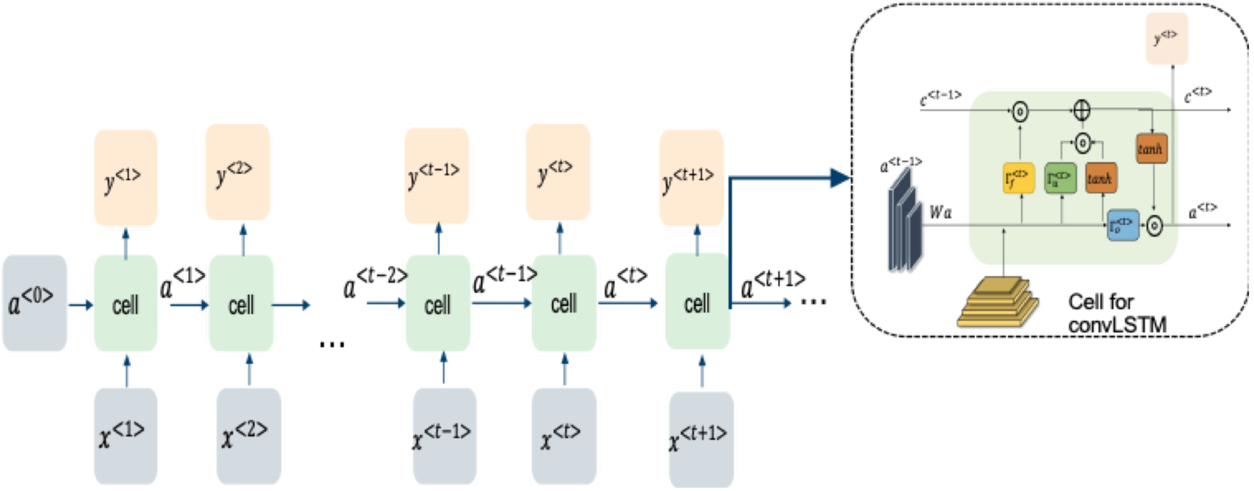
EXPERIMENT SETTING

Models

VAE | **ConvLSTM** | ConvLSTM GAN | Conditional VAE | Conditional GAN | SAVP

Simple

Advanced



Convolutional Long Short Term Memory (ConvLSTM) performs convolution operator in each LSTM cell.

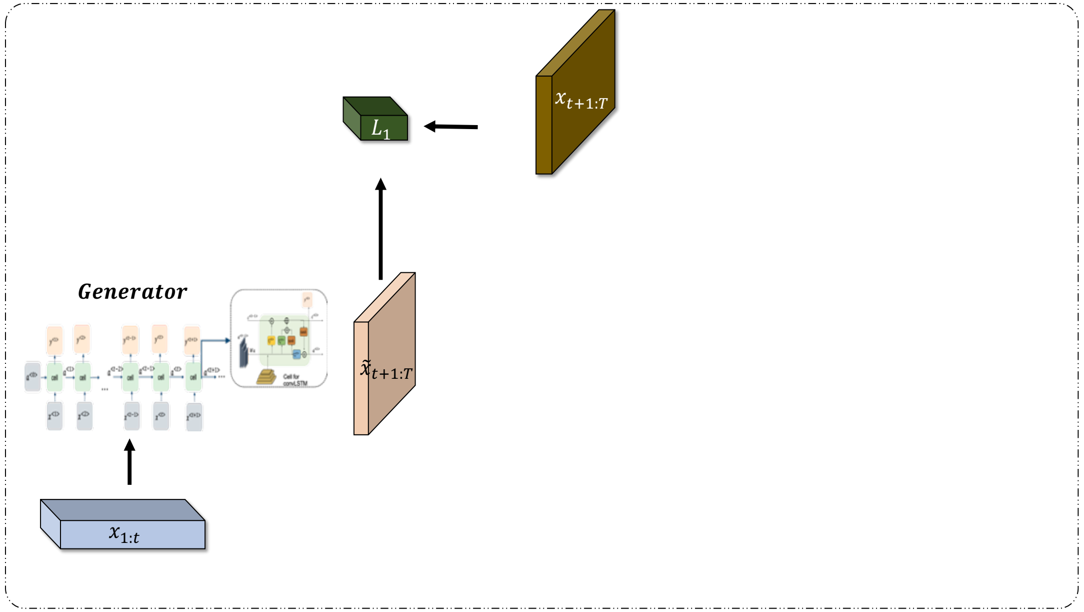
Shi, Xingjian et al. 2015. “Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting.” In Advances in {{Neural Information Processing Systems}} 28, eds. C Cortes et al. Curran Associates, Inc., 802–10.

