

ESTIMATING LANDFALLING HURRICANE WAVE CHARACTERISTICS WITH SURROGATE MODELLING

Du Tianju (Tim)

Supervisors:

**Dr. Lluis Via Estrem,
Prof. Matthew Piggott**

OVERVIEW

1. Introduction: Florida's Coastline & MIKE 21
2. Problem Description and Objectives
3. Exploratory Data Analysis
4. Modelling
5. Results
6. Conclusion

FLORIDA'S COASTLINE

Tropical Cyclones

Tropical Cyclones pose significant economic risks :

Recent hurricanes, like Irma in 2017, caused over \$5.88 billion in damages

Storm surge-induced **wave height** is crucial for predicting potential hazards.



Irma Affecting South Florida (NASA, 2017)

MIKE21 : COASTAL ENVIRONMENTS



Use synthetic tracks

Hydrodynamics Model

From tides and wind stress to
Water current & Free Surface
Elevation (FSE)

Spectral Wave Model

From Wind stress data, SSE and
Water current to **Wave height**
and Peak wave period

Time Taken

Hydrodynamics Model for South Florida: ~20hrs/Storm (fine mesh)
Hydrodynamics Model for East Florida: ~2hrs/storm

Spectral Wave Model for South Florida: ~110hrs
Spectral Wave Model for East Florida: ~25hrs
Total simulation Time: ~7 days

PROBLEM AND OBJECTIVES

Problem:

- **High computational cost** of hydrodynamic models (MIKE21)

Objectives:

- A surrogate model using deep learning to predict wave heights more efficiently
- Leverage the speed and efficiency of DL models to provide **faster** predictions with accuracy comparable to traditional methods
- Create user-friendly pipeline to handle **different** areas

THE DATA

Raw Data

Florida South csv

MIKE21 inputs and outputs

97 time-step length time series with interval 0.5 hour

432 storms for **South Florida**

288 grid points

grid_dfsu_data_M08FLS_1178

| Hs | Hmax | Tp | Wspd | Wdir | FSE | TWD | time | pt_id | track_id |
|-----------|-----------|-----------|-----------|----------|-------------|-----------|---------------------|-------|----------|
| 0.2742777 | 0.5664476 | 2.4592955 | 1.2482222 | 103.2401 | 0.047367822 | 37.580235 | 1962-08-24 22:00:00 | 1 | 1178 |

⋮

Florida East csv

MIKE21 inputs and outputs

97 time-step length time series with interval 0.5 hour

330 storms for **East Florida**

88 grid points

grid_dfsu_data_M09FLE_1178

| Hs | Hmax | Tp | Wspd | Wdir | FSE | TWD | time | pt_id | track_id |
|-----------|-----------|-----------|-----------|----------|----------|-----------|---------------------|-------|----------|
| 0.2742777 | 0.5364476 | 2.4592955 | 1.2482222 | 103.2401 | 0.047822 | 37.580235 | 1962-08-24 22:00:00 | 1 | 1178 |

