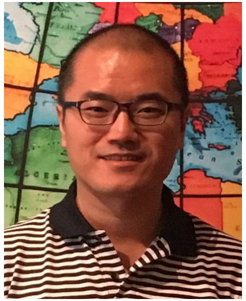
Widening the Time Horizon: Predicting the Long-Term Behavior of Chaotic Systems

Yong Zhuang | Matthew Almeida | Wei Ding | Patrick D Flynn | Shafiqul Islam | Ping Chen

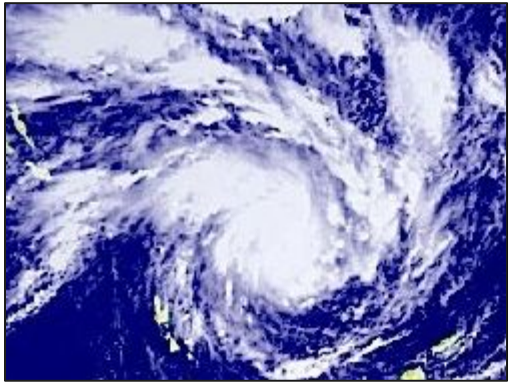
UMass Boston

Tufts University

Grants:
IIS-2008202
IIS-2008276

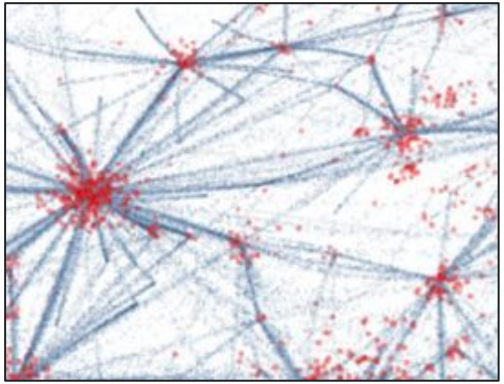
Motivation

Chaotic systems are found across many fields of study:



Climatology

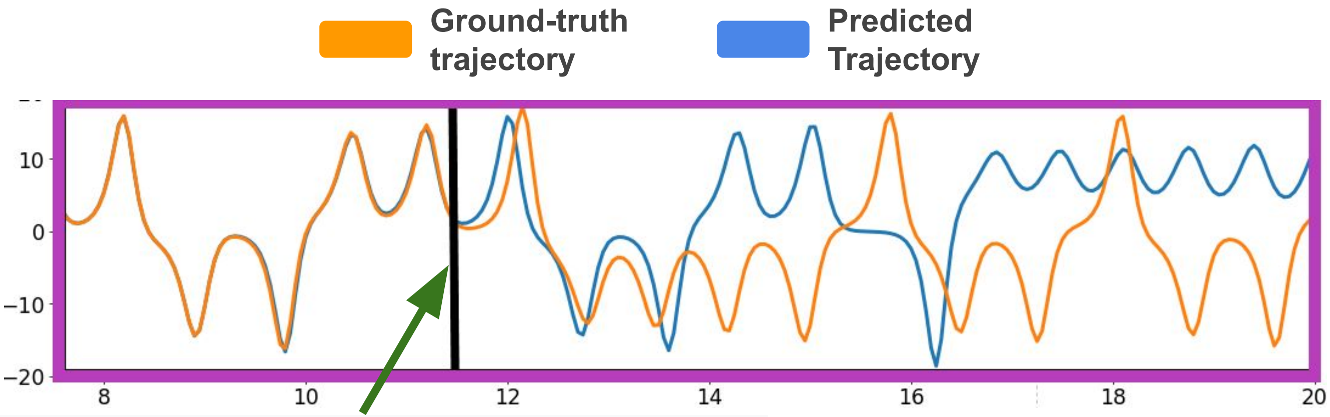
Biology



Virology

Accurate long-term forecasts can be very helpful in understanding such systems, warning of impending disasters, and making long-term plans.

Prediction Horizon

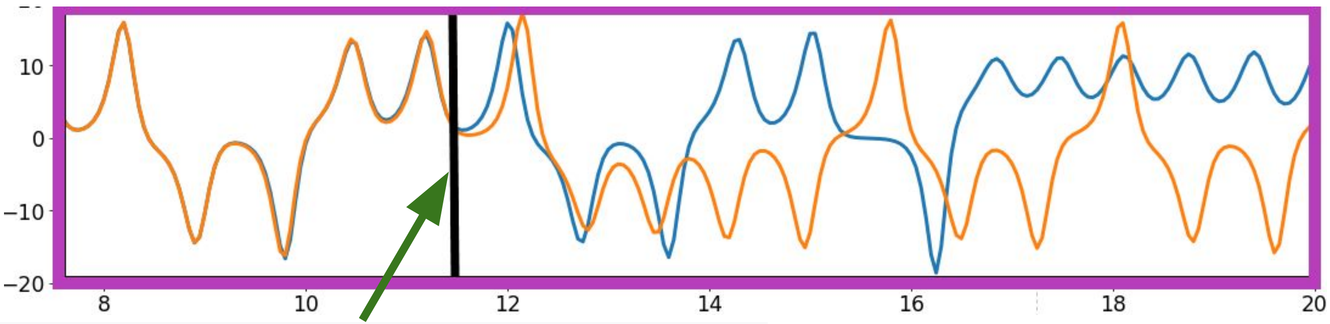


Prediction Horizon: How far ahead a model predicts the future.