EXPERIMENT SETTING

Models

VAE | ConvLSTM | **ConvLSTM GAN** | Conditional VAE | Conditional GAN | SAVP

Simple → Advanced

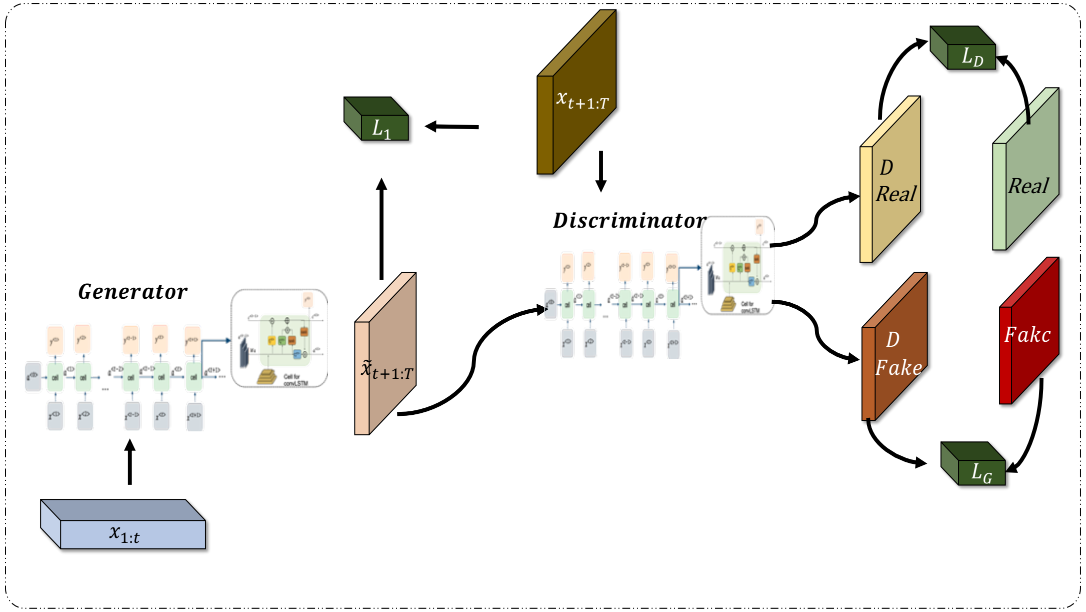


EXPERIMENT SETTING

Models

VAE | ConvLSTM | **ConvLSTM GAN** | Conditional VAE | Conditional GAN | SAVP

Simple → Advanced



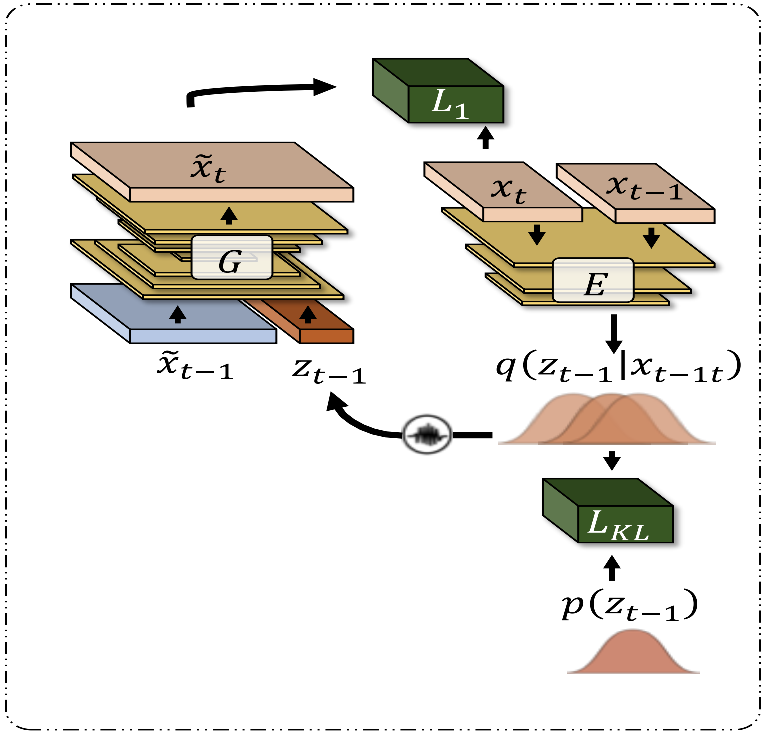
JÜLICH Forschungszentrum

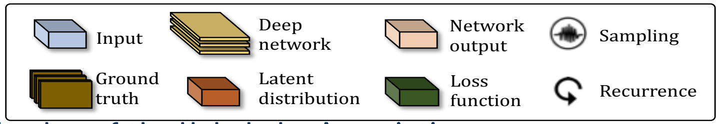
EXPERIMENT SETTING

Models

VAE | ConvLSTM | ConvLSTM GAN | **Conditional VAE** | Conditional GAN | **SAVP**

Simple → Advanced





$$\lambda_1 \ \mathcal{L}_1(G,E) + \lambda_{kl} \ \mathcal{L}_{kl}(E)$$

Lee AX, Zhang R, Ebert F, Abbeel P, Finn C, Levine S. Stochastic adversarial video prediction. arXiv preprint arXiv:1804.01523. 2018 Apr 4.