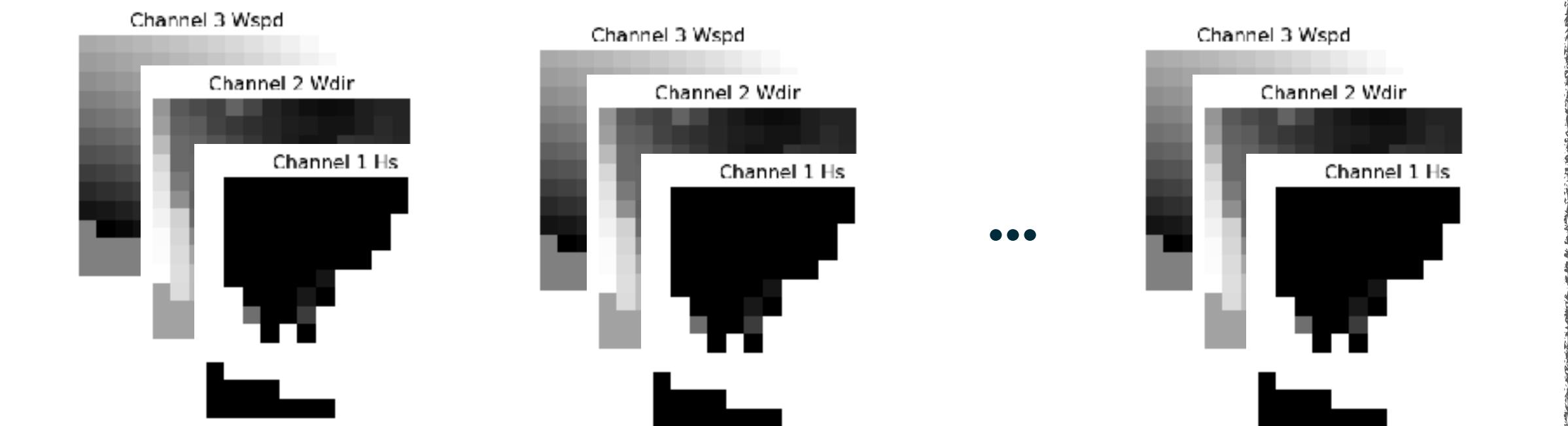
⋮

Pre-Processed Data

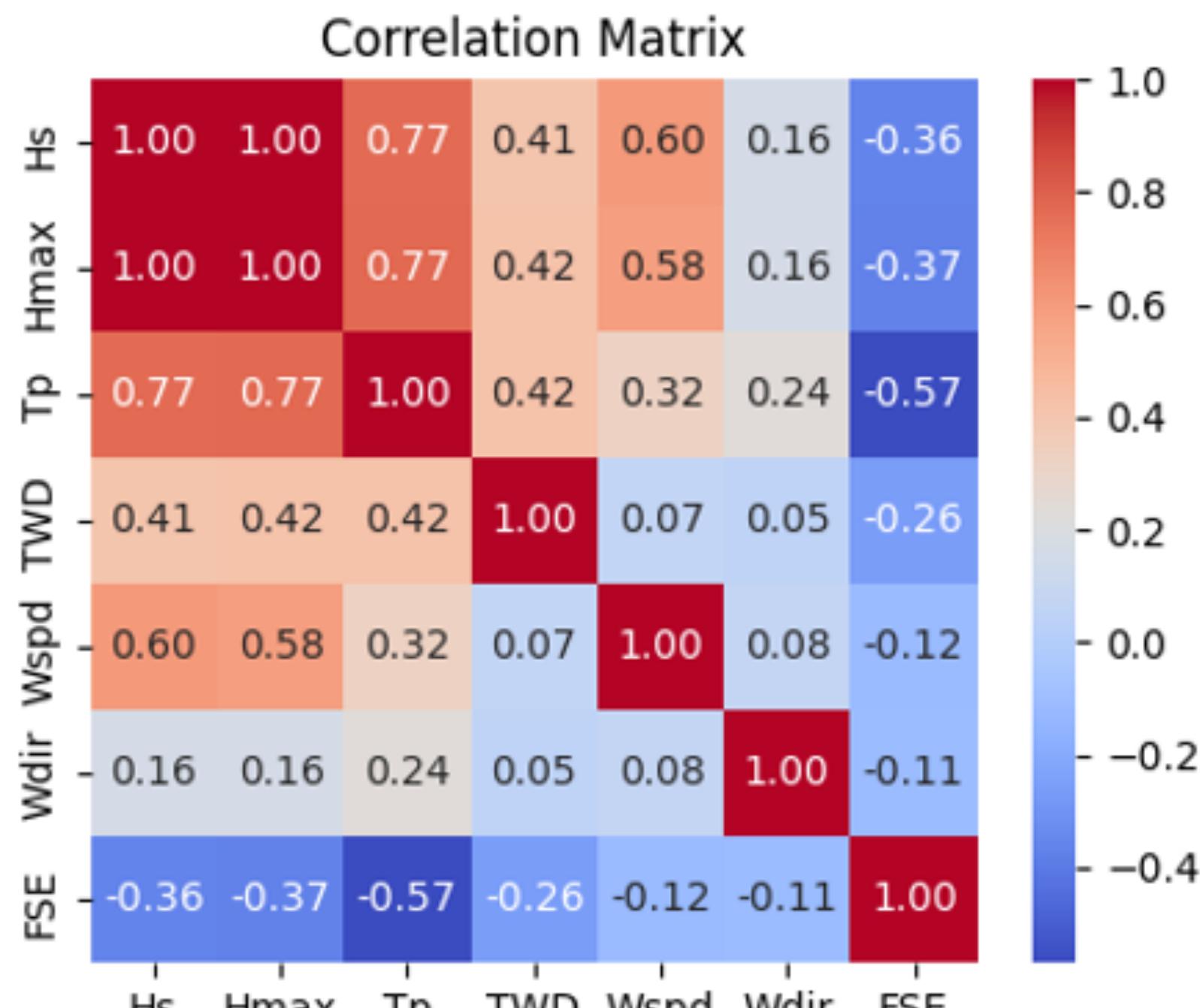
images_format.npy

MIKE21 inputs and outputs to 2D shape

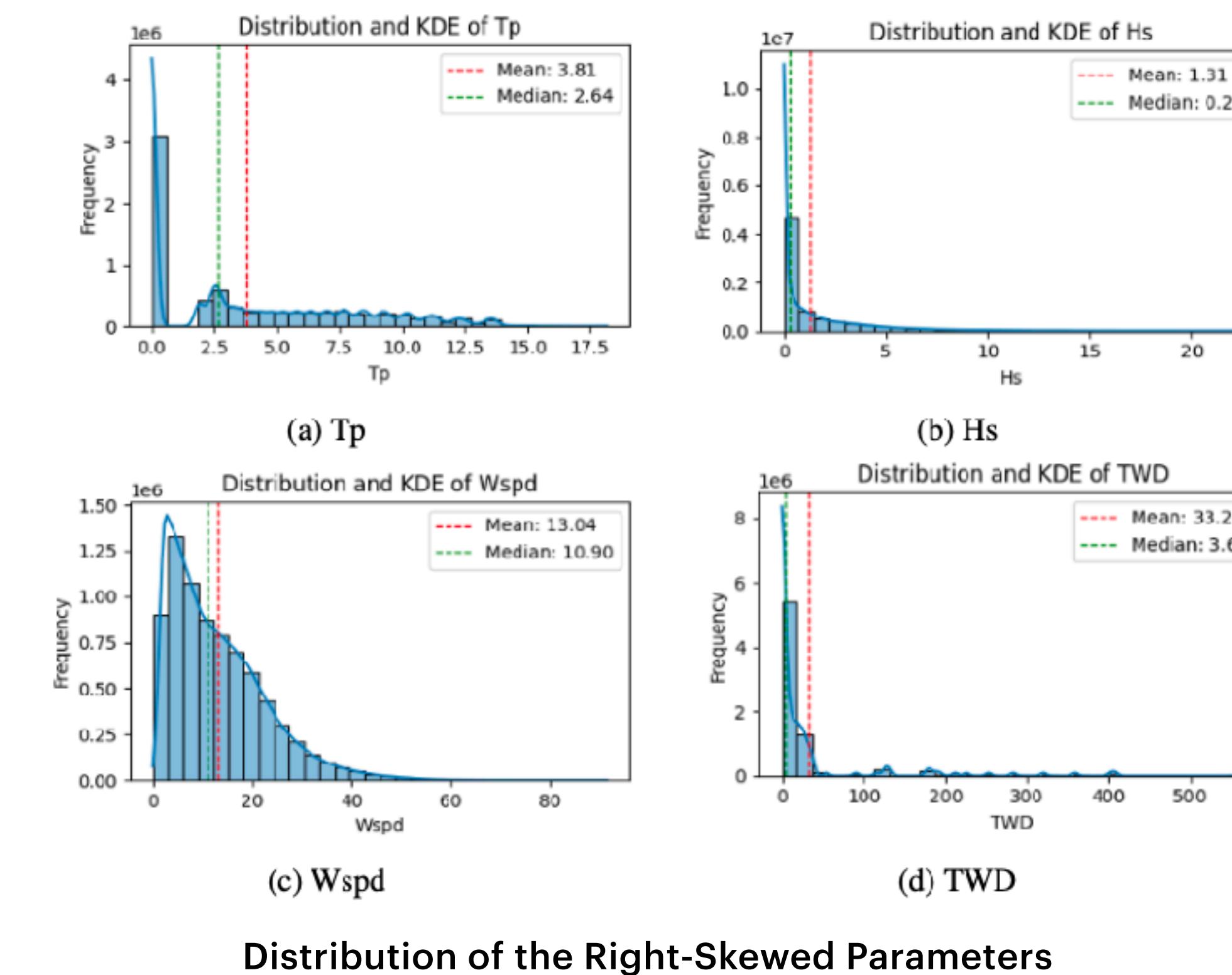
Timestep: 0.5 hour



EXPLORATORY DATA ANALYSIS: CORRELATION AND DISTRIBUTION

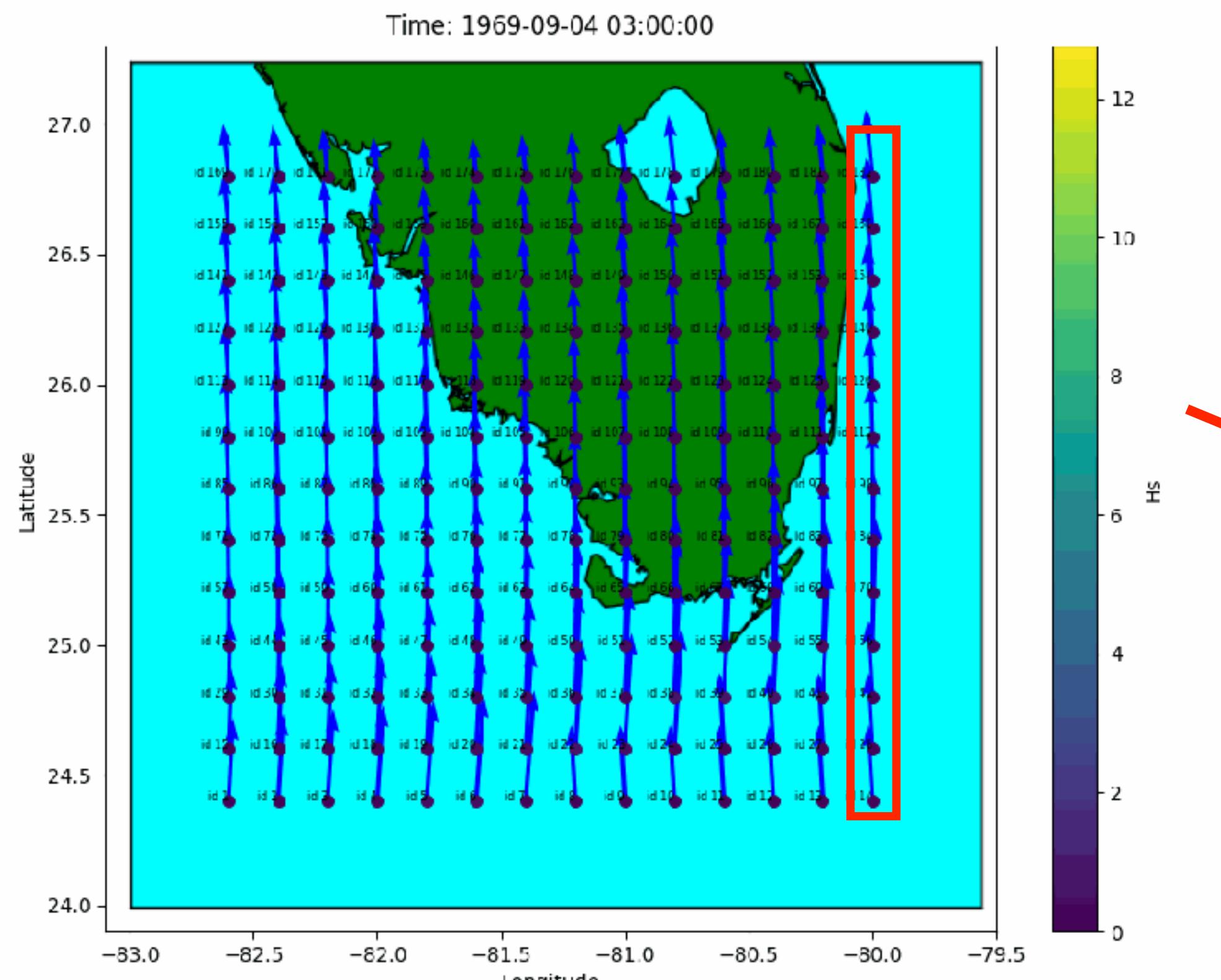


Correlation Matrix of MIKE21 input
and output csv files

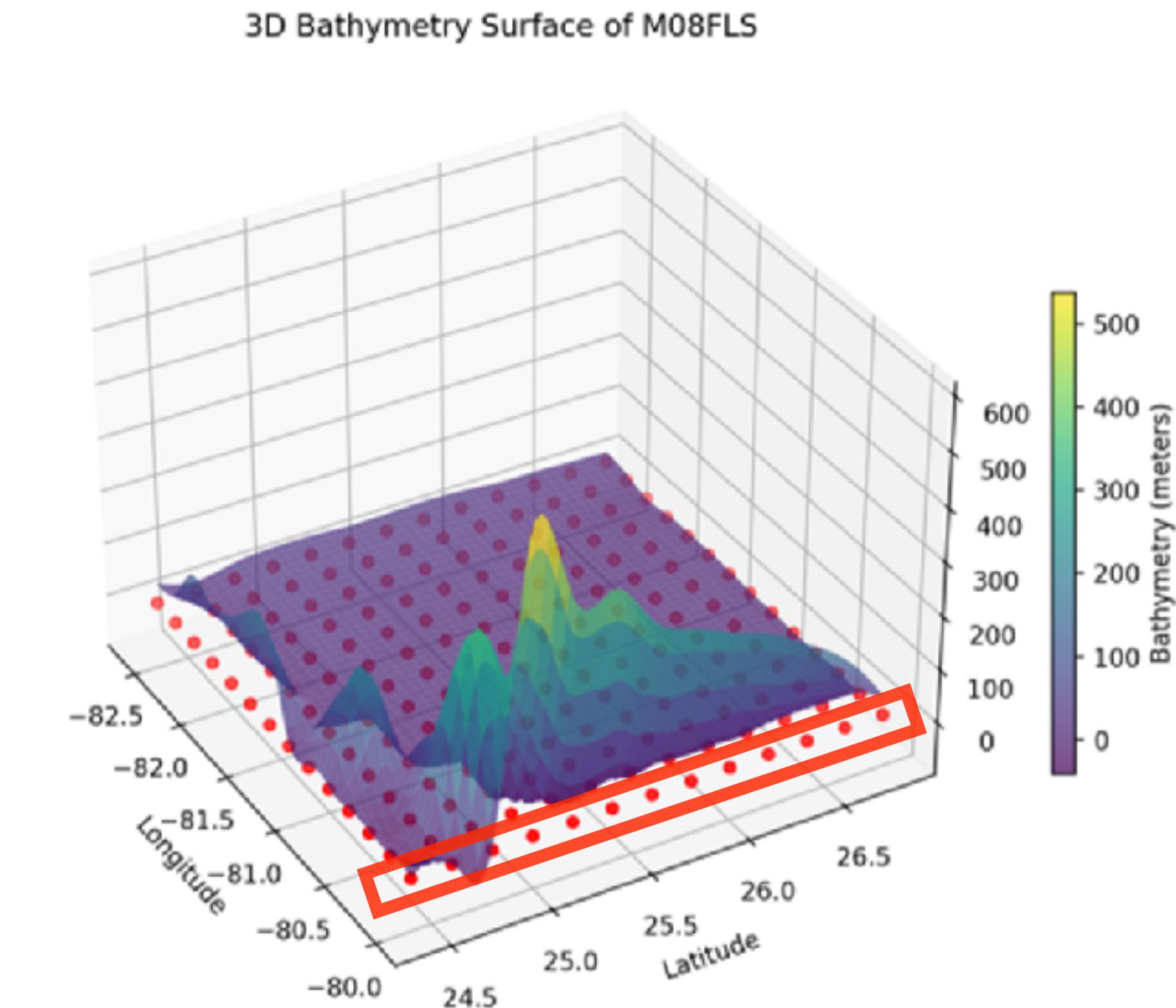


Hs: Significant Wave Height
 Wspd: Wind Speed
 Wdir: Wind Direction
 TWD: Total Water Depth
 FSE: Free Surface Elevation
 Tp: Peak Wave Period

EXPLORATORY DATA ANALYSIS: GEOGRAPHIC INSIGHTS



The Arrangement of computational points in the
MIKE 21 model's input and output data



3D Bathymetry Surface of South
Florida

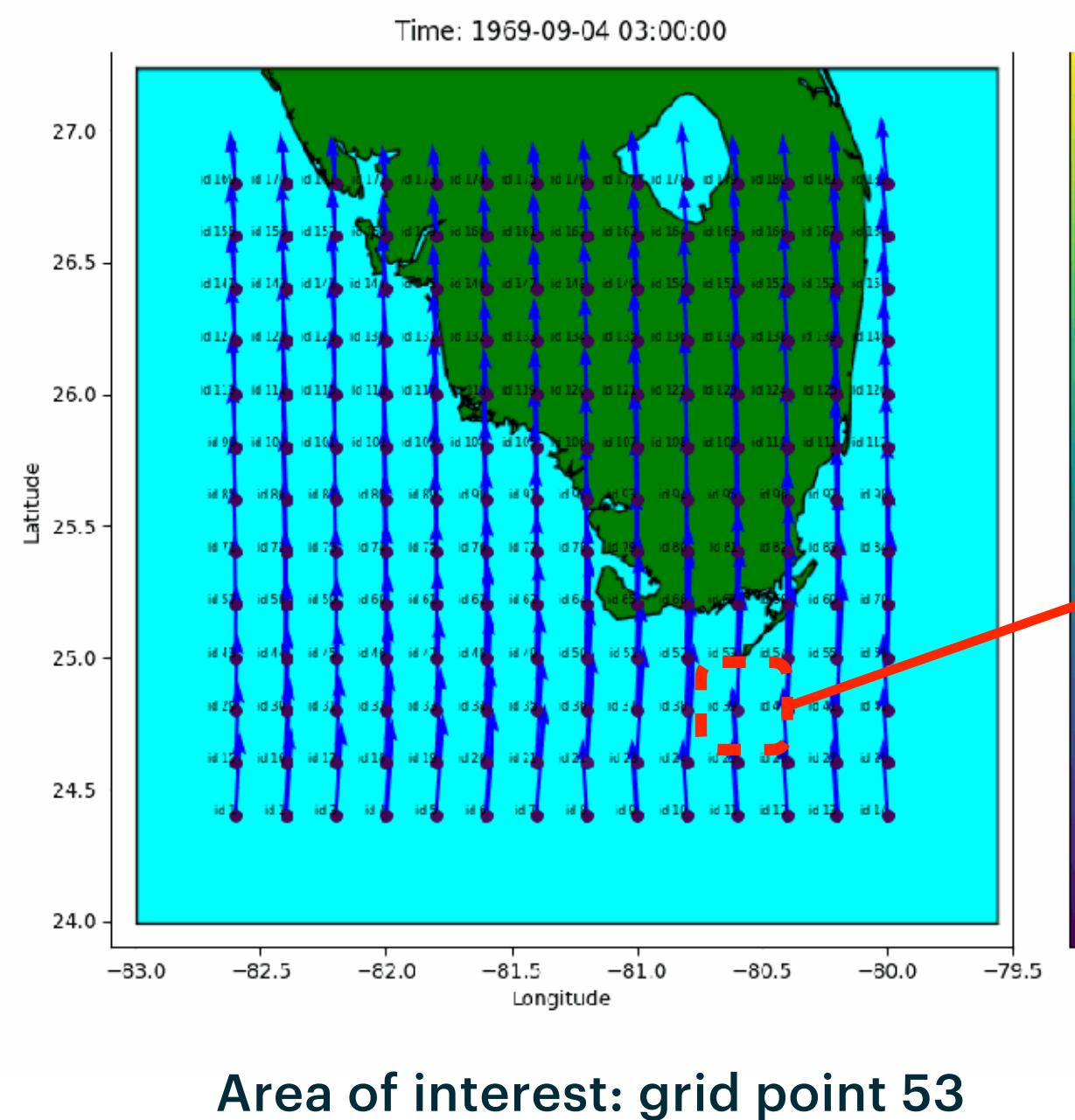
MODEL 1:Baseline Model LSTM

MODLE 2: CNN+LSTM Model (spatial improvement)

