

# **Introduction to Physics informed machine learning**

Dr. Suchita Kulkarni

# What is physics informed machine learning?

- Machine learning: art of making data driven decisions
  - Works incredibly well in presence of lot of data
  - May inherit data biases, may not extrapolate well
  - Explainability is challenge: feels like a blackbox
- Physics informed machine learning (PIML)
  - Art of injecting known knowledge of physical systems into machine learning
  - Suffers less from data biases, and extrapolates better
  - Contains built-in explainability

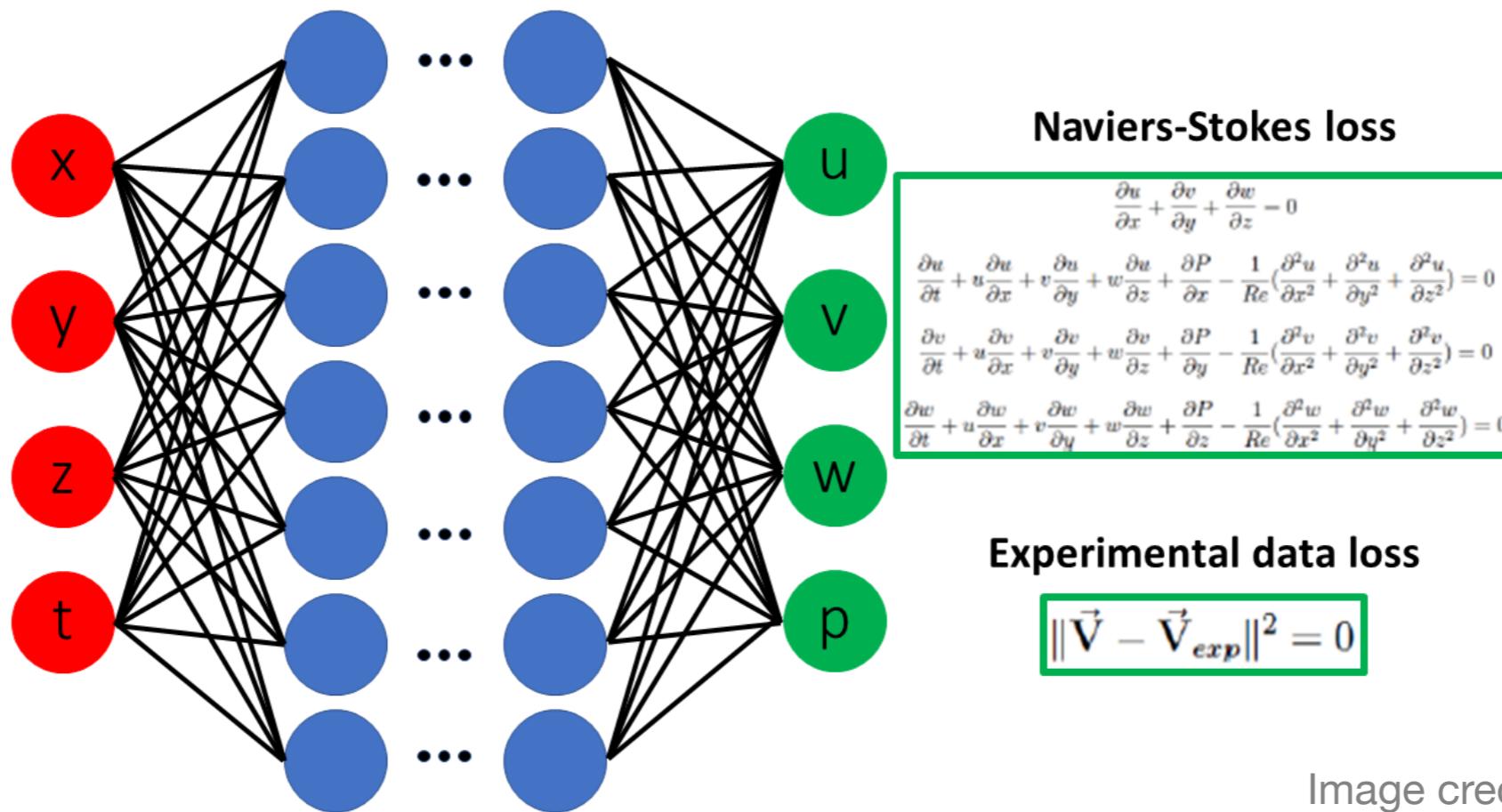


Image credit: Riccardo Munafò

# PIML success stories

## How Google's DeepMind tool is 'more quickly' forecasting hurricane behavior

'Less expensive and time consuming' model helps with fast and accurate predictions, possibly saving lives and property

arXiv:2507.11589

Guardian



## EINSTEIN FIELDS: A NEURAL PERSPECTIVE TO COMPUTATIONAL GENERAL RELATIVITY

Sandeep S. Cranganore<sup>\* 1</sup>

Andrei Bodnar<sup>\* 2</sup>  
Johannes Brandstetter<sup>1,3</sup>

Arturs Berzins<sup>\* 1</sup>

<sup>\*</sup>Equal contribution

<sup>1</sup>LIT AI Lab, Institute for Machine Learning, JKU Linz, Austria

<sup>2</sup>University of Manchester, United Kingdom

<sup>3</sup>Emmi AI GmbH, Linz, Austria

{cranganore, berzins, brandstetter}@ml.jku.at  
andrei.bodnar@student.manchester.ac.uk

## AURORA: A FOUNDATION MODEL OF THE ATMOSPHERE

Cristian Bodnar<sup>\* 1</sup>, Wessel P. Bruinsma<sup>\* 1</sup>, Ana Lucic<sup>\* 1</sup>, Megan Stanley<sup>\* 1</sup>,  
Johannes Brandstetter<sup>3, †</sup>, Patrick Garvan<sup>1</sup>, Maik Riechert<sup>1</sup>, Jonathan Weyn<sup>2</sup>, Haiyu Dong<sup>2</sup>,  
Anna Vaughan<sup>4</sup>, Jayesh K. Gupta<sup>5, †</sup>, Kit Tambiratnam<sup>2</sup>, Alex Archibald<sup>4</sup>, Elizabeth Heider<sup>1</sup>,  
Max Welling<sup>6, †</sup>, Richard E. Turner<sup>1, 4</sup>, and Paris Perdikaris<sup>1</sup>

<sup>1</sup>Microsoft Research AI for Science

<sup>2</sup>Microsoft Corporation

<sup>3</sup>JKU Linz

<sup>4</sup>University of Cambridge

<sup>5</sup>Poly Corporation

<sup>6</sup>University of Amsterdam

<sup>\*</sup>Equal contribution    <sup>†</sup>Work done while at Microsoft Research

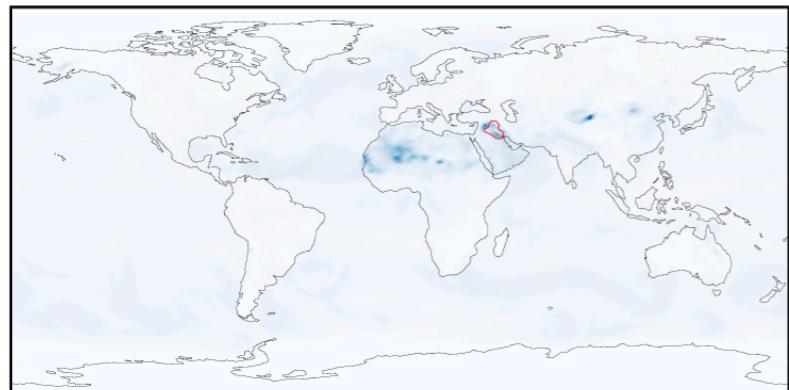
arXiv:2405.13063

# PIML success stories

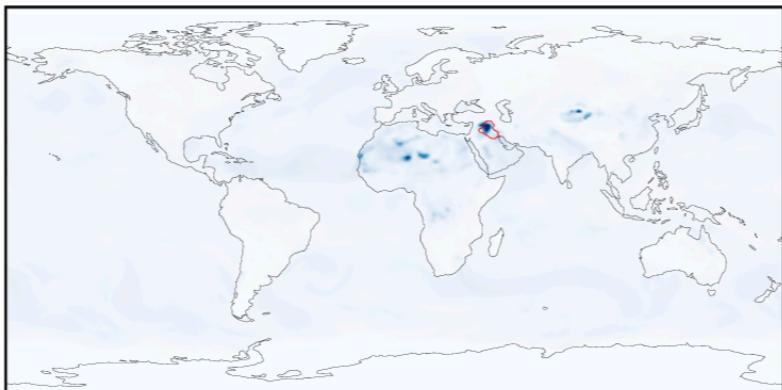
Aurora see storms brewing a day in advance

arXiv:2405.13063

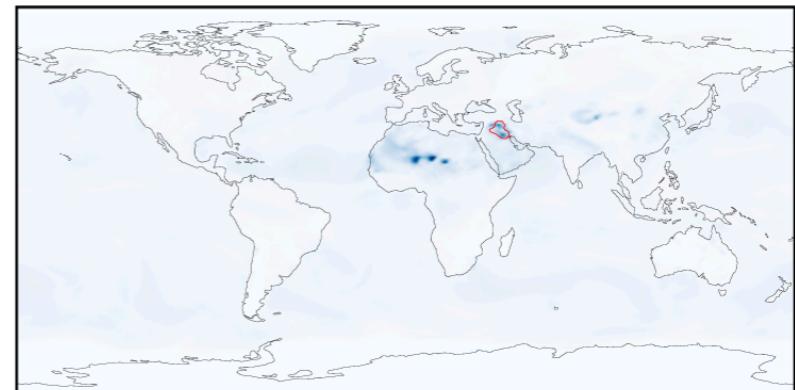
2022-06-12 12 UTC



2022-06-13 00 UTC

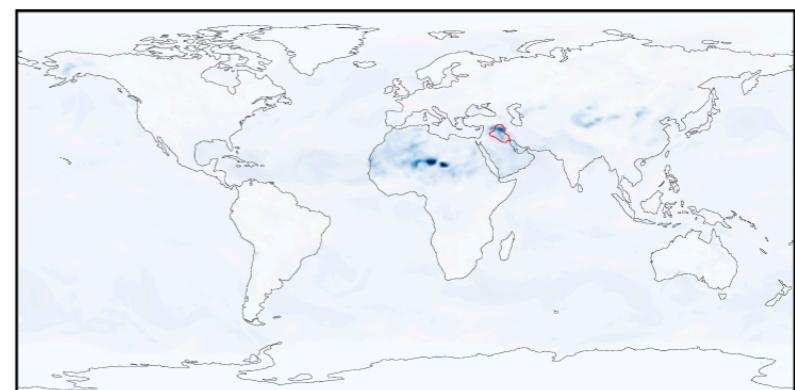
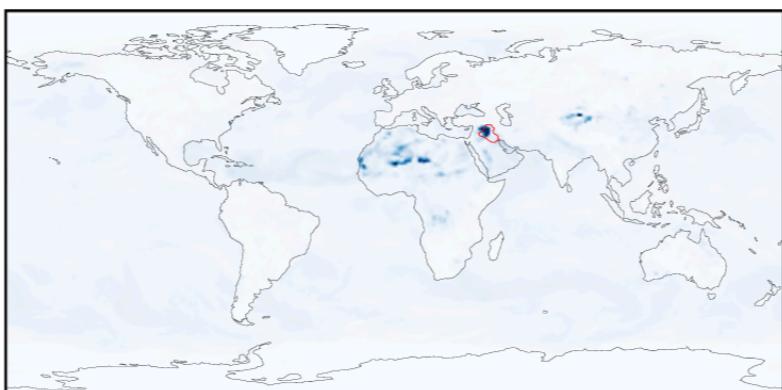
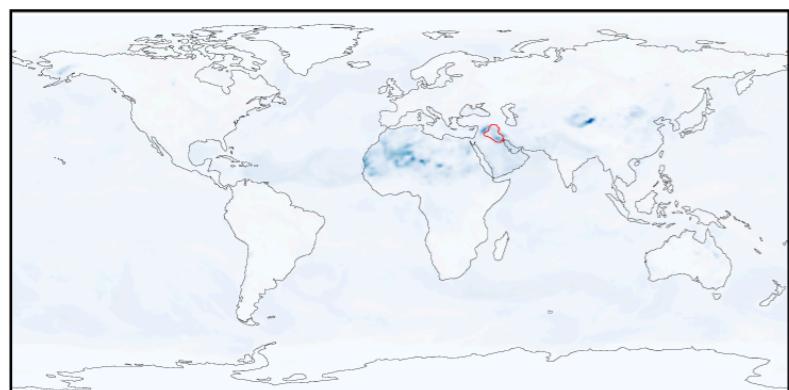


