TidalMark: AI-Driven Forecasts for Next-Gen Water Level Monitoring

Lucas Raicu
Advisors: Daniel Grzenda, Prof. Kyle Chard
August 5th, 2025

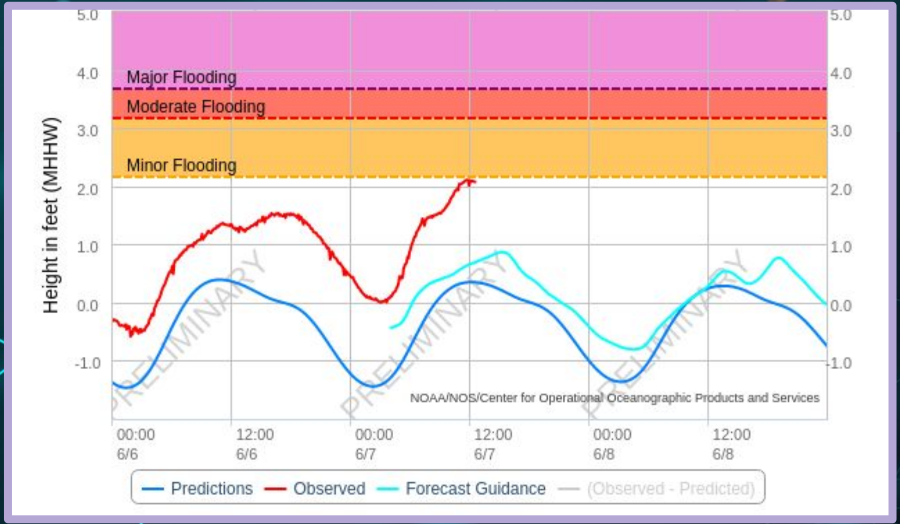
About me

- Rising Senior at Glenbrook South High School (GBS)
- Coursework
  - Honors CS and AP CS A
  - Fundamentals of AI through Oakton College
- Founder of CS Honor Society at GBS
- Prior projects
  - CRYPTEX: fine-grained CRYPTocurrency datasets EXploration [GCASR'24]
  - Slow Turtle BackTesting with CRYPTEX [Tech Report '25]
  - CryptoHeat: A Smart, Cost-Efficient Personal Heater Powered by Blockchain Mining [GCASR'25]

https://lucasraicu.github.io

Problem Statement

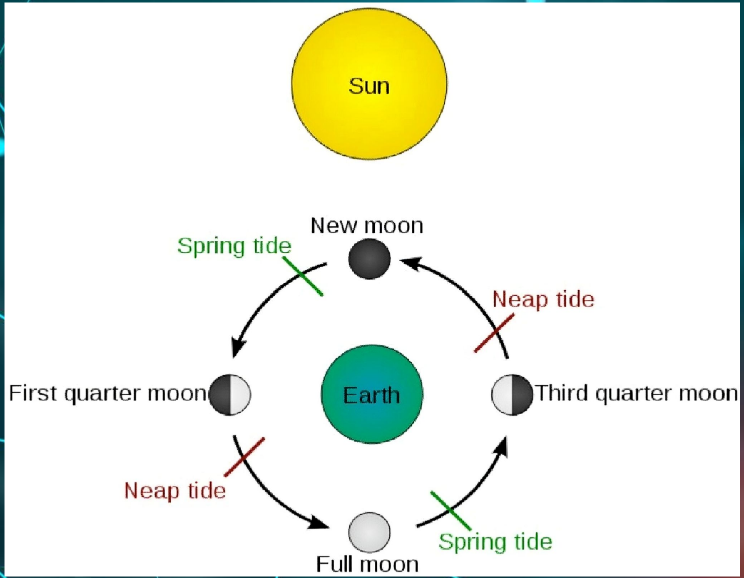
Coastal flood forecasting remains hampered by traditional models' inability to capture complex spatiotemporal dynamics, limiting lead time and accuracy.

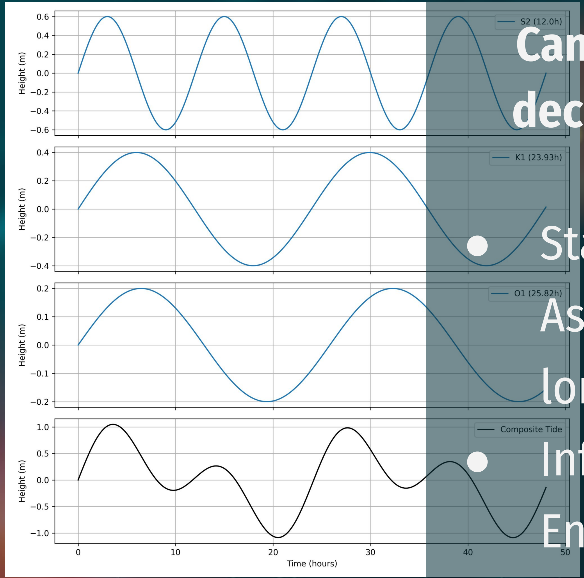


THE UNIVERSITY OF CHICAGO

globus labs

State-of-the-art Solution
Harmonic Analysis





Can predict water level decades into the future

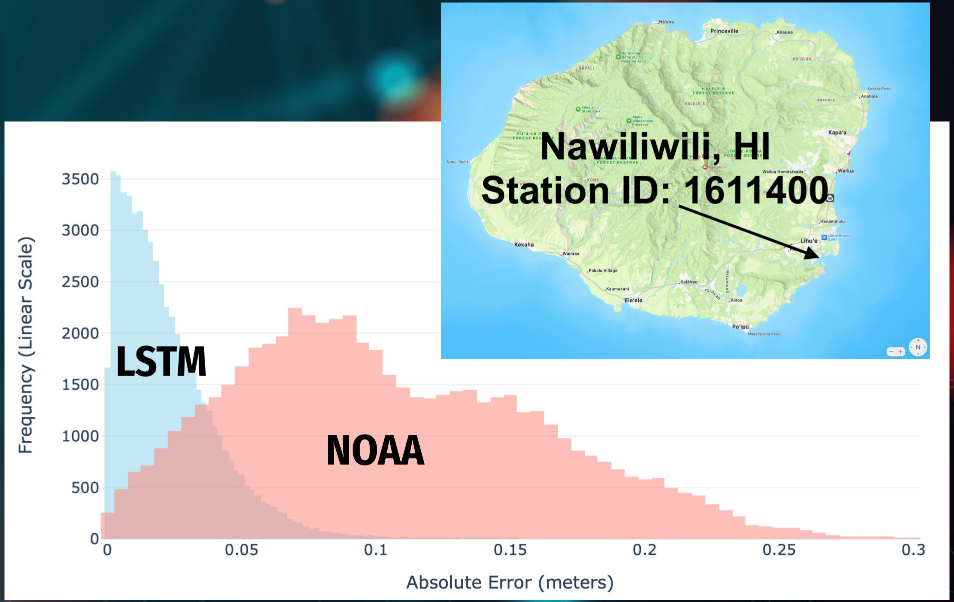
- Stationarity and Linearity Assumptions (e.g. long-term sea-level rise)
- Inflexibility under Rapid Environmental Change