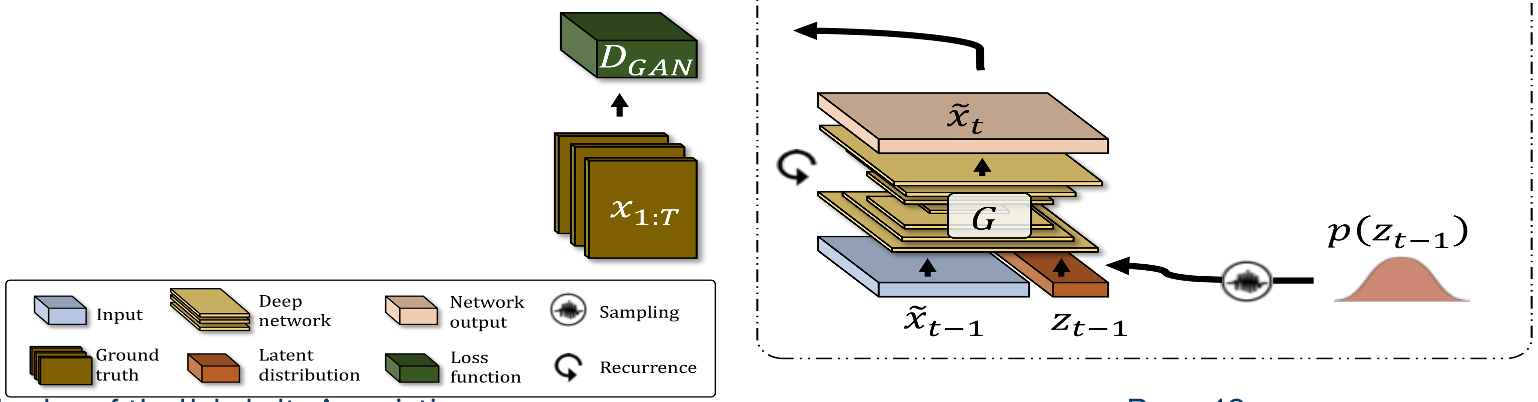
EXPERIMENT SETTING

Models

VAE | ConvLSTM | ConvLSTM GAN | Conditional VAE | **Conditional GAN** | SAVP

Simple → Advanced

$$G^* = \arg\min_{G}\max_{D}\mathcal{L}_{GAN}(E,D)$$



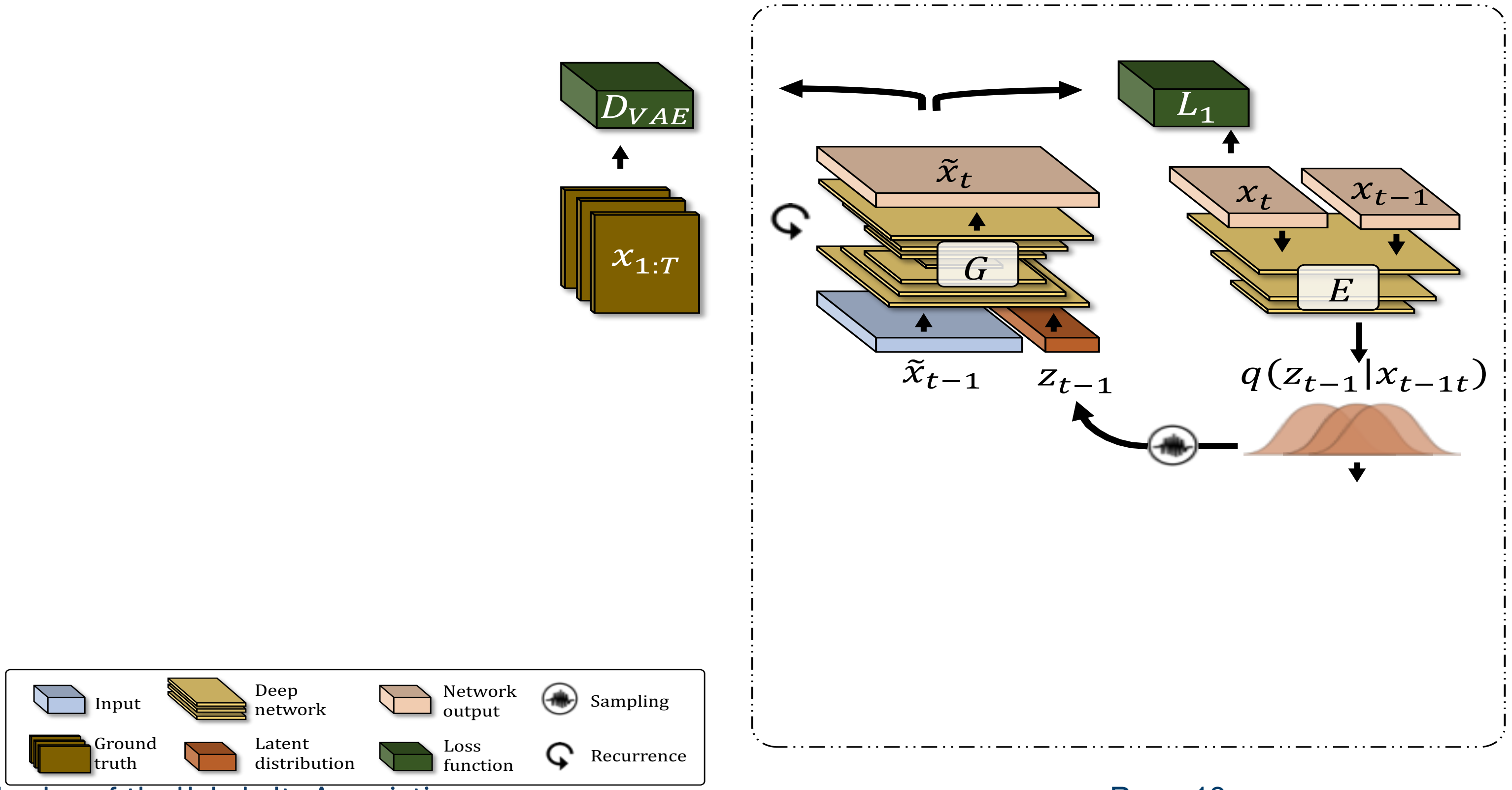
Lee AX, Zhang R, Ebert F, Abbeel P, Finn C, Levine S. Stochastic adversarial video prediction. arXiv preprint arXiv:1804.01523. 2018 Apr 4.

