

Analog Ensemble Probabilistic Forecasting Using Deep Generative Models (Case study: Wind Speed)

Behrooz Roozitalab

The University of Iowa, SIParCS intern



July 31, 2019



Acknowledgments



Mentor:

- Alessandro Fanfarillo (RAL)

Co-Mentor:

- Davide Del Vento

Collaborators:

- Guido Cervone (Penn State University)
- Weiming Hu (Penn State University)

CISL OUTREACH, DIVERSITY, AND EDUCATION (CODE)

- AJ Lauer
- Virginia Do
- Elliott Foust
- Blake Lewis (Intern)

Administrative supports admins:

- Shilo Hall
- Kristi Hartsock



Predictions uncertainty has to be addressed!

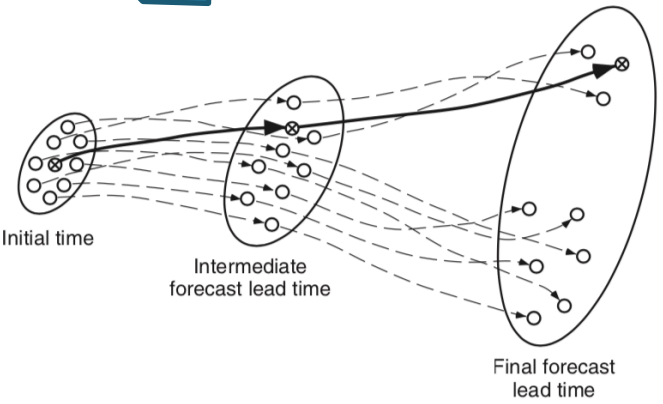
NWP¹ Forecast

Initial & Boundary conditions
(scientific & data uncertainty)

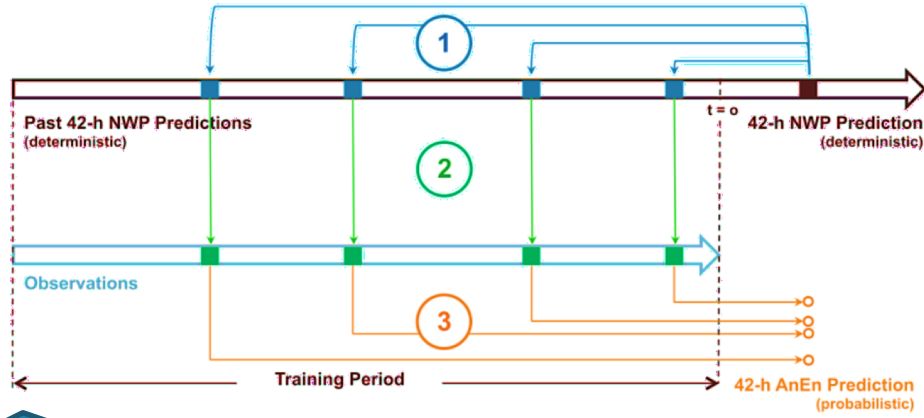
Physics
(science)

Dynamics
(science)

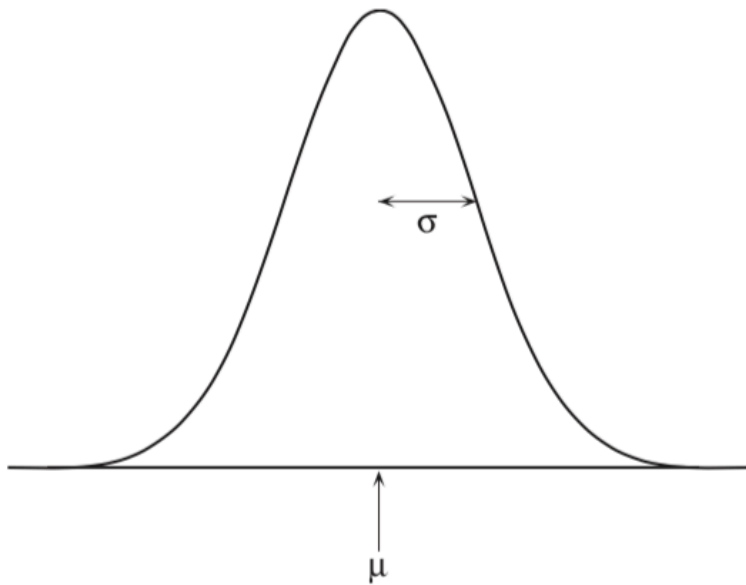
Ensemble Forecasting²



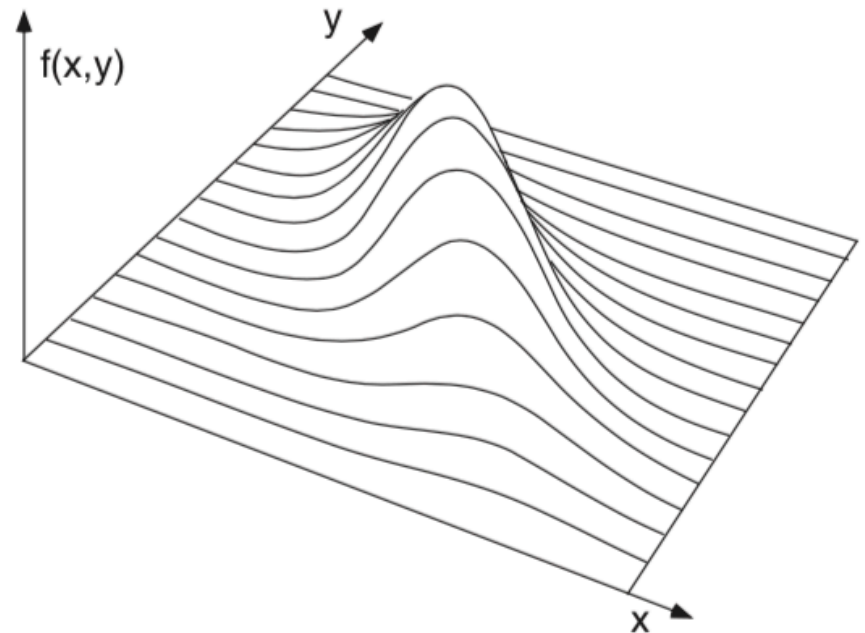
Analog Ensemble³



But we can use the conditional probability distribution instead of the huge memory-consuming dataset.

