Deep Unsupervised Learning for Climate Informatics

Claire Monteleoni

University of Colorado
Boulder

October 2012: Hurricane Sandy – Reuters

October 2019: Kincade Fire, California – AFP

January 2014: Drought, Folsom Lake – California Department of Water Resources


November 2019: Flooding, Venice, Italy: Reuters

Machine learning can shed light on climate change.

Climate Informatics

2011 First International Workshop on Climate Informatics (New York City)

2013 "Climate Informatics" book chapter [Monteleoni et al. 2013]

→ In the first 5 years: participants from over 19 countries and 30 U.S. states

2020 10th International Conference on Climate Informatics &
6th Climate Informatics Hackathon (online/Oxford, UK)

2021 7th Climate Informatics Hackathon, September 7-10 (online/Boulder, CO)
KITP Conf. on Machine Learning for Climate (November 1-4, Santa Barbara)

2022 **11th International Conference on Climate Informatics** (NOAA, Asheville, NC)

ENVIRONMENTAL DATA SCIENCE

An interdisciplinary, open access journal dedicated to the potential of artificial intelligence and data science to enhance our understanding of the environment, and to address climate change

Data and methodological scope: Data Science broadly defined, including:
Artificial Intelligence; Machine Learning; Computer Vision; Data Mining; Statistics; Econometrics

Environmental scope, includes:
Water cycle, atmospheric science (including air quality, climatology, meteorology, atmospheric chemistry & physics, paleoclimatology)
Climate change (including carbon cycle, transportation, energy, and policy)
Sustainability and renewable energy (the interaction between human processes and ecosystems, including resource management, transportation, land use, agriculture and food)
Biosphere (including ecology, hydrology, oceanography, glaciology, soil science)
Societal impacts (forecasting, mitigation, and adaptation, for environmental extremes and hazards)
Environmental policy and economics

OPEN ACCESS CAMBRIDGE UNIVERSITY PRESS @envdatascience

Launched December 2020, open for submissions: www.cambridge.org/eds

Climate Change: Challenges for ML

[Banerjee & Monteleoni, Invited Tutorial, NeurIPS, 2014]

1. Past: Paleo-climate reconstruction

   What was the climate before we had thermometers?

2. Local: Climate downscaling

   What climate can I expect in my own backyard?

3. Future: Climate model ensembles

   How to reduce uncertainty on future predictions?

4. Spatiotemporal: Space and time

   How to capture dependencies over space and time?

5. Tails/impacts: Extreme events

   What are extreme events and how will climate change affect them?

6. Other problems

   Data-rich playground with many opportunities for ML to have an impact!

Today

- Focus on methods in Unsupervised Deep Learning
  - Along with semi-supervised and self-supervised variants

- Case-studies
  - Detection of avalanches
  - Downscaling of temperature and precipitation

Unsupervised Deep Learning

- Supervised DL. Prediction loss is a function of the label, y, and the network's output on input x.

  Network output

  $$f_W(x) = \hat{y}$$

  Loss function

  $$\mathcal{L}(\hat{y}, y)$$

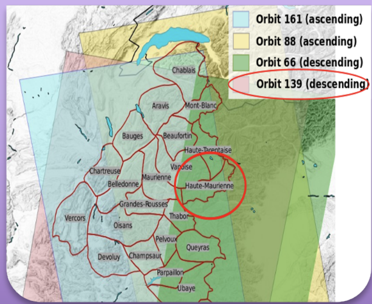
- Unsupervised DL. Prediction loss is only a function of x, and the network's output on input x. There is no label, y.

  Network output

  $$f_W(x) = \hat{x}$$

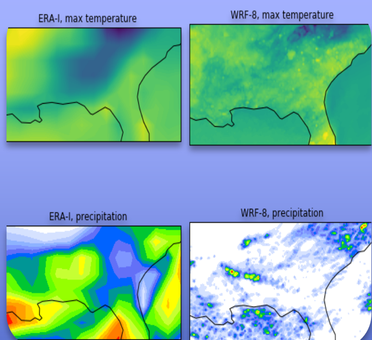
  Loss function

  $$\mathcal{L}(\hat{x}, x)$$



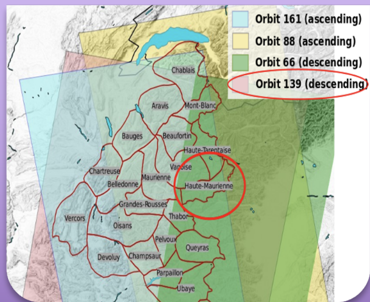
Semi-supervised DL

- Avalanche detection



Unsupervised DL

- Temp. and precip. downscaling



{Semi, Un}-supervised DL

- Avalanche detection

ML for extreme events

- How to define diverse, multivariate extreme events?
  - Unsupervised topic modeling approach

  [Tang & Monteleoni, CI 2014; IEEE CISE 2015]

- Hurricane track prediction

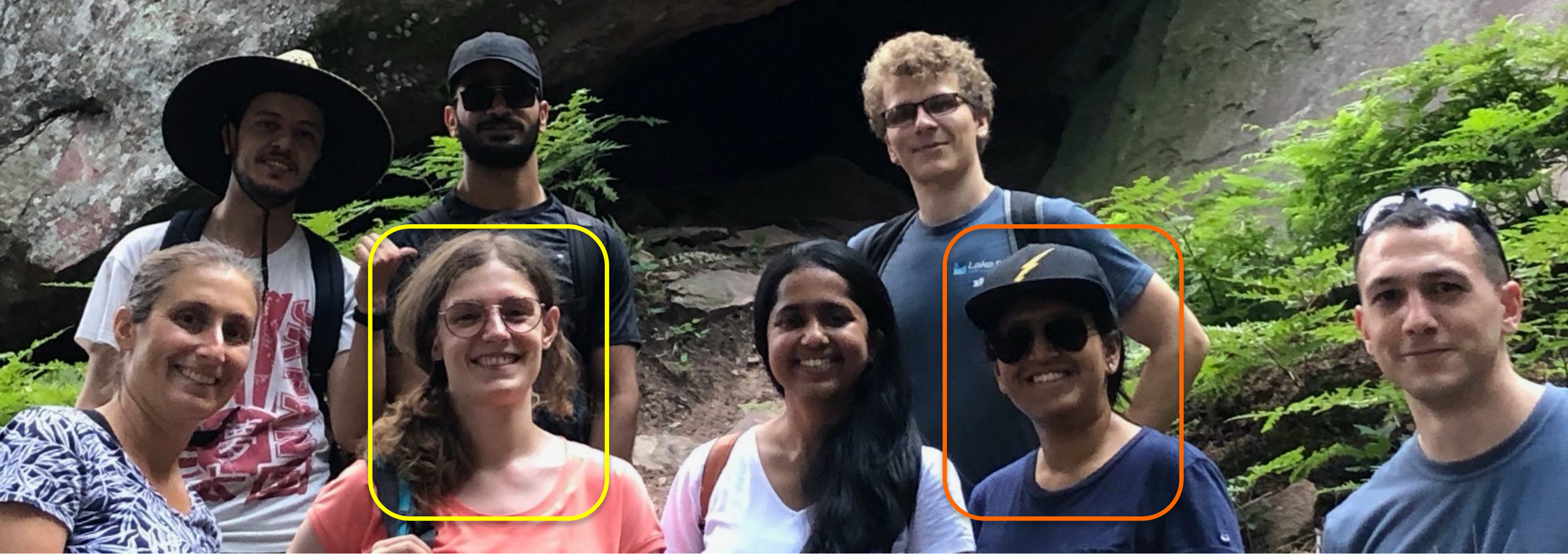
  [Giffard-Roisin et al., CI 2018; Frontiers 2020]

- Avalanche detection

  [Sinha et al., CI 2019; CI 2020]

Avalanche Detection

Img src : https://www.nytimes.com/2017/01/19/world/europe/italy-avalanche.html



Detecting Avalanche Deposits using VAE on Sentinel-1 Imagery

Saumya Sinha*, Sophie Giffard-Roisin*, Fatima Karbou, Michael Deschatres, Anna Karas, Nicolas Eckert, Cécile Coléou, Claire Monteleoni