EXPERIMENT SETTING

Models

VAE	ConvLSTM	ConvLSTM GAN	Conditional VAE	**Conditional GAN**	SAVP

Simple → Advanced

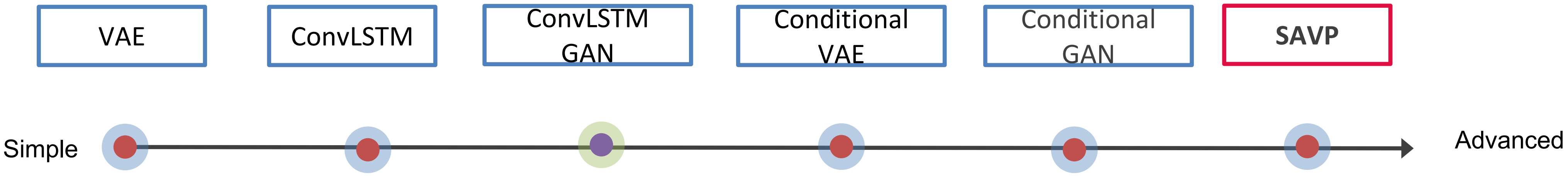


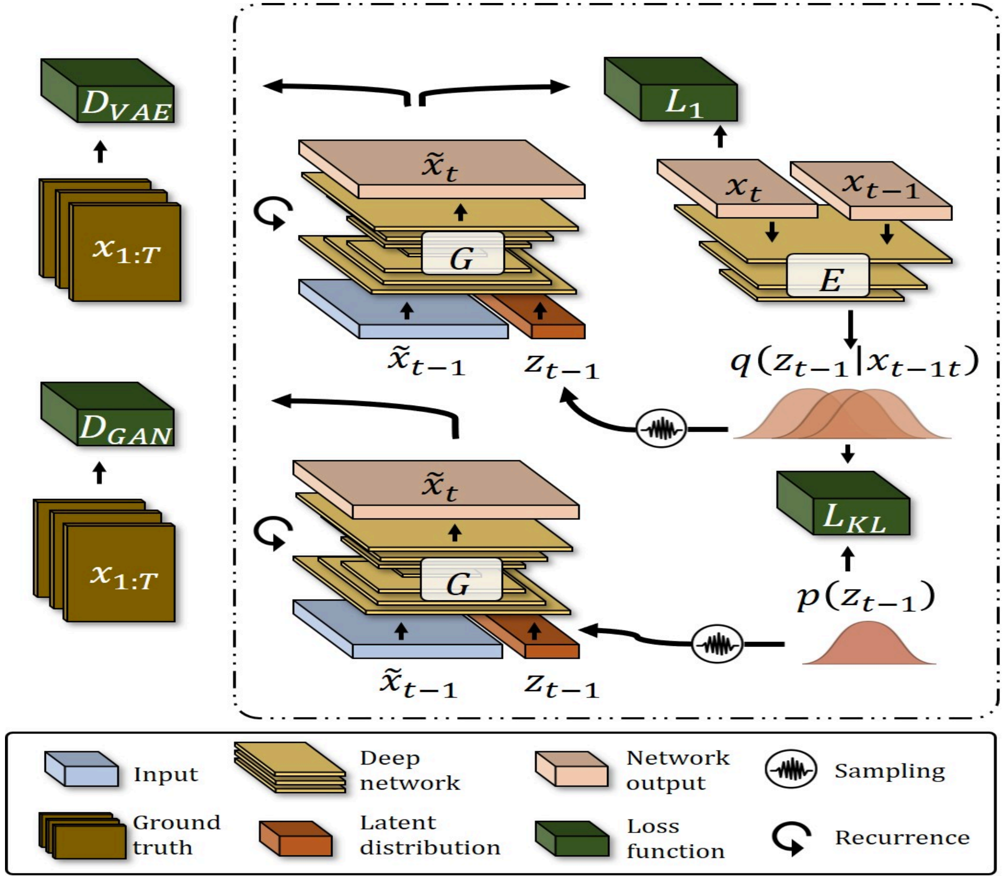
$$G^* = \arg\min_{G}\max_{D} \mathcal{L}_{GAN}^{VAE}(E, D)$$

Lee AX, Zhang R, Ebert F, Abbeel P, Finn C, Levine S. Stochastic adversarial video prediction. arXiv preprint arXiv:1804.01523. 2018 Apr 4.

