CLIMATEHACK.AI

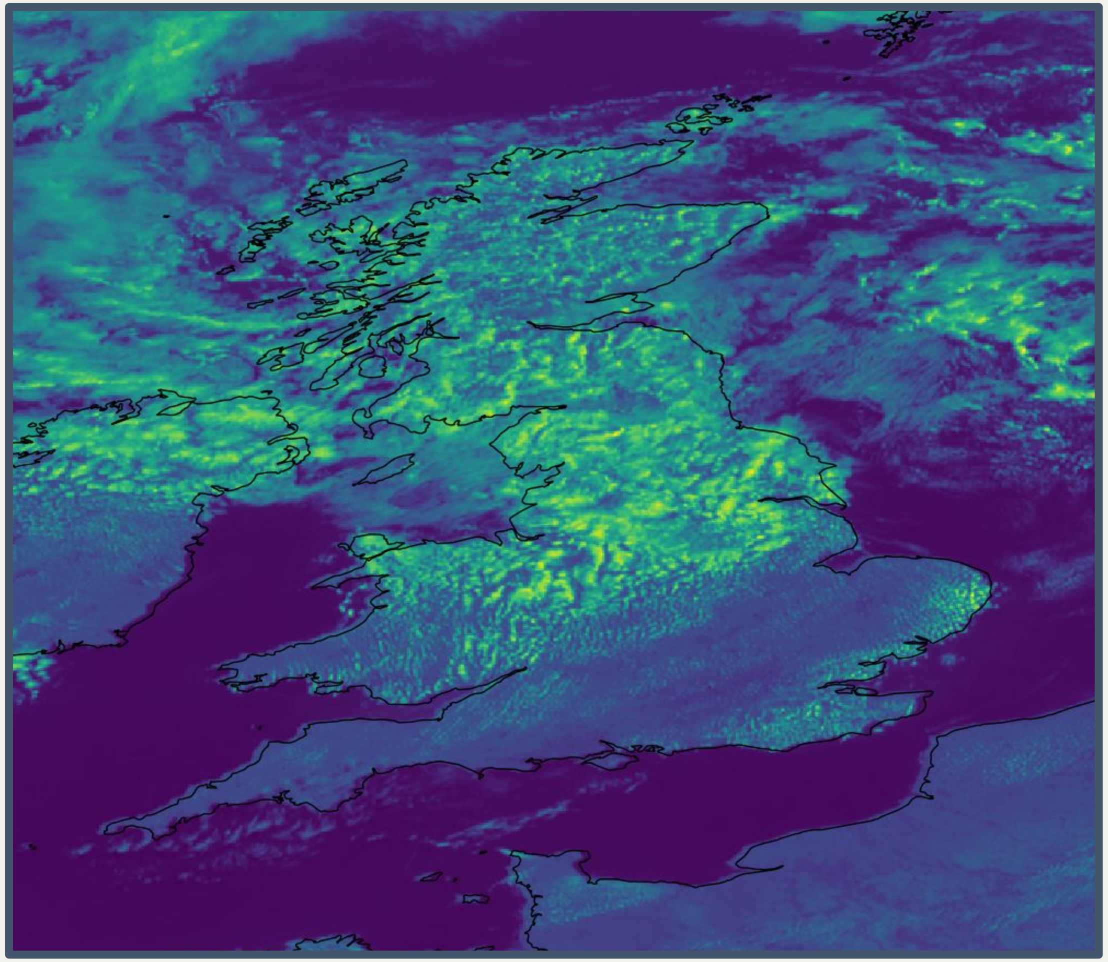
Richard Hanxu, Andy Cai, Brayden Zhang

UNIVERSITY OF TORONTO

THE CHALLENGE

“Develop a model for site-level PV forecasting over the next four hours that is both accurate and performant”

DATASET

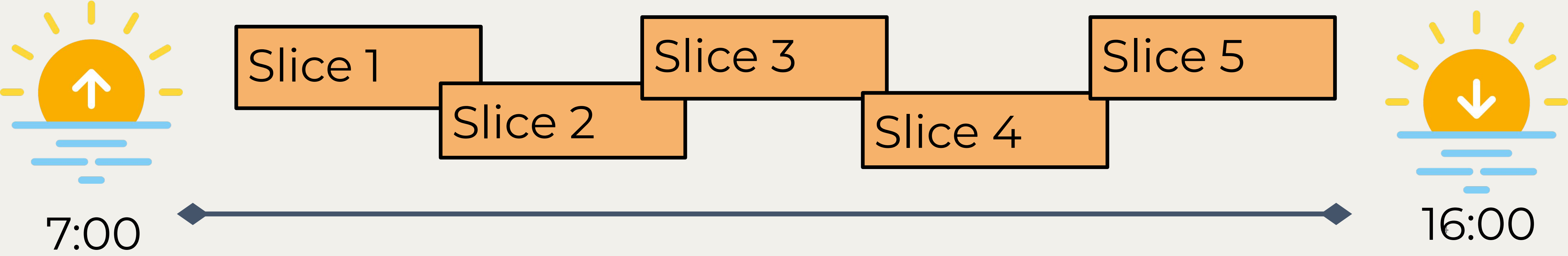


Useful Features

- Past PV
- Weather
  - Diffusive short wave radiation (aswdifd_s)
  - Direct short wave radiation (aswdir_s)
  - Cloud cover % (clch, clcm, clct
  - Relative humidity % (relhum_2m)
- Time
  - Only evaluated between sunrise and sunset

Data Generation

Take 5 4-hour time slices throughout each day, for every date of the year, for every site location