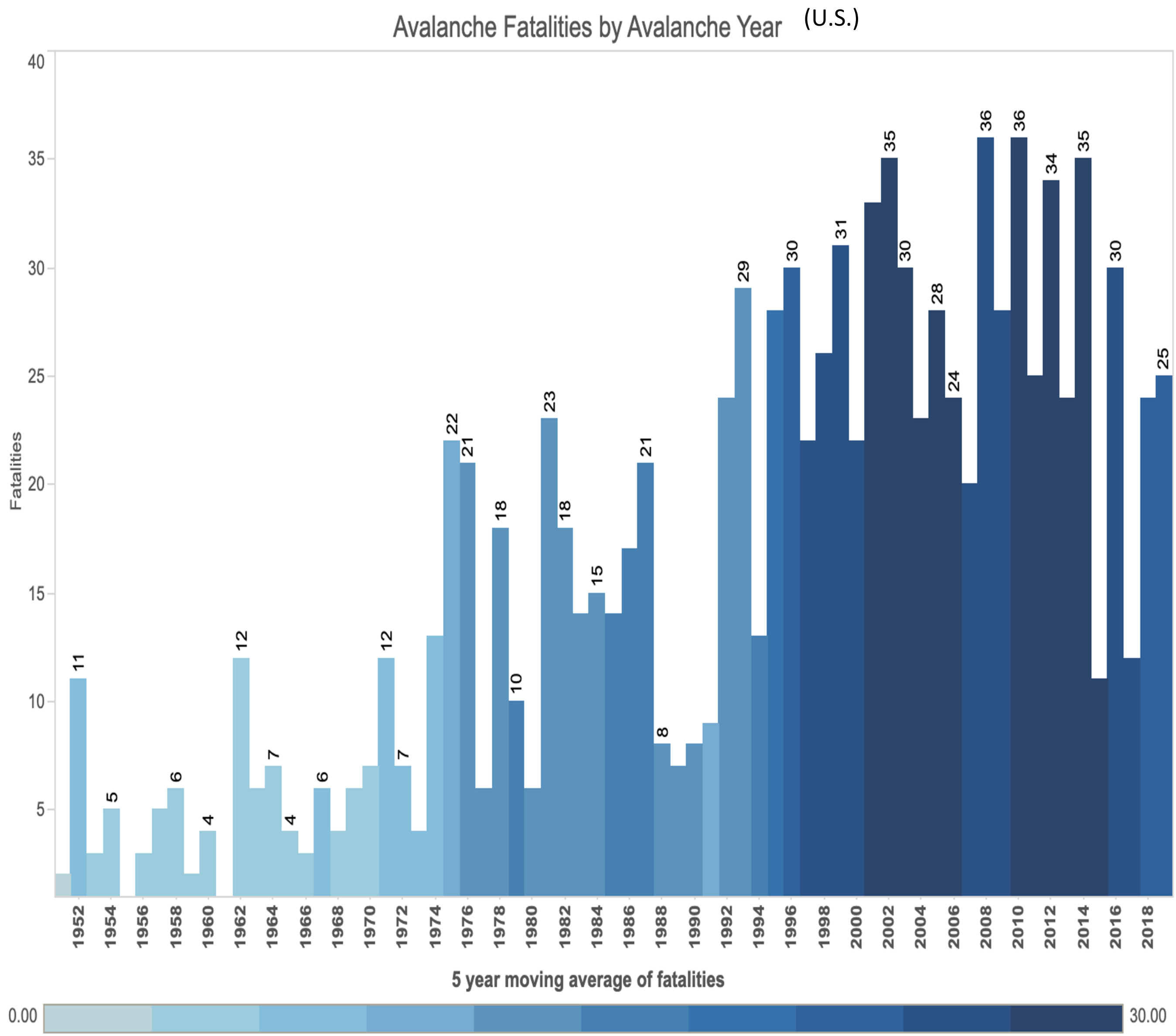
Variational autoencoder anomaly-detection of avalanche deposits in satellite SAR imagery. *Proc. 10th International Conference on Climate Informatics (CI) 2020*

Boulder

*Co-first-authorship

METEO FRANCE

Avalanche Fatalities by Avalanche Year



Img src : https://avalanche.state.co.us/accidents/statistics-and-reporting/

Climate change impacts

Img src: https://www.independent.co.uk/news/world/europe/mont-blanc-deaths-french-alps-avalanche-skiing-a9227461.html

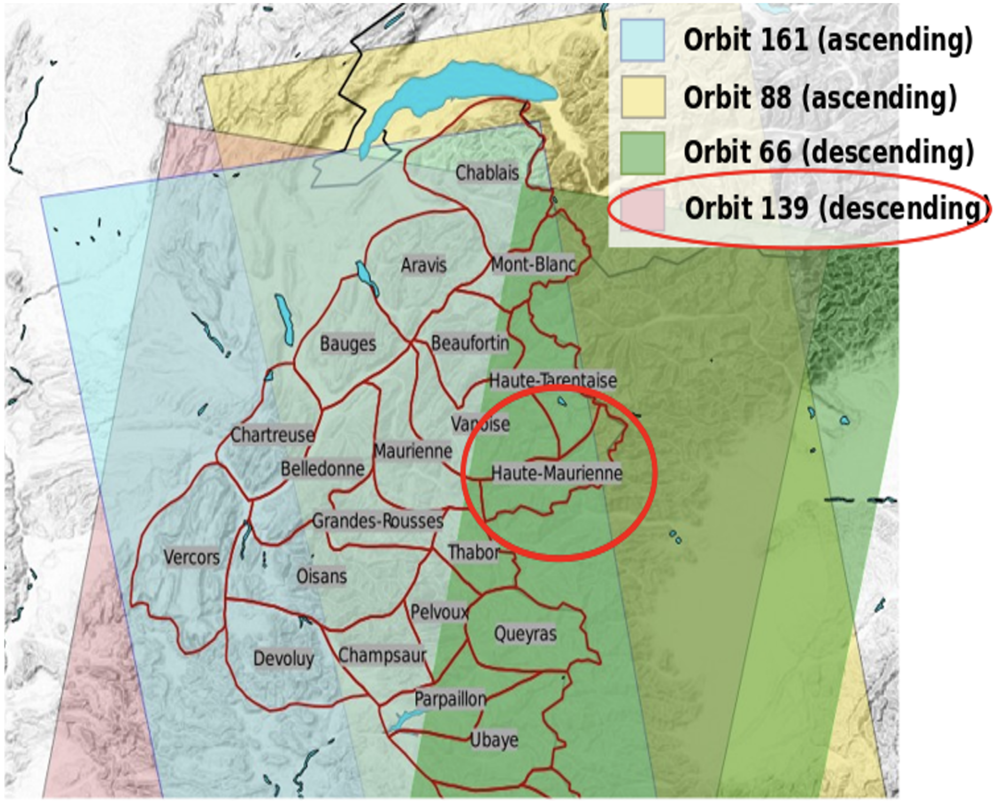
Challenges for Machine Learning

- Severe class imbalance
  - Avalanches are rare events
- Ground-truth labeled data difficult to obtain
  - Terrain accessibility
  - Weather conditions
  - Danger of avalanches

Avalanche detection problem

- **Limited** in-situ ground-truth measurements
  - Météo-France

- **Unlabeled** SAR imagery
  - Monitoring French Alps in 2017-2018
  - Sentinel-1A and 1B satellites
  - 4 features:
    - Backscatter coefficients at present and previous time
    - Topological features: Slope & Angle



Our Approach

① Treat an avalanche as a rare event, or an anomaly

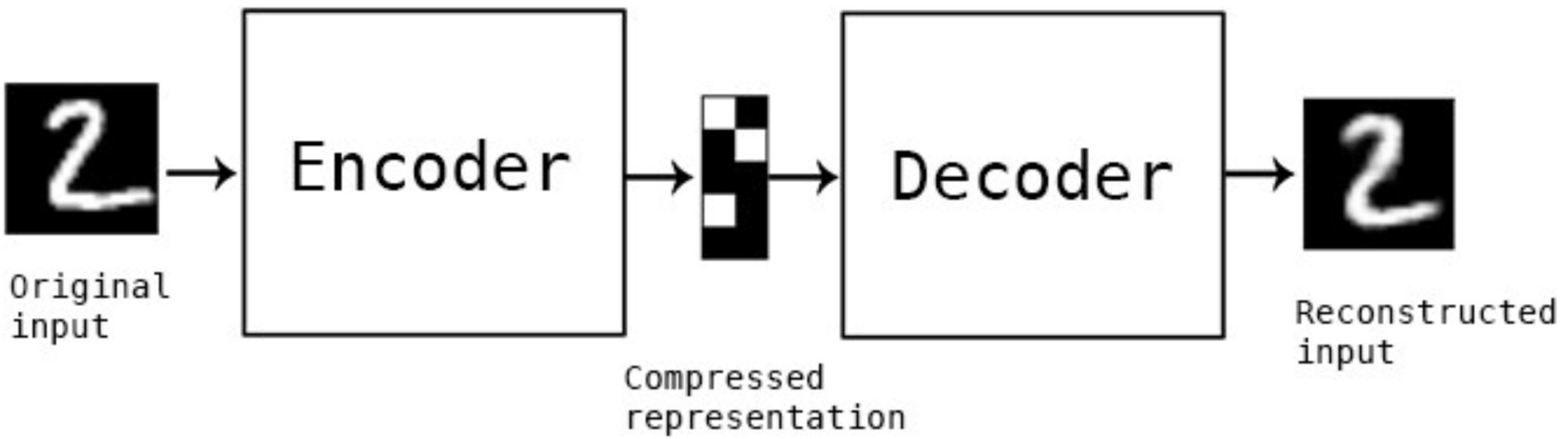
② Train a variational autoencoder (VAE) on the negative examples

