NeurIPS | 2023
Thirty-seventh Conference on Neural Information Processing Systems

UCLA

ClimateLearn: Benchmarking Machine Learning for Weather and Climate Modeling


Tung Nguyen


Jason Jewik

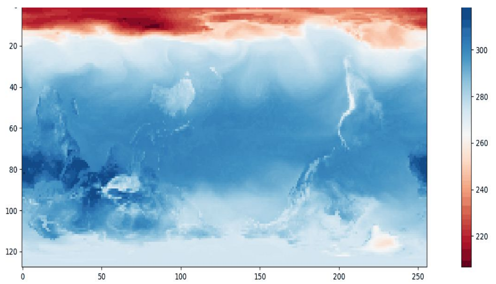

Hritik Bansal


Prakhar Sharma

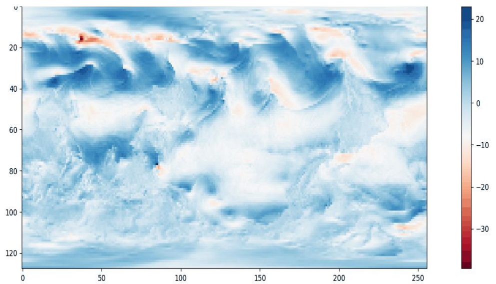

Aditya Grover

Background and motivation

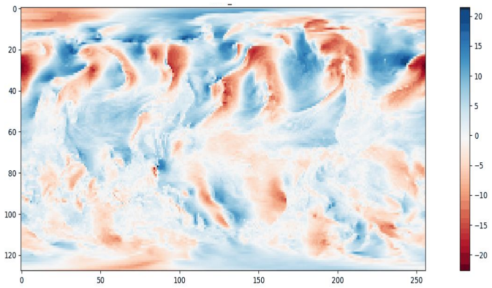
- ❖ Climate modeling is fundamental in understanding atmospheric, oceanic, and land processes
  - ➢ Short-term weather forecasts: given current weather conditions, predict weather a few days later
  - ➢ Long-term climate projections: predict average condition changes over annual/decadal scales
- ❖ Current applications of machine learning to climate modeling
  - ➢ Lack of standardization (*e.g.*, train-test splits, data augmentation)
  - ➢ Few end-to-end approaches