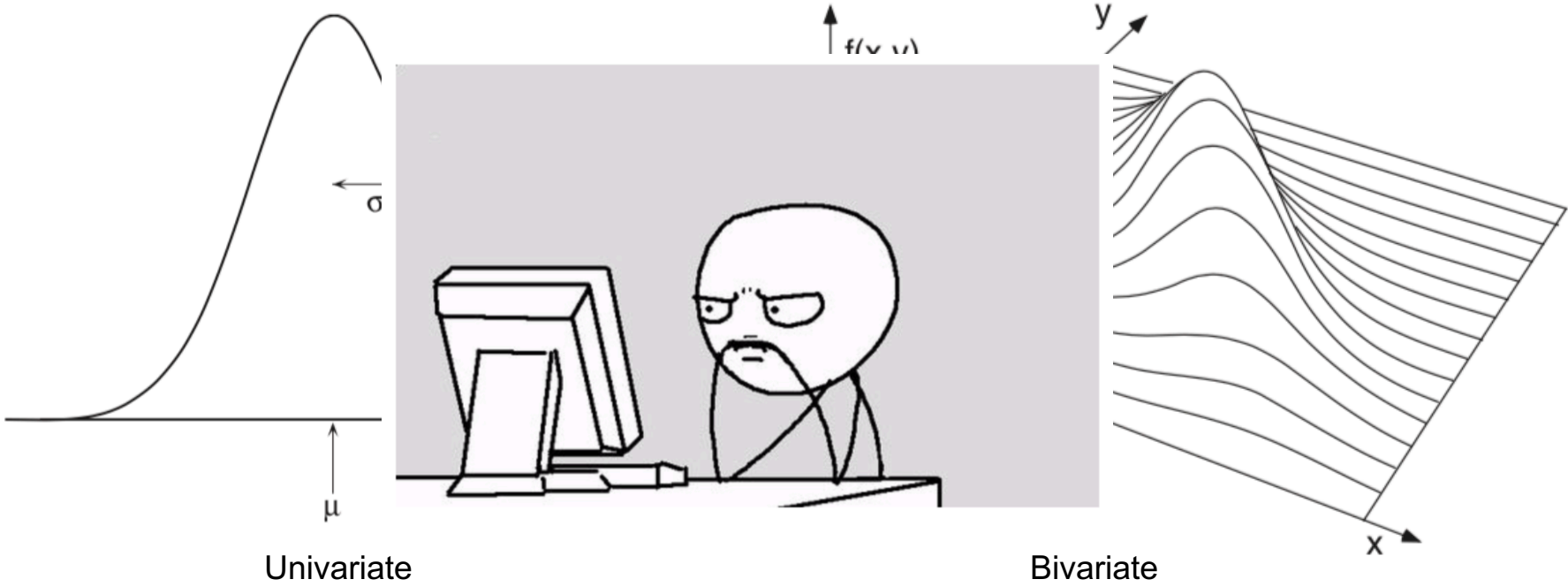


Univariate



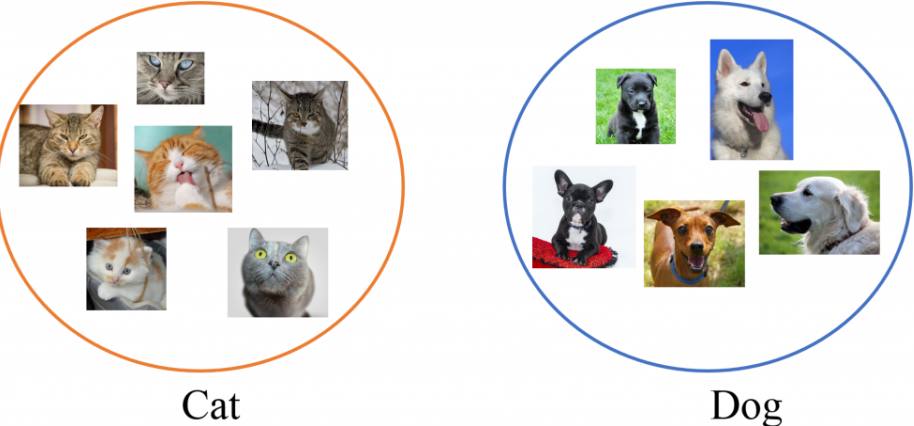
Bivariate

But we can use the conditional probability distribution instead of the huge memory-consuming dataset.



Conditional Variational AutoEncoder (CVAE) is a Generative not a Discriminative machine learning (ML) model.