③ Threshold the VAE's reconstruction error to classify a new image

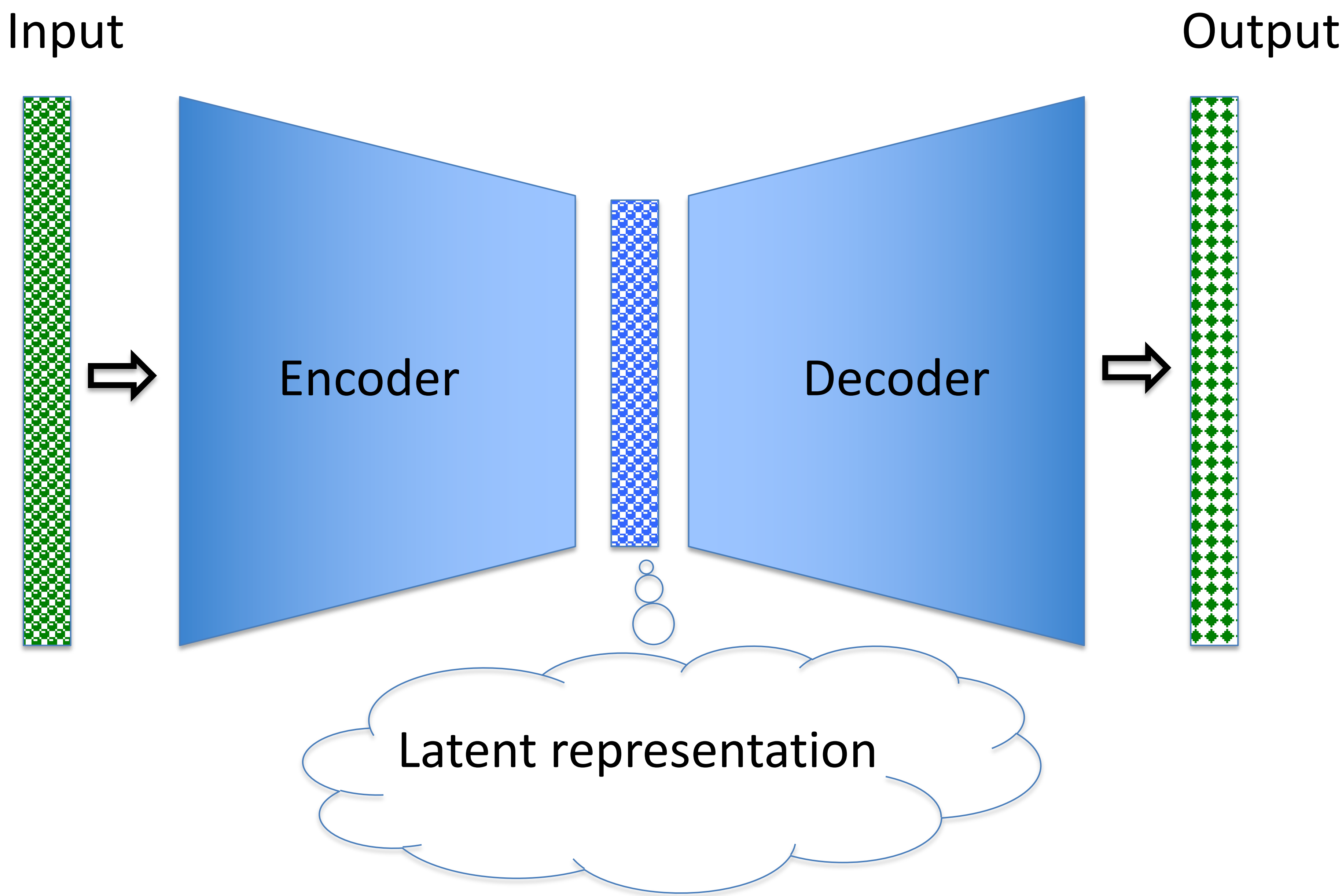
- When labeled data is scarce, the VAE can instead be trained **without** supervision!

What is an Auto-encoder?

- Train a neural network in an unsupervised setting
  - Use the unlabeled data both as input, and to evaluate the output
- After training, the bottleneck layer will be a compact representation of the input distribution

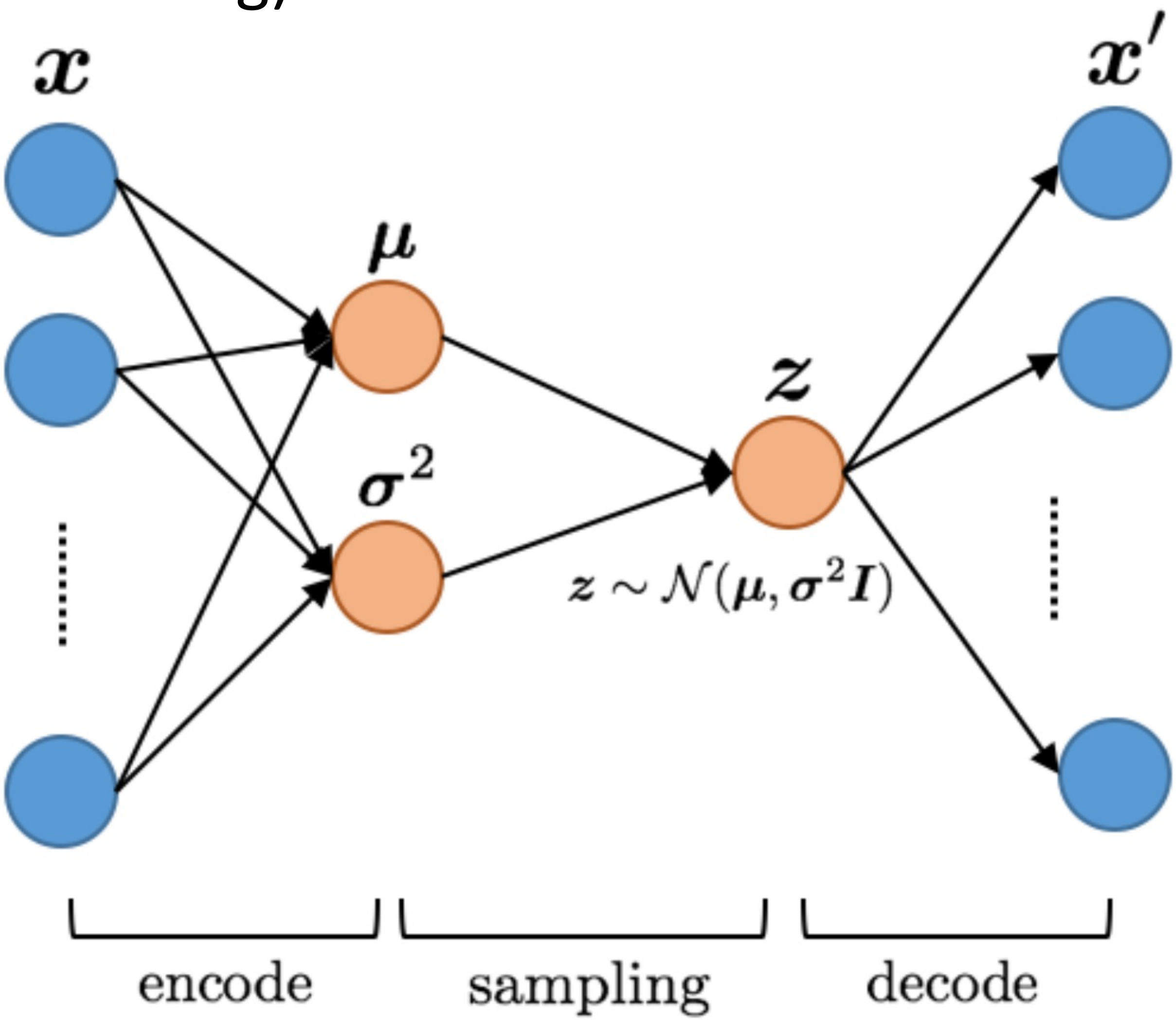


Autoencoder: The parameters of the encoder and decoder networks are trained to make the output approximate the input. After training on many input examples, the parameters of the bottleneck layer form a compact representation of the input distribution.