Our Approach

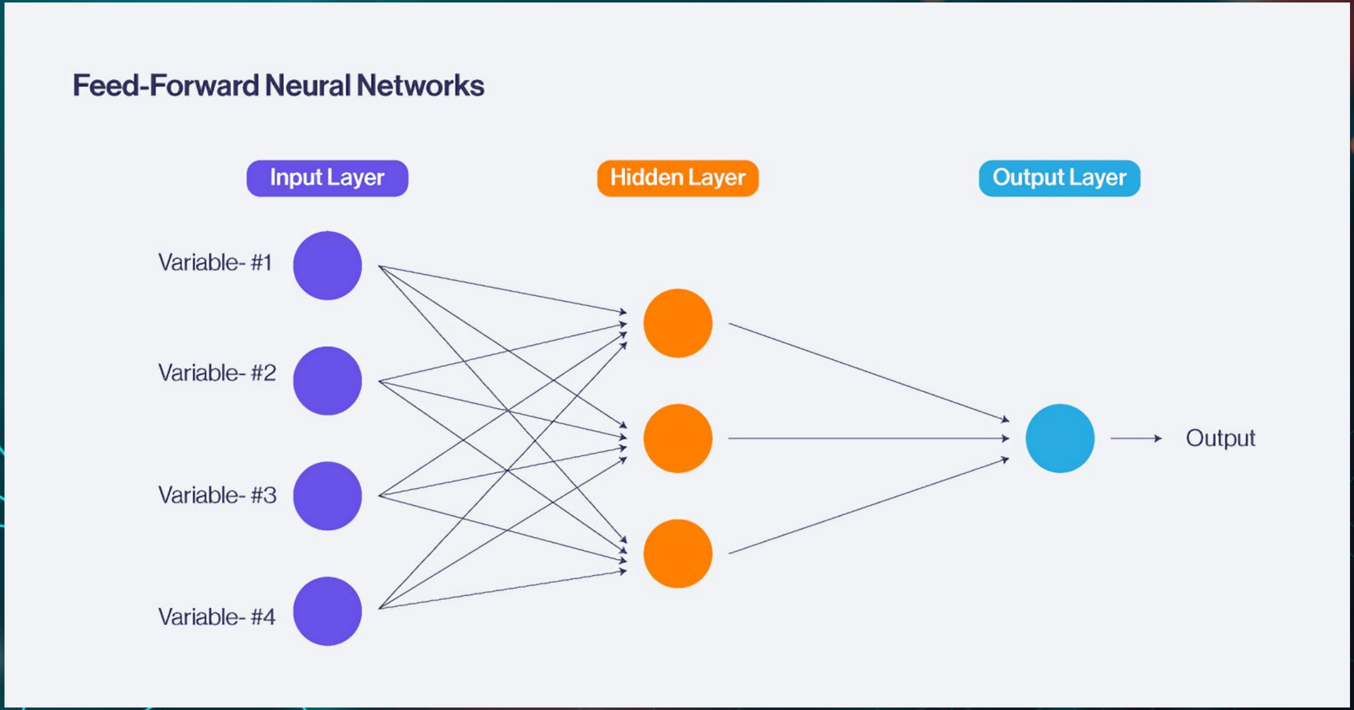
Temporal-spatial Water-level Data

Artificial Intelligence

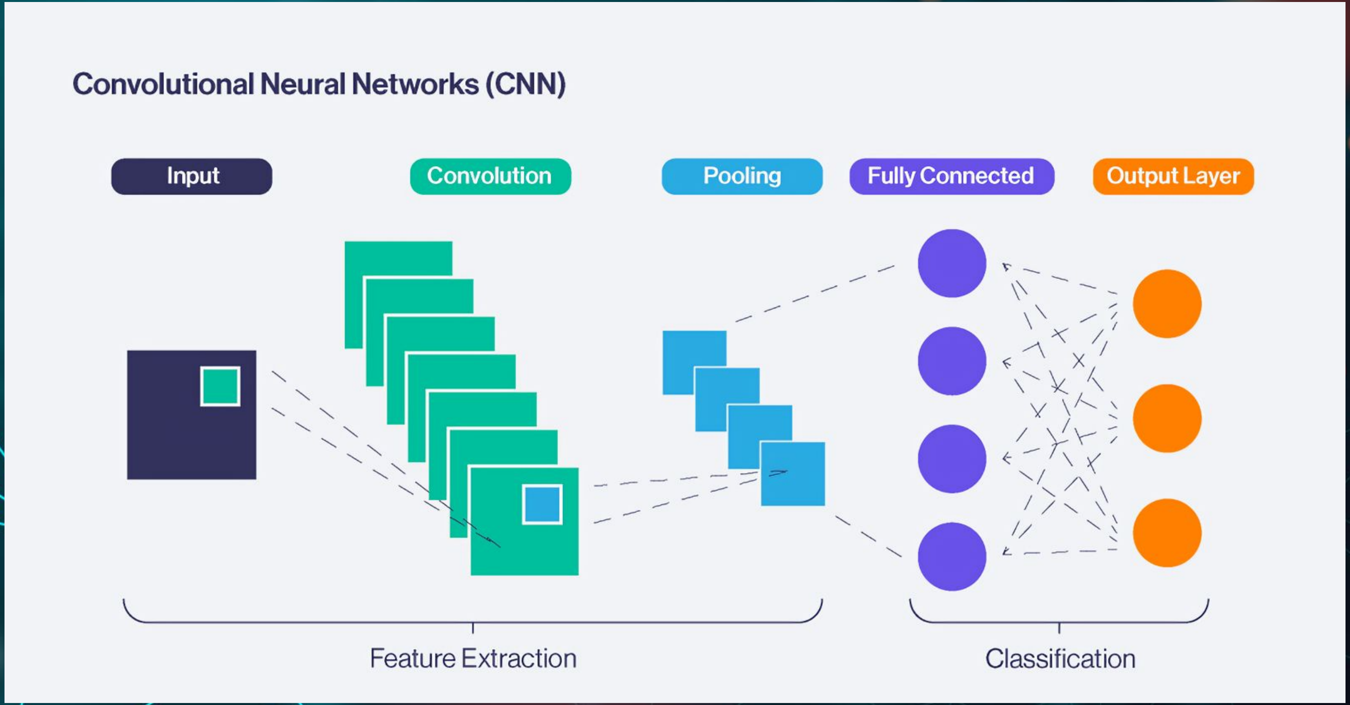
Accurate Future Trend Predictions



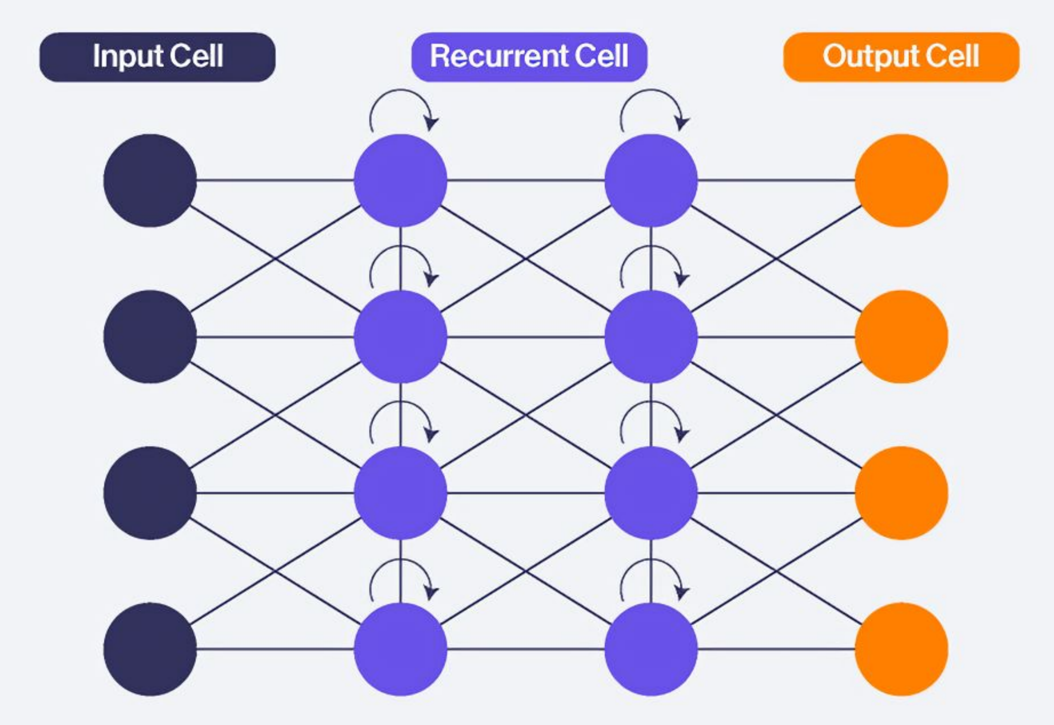
Machine Learning as an Option



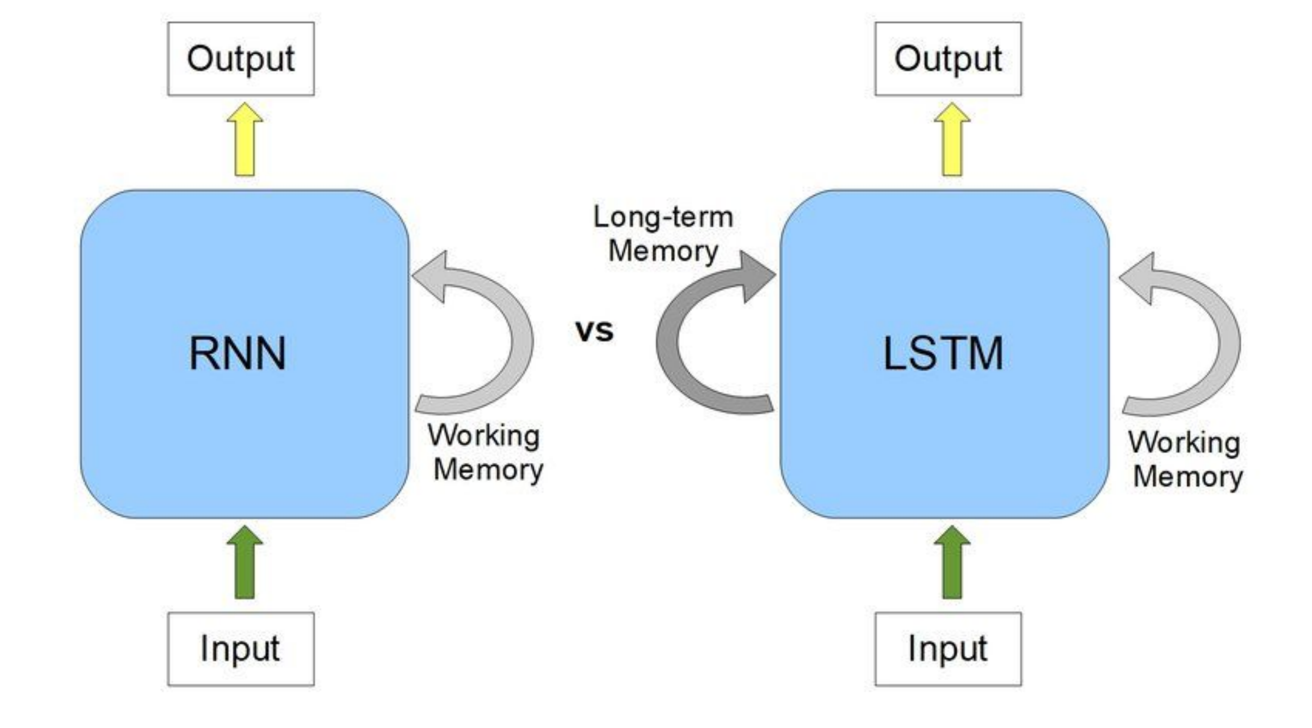
Machine Learning as an Option (cont.)



Recurrent Neural Networks (RNN)

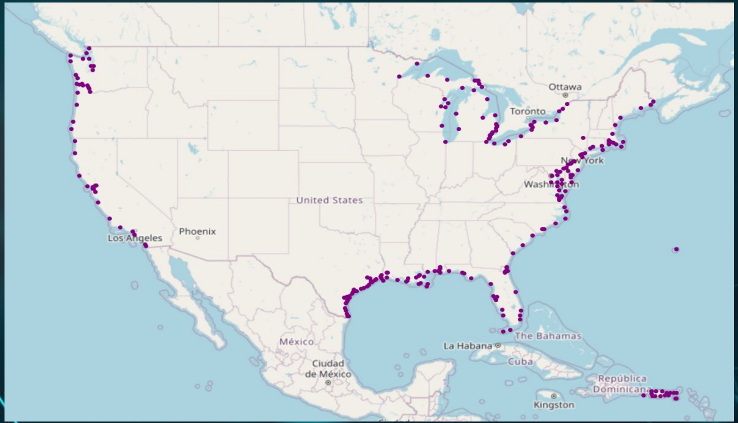


Simple RNN vs LSTM



NOAA Dataset

- Dataset size: 82 gigabytes
  - 217 coastal water level stations (w/ full Metadata)
  - Over 127 million data points (6-minute granularity)
  - Raw & predicted data samples
  - Focused on one station (Nawiliwili, HI - Station ID: 1611400)



THE UNIVERSITY OF CHICAGO

globus labs

TidalMark Framework

- Sequential models to capture long-term-dependencies in water-levels
- Performed a single-shot walk-forward validation
  - 80% training data, 10% validation, 10% testing
- Params: seq_len, batch_size, learning_rate, hidden_size, num_layers
- Output sequence of 7 days (logs at 1, 3, 5, 7)
- Visualization
  - Absolute Error histogram
  - Train vs Validation Loss
  - Correlation Heatmap
  - Hyper-Parameter comparisons through Box, Scatter, and Violin Plots

Data Format

Stations Info (tsv)

~/Documents/summer-2025/fountain/data/processed/noaa/stations/stations.tsv   lucas

tidal	greatlakes	shefcode	state	timezone	timezonecorr	id	name	lat	lng	affiliations	tideType	established	removed	origyear	missing_percent
True	False	NWWH1	HI	HAST	-10	1611400	Nawiliwili	21.9544	-159.3561	NWLON	Mixed	1954-11-24 00:00:00.0		1991-02-16 00:00:00.0	1.6903284477206767e-06
True	False	OOUH1	HI	HAST	-10	1612340	Honolulu	21.303333	-157.864528	NWLON	Mixed	1905-01-01 00:00:00.0		1989-01-20 00:00:00.0	0.0
True	False	PRHH1	HI	HAST	-10	1612401	Pearl Harbor	21.3675	-157.963898	PORTS	Mixed	2023-05-31 00:00:00.0		2023-05-31 00:00:00.0	
True	False	MOKH1	HI	HAST	-10	1612480	Mokuoloe	21.433056	-157.79	NWLON	Mixed	1957-05-03 00:00:00.0		1989-01-21 00:00:00.0	0.0
True	False	KLIH1	HI	HAST	-10	1615680	Kahului, Kahului Harbor	20.895	-156.469167	NWLON	Mixed	1946-12-19 00:00:00.0		1989-01-29 00:00:00.0	0.0

- Data pre-processing
  - Filter (verify, station ID)
  - Sort (time)

Sorted & Filtered Station Data (.tsv) - Nawiliwili

time	value	sigma	quality	inferred	flat	roc	threshold	station	datum
2018-01-01 00:00:00	0.273	0.002	v	0	0	0	0	1611400	mllw
2018-01-01 00:06:00	0.278	0.003	v	0	0	0	0	1611400	mllw
2018-01-01 00:12:00	0.277	0.003	v	0	0	0	0	1611400	mllw
2018-01-01 00:18:00	0.276	0.006	v	0	0	0	0	1611400	mllw

Table of Contents

- Parameters (Seq Len vs Output Len)
- Models comparable to LSTMS
- Hyper-Parameter Sweep across Params and Model-bases
  - Similar performance and limited Parameter Variability
- Base LSTM large sweep to determine parameter correlations
- Best config to compare with multivariate approaches
  - Neighbor, NOAA, and Constituents / Workflow Diagram
- Challenges
  - Too many open file handle, out of memory errors, long computational time
- Future Work ⇒ Dynamic Graph Neural Networks and River Predictions

Hyperparameter Search

LSTM Large Sweep –
120 configurations

- Seq_Len: [7, 14]
- BS: [32, 64, 128, 256, 512]
- LR: [1e-3, 1e-4, 1e-5]
- HS: [32, 64]
- Num_Layers: [1,2]

Many Models Sweep –
180 configurations

- Seq_Len: [7]
- BS: [32, 64, 128]
- LR: [1e-3, 5e-4, 1e-4]
- HS: [32, 64]
- Num_Layers: [1,2]
- Models: [LSTM, BILSTM, CONV-LSTM, GRU, ATTN]

Multivariate Experiments –
Input Dim: Primary Station +

- Neighbor
- NOAA
- 37 Harmonic Constituents

Hyperparameter Search #3

Hidden Size
[32, 64]

Models
[LSTM]

Num Layers
[1, 2]

120

Learning Rate
[1e-3, 1e-4, 1e-5]

Sequence Length
[7, 14]

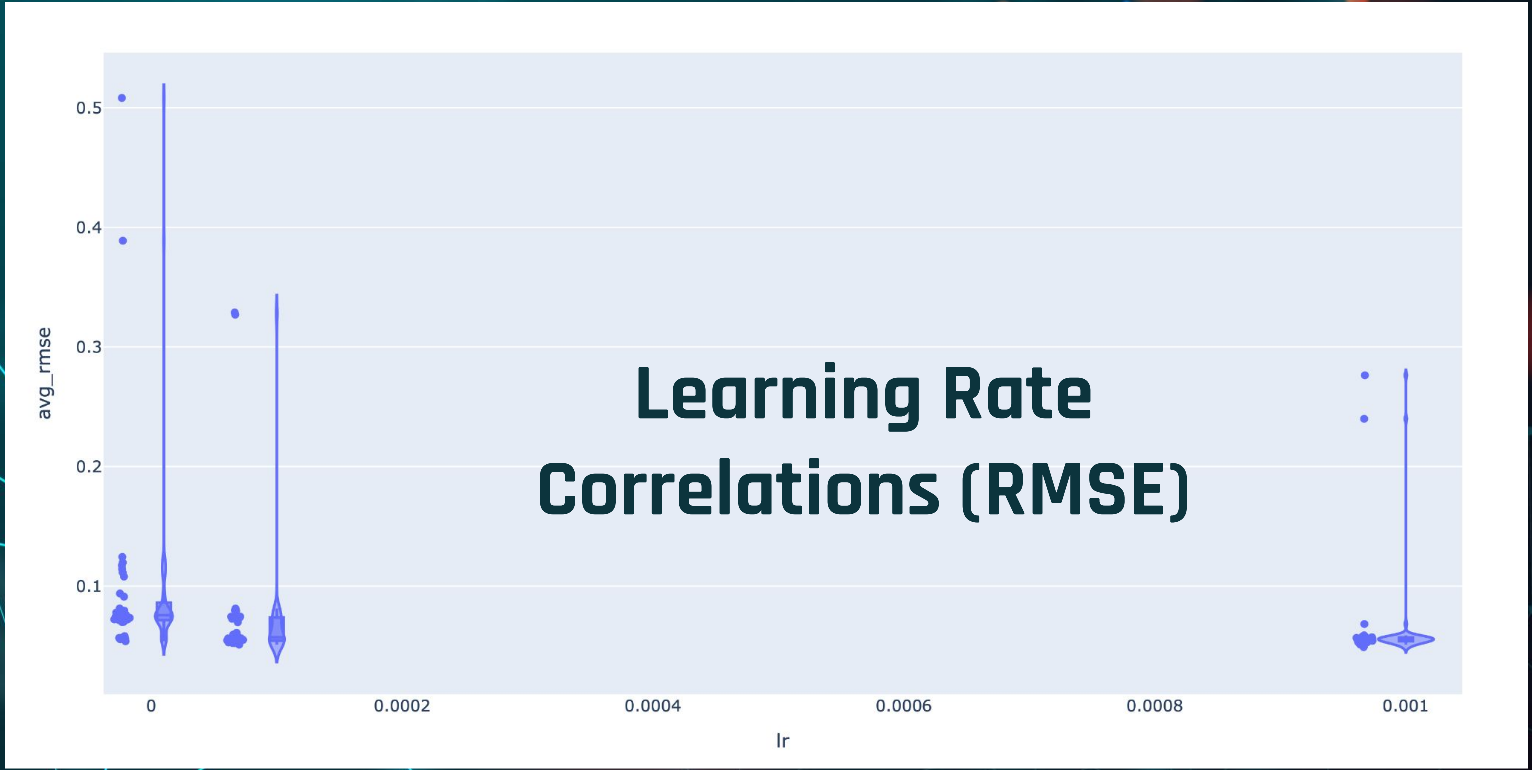
Batch Size
[32, 64, 128, 256, 512]

THE UNIVERSITY OF CHICAGO

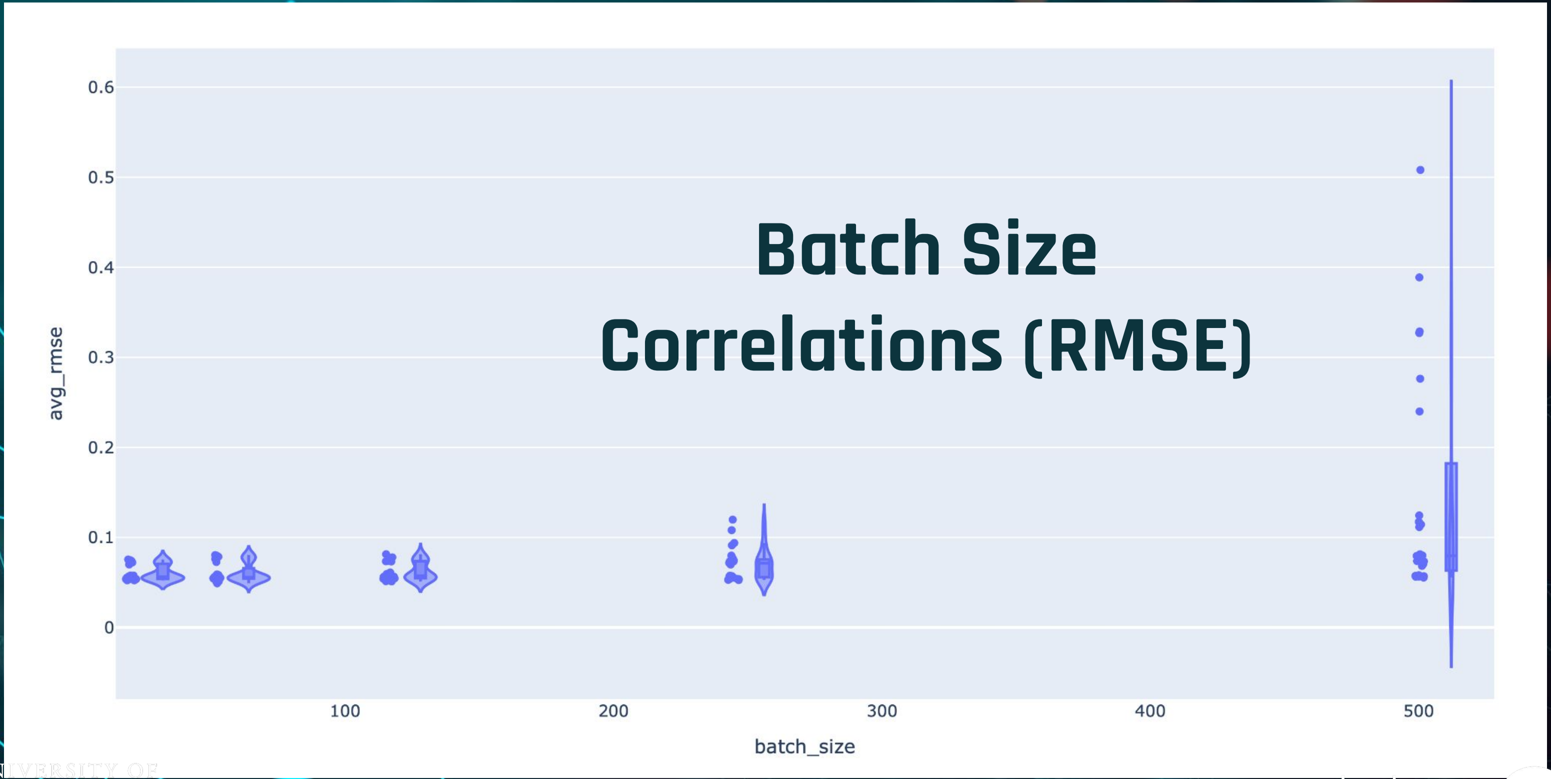
globus labs

How can I visualize hyperparameter sweeps?