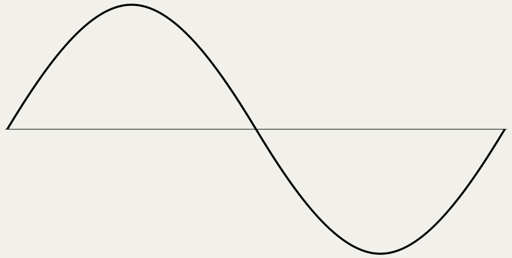


Example

DATA PREPROCESSING

For Time Variables:

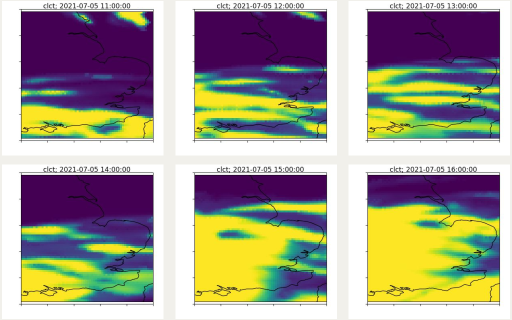
Months / hours →



Sinusoidal Encoding

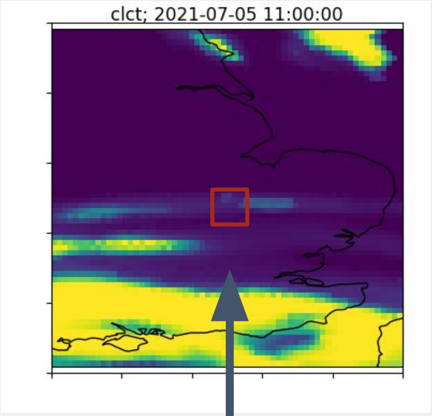
For All Weather Variables:

E.g. Cloud Cover Total (clct) →



$$\frac{\text{[maps]} - \mu_w}{\sigma_w}$$

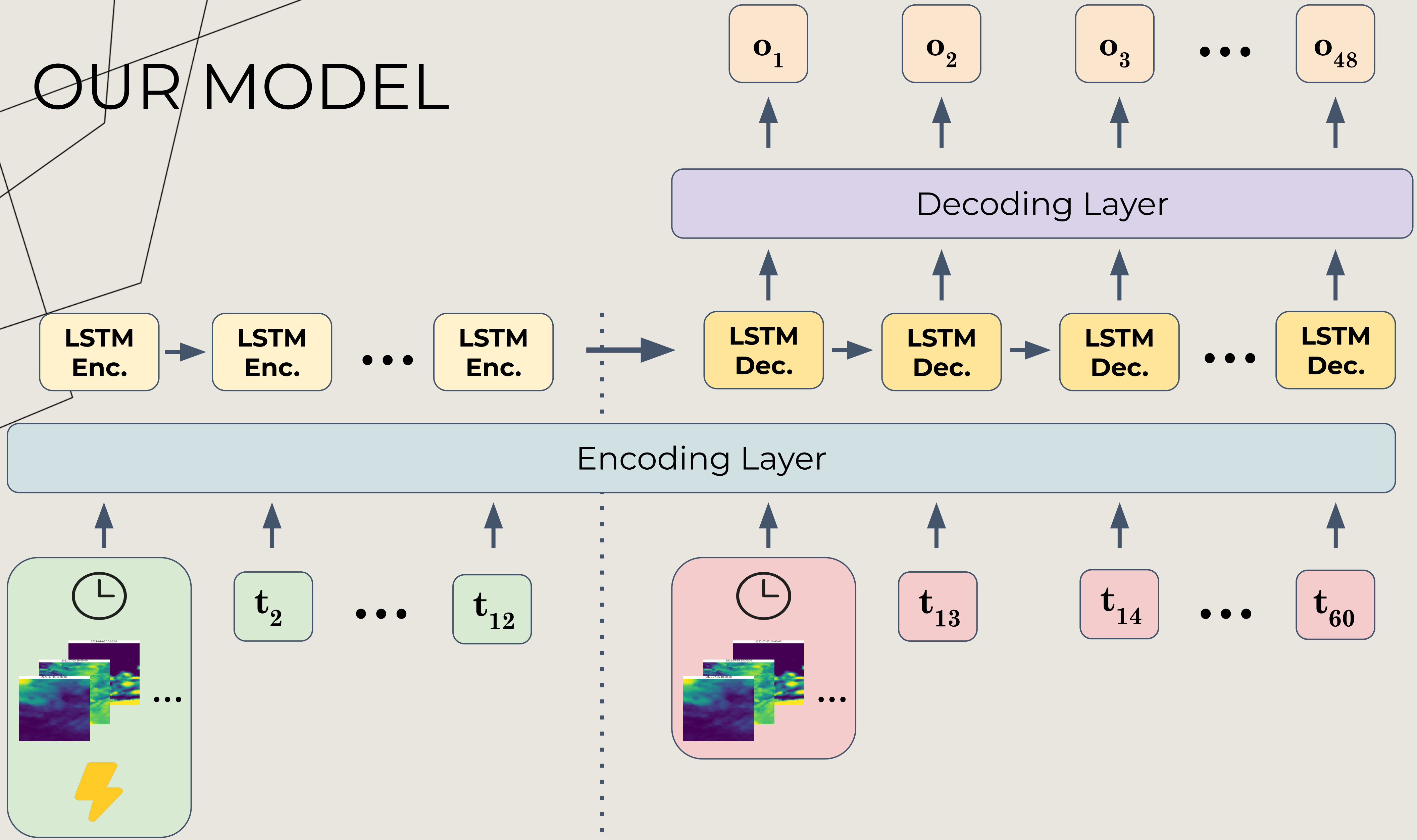
→



Centre 2 x 2

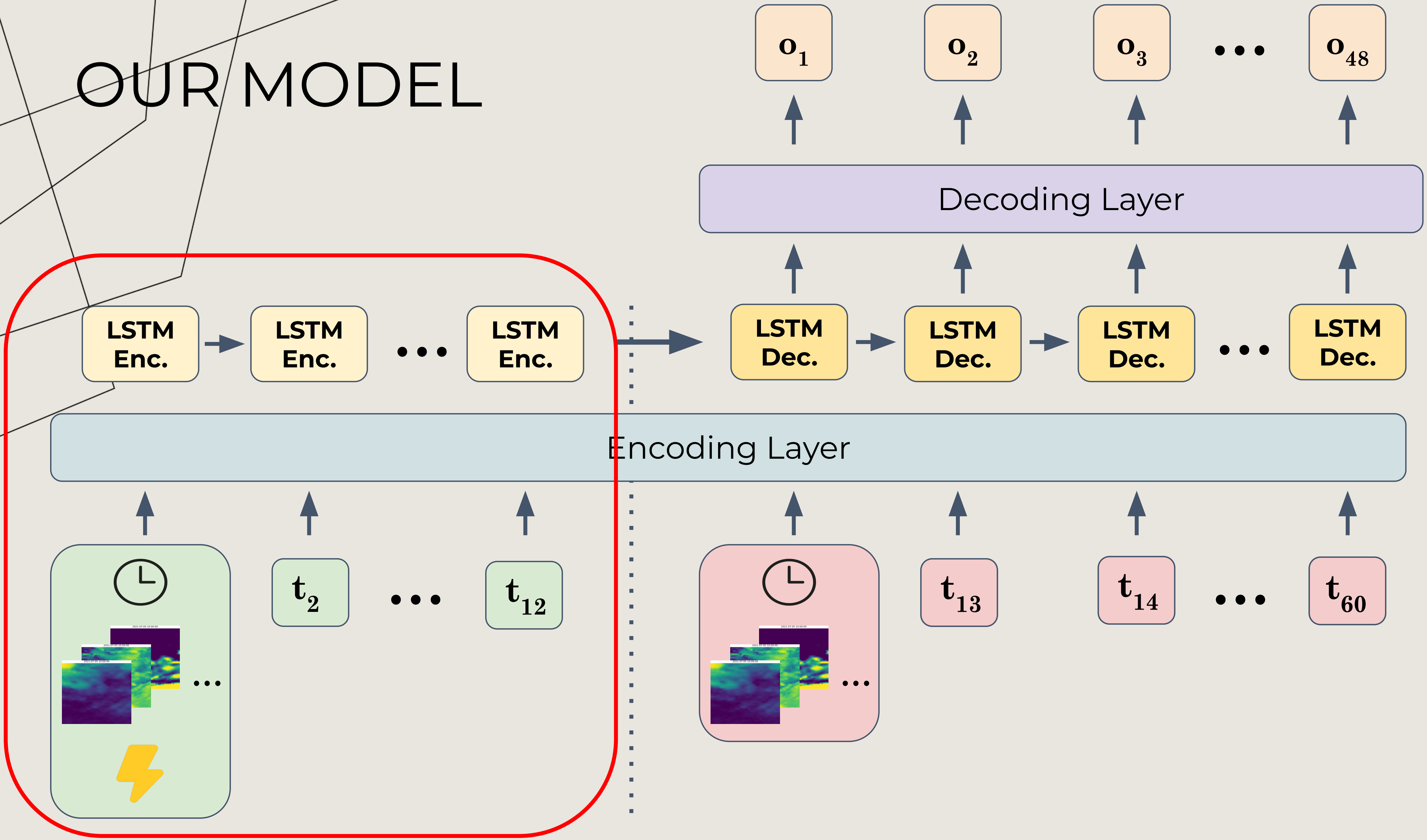
μ_w, σ_w calculated over all selected data points in dataset

