Automatic recognition of anomalous patterns in discharges by recurrent neural networks

Gonzalo Farias[1]
Ernesto Fabregas[2]
Sebastián Dormido-Canto[2]
Jesús Vega[3]
Sebastián Vergara[1]

May 13-17 2019 / Daejeon, KOREA

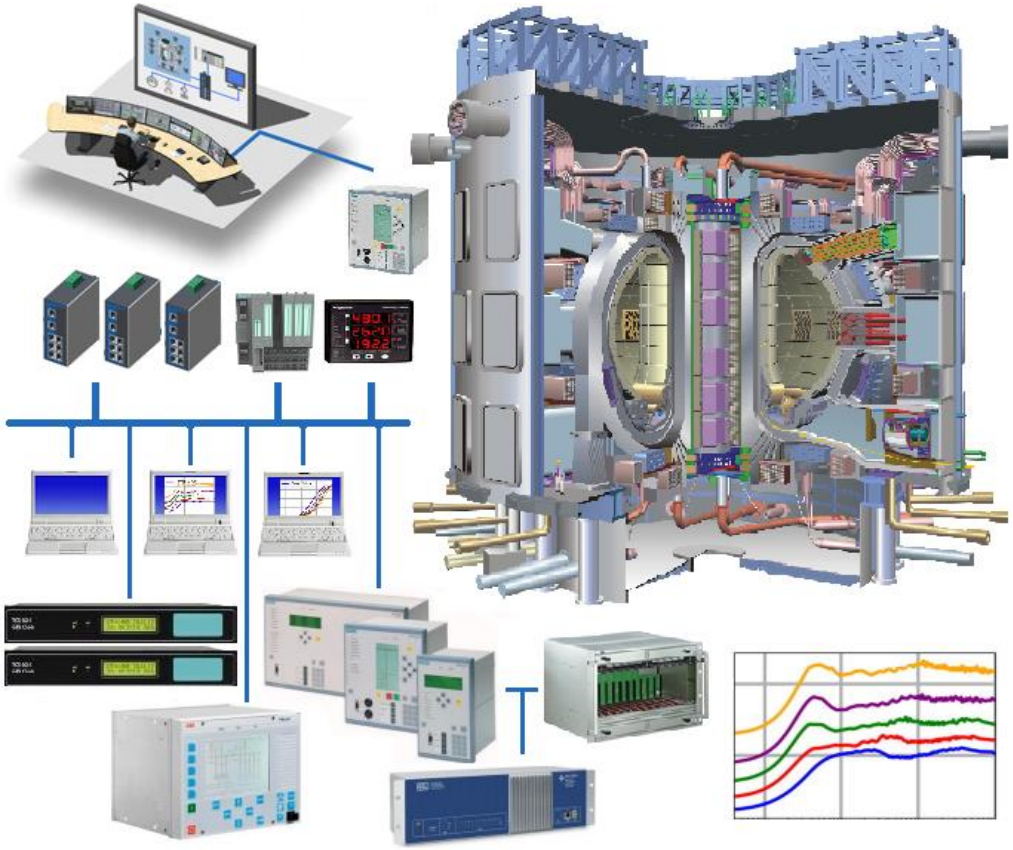
1 PONTIFICIA UNIVERSIDAD CATOLICA DE VALPARAISO
2 UNED
3 GOBIERNO DE ESPAÑA MINISTERIO DE ECONOMÍA Y COMPETITIVIDAD
Ciemat Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas

Outline

- **Introduction**
- Background
  - Anomaly Detection
- Proposed Solution
  - Recurrent Neural Networks (LSTM)
- Results
- Summary

Introduction

- The experiments generate huge quantities of data. It is estimated that only 10% of this data is analyzed.