LSTM Large Sweep Results



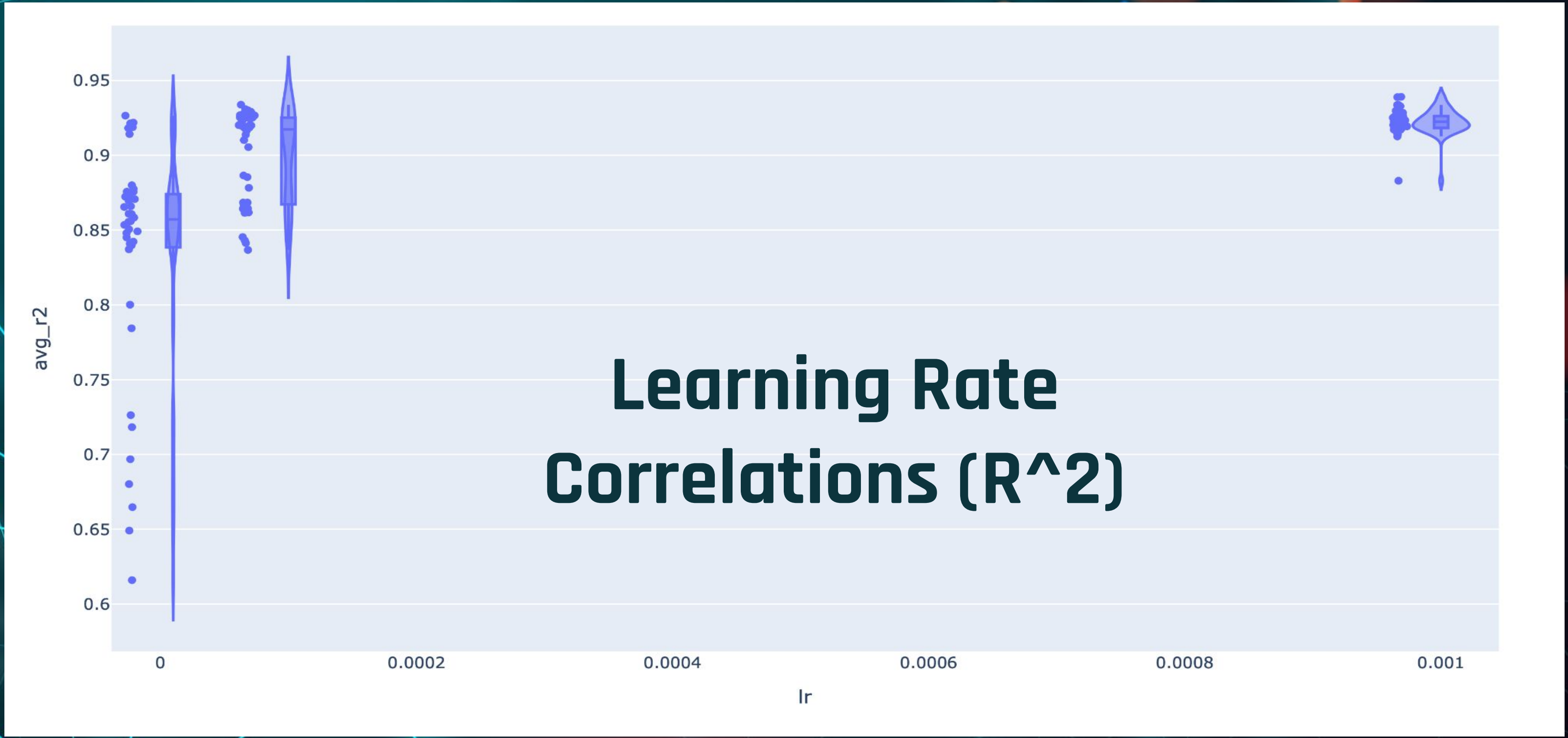
LSTM Large Sweep Results (cont.)



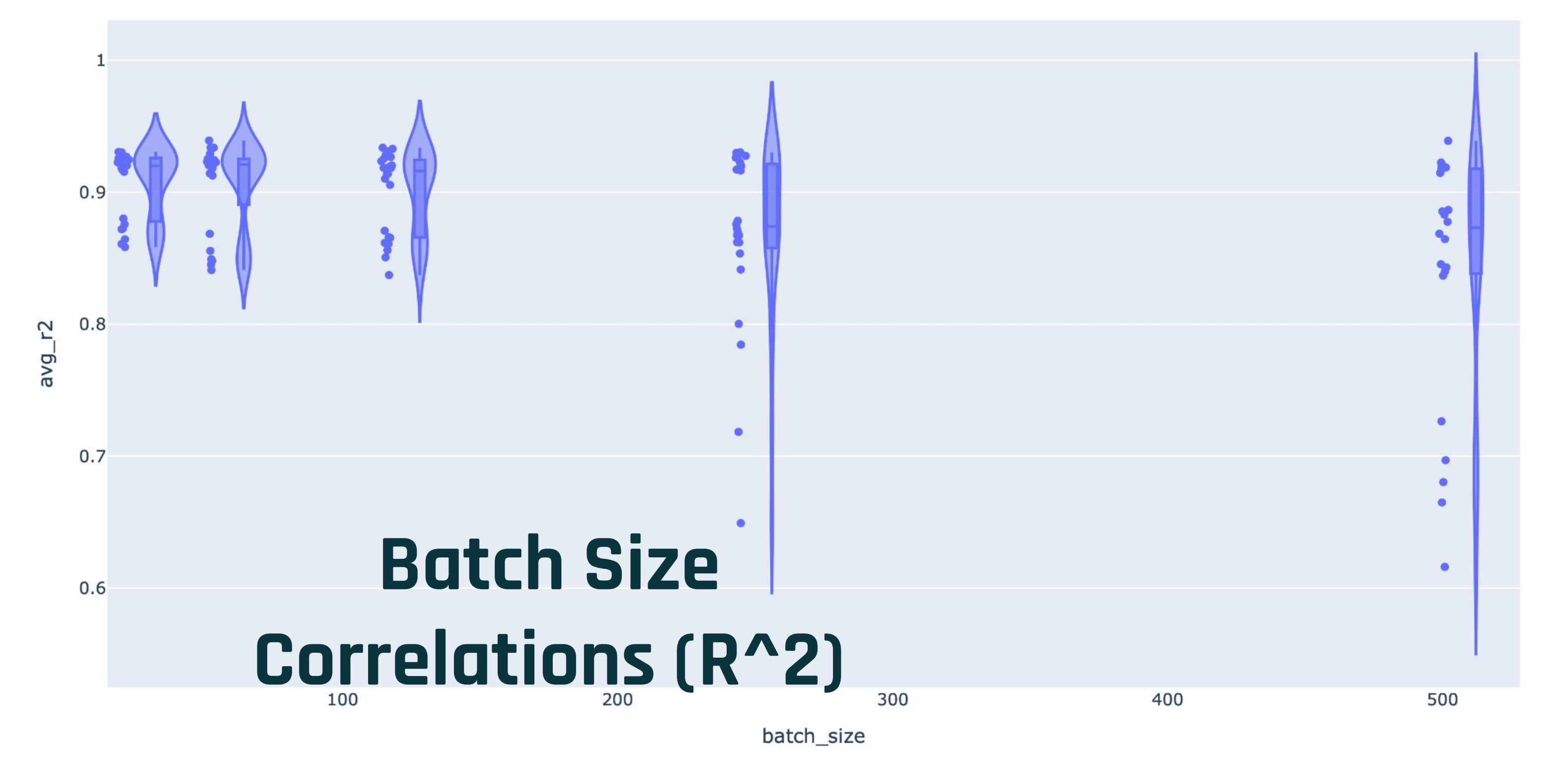
THE UNIVERSITY OF CHICAGO

globus labs

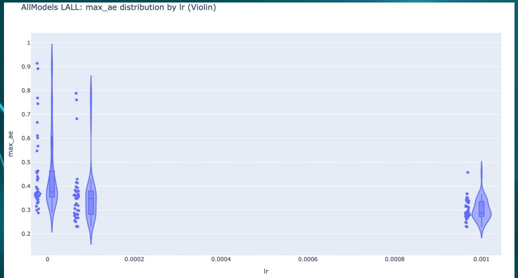
LSTM Large Sweep Results

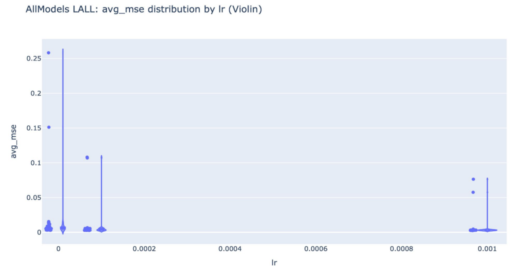


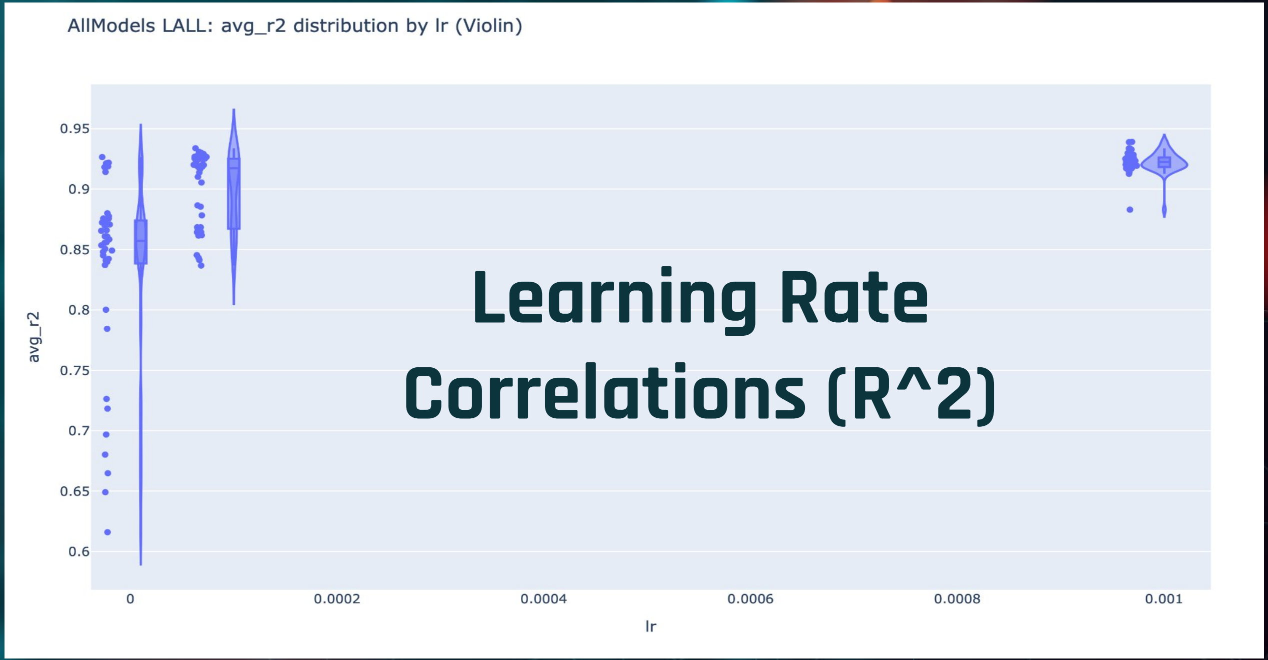
LSTM Large Sweep Results (cont.)



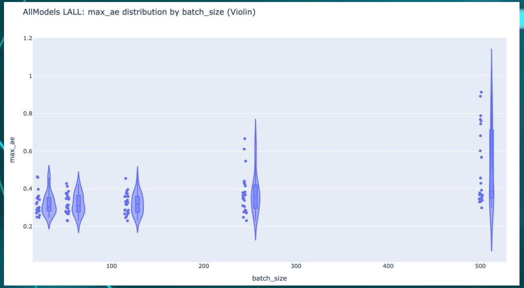
LSTM Large Sweep Results (cont.)

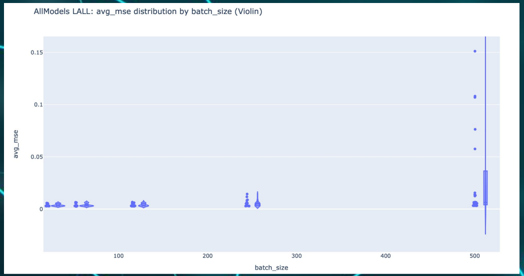


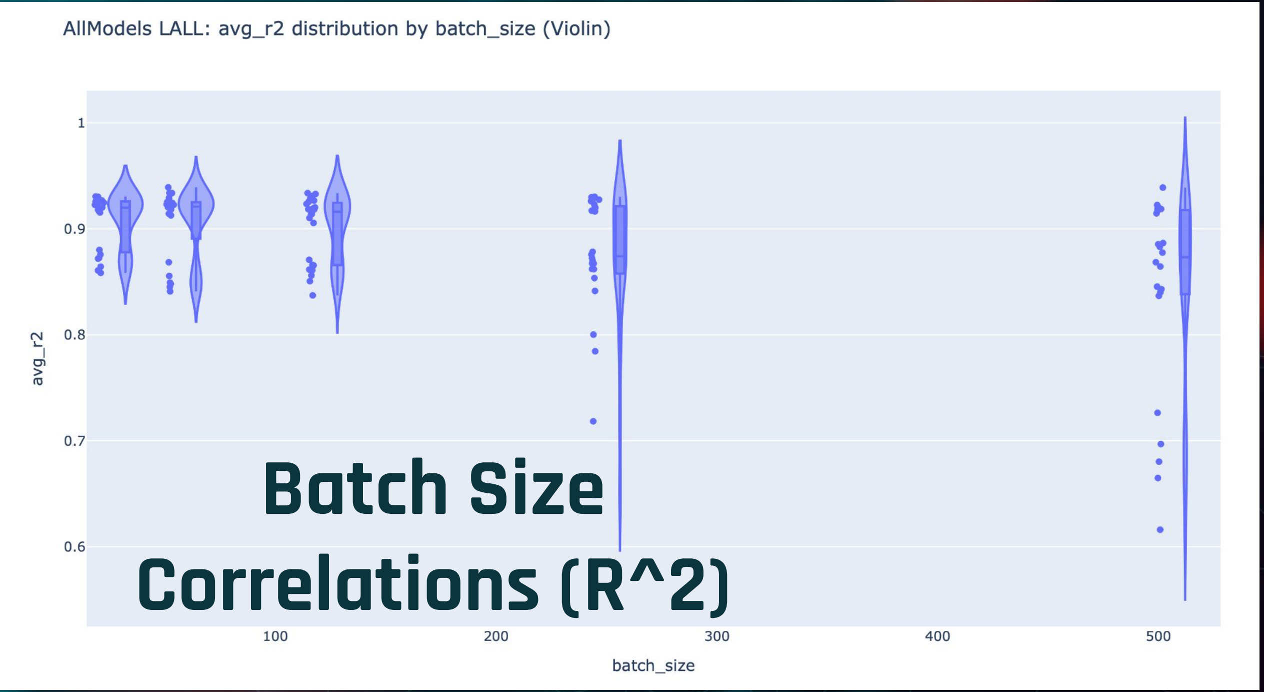




LSTM Large Sweep Results (cont.)

