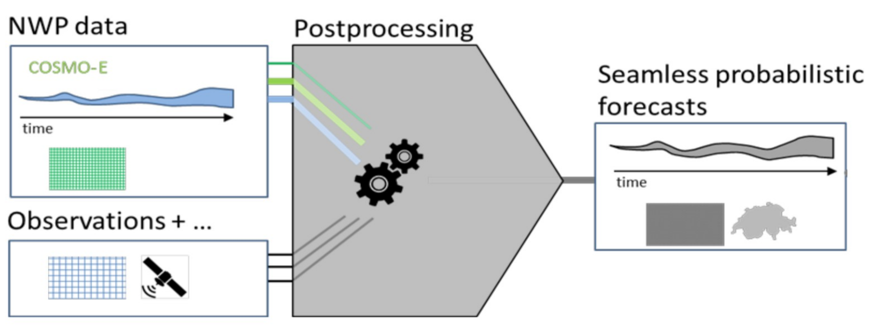
Post-Processing Cloud Cover Forecasts Using Deep Learning

Yinghao Dai, Claudio Ferrari, Rhea Sukthanker
Supervised by Dr. Stephan Hemri (MeteoSwiss), Dr. Prof. Ce Zhang (ETH) and Dr. Prof. Andreas Krause (ETH)

Data Science Lab 2019

1 Project Goals

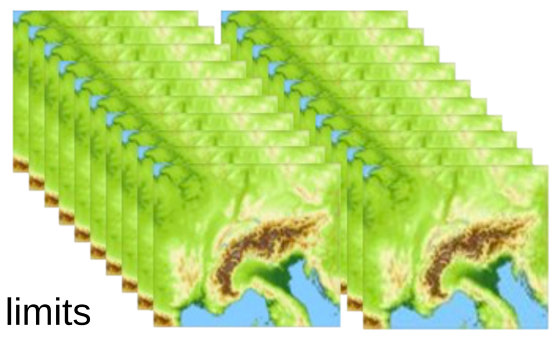


- Numerical weather prediction (NWP) models describe current atmospheric state
- The resulting predictions of cloud coverage suffer from systematic biases
- Post-processing is used to reduce these biases, currently done manually

Our objective
- Introduce machine learning based post-processing
- Provide spatial, probabilistic and seamless cloud coverage prediction
- Reduce forecast prediction error

2 Dataset

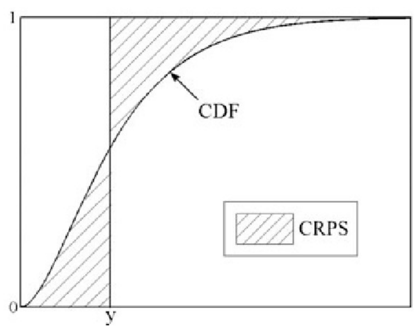
Input
- COSMO-E weather simulation
  - Twice-daily (initialization times 00 and 12)
  - Hourly cloud cover predictions for 120h into the future (so called lead times)
  - 21 ensemble simulations running concurrently to model forecast uncertainty
  - We subsampled the data due to computational limits
  - 2014/02 - present



Labels
- Hourly cloud coverage data derived from satellite images at 5x5km
- 2014 – 2018