Widen the Prediction Horizon

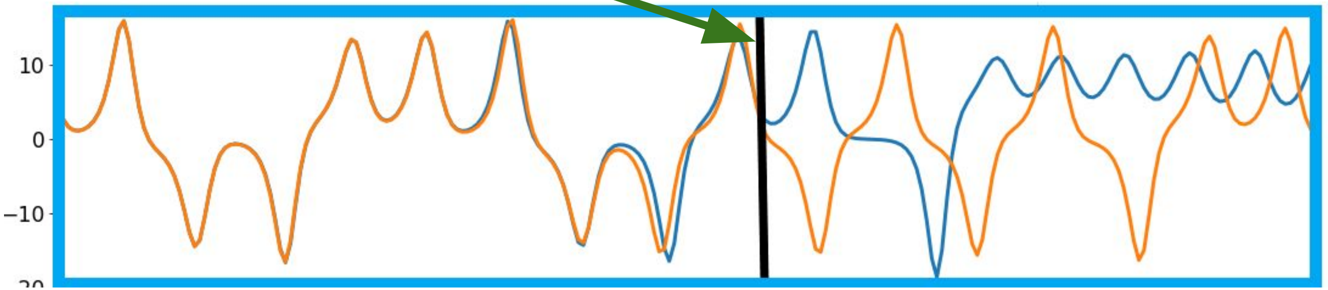
Ground-truth trajectory

Predicted Trajectory



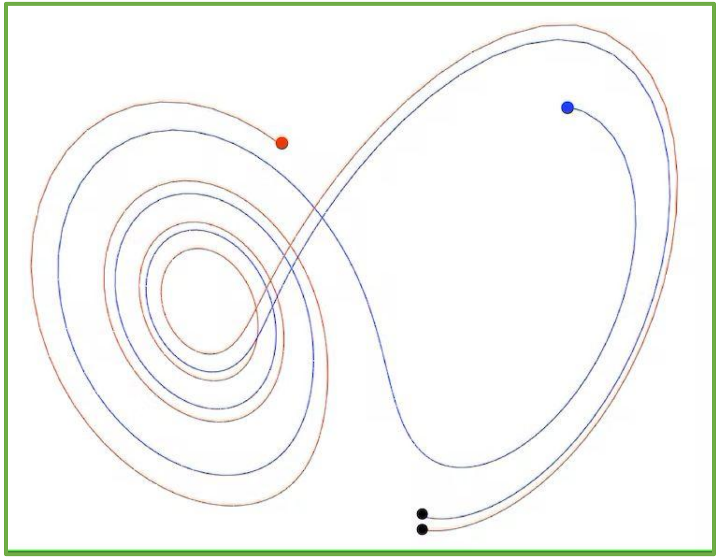
Prediction Horizon: How far ahead a model predicts the future.

Can the prediction horizon of this model be widened?



Challenge

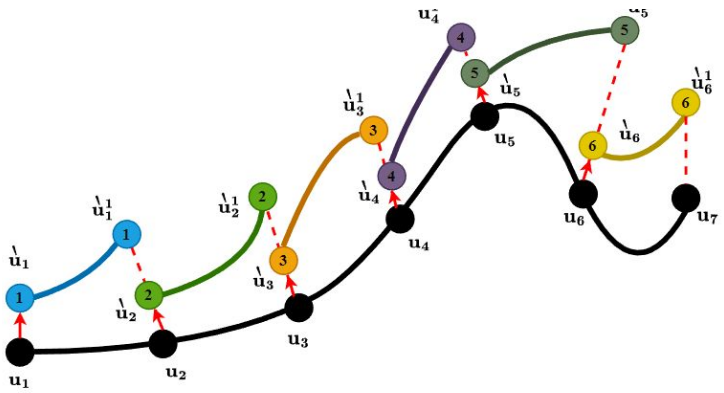
Sensitive dependence on initial conditions



Two segments of the three-dimensional evolution of two trajectories for the same period of time in the Lorenz attractor starting at two close initial points. The initial differences will soon behave quite differently- predictive ability is rapidly lost.

The maximal lyapunov exponent

Maximal Lyapunov Exponent

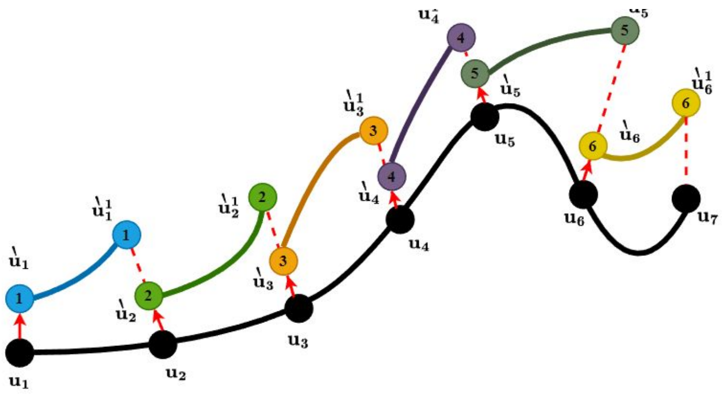


$$\lambda_{max} = \mathbb{E}\left[\log \frac{\|\grave{\mathbf{u}}_t^1 - \mathbf{u}_{t+1}\|}{\|\grave{\mathbf{u}}_t - \mathbf{u}_t\|}\right]$$

It can be understood as the expected logarithm of the growth rate of unit errors near a strange attractor.

Lyapunov Horizon Loss

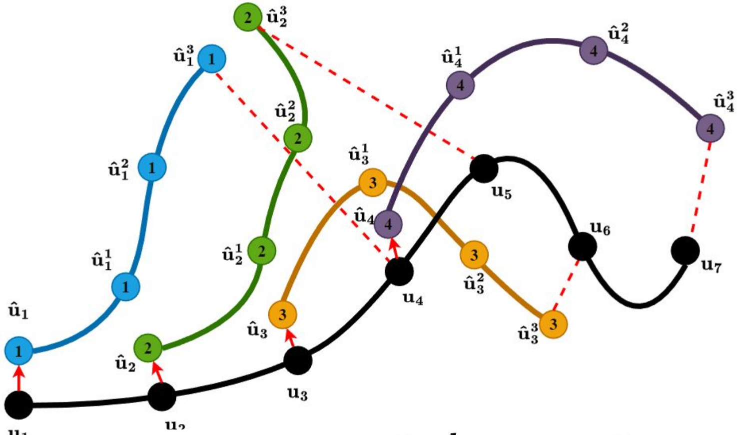
Maximal Lyapunov Exponent



$$\lambda_{max} = \mathbb{E}\left[\log \frac{\|\grave{\mathbf{u}}_t^1 - \mathbf{u}_{t+1}\|}{\|\grave{\mathbf{u}}_t - \mathbf{u}_t\|}\right]$$

It can be understood as the expected logarithm of the growth rate of unit errors near a strange attractor.

Lyapunov Horizon Loss

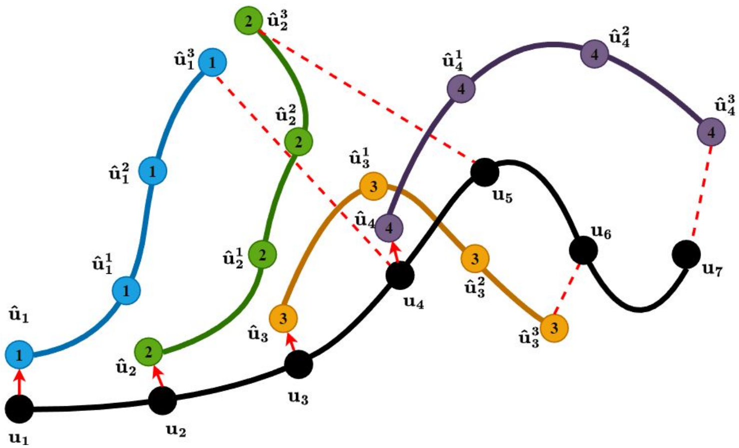


$$\lambda_{loss} = \sum \log \frac{\|\hat{\mathbf{u}}_t^k - \mathbf{u}_{t+k}\|}{\|\hat{\mathbf{u}}_t - \mathbf{u}_t\|}$$

We track how errors at each time step on the trajectory evolve from step t to the time horizon t+k.