EXPERIMENT SETTING

Models





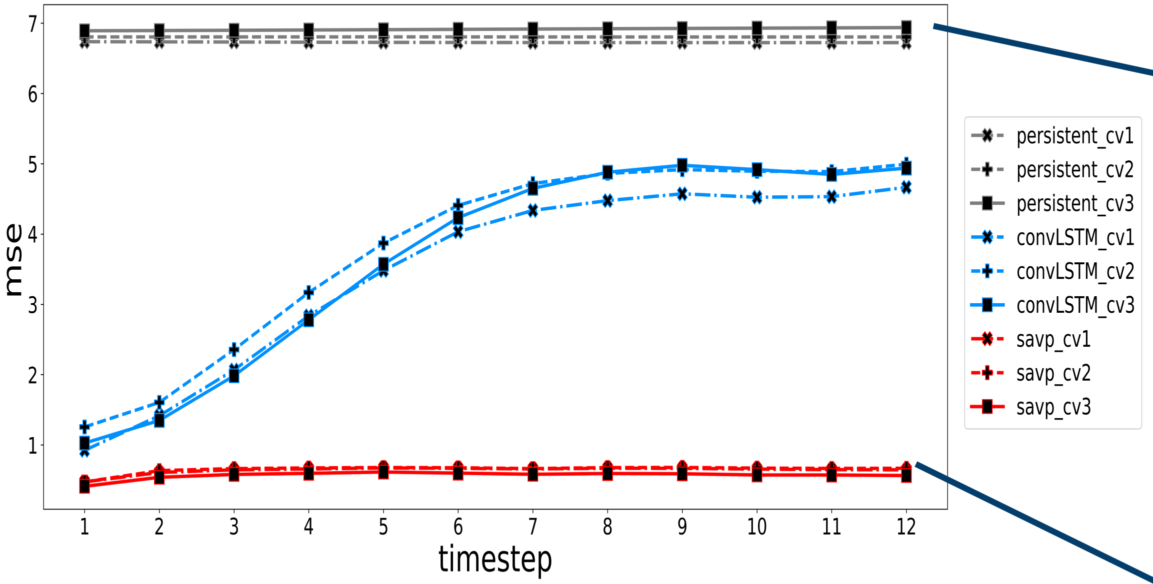
$$G^* = \arg\min_{D}\max_{G} \lambda_1 \mathcal{L}_1(G,E) + \lambda_{kl}\,\mathcal{L}_{kl}(E) + \mathcal{L}_{GAN}(E,D) + \mathcal{L}_{GAN}^{VAE}(E,D)$$

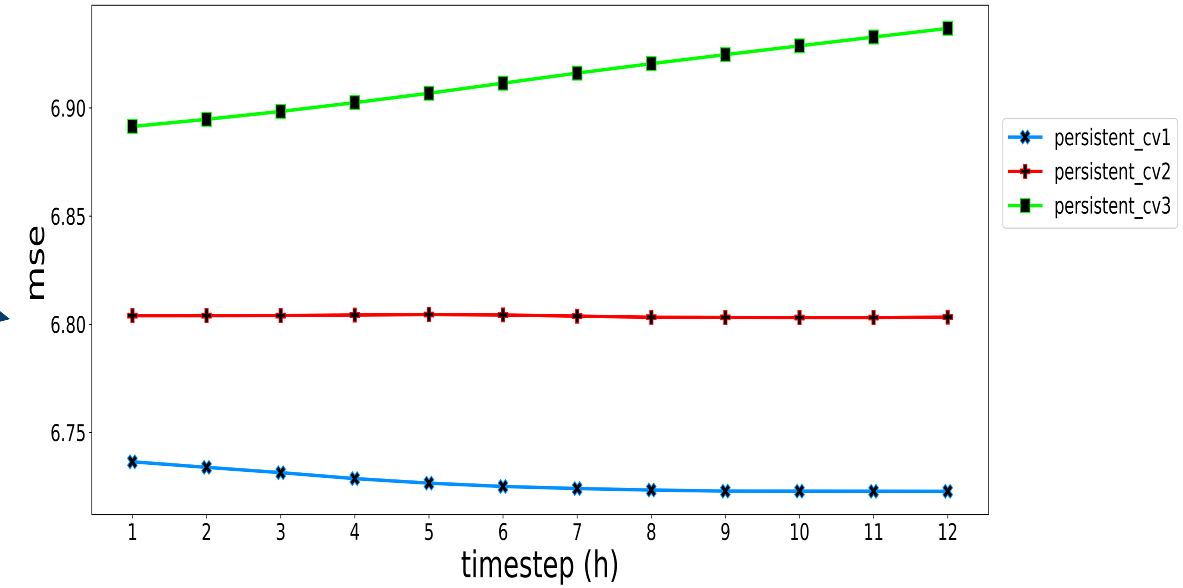
Stochastic adversarial video prediction (SAVP)

Lee AX, Zhang R, Ebert F, Abbeel P, Finn C, Levine S. Stochastic adversarial video prediction. arXiv preprint arXiv:1804.01523. 2018 Apr 4.