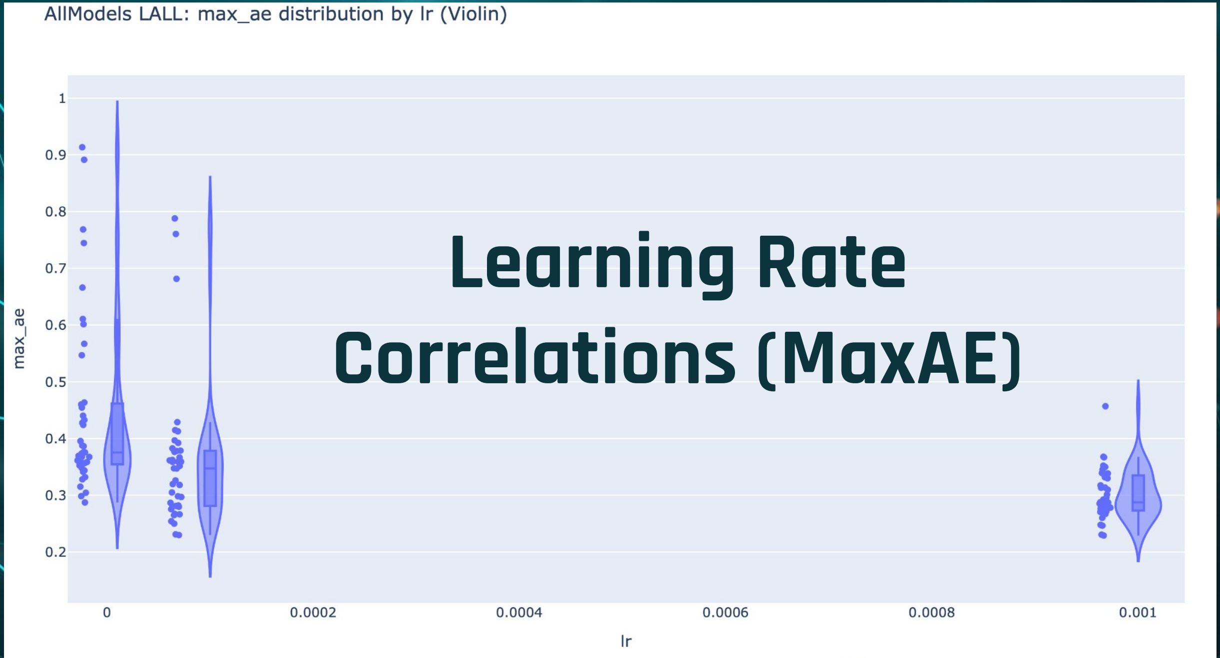


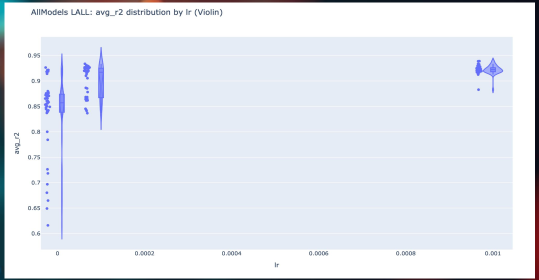


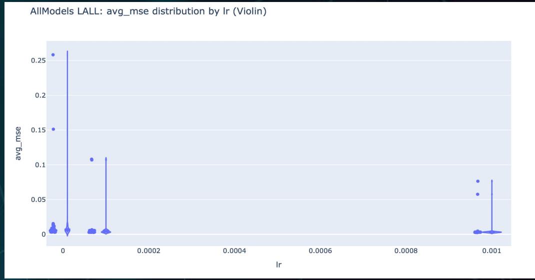


LSTM Large Sweep Results (cont.)

Learning Rate Correlations (MaxAE)



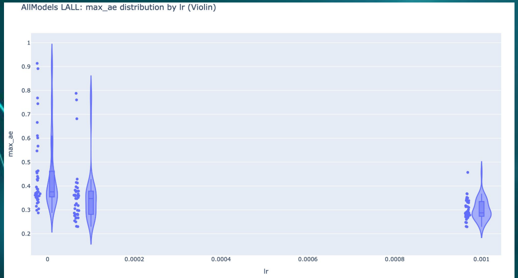


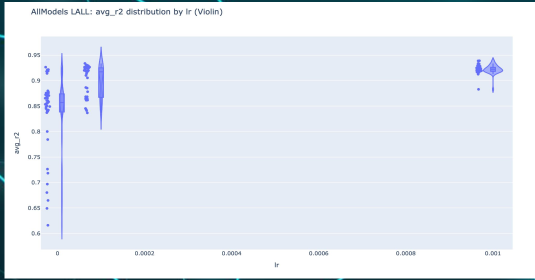


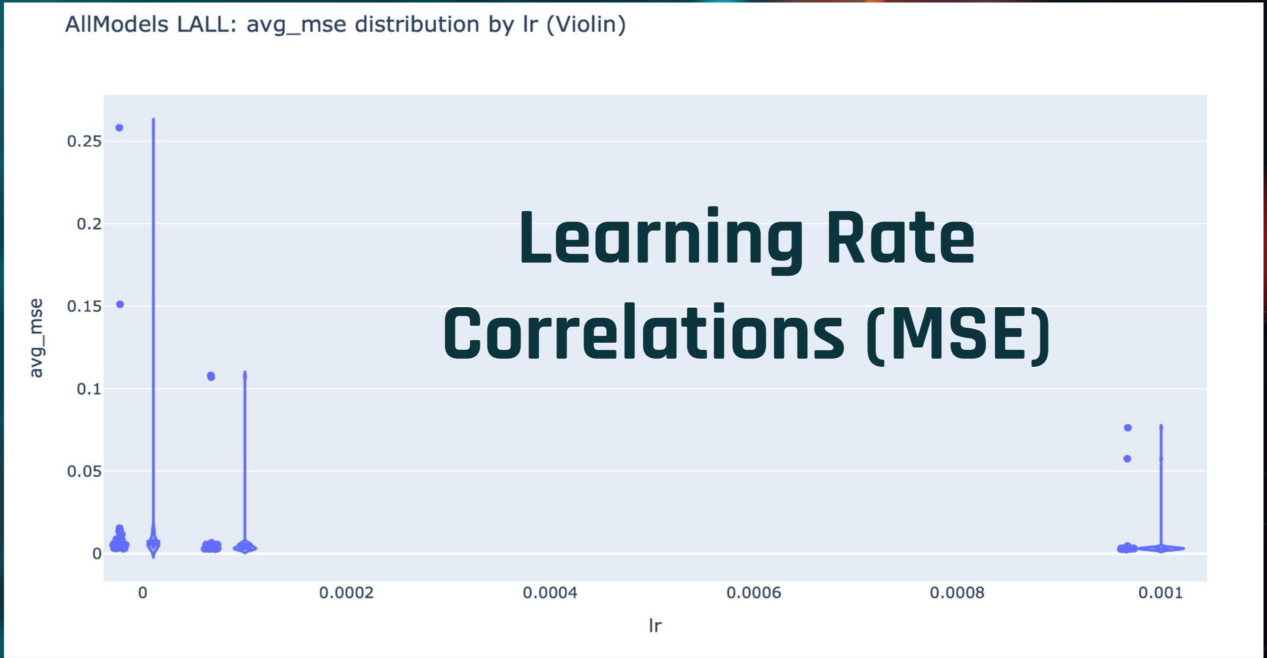
THE UNIVERSITY OF CHICAGO

globus labs

LSTM Large Sweep Results (cont.)





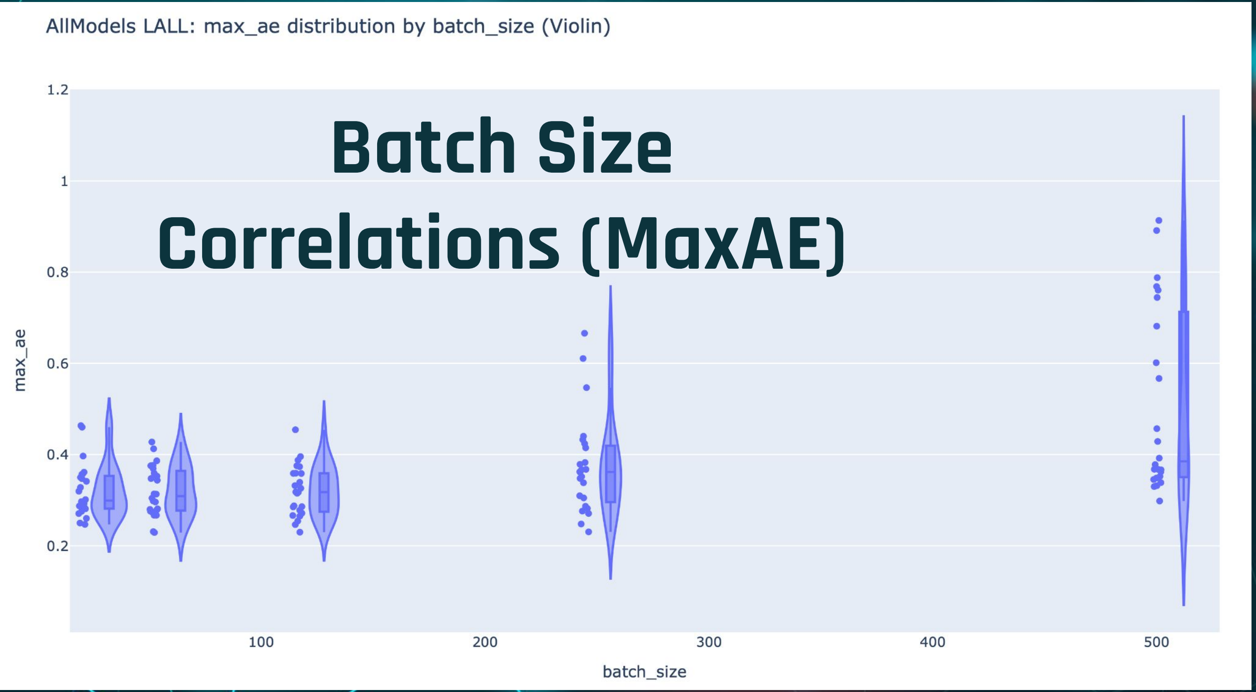


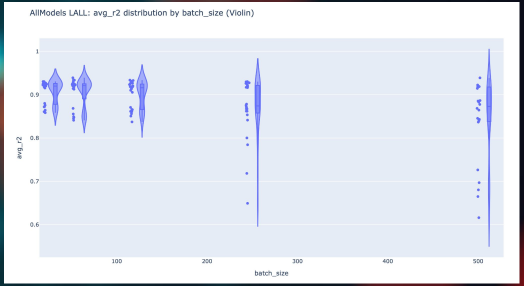
Learning Rate Correlations (MSE)

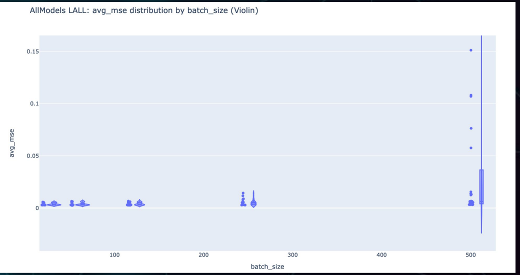
THE UNIVERSITY OF CHICAGO

globus labs

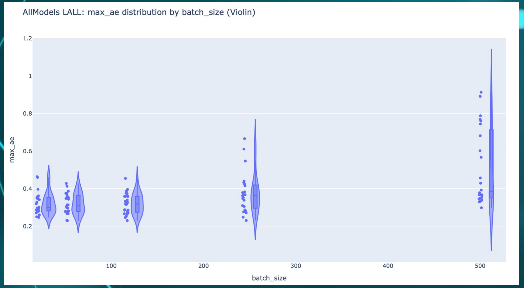
LSTM Large Sweep Results (cont.)

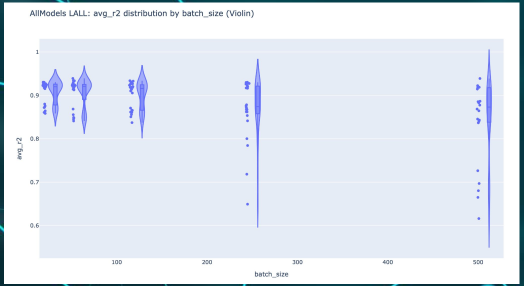






LSTM Large Sweep Results (cont.)