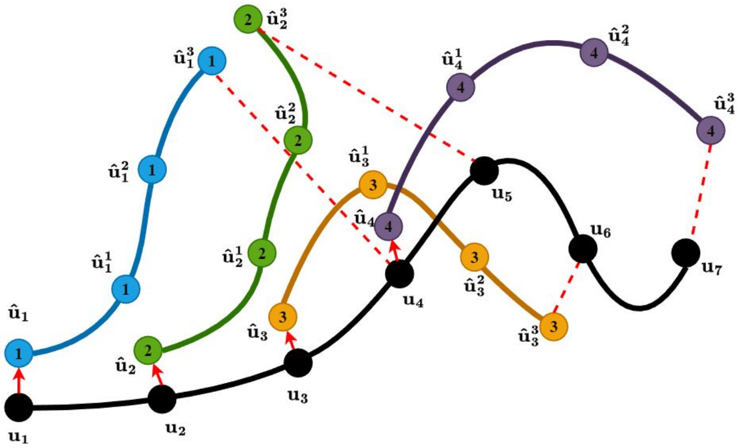
Error Trajectory Tracing Architecture (ETT)

Lyapunov Horizon Loss

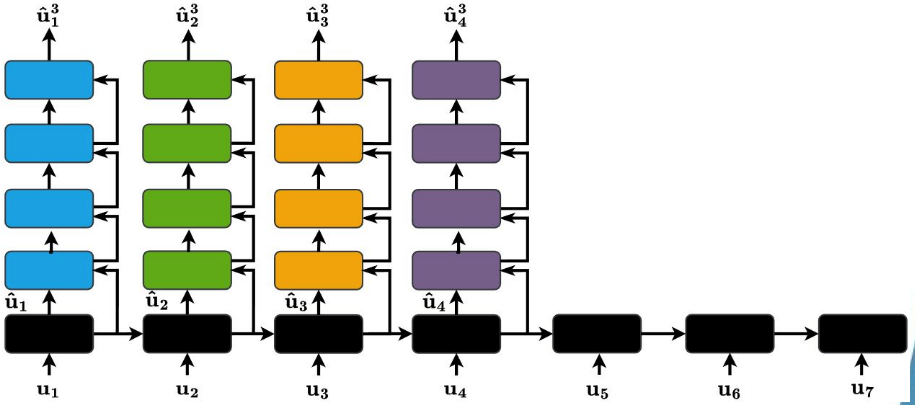


Error Trajectory Tracing Architecture (ETT)

Lyapunov Horizon Loss

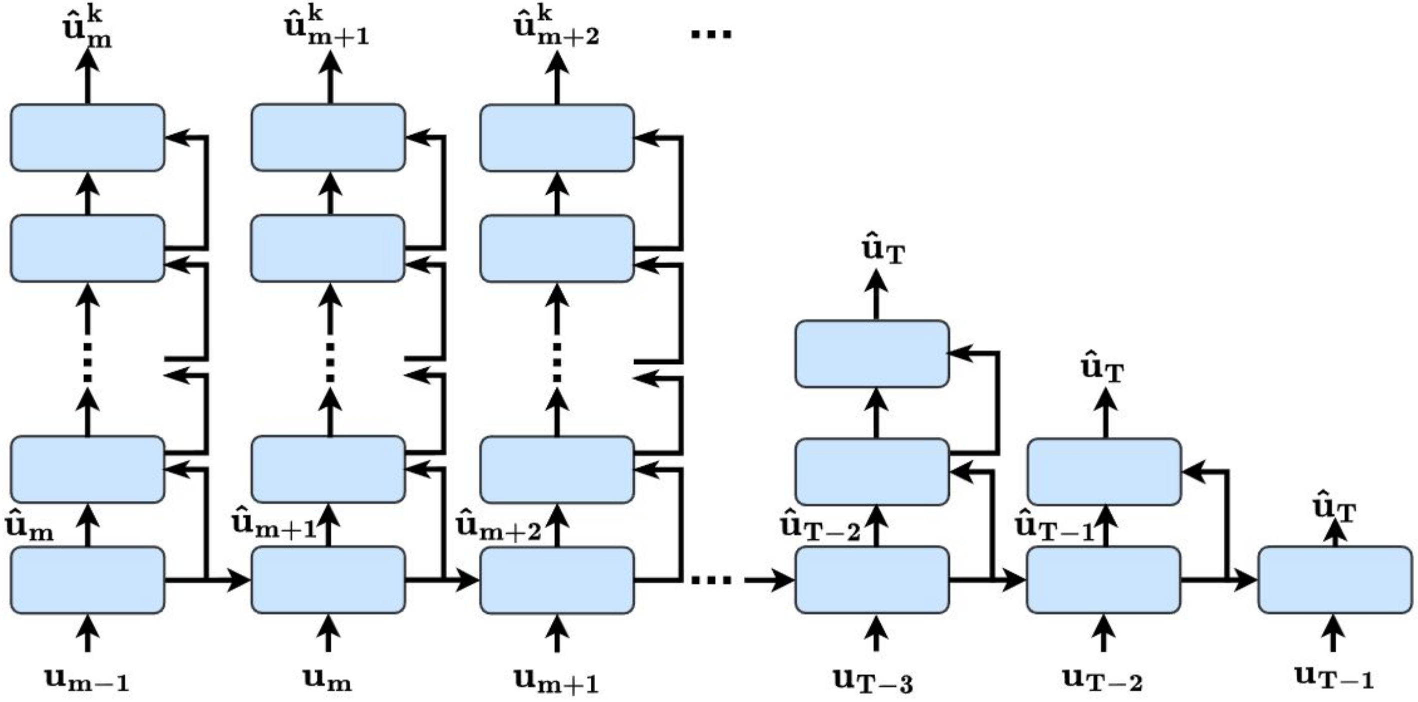


Modeling via a recurrent architecture



UMASS BOSTON

Error Trajectory Tracing Architecture (ETT)