A shot of few seconds can generate huge quantity of data:

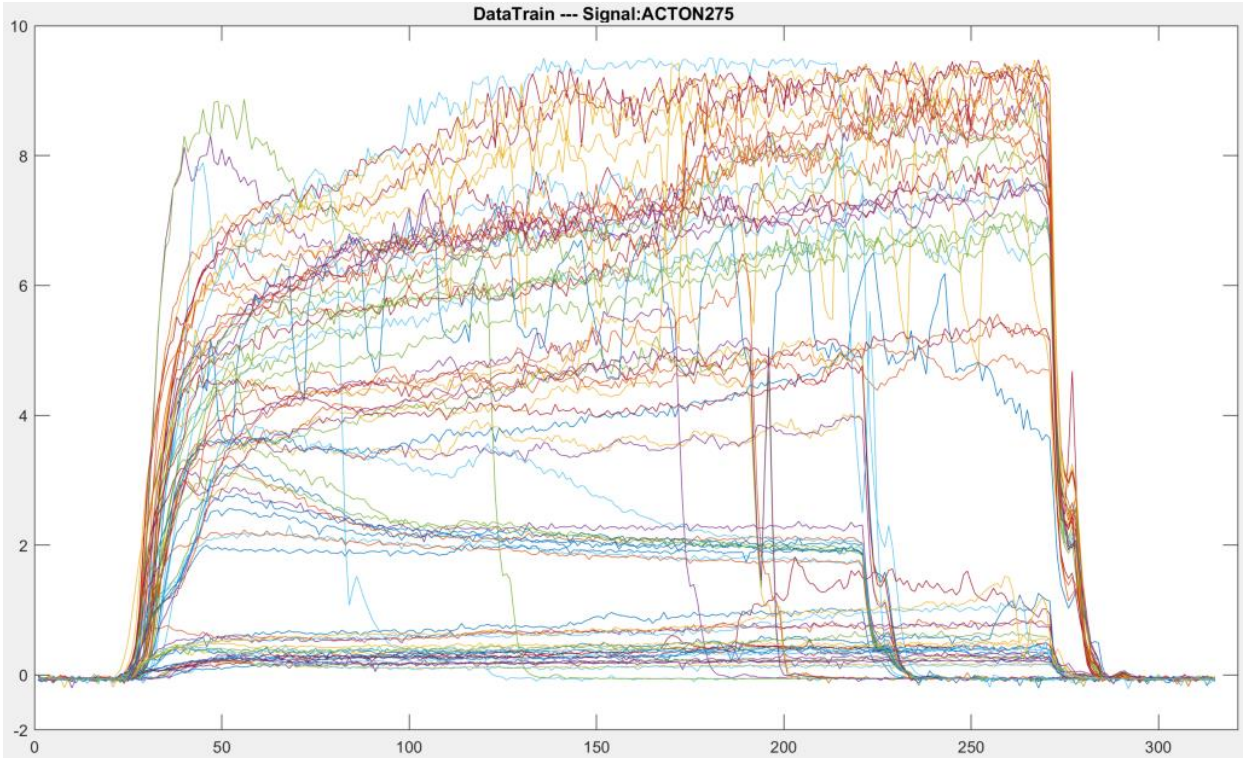
- **TJ-II** device has +1000 channels of measurements.
- A shot in **JET** can take around 10 seconds (**10 GB/shot**. around 100 TB/year).
- **ITER** could generate **1 TB/shot**. around 1 PB/year.

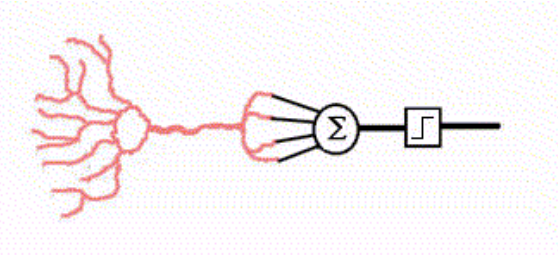
Outline

- Introduction
- **Background**
  - Anomalies
- Proposed Solution
  - Recurrent Neural Networks (LSTM)
  - Anomaly Detection
- Results
- Summary