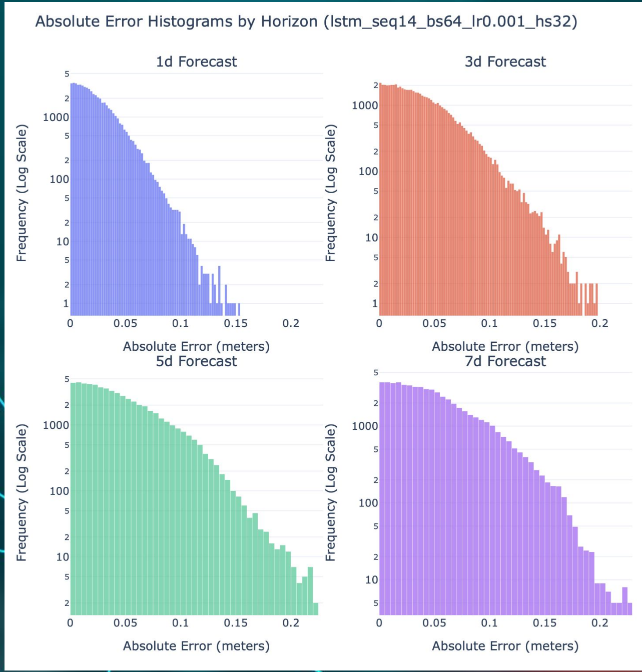


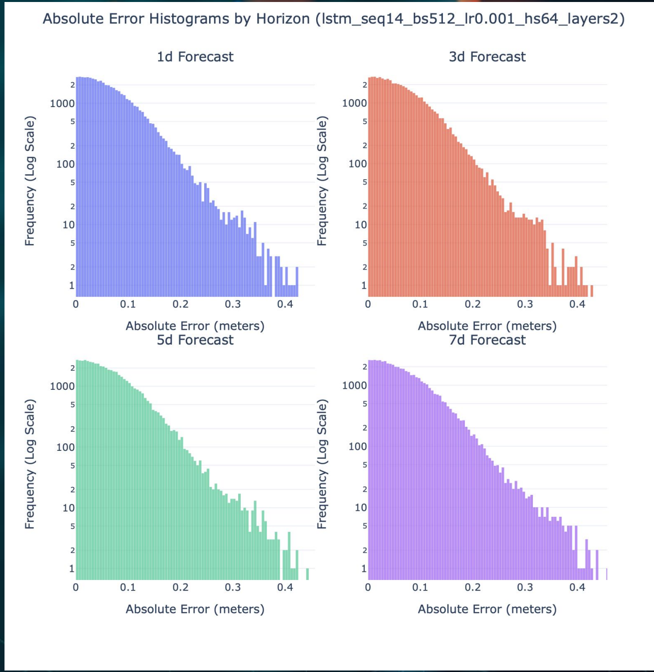


THE UNIVERSITY OF CHICAGO

globus labs

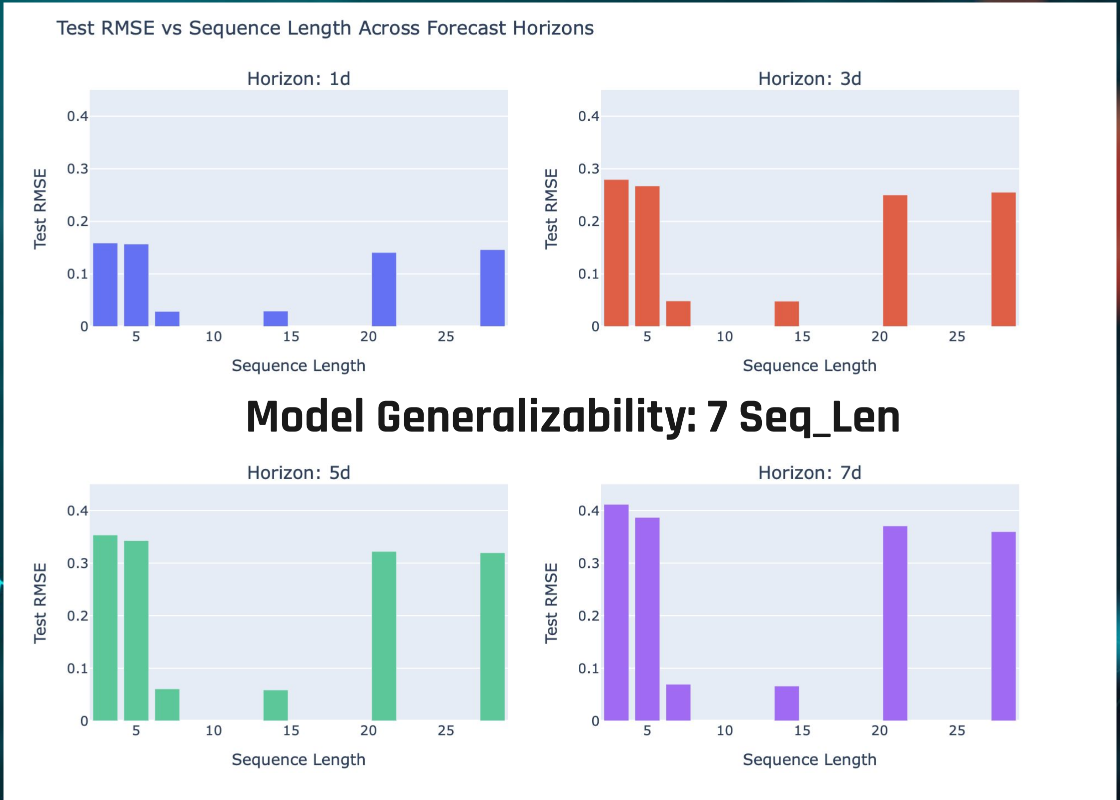
Best configurations (R^2)

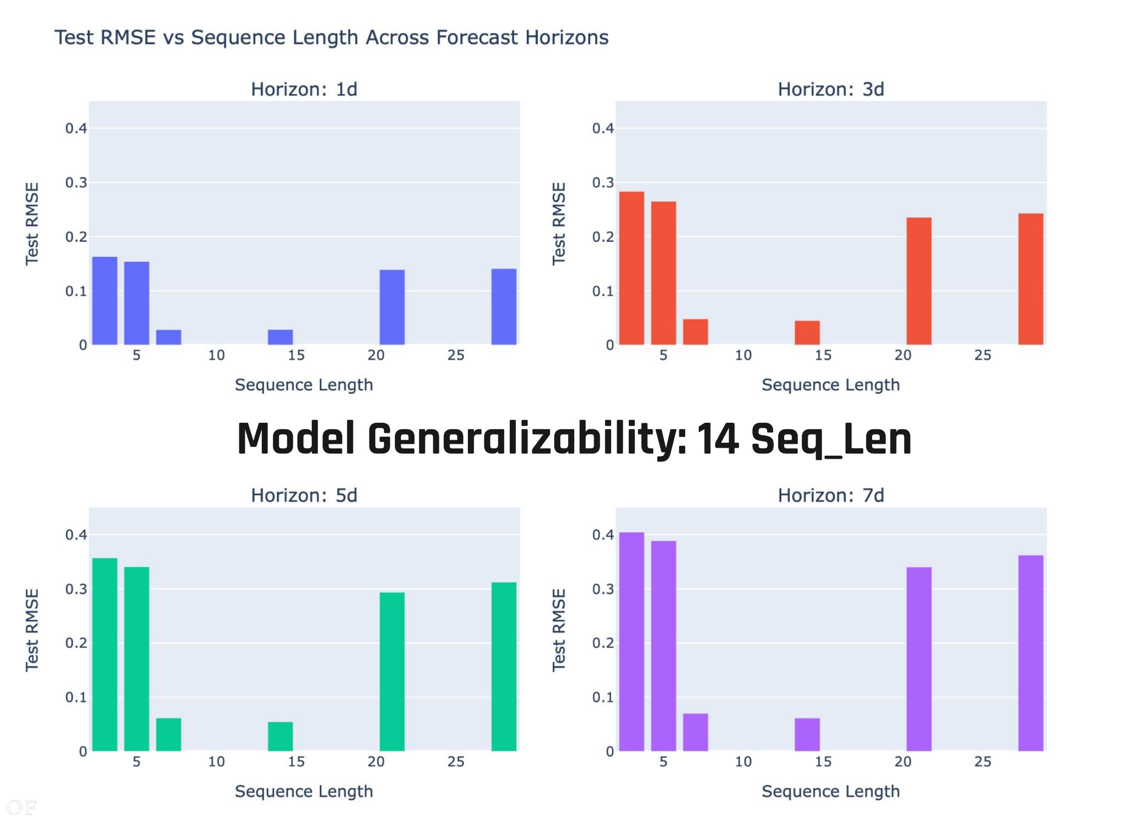




Model	R^2	MSE	MAE	MaxAE
LSTM_L1	0.939	0.002	0.038	0.229
LSTM_L2	0.939	0.058	0.204	0.368

Model Generalizability: 7 Seq_Len





Model Generalizability: 14 Seq_Len

THE UNIVERSITY OF CHICAGO

globus labs

Takeaway

Hyperparameter Search

LSTM Large Sweep –
120 configurations

- Seq_Len: [7, 14]
- BS: [32, 64, 128, 256, 512]
- LR: [1e-3, 1e-4, 1e-5]
- HS: [32, 64]
- Num_Layers: [1,2]

Many Models Sweep –
180 configurations

- Seq_Len: [7]
- BS: [32, 64, 128]
- LR: [1e-3, 5e-4, 1e-4]
- HS: [32, 64]
- Num_Layers: [1,2]
- Models: [LSTM, BILSTM, CONV-LSTM, GRU, ATTN]

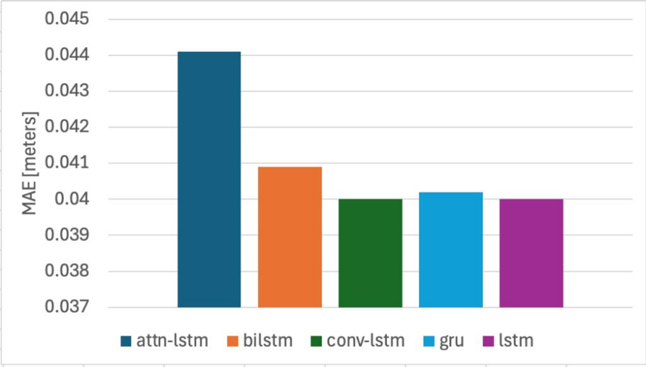
Multivariate Experiments –
Input Dim: Primary Station +

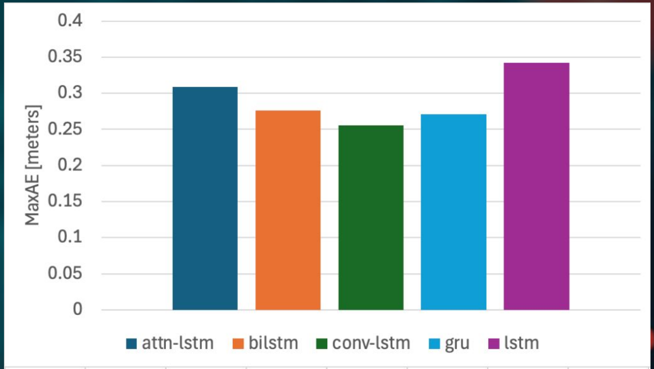
- Neighbor
- NOAA
- 37 Harmonic Constituents

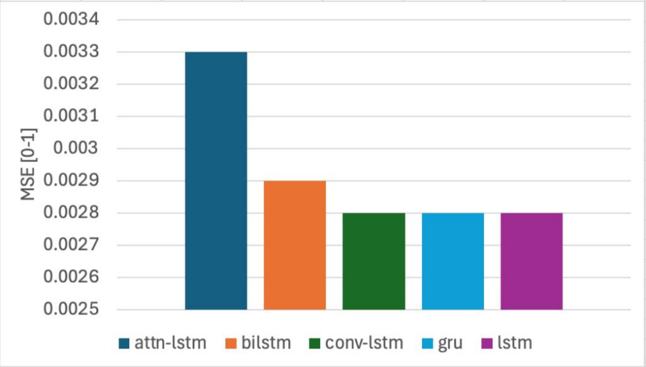
Variations on LSTM Models

Bidirectional LSTM	Convolutional LSTM
Gated Recurrent Unit	Attention-based LSTM

Many Models Sweep Results - Model vs Metrics







THE UNIVERSITY OF CHICAGO

globus labs

Many Models Sweep

Hidden Size
[32, 64]

Models
[LSTM, BILSTM, Conv-LSTM, GRU, ATTN]

Num Layers
[1, 2]

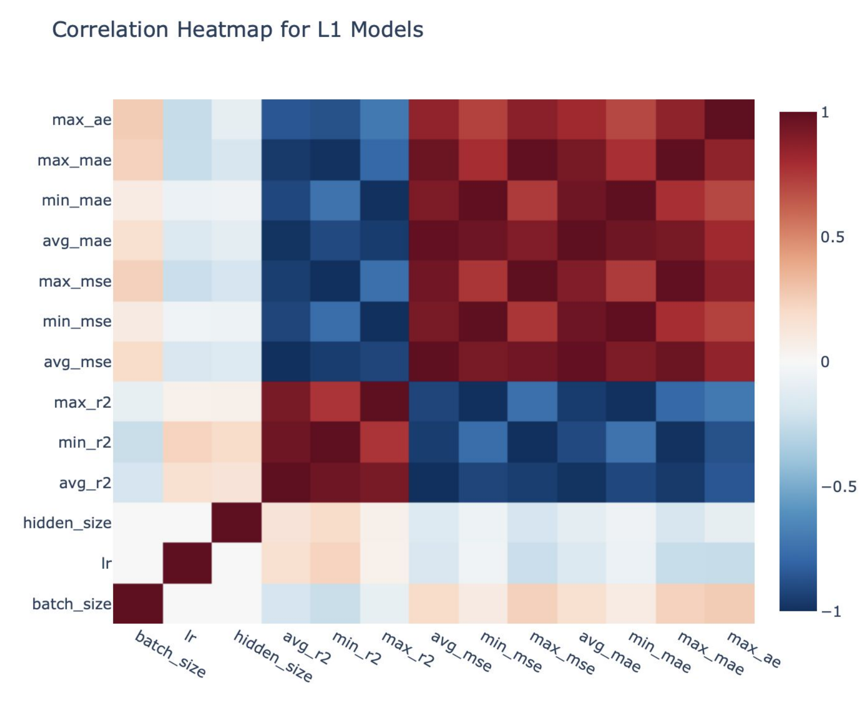
180

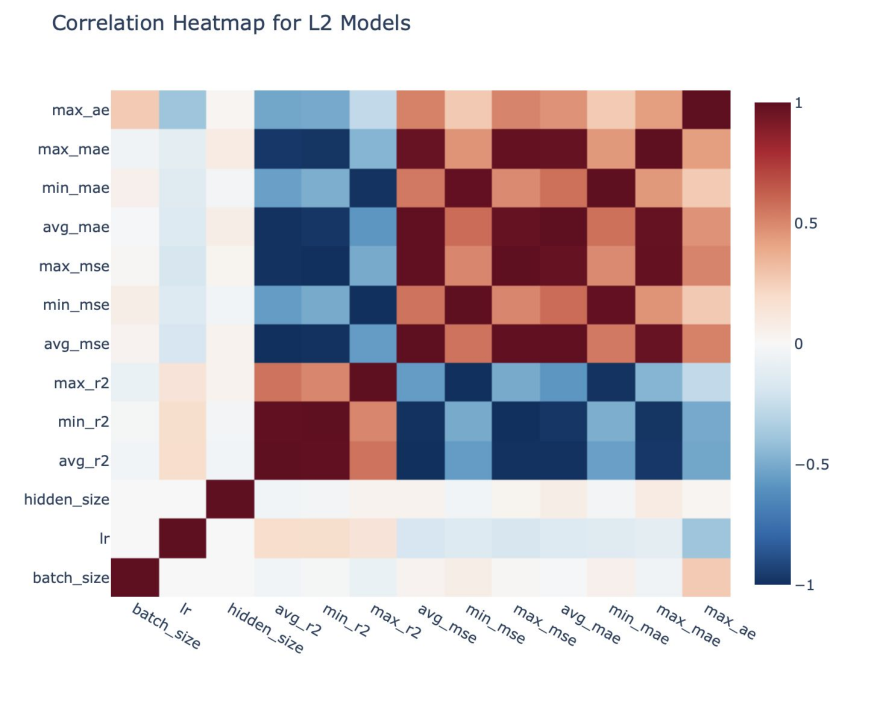
Learning Rate
[1e-3, 5e-4, 1e-4]

Sequence Length
[7]

Batch Size
[32, 64, 128]

Many Models Sweep Results





THE UNIVERSITY OF CHICAGO

globus labs

Many Models Sweep Results (cont.)

	BEST	Model	R^2	MSE	MAE	MaxAE	seq	batch	lr	hidden	layers
Horizon 1d	R^2	conv-lstm	0.980	0.001	0.022	0.133	7	64	0.0005	32	1
	MAE	conv-lstm	0.980	0.001	0.022	0.133	7	64	0.0005	32	1
	MaxAE	lstm	0.978	0.001	0.023	0.116	7	32	0.001	32	2
Horizon 3d	R^2	lstm	0.942	0.002	0.038	0.173	7	64	0.001	32	1
	MAE	conv-lstm	0.942	0.002	0.037	0.206	7	64	0.0005	32	1
	MaxAE	bilstm	0.940	0.002	0.039	0.169	7	32	0.0005	32	1
Horizon 5d	R^2	gru	0.910	0.004	0.047	0.251	7	64	0.0001	64	2
	MAE	lstm	0.910	0.004	0.047	0.227	7	128	0.0005	32	1
	MaxAE	gru	0.903	0.004	0.049	0.219	7	32	0.001	32	1
Horizon 7d	R^2	conv-lstm	0.891	0.004	0.053	0.251	7	64	0.001	32	1
	MAE	conv-lstm	0.891	0.004	0.053	0.251	7	64	0.001	32	1
	MaxAE	gru	0.880	0.005	0.056	0.222	7	32	0.001	32	1

Many Models Sweep Results (cont.)