Loss Function

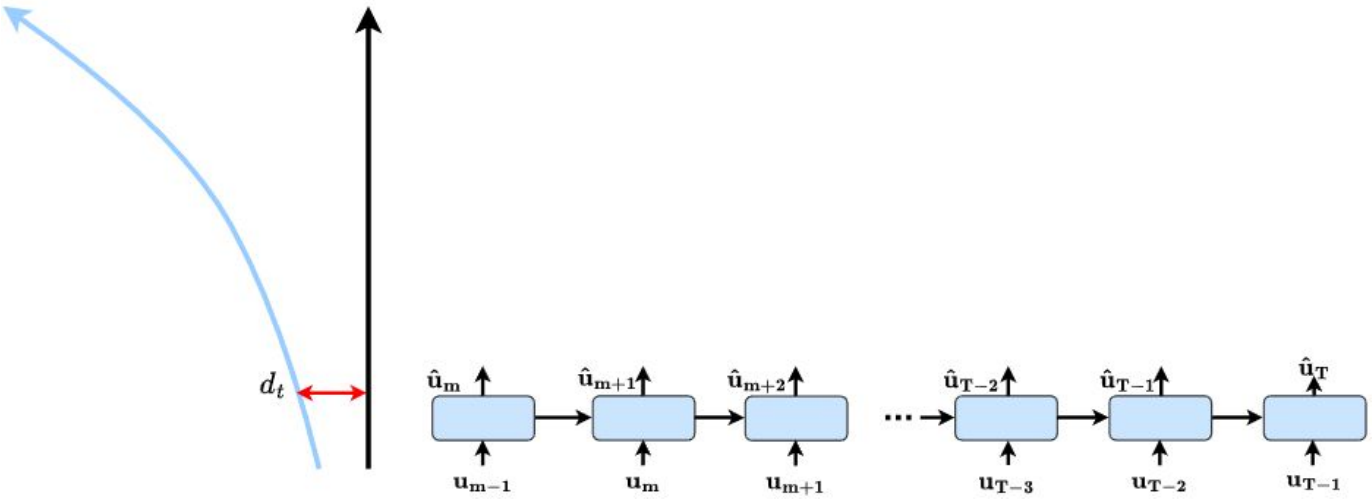
$$\theta^* = \arg\min_{\theta}(\mathcal{L} + \lambda_{loss})$$
$$= \arg\min_{\theta} \sum_{i=1}^{N} \left[\sum_{t=1}^{T} (\hat{\mathbf{u}}_t^{(i)} - \mathbf{u}_t^{(i)})^2 + \sum_{t=1}^{T-k} \lambda_t^{(i)} \right]$$
$$= \arg\min_{\theta} \sum_{i=1}^{N} \left[\sum_{t=1}^{T} (\hat{\mathbf{u}}_t^{(i)} - \mathbf{u}_t^{(i)})^2 + \sum_{t=1}^{T-k} \log \frac{\|\hat{\mathbf{u}}_t^{(i),k} - \mathbf{u}_{t+k}^{(i)}\|}{\|\hat{\mathbf{u}}_t^{(i)} - \mathbf{u}_t^{(i)}\|} \right]$$
$$= \arg\min_{\theta} \sum_{i=1}^{N} \left[\sum_{t=1}^{T} (d_t^{(i)})^2 + \sum_{t=1}^{T-k} \log \frac{\left\| d_t^{(i),k} \right\|}{\left\| d_t^{(i)} \right\|} \right]$$

Training procedure

$$\theta^* = \arg\min_{\theta}(\mathcal{L} + \lambda_{loss})$$
$$= \arg\min_{\theta} \sum_{i=1}^{N} \left[\sum_{t=1}^{T} (\hat{\mathbf{u}}_t^{(i)} - \mathbf{u}_t^{(i)})^2 + \sum_{t=1}^{T-k} \lambda_t^{(i)} \right]$$
$$= \arg\min_{\theta} \sum_{i=1}^{N} \left[\sum_{t=1}^{T} (\hat{\mathbf{u}}_t^{(i)} - \mathbf{u}_t^{(i)})^2 + \sum_{t=1}^{T-k} \log \frac{\|\hat{\mathbf{u}}_t^{(i),k} - \mathbf{u}_{t+k}^{(i)}\|}{\|\hat{\mathbf{u}}_t^{(i)} - \mathbf{u}_t^{(i)}\|} \right]$$
$$= \arg\min_{\theta} \sum_{i=1}^{N} \left[\sum_{t=1}^{T} (d_t^{(i)})^2 + \sum_{t=1}^{T-k} \log \frac{\left\| d_t^{(i),k} \right\|}{\left\| d_t^{(i)} \right\|} \right]$$

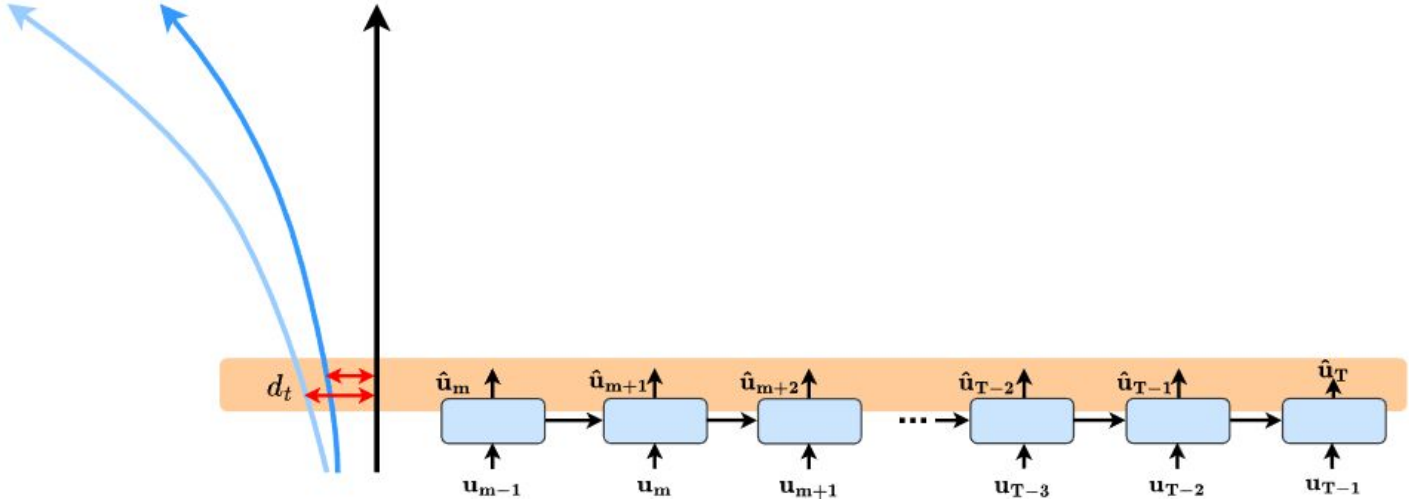
- **Initialize base model**
- **Train base model**
- **Build ETT**
- **Train model for horizon K**