2022-06-13 12 UTC

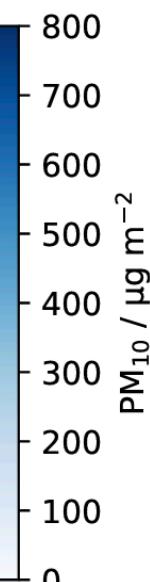


Aurora

2022-06-12 12 UTC



CAMS analysis

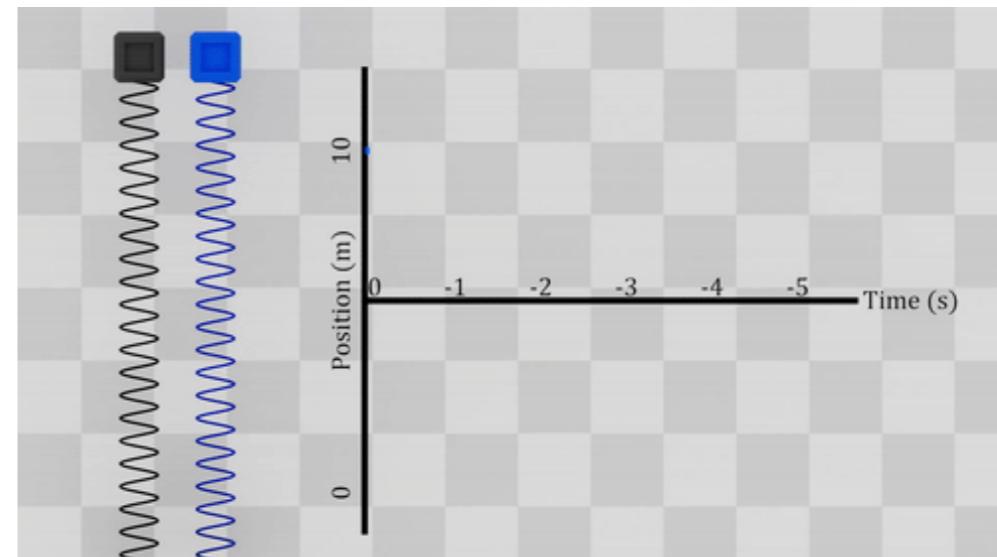
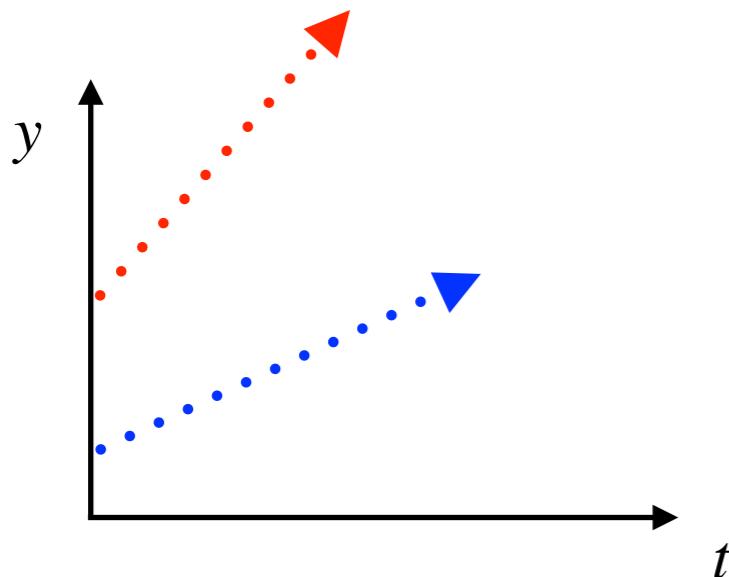


# The physics

- (Partial) differential equations are the backbone of physics
  - Tell us how physical systems change across space and time

$$\frac{dy}{dt} = v \Rightarrow y - y_0 = v \times t$$

$$\frac{d^2y}{dt^2} + \omega^2 y = 0$$



Source wikipedia

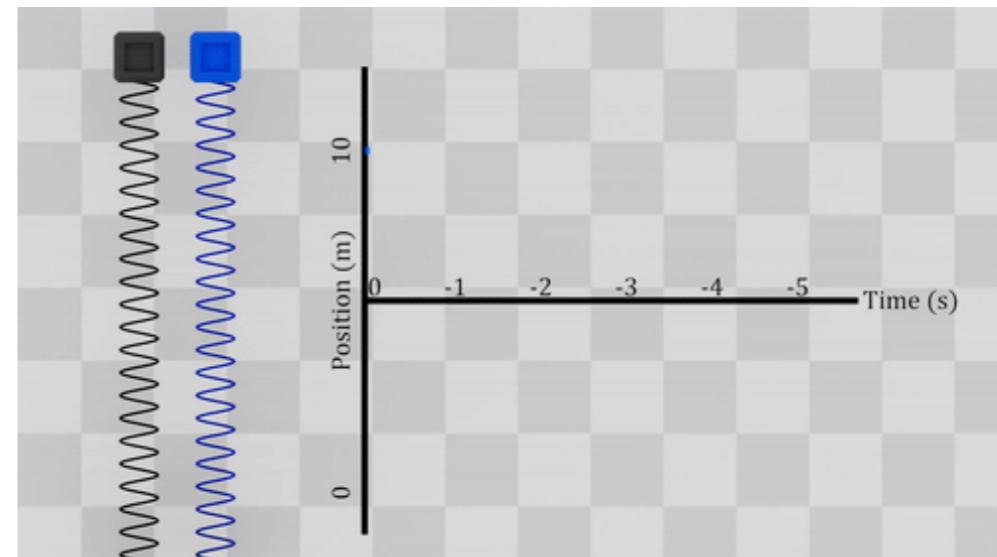
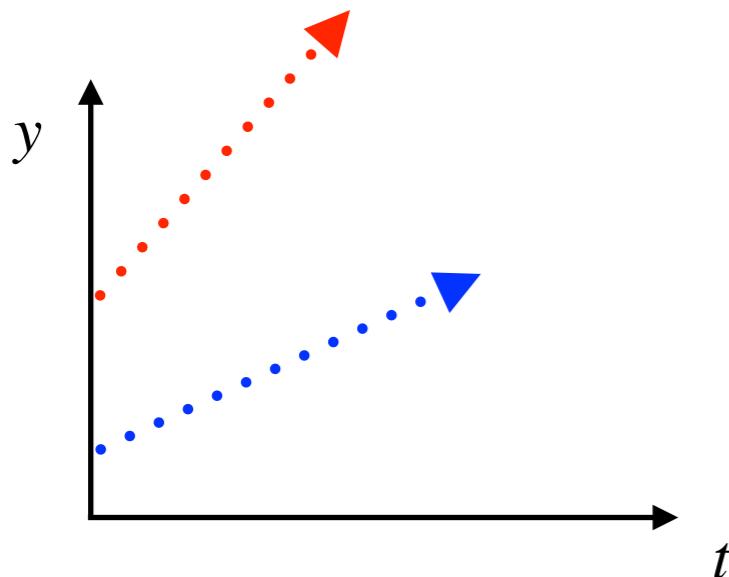
- But apart from partial differential equations, there is a lot of other physics information available
  - Symmetries (not covered today)
  - Conserved quantities (not covered today)

# The physics

- (Partial) differential equations are the backbone of physics
  - Tell us how physical systems change across space and time

$$\frac{dy}{dt} = v \Rightarrow y - y_0 = v \times t$$

$$\frac{d^2y}{dt^2} + \omega^2 y = 0$$

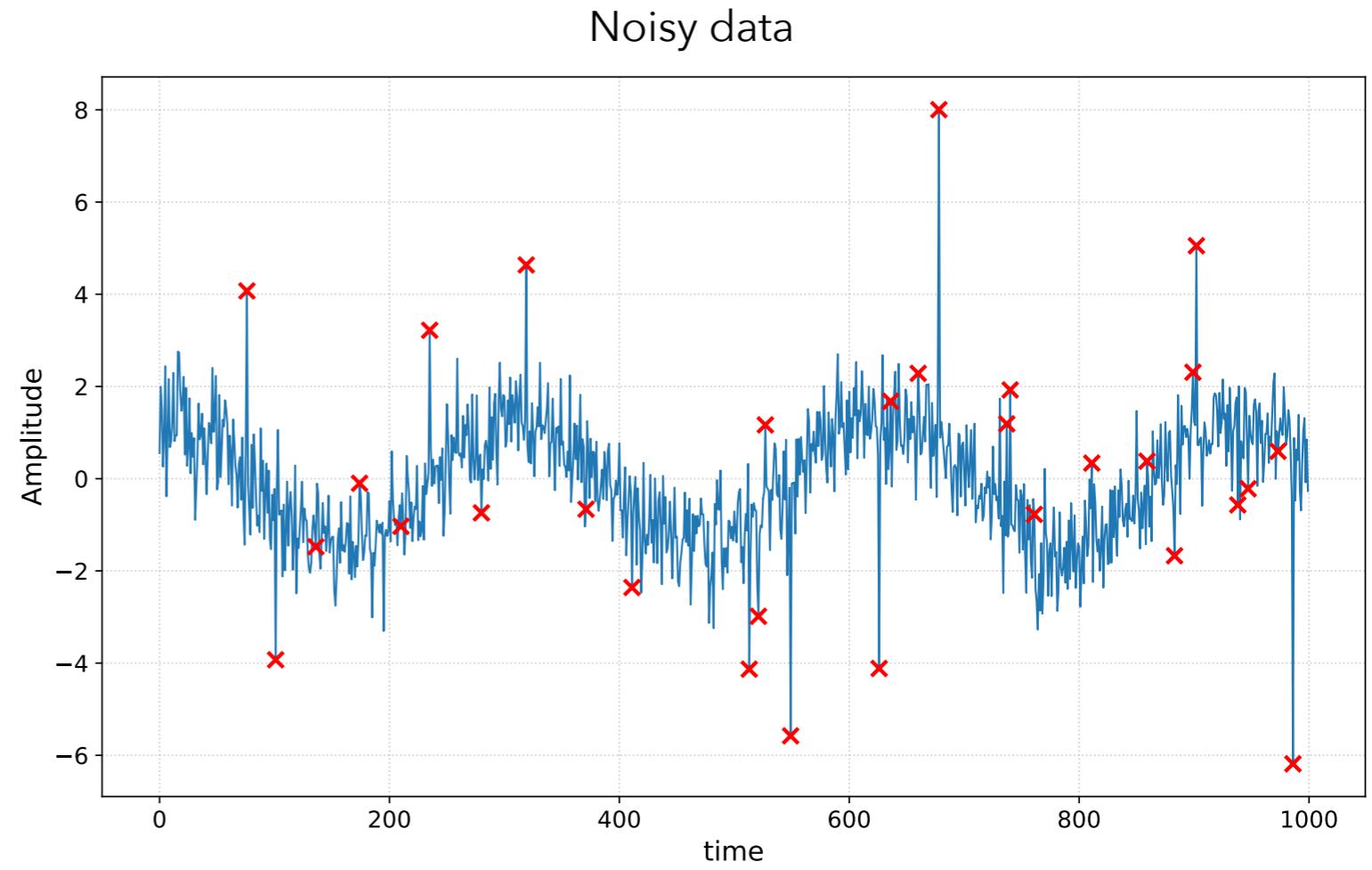
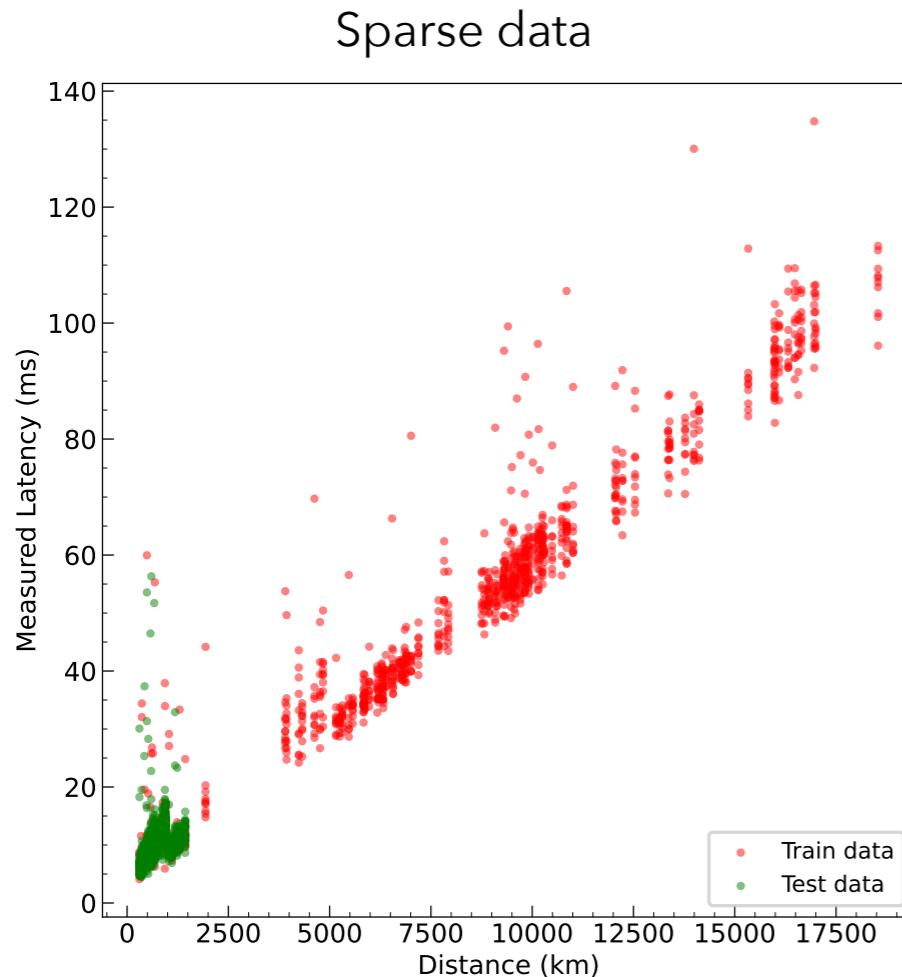


Source wikipedia

- But apart from partial differential equations, there is a lot of other physics information available
  - Symmetries (not covered today)
  - Conserved quantities (not covered today)