Background

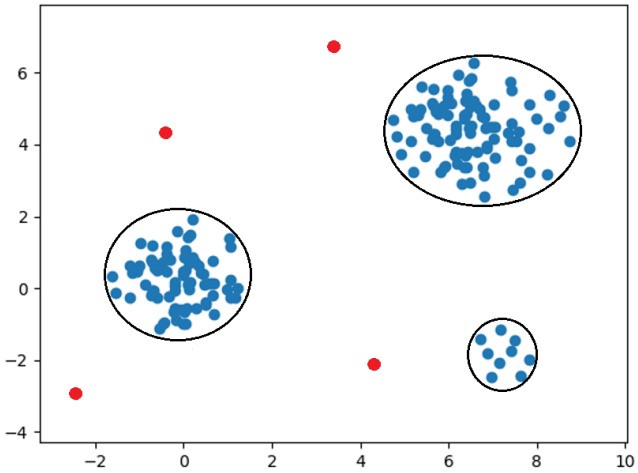
- The idea is to use Artificial Intelligence to deal with fusion data.
- Create systems that allow specialists to analyze and interpret data more quickly and efficiently than manually.



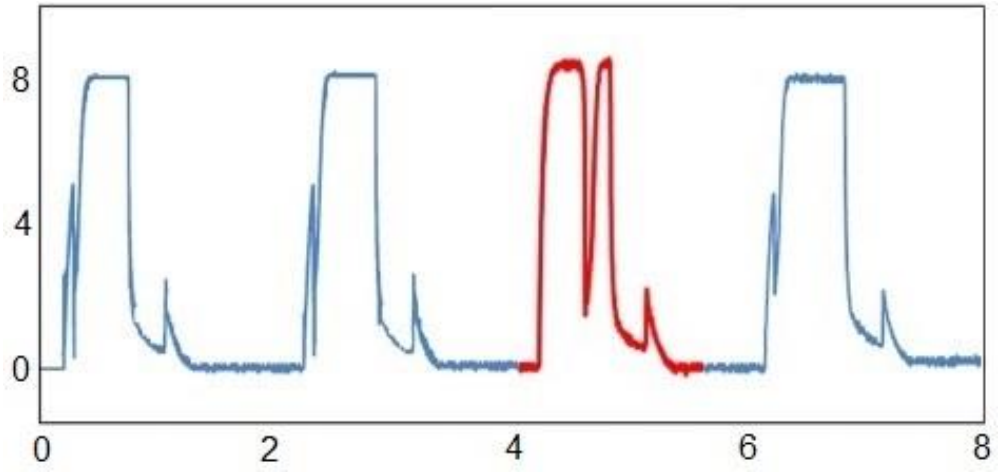


Background – Anomalies

- **Anomaly**: Something that deviates from what is standard, normal, or expected.
- One type of anomaly is known as '**outlier**', which is a value located outside of the normal class.
- Other type of anomaly is an anomalous behavior, which is a **periodic collapsing phenomenon in time series**.