Training procedure: Horizon Forcing



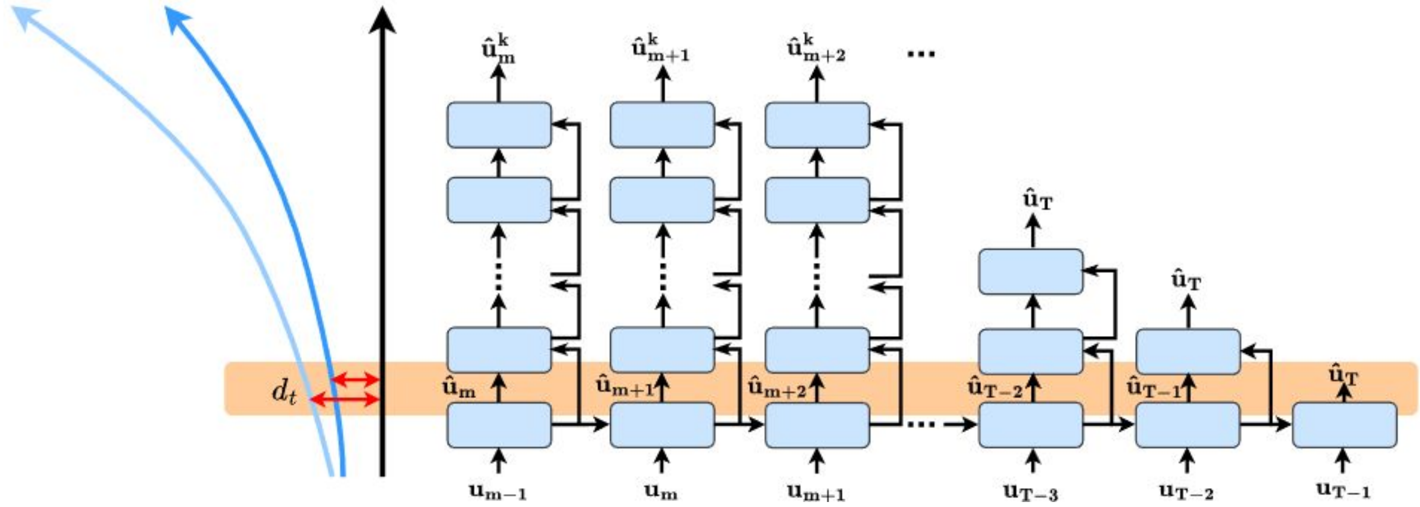
- Initialize base model <<<
- Train base model
- Build ETT
- Train model for horizon K

Training procedure: Horizon Forcing



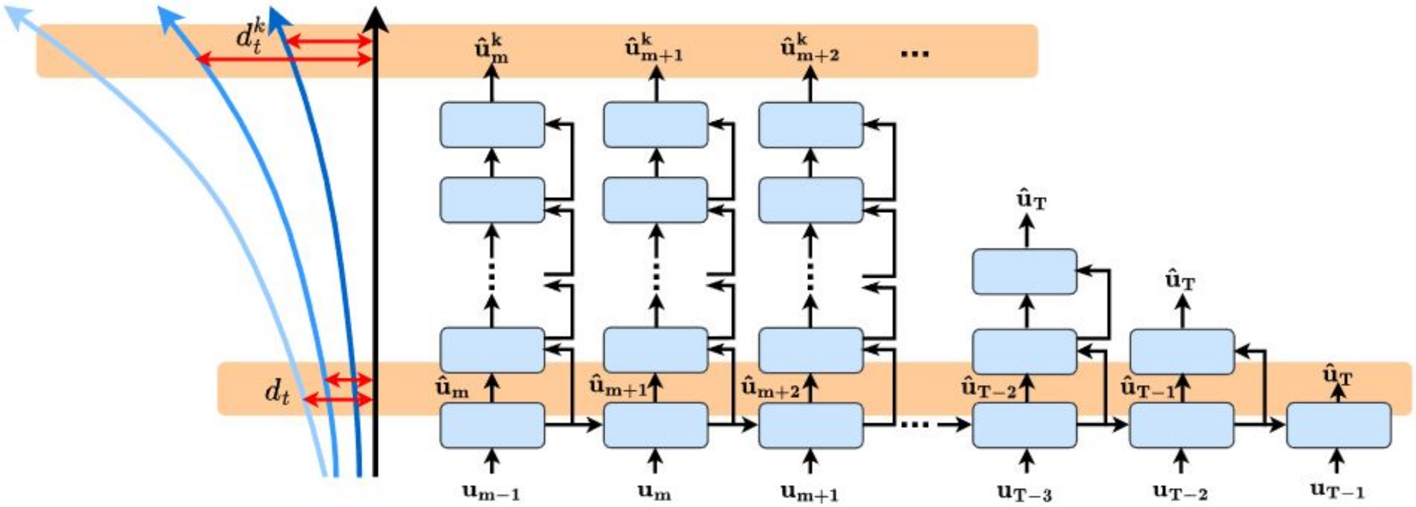
- Initialize base model
- Train base model <<<
- Build ETT
- Train model for horizon K

Training procedure: Horizon Forcing



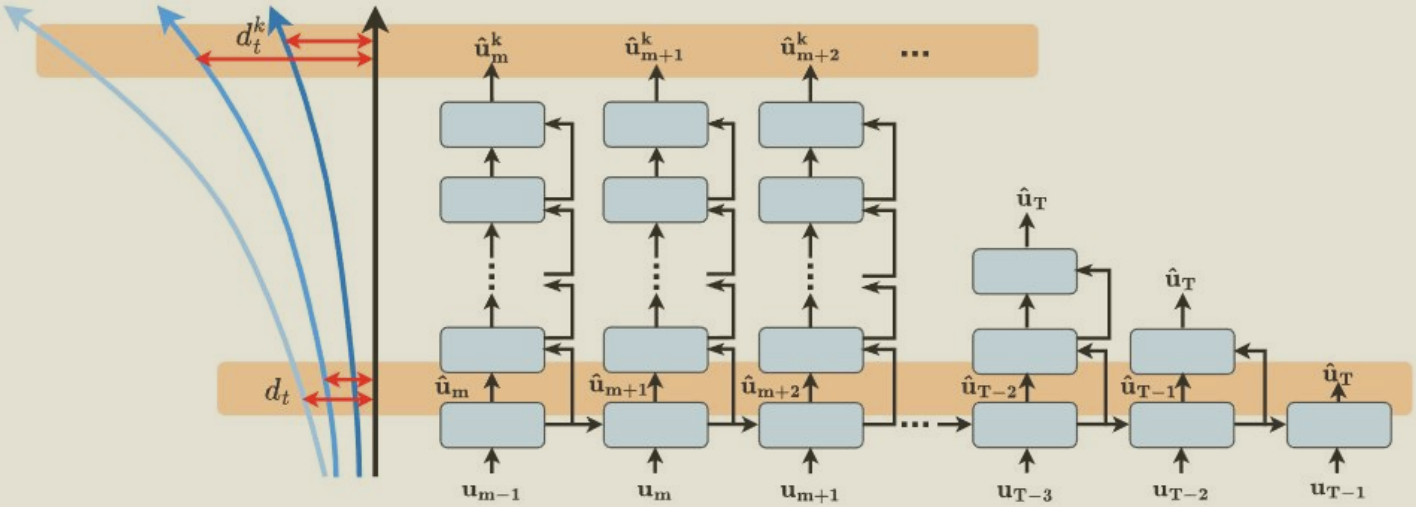
- Initialize base model
- Train base model
- **Build ETT <<<**
- Train model for horizon K