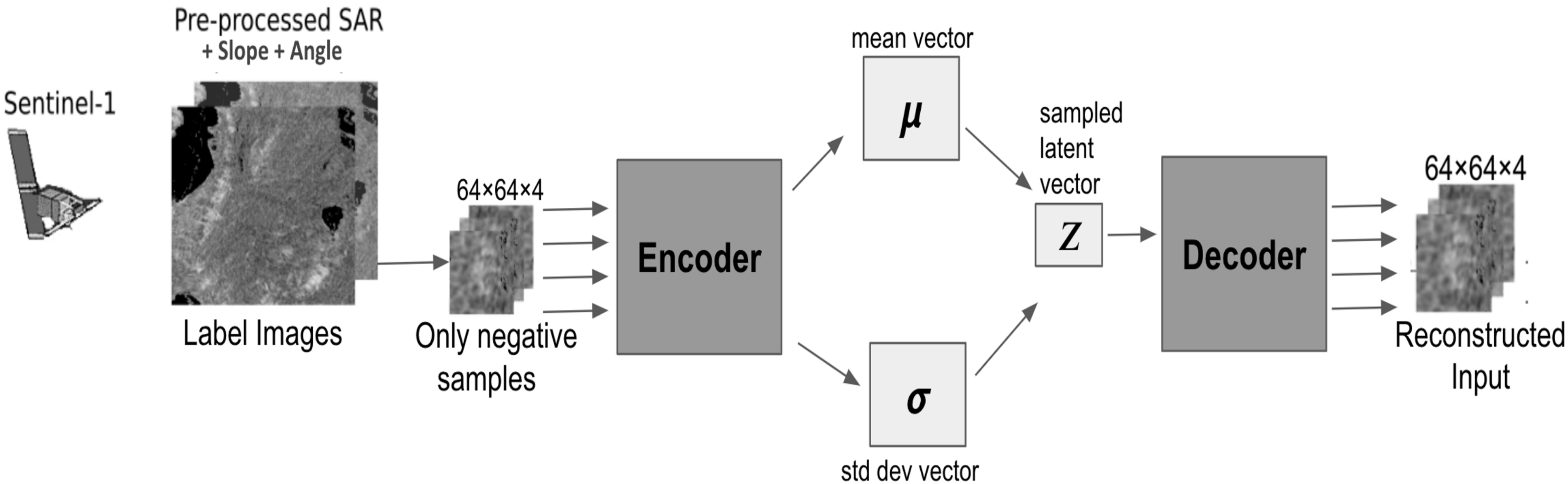


Variational Autoencoder (VAE)

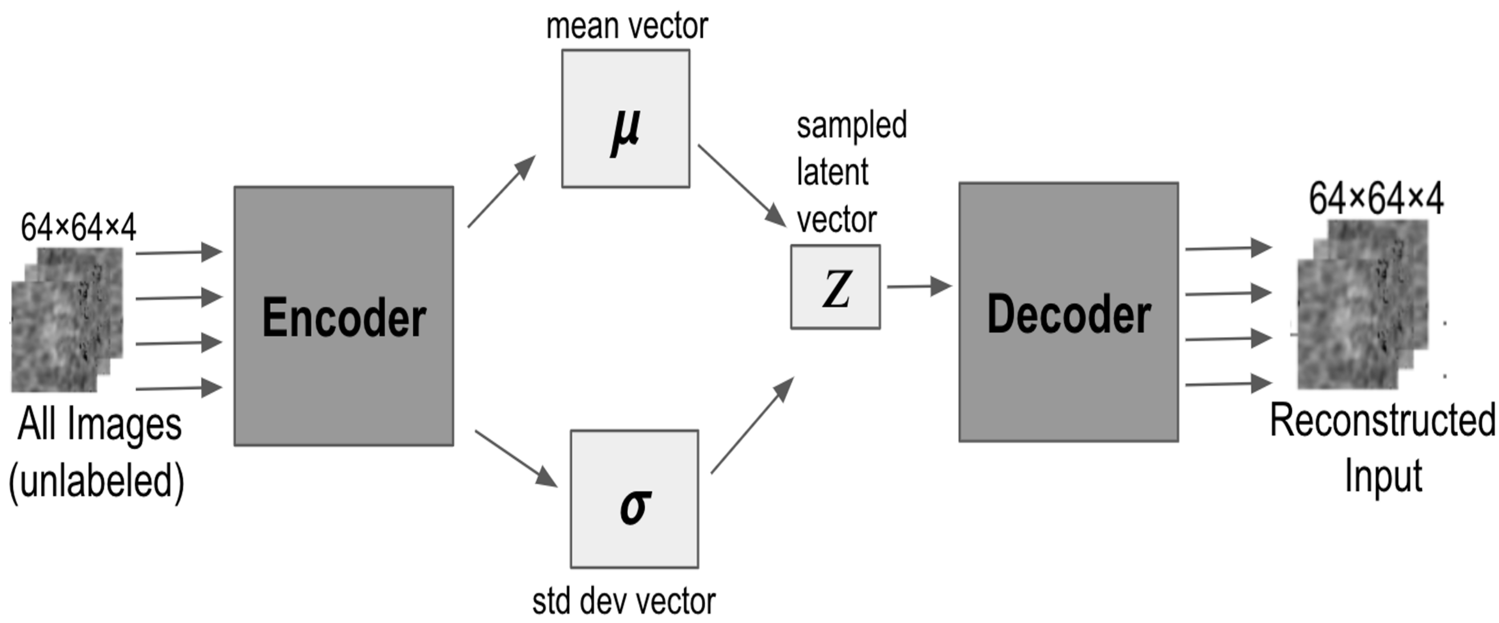
Learn a distribution over latent representations (instead of a single encoding).