Clustering

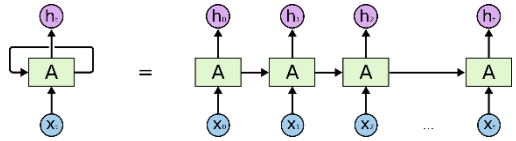


Pattern recognition

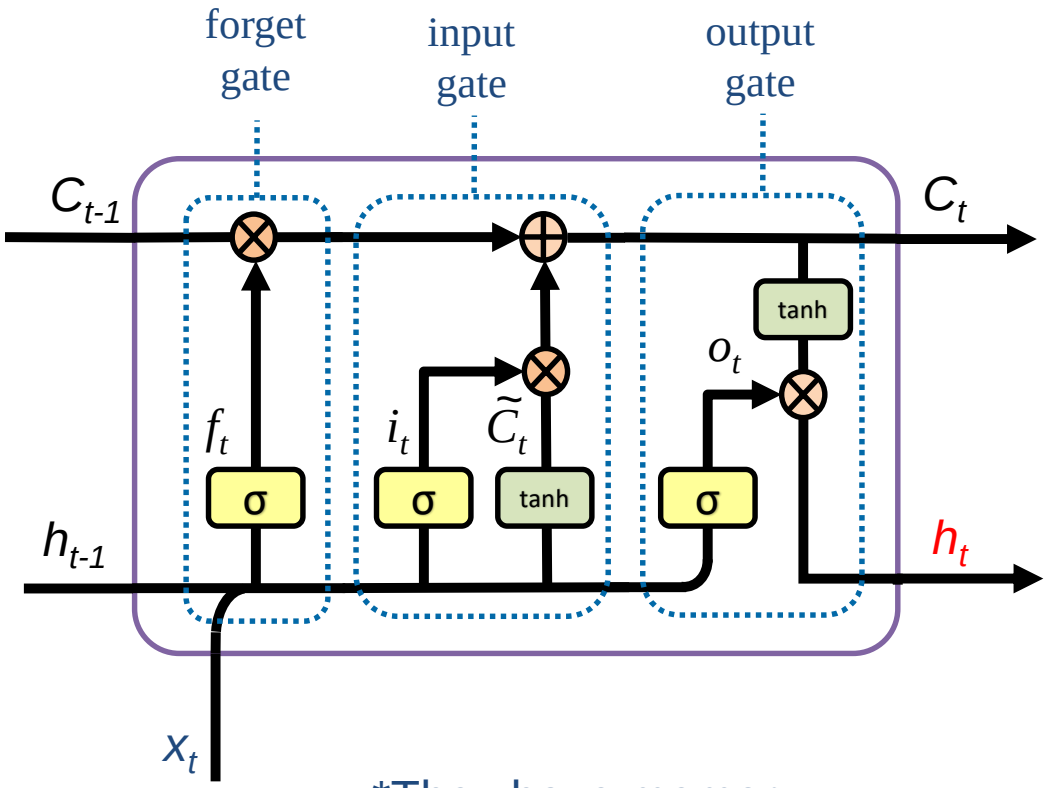
Outline

- Introduction
- Background
  - Anomalies
- **Proposed Solution**
  - Recurrent Neural Networks (LSTM)
  - Anomaly Detection
- Results
- Summary

Proposed Solution – LSTM



- Recurrent Neural Network – Long Short Term Memory (LSTM)