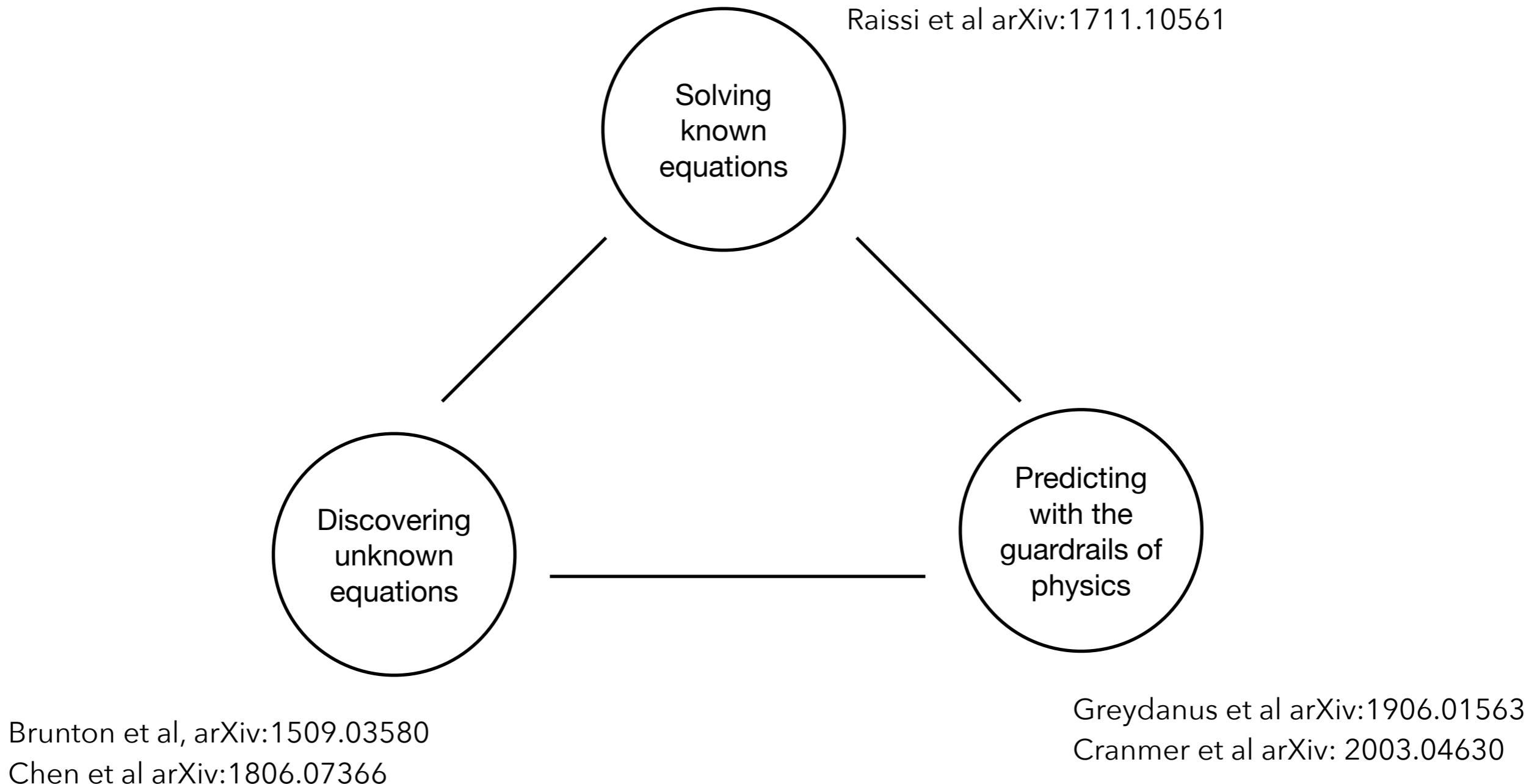
# The problem

- Real life data is hardly ever clean
- Real life data is incomplete or sparse or anomalous



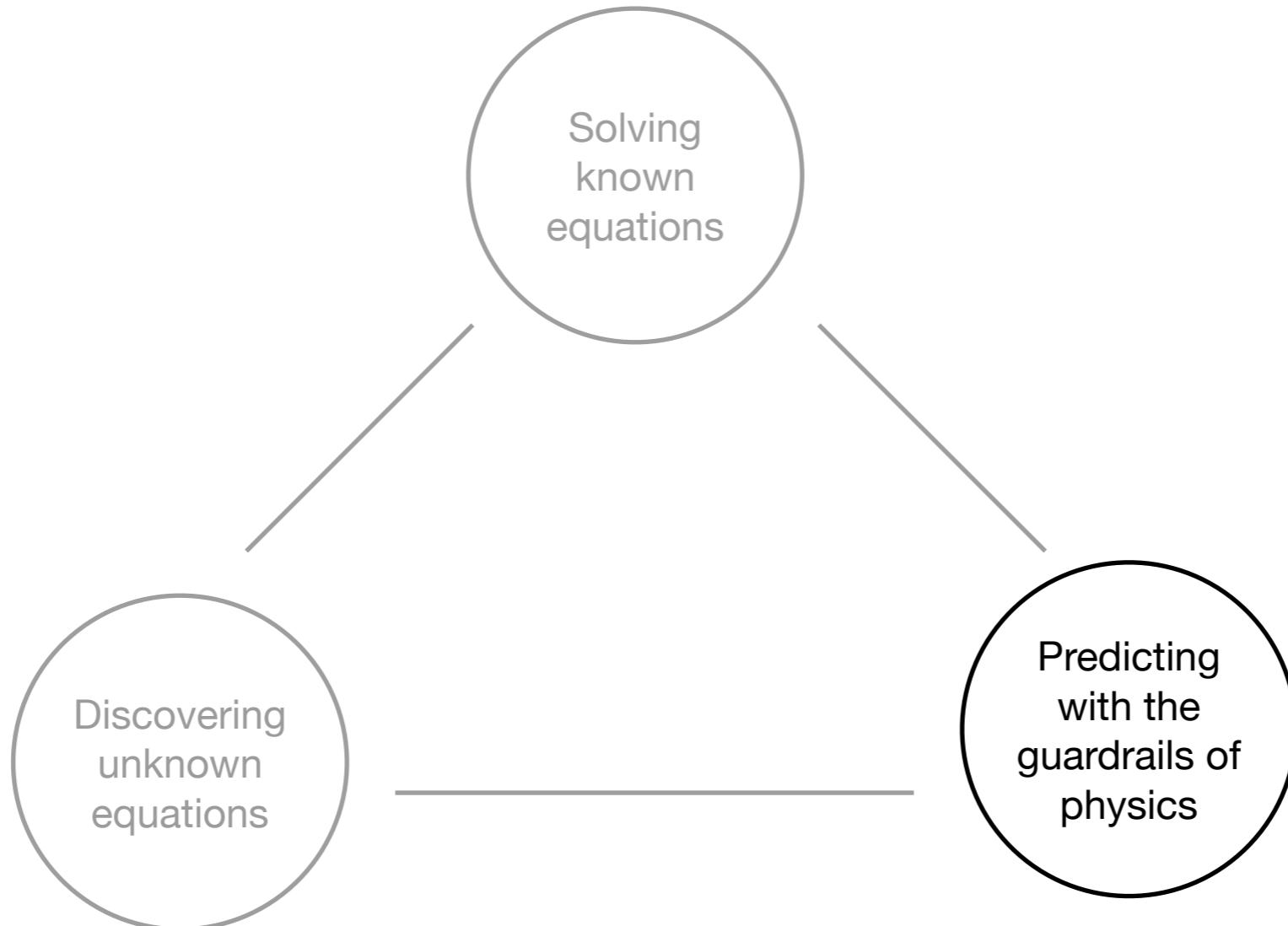
- Often solving differential equations is not sufficient, they need to be aided with some machine learning techniques and this leads to the idea of physics inspired machine learning

# The many aspects of PIML



# Plan for today

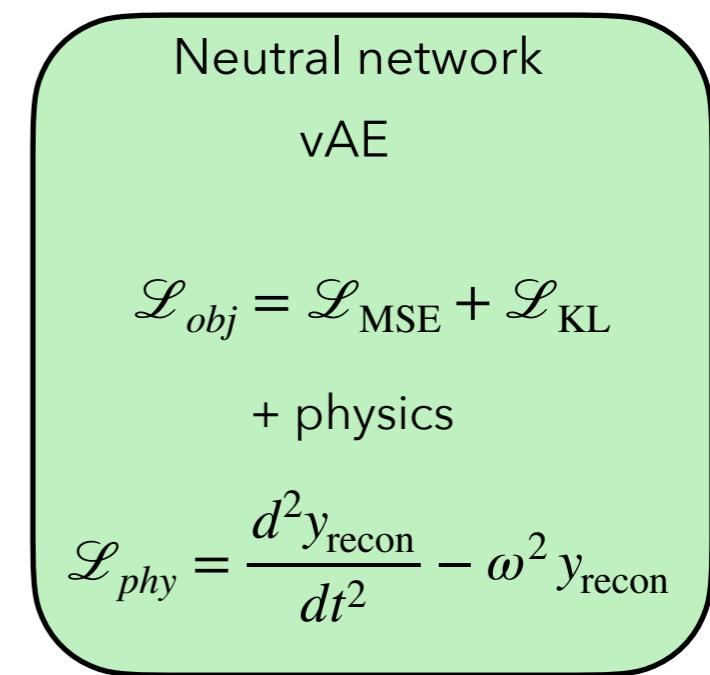
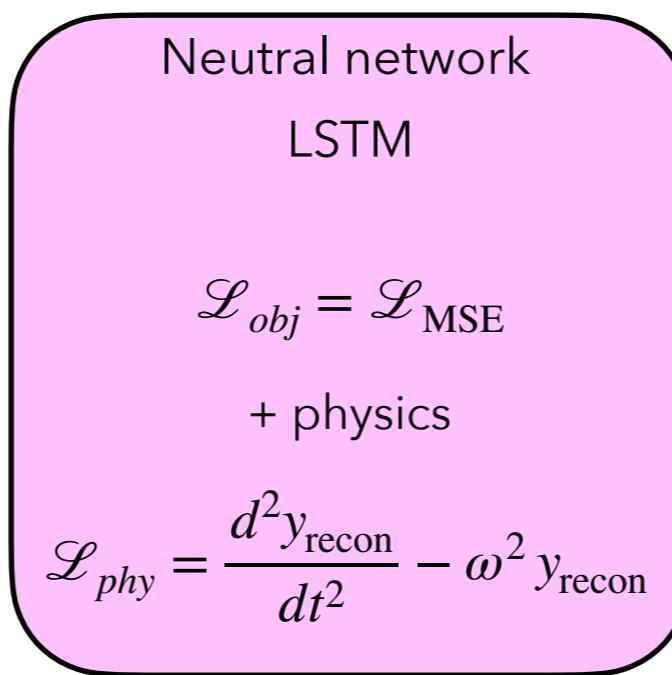
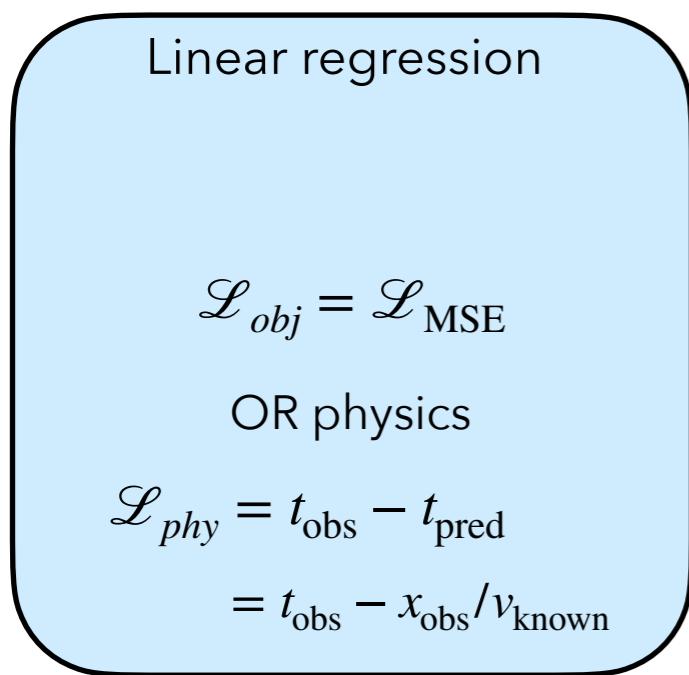
---



Use controlled synthetic signals to isolate the effect of the physics constraints. In a real production setting substitute your domain's signals and keep the same modelling principles.