Model	Layers	seq	batch	lr	hidden	layers	R^2	MSE	MAE	MaxAE
gru	2	7	64	0.0001	64	2	0.930	0.003	0.040	0.271
conv-lstm	1	7	64	0.001	32	1	0.930	0.003	0.040	0.256
lstm	1	7	128	0.0005	32	1	0.929	0.003	0.040	0.342
gru	1	7	32	0.0005	64	1	0.928	0.003	0.041	0.286
bilstm	1	7	32	0.001	32	1	0.926	0.003	0.041	0.276
conv-lstm	2	7	128	0.001	32	2	0.926	0.003	0.041	0.259
lstm	2	7	64	0.0005	32	2	0.924	0.003	0.042	0.273
bilstm	2	7	128	0.001	32	2	0.923	0.003	0.042	0.265
attn-lstm	1	7	32	0.001	64	1	0.916	0.003	0.044	0.309

Hyperparameter Search

LSTM Large Sweep –
120 configurations

- Seq_Len: [7, 14]
- BS: [32, 64, 128, 256, 512]
- LR: [1e-3, 1e-4, 1e-5]
- HS: [32, 64]
- Num_Layers: [1,2]

Many Models Sweep –
180 configurations

- Seq_Len: [7]
- BS: [32, 64, 128]
- LR: [1e-3, 5e-4, 1e-4]
- HS: [32, 64]
- Num_Layers: [1,2]
- Models: [LSTM, BILSTM, CONV-LSTM, GRU, ATTN]

Multivariate Experiments –
Input Dim: Primary Station +

- Nearest Neighbor
- NOAA Predictions
- 37 Harmonic Constituents

Hyperparameter Search

LSTM Large Sweep –
120 configurations

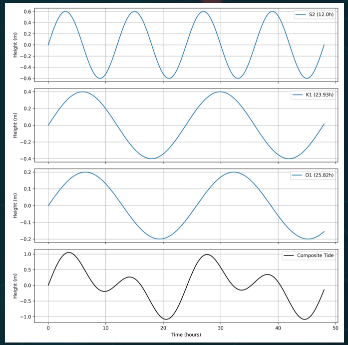
- Seq_Len: [7, 14]
- BS: [32, 64, 128, 256, 512]
- LR: [1e-3, 1e-4, 1e-5]
- HS: [32, 64]
- Num_Layers: [1,2]

Many Models Sweep –
180 configurations

- Seq_Len: [7]
- BS: [32, 64, 128]
- LR: [1e-3, 5e-4, 1e-4]
- HS: [32, 64]
- Num_Layers: [1,2]
- Models: [LSTM, BILSTM, CONV-LSTM, GRU, ATTN]

Multivariate Experiments –
Input Dim: Primary Station +

- Nearest Neighbor
- NOAA Predictions
- 37 Harmonic Constituents

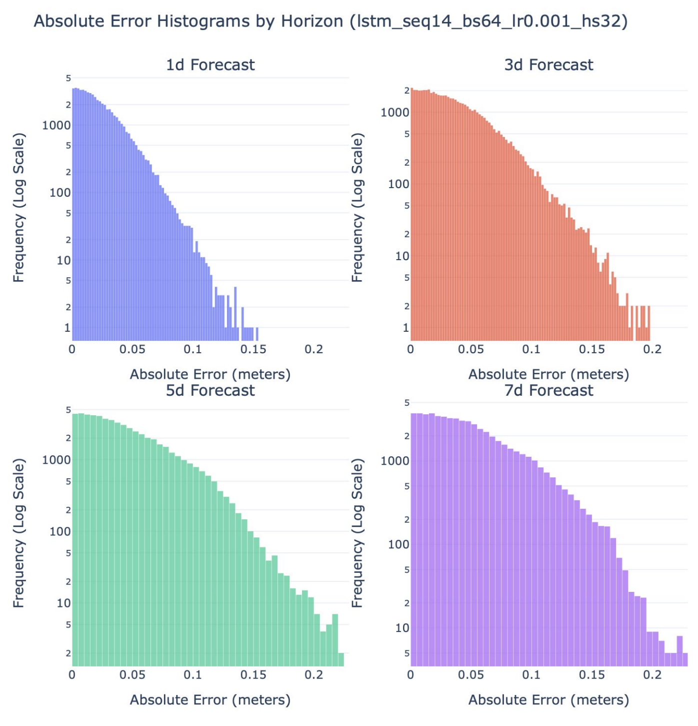


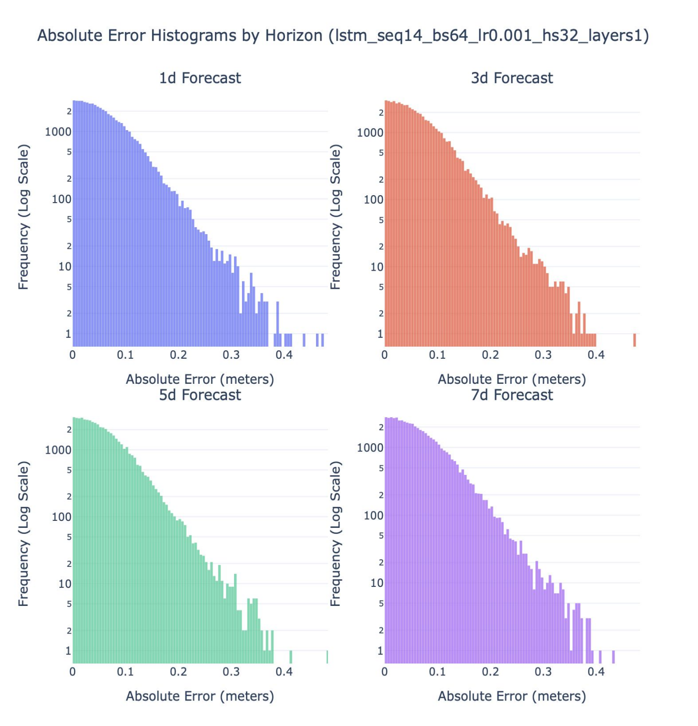
Other Approaches

- Current Univariate model
    - Input dimension: main water-level
    - Output dimension: main water-level
- Multivariate models
    - Possible Input Dims
        - Nearest neighbor + main water-level
        - NOAA predictions + main water-level
        - 37 Harmonic constituents + main water-level
    - Output Dim: main water-level