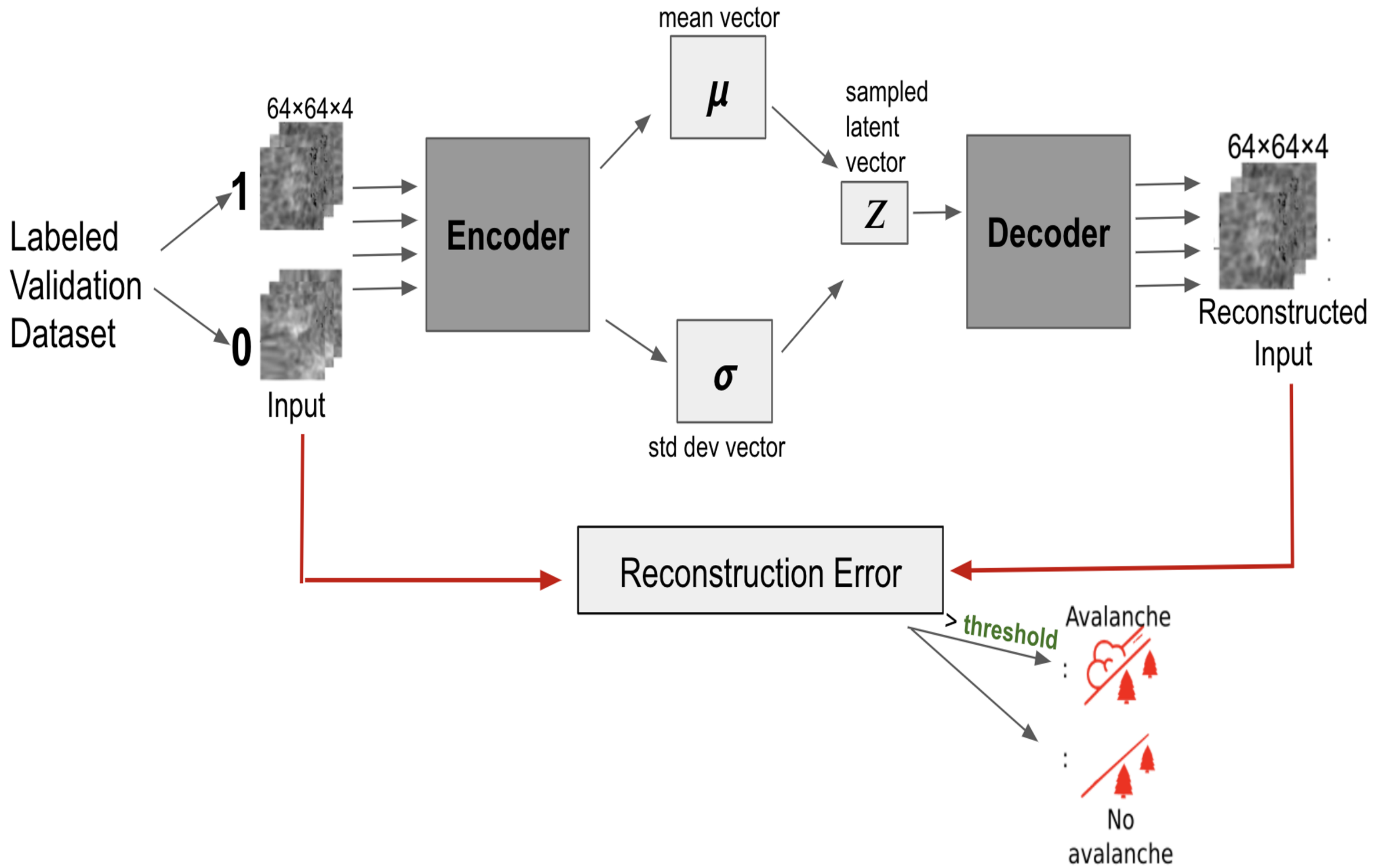


VAE trained on negative examples

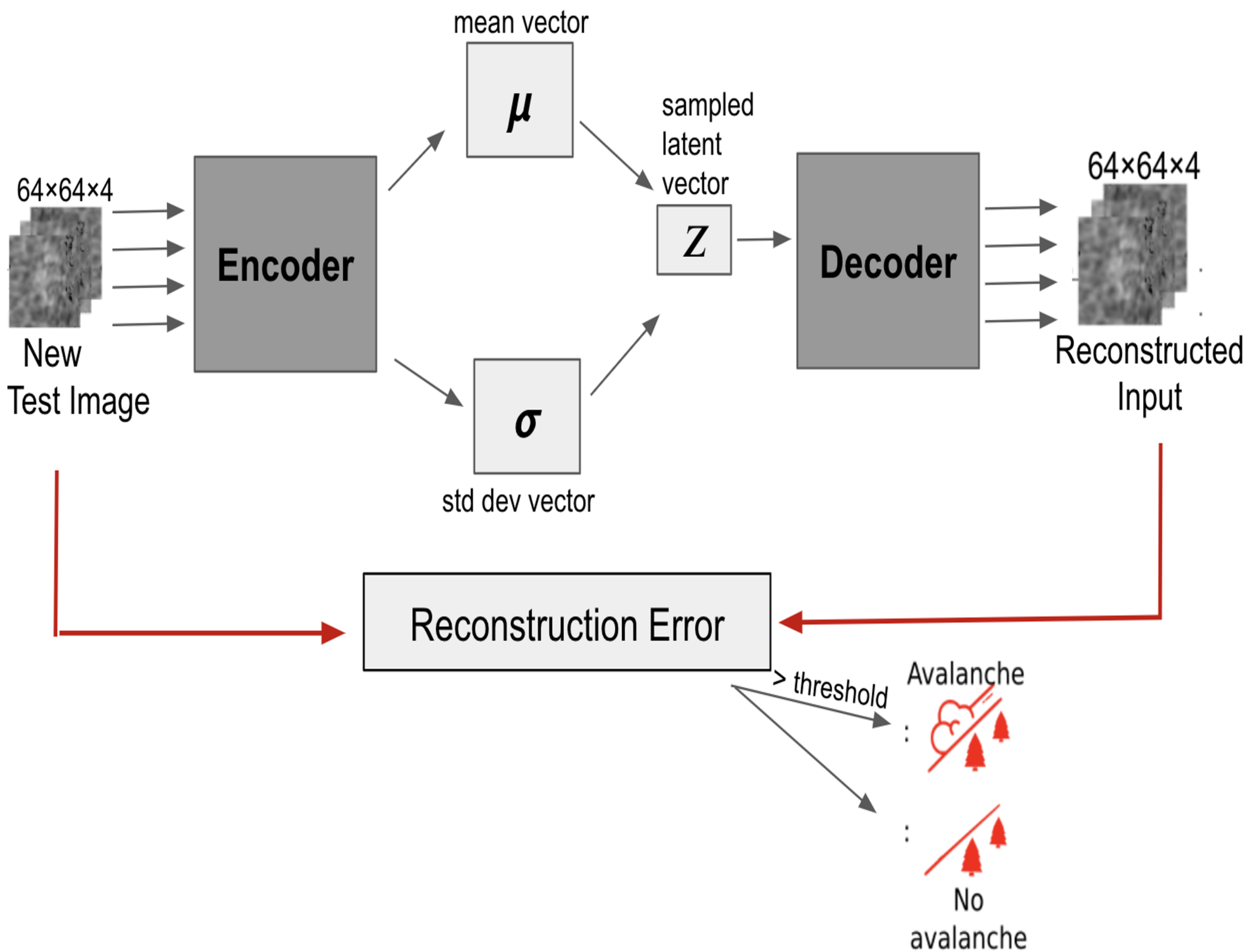


VAE trained on ALL examples



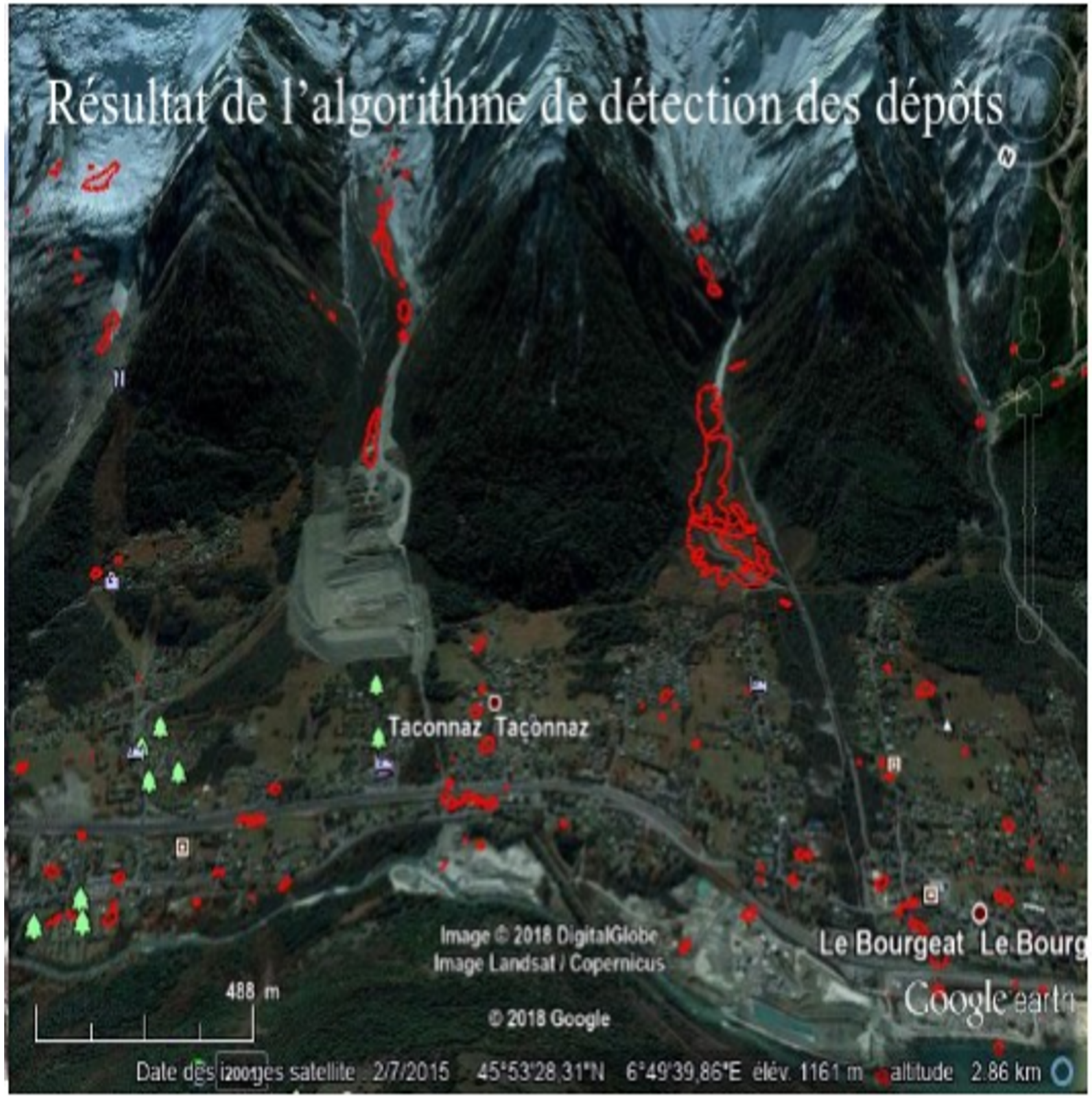


Tuning the threshold for Anomaly Detection



Anomaly Detection Pipeline

Baseline method: Thresholding



[Karbou et al., International Snow Science Workshop 2018; EGU 2018]

Evaluation

One of the most susceptible mountain chains (out of the 18 in “All Alps”)