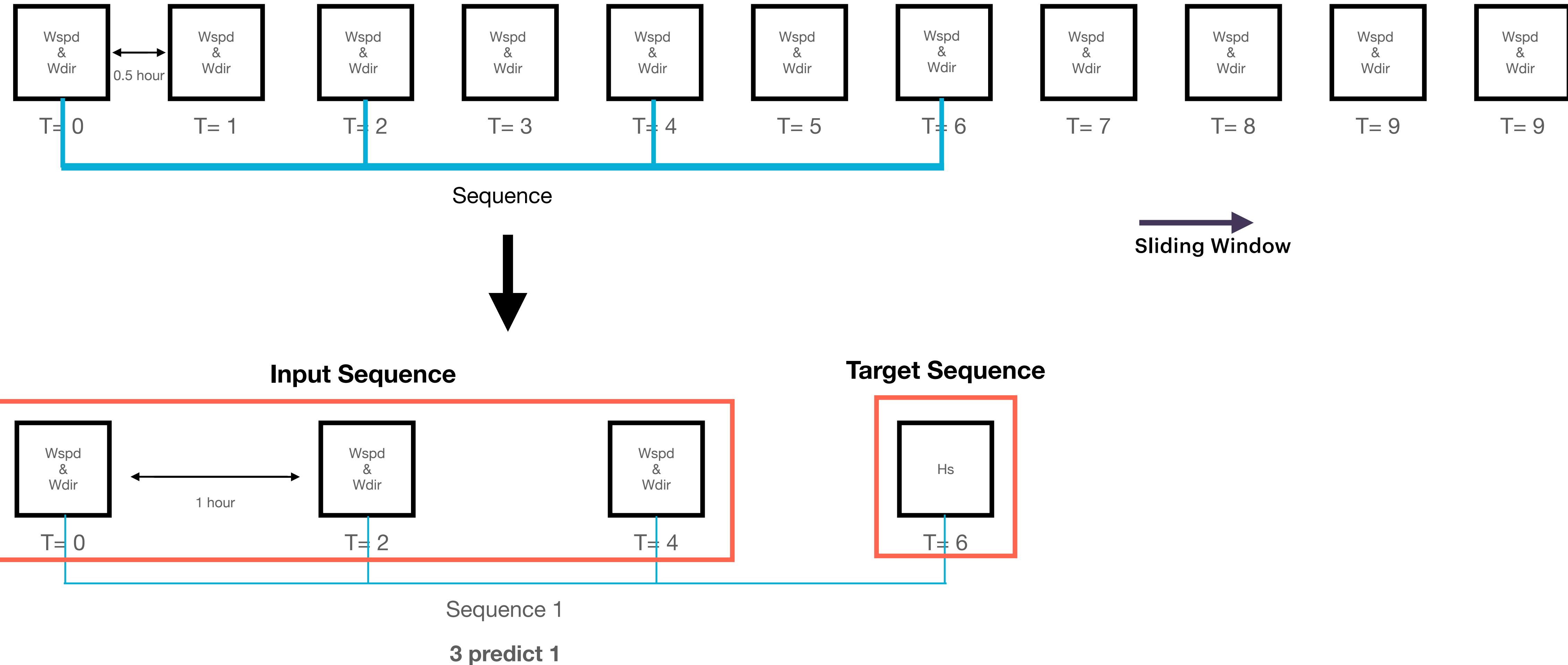
MODEL 3: Latent Space Model (peak value improvement)

MODELLING APPROACH 1: 1D LSTM

OBJECTIVE: Forecast Significant Wave Height at a specific grid point (53) using Wind speed and direction with Long short-term memory LSTM Model for Time Series Forecasting



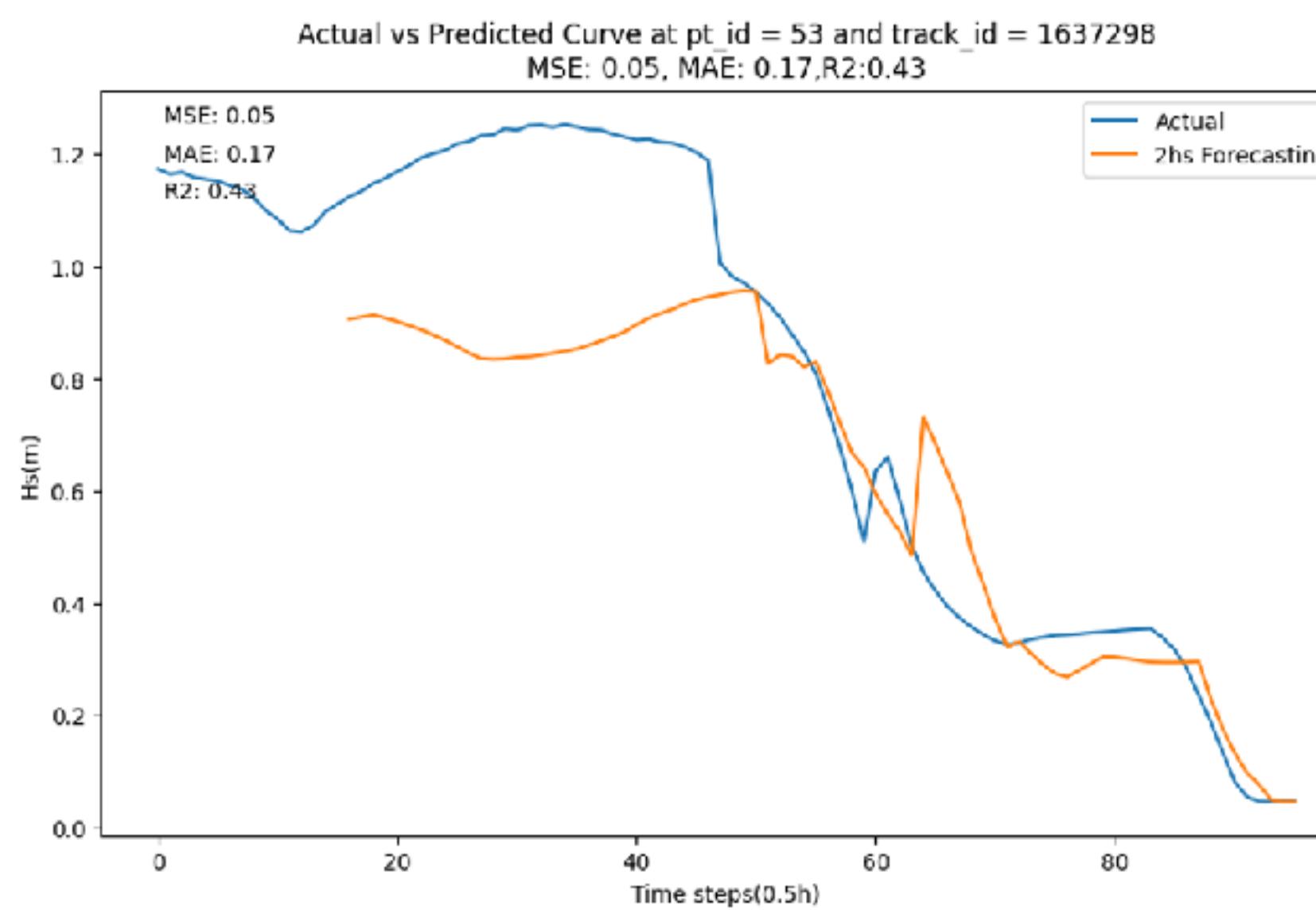
BASELINE MODEL 1D LSTM: DATA LOADER CONSTRUCT



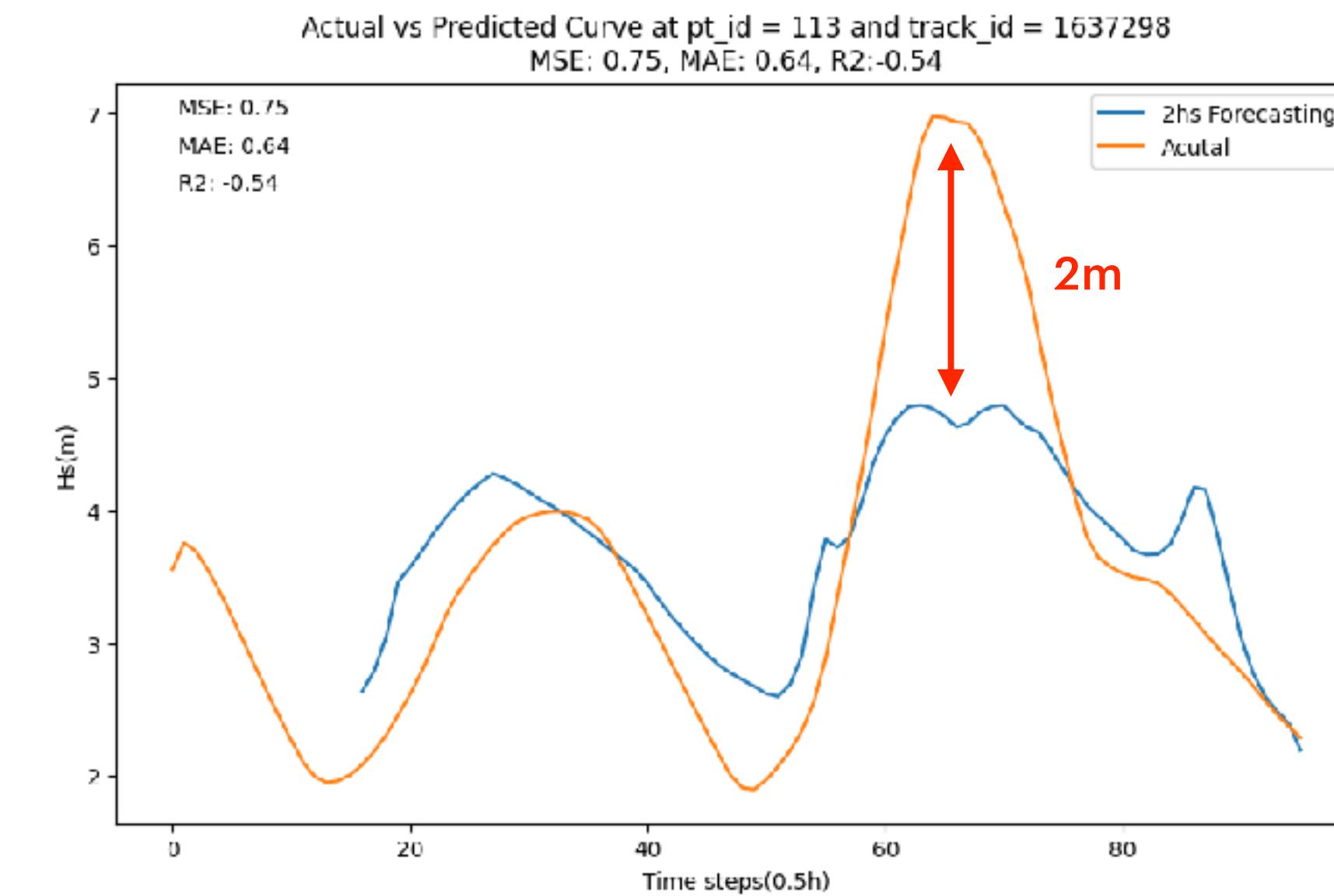
RESULTS AND LIMITATIONS

Performance evaluate metric

| Forecasting Time (Hidden Size of LSTM) | MSE (m^2) | CoE | R ² |
|--|---------------|-------|----------------|
| 2 hours (256) | 0.0147 | 0.971 | 0.89 |
| 4 hours (256) | 0.0233 | 0.947 | 0.81 |
| 6 hours (256) | 0.0372 | 0.936 | 0.77 |
| 8 hours (512) | 0.0451 | 0.899 | 0.69 |