RESULTS

Persistent, ConvLSTM, SAVP





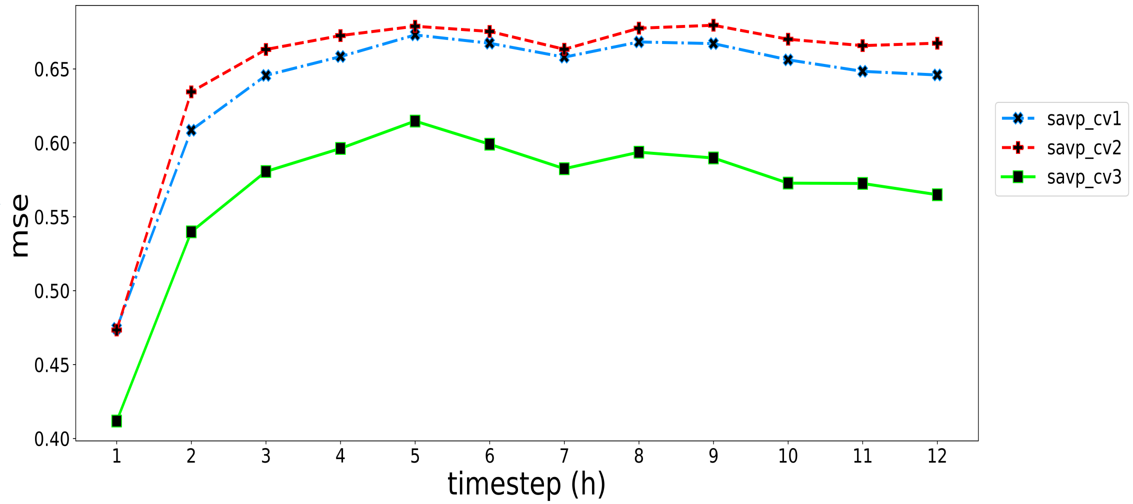


Fig. Mean square error (MSE) for persistent analysis and convLSTM, and stochastic adversarial video prediction architectures on the there-fold test data from Dataset 1

RESULTS

ConvLSTM and SAVP

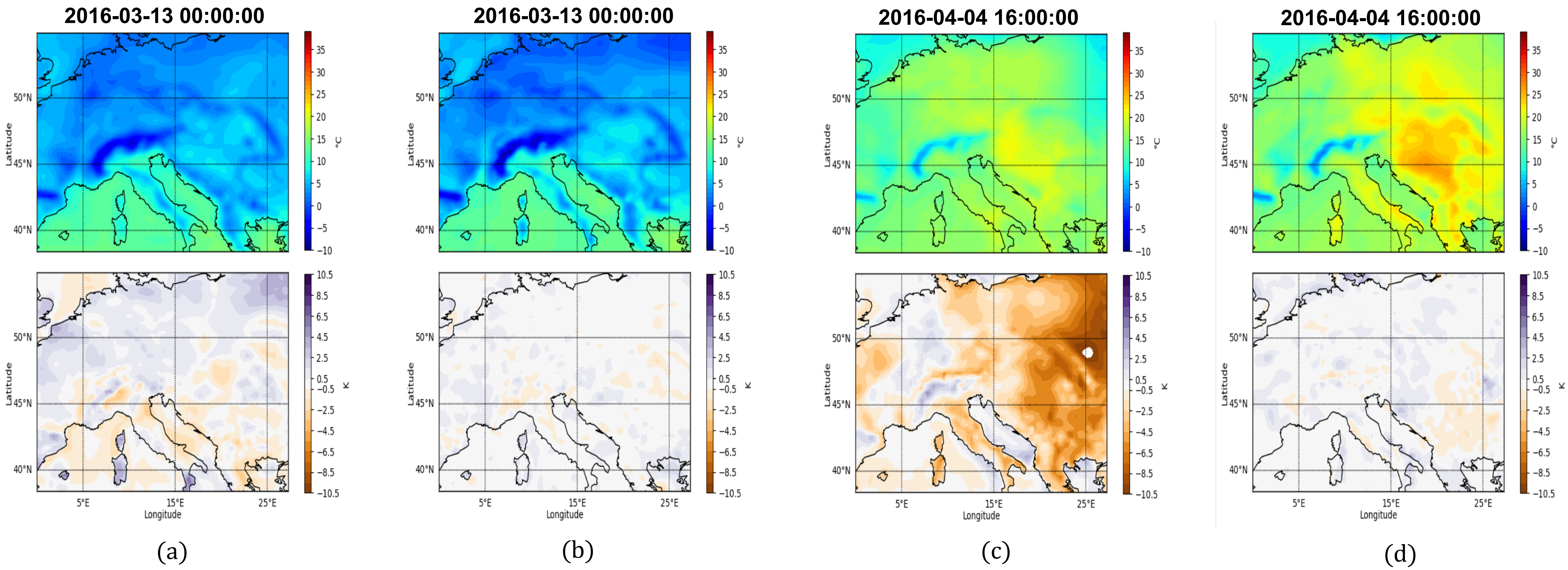


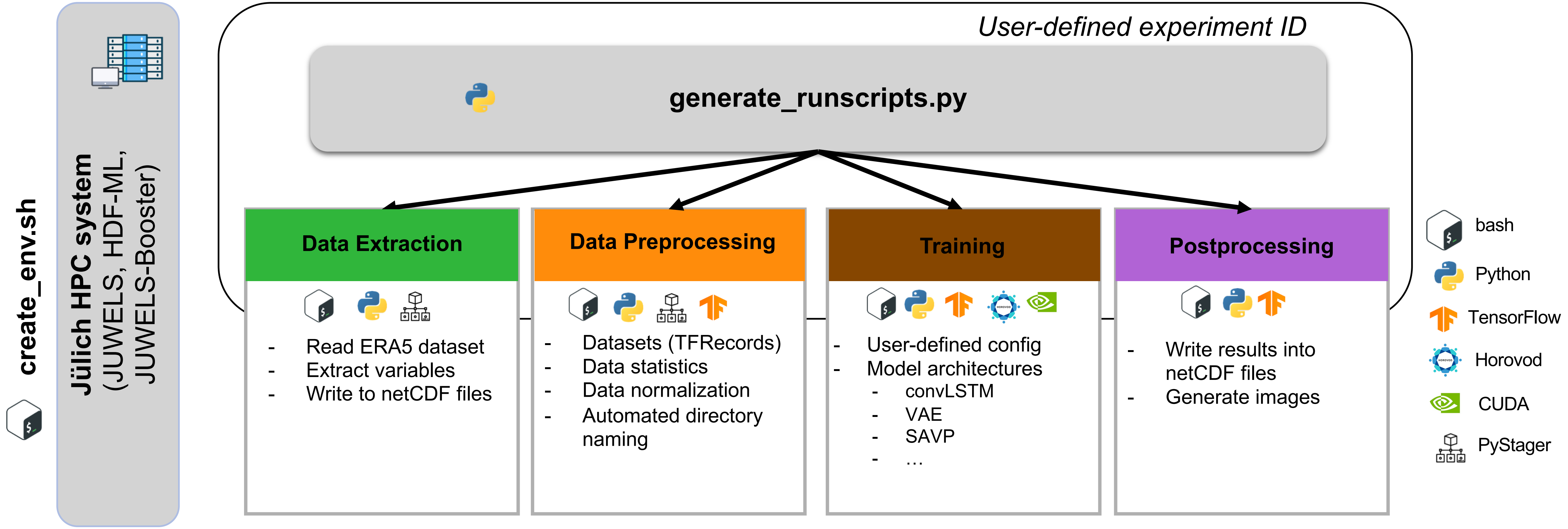
Fig. 2m temperature prediction with the 12th -hour lead time (1st row) and the temperature differences between forecast and ground truth (2nd row)