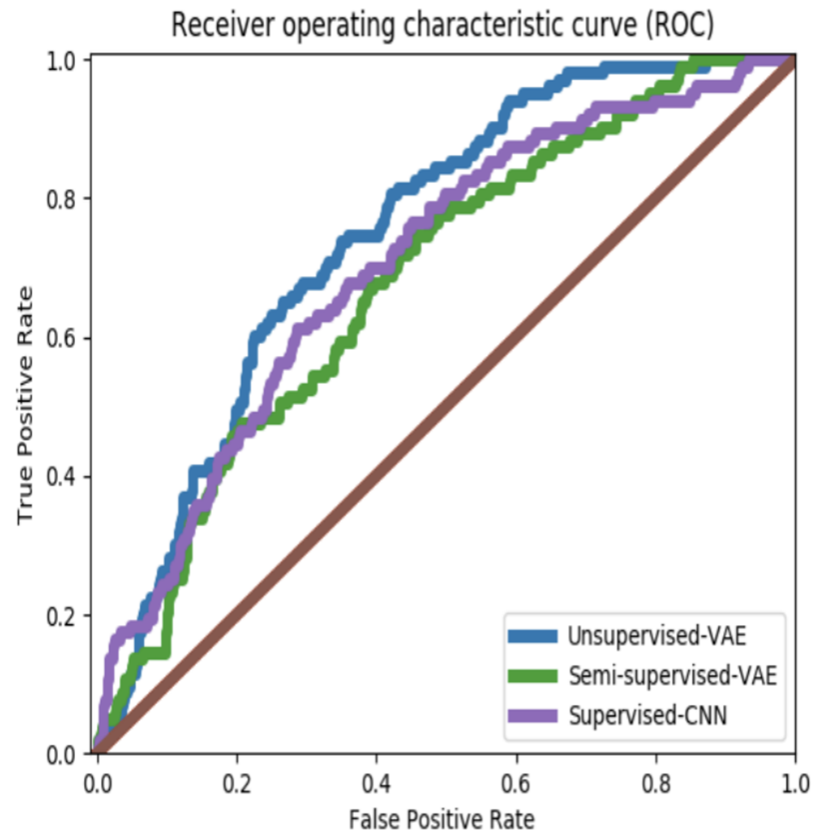
	All Alps		Haute Maurienne	
	Balanced Accuracy	**F1-score**	**Balanced Accuracy**	**F1-score**
Baseline	0.58	0.05	0.58	0.12
Supervised - CNN	0.53	0.10	0.53	0.12
Semi-supervised - VAE	0.59	0.11	0.6	0.23
Unsupervised - VAE	0.69	0.14	0.68	0.26

- Held-out test set: 6,498 labeled examples
- Baseline method from avalanche-detection literature: Thresholding [Karbou et al., ISSW 2018; EGU 2018]
- Supervised-learning benchmark method: Convolutional Neural Network (CNN) trained on artificially balanced dataset [Sinha et al., Climate Informatics 2019]

Evaluation

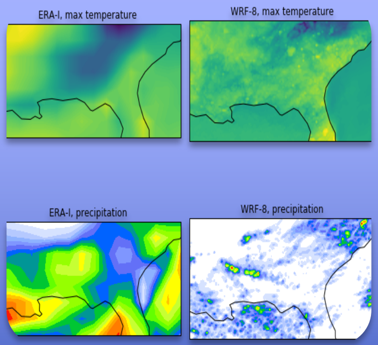
Method	AUC ROC
Supervised - CNN	70.7
Semi-supervised - VAE	68.3
Unsupervised - VAE	75

ROC Curves for Haute Maurienne region



Outlook

- Provided a semi-supervised approach to detecting rare events when labeled data is limited
- Can be viewed as a form of virtual sensor
- Next step: forecasting



{Un, Self}-supervised DL

- Temp. and precip. downscaling

Unsupervised DL for Downscaling

Brian Groenke's Masters Thesis, CU Boulder, May 2020

with help from Luke Madaus, Jupiter Intelligence

- Downscaling: Classic problem in climate & meteorology
  - Goal: use coarse-scale spatiotemporal data to infer values at finer scales
- Field of statistical downscaling, existing work:
  - Supervised learning methods
  - Provide point predictions
- Generative downscaling is largely open

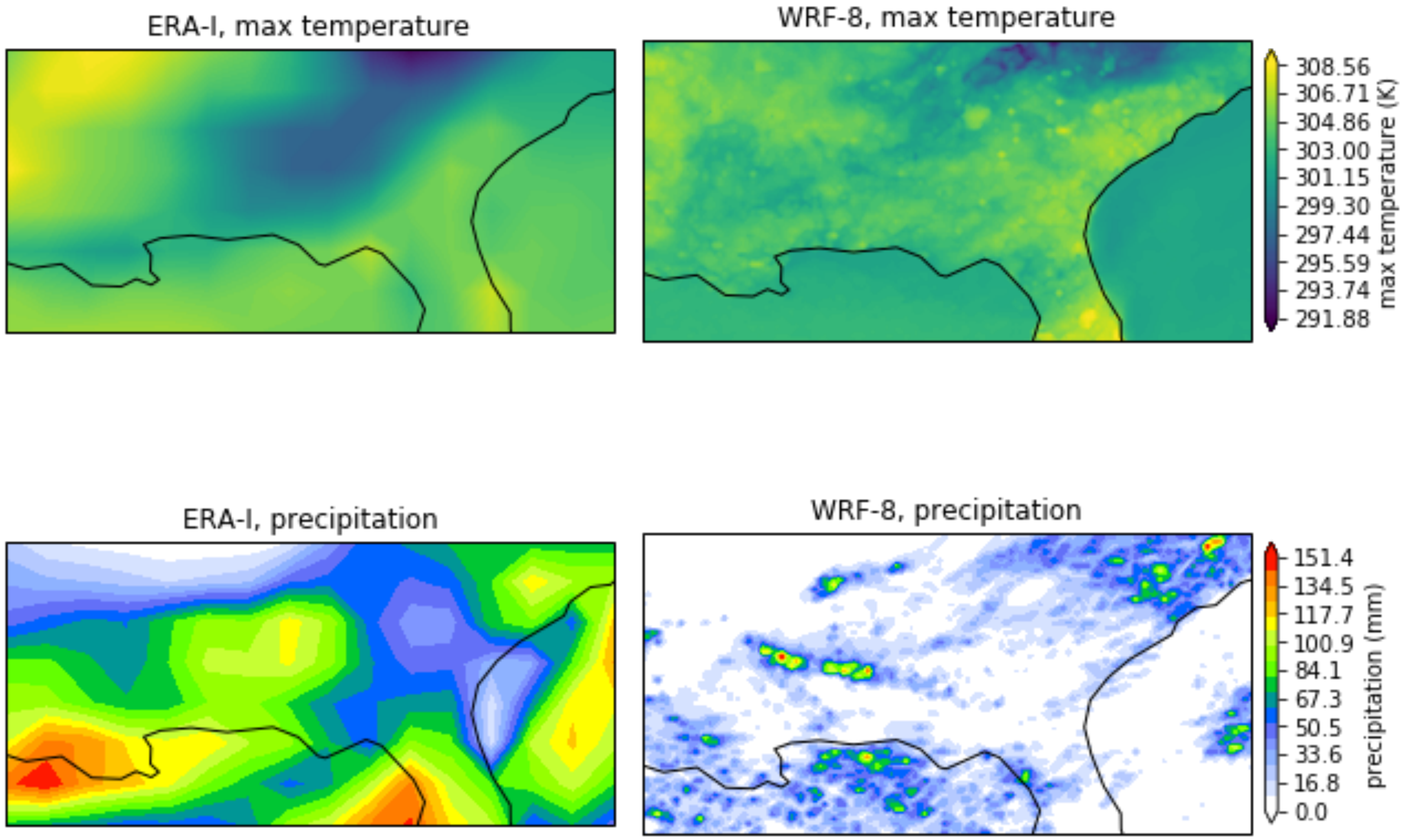
Unsupervised DL for Downscaling

[Groenke, Madaus & Monteleoni. Climalign: Unsupervised statistical downscaling of climate variables via normalizing flows. *Proc. 10th International Conference on Climate Informatics (CI) 2020*]

- Cast downscaling as the ML task of domain alignment
- Extend deep unsupervised domain alignment
  - AlignFlow [Grover et al., AAAI 2020]
  - Glow normalizing flow [Kingma & Dhariwal, NeurIPS 2018]
  - Self-supervision via geographic alignment of both domains
- Obtain generative model for downscaling

Downscaling: training data

ERA: reanalysis data, 1° resolution; WRF: numerical weather model prediction, $\frac{1}{8}$° resolution