Prediction on the same grid point



Prediction on a grid point 113

LIMITATIONS

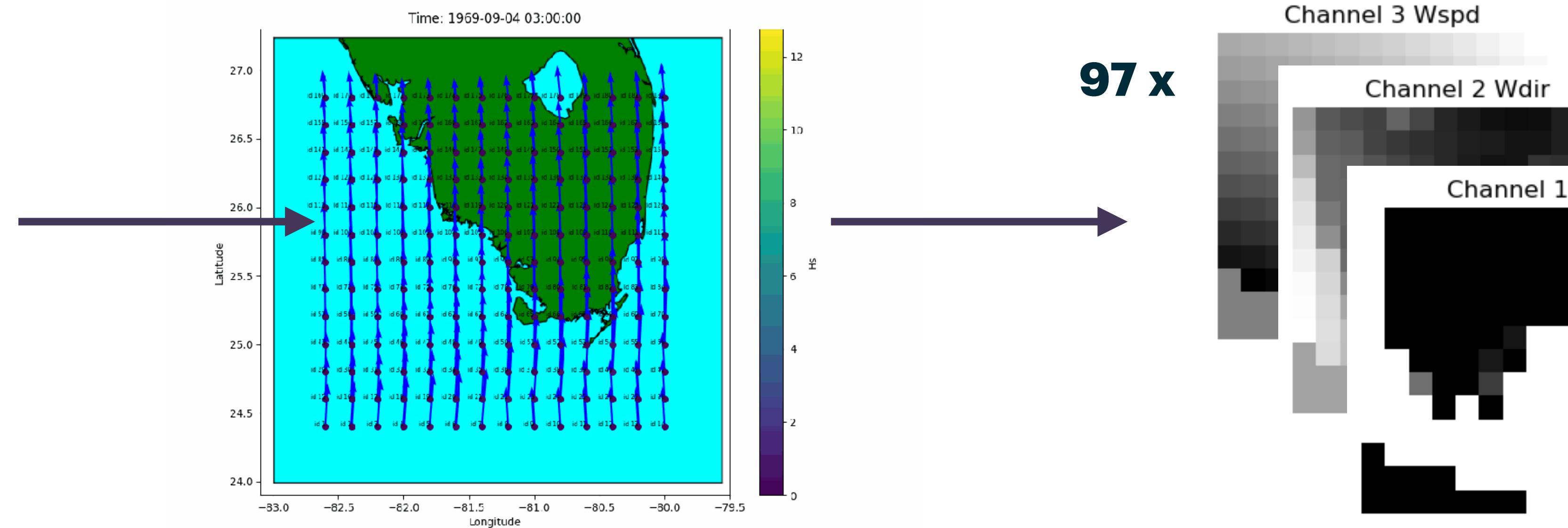
- Lack of **spatial feature learning**
- Poor **peak value predictions**, particularly for extreme events.
- Inability to generalise across varying **bathymetry** and **storm characteristics**

MODELLING APPROACH 2: 2D CNN+LSTM

Objective: Incorporate spatial features by converting grid data into image format.

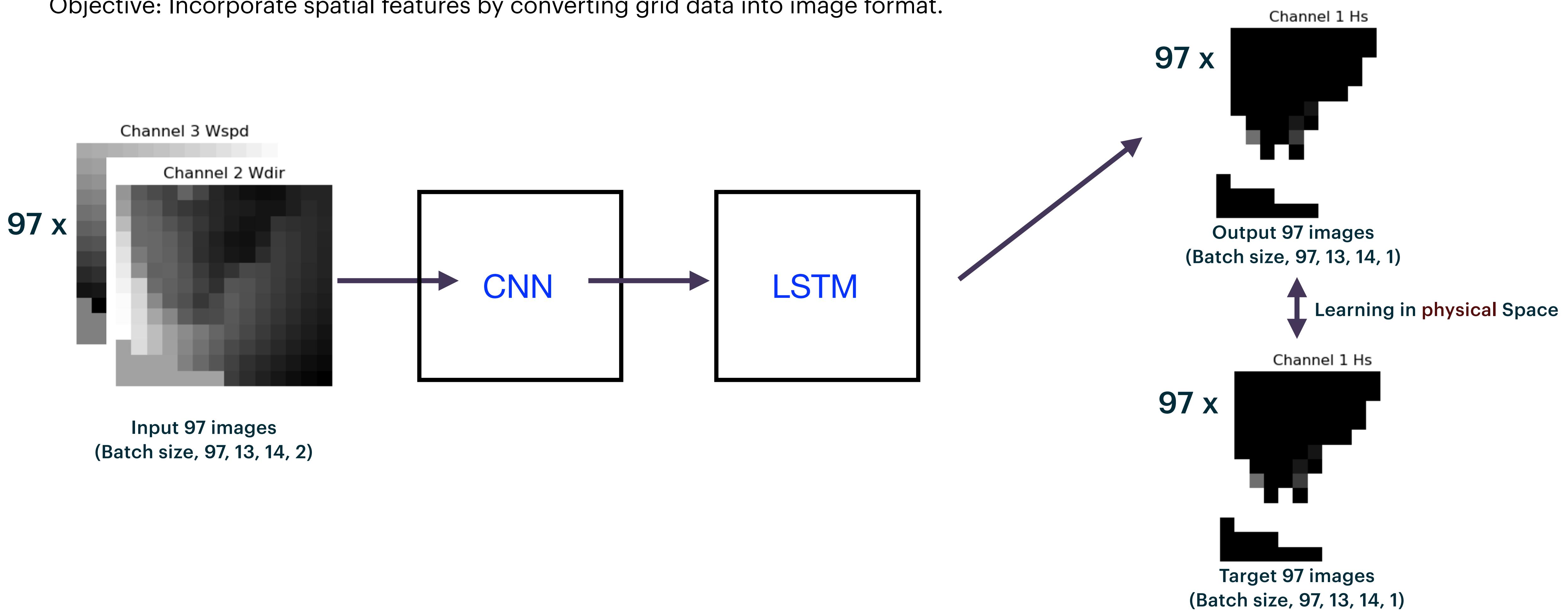
Convert 1D input to 2D image input, The figure illustrates the process of transforming numerical data into images, where each pixel represents a grid point, and each channel corresponds to a different variable. Each time-step represents a image.

CSV files



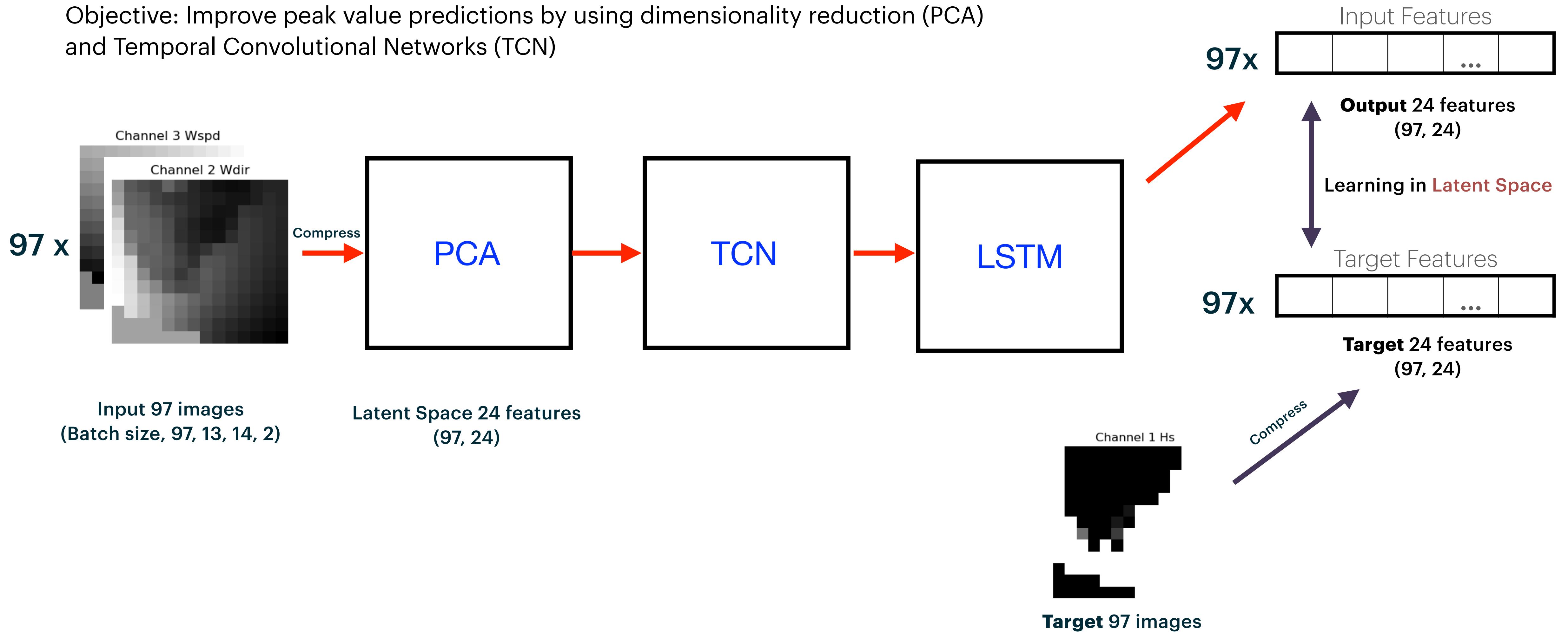
MODELLING APPROACH 2: 2D CNN+LSTM

Objective: Incorporate spatial features by converting grid data into image format.