OUR MODEL



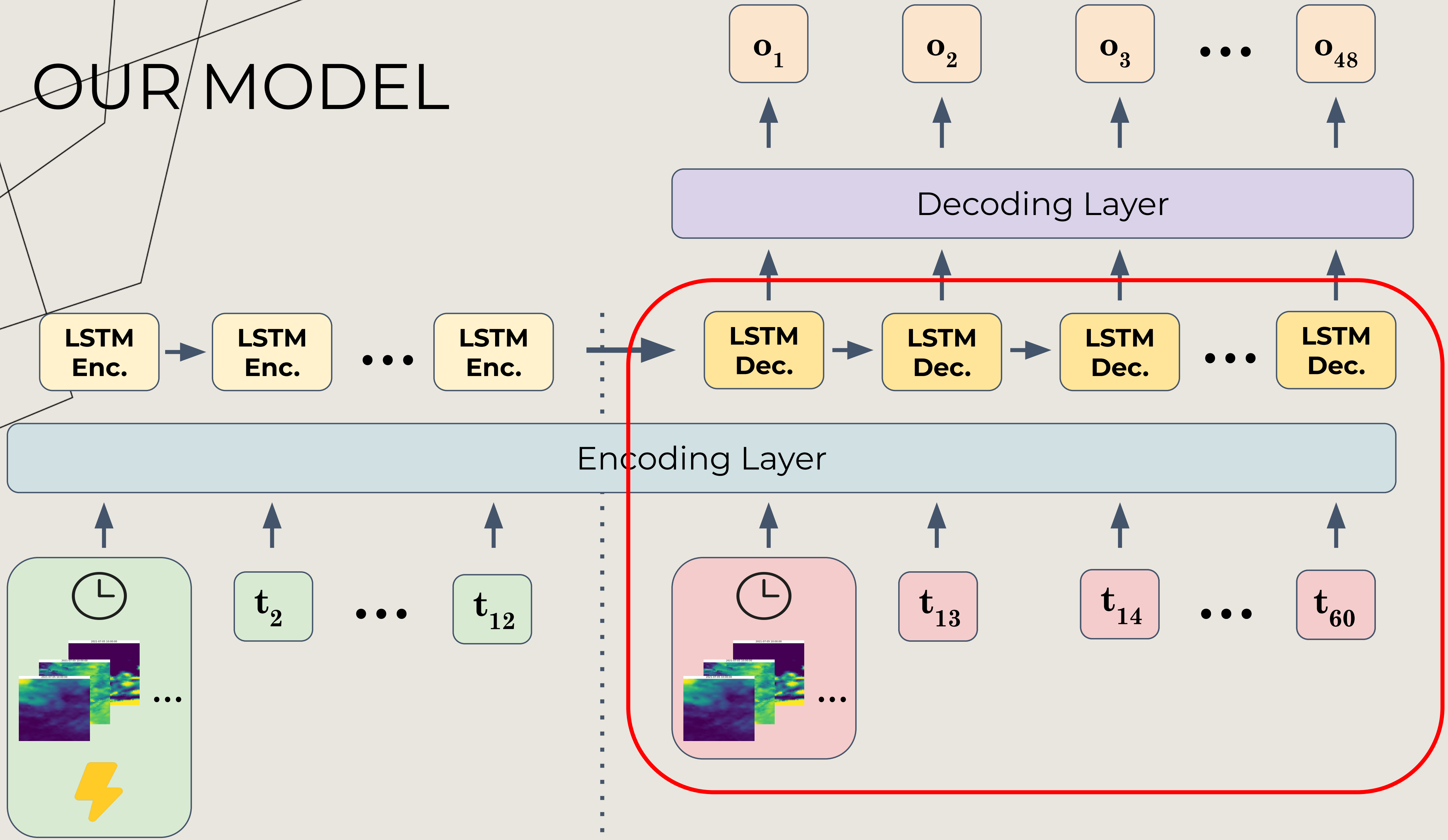
o1
o2
o3
• • •
o48
Decoding Layer
LSTM Enc.
LSTM Enc.
• • •
LSTM Enc.
LSTM Dec.
LSTM Dec.
LSTM Dec.
• • •
LSTM Dec.
Encoding Layer
t2
• • •
t12
t13
t14
• • •
t60

OUR MODEL



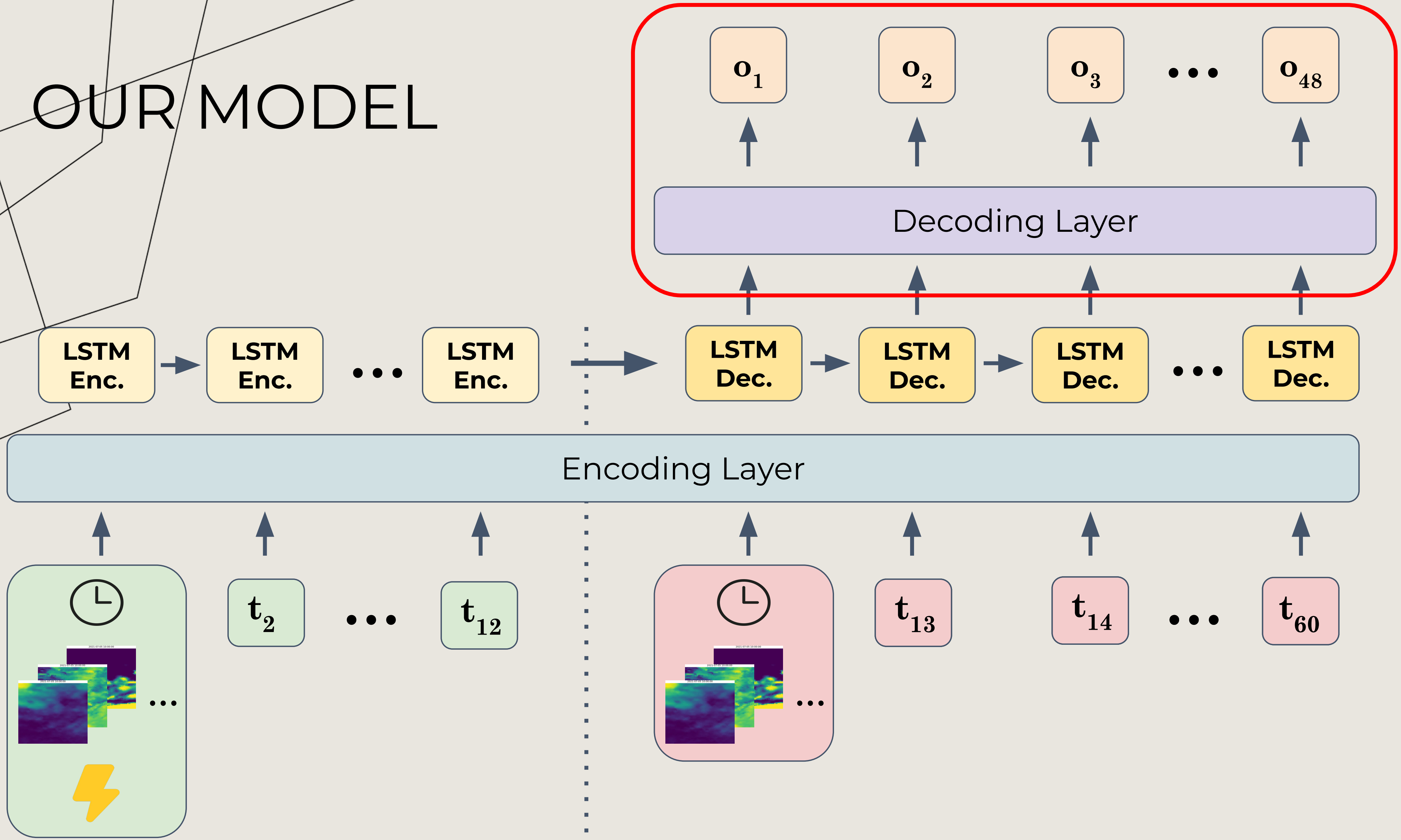
o1
o2
o3
...
o48
Decoding Layer
LSTM Enc.
LSTM Enc.
...
LSTM Enc.
LSTM Dec.
LSTM Dec.
LSTM Dec.
...
LSTM Dec.
Encoding Layer
t2
...
t12
t13
t14
...
t60