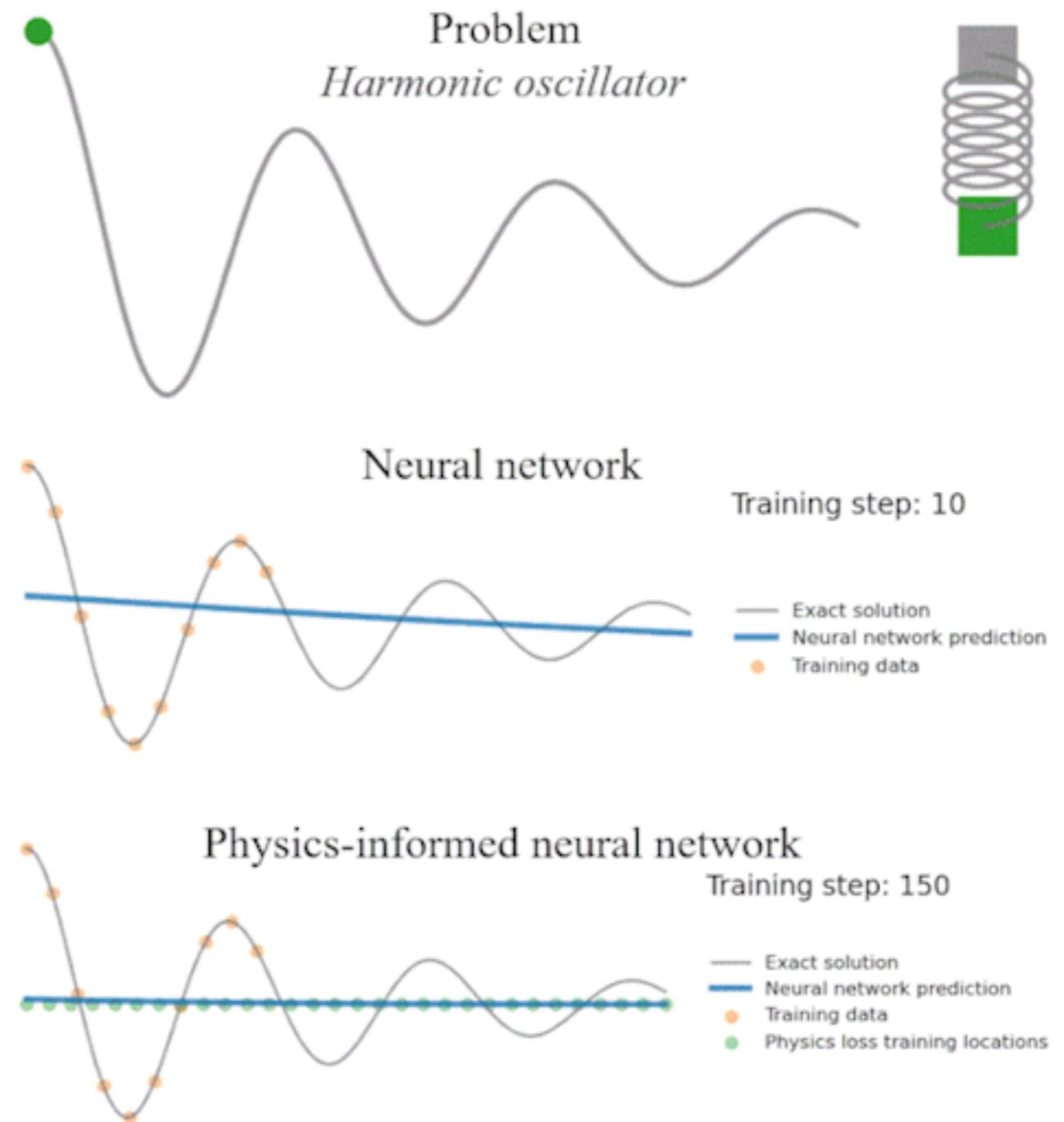
# Predictions with physics guardrails

- The central objective of any machine learning project is minimisation of loss function
- For example
  - MSE: mean squared loss, difference between observed and reconstructed data
    - $\mathcal{L}_{\text{MSE}} = \frac{1}{N} ||y_{\text{recon}} - y_{\text{obs}}||$
  - Kullback - Leibler divergence: measure of difference between two probability distributions
    - $\mathcal{L}_{\text{KL}} = \mathcal{D}_{\text{KL}}(p(x) || q(x)) = \int p(x) \log \frac{p(x)}{q(x)} dx$



# Why does physics help?

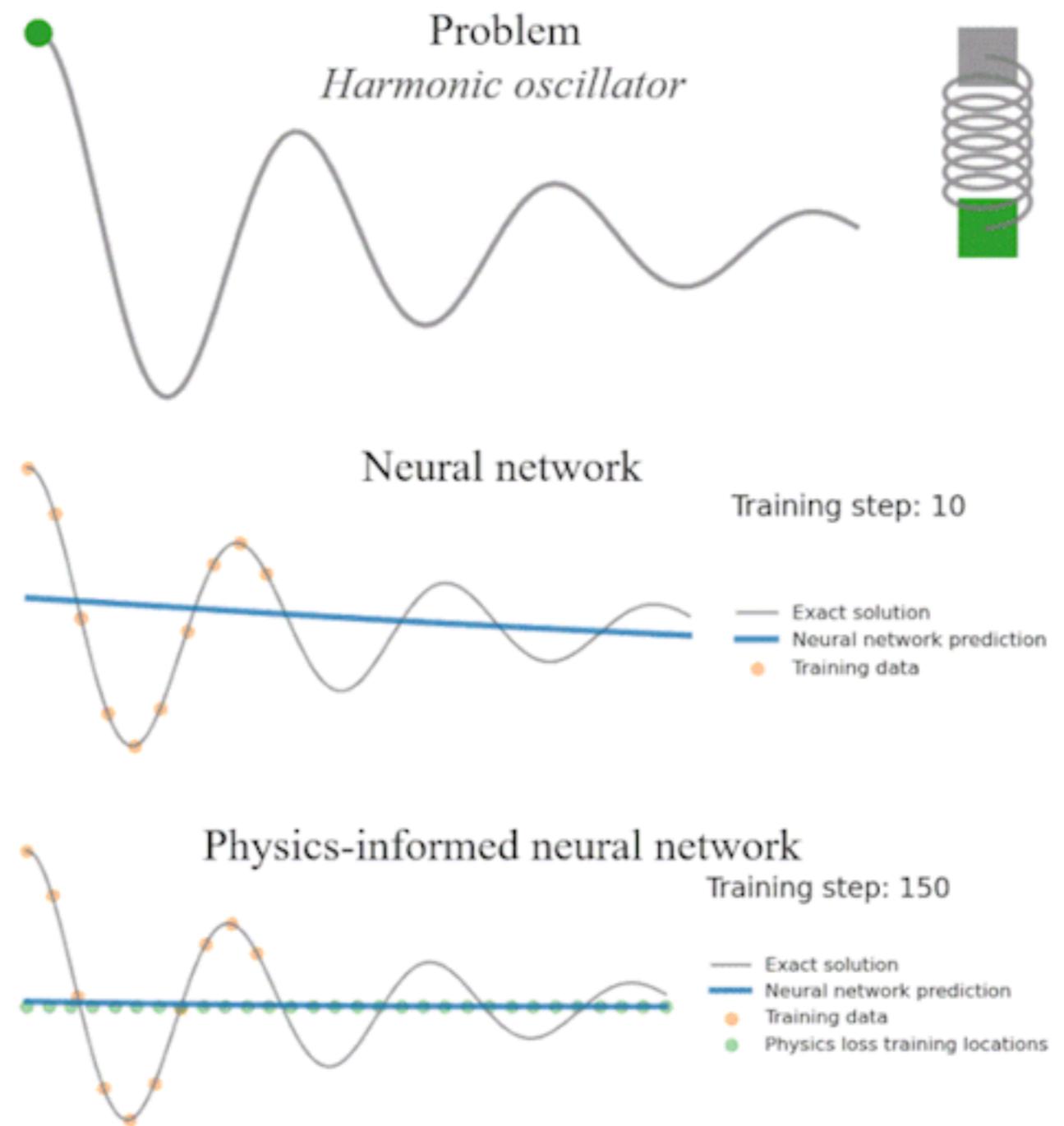
- Regularization: Physics loss acts as a regularizer that restricts the hypothesis space to physically plausible solutions. Without it, the network can fit noise or learn spurious correlations.
- Extrapolation: Pure data-driven models interpolate well but extrapolate poorly. Physics constraints ensure the model respects governing equations even outside the training domain.



Simulation credit: <https://github.com/benmoseley/harmonic-oscillator-pinn>

# Why does physics help?

- Regularization: Physics loss acts as a regularizer that restricts the hypothesis space to physically plausible solutions. Without it, the network can fit noise or learn spurious correlations.
- Extrapolation: Pure data-driven models interpolate well but extrapolate poorly. Physics constraints ensure the model respects governing equations even outside the training domain.



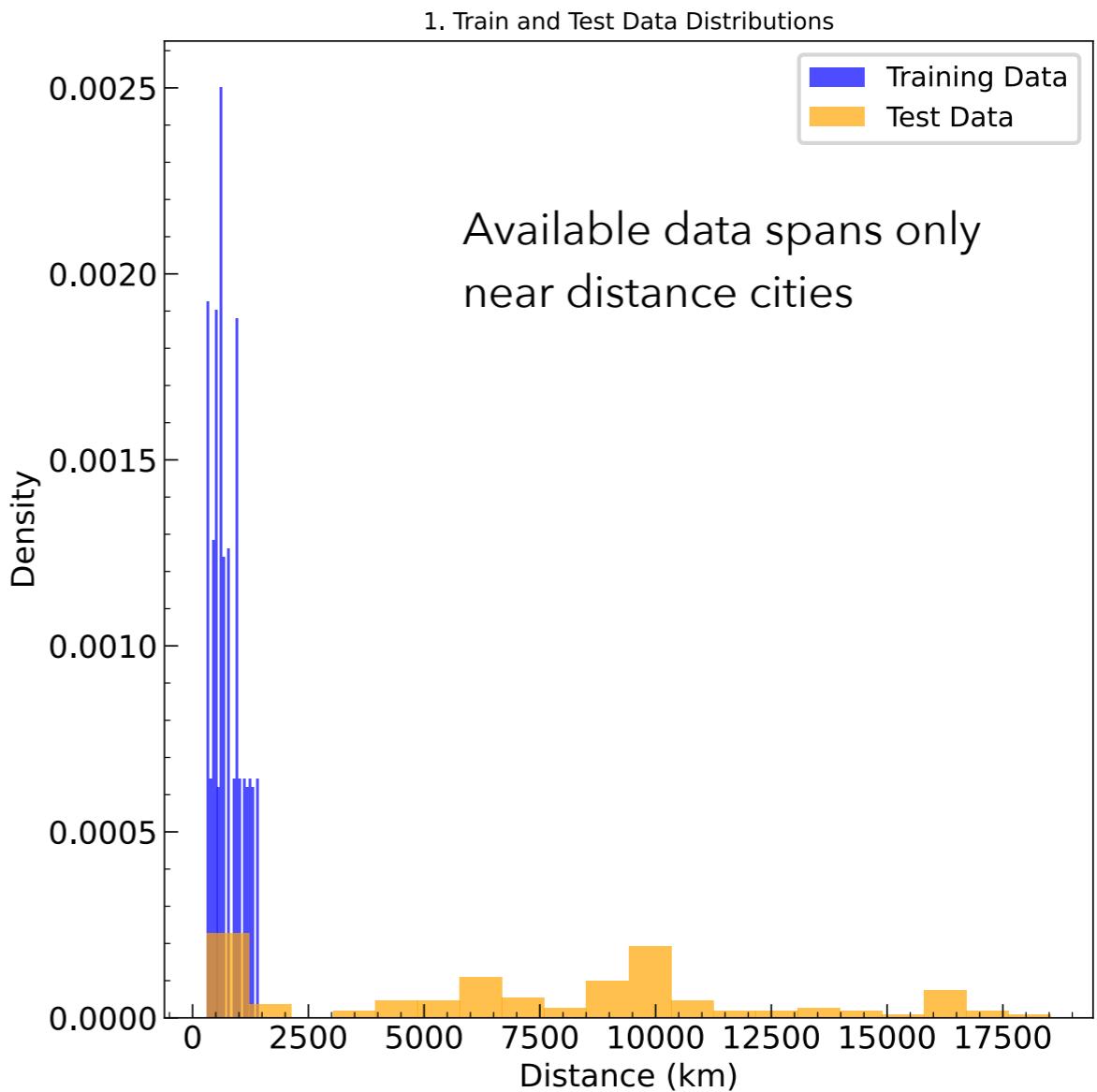
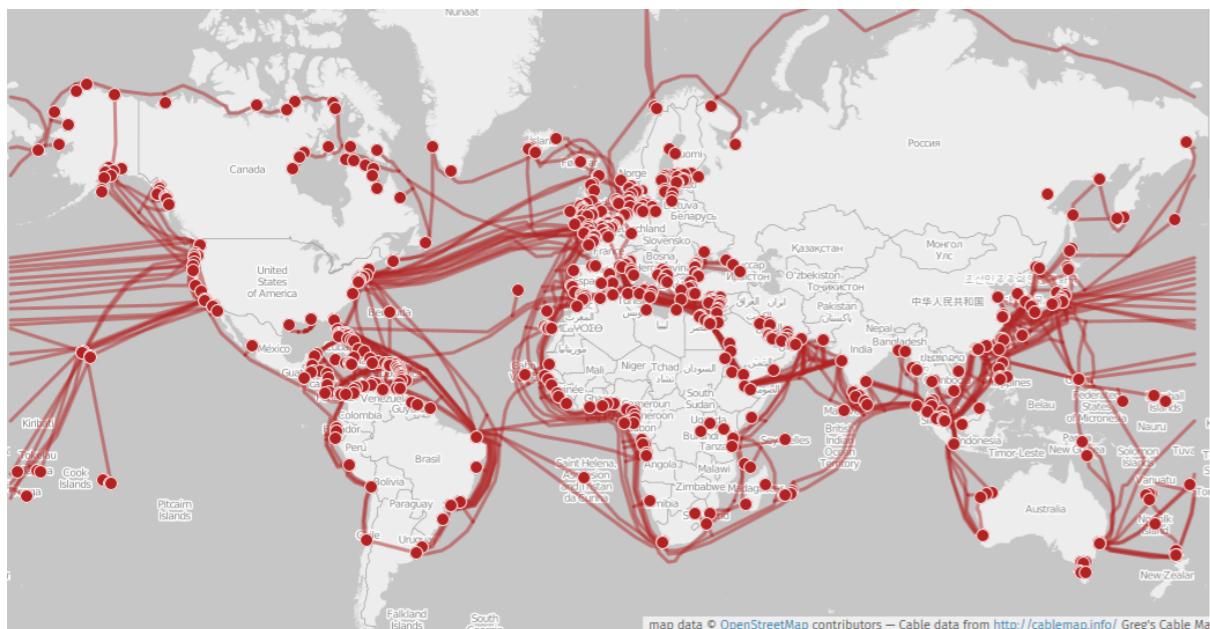
Simulation credit: <https://github.com/benmoseley/harmonic-oscillator-pinn>

---

Part - I  
Physics informed latency prediction  
Contrasting a learned model with a pure physics baseline

# Physics informed latency prediction

- Case: We've been given data of time taken for internet signals to reach from point A to point B. We know the geographical distance between the two cities but not the length of the fibre. We are supposed to flag anomalous network communication.