*They have memory

Forget gate layer

$$f_t = \sigma\left(W_f\,[h_{t-1}, x_t] + b_f\right)$$

Input gate layer

$$i_t = \sigma\left(W_i\,[h_{t-1}, x_t] + b_i\right)$$

$$\tilde{C}_t = tanh\left(W_c\,[h_{t-1}, x_t] + b_c\right)$$

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t$$

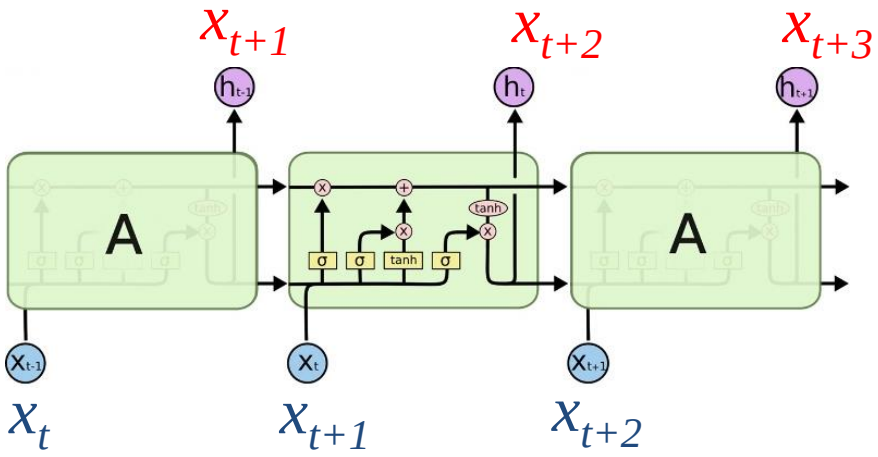
Output gate layer

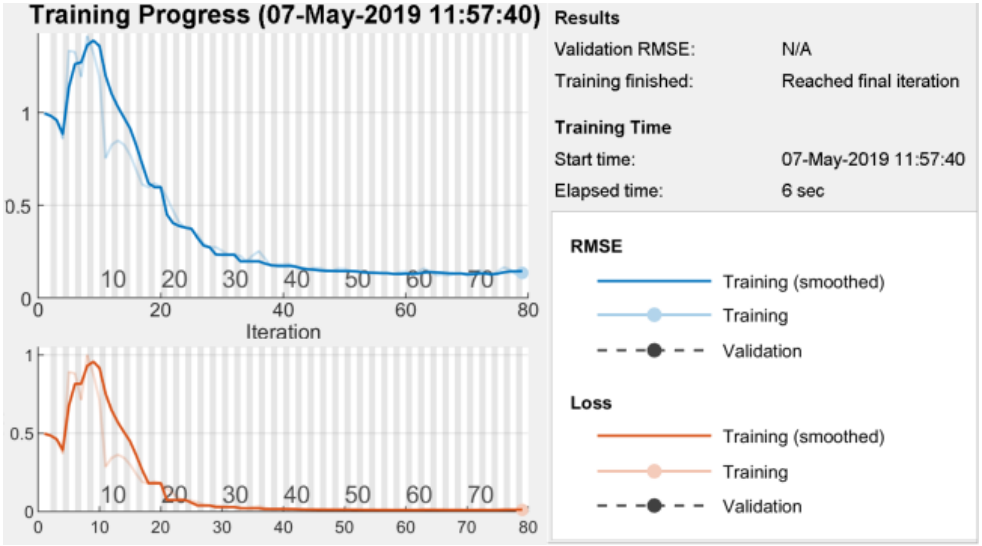
$$o_t = \sigma\left(W_o\,[h_{t-1}, x_t] + b_o\right)$$

$$h_t = o_t * tanh\left(C_t\right)$$

* https://colah.github.io/posts/2015-08-Understanding-LSTMs/

LSTM – Forecasting (training stage)





Forget gate layer

$$f_t = \sigma\left(\boldsymbol{W_f}[h_{t-1}, x_t] + \boldsymbol{b_f}\right)$$

Input gate layer

$$i_t = \sigma\left(\boldsymbol{W_i}[h_{t-1}, x_t] + \boldsymbol{b_i}\right)$$

$$\tilde{C}_t = tanh\left(\boldsymbol{W_c}[h_{t-1}, x_t] + \boldsymbol{b_c}\right)$$