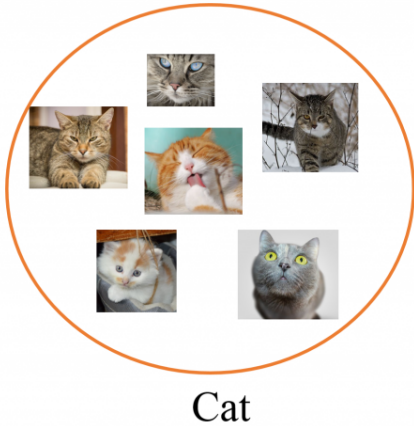
Discriminative model



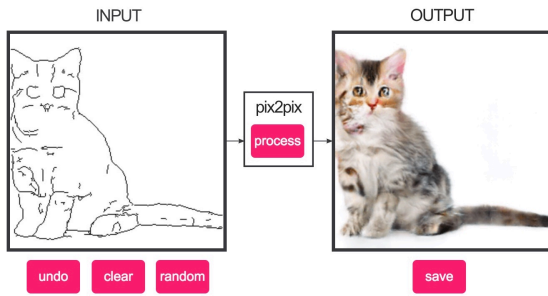
Is it a cat?



Generative model

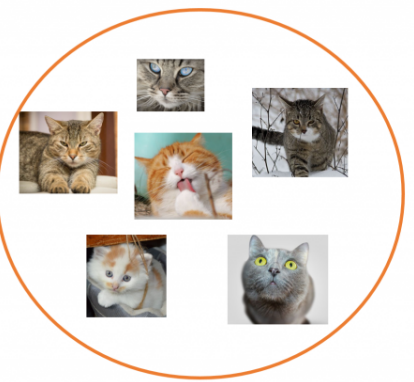


Give me a cat image

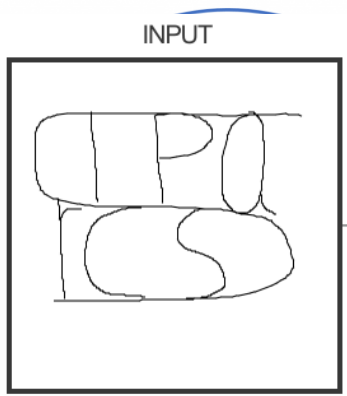


Conditional Variational AutoEncoder (CVAE) is a Generative not a Discriminative machine learning (ML) model.

Discriminative model



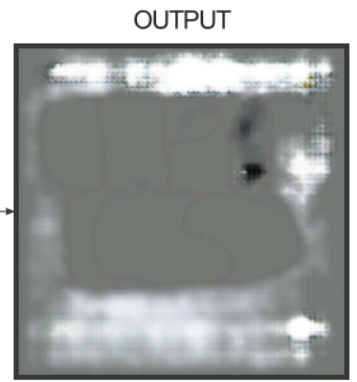
Cat



undo clear random

pix2pix
process

Generative model



save

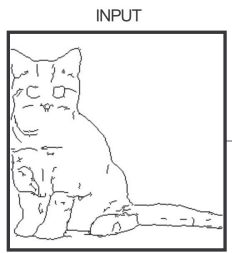


Cat

Is it a cat?