(a) The **best** 2m temperature (°C) prediction (2016-03-13 00:00:00) by **ConvLSTM**; (c) The **worst** 2m temperature (°C) prediction (2016-04-04 16:00:00) by **ConvLSTM**

(b) The 2m temperature (°C) prediction (2016-03-13 00:00:00) by **SAVP**; (d) The 2m temperature (°C) prediction (2016-04-04 16:00:00)) by **SAVP**

STRUCTURE OF AMBS WORKFLOW

Atmospheric Machine learning Benchmarking System (AMBS)



SCALABILITY

Scale deep learning on JUWELS Booster

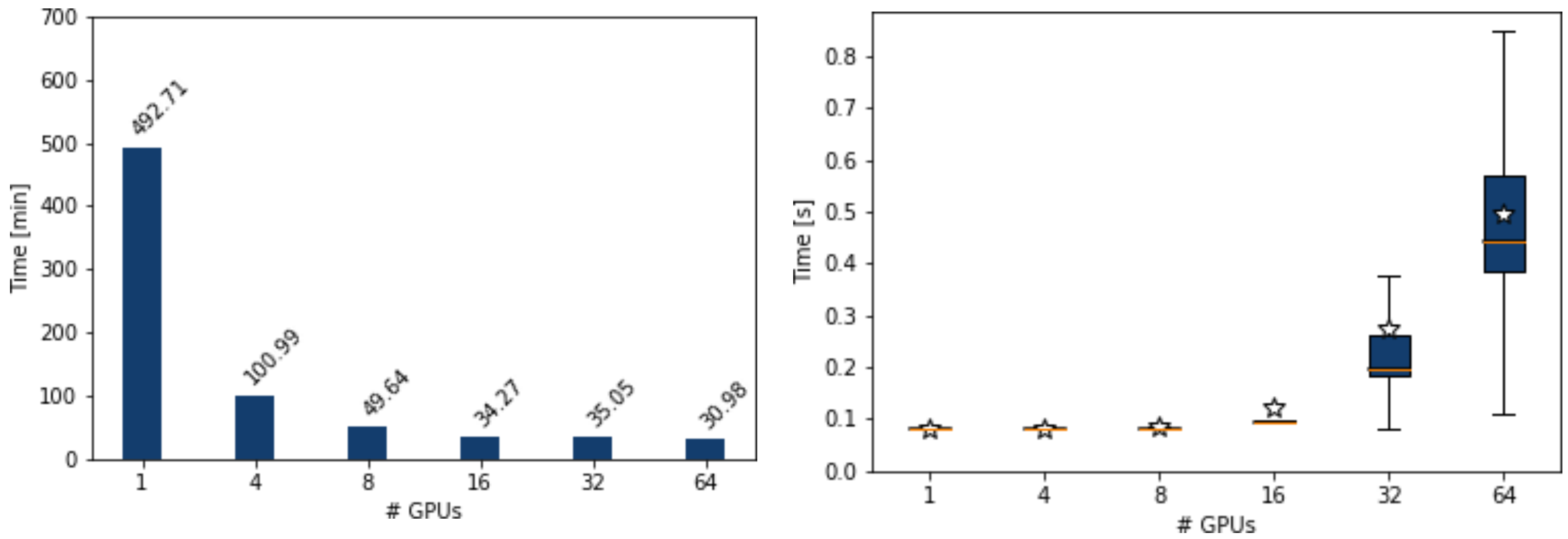


Fig. Total training time on ConvLSTM in minutes (left) and box whisker plot of iteration time in seconds (right). The star in the box whisker plot denotes the averaged iteration time, while its median is highlighted by an orange line.

CONCLUSIONS

Key results:

1. ConvLSTM performs better than persistent and VAE.
2. Advanced conditional VAE, conditional GAN and SAVP improve the forecasting performance significantly compared to the ConvLSTM.