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t$$

Output gate layer

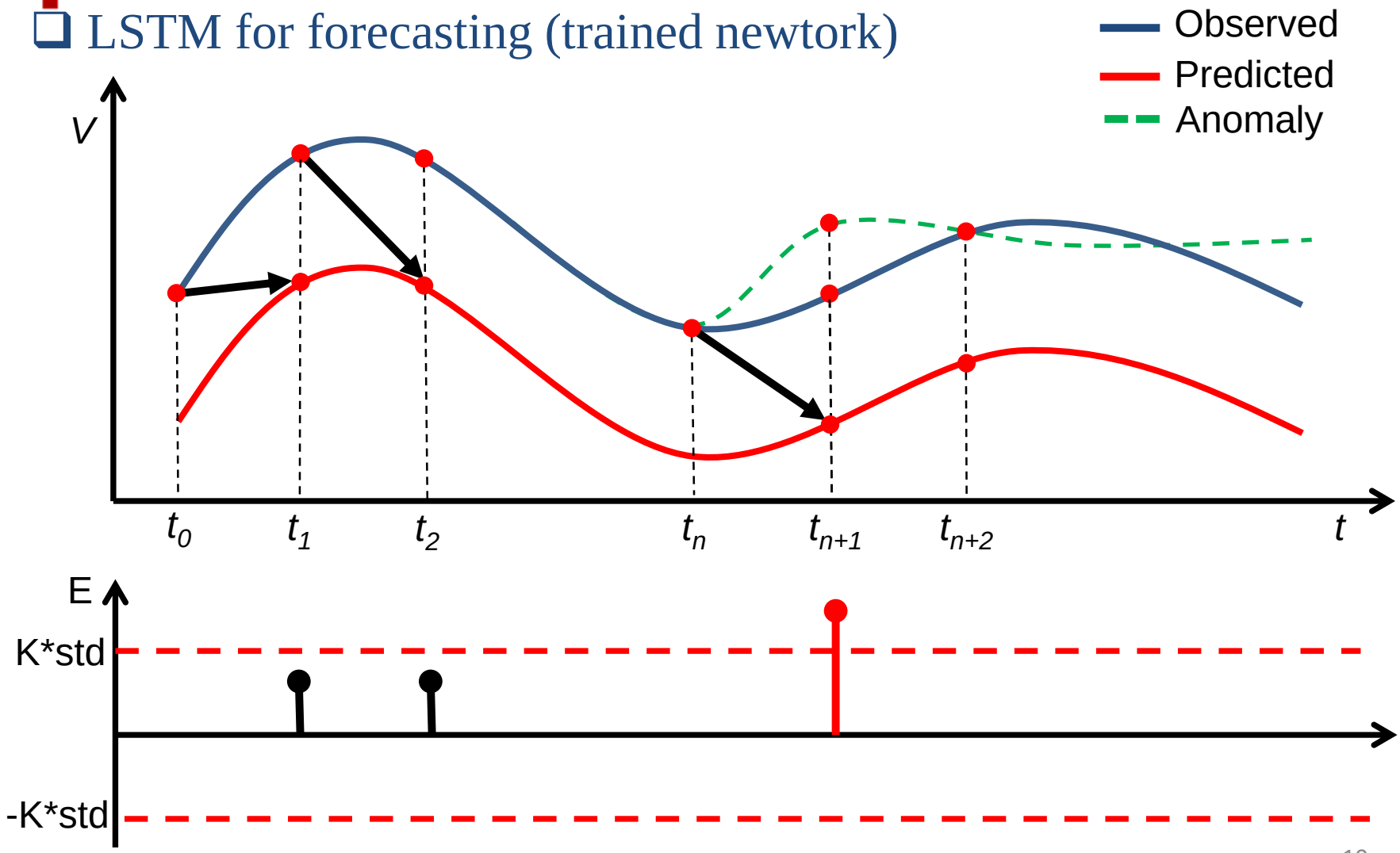
$$o_t = \sigma\left(\boldsymbol{W_o}[h_{t-1}, x_t] + \boldsymbol{b_o}\right)$$

$$h_t = o_t * tanh\left(C_t\right)$$

*It adjusts the bias and weights to learn the shape of the waveform

LSTM – Forecasting (test stage)

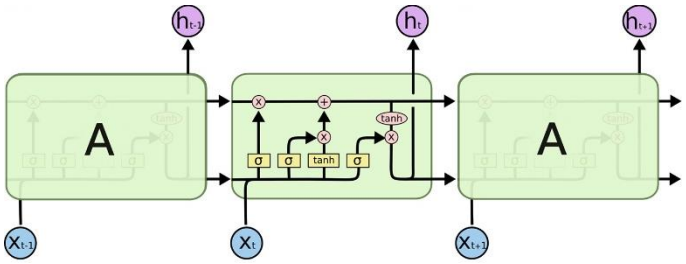
- LSTM for forecasting (trained newtork)

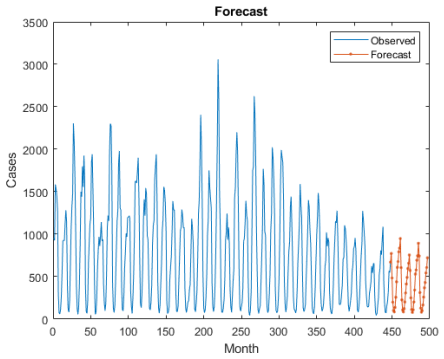


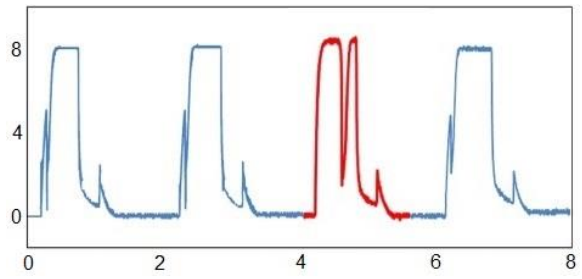
Goals

- **General Goal**
  - Anomaly detection using Recurrent Neural Network (LSTM - Long Short Term Memory).