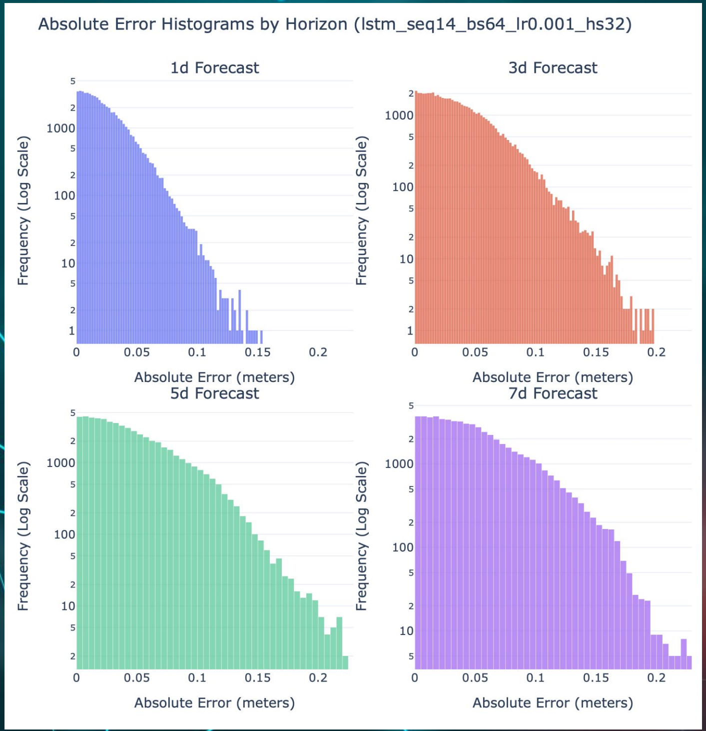
Preliminary (1 Layer) Results

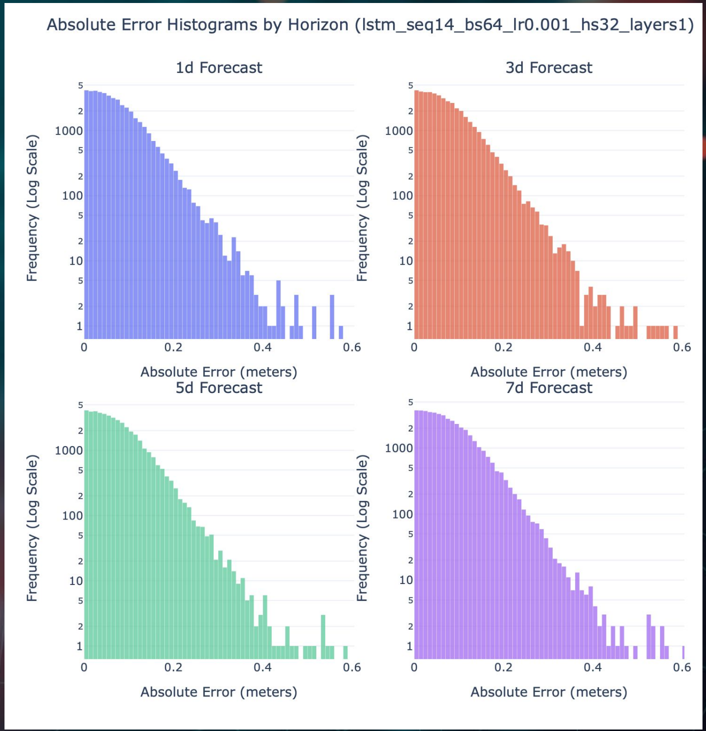
Univariate vs Nearest Neighbor



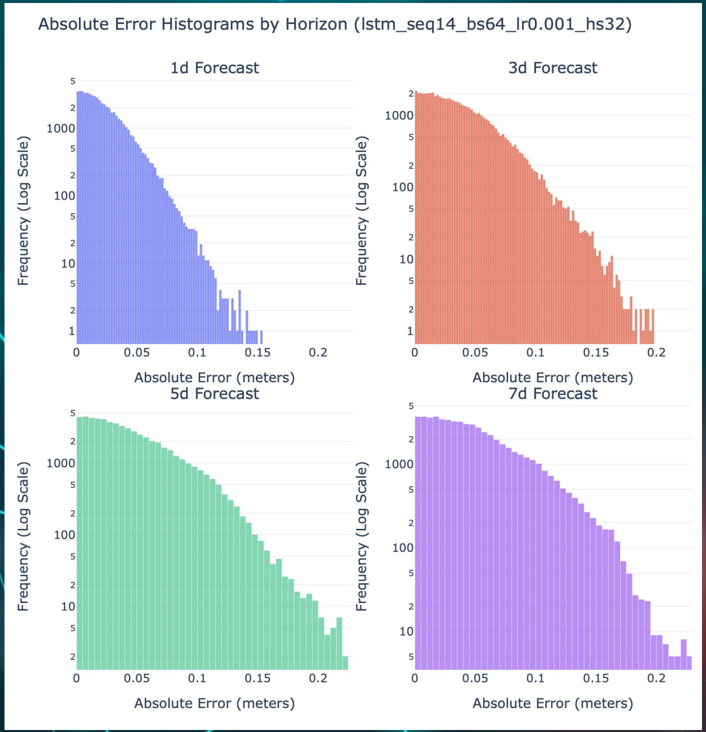


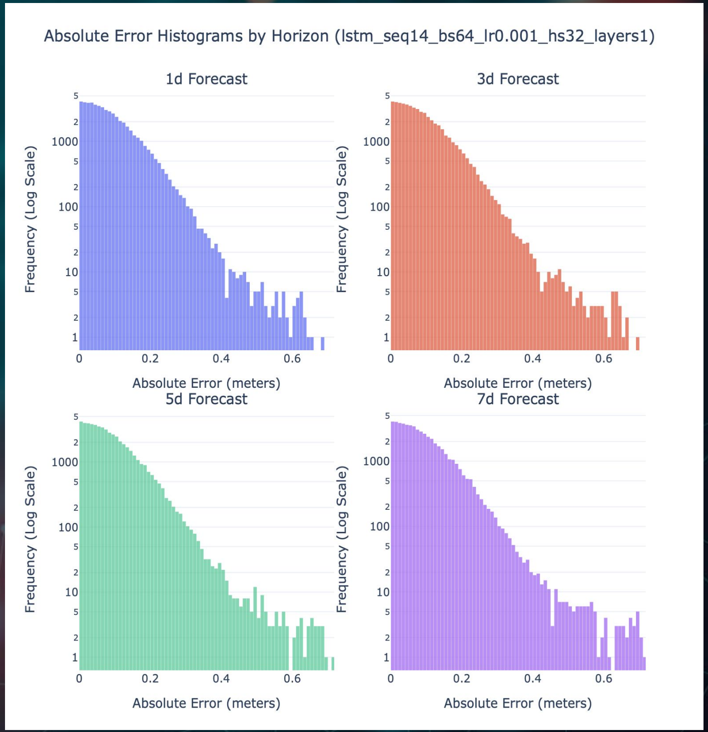
Univariate vs NOAA predictions





Univariate vs Constituents



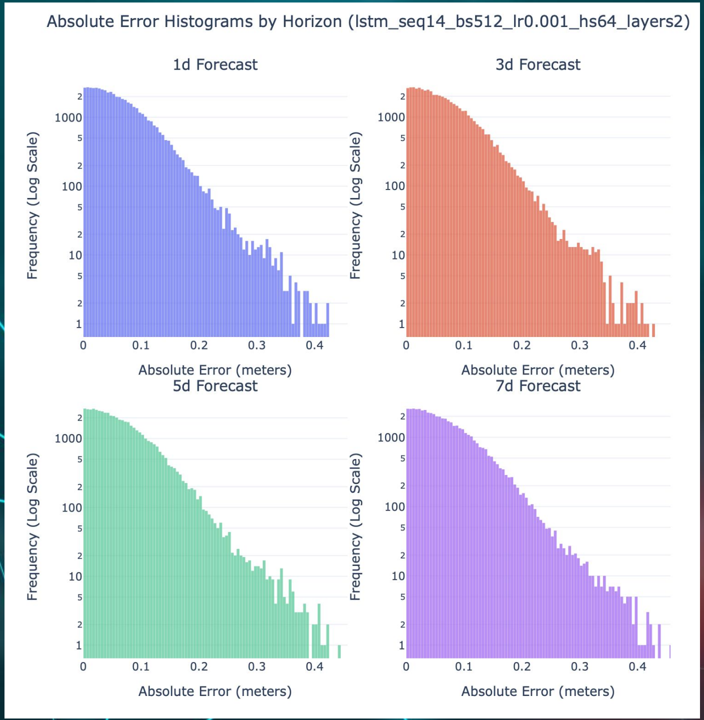


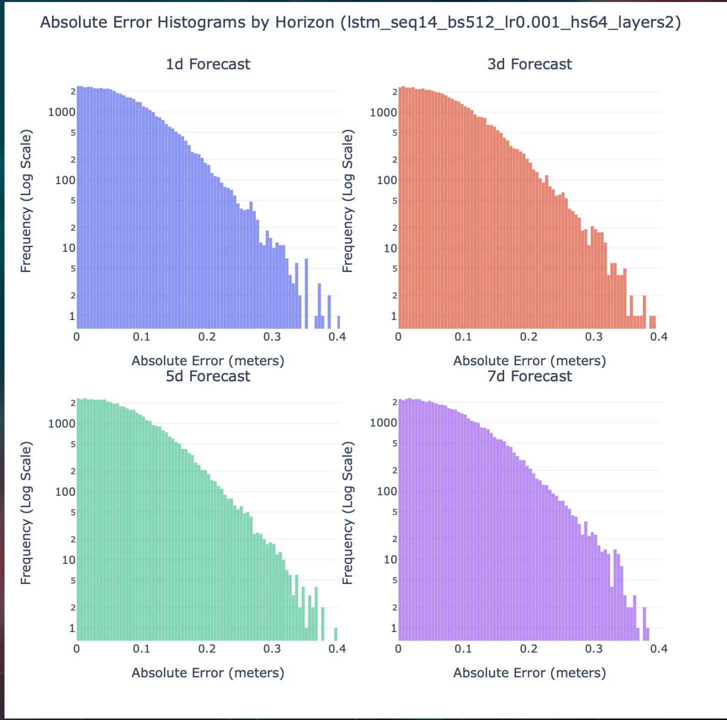
THE UNIVERSITY OF CHICAGO

globus labs

Preliminary (2 Layers) Results

Univariate vs Nearest Neighbor

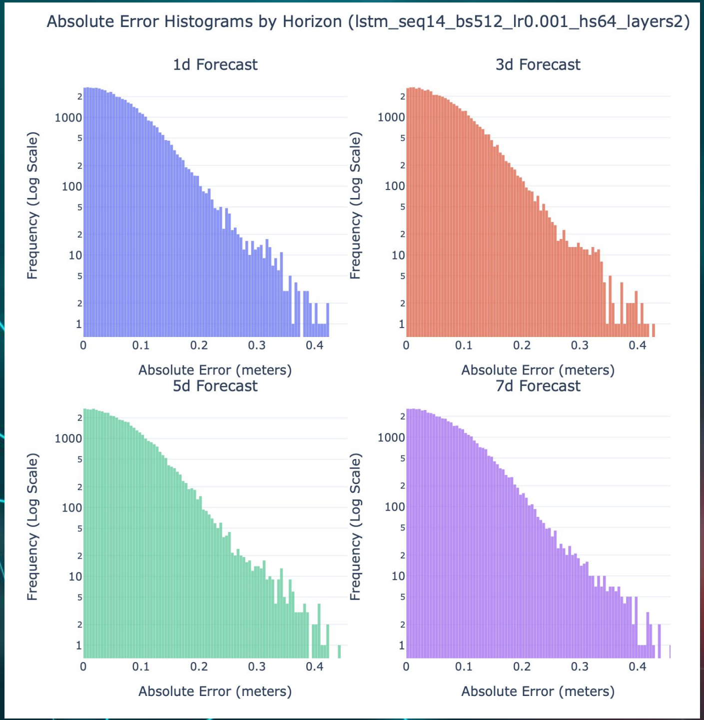


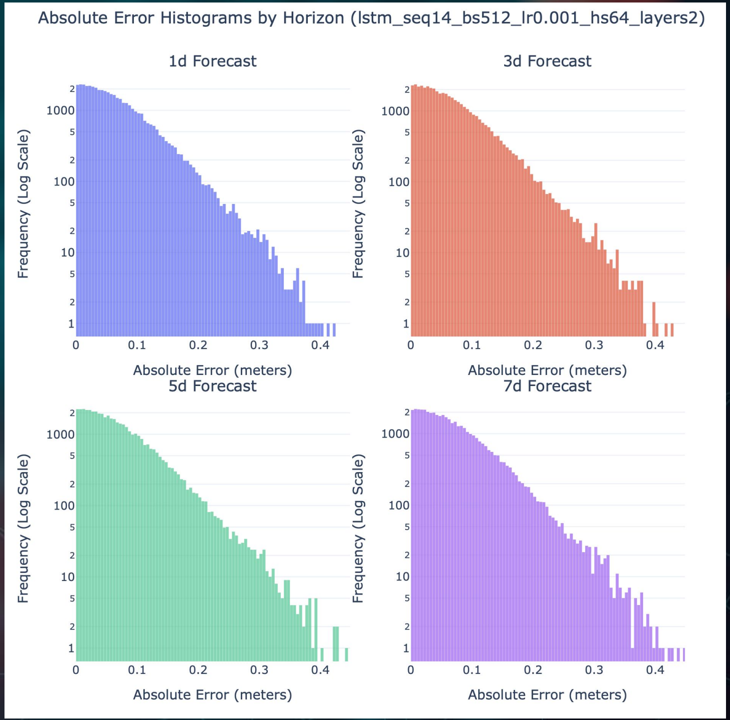


THE UNIVERSITY OF CHICAGO

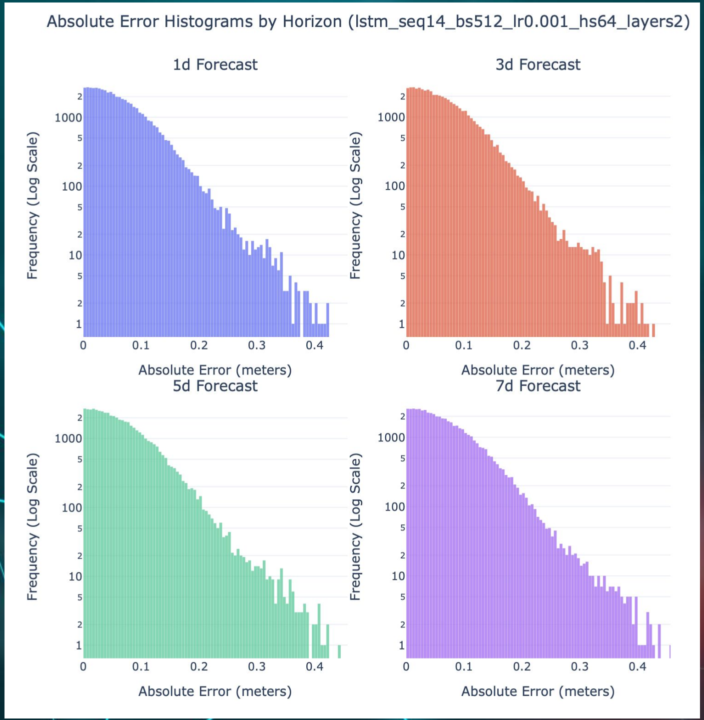
globus labs

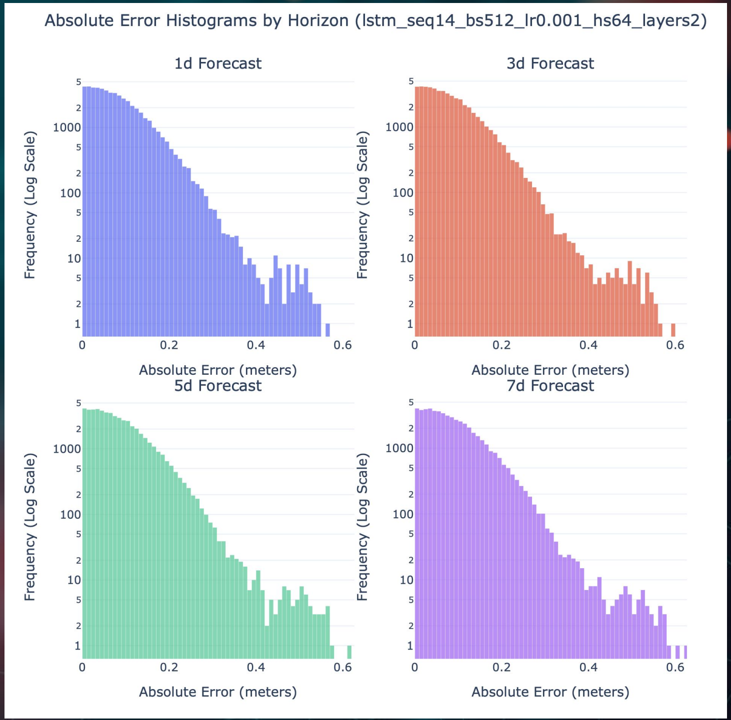
Univariate vs NOAA predictions





Univariate vs Constituents





Takeaway

Other Approaches Summary

- Multivariate models
  - Decreased performance
  - Takes longer to compute
- Might need to do hyperparameter sweep for multivariate models

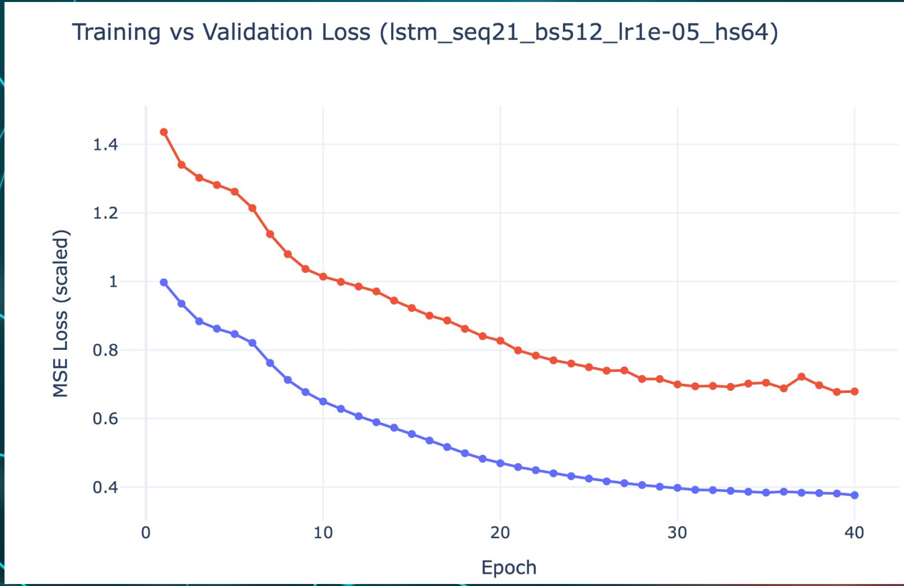
THE UNIVERSITY OF CHICAGO

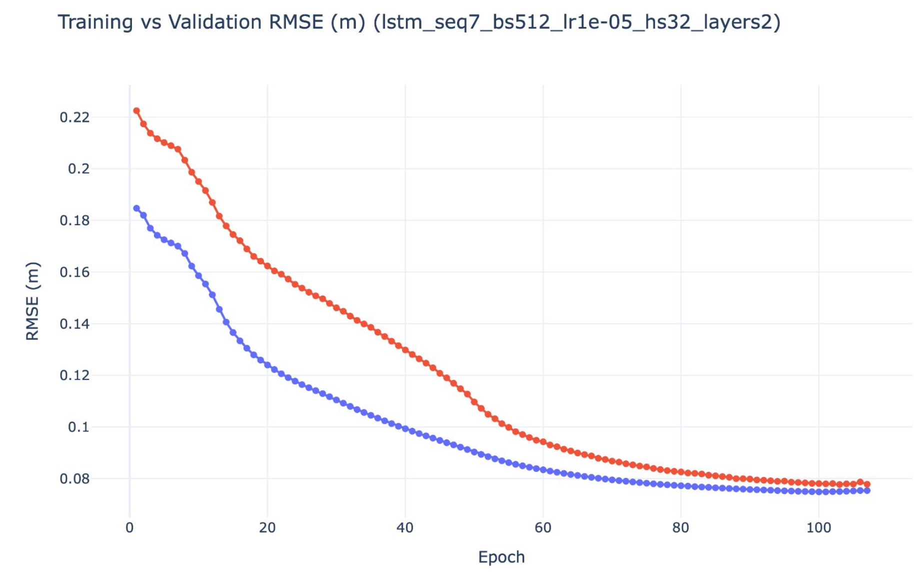
globus labs

Big Takeaways Summary

- Average r^2 ↗ when …
- Learning rate is **VERY IMPORTANT** ⇒ 0.001
- Epoch limit must be set high ⇒ 200
- Must have patience ⇒ 5
- Num layers ⇒ 1 or 2
- Correlation between input sequence length and r^2
- Different variations of LSTM models perform similarly
- Adding more data (neighbors, NOAA) does not improve performance and has more computational overheads

Epoch Success!





THE UNIVERSITY OF CHICAGO

globus labs

Challenges

- Too many open file handles & out of memory errors
    - Large hyper-parameter combinations
    - Persistent workers
- Long computation time
- Fingerprint SSH Authentication across instances

Solutions

- ulimit, /etc/security/limits.conf
- With context in Python for file open/close
- Sharing_strategy("file_system") across processors to prevent overuse of file handles
- Expandable Segments to allocate more memory for computing
- Multiple Instances
- Creation of new floating IP

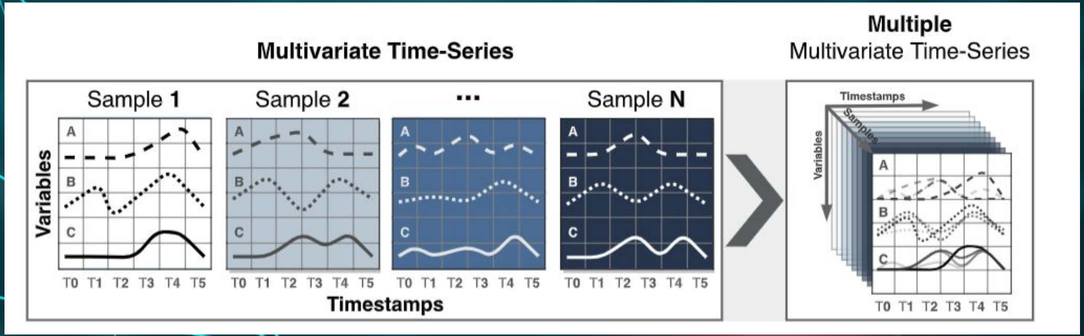
What I learned over the summer?

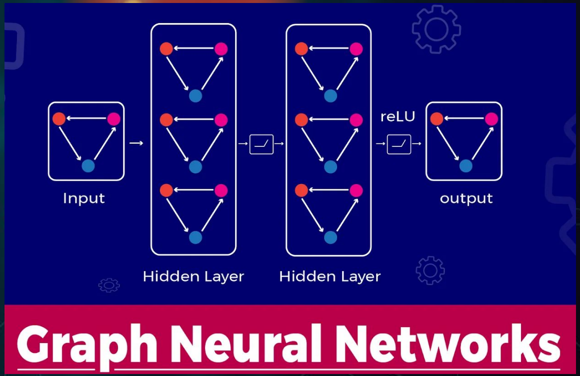
- Software engineering
  - Git, Python, PyTorch, BASH
- Deep Learning Models
- Limitations of personal laptops for AI-focused workloads
- Concurrency
  - Chameleon, ssh , rsync

THE UNIVERSITY OF CHICAGO

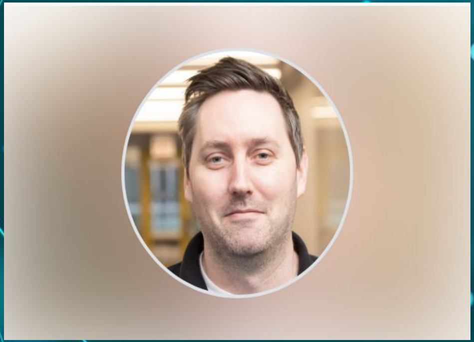
globus labs

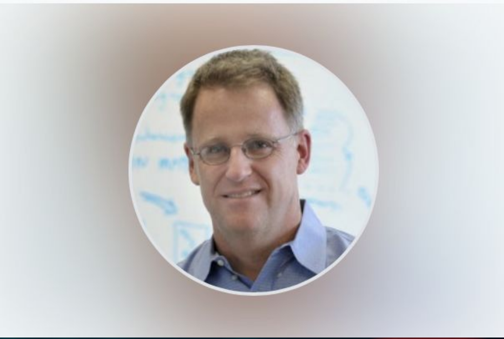
Future Work

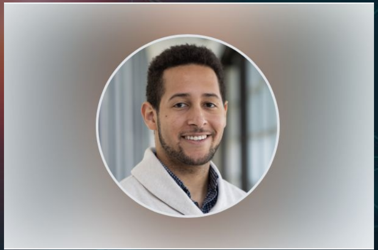

Multivariate Time-Series
Multiple
Multivariate Time-Series
Sample 1
Sample 2
...
Sample N
Variables
A
B
C
T0 T1 T2 T3 T4 T5
Timestamps
Samples

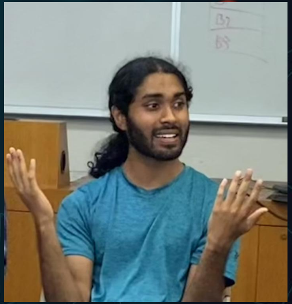

Input
Hidden Layer
Hidden Layer
reLU
output
Graph Neural Networks

Thank you!