Give me a cat image



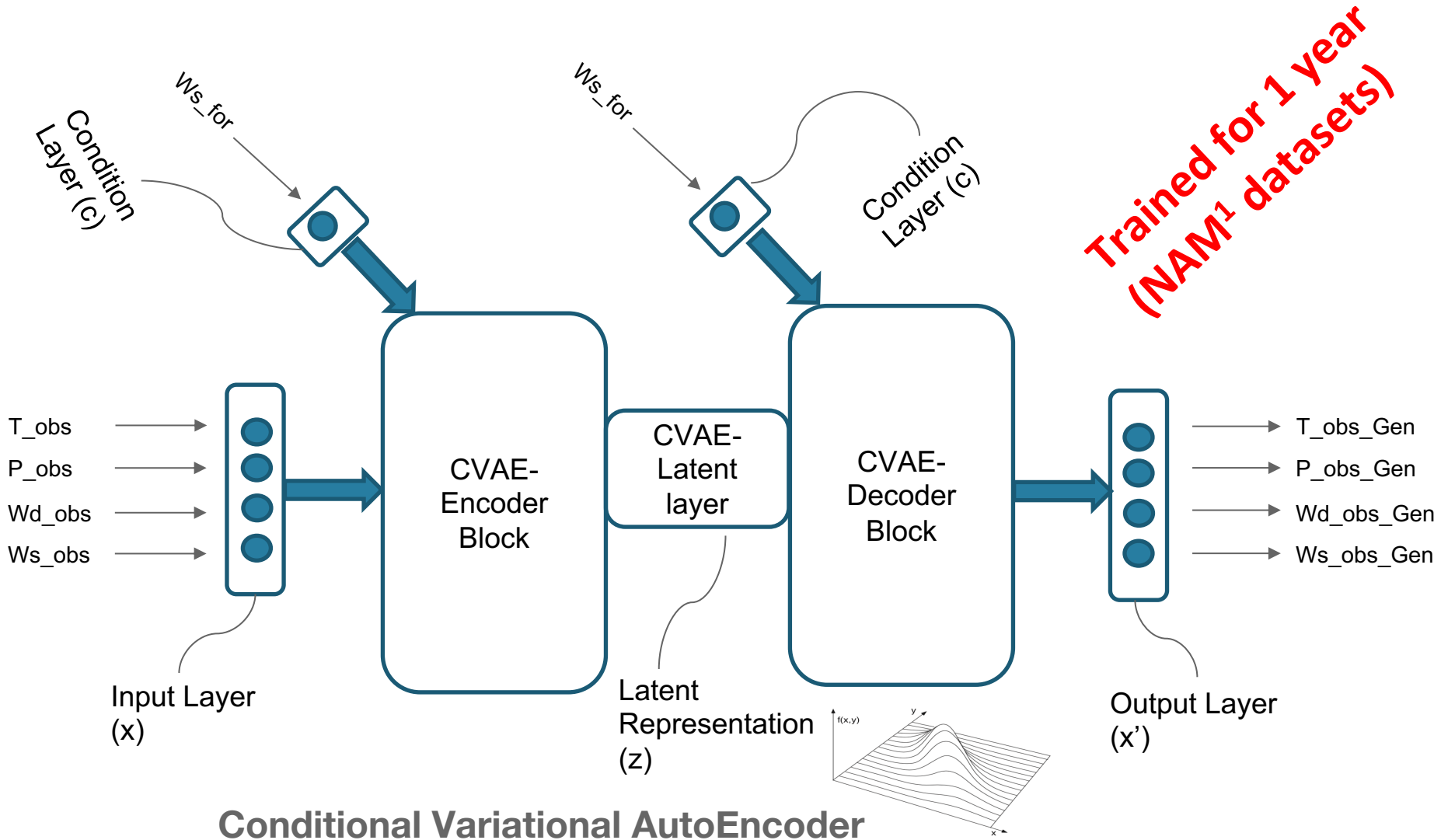
undo clear random

pix2pix
process

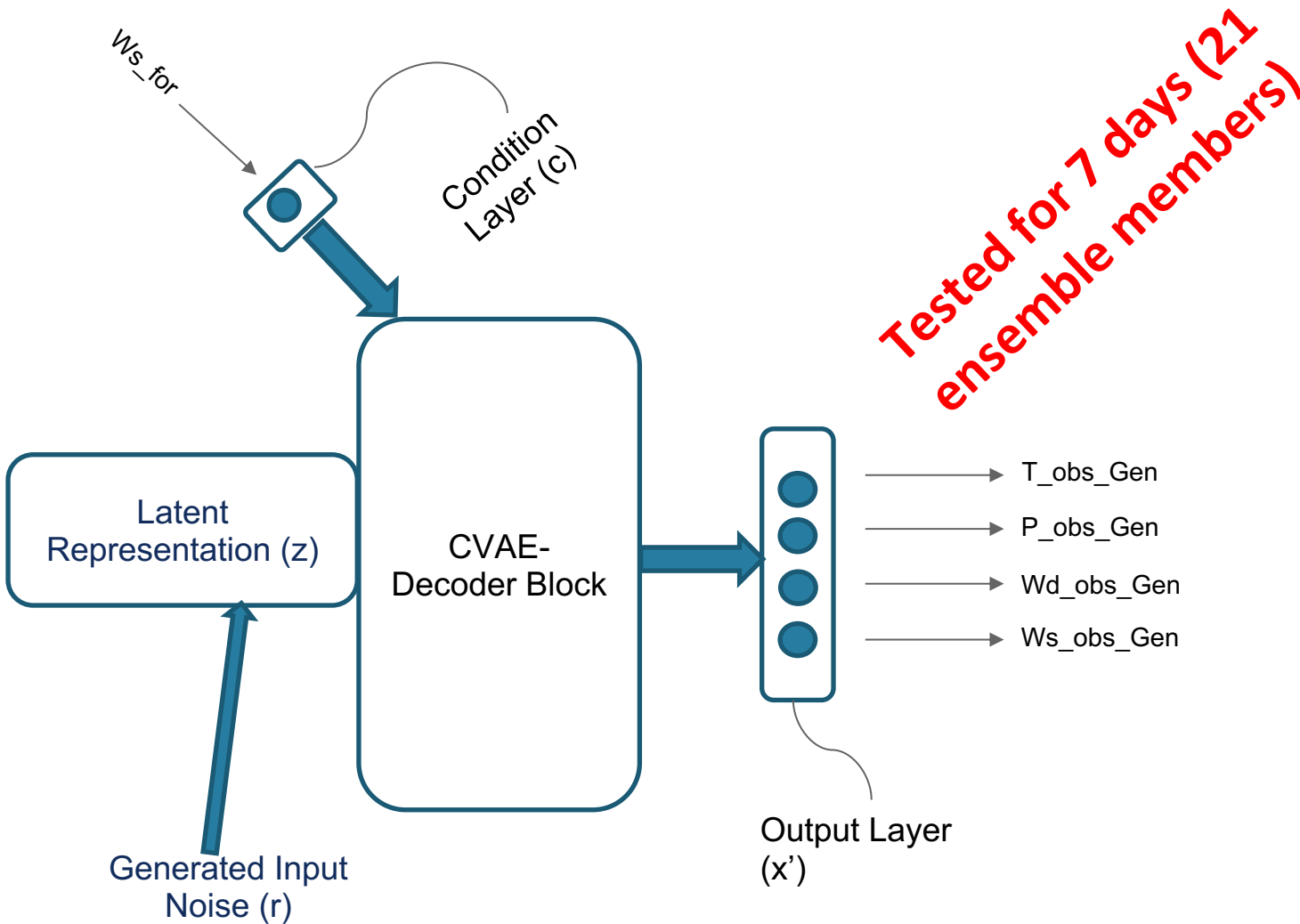


save

The performance of a ML model significantly depends on the Architecture and Hyperparameters.



Conditional PDF is stored in the Latent Representation layer.

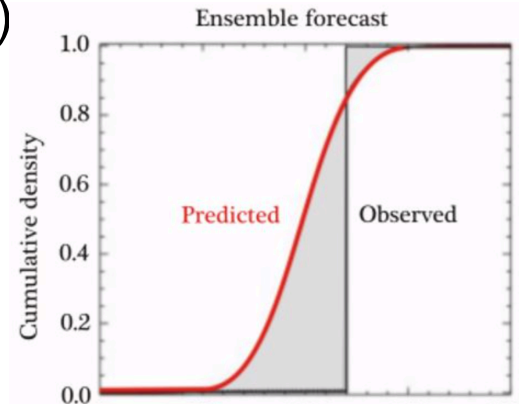


How to evaluate performance of probabilistic forecasts?