3 Metrics

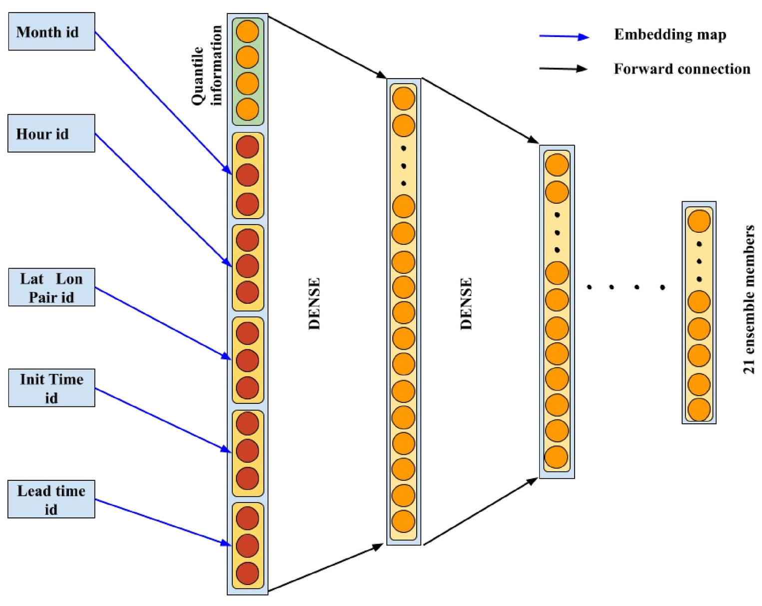
Continuous ranked probability score (CRPS)
- Error metric for forecast distribution vs. scalar observation
- We evaluate the mean CRPS by season
- CDF in our case a step function since we output 21 members of an empirical distribution



PIT histograms
- Compute ranks of observations w.r.t. 21 ensemble predictions
- Plot histogram of rank frequencies
- For a well calibrated forecast model, distribution should be uniform

4 Model

- For every point we extracted 7 quantiles from the 21 ensemble predictions
- Context information embedded into dense vectors (hour of day, latitude, longitude, initialization time, lead time)
- 7 quantiles and context embedding vector concatenated
- Dense layers with ReLU activation
- Output 21 ensemble members