Training procedure: Horizon Forcing



- Initialize base model
- Train base model
- Build ETT
- **Train model for horizon K <<<**

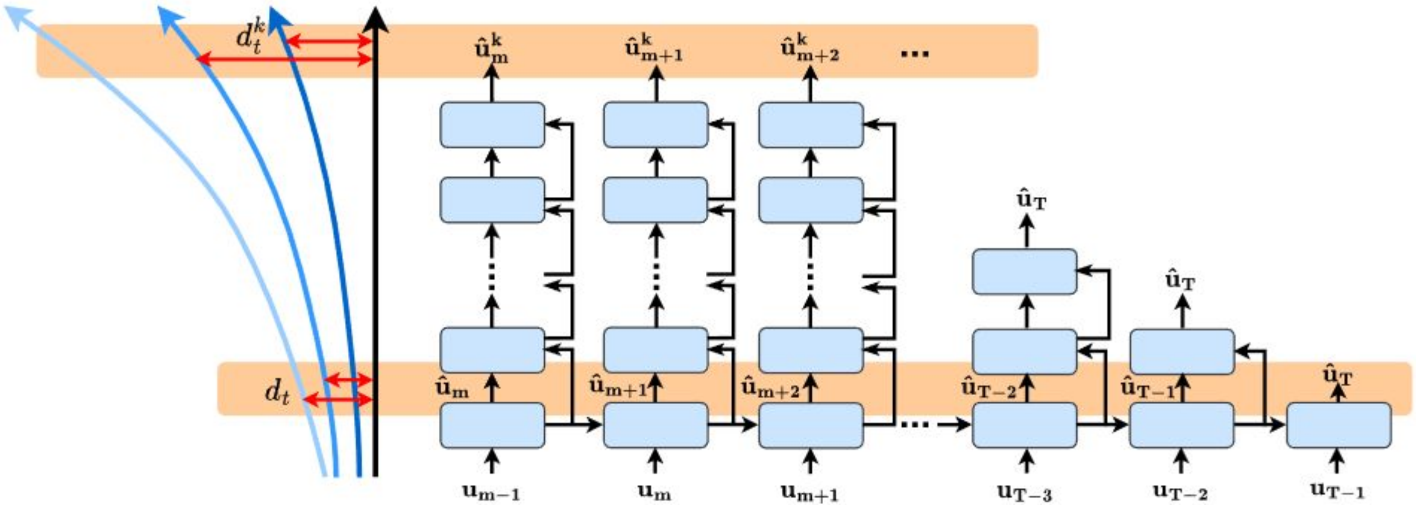
Training procedure: Horizon Forcing



- Initialize base model
- Train base model
- Build ETT
- **Train model for horizon K <<<**

How to train model for bigger horizon?

Training procedure: Horizon Forcing



- Initialize base model
- Train base model
- Build ETT
- Train model for the **next** horizon K

Repetition

Experiments Setting

Metrics: Prediction horizon and Expectation

Evaluation Metric

Prediction Horizon: $\mathbb{P}_M = \min_j \left\{ j \;\middle|\; \frac{1}{N} \sum_{i=1}^{N} M\left(\mathbf{u}_j^{(i)}, \hat{\mathbf{u}}_j^{(i)}\right) > \delta_M \right\}$

Time step

Threshold

Expectation: $\mathbb{E}_M = \frac{1}{N \times T} \sum_{i=1}^{N} \sum_{j=1}^{T} M\left(\mathbf{u}_j^{(i)}, \hat{\mathbf{u}}_j^{(i)}\right)$

UMASS BOSTON

Experiments Setting

Three metrics for forecasting analysis

Metric	Expression	Symbols: Expectation & Prediction Horizon
Root Mean Squared Error (RMSE)	$\gamma\left(\mathbf{u}_t^{(i)}, \hat{\mathbf{u}}_t^{(i)}\right) = \sqrt{\frac{1}{D}\sum_{d=1}^{D}\left\lVert \hat{\mathbf{u}}_{t,d}^{(i)} - \mathbf{u}_{t,d}^{(i)}\right\rVert^2}$	$\mathbb{E}_\gamma ; \mathbb{P}_\gamma$
Mean Normalized Error (MNE)	$\zeta\left(\mathbf{u}_t^{(i)}, \hat{\mathbf{u}}_t^{(i)}\right) = \frac{1}{D}\sum_{d=1}^{D}\frac{\left\lVert \hat{\mathbf{u}}_{t,d}^{(i)} - \mathbf{u}_{t,d}^{(i)}\right\rVert}{\left\lVert \mathbf{u}_{t,d}^{(i)}\right\rVert}$	$\mathbb{E}_\zeta ; \mathbb{P}_\zeta$
Symmetric Mean Absolute Percent Error (SMAPE)	$\mu\left(\mathbf{u}_t^{(i)}, \hat{\mathbf{u}}_t^{(i)}\right) = \frac{1}{D}\sum_{d=1}^{D}\frac{\left\lVert \hat{\mathbf{u}}_{t,d}^{(i)} - \mathbf{u}_{t,d}^{(i)}\right\rVert}{\left\lVert \hat{\mathbf{u}}_{t,d}^{(i)}\right\rVert + \left\lVert \mathbf{u}_{t,d}^{(i)}\right\rVert}$	$\mathbb{E}_\mu ; \mathbb{P}_\mu$