Temperature

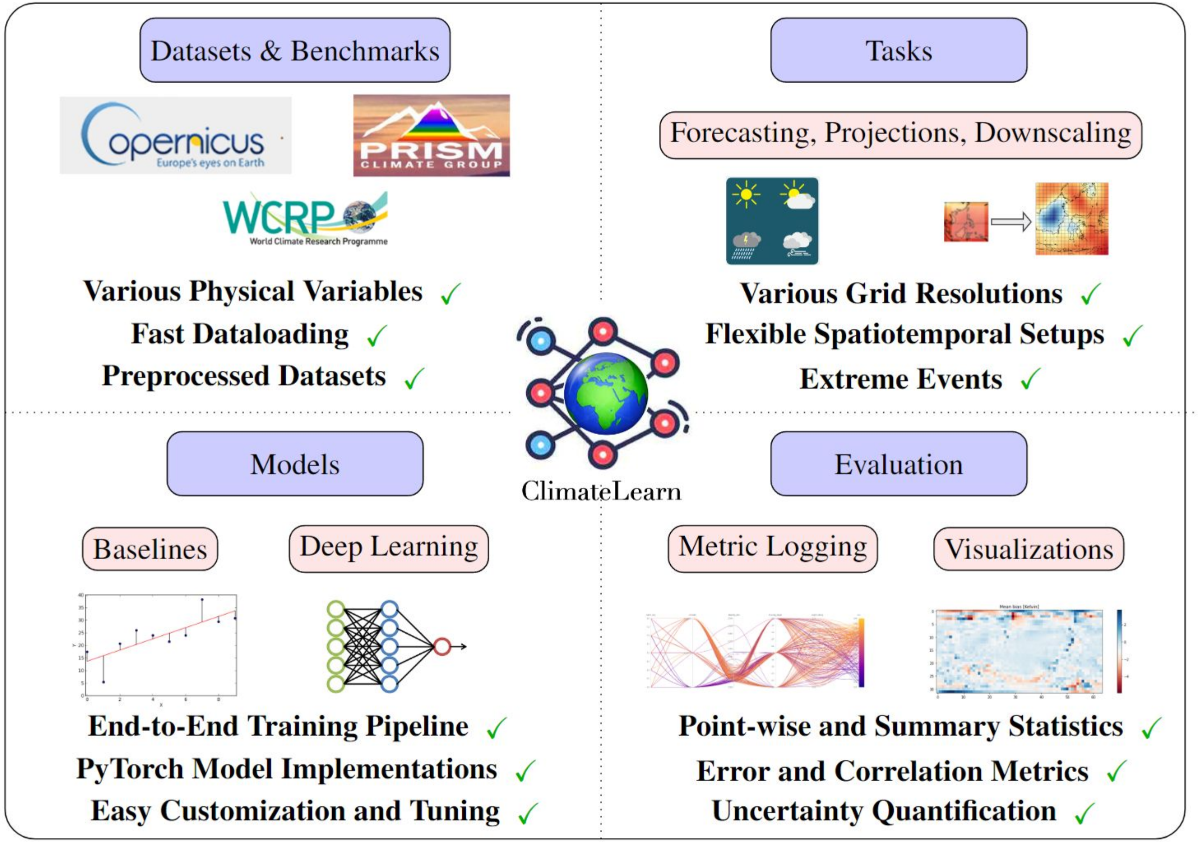


U-component of wind

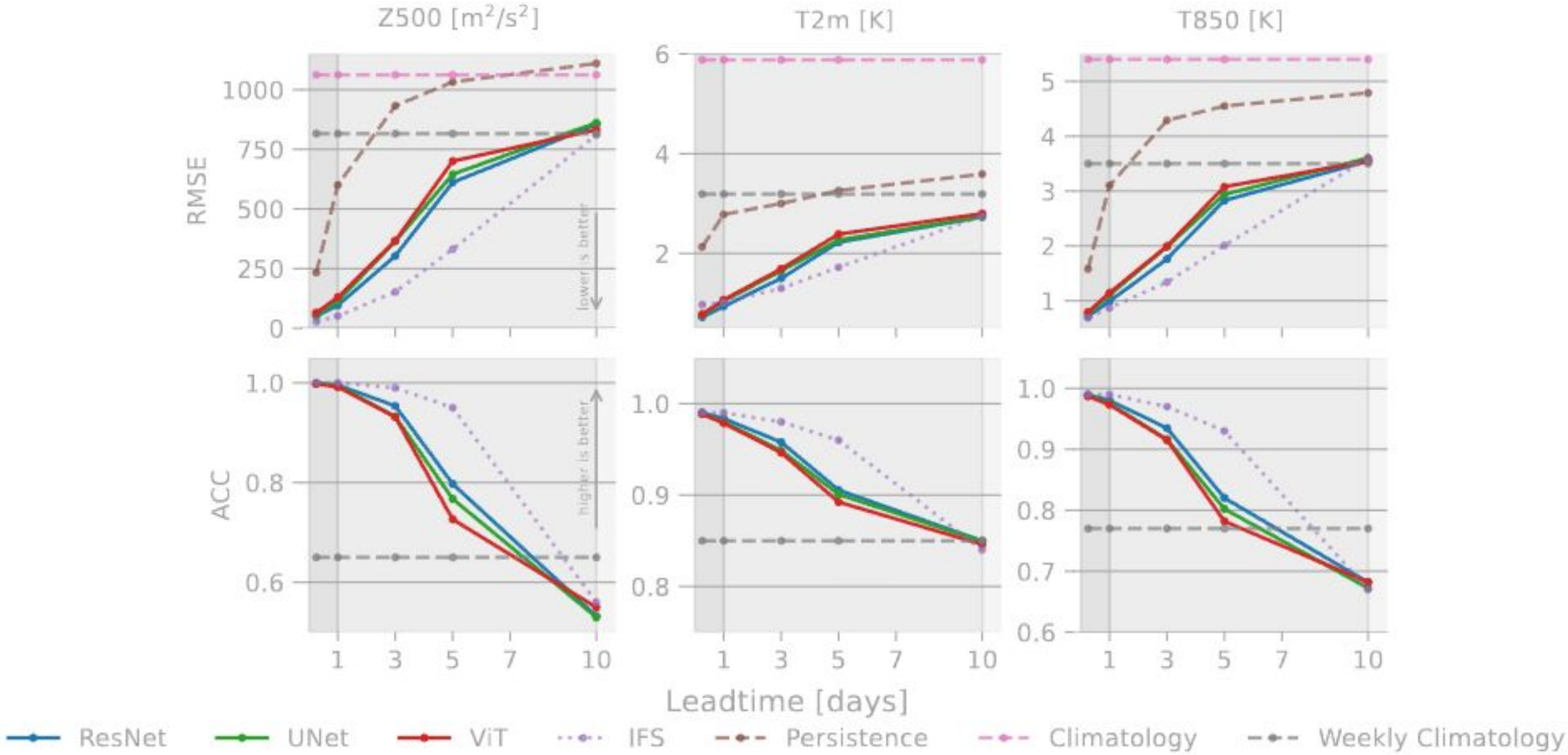


V-component of wind

Key features of our library



Comparison of simple baselines and deep learning methods



Example: Downloading Data

```
import climate_learn as cl

cl.data.download_weatherbench(
    dst="/home/user/climate-learn/geopotential",
    dataset="era5",
    variable="geopotential",
    resolution=5.625  # optional, default is 5.625
)
```