Current work:

1. Block bootstrap method to estimate the significance of the test statistic.
2. Sensitivity analysis:
   - Test sensitivity to the size of training data
   - Test sensitivity to the input variable
   - Test input regions

ACKNOWLEDGEMENT

IntelliAQ

European Research Council

Established by the European Commission

Funding is provided through ERC Advanced grant ERC-2017-ADG #787576 by Martin Schultz

DeepRain

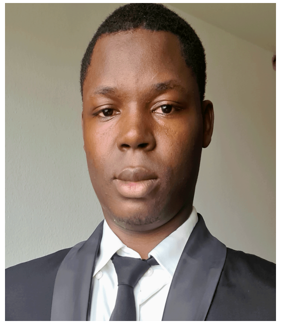
DeepRain is funded by the Bundesministerium fuer Bildung und Forschung (BMBF),
under grant agreement 01 IS18047A

Our team

Bing Gong

Michael Langguth

Amirpasha Mozafarri



Karim Mache

Yan Ji

Scarlet Stadtler

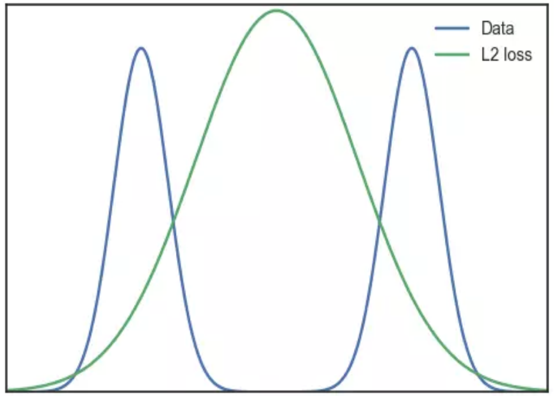
Martin Schultz (Supervisor)

THANK YOU

DEEP LEARNING ARCHITECTURES

Downsides of state-of-the-art architectures for video prediction

Problem 1: Unrealistic images



$$L2 = \sum_{i=1}^{n} (y - \hat{y})^2 \quad L1 = \sum_{i=1}^{n} |y - \hat{y}|$$

Adversarial loss

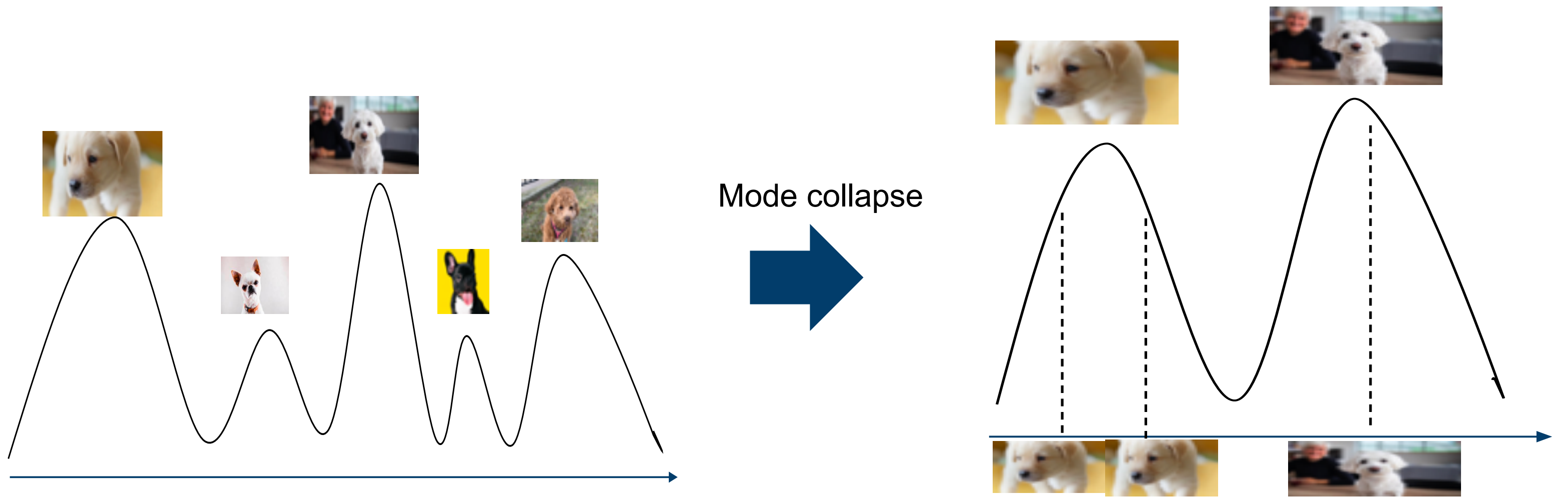
$$\mathcal{L}_{GAN}(G, D) = \mathrm{E}_{x_{1:T}}[log D(x_{1:T})] + \mathrm{E}_{x_{1:T, z \sim p(z_t)|_{t=0}^{T-1}}}[\log(1 - D(G(x_0, z_{0:T-1})))]$$

- For RNN and VAE, if each pixel follows multi-modal distribution, then using L2 Norm or L1 Norm will average the loss function. It means to produce the mean of image of all possible futures, as the global optimum. This will produce unrealistic prediction images.

DEEP LEARNING ARCHITECTURES

Downsides of state-of-the-art architectures for video prediction

Problem 2: Mode collapse



- GAN based on adversarial learning can produce more realistic images than RNN and VAE with L2/L1 loss. On the downside, it potentially only produces a limited diversity of output (e.g. limited number of dog breeds), which is commonly called ***mode collapse***.