- **Consistency:**

- Continuous Ranked Probability Score (CRPS)

$$CRPS = \frac{1}{N} \sum_{i=1}^N \int_{-\infty}^{\infty} \{F_i^f(x) - F_i^a(x)\}^2 dx$$



- **Reliability:**

- Rank Histogram (RH)

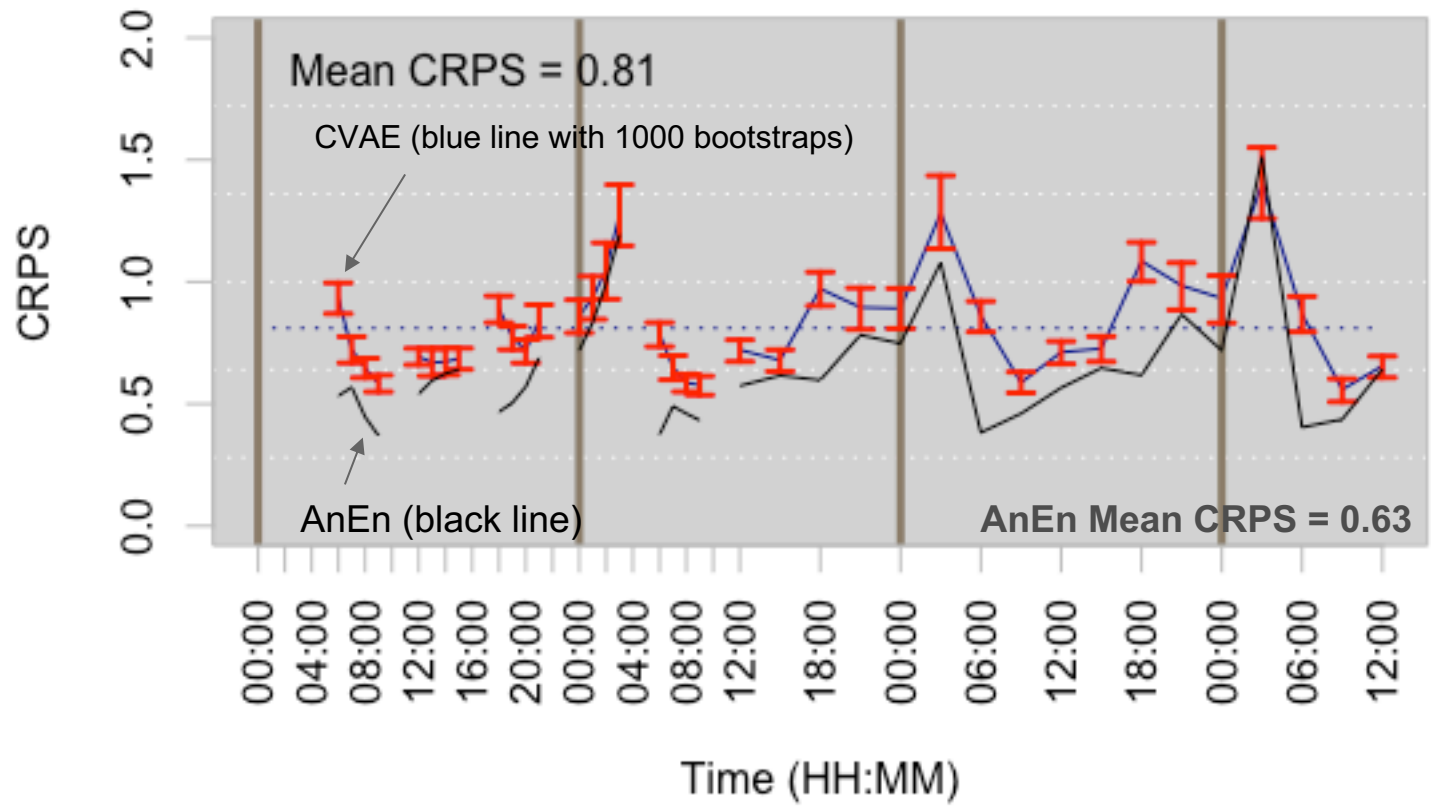
An observed value will be ranked based on its corresponding ensemble members and the results after giving ranks to all the observed values will be presented



- Dispersion

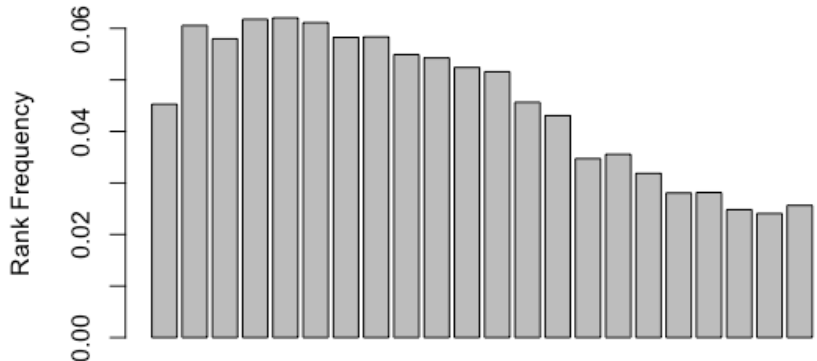
$$E[(x - x^{\wedge})^2] = \frac{m + 1}{1m} E[s^2]$$

CRPS for CVAE is comparable to AnEn.