- **Specific Goal**
  - The LSTM Neural Network learns the waveform to detect anomalies through forecasting.



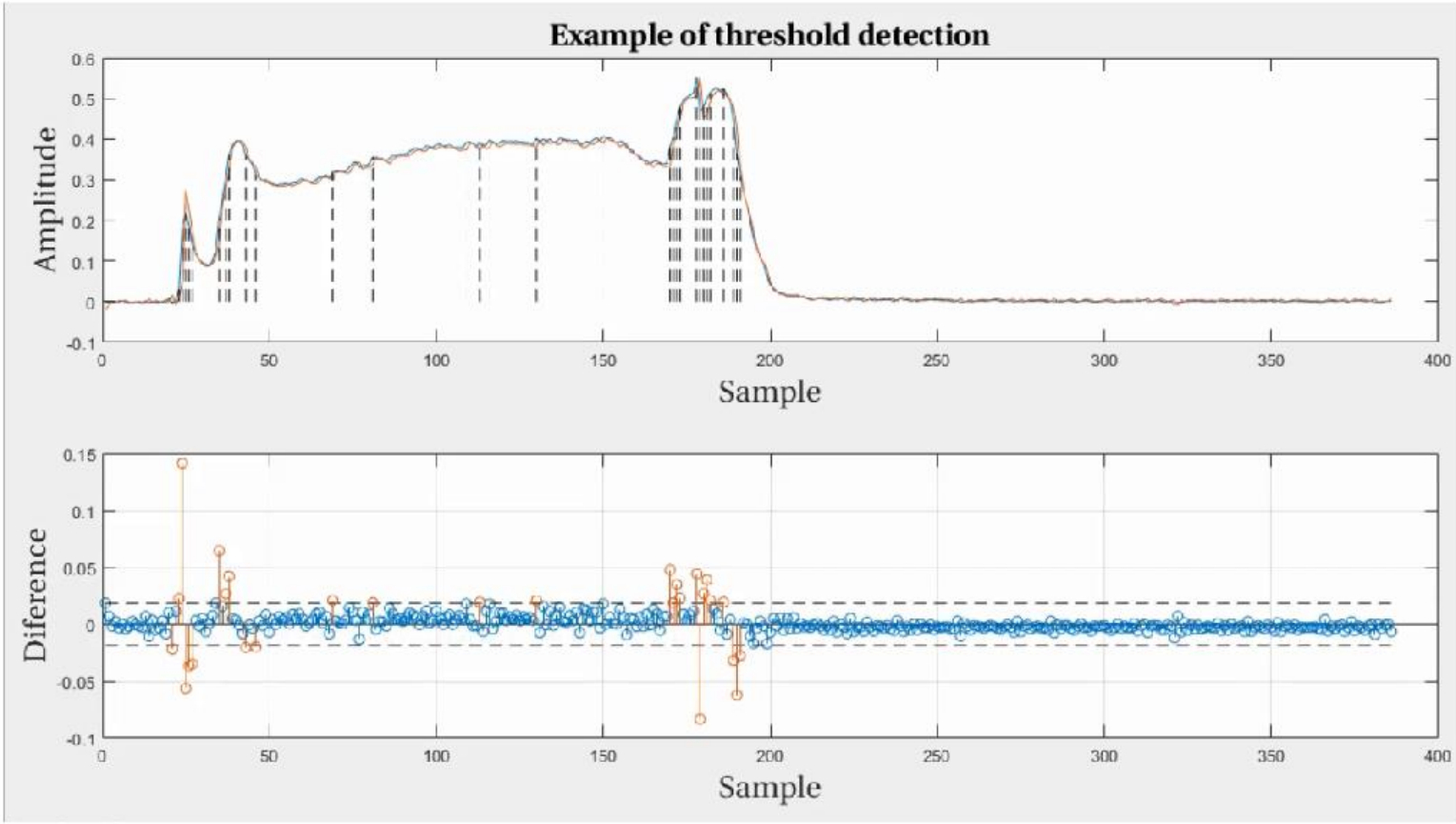