OUR MODEL



o1
o2
o3
...
o48
Decoding Layer
LSTM Enc.
LSTM Enc.
...
LSTM Enc.
LSTM Dec.
LSTM Dec.
LSTM Dec.
...
LSTM Dec.
Encoding Layer
t2
...
t12
t13
t14
...
t60

OUR MODEL



o_1
o_2
o_3
...
o_{48}
Decoding Layer
LSTM Enc.
LSTM Enc.
...
LSTM Enc.
LSTM Dec.
LSTM Dec.
LSTM Dec.
...
LSTM Dec.
Encoding Layer
...
t_2
...
t_{12}
...
t_{13}
t_{14}
...
t_{60}

Model Training

Loss Function: **MAE**

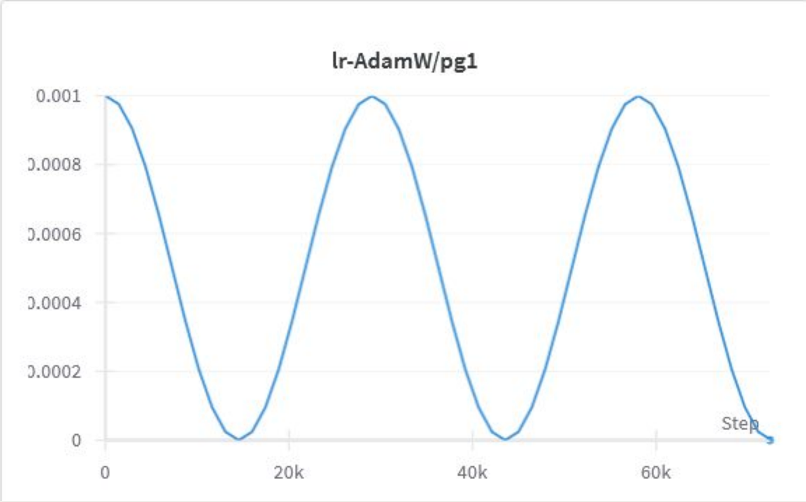
Optimizer: **Adam**

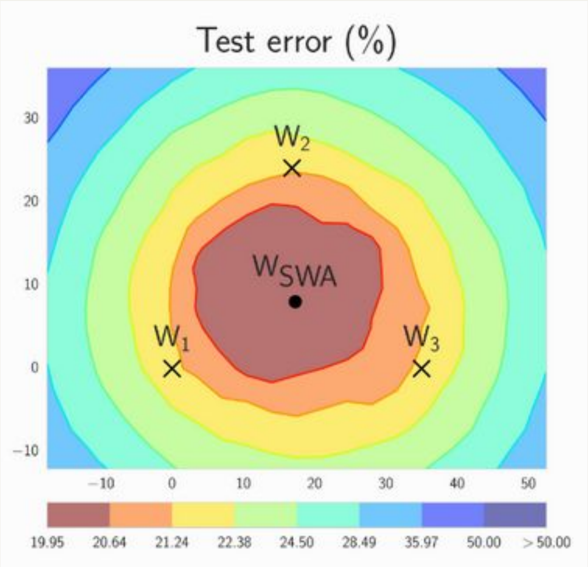
Training Tricks:

- *Cyclic learning rate scheduling*
- *Stochastic Weight Averaging*

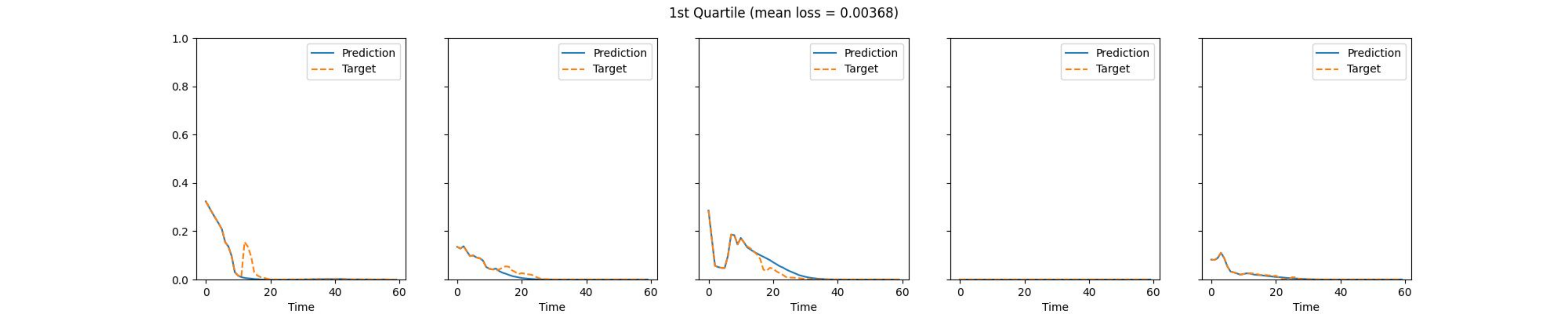
Epochs: 12 ⇒ Training Time: ~7 hours

Hardware: Trained on the Cloud – AWS/Vast.ai

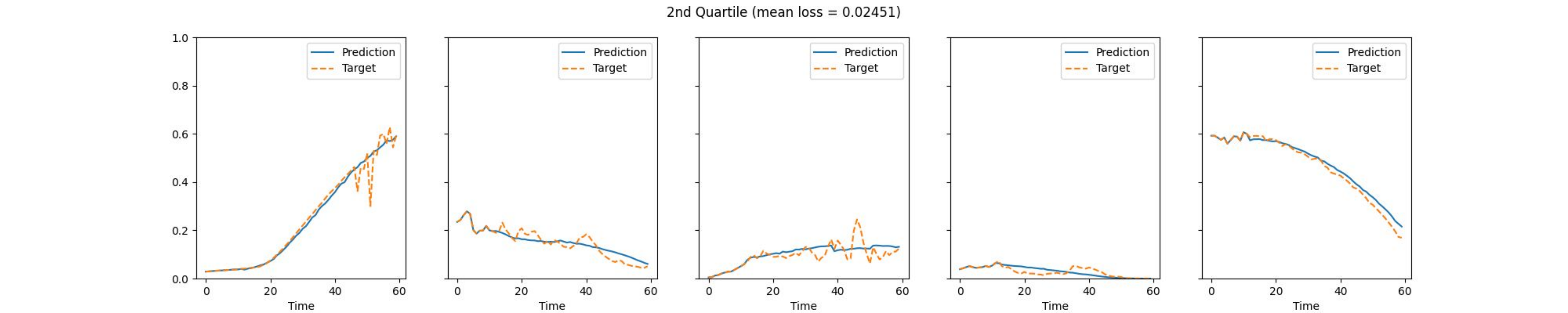




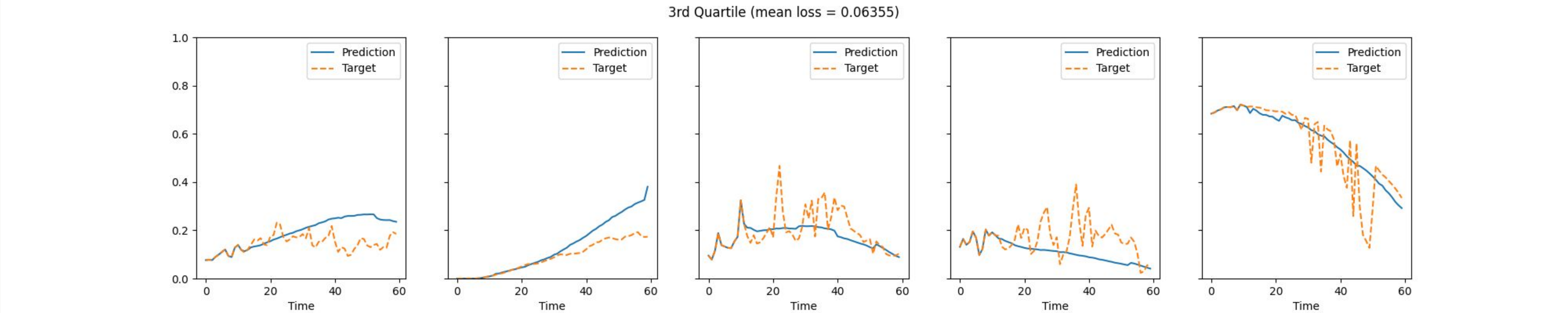
Model Performance: 1st Quartile



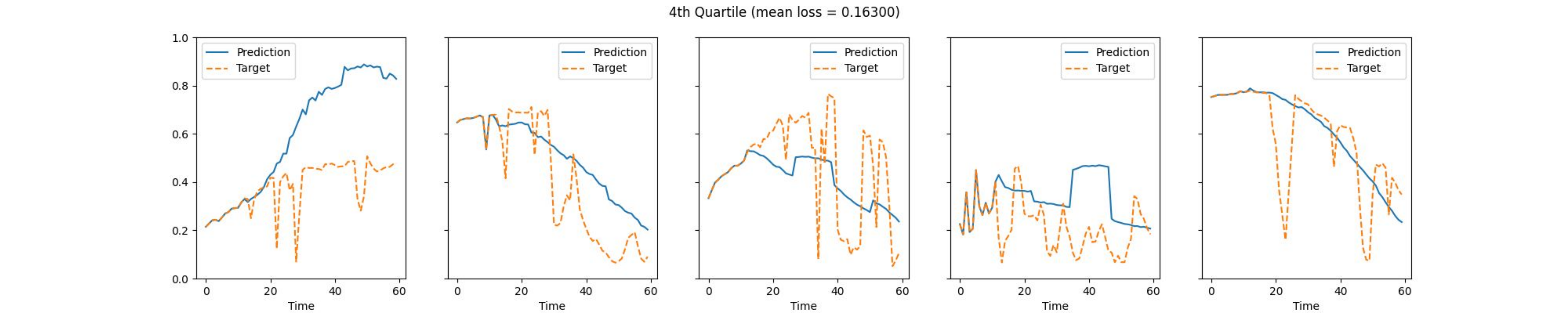
Model Performance: 2nd Quartile



Model Performance: 3rd Quartile



Model Performance: 4th Quartile