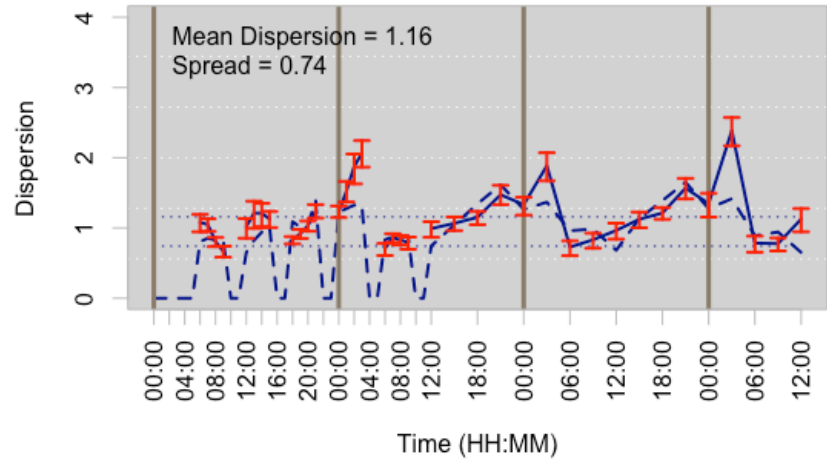
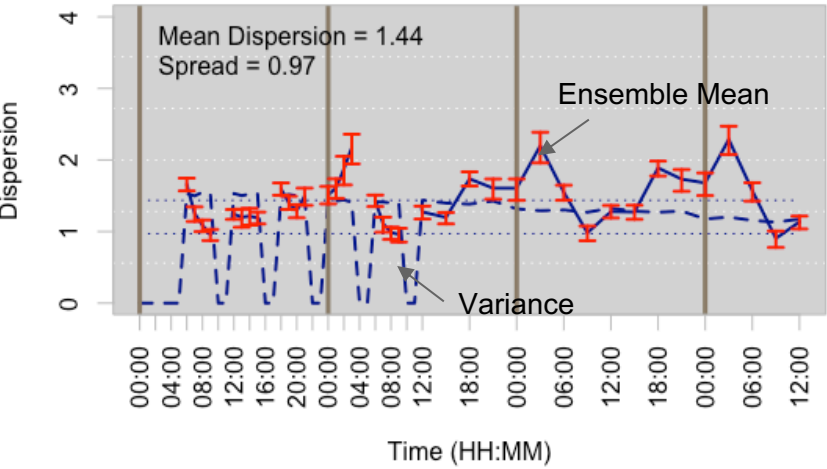
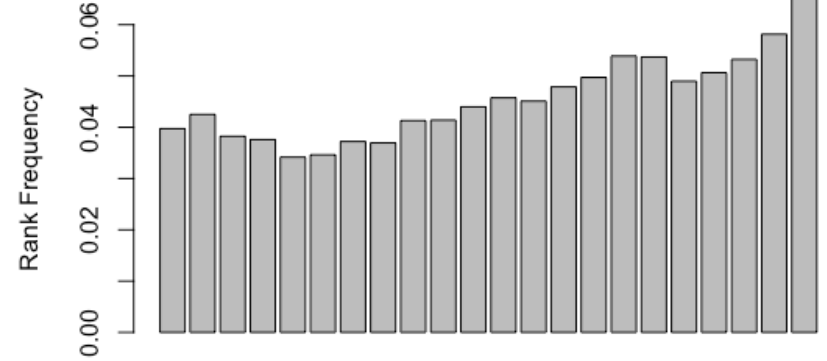


CVAE probabilistic forecasts are as reliable as AnEn but with higher Bias.