Downscaling as domain alignment

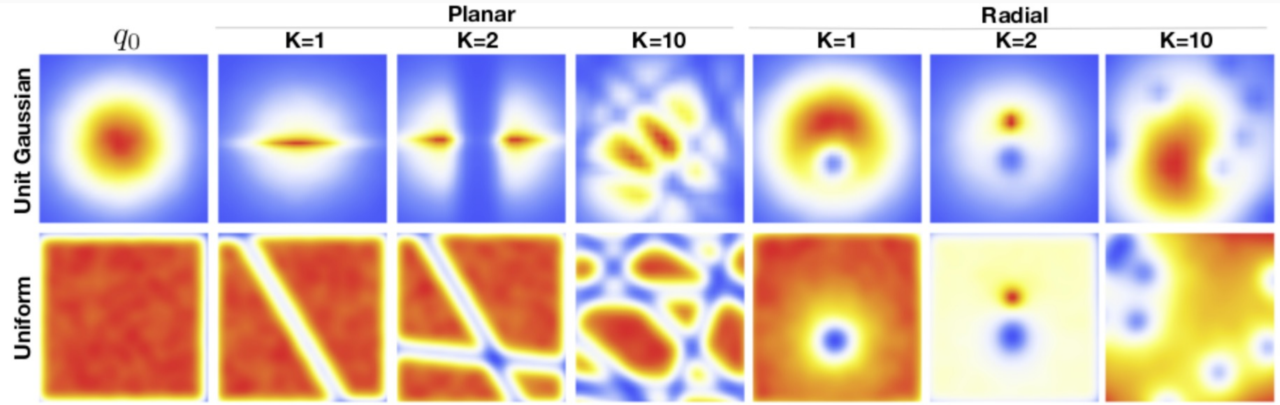
- Domain alignment task: given random variables X, Y, learn a mapping f: X → Y such that, for any $x_i \in X$ and $y_i \in Y$,
  $f(x_i) \sim P_Y$ and $f^{-1}(y_i) \sim P_X$

- Downscaling as domain alignment
  - Learn the joint PDF over X and Y, by assuming conditional independence over a shared latent space Z

$$P_{XY}(x,y) = \int_{z \in Z} P_{XYZ}(x,y,z)dz = \int_{z \in Z} P(x|z)P(y|z)P_Z(z)dz$$

  - Model $P(x|z), P(y|z)$ using AlignFlow [Grover et al. 2020]
  - Starting with a simple prior on P_Z, learn normalizing flows
  - No pairing between x and y examples needed!

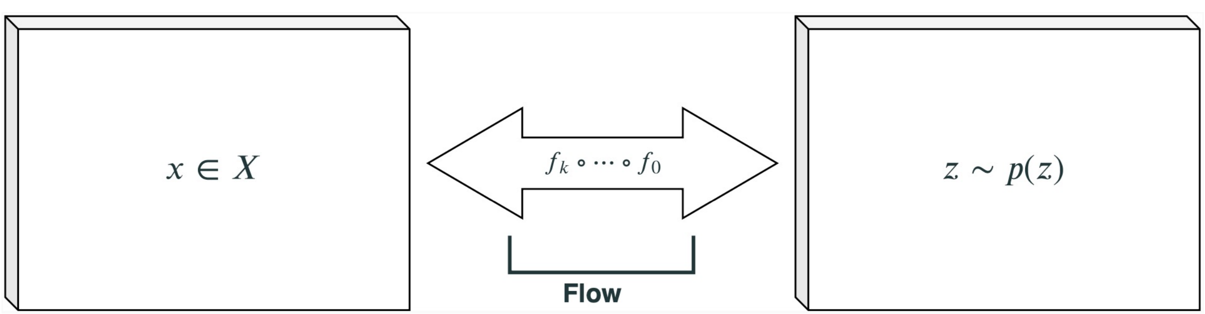
Normalizing Flows



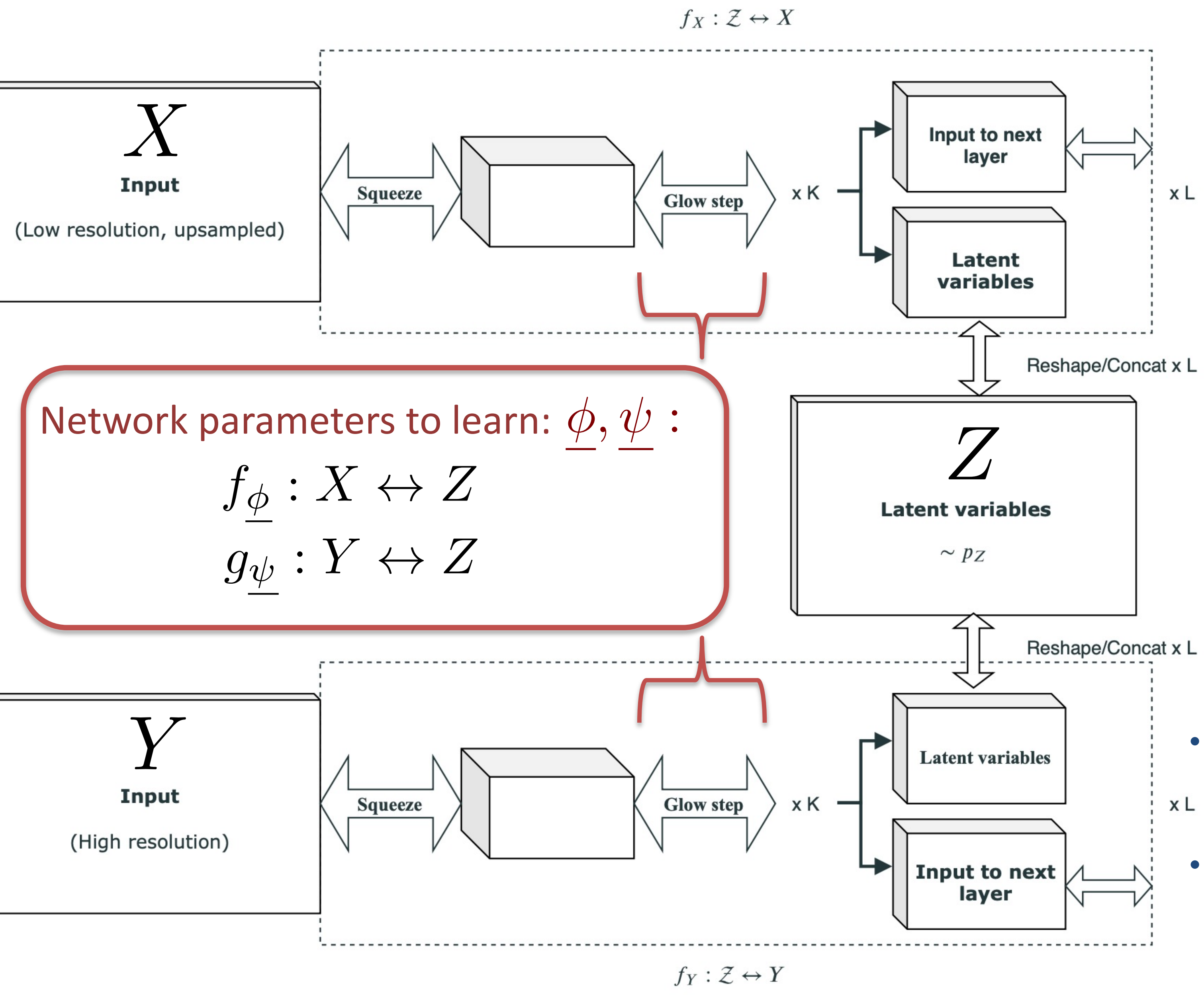
[Rezende & Mohamed, 2015]

Learn a series of invertible transformations, $\{f_i\}$, from a simple prior on Z, to allow for more informative distributions on the latent space:

$$z_k = f_k \circ f_{k-1} \circ \cdots \circ f_1(z_0)$$



ClimAlign architecture



Network parameters to learn: $\underline{\phi}, \underline{\psi}$:

$$f_{\underline{\phi}} : X \leftrightarrow Z$$
$$g_{\underline{\psi}} : Y \leftrightarrow Z$$

- Architecture follows AlignFlow [Grover et al., 2020]
- Normalizing flow: Glow [Kingma & Dhariwal, 2018]