Experiments Setting

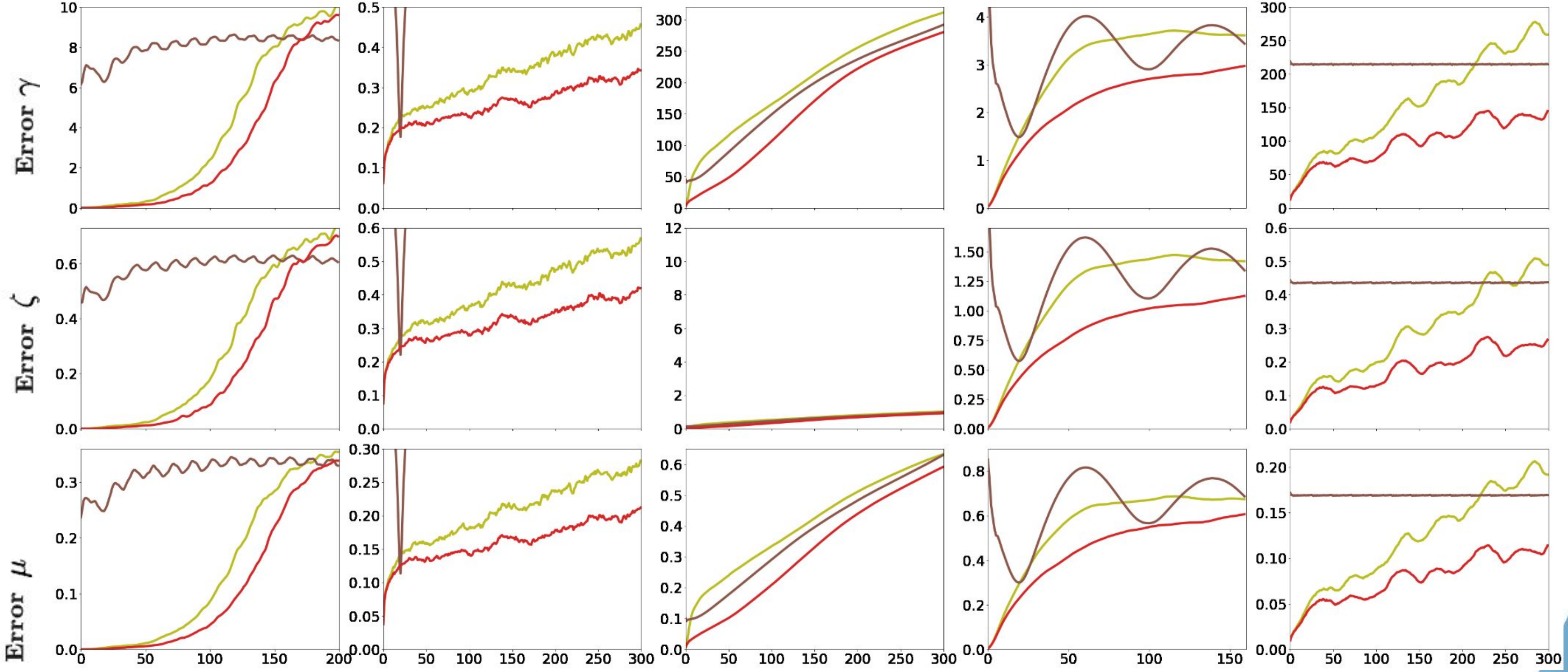
Five time series prediction tasks with chaotic characteristics.

Data sets	# of vars	# of training	# of testing	# of forecasting steps	δγ	δζ	δμ
Lorenz	3	8,500	1,201	200	3.11	0.23	0.11
Roaming Worm	5	5,000	1,649	160	2.27	0.86	0.46
Accelerometer	3	5,500	801	300	0.29	0.36	0.18
Gait Force	6	5,600	1,023	300	158.69	0.50	0.26
Electricity	1	5,000	1,081	300	121.54	0.22	0.06