CVAE

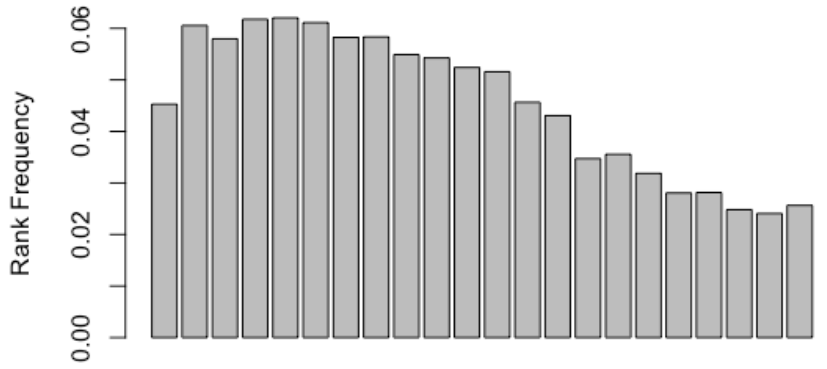


AnEn

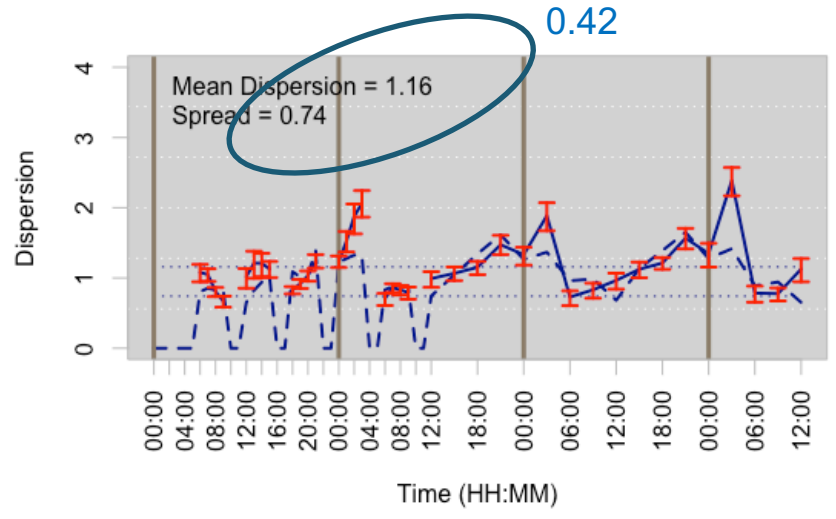
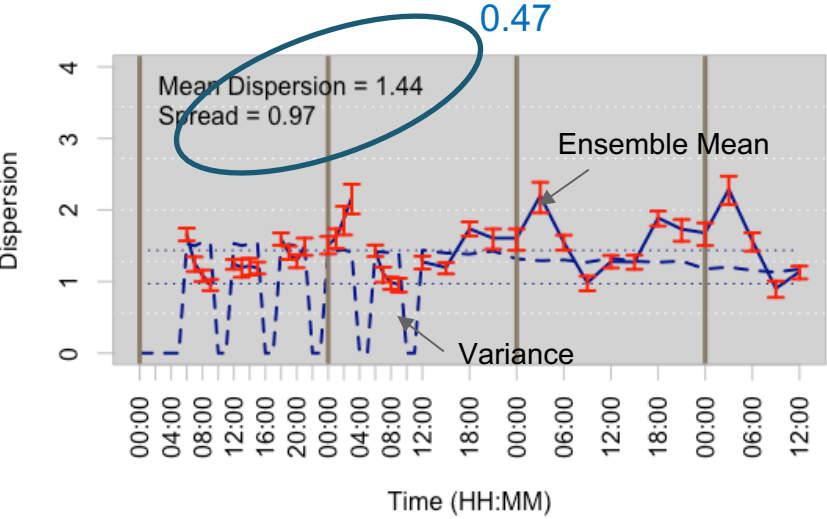
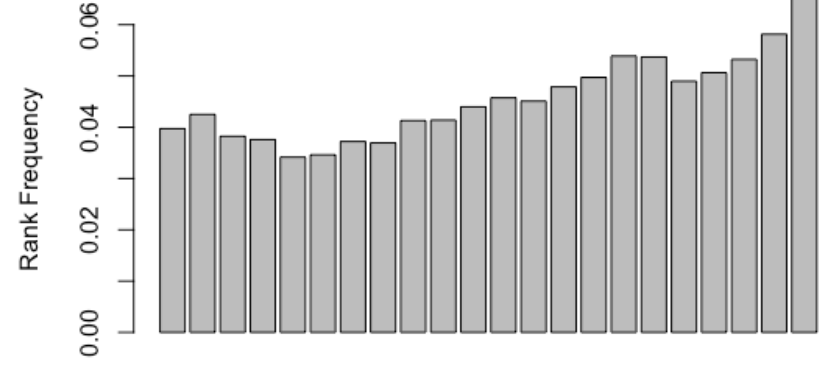


CVAE probabilistic forecasts are as reliable as AnEn but with higher Bias.

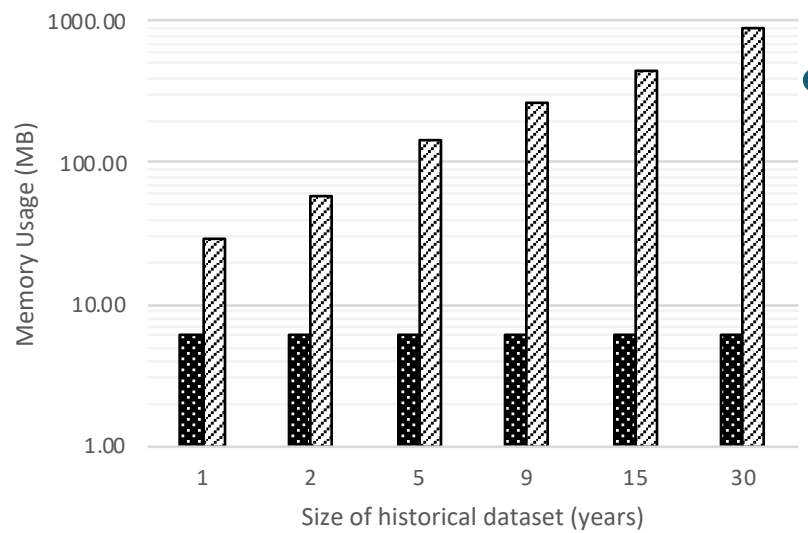
CVAE



AnEn

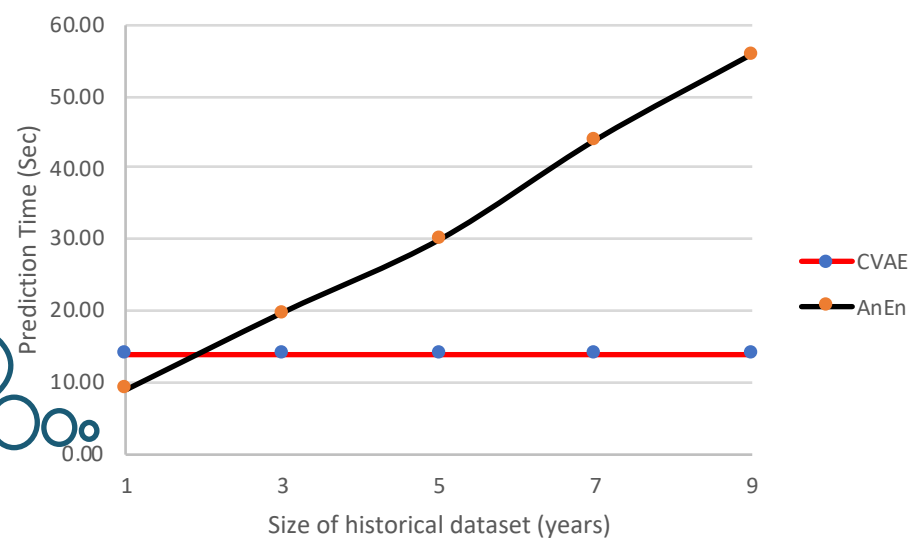


CVAE is highly memory- and time-efficient comparing to AnEn.