Comparison with supervised benchmarks

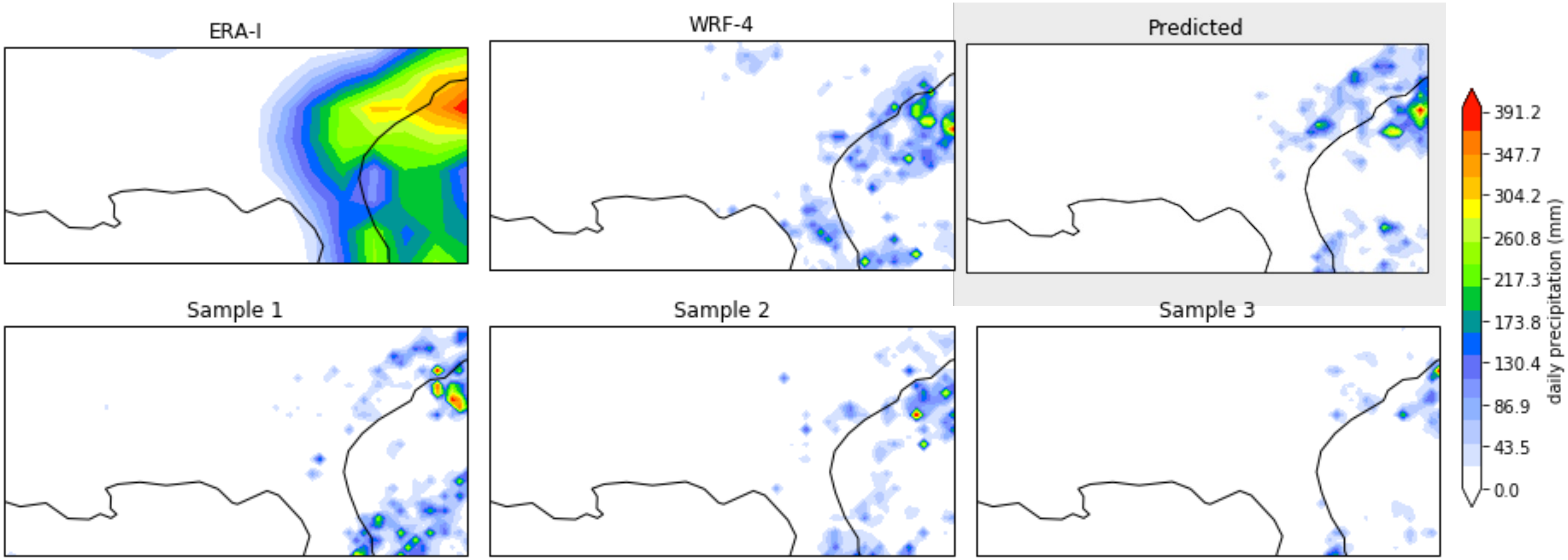
Daily Max Temperature

Region	Method	RMSE	Bias	Corr
SE-US	BCSD	1.51 $\pm$ 0.15	-0.02 $\pm$ 0.21	**0.93** $\pm$ 0.05
	BMD-CNN	**1.30** $\pm$ 0.12	0.03 $\pm$ 0.13	0.90 $\pm$ 0.05
	ClimAlign (ours)	1.56 $\pm$ 0.13	**-0.005** $\pm$ 0.22	0.87 $\pm$ 0.06
P-NW	BCSD	1.54 $\pm$ 0.23	**0.01** $\pm$ 0.10	**0.95** $\pm$ 0.03
	BMD-CNN	**1.25** $\pm$ 0.14	-0.06 $\pm$ 0.05	0.93 $\pm$ 0.02
	ClimAlign (ours)	1.58 $\pm$ 0.18	0.03 $\pm$ 0.15	0.89 $\pm$ 0.04

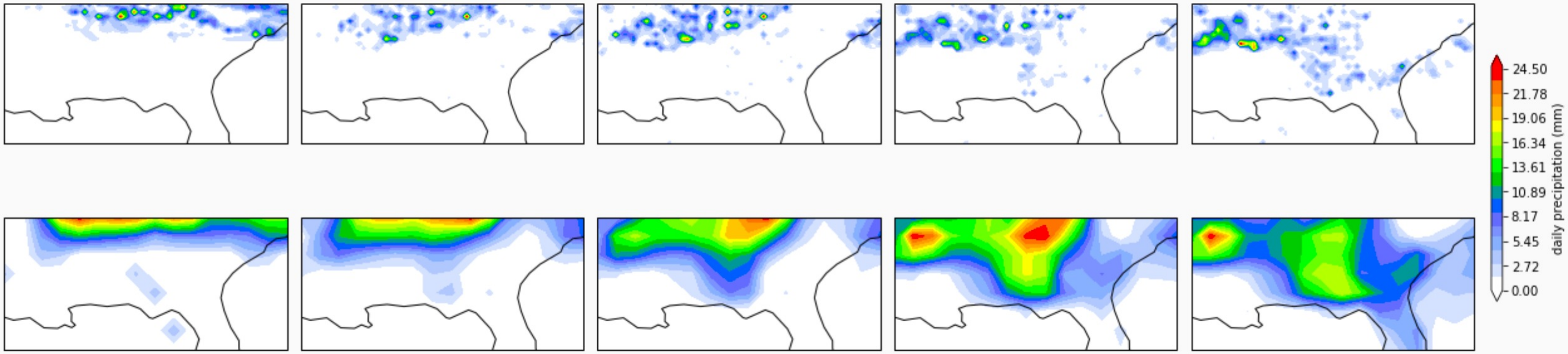
Daily Precipitation

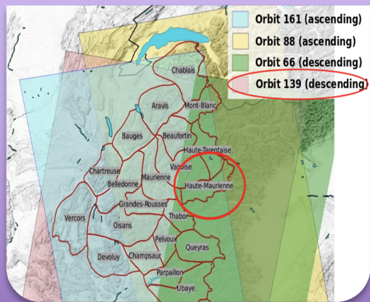
Region	Method	RMSE	Bias	Corr
SE-US	BCSD	27.32 $\pm$ 5.0	0.95 $\pm$ 1.4	0.39 $\pm$ 0.07
	BMD-CNN	**14.11** $\pm$ 2.18	-0.23 $\pm$ 0.47	**0.50** $\pm$ 0.10
	ClimAlign (ours)	18.40 $\pm$ 2.64	**0.08** $\pm$ 0.86	0.42 $\pm$ 0.07
P-NW	BCSD	8.90 $\pm$ 2.30	0.41 $\pm$ 0.26	0.61 $\pm$ 0.06
	BMD-CNN	**5.77** $\pm$ 0.72	**-0.18** $\pm$ 0.61	**0.70** $\pm$ 0.03
	ClimAlign (ours)	7.33 $\pm$ 0.69	0.54 $\pm$ 0.54	0.67 $\pm$ 0.03

Point prediction example



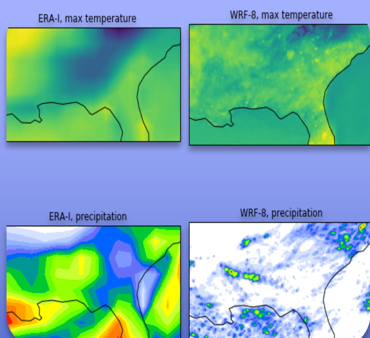
Interpolation example





Semi-supervised DL

- Avalanche detection



Unsupervised DL

- Temp. and precip. downscaling

Challenges/Bottlenecks in Climate Informatics

- Data challenges
  - Limited labeled data: unsupervised learning, dimensionality reduction
  - Class imbalance: e.g., extreme events are rare by definition!
  - Data is limited along the time dimension. Can we substitute data diversity and granularity over space?
  - Spatiotemporal data: self-supervised learning of spatiotemporal features?
- Scale resolution challenges
  - Downscaling spatiotemporal data fields
  - Climate model parameterization problems
- Non-stationarity
  - Climate *change* means we cannot assume i.i.d. data!
  - ML models need to adapt over time, and space
- Interpretability
  - Communities need trustworthy ML models and forecasts

Thank you!

Many thanks to:

Arindam Banerjee, *University of Illinois Urbana-Champaign*
Cécile Coléou, *Météo-France & CNRS*
Michael Dechartre, *Irstea, Université Grenoble Alpes*
Nicolas Eckert, *Irstea, Université Grenoble Alpes*
Sophie Giffard-Roisin, *IRD Grenoble*
Brian Groenke, *Alfred Wegener Institute, Potsdam*
Anna Karas, *Météo-France & CNRS*
Fatima Karbou, *Météo-France & CNRS*
Luke Madaus, *Jupiter Intelligence*
Saumya Sinha, *University of Colorado Boulder*

AI for Earth grants
Join hundreds of grantees and empower your work with an AI for Earth grant.
Microsoft | AI

Resources

- Climate Informatics: www.climateinformatics.org
  - Conferences, hackathons, resources
- Climate Informatics News: Google Group
  - Announcements, job listings, events, community
- The 10th International Conference on Climate Informatics (CI 2020)
  - Videos: https://slideslive.com/climateinformatics/climate-informatics-2020
  - Proceedings: https://dl.acm.org/doi/proceedings/10.1145/3429309
- ClimAlign code: github.com/bgroenks96/generative-downscaling
- Climate Informatics Hackathon: storm intensity forecasting

  github.com/ramp-kits/storm_forecast