Advantages & Limitations

Advantages

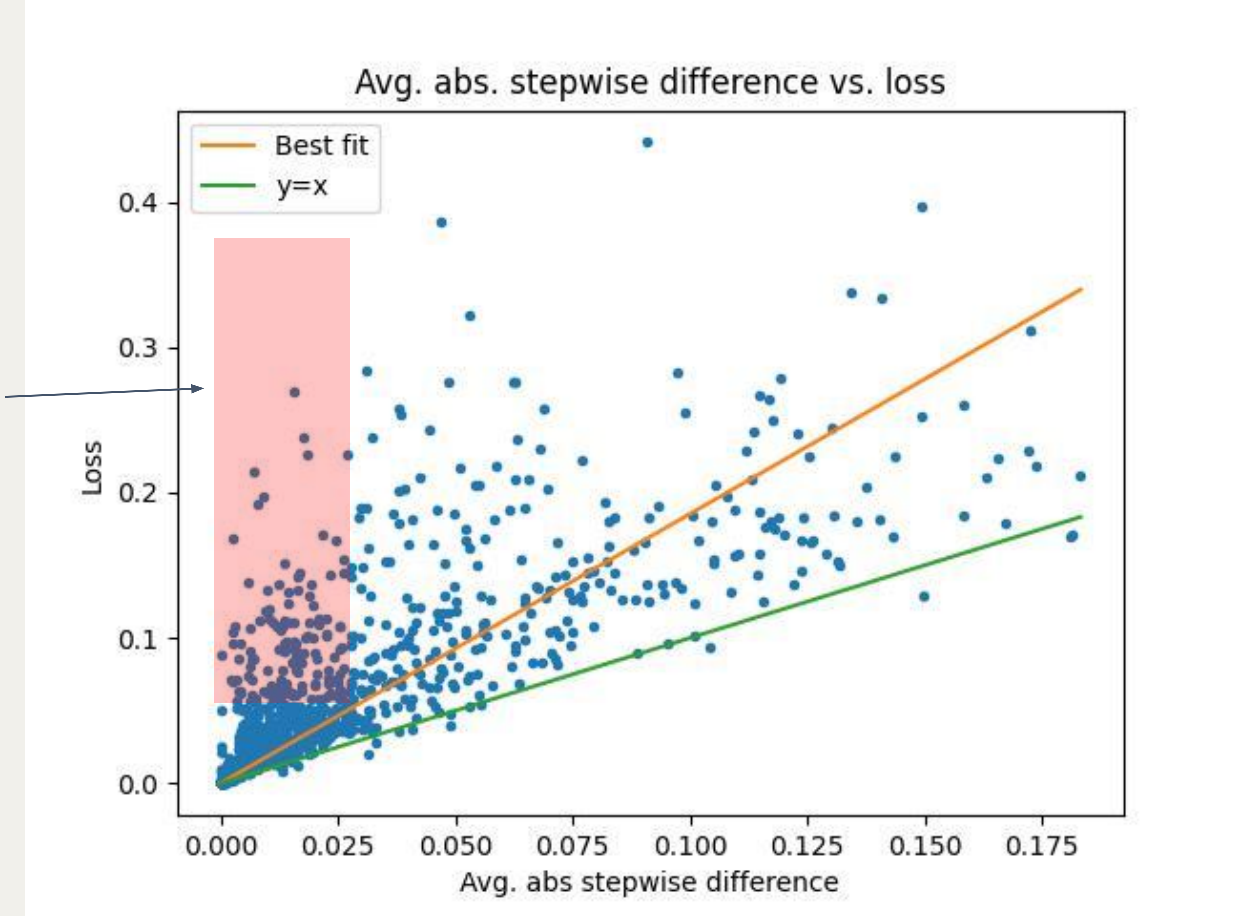
1. **Lightweight/Highly Deployable**
   a. Faster inference
   b. Lower Training Time
   c. Compact
2. Simplicity

Limitations

1. Only uses weather features, which has low time resolution
2. Is poorer at predicting volatile targets

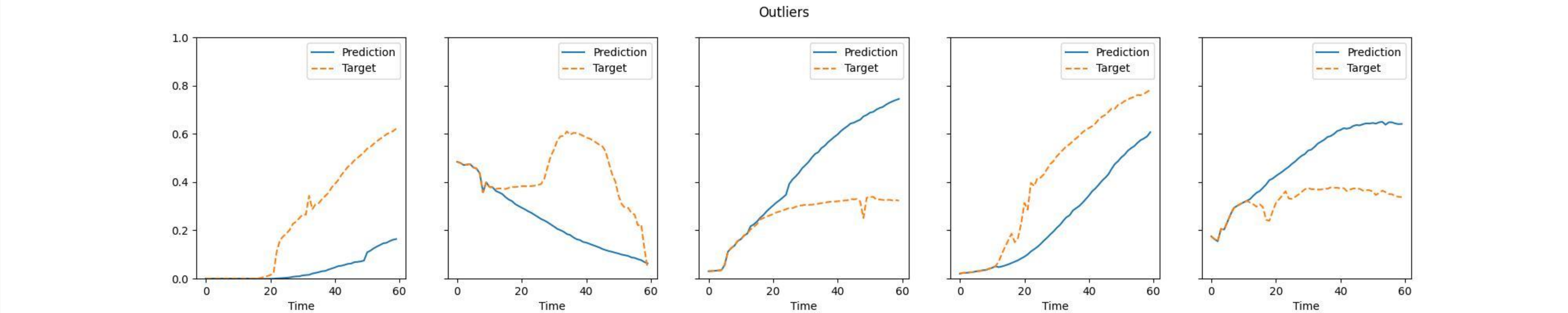
Advantages & Limitations

Region with unexpectedly poor performance



"Volatility"