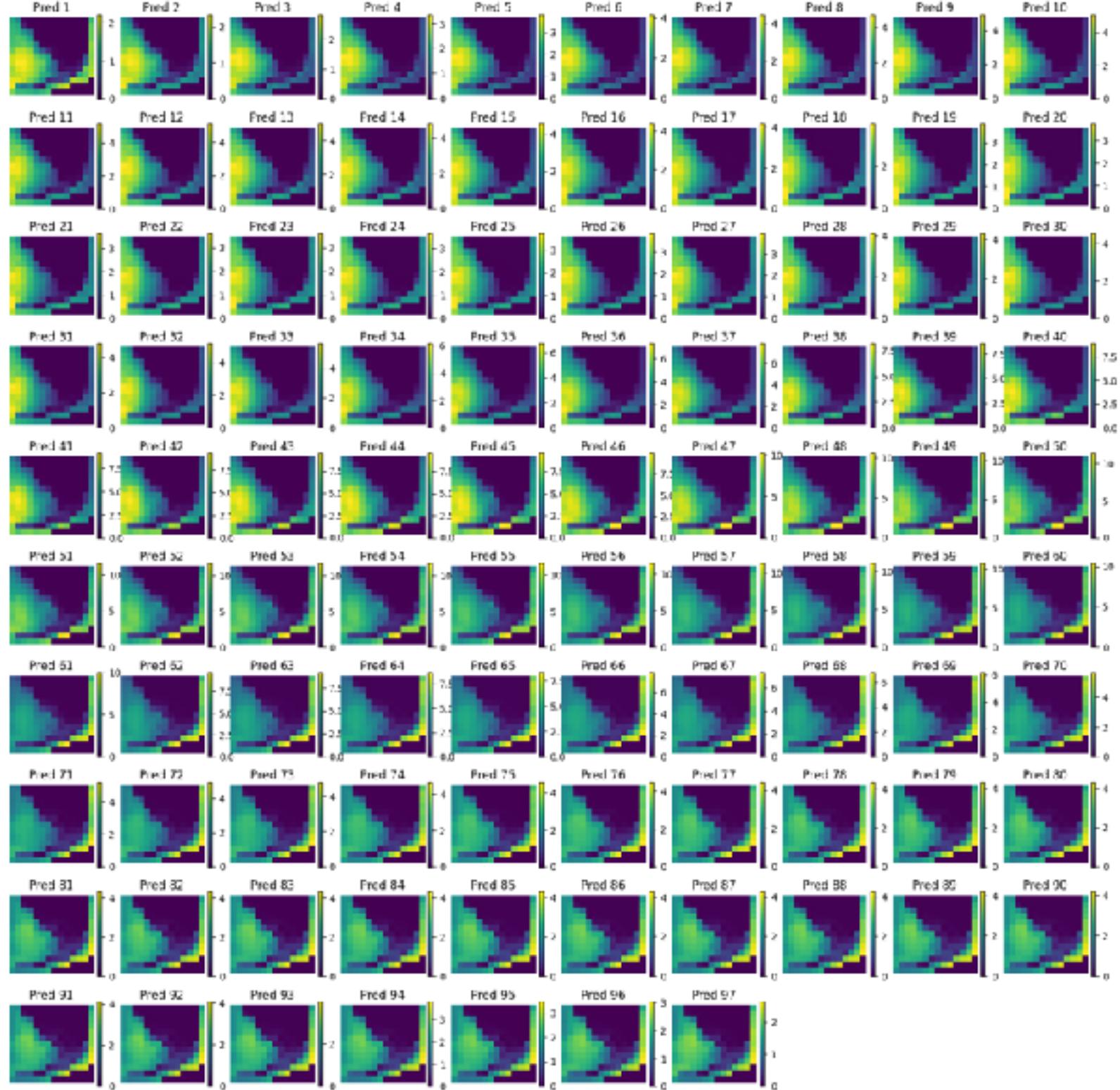
MODELLING APPROACH 3: 2D PCA-TCN-LSTM

Objective: Improve peak value predictions by using dimensionality reduction (PCA) and Temporal Convolutional Networks (TCN)