✓ CVAE needs a constant model to be loaded
✗ AnEn has to keep a new dataset in memory

✓ CVAE uses a constant model
✗ AnEn has to search the physical dataset



Summary

- CVAE is a probabilistic machine learning model that can be used for probabilistic forecasts.
- Probabilistic forecasts evaluation for CVAE shows consistent and reliable performance of the model.
- CVAE significantly saves computational resources.

Future work

- Training the model with more features.
- Tuning the model to be applicable to bigger datasets.
- Testing the model with different datasets.

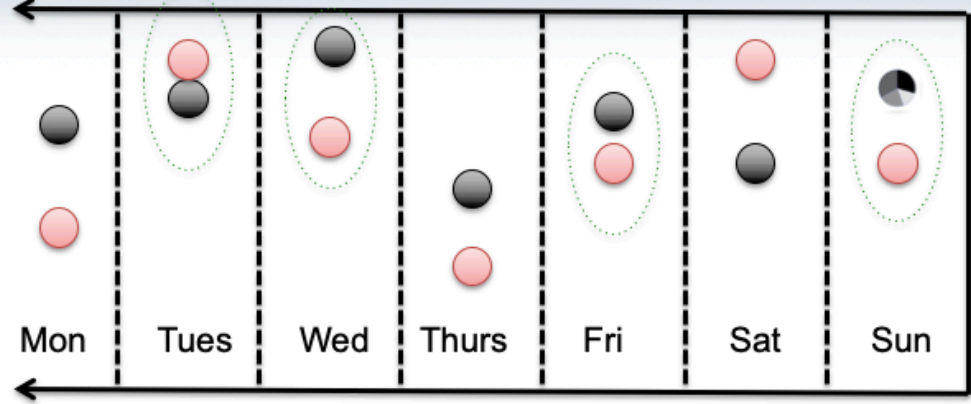
Thanks for your time.



Extra Slides

Analog Ensemble (AnEn) method is memory consuming!

Analog Ensemble Technique



$$\|F_t, A_{t'}\| = \sum_{i=1}^{N_v} \frac{w_i}{\sigma_{f_i}} \sqrt{\sum_{j=-\tilde{t}}^{\tilde{t}} (F_{i,t+j} - A_{i,t'+j})^2}$$

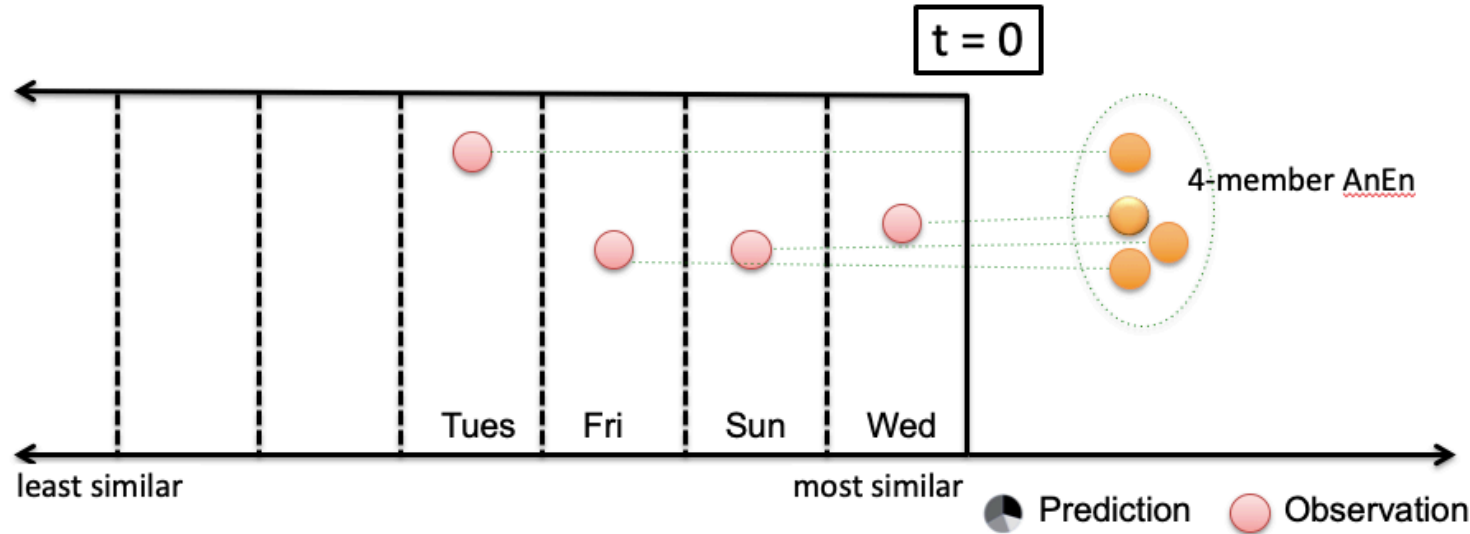


Figure credited to: Laura Clemente-Harding, Weiming Hu, Parallel Analog Ensemble Forecasts with Ensemble Toolkit on HPC, 2019 Software Engineering Assembly, NCAR, Boulder, CO, <https://sea.ucar.edu/event/parallel-analog-ensemble-forecasts-ensemble-toolkit-hpc>

Kullback-Leibler (KL) loss function

$$-D_{KL}(Q_{\phi}(z|x)||P_{\theta}(z)) = 0.5 \sum_{j=1}^J (1 + \log(\sigma_j)^2 - (\mu_j)^2 - (\sigma_j)^2)$$