5 Results

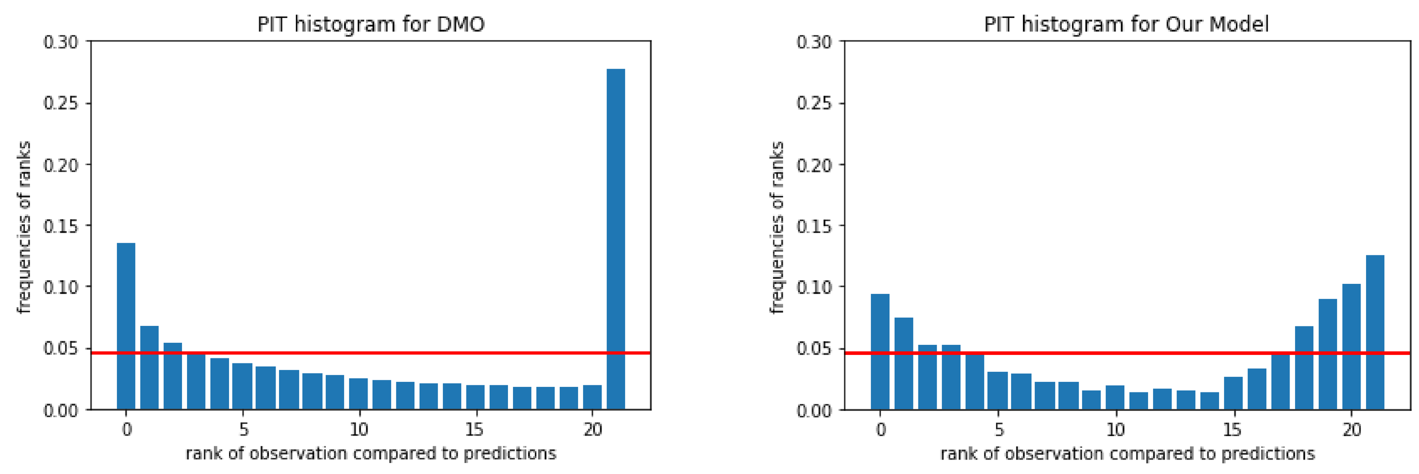


Fig. 1. PIT histograms of DMO vs. Our Model

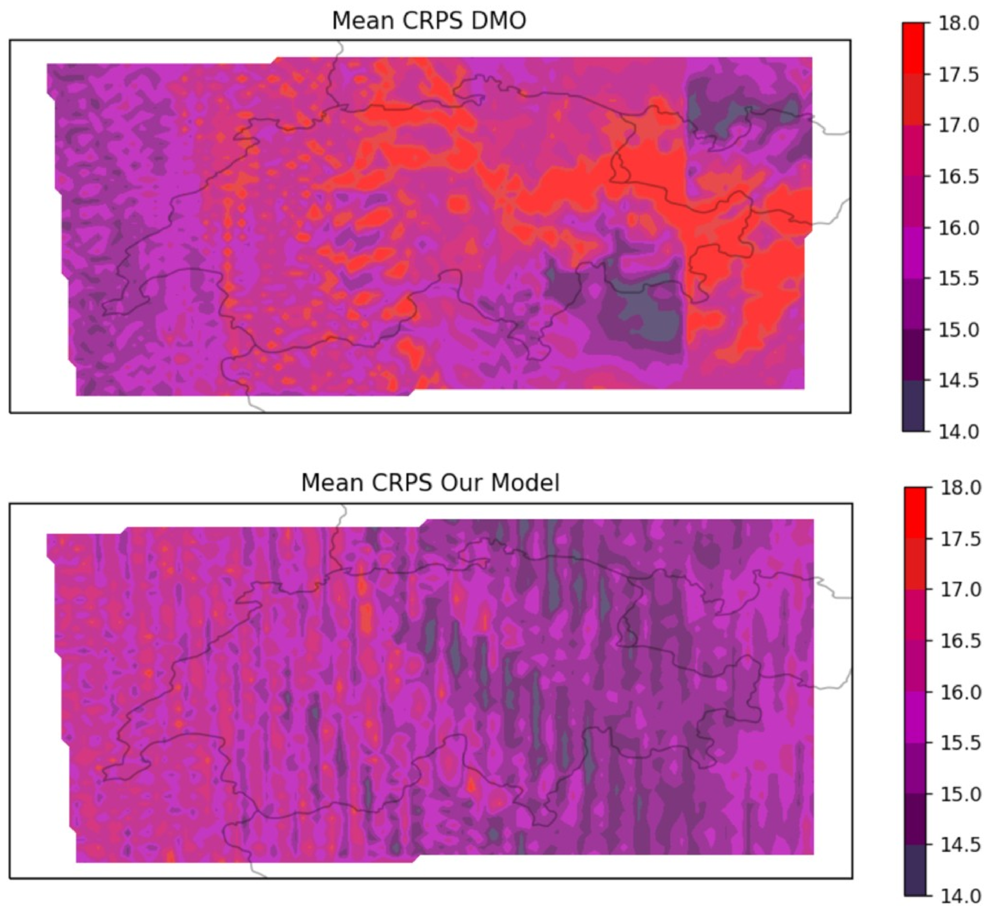


Fig. 2. Mean CRPS by location of DMO vs. Our Model

Mean CRPS	winter	spring	summer	fall	total
DMO	17.39	15.21	17.62	17.90	17.03
Our Model	**15.49**	**13.91**	**16.56**	**16.52**	**15.62**

Table 1. Mean CRPS by season of DMO vs. Our Model

6 Discussion

Results
- Significant reduction in forecast prediction error
- Predictions better calibrated
  - PIT histogram more closely resembles a uniform distribution
- Model produces interesting patterns in map plots

Future work
- Incorporate more features of the COSMO-E model
  - Humidity, temperature, atmospheric pressure, etc.
- Investigate patterns in map plots
- More interpretable model could give insights into systematic biases of DMO

7 References

1. Spirig, C., Hemri, S., Bhend, J., Rajczak, J., & Liniger, M. (2019, April). *Development of an operation postprocessing suite at MeteoSwiss.* EGU General Assembly 2019, Vienna.
2. Shi, L., Ma, H., & Lin, D.K. (2016). Process Capability Analysis via Continuous Ranked Probability Score. *Quality and Reliability Eng. Int.*, *32*, 2823-2834.

Partners Bundesamt für Meteorologie und Klimatologie MeteoSchweiz · DS3Lab ETH ZURICH · Learning & Adaptive Systems