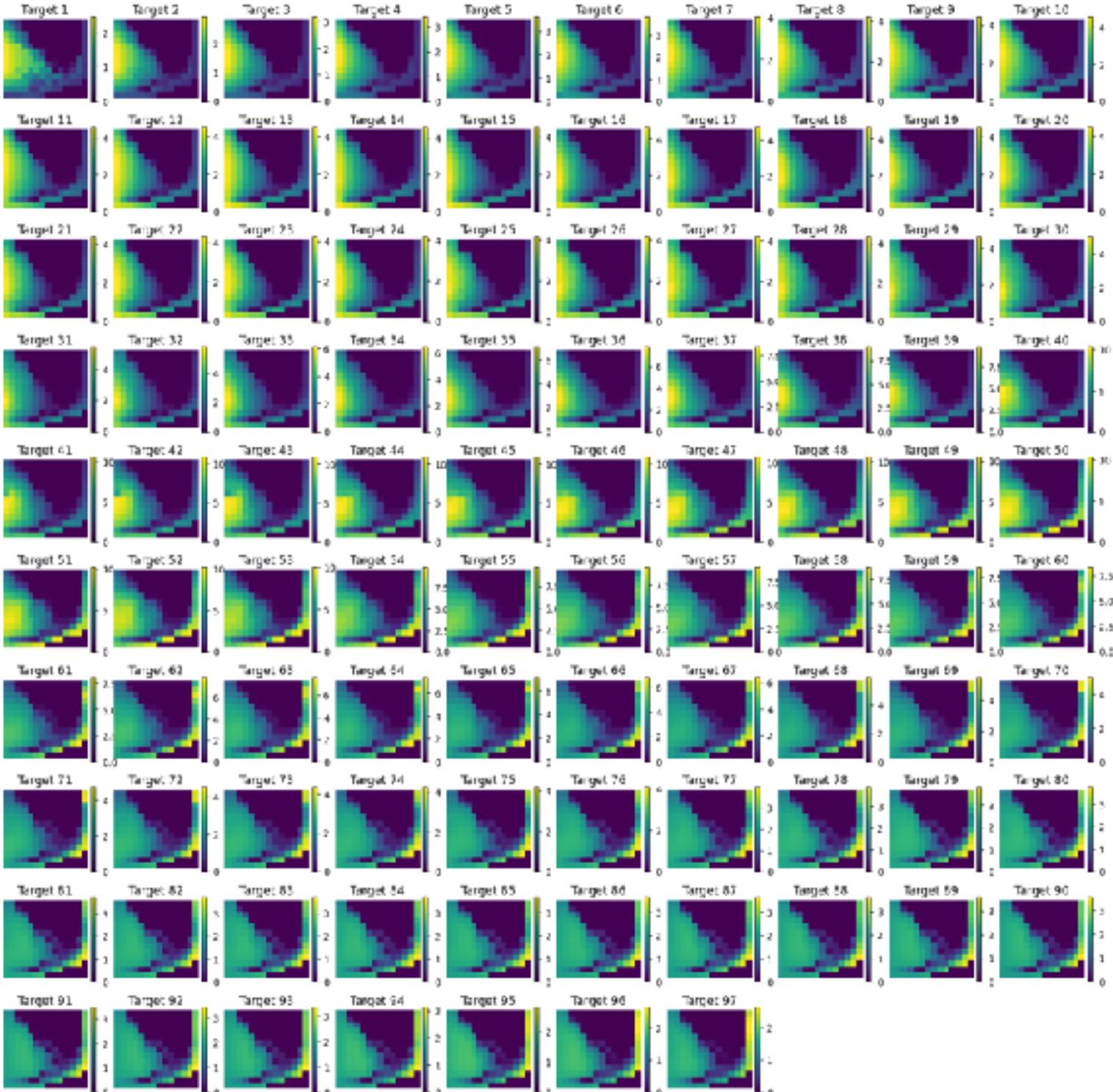


DECOMPRESSED IMAGES

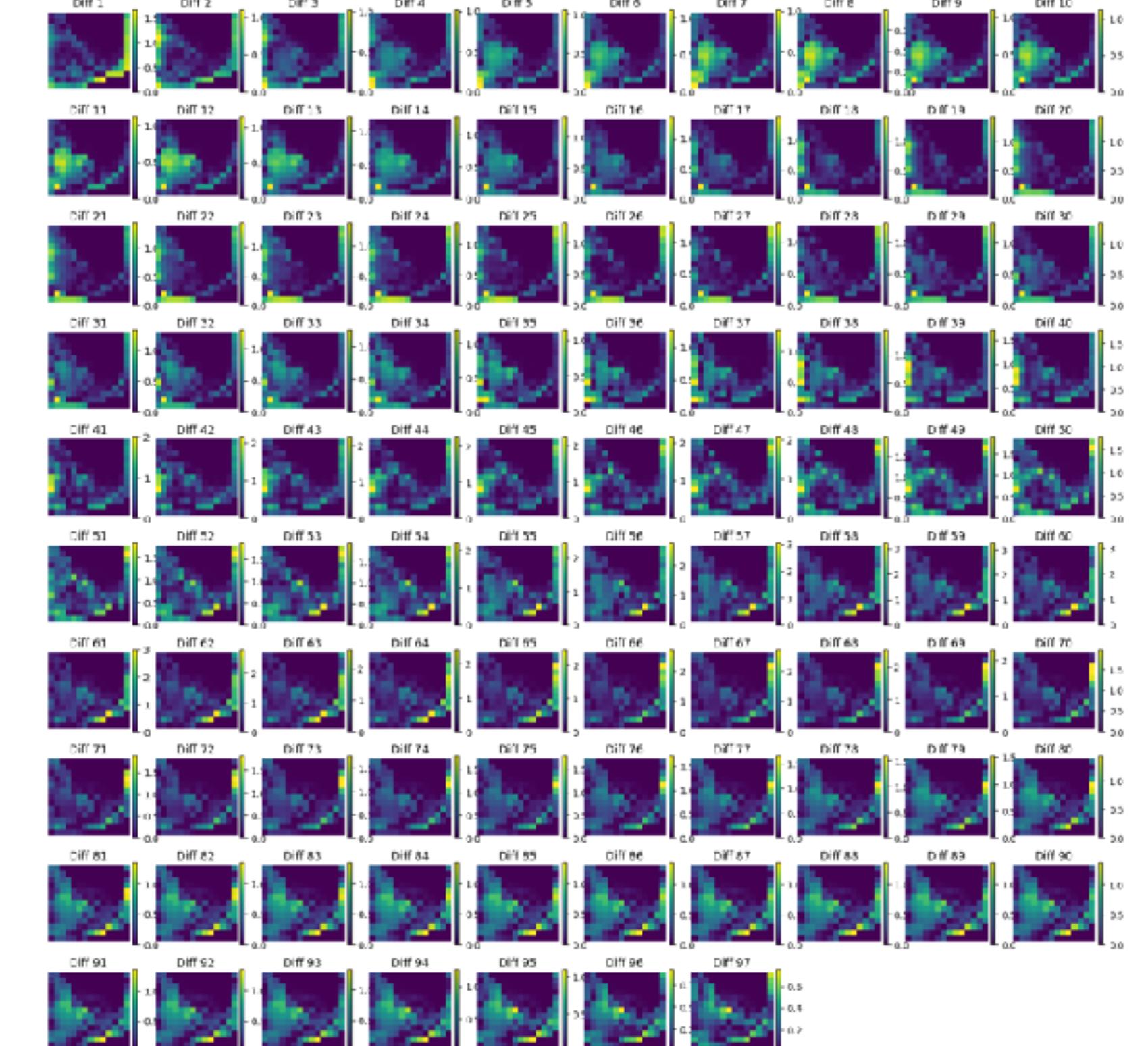
Predictions



Targets



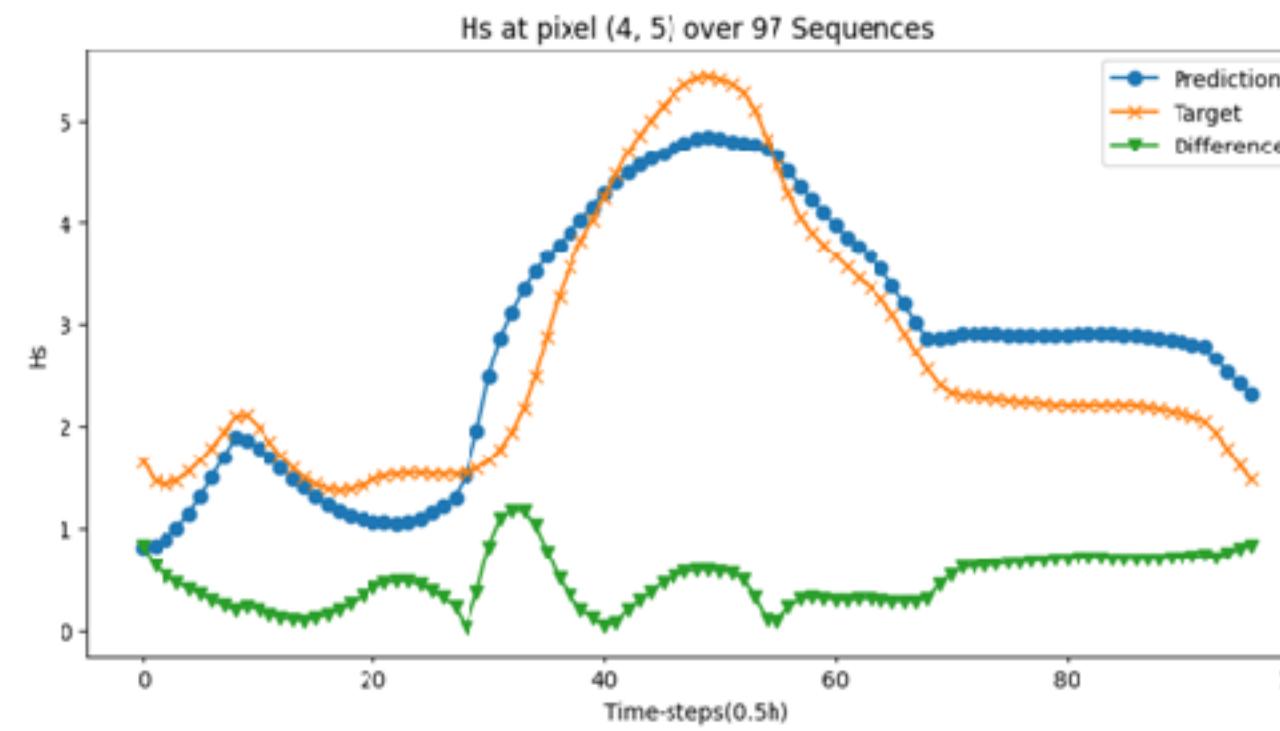
Differences



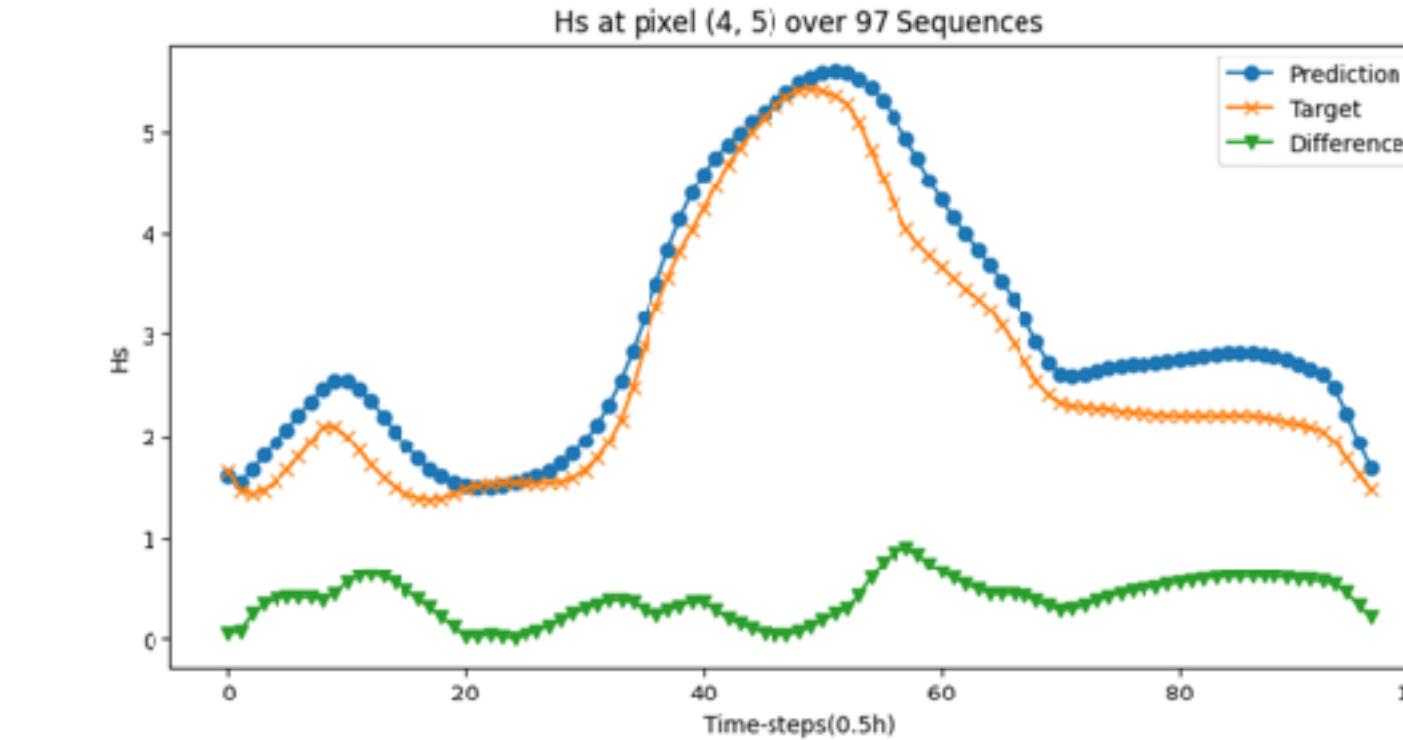
PHYSICAL VS LATENT MODEL

Physical: CNN + LSTM

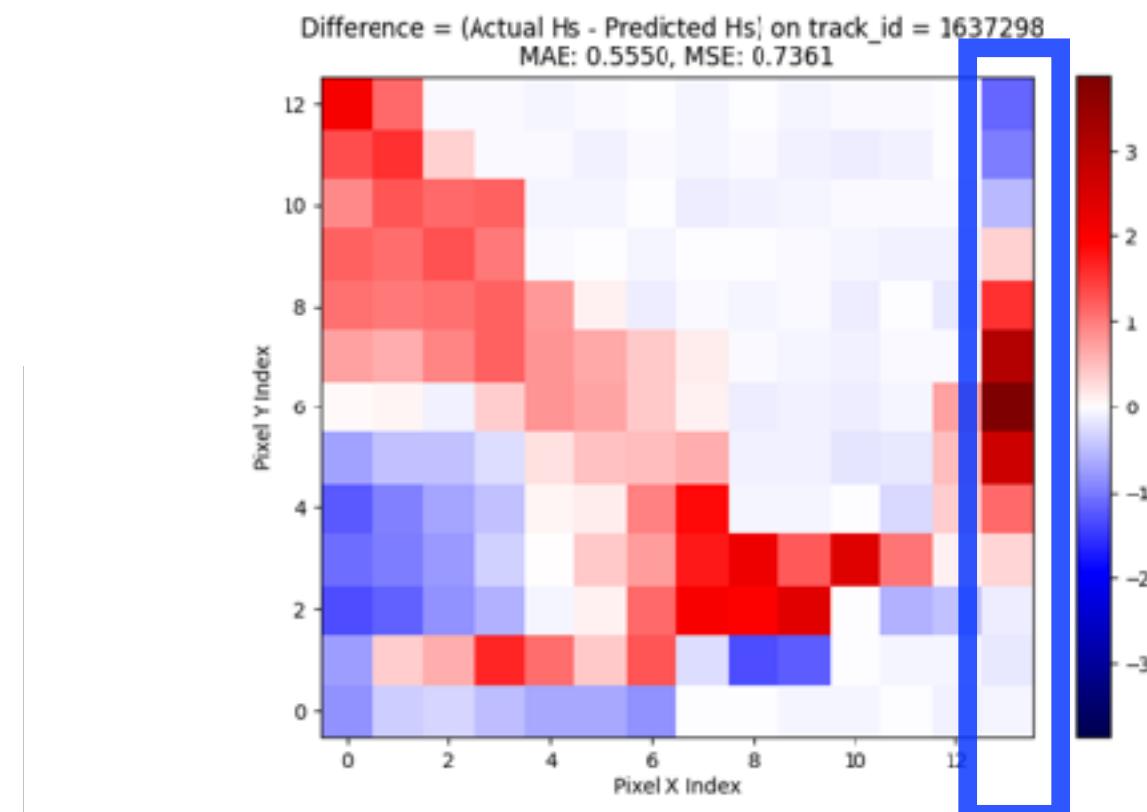
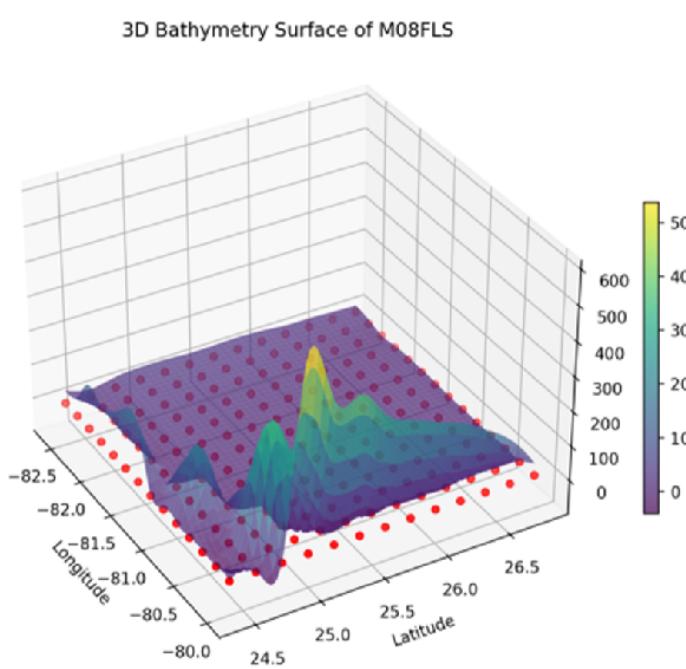
**Whole Track
prediction**