# Physics informed latency prediction

- Case: We've been given data of time taken for internet signals to reach from point A to point B. We know the geographical distance between the two cities but not the length of the fibre. We are supposed to flag anomalous network communication.

## Physics model (no ML)

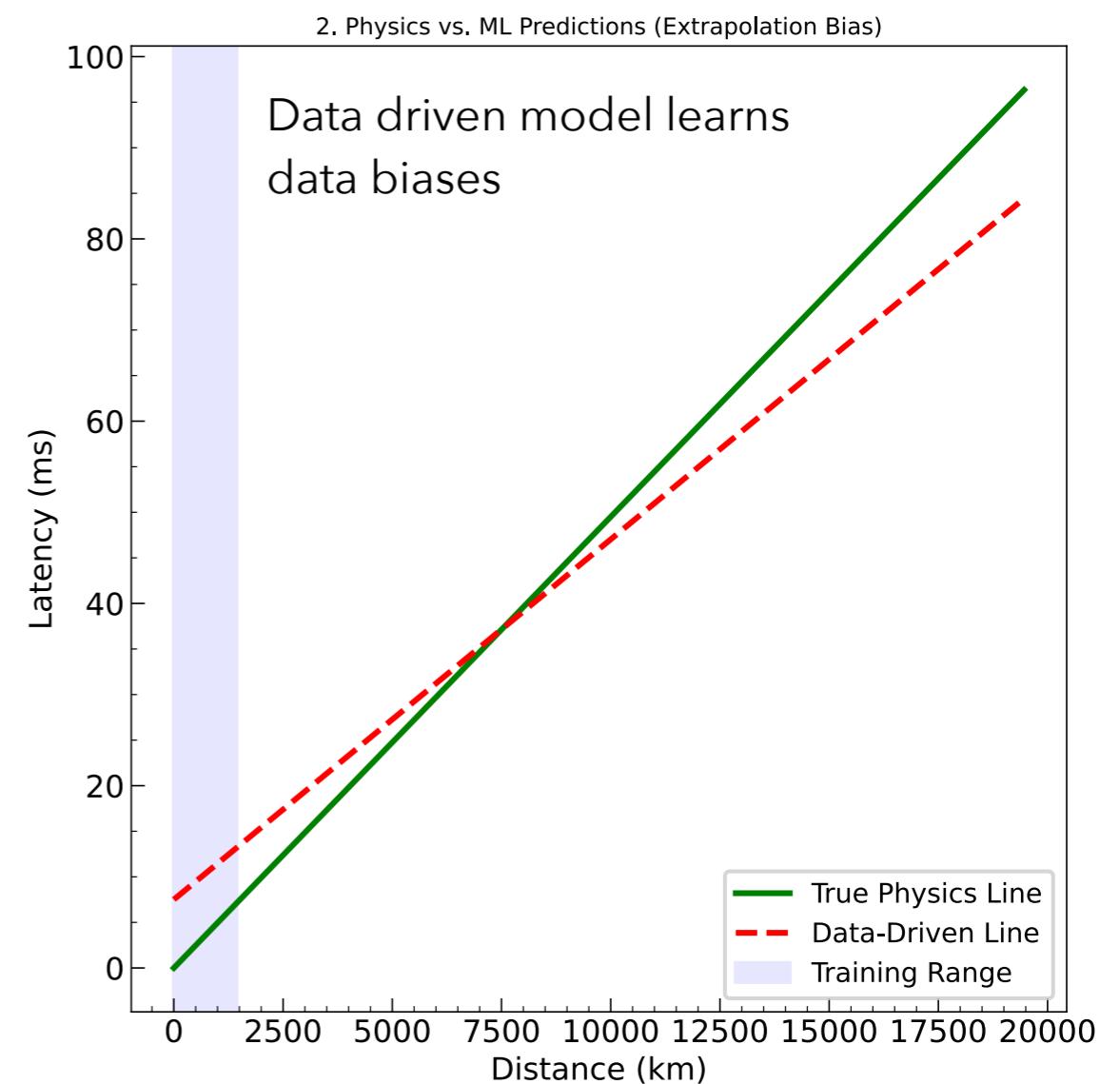
$$\frac{dx}{dt} = v \Rightarrow \text{latency}_{\text{pred}} = \frac{x_{\text{city}_1} - x_{\text{city}_2}}{\text{speed of light in fibre}}$$

If  $\text{latency}_{\text{pred}} < \text{latency}_{\text{measured}}$  flag anomaly

Find quantitative estimate of this inequality via threshold tuning of precision and recall

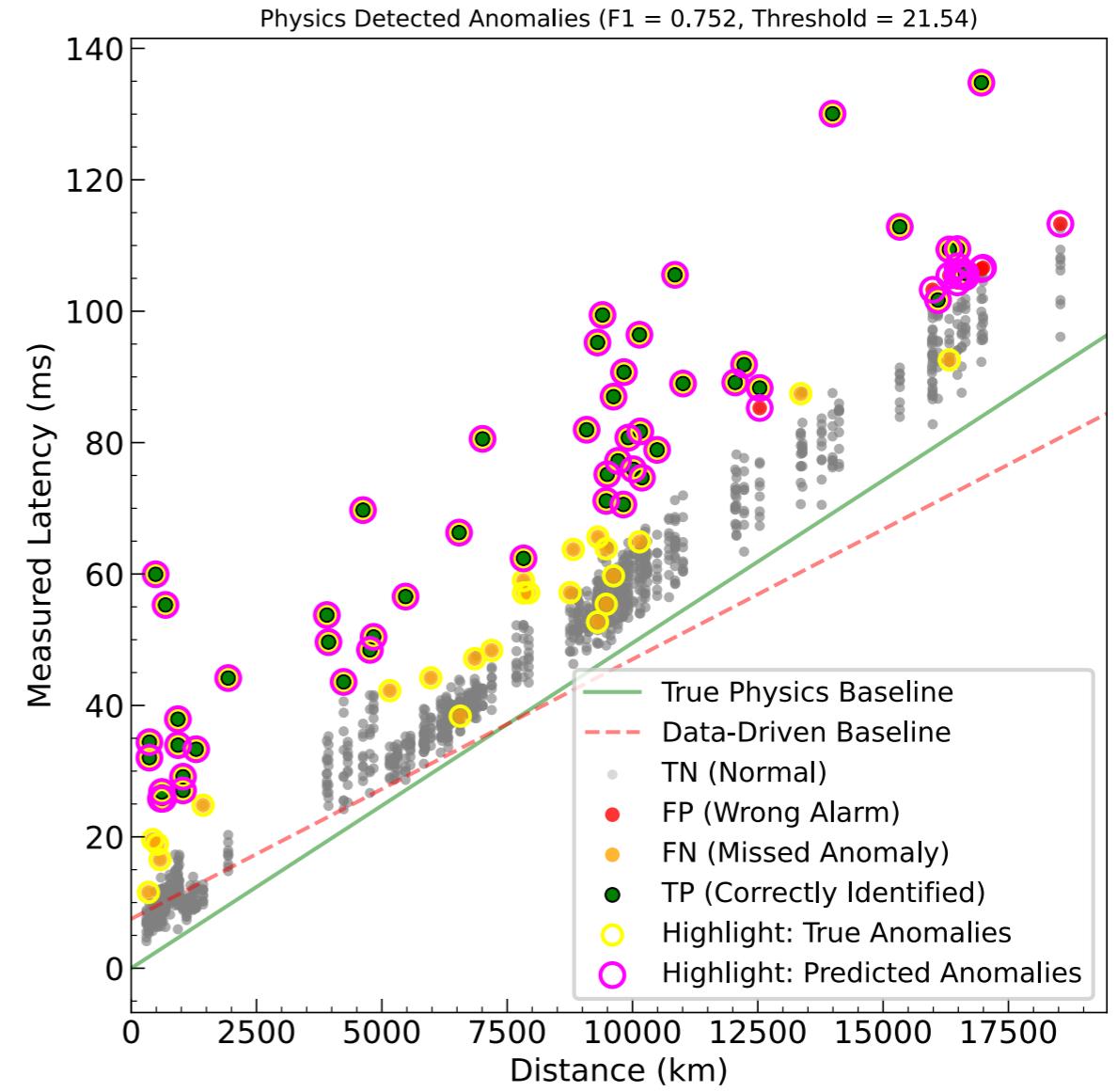
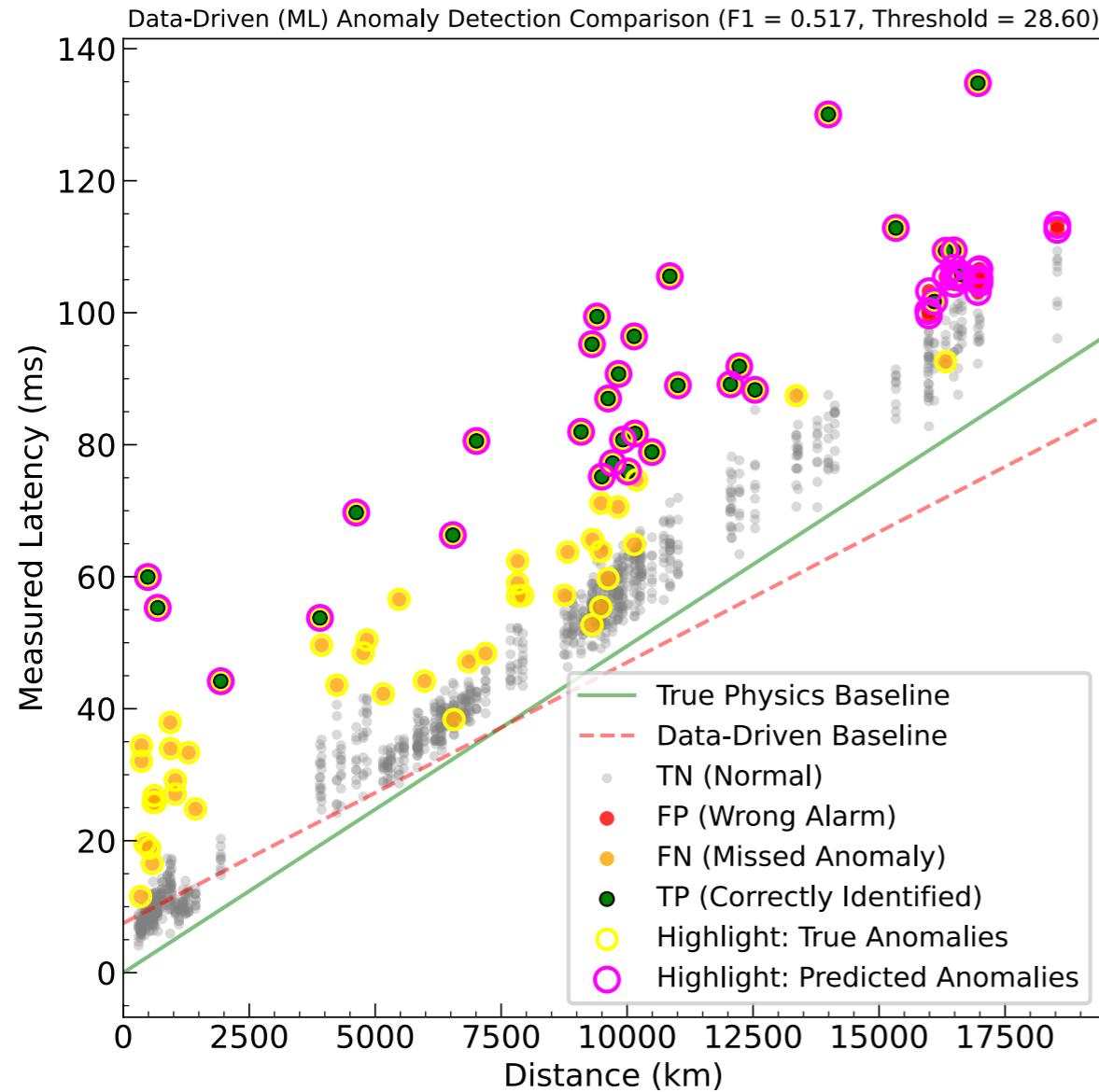
## ML model

Fit linear regression, find optimal anomaly detection threshold by tuning of precision and recall