Advantages & Limitations



Other Ideas Tried

Curve Fitting

- Fit a quartic curve to data per each month
- *Drawback: Does not account for imagery/site data*

K-Nearest-Neighbors

- Match data that is “similar” in various features (month, time, weather, hrv, etc.)
- *Drawback: Computational time increases linearly with number of dimensions and data-points used; can become very large.*

Sunlight Baseline

- Compute sun angle at time with lat., lon; combine with orientation and tilt to approximate amount of sunlight on solar panel
- *Drawback: Does not account for imagery/site data*

Expectation Maximization

- Used the same LSTM architecture but also predicted uncertainty; tried variety of distributions (normal, laplace, truncated normal/laplace)
- *Drawback: Performance was worse than normal LSTM*

Future Improvements

- Explore if larger model size (e.g. more combinations of variables) leads to better performance
  - Utilize more weather variables
- Explore optimal crop location and size
  - most important region not necessarily centered on site due to location of sun in sky
  - larger crops necessary for higher altitude features due to light diffusion
- Incorporate HRV with weather features (e.g. wind speed) to get better time resolution using optical flow

Future Improvements

- Investigate importance of aerosol and non-HRV data
- Predict difference from sunlight baseline
  - Force model to predict high time resolution variations
  - Handle smooth prediction with basic model
- Improve expectation maximization models so that performance is comparable to simpler model
  - Could be used to generate ensemble (take least uncertain predictions at each time step)

Overall Insights

- Importance of feature engineering & normalization
- Difficulty of determining influential parameters – importance of time ⇒ LSTM

Skills Developed

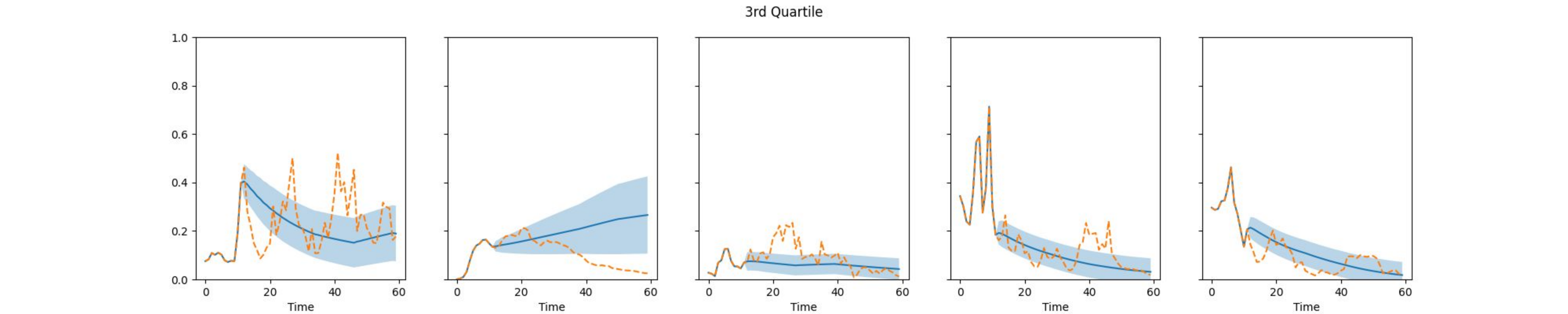
- Data pre-processing
- Setting up organized training modules and designing efficient training methods
- Implementing and evaluating different DL models in PyTorch

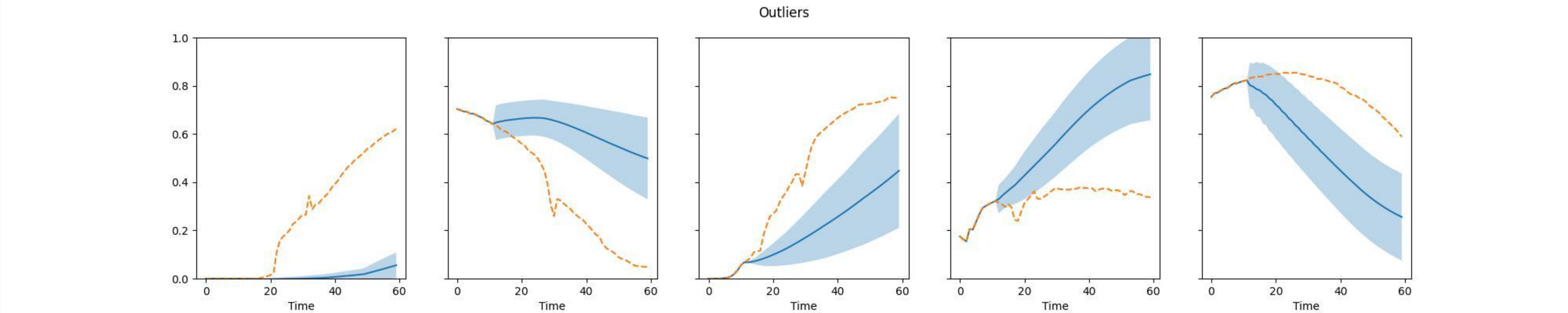
Thank You!

richard.hanxu@mail.utoronto.ca

andy.cai@mail.utoronto.ca

brayden.zhang@mail.utoronto.ca


3rd Quartile
1.0
0.8
0.6
0.4
0.2
0.0
0
20
40
60
Time


Outliers
1.0
0.8
0.6
0.4
0.2
0.0
0
20
40
60
Time