Latent: PCA-TCN-LSTM

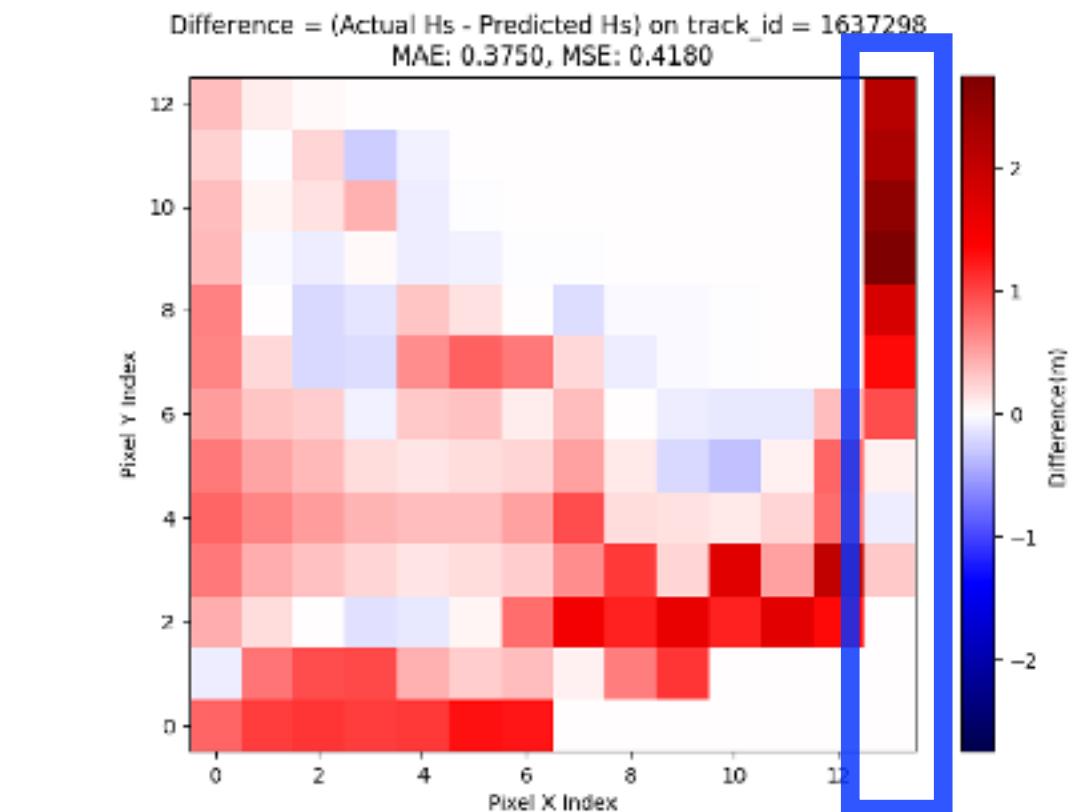


**Peak value
prediction of all grid points**



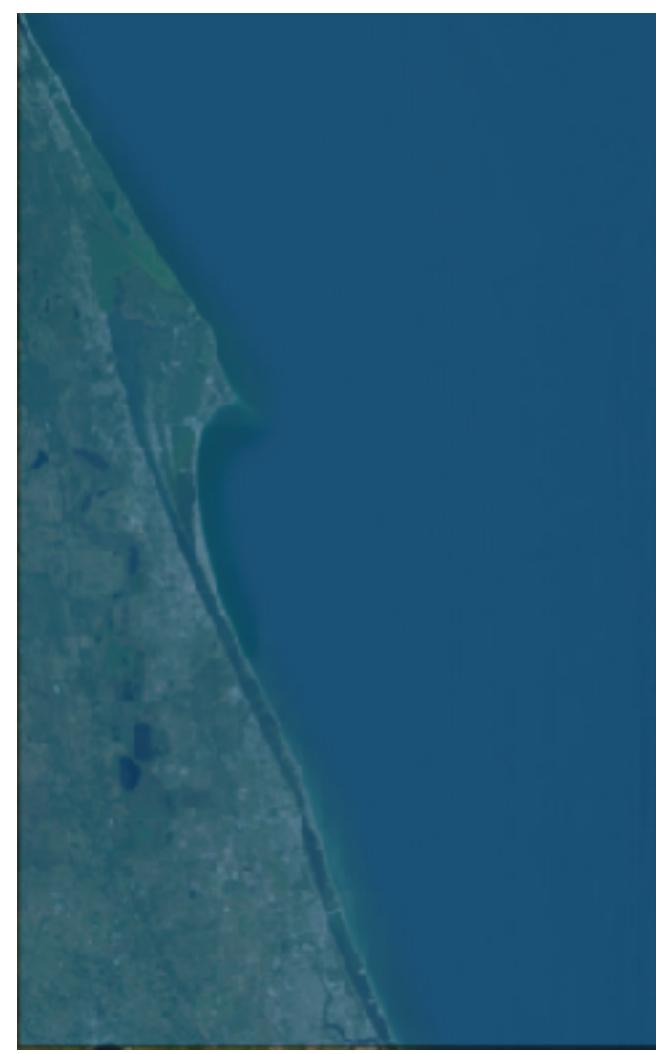
MAE: 0.74 m
MSE: 0.55 m²

Difference between actual and predicted peak wave values

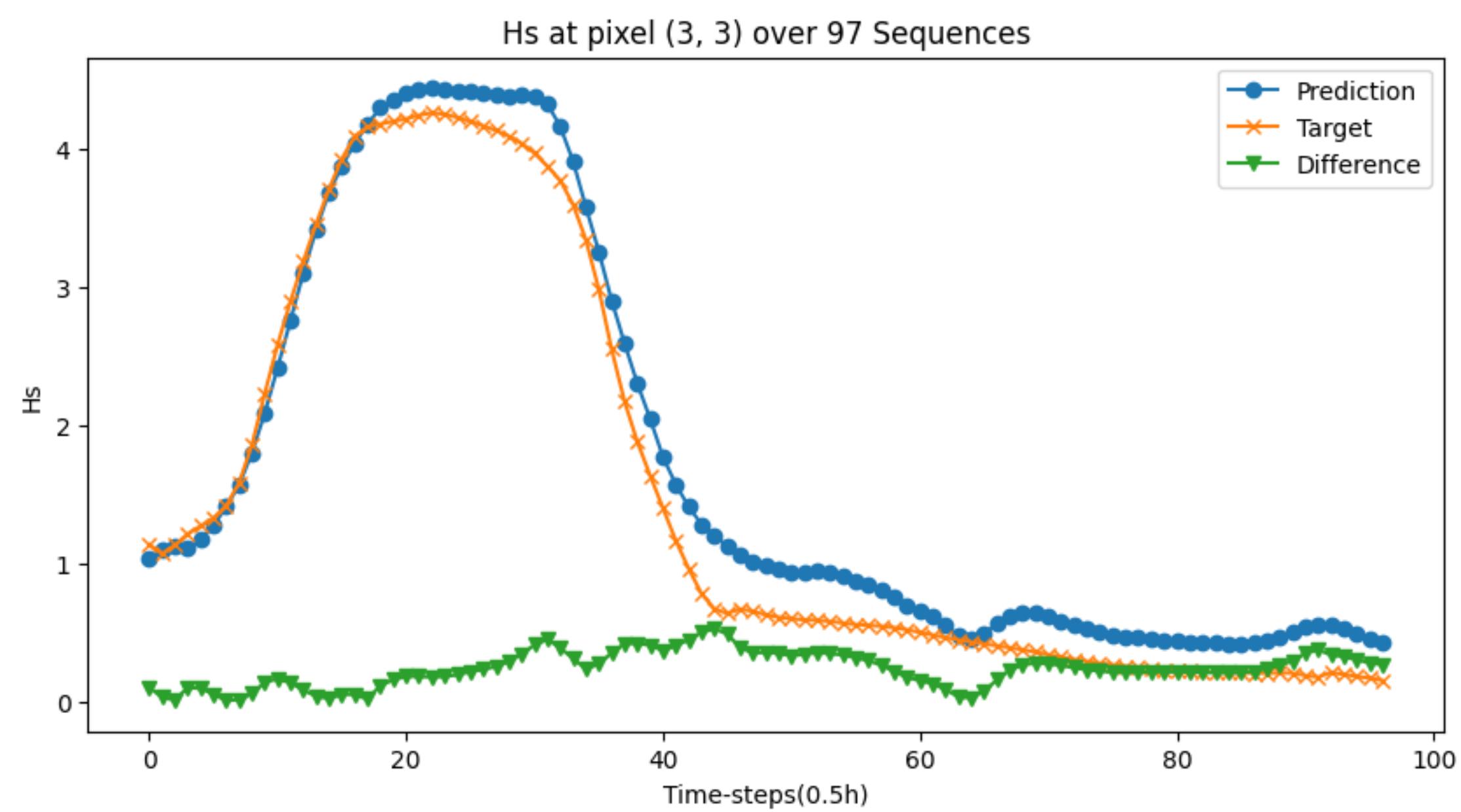


MAE: 0.38 m
MSE: 0.41 m²

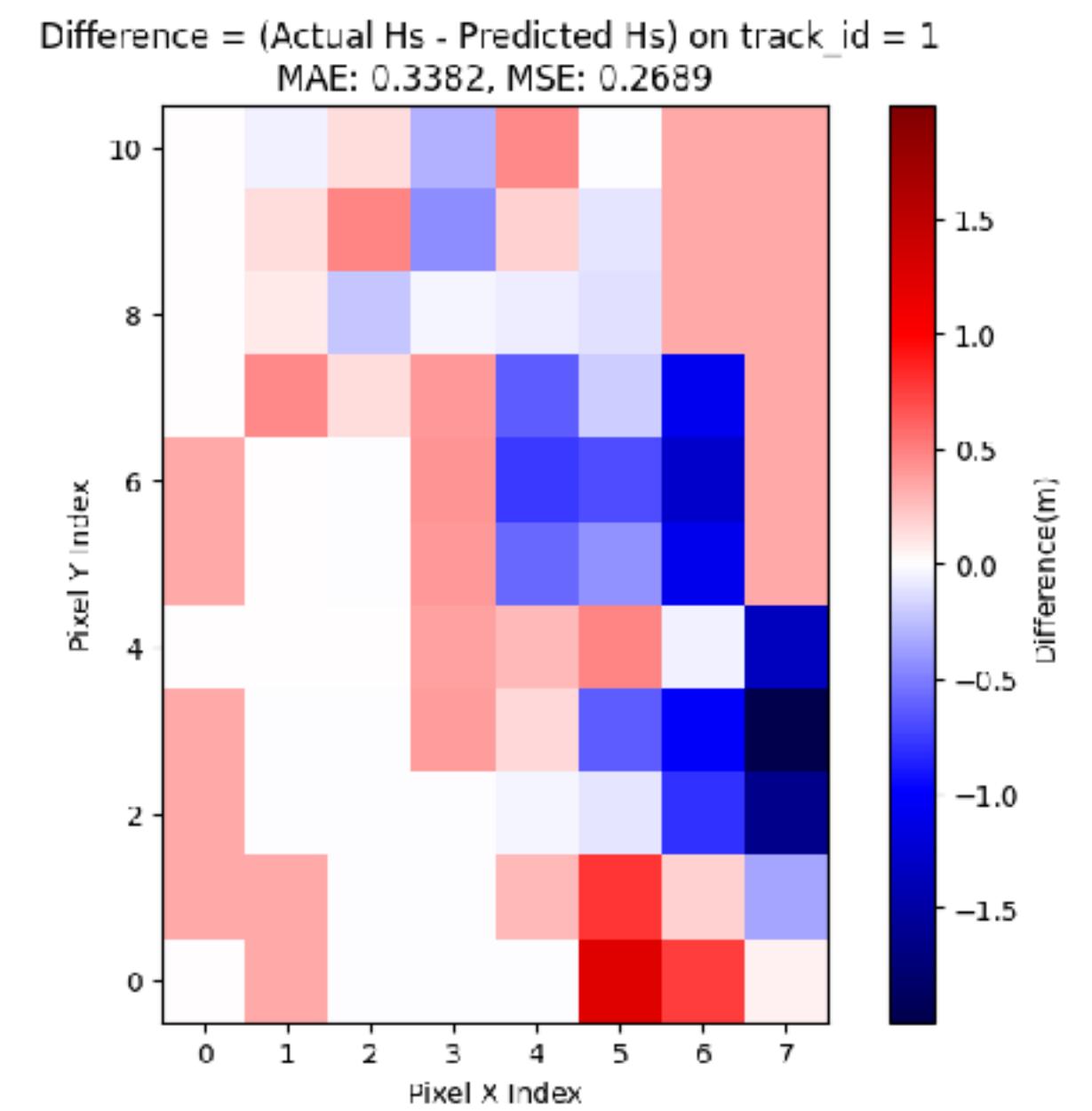
TESTING PERFORMANCE ON EAST FLORIDA



**Area of interest
East Florida**



Whole track prediction



Peak values differences

CONCLUSION

Deep Learning Potential:

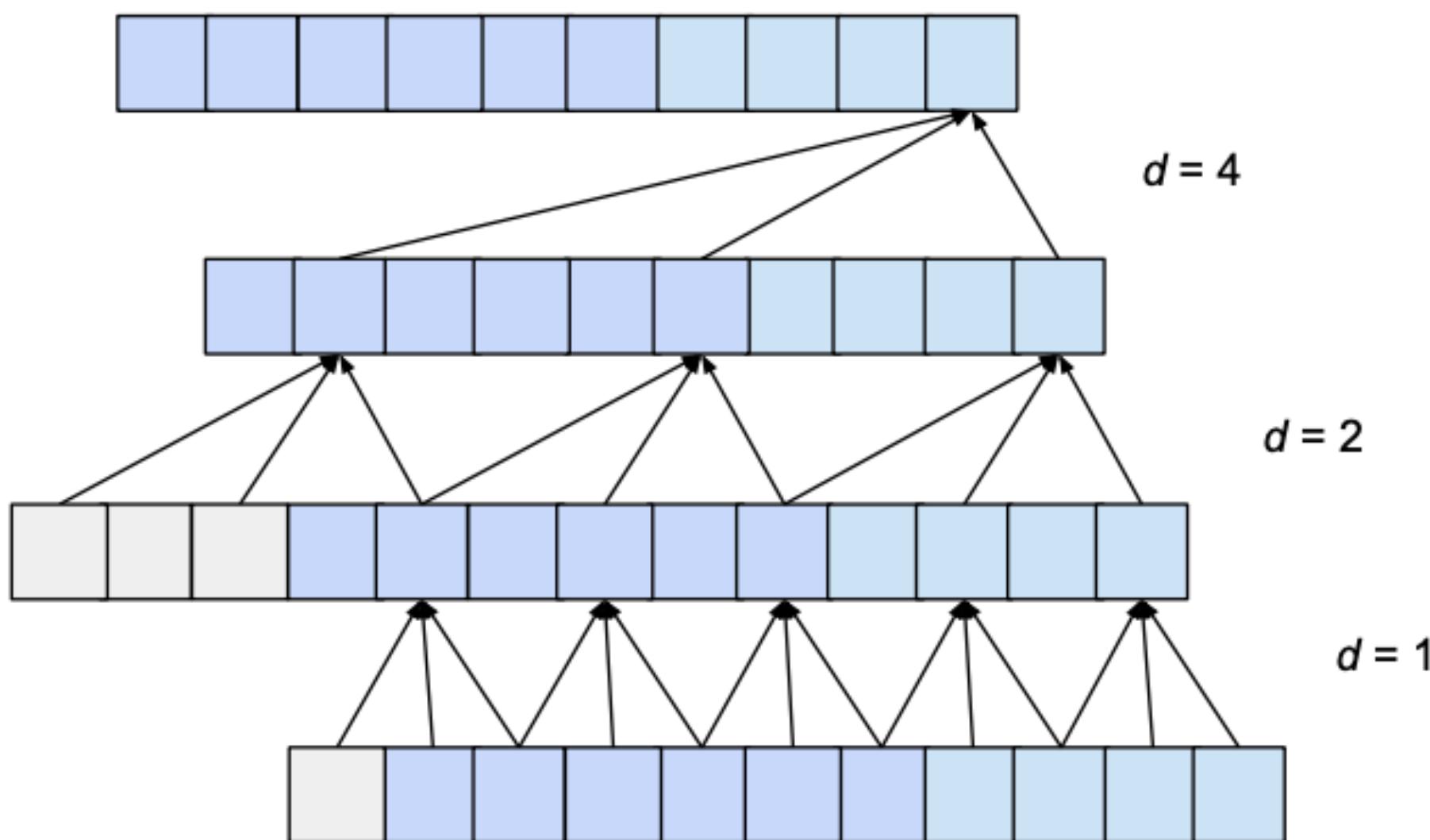
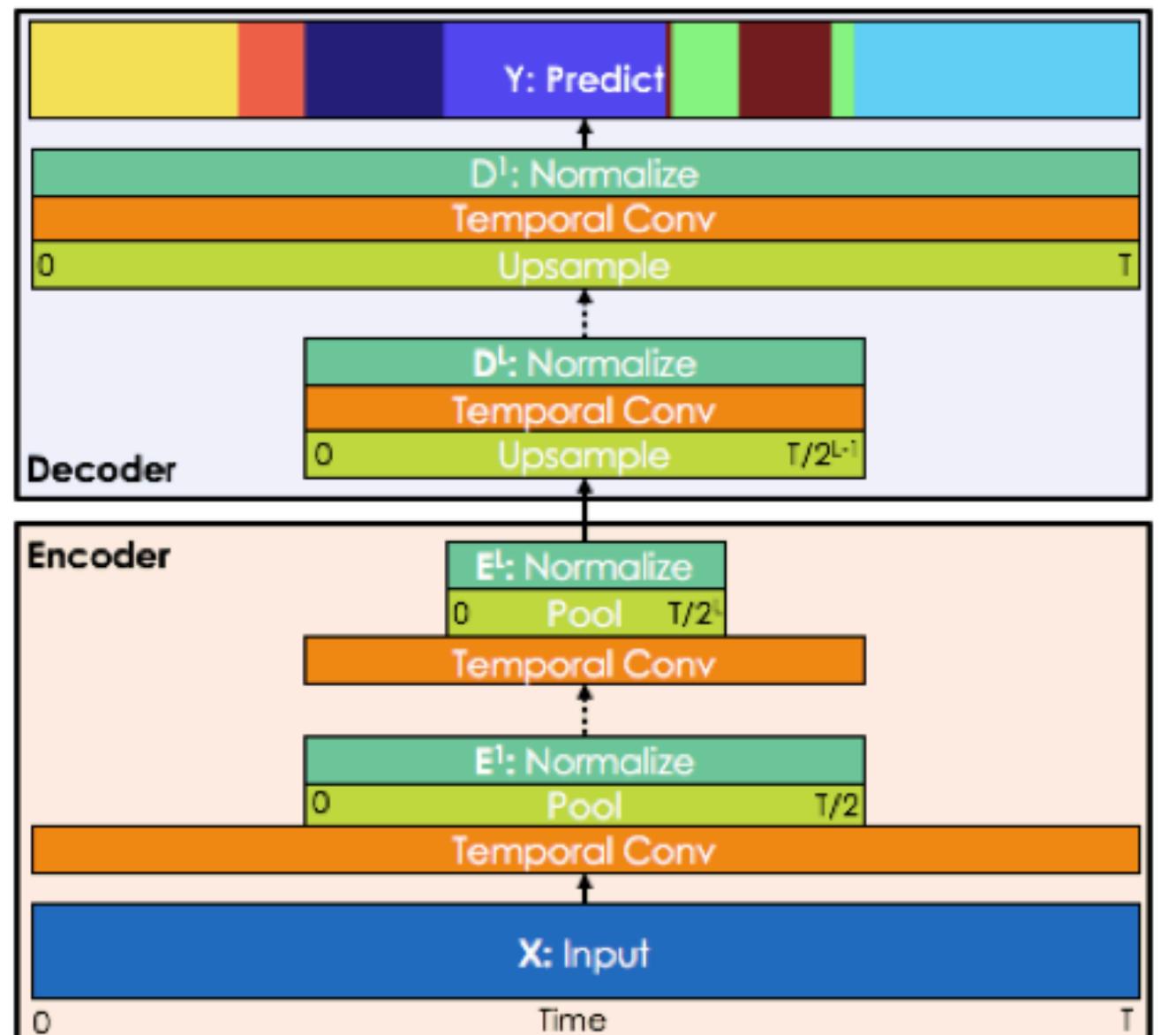
- Demonstrated the effectiveness of deep learning models in predicting significant wave heights (Hs) from tropical cyclones, offering a faster alternative (<5 mins) to MIKE21.

Model Progression:

- Baseline LSTM struggled with generalisation and skewed data.
- CNN-LSTM improved spatial feature learning but faced difficulties with peak Hs predictions.
- PCA-TCN-LSTM: Major improvement, leveraging PCA for data compression and focusing on critical features, excelling in extreme wave height predictions.

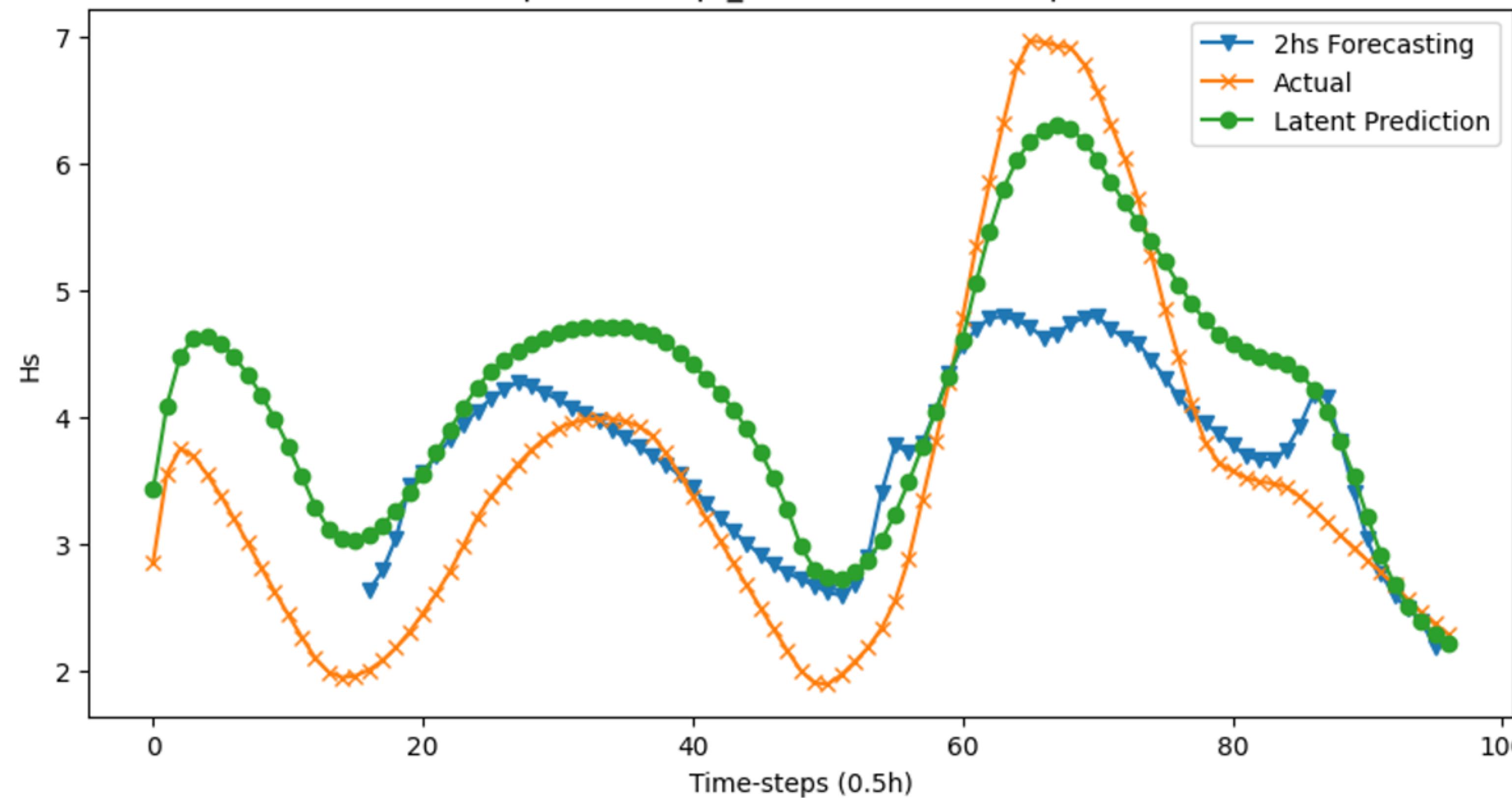
Q&A

TCN



Lea, 2016

Hs at pixel (8, 0) pt_id =113 over 97 Sequences



Regression vs Forecasting