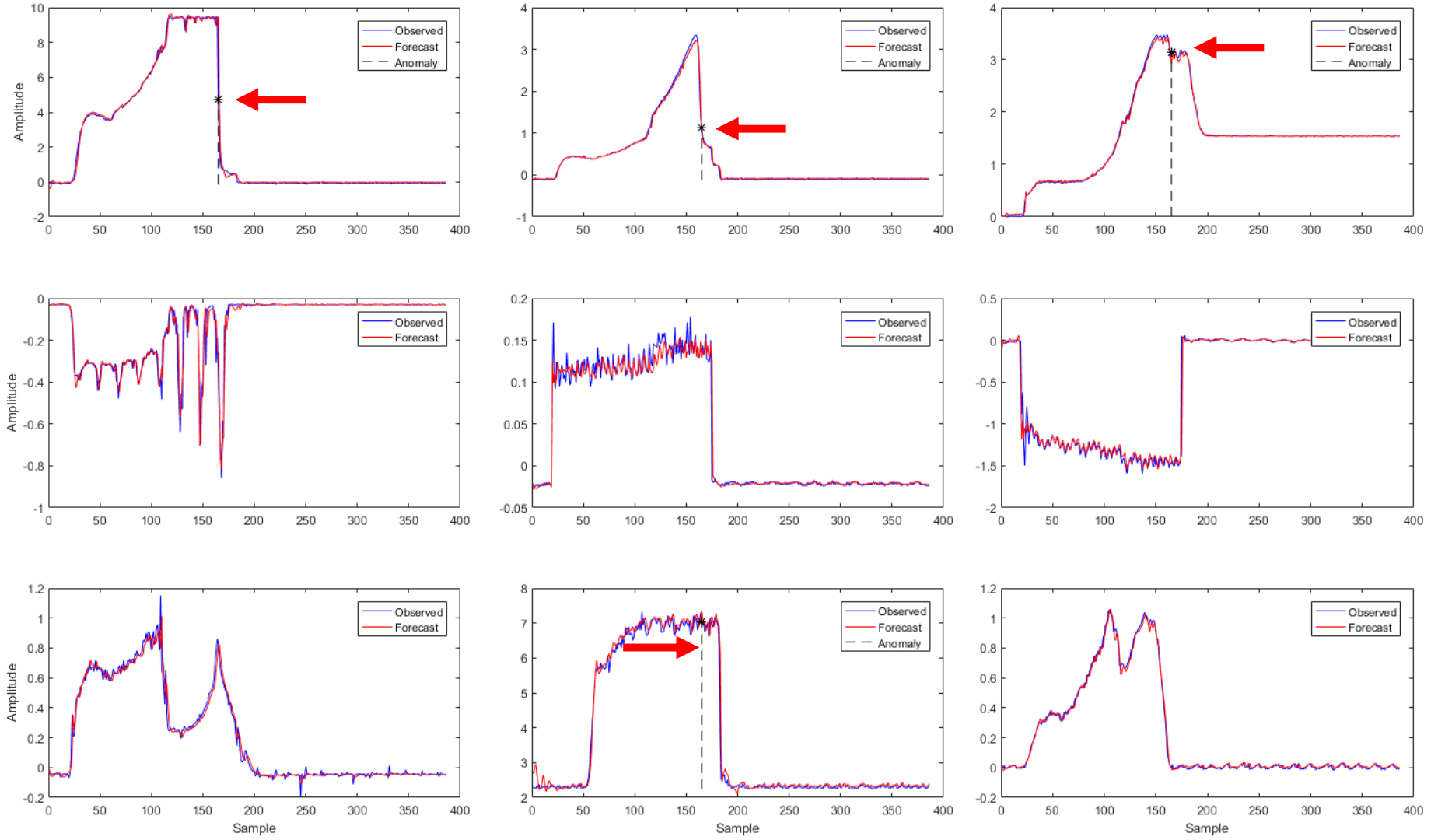
* https://colah.github.io/posts/2015-08-Understanding-LSTMs/

Anomaly Detection – Threshold (th=k*std)