# Physics informed latency prediction

- When generalised, physics driven model captures anomalies better



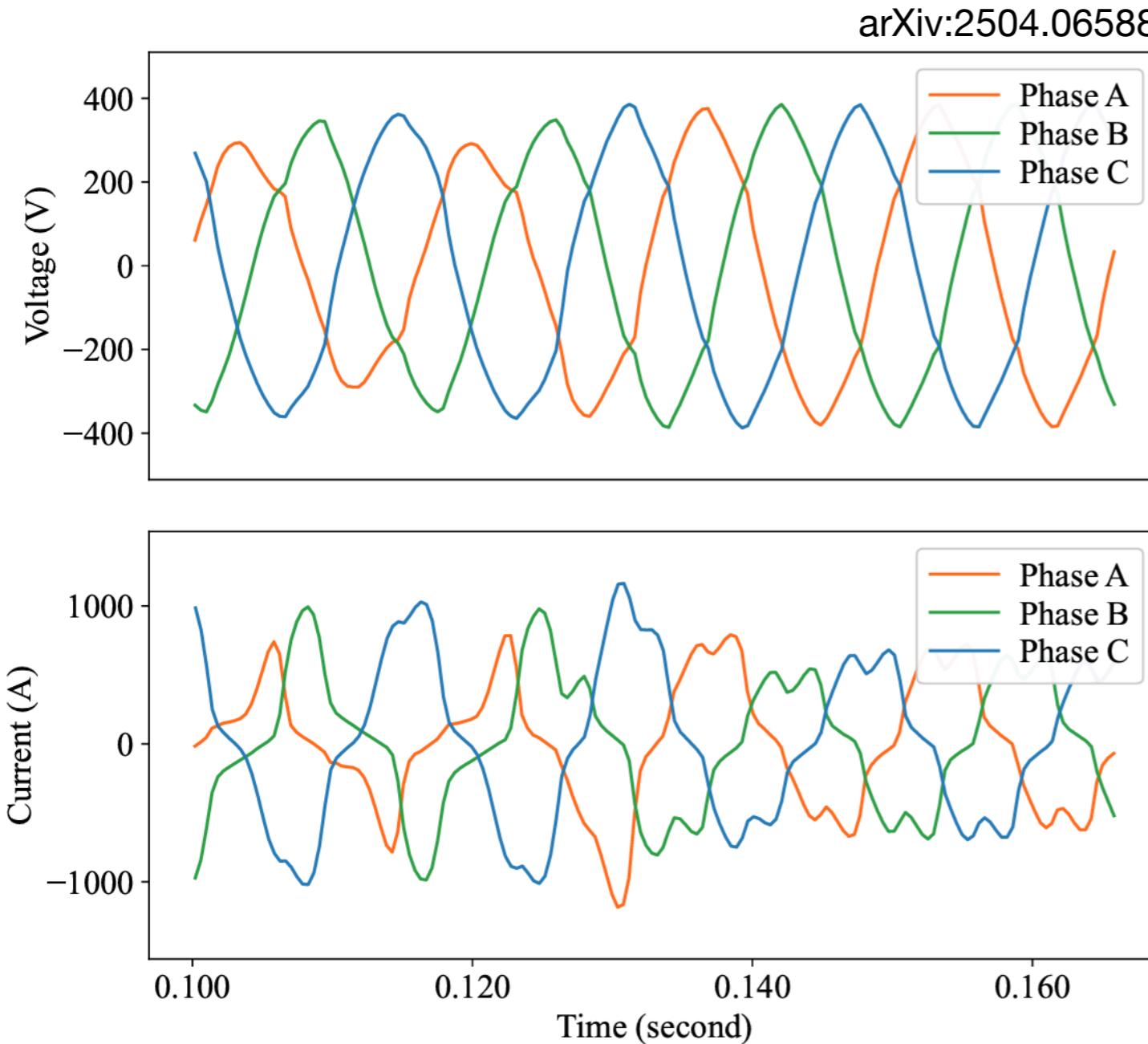
- Learn more at: <https://physics-informed-latency-pred.streamlit.app/>

---

Part - II  
Anomaly detection in time series data  
LSTM behaviour under simple harmonic dynamics

# Anomaly detection in simple harmonic oscillator

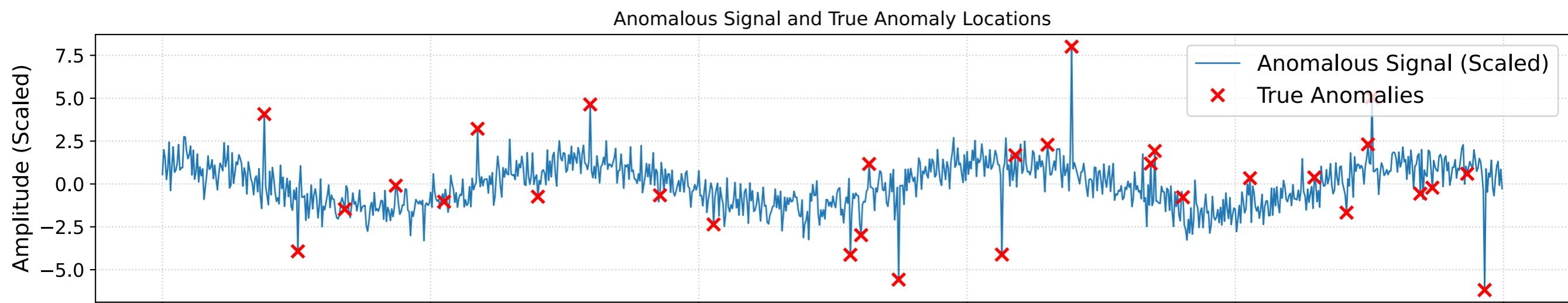
- Case: Many systems have oscillatory output
- Example: voltage and currents in power-grid



- Real-world synchrophasor dataset i.e. high resolution time synchronised power system measurements
- Three-phase voltage and current injection waveform measurement
- Shows phase imbalance, harmonic distortion, non-unitary power factor, and a transient event
- Also presents an important example of oscillatory systems

# Anomaly detection in simple harmonic oscillator

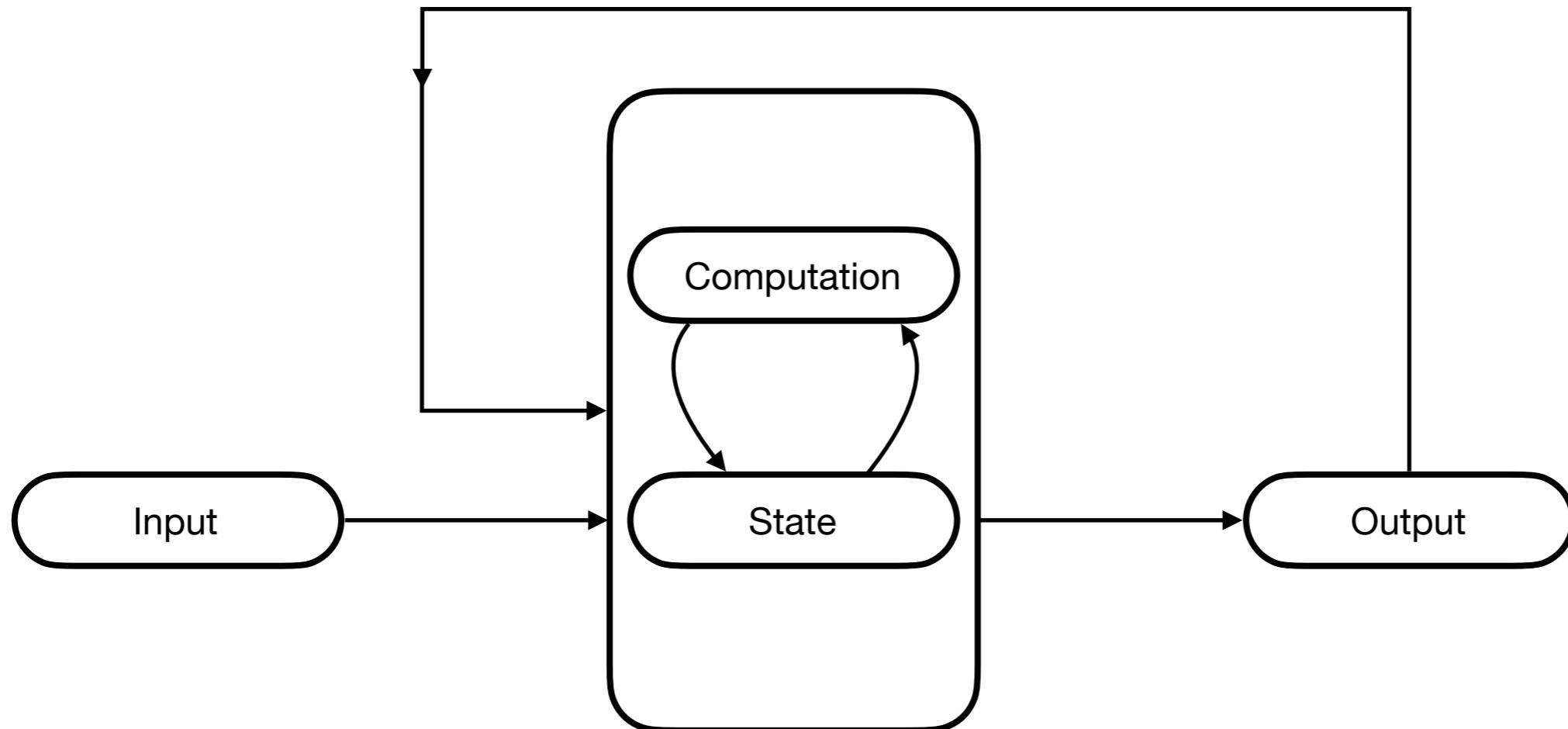
- Dataset
  - Train data: clean harmonic oscillator signal with known frequency, frequency governs the trajectory
  - Test data: noisy signal containing anomalies but obeying same frequency
- Our aim is to identify anomalies i.e. how many times does the data deviate from the ideal trajectory



- Given signal contains a fixed frequency and a number of anomalies of variable strength
  - Real life situation will contain noise, drift, unknown frequency etc.
  - These problems can be mitigated using additional processing, e.g., smoothing within windows, making frequency a learnable parameter or using fast-Fourier transform to estimate frequency etc. These will not be discussed here.

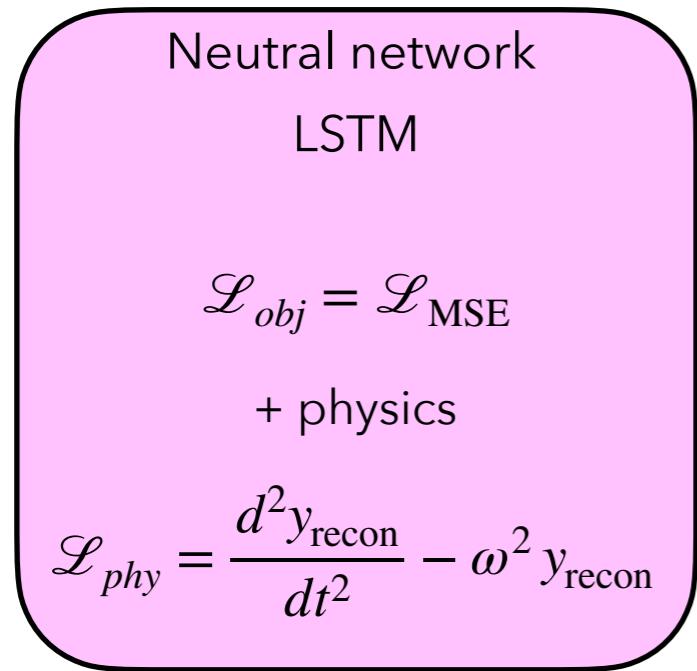
# Anomaly detection in simple harmonic oscillator

- Our aim is to identify anomalies i.e. how many times does the data deviate from the ideal trajectory
  - Use LSTM: type of recurrent neural network that allows long term memory relevant for time series data



# Loss functions

- Loss functions get extended by physics informed loss

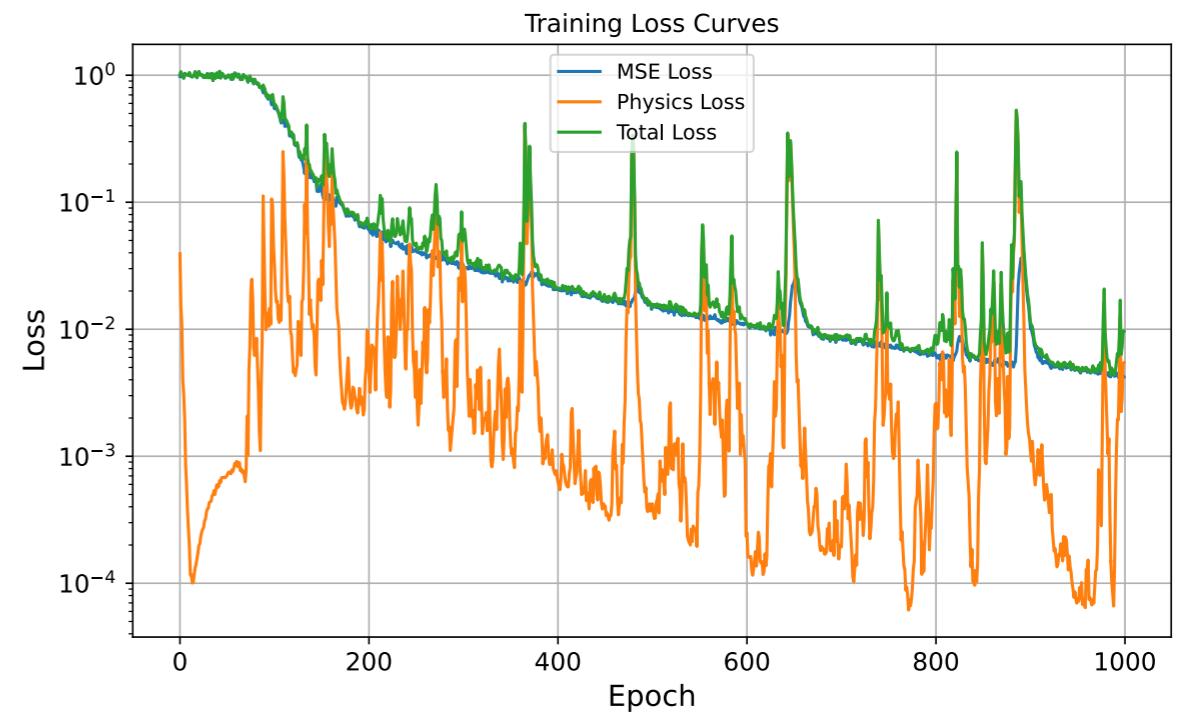
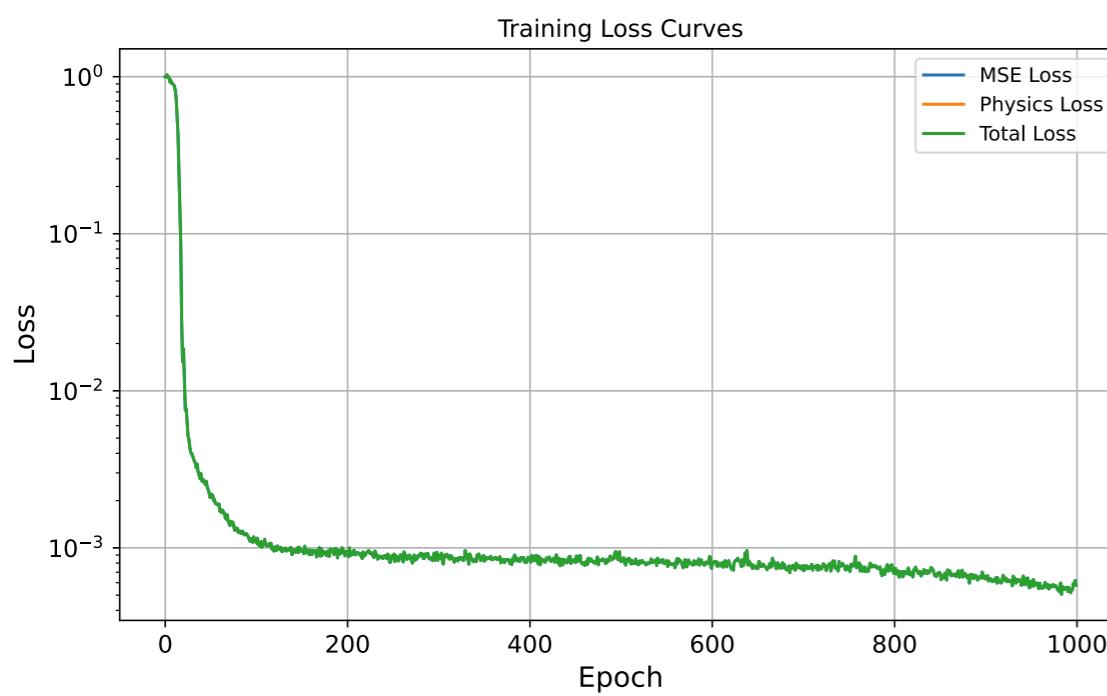