Ablation Study: Horizon-Forced Model VS Teaching Forcing VS ETT Direct

Effectivity checking

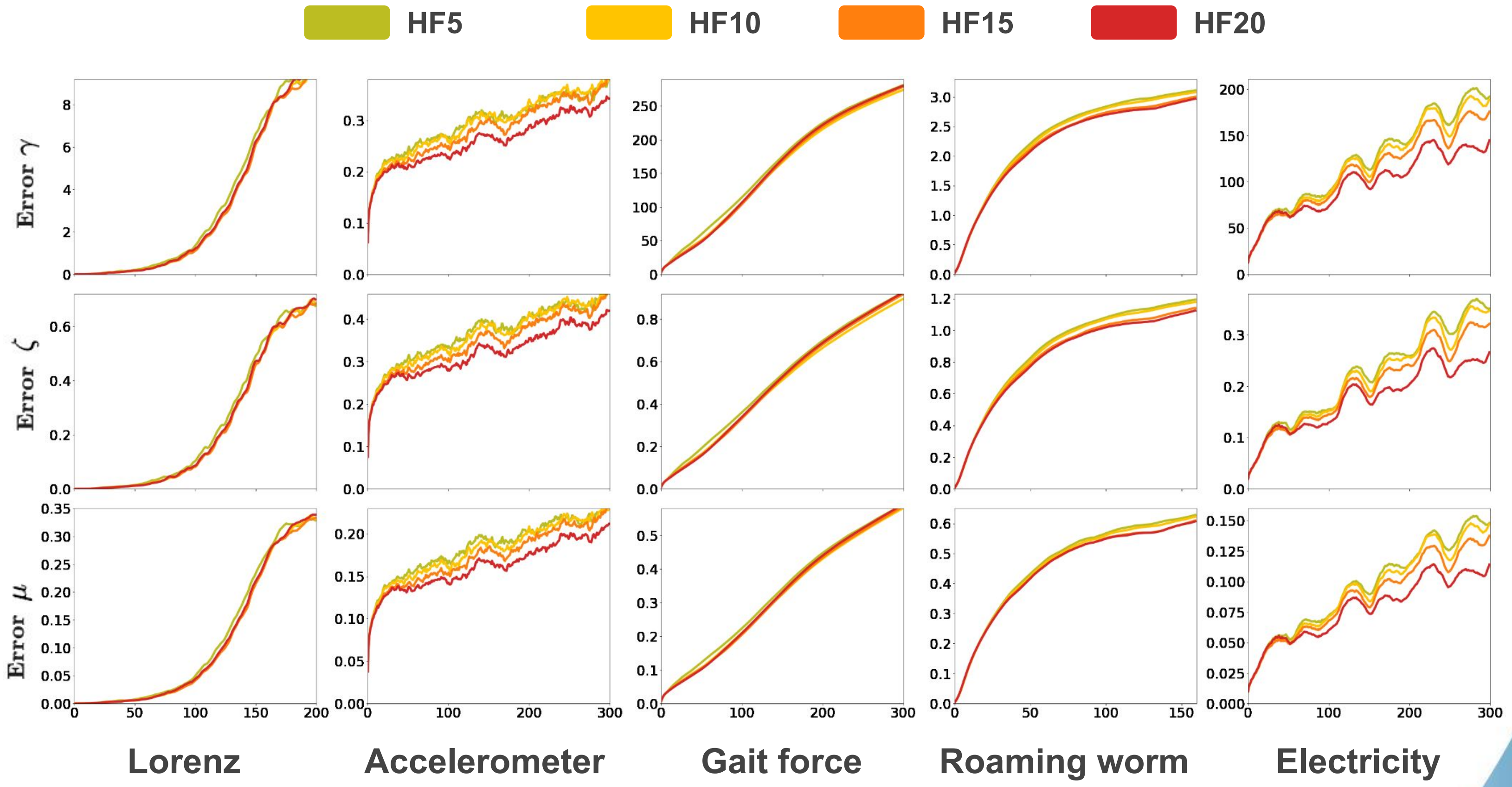
Teacher Forcing | ETT Direct | HF20



Lorenz | Accelerometer | Gait force | Roaming worm | Electricity

UMASS BOSTON

Ablation Study: Choice of Horizon

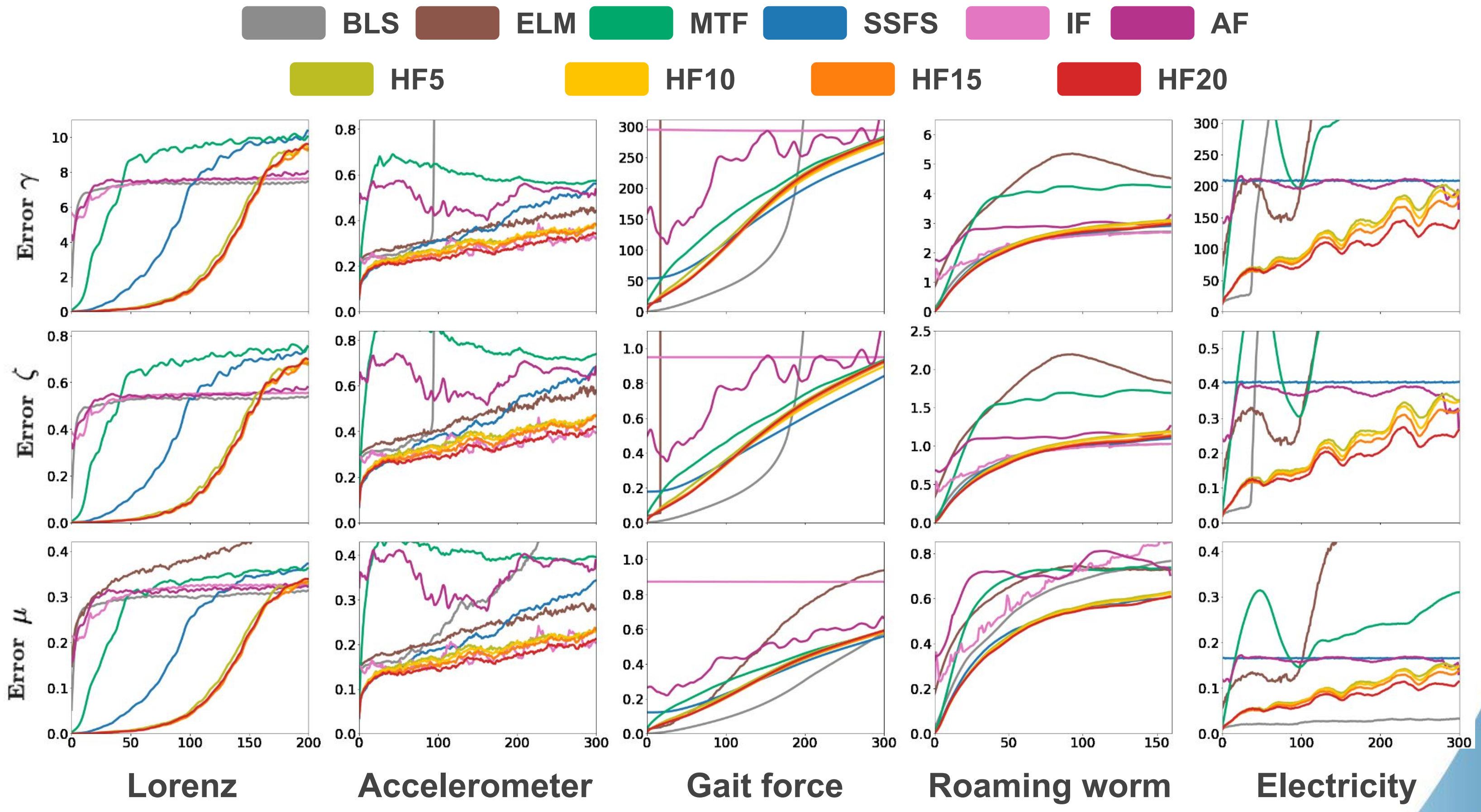


Experiments Setting

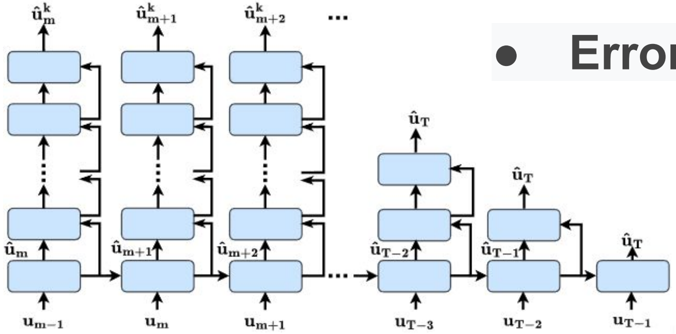
Six competitors for long-term behavior prediction

- Broad Learning System (BLS)
- Extreme learning machine (ELM)
- Scheduled Sampling - Full Schedule(SSFS)
- Music Transformer (MTF)
- Informer (IF)
- AutoFormer (AF)

Benchmarking



CONCLUSION



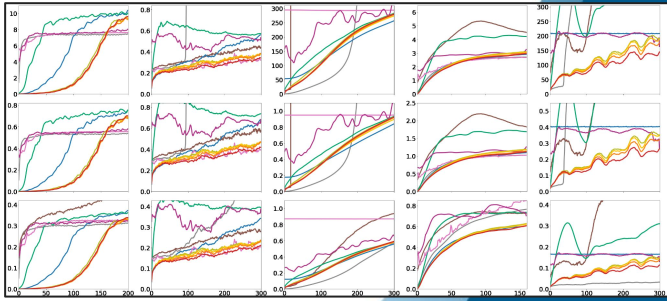
- Error Trajectory Tracing (ETT)

- Horizon Forcing (HF)

- Initialize base model
- Train base model
- Build ETT
- Train model for the **next** horizon K

Repetition

- Extensive experiments on 5 data sets



Thank You!
Questions & Answers