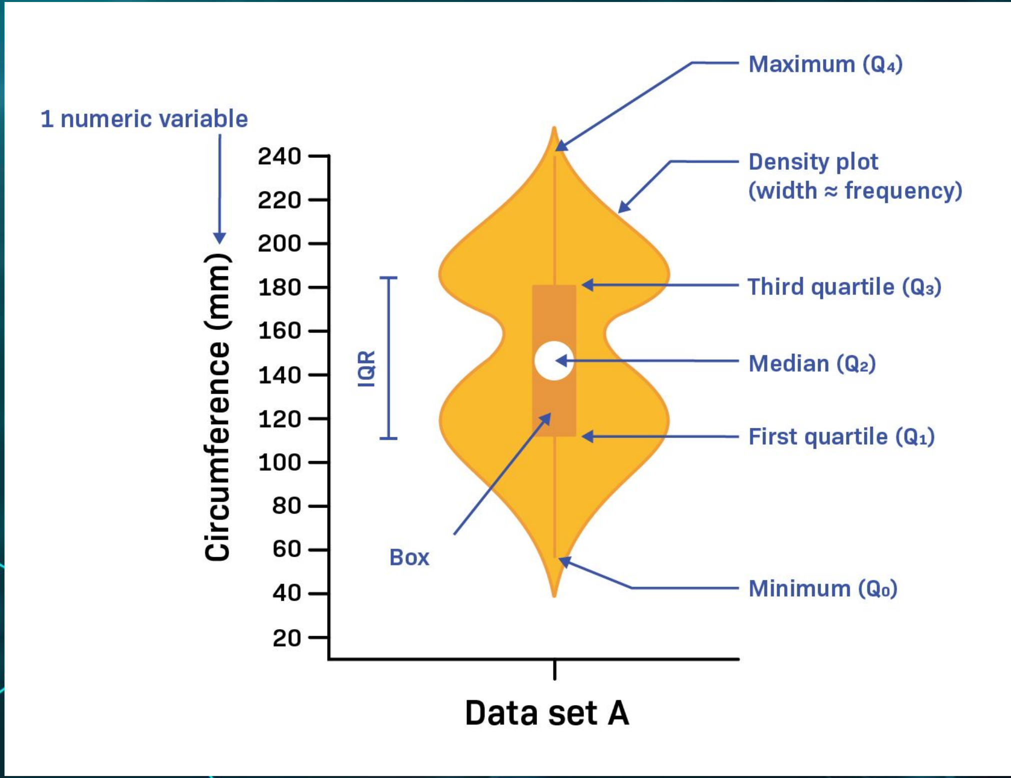
globus labs

Questions

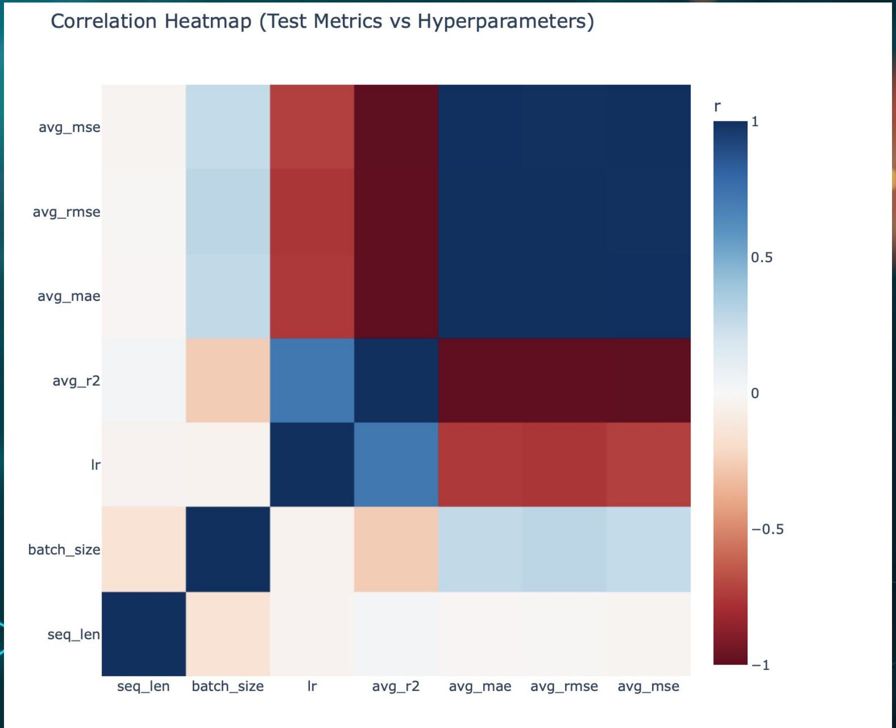
Testbed

- Preliminary Testbed
  - Hardware: MacOS, 6-core Intel CPU 2.6GHz, 16GB RAM
  - Time to train LSTM with 500K 6-min samples for two epochs: 2hrs
- Chameleon (three instances)
  - Intel 24-core CPU 2.6GHz, 192GB RAM, 400GB SSD
  - Nvidia 4608-core RTX 6000 GPU, 24GB RAM
  - Time to train LSTM with 500K 6-min samples for 200 epochs: 2hrs **(100X speedup)**

Violin Plots



False Start Results (cont.)



THE UNIVERSITY OF CHICAGO

globus labs

False Start

Hidden Size
[64]

Models
[LSTM]

Num Layers
[1]

27

Learning Rate
[1e-4, 1e-5, 1e-6]

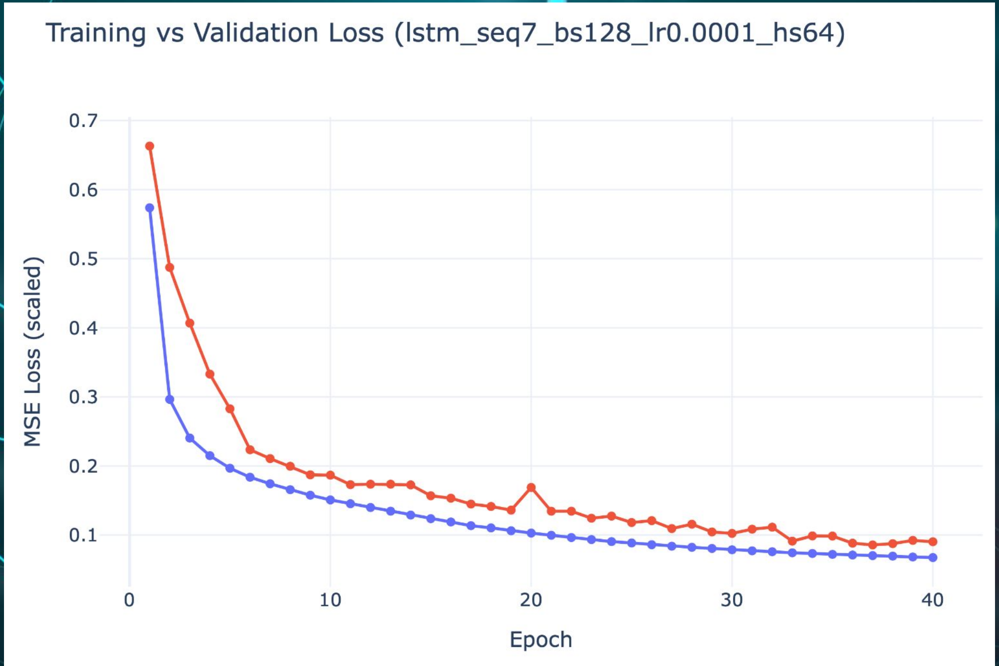
Sequence Length
[7, 14, 21]

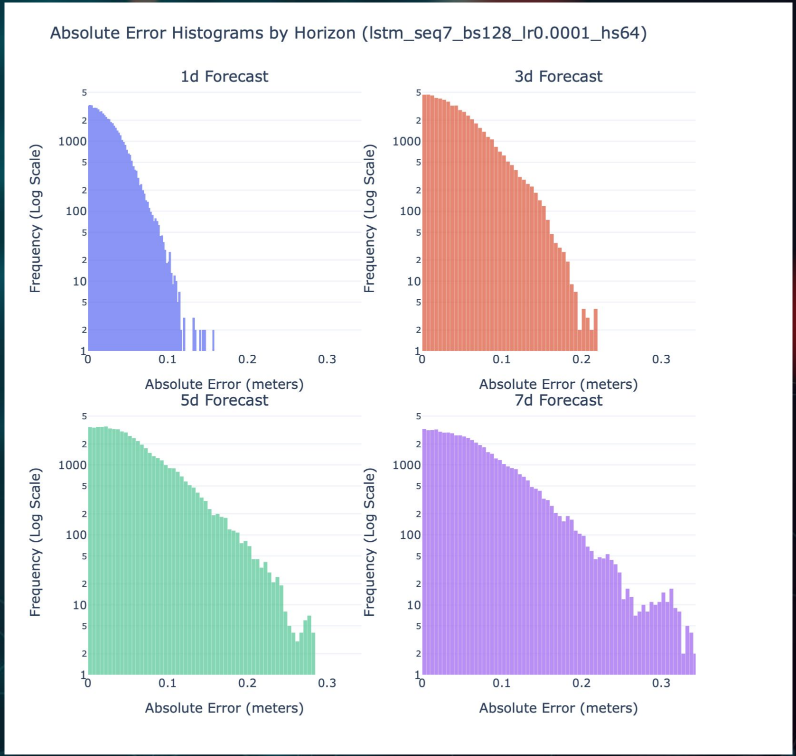
Batch Size
[128, 256, 512]

False Start Results

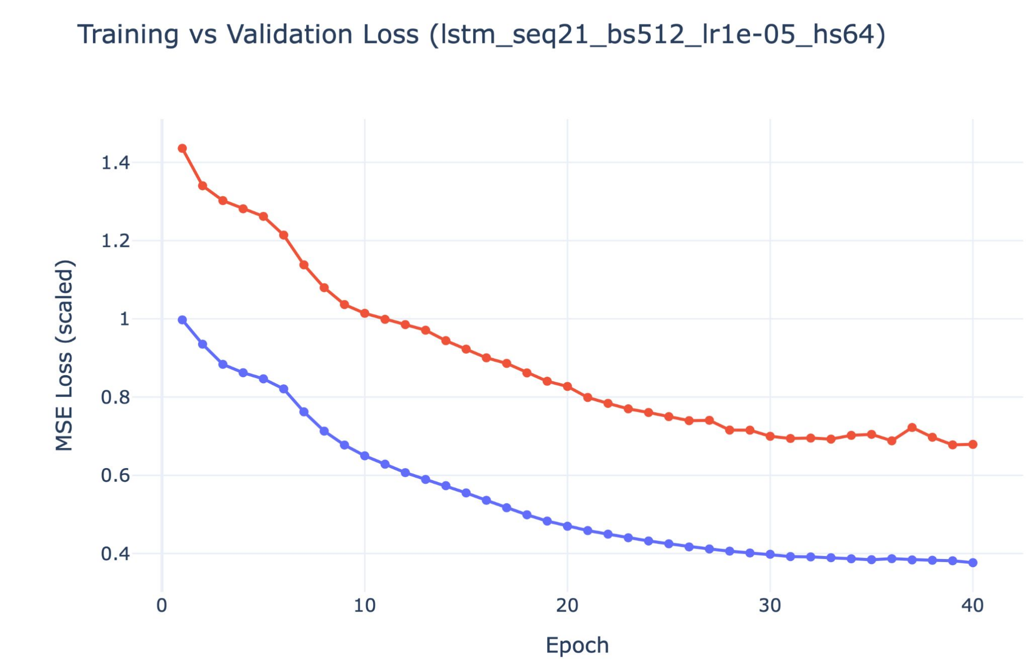
model	seq	batch	lr	hidden	r2	mae	rmse	mse
lstm	7	128	0.0001	64	0.913	0.044	0.056	0.003
lstm	14	256	0.0001	64	0.879	0.051	0.066	0.005
lstm	7	256	0.0001	64	0.875	0.052	0.067	0.005
lstm	21	256	0.0001	64	0.866	0.054	0.069	0.005
lstm	14	128	0.0001	64	0.862	0.055	0.070	0.005
lstm	21	128	0.0001	64	0.859	0.055	0.071	0.006
...	...	...	...	...	...	...	...	...
lstm	14	128	1.00E-06	64	0.505	0.105	0.130	0.020
lstm	7	256	1.00E-06	64	0.456	0.112	0.138	0.022
lstm	14	256	1.00E-06	64	0.426	0.114	0.141	0.023
lstm	21	256	1.00E-06	64	0.422	0.114	0.142	0.023
lstm	7	512	1.00E-06	64	0.336	0.125	0.157	0.026
lstm	14	512	1.00E-06	64	0.299	0.131	0.164	0.028

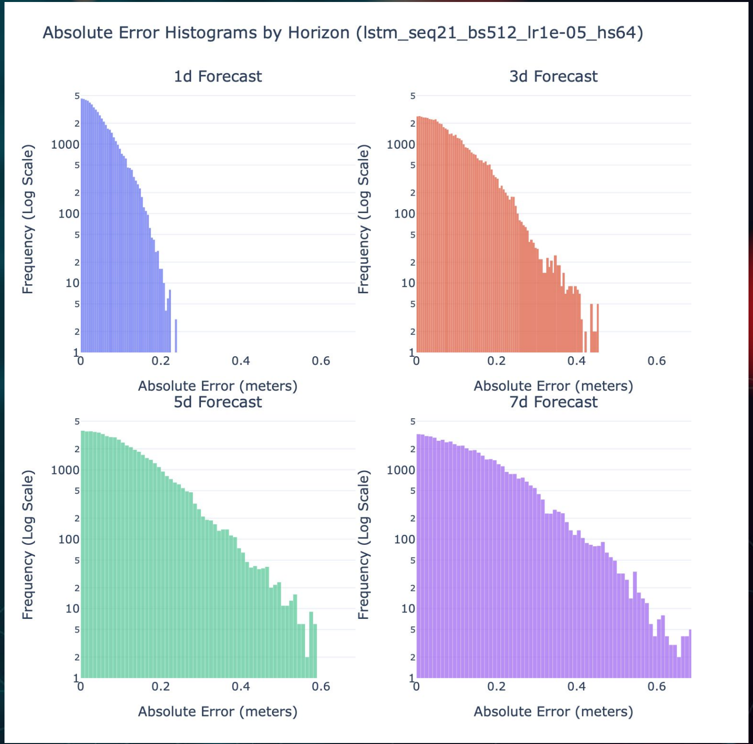
False Start
Results (cont.)





False Start Results (cont.)





False Start: What did we learn?

- Average r^2 increases when …
- Epoch limit
- Patience – early stopping period
- Num layers
- Correlation between input sequence length and r^2
- Different models?
- Multiple input dimensions?