⇒

$$\mathcal{L}_{total} = \mathcal{L}_{obj} + \beta \times \mathcal{L}_{physics}$$

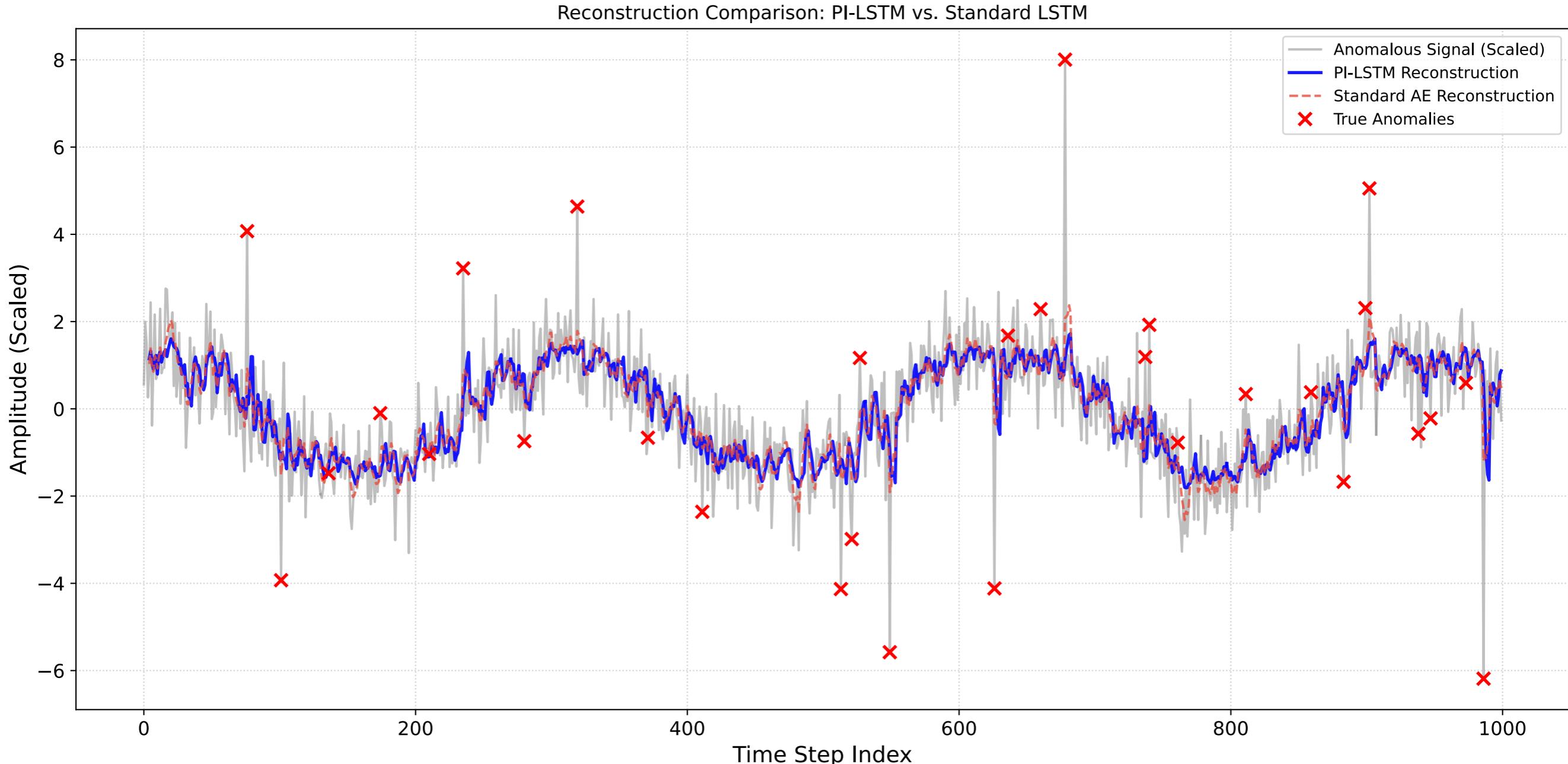
$\beta$ : additional hyper-parameter, needs tuning  
(for today, manual tuning)

Large  $\beta$ : physics loss dominates, trivial solutions  
Small  $\beta$  too little physics importance



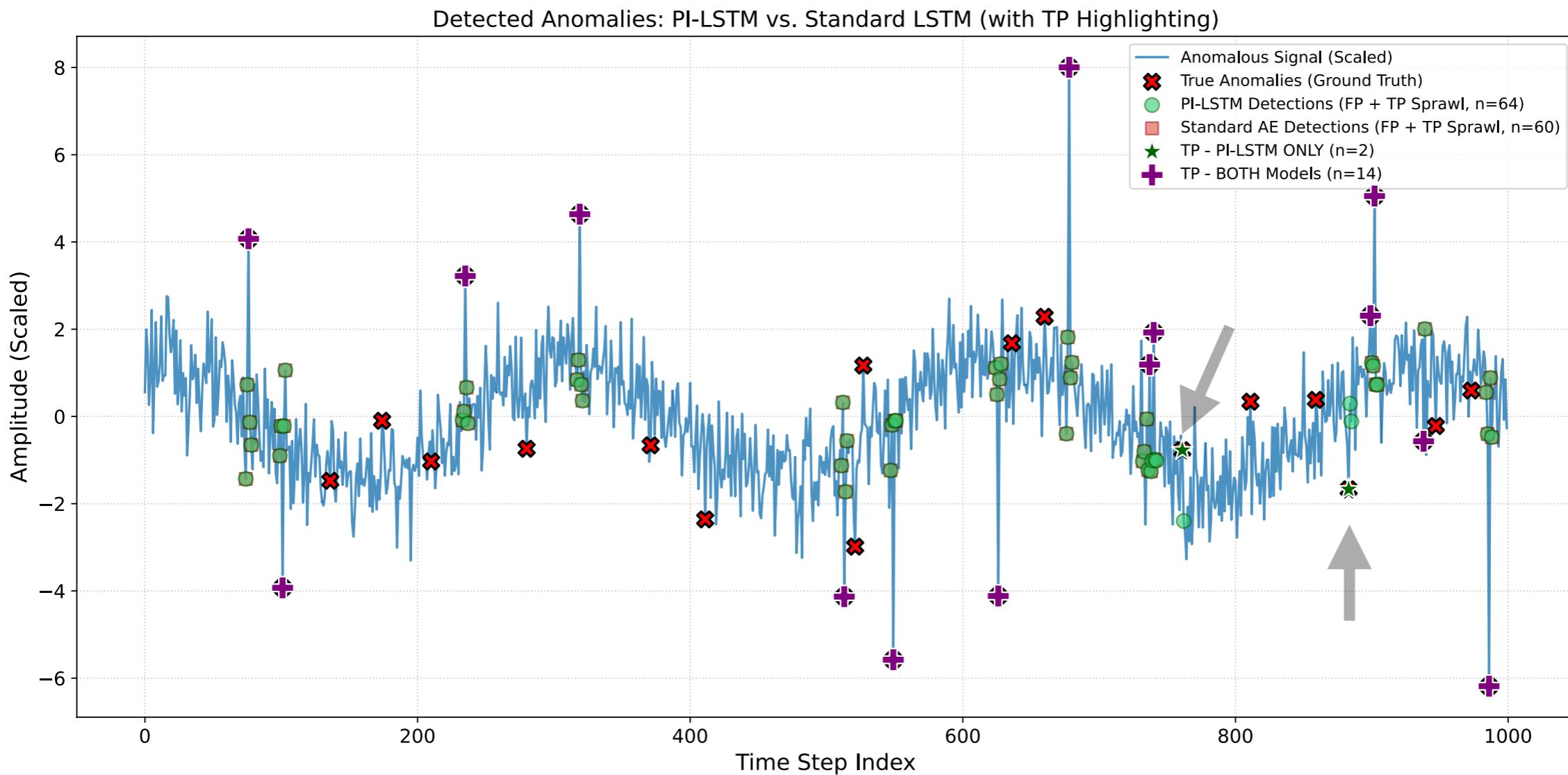
# Anomaly detection in simple harmonic oscillator

- Trained on clean simple harmonic oscillator signal, tested on noisy signal containing anomalies



- Standard LSTM introduces distortions in order to recreate test signal

# Anomaly detection in simple harmonic oscillator



Model	Unique Anomalies Detected (TP)	Total Windows Flagged	F1-Threshold
PINN ( $\beta = 0.01$ )	16	64	1.41
Standard AE ( $\beta = 0.00$ )	14	60	1.23

14% more anomalies  
and 15% higher  
detection threshold  
with PINN

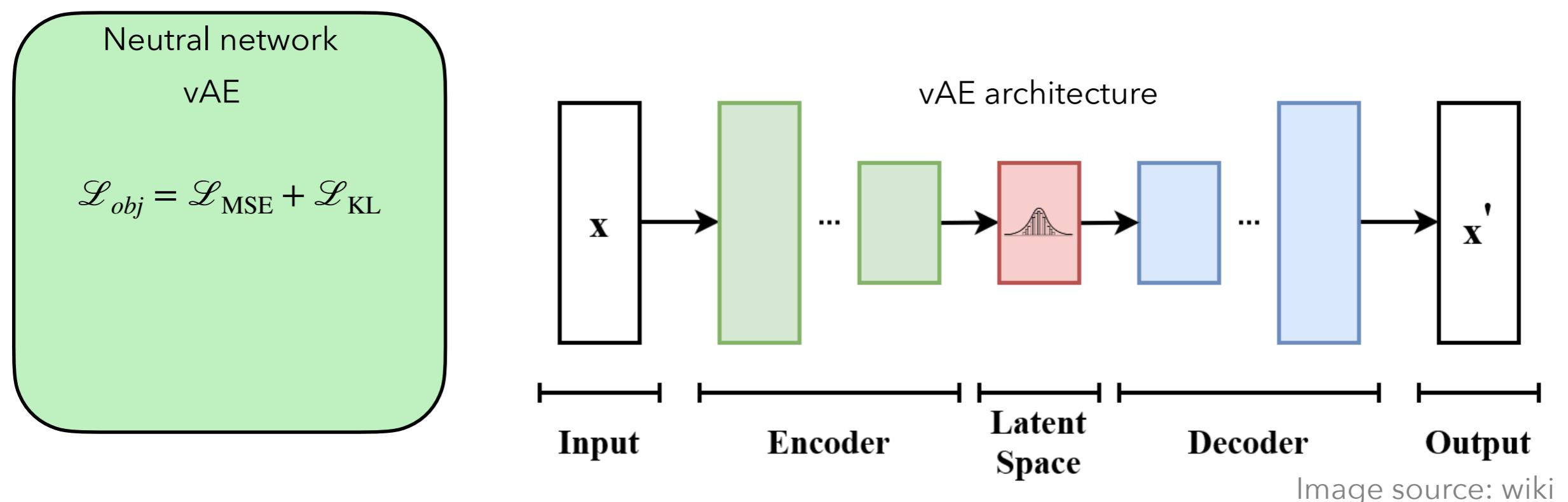
- Learn more at: <https://pinnlearning.streamlit.app/>

---

Part - III  
Generative modelling with structure  
VAEs that learn to produce physically valid signals

# Signal generation capabilities

- Case: Can we combine generative AI with physics information to create more accurate and controlled simulation?
- Aim: Demonstrate with a simple harmonic oscillator
- Use VAE: provides an interpretable latent space, that allows generative capabilities and physics integration



$\beta$ : additional hyper-parameter, needs tuning

Large  $\beta$ : physics loss dominates, trivial solutions

Small  $\beta$  too little physics importance

# Signal generation capabilities

- Case: Can we combine generative AI with physics information to create more accurate and controlled simulation?
- Aim: Demonstrate with a simple harmonic oscillator
- Use VAE: provides an interpretable latent space, that allows generative capabilities and physics integration

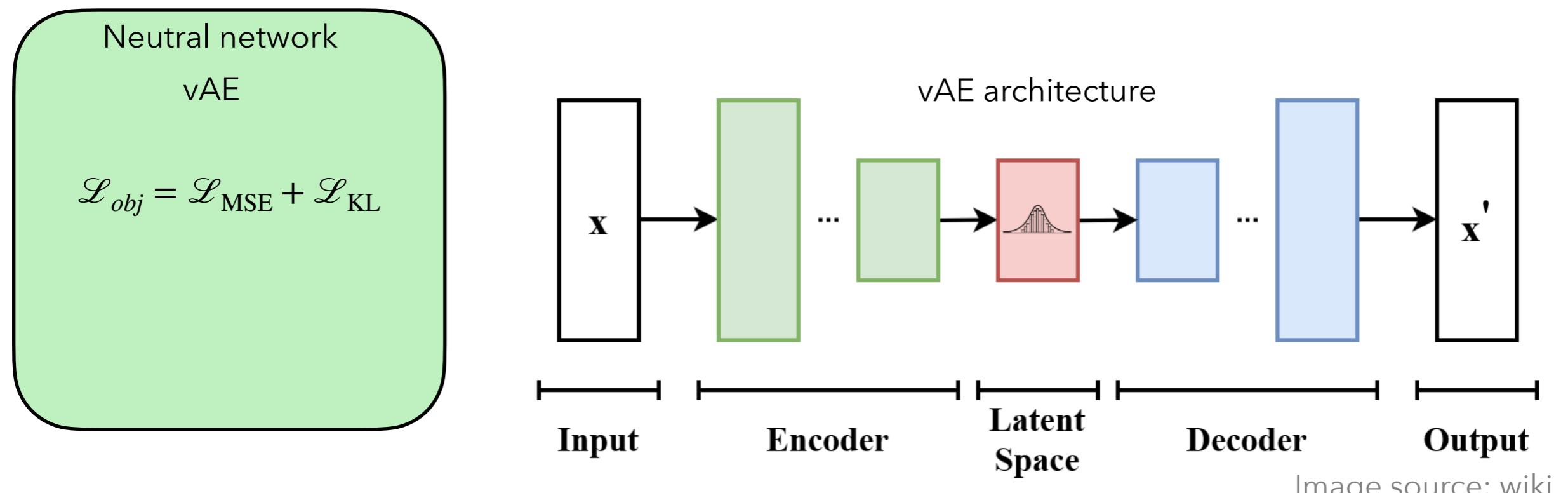


Image source: wiki

$$\Rightarrow \quad \mathcal{L}_{total} = \mathcal{L}_{obj} + \beta \times \mathcal{L}_{physics}$$

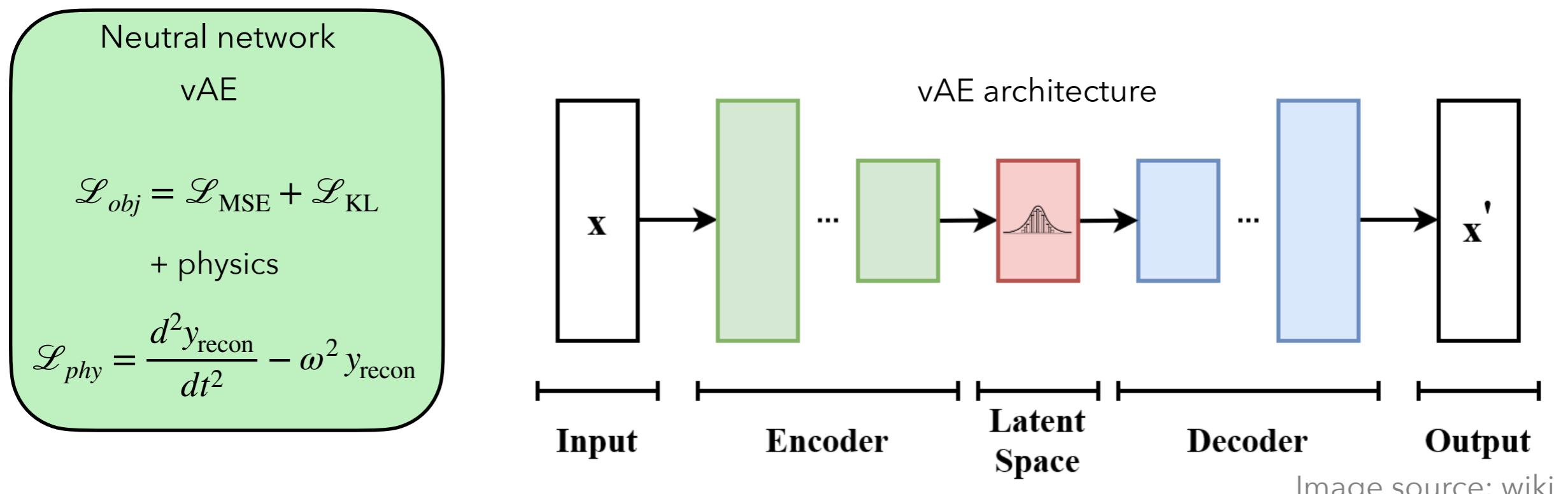
$\beta$ : additional hyper-parameter, needs tuning

Large  $\beta$ : physics loss dominates, trivial solutions

Small  $\beta$  too little physics importance

# Signal generation capabilities

- Case: Can we combine generative AI with physics information to create more accurate and controlled simulation?
- Aim: Demonstrate with a simple harmonic oscillator
- Use VAE: provides an interpretable latent space, that allows generative capabilities and physics integration



$\Rightarrow$

$$\mathcal{L}_{total} = \mathcal{L}_{obj} + \beta \times \mathcal{L}_{physics}$$

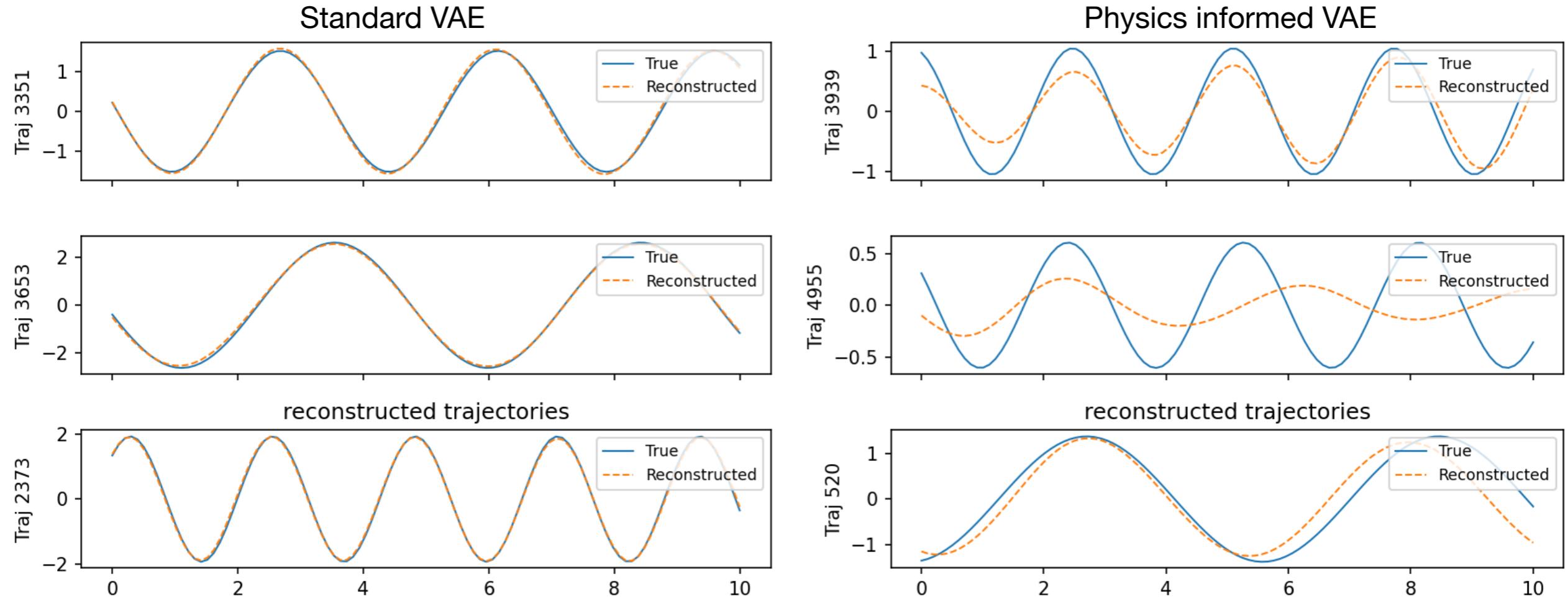
$\beta$ : additional hyper-parameter, needs tuning

Large  $\beta$ : physics loss dominates, trivial solutions

Small  $\beta$  too little physics importance

# Training a physics informed VAE

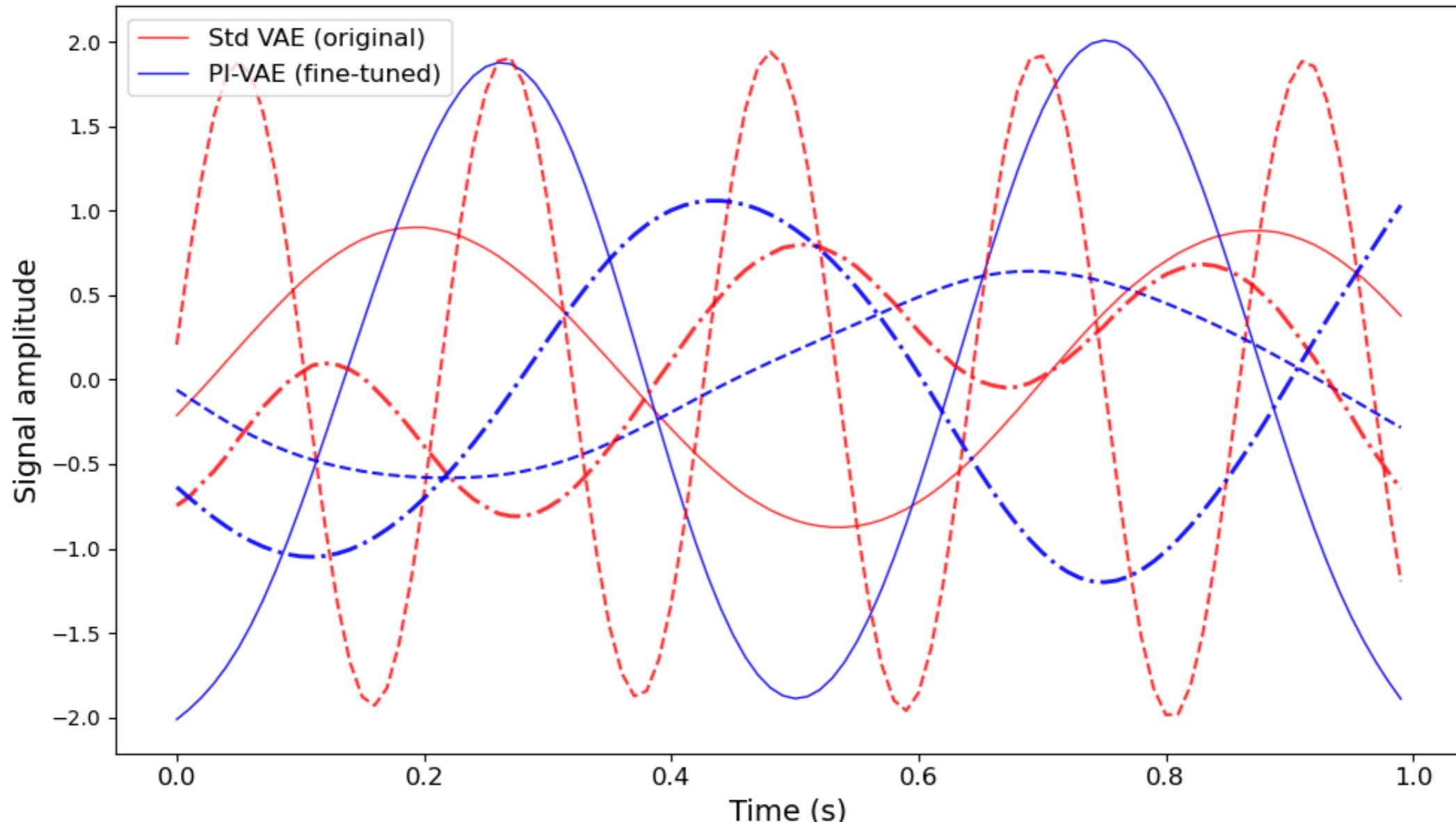
- Train on a clean simple harmonic oscillator signal of variable frequency, amplitude and phase
- Aim: use trained model to generate new simple harmonic oscillator signals



- The physics informed VAE appears to have worse reconstruction performance
- Sacrifice on reconstruction during training for physics consistency, can be improved by fine-tuning hyper-parameter  $\beta$ , here no fine-tuned  $\beta$ : chosen value 'good enough' not 'the best'

# Signal generation

- Small fine-tuning (minimise physics loss only) of PI-VAE post training as  $\beta$  was not optimised
- Procedure: Sample  $z \sim \mathcal{N}(0, I)$   $\rightarrow$  Signal  $x = \text{Decoder}(z)$
- Standard VAE doesn't do very good with signal generation



- PI-VAE has a more stable performance in signal generation

# Conclusions

- Physics constraints = better extrapolation and data efficiency when equations are known
- Trade-off: beta tuning matters, wrong physics hurts more than helps
- PIML works best for: sparse data, known governing equations, need for reliability

<https://qr.link/4khP9z>



- Happy to discuss and connect