Example: Preprocessing Data

```python
from climate_learn.data.processing.nc2npz import convert_nc2npz

convert_nc2npz(
    root_dir="/home/user/climate-learn",
    save_dir="/home/user/climate-learn/processed",
    variables=["geopotential"],
    start_train_year=1979,
    start_val_year=2015,
    start_test_year=2017,
    end_year=2018,
    num_shards=16
)
```

Example: Loading Data

```
dm = cl.data.IterDataModule(
    task="direct-forecasting",
    inp_root_dir="/home/user/climate-learn/processed",
    out_root_dir="/home/user/climate-learn/processed",
    in_vars=["geopotential"],
    out_vars=["geopotential"],
    src="era5",
    history=3,
    window=6,
    pred_range=72,
    subsample=6,
    batch_size=128,
    num_workers=8,
)
```

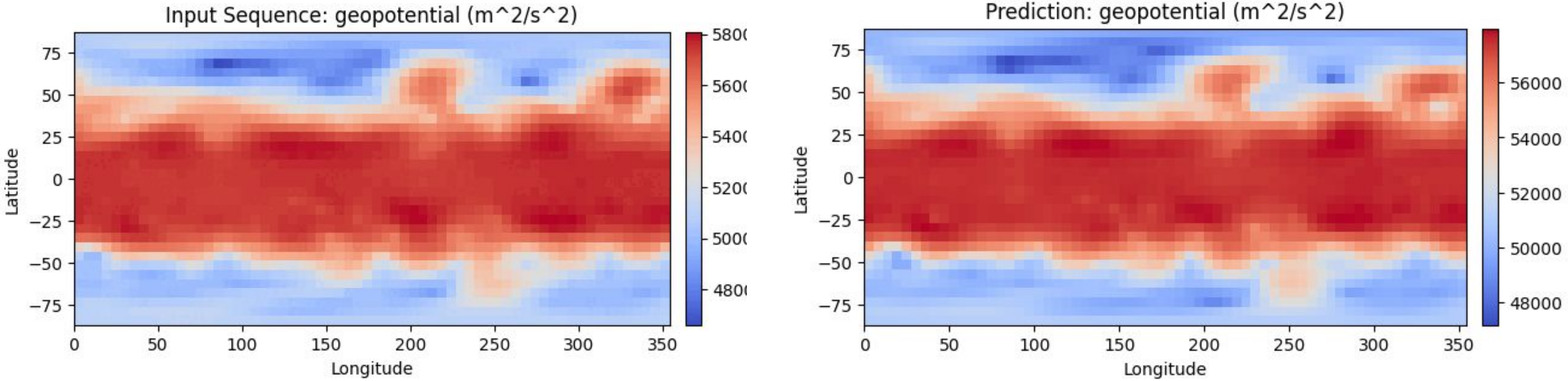
Example: Instantiating a Model

```
model = cl.load_forecasting_module(
    data_module=dm,
    model="resnet",
    model_kwargs={"n_blocks": 4, "history": 5},
    optim="adamw",
    optim_kwargs={"lr": 5e-4},
    sched="linear-warmup-cosine-annealing",
    sched_kwargs={"warmup_epochs": 5, "max_epochs": 50}
)
```

Example: Visualizing inputs and outputs

```
denorm = model.test_target_transforms[0]
in_graphic = cl.utils.visualize_at_index(
    model,
    dm,
    in_transform=denorm,
    out_transform=denorm,
    variable="geopotential",
    src="era5",
    index=0
)
HTML(in_graphic.to_jshtml())
```

Example: Visualizing inputs and outputs (cont'd)



Output of the previous slide's code.

Resources

- Github: https://github.com/aditya-grover/climate-learn
- Quickstart Colab: https://colab.research.google.com/drive/1LcecQLgLtwaHOwbvJAxw9UjCxfM0RMrX?usp=sharing
- Documentation: https://climatelearn.readthedocs.io/en/latest
- Future Work:
  - Expanding catalog of data sources
  - Building a hub of pre-trained deep neural networks for climate modeling problems
  - Incorporating physics-informed and other hybrid neural network architectures