DEEP LEARNING FOR WEATHER FORECASTS

SECOND SYMPOSIUM ON AISIS, OCTOBER 11-15, 2021

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

EARTH SYSTEM DATA EXPLOTATION (ESDE) GROUP, JUELICH SUPERCOMPUTER CENTER

INTRODUCTION OF PRESENTATOR

Yan Ji

Ph.D Student of Nanjing University of Information Science & Technology, China
Supervisor: Prof. Dr. Xiefei Zhi

Now in the ESDE group led by Dr. Martin Schultz of JSC, FZJ

Research filed:

- Applications of deep learning in weather forecasting;
- Downscaling of numerical weather prediction models;
- Ensemble weather forecasting;

E-mail: y.ji@fz-juelich.de

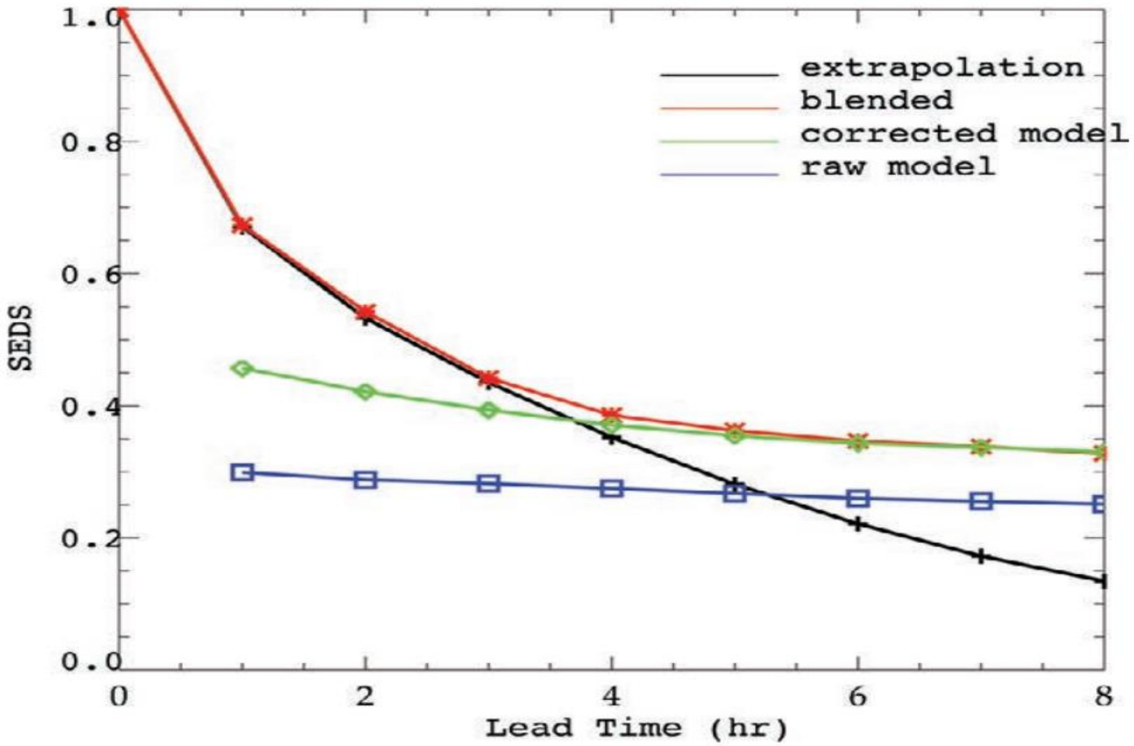
OUTLINE

- Motivation
- Deep learning architectures
- Verification methods
- Applications
  - 2m Temperature forecasts
  - Precipitation nowcasting
- Conclusion

MOTIVATION

- **Requirements** of convection-permitting NWP
- **Advantages** of statistical models in short-term forecasting [2]

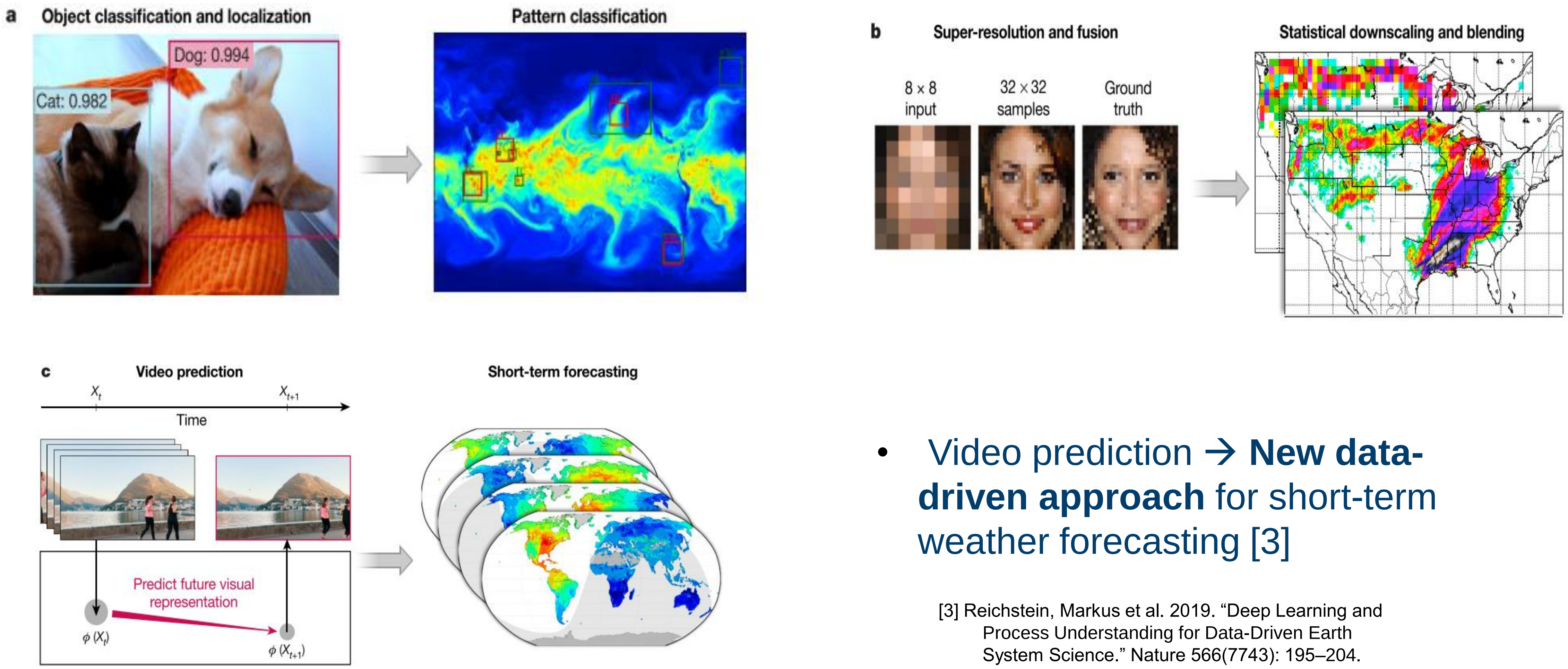
Symmetric extreme dependency score
of vertical integrated liquid (VIL)



[2] Sun, Juanzhen, et al. "Use of NWP for nowcasting convective precipitation: Recent progress and challenges." Bulletin of the American Meteorological Society 95.3 (2014): 409-426.

JÜLICH
Forschungszentrum

MOTIVATION

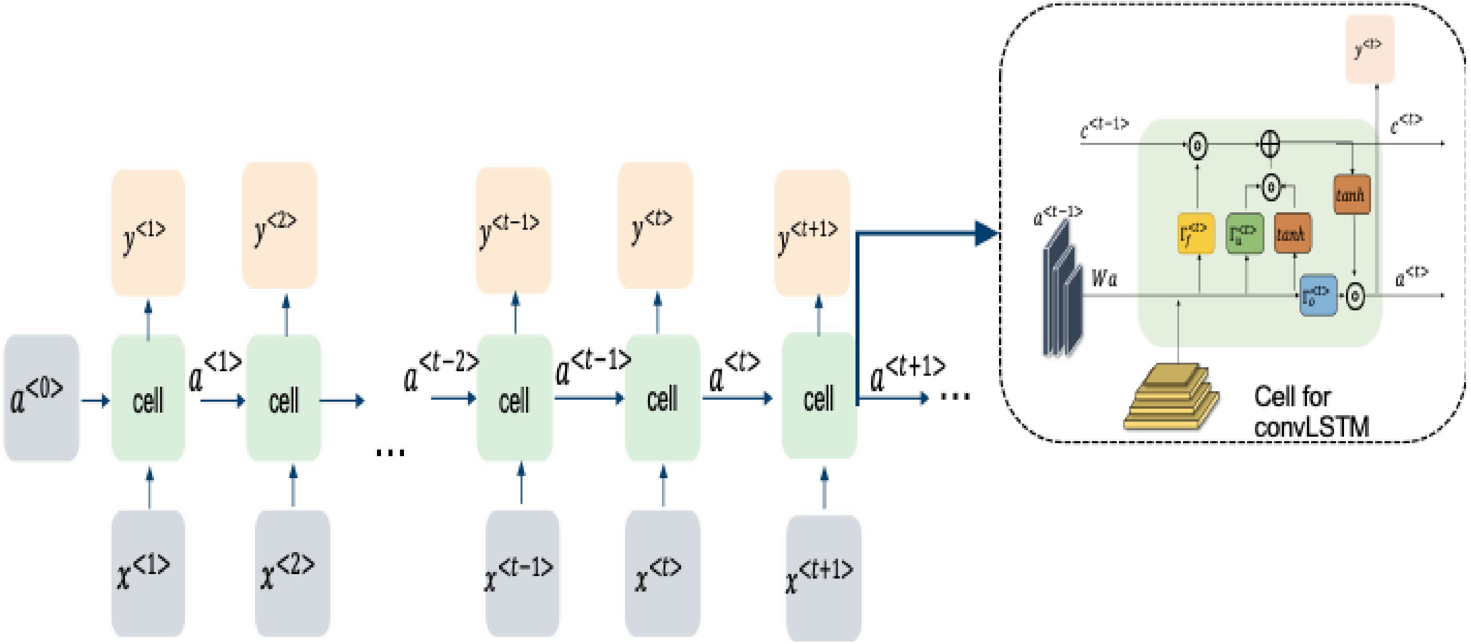


- Video prediction → **New data-driven approach** for short-term weather forecasting [3]

[3] Reichstein, Markus et al. 2019. "Deep Learning and Process Understanding for Data-Driven Earth System Science." Nature 566(7743): 195–204.

DEEP LEARNING ARCHITECTURES

Models - Convolutional LSTM (ConvLSTM)



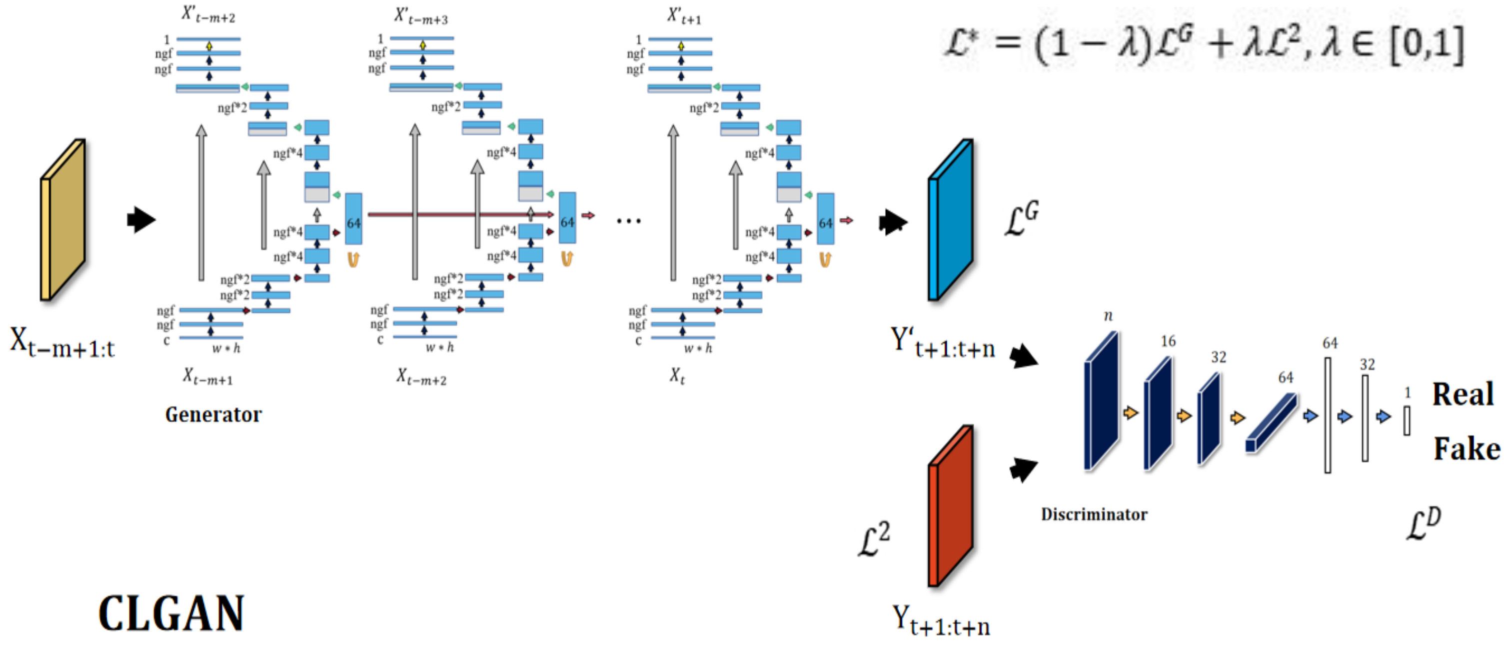
Convolutional Long Short Term Memory (ConvLSTM) [4] performs convolution operator in each LSTM cell.

[4] Xingjian, S. H. I., et al. "Convolutional LSTM network: A machine learning approach for precipitation nowcasting." Advances in neural information processing systems. 2015.

JÜLICH Forschungszentrum

DEEP LEARNING ARCHITECTURES

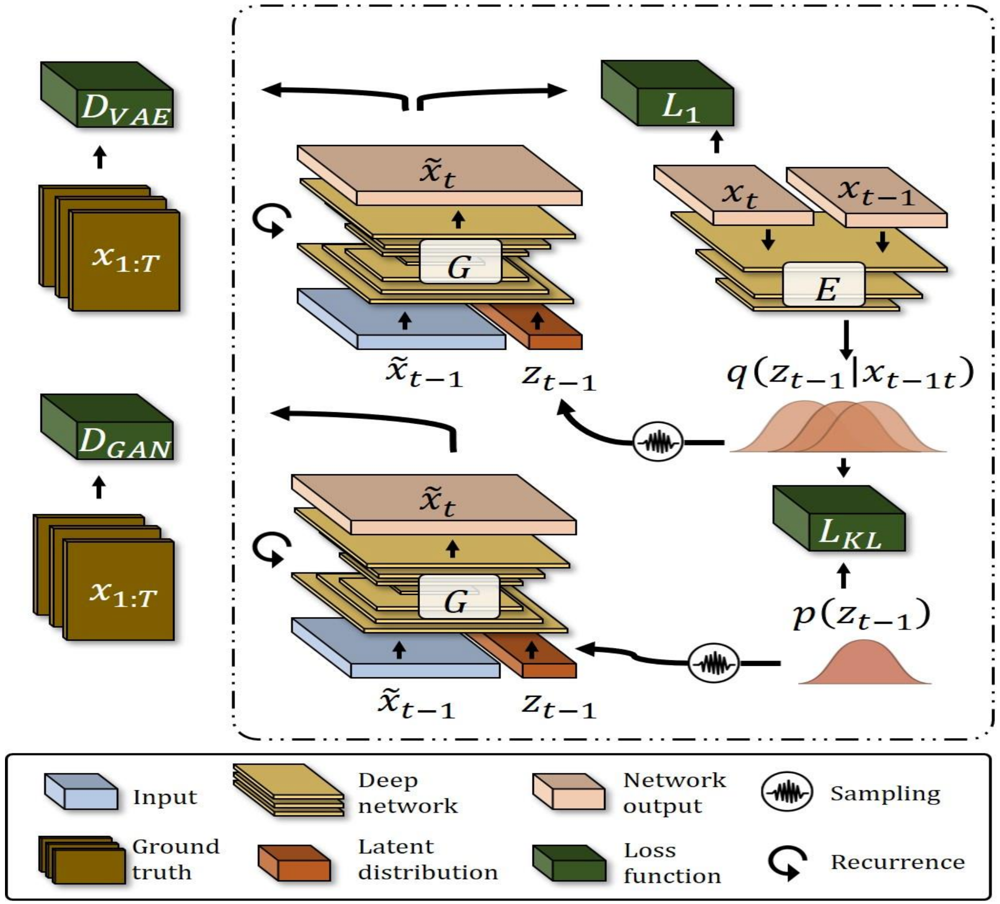
Models - ConvLSTM GAN (CLGAN)



- **CLGAN** extends the **ConvLSTM** and **U-Net** by introducing the **adversarial loss.**
- **CLGAN** attempts to learn the spatiotemporal correlations and fine-features to obtain sharp predictions

DEEP LEARNING ARCHITECTURES

Models - Stochastic adversarial video prediction (SAVP)



Combines Generative adversarial nets (GAN) and variational autoencoder (VAE) and ConvLSTM to enable high-quality stochastic video prediction [5]

[5] Lee, Alex X., et al. "Stochastic adversarial video prediction." arXiv preprint arXiv:1804.01523 (2018).

VERIFICATION METHODS

- **Methods for forecasts of continuous variables**

  RMSE (Root mean square error)

  ACC (Anomaly correlation coefficient)

$$RMSE = \sqrt{\frac{\sum_i^N (y_i - o_i)^2}{N}}$$

$$ACC = \frac{\sum_{i=1}^N y_i' o_i'}{\sqrt{\sum_{i=1}^N (y_i')^2 \sum_{i=1}^N (o_i')^2}}$$

- **Methods for dichotomous (yes/no) forecasts**

  CSI (Critical Success Index)

Contingency Table

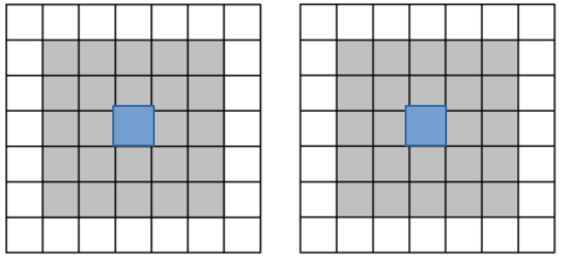
		Observed yes	Observed no	Total
Forecast	yes	*hits*	*false alarms*	*forecast yes*
	no	*misses*	*correct negatives*	*forecast no*
Total		*observed yes*	*observed no*	*total*

$$CSI = \frac{hits}{hits + misses + falsealarms}$$

- **Methods for spatial forecasts**

  FSS (Fractions skill score)

  MODE (the Method for Object-based Diagnostic Evaluation)

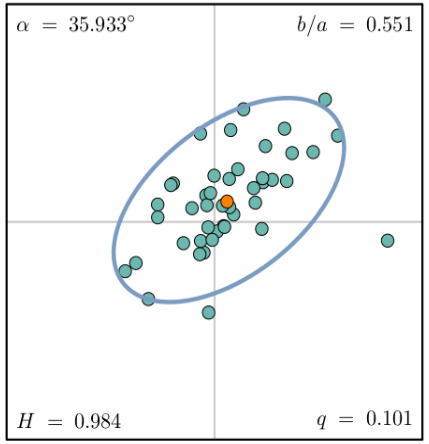


observation (fuzzy verification)

matched forecast (fuzzy verification)

$$\mathrm{FBS} = \frac{1}{N}\sum_N (\langle P_y \rangle_s - \langle P_x \rangle_s)^2$$

$$FSS = 1 - \frac{FBS}{FBS_{worst}}$$



APPLICATION 1

2m Temperature forecasts

Training data: ERA5 reanalysis data

- Time resolution: Hourly
- Spatial resolution: 0.3 degrees
- Inputs: previous 12 hours
- Outputs: next 12 hours
- Region: Central Europe
- Input variables:
  2m Temperature, 850 hPa temperature,
  Cloud cover

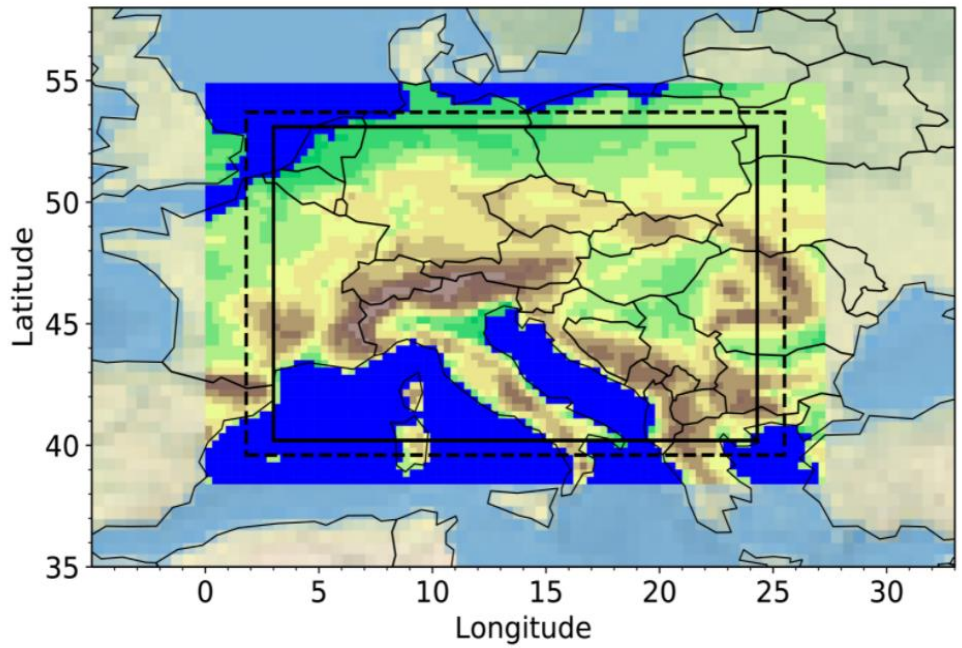


Fig. Topographic height of study region

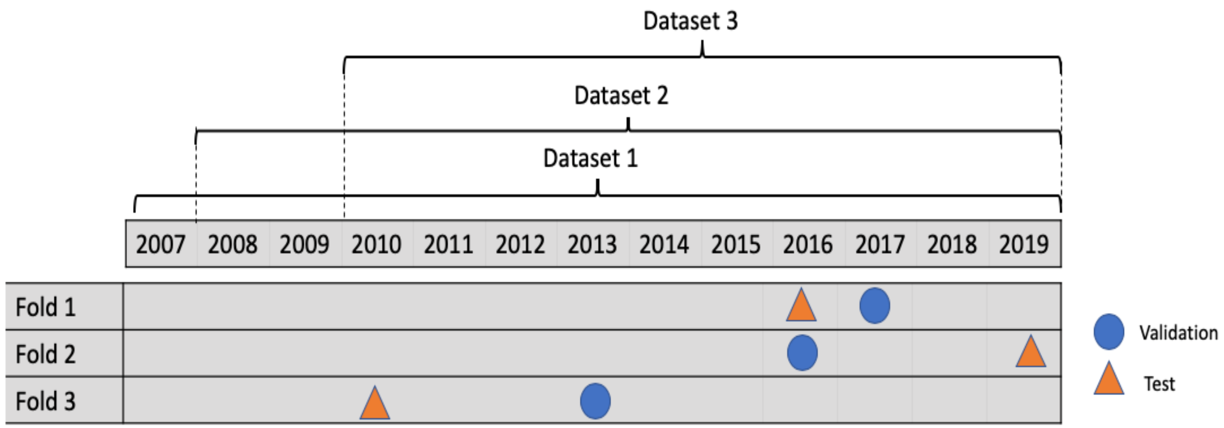


Fig. Three-fold cross validation strategies for the 13 years data from 2007 to 2019

RESULTS

2m Temperature forecasts

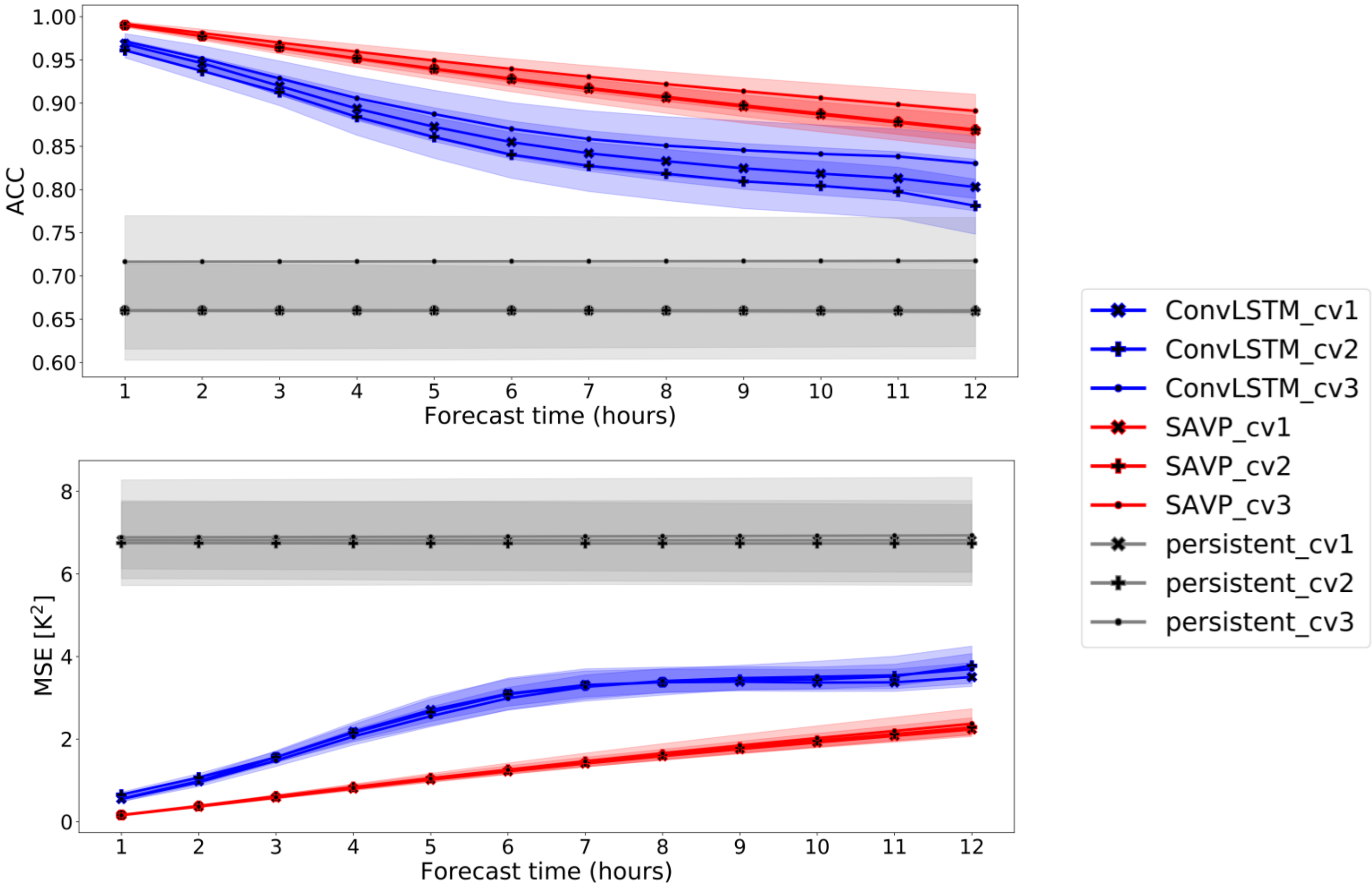


Fig. Mean scores [(a) MSE, (b) ACC] across lead times for 2 m temperature over verifi-cation periods with 95% bootstrap confidence intervals (shading area) on three cross valiation datasets

- SAVP shows remarkable improvements in terms of MSE and keep a higher consistency with observations than ConvLSTM in terms of ACC

RESULTS

2m Temperature forecasts

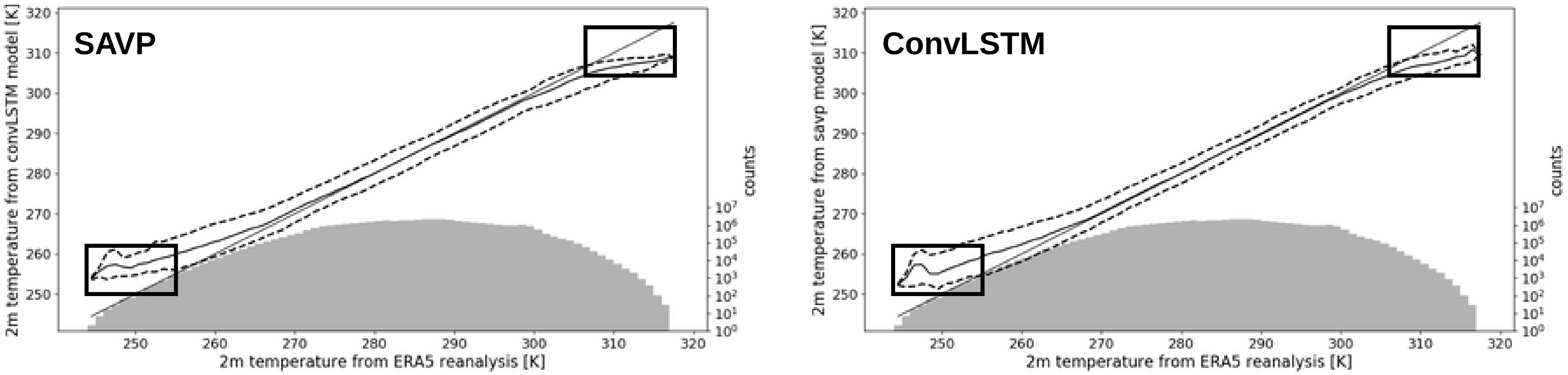


Fig. Conditional quantile plots in terms of calibration-refinement factorization for 2 m temperature forecasts with the lead time of 1hand 12h using ConvLSTM and SAVP model over 2016. The solid straight line is the 1:1 rference line, the dashed lines are 10th and 90th quantiles respectively and the solid line is the median.

- SAVP behaves better in forecsat high/low temperature

JÜLICH Forschungszentrum

RESULTS

2m Temperature forecasts

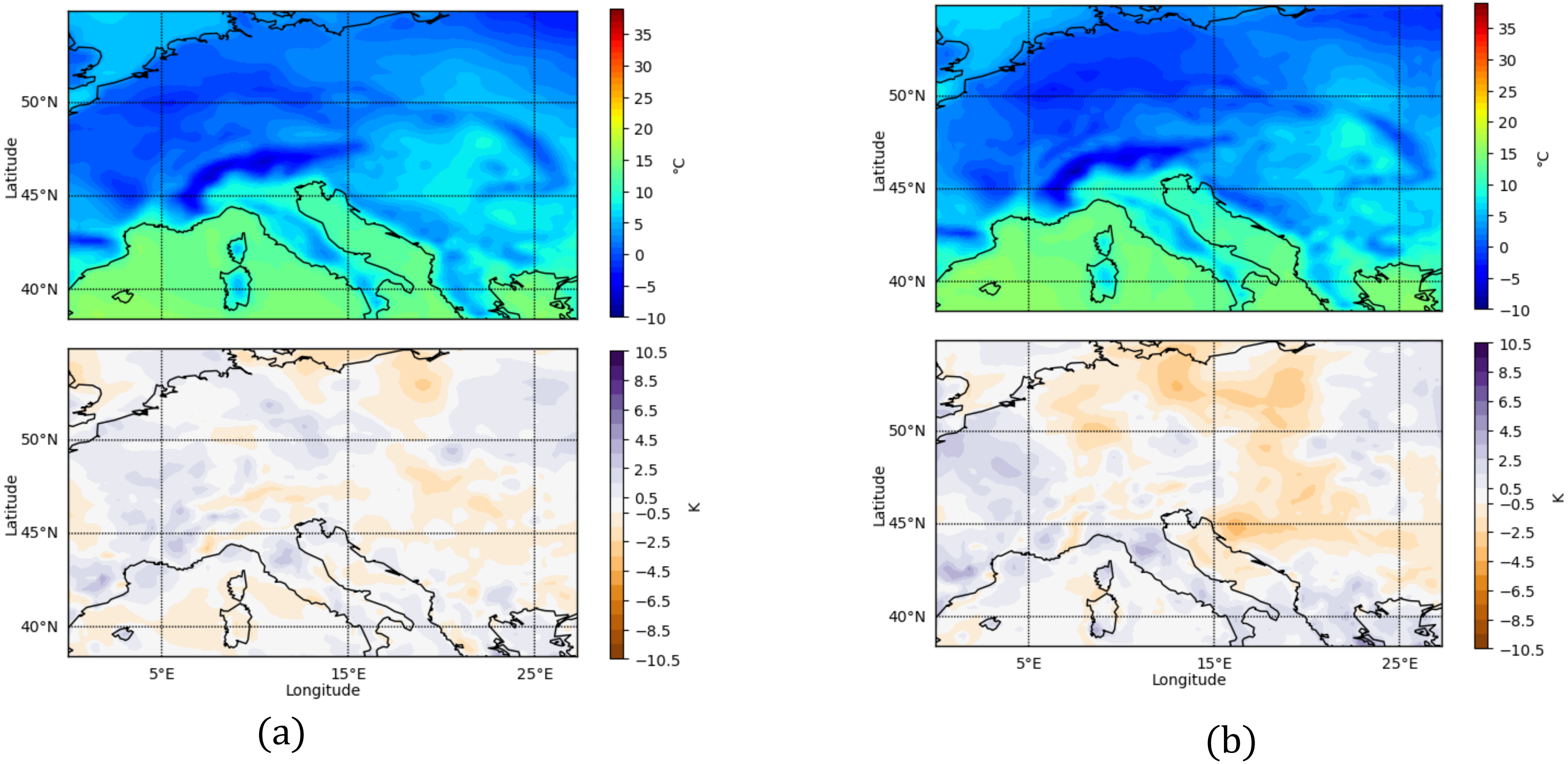


Fig. 2m temperature prediction with the 12th -hour lead time (1st row) and the temperature differences between forecast and ground truth (2nd row)
(a) The **best** 2m temperature () prediction (2019-04-13 14:00:00) by **SAVP**
(b) The 2m temperature () prediction (2019-04-13 14:00:00) by **ConvLSTM;**

RESULTS

2m Temperature forecasts

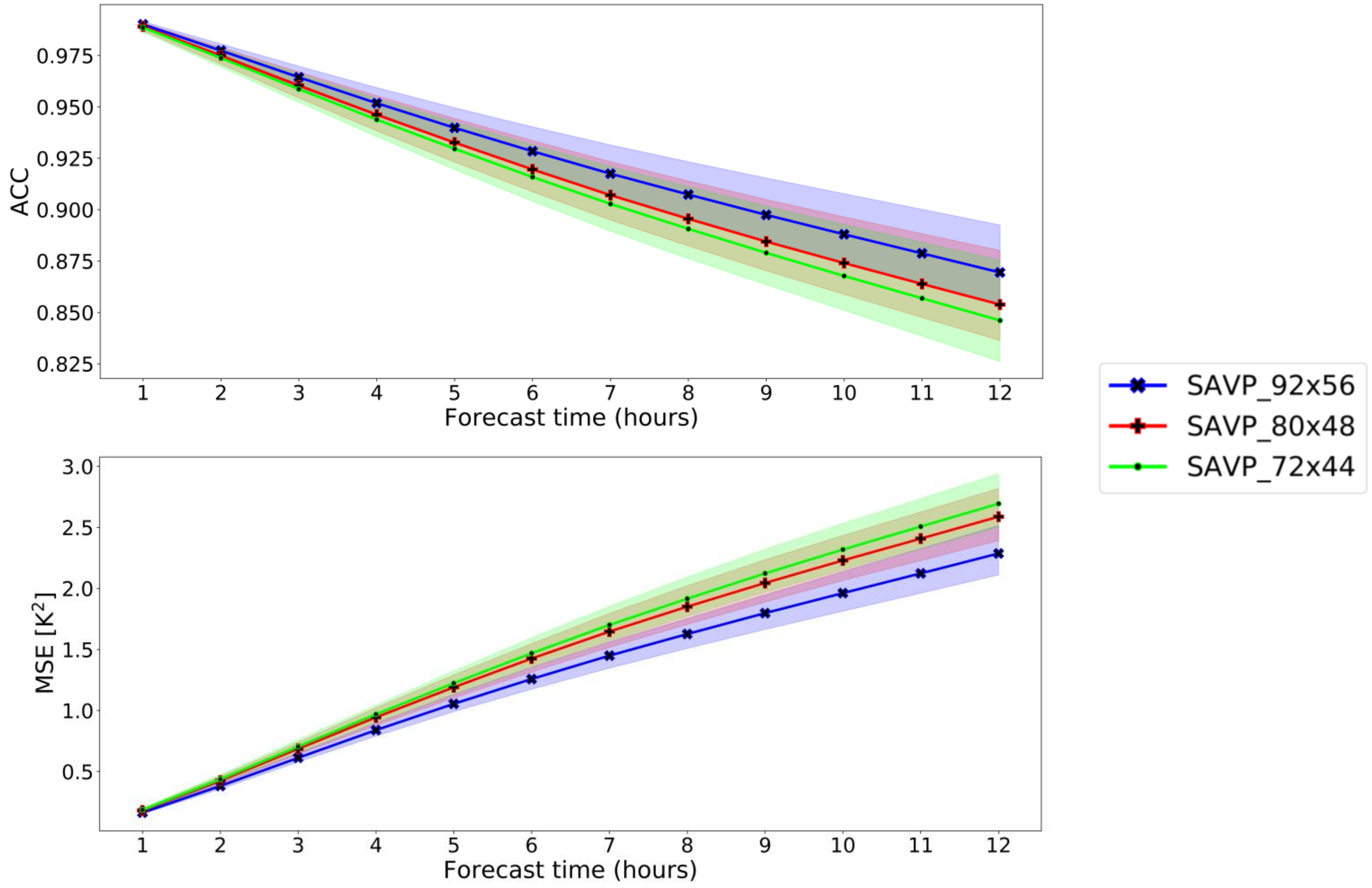
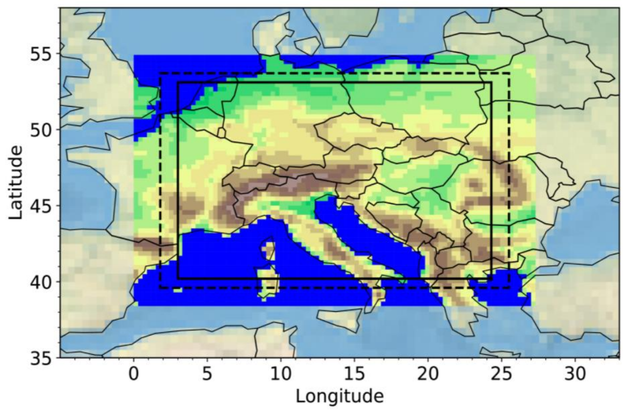


Fig. Mean scores [(a) MSE, (b) ACC] across lead times for 2 m temperature over verifi-cation periods with 95% bootstrap confidence intervals (shading area) on different size of study regions



- The size of study region shows influence in results where the experiment in a larger region behaves better.

- This is probably because large-scale features can be captured by introducing large size of the data; or due to smaller marine regions in the smaller target domain, where the temperature has higher spatio-temporal variability that hardly be captured by DL

APPLICATION 2

Precipitation nowcasting

Training data: observations from AWS

- Time resolution: 10-Minutes
- Spatial resolution: 0.125 degrees
- Inputs: previous 12 frames (120 minutes)
- Outputs: next 12 frames (120 minutes)
- Region: Guizhou, China
- Training period: 2015~2018
  Testing period: 2019
- Data preprocessing:
  I. Bilinear interpolation
  II. Log transformation
  III. Maxmin normalization

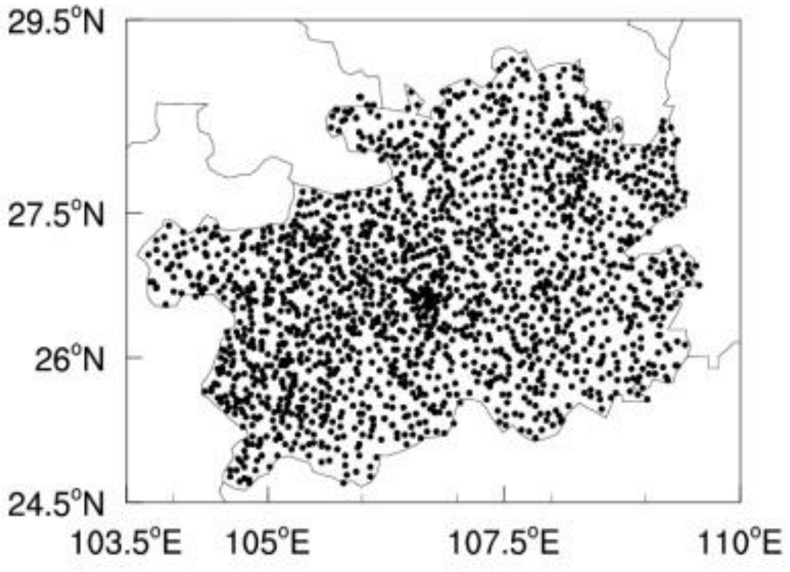


Fig. The spatial distribution of AWS over Guizhou

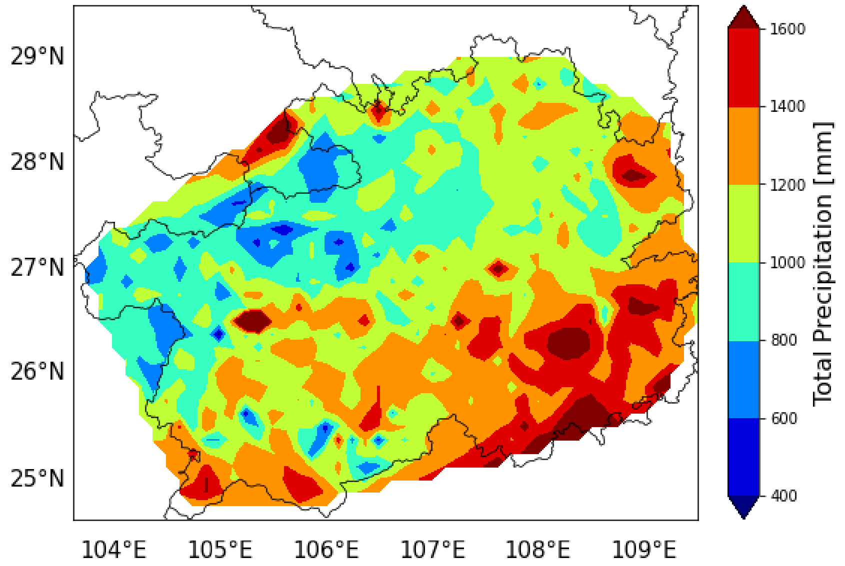


Fig. Annual average cumulative precipitation in Guizhou from 2015 to 2019

JÜLICH
Forschungszentrum

RESULTS

Precipitation nowcasting

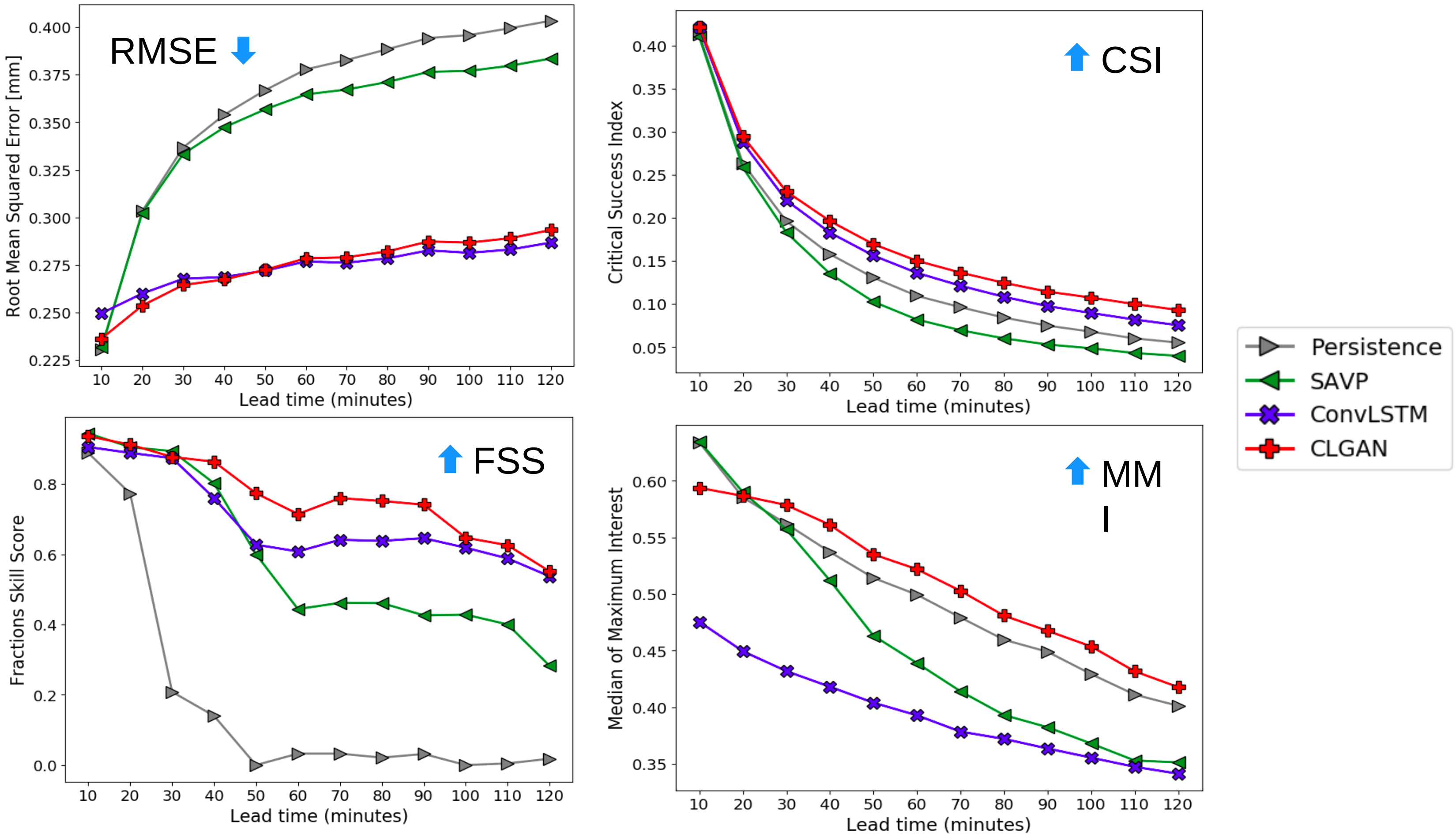
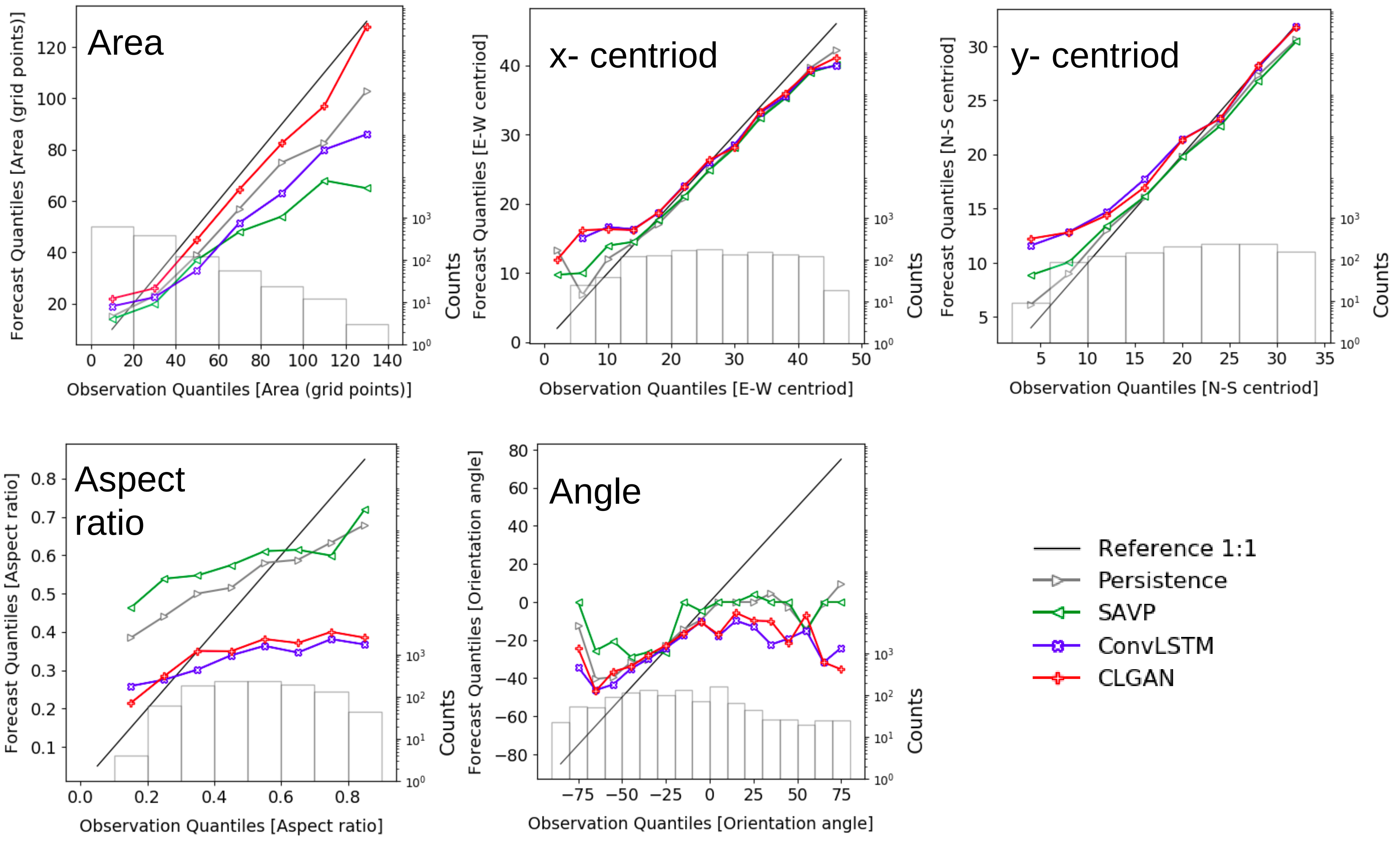


Fig. Mean scores [RMSE, CSI, FSS and MMI] across lead time for precipitation rate [mm/(10 minutes)] over the verification period using different models. The threshold T of CSI is 1 mm. The threshold T of FSS is 1 mm and the scale radius s is 3.

- All the video prediction models reduce the forecast errors compared with the persistence forecast in terms of RMSE, especially the ConvLSTM and CLGAN models
- CLGAN behave better in evaluation metrics for dichotomous and spatial forecasts in terms of CSI, FSS and MMI

RESULTS

Precipitation nowcasting



- These precipitation attributes are computed following the method of object-based diagnostic evaluation
- Video prediction models have advantages in forecasting precipitation area and position
- Forecasts of precipitation shape need improvements

JÜLICH
Forschungszentrum

RESULTS

Precipitation nowcasting

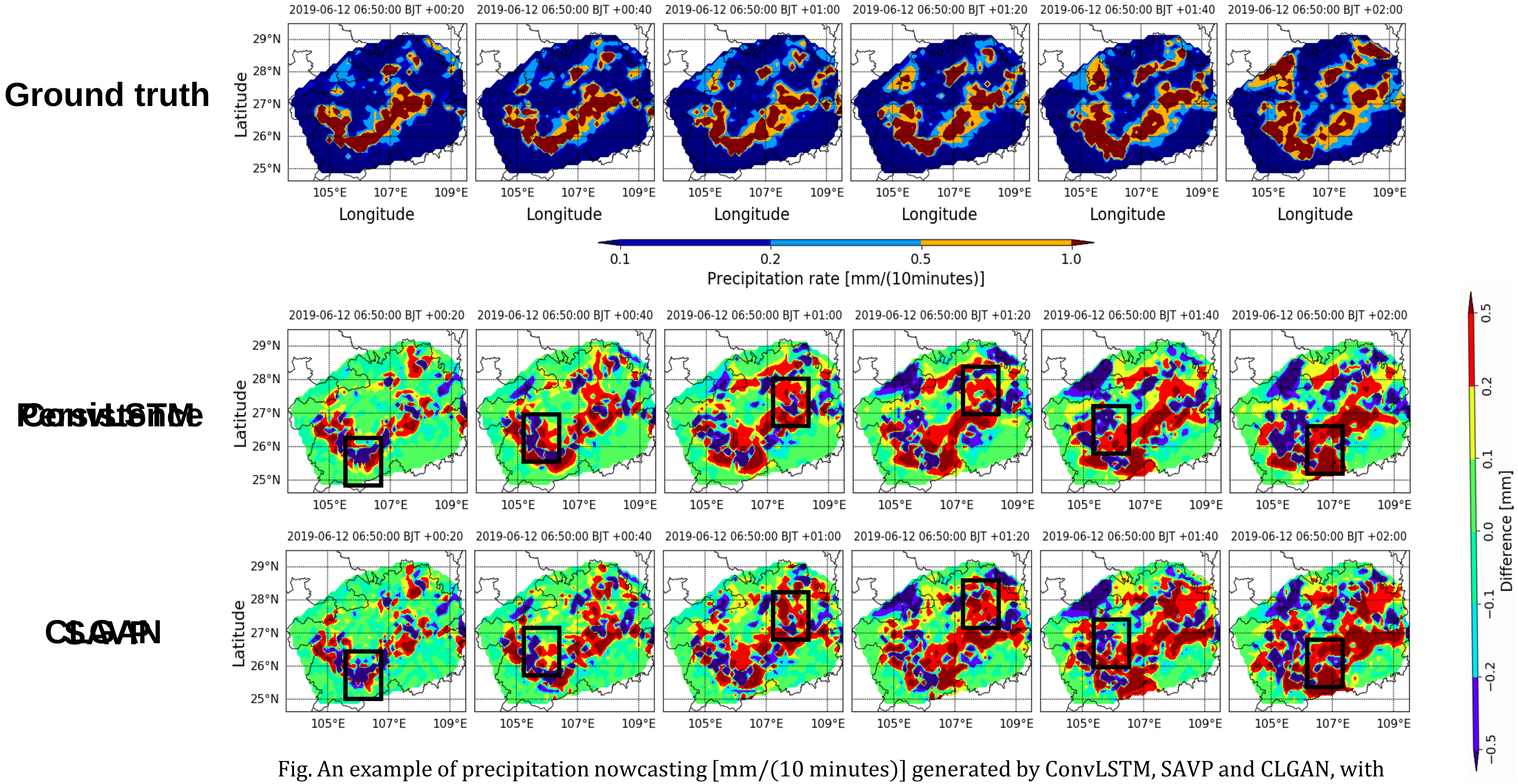
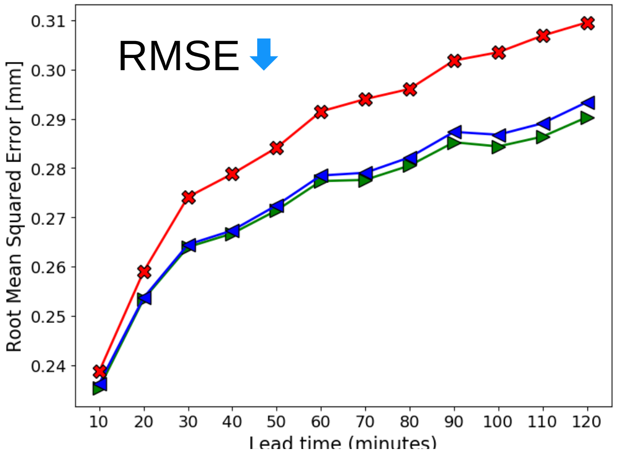
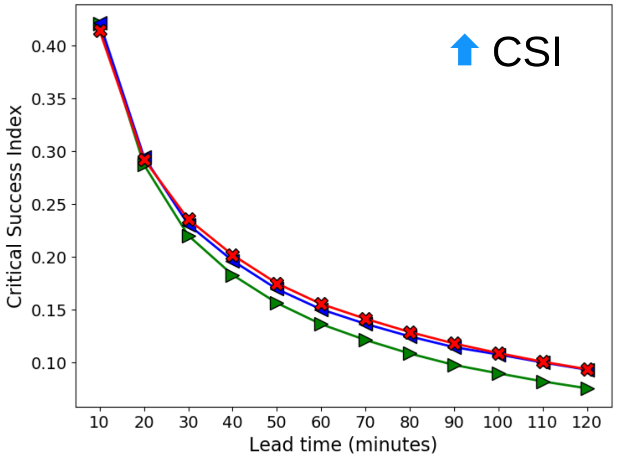


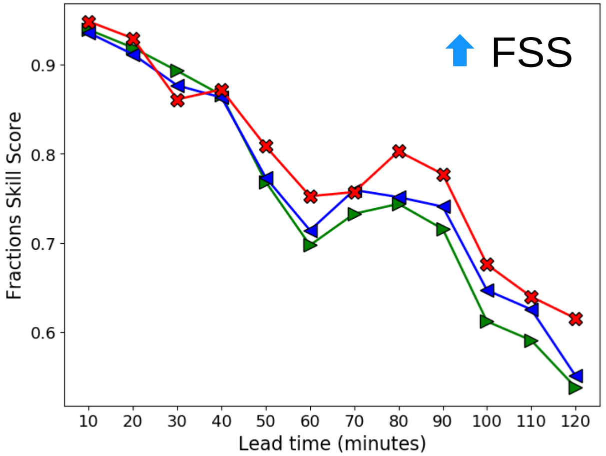
Fig. An example of precipitation nowcasting [mm/(10 minutes)] generated by ConvLSTM, SAVP and CLGAN, with ground truth. The initial time is June 12th 2019 06:50 (BJT) and the lead time is every 20 minutes up to 2 hours

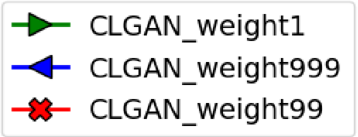
RESULTS

Precipitation nowcasting









$$\mathcal{L}^{*} = (1-\lambda)\mathcal{L}^{G} + \lambda\mathcal{L}^{2}, \lambda \in [0,1]$$

Fig. Mean scores [RMSE, CSI, FSS and MMI] across lead time for precipitation rate [mm/(10 minutes)] over the verification period using different models. The threshold T of CSI is 1 mm. The threshold T of FSS is 1 mm and the scale radius s is 3.

- When enlarging the weight of adversarial loss, the forecast error will increase in terms of RMSE while the scores for dichotomous and spatial forecasts improve in terms of CSI and FSS.
- The adversarial loss has more relationships with generating detailed structures.

CONCLUSION

- Video prediction models can achieve remarkable improvements in short-term weather forecasting.
- The SAVP model shows the best performance in 2 meter temperature forecasting. The size of study region shows influence in results where the experiment in a larger region behaves better. The predictors show a significant effect on forecast performance and adding other variables can improve the skills.
- The CLGAN model behaves better in precipitation nowcasting. Video prediction models have advantages in forecasting precipitation area and position while the forecasts of precipitation shape need improvements.
- Tuning the hyperparameters to balance adversarial loss and l2 loss is important when using GAN-based model in different applications.

ACKNOWLEDGEMENT

MAELSTROM

The project is funded by the European Union's Horizon 2020 research and innovation programme and the EuroHPC Joint Undertaking under Grant Agreement 955513 and co-funded by the German Ministry of Education and Research (BMBF) under funding reference 16HPC029.

DeepRain

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

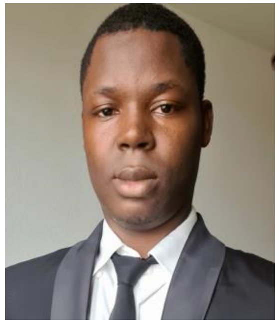
Our team

Bing Gong

Michael Langguth

Amirpasha Mozafarri

Yan Ji



Karim Mache

Martin Schultz (Supervisor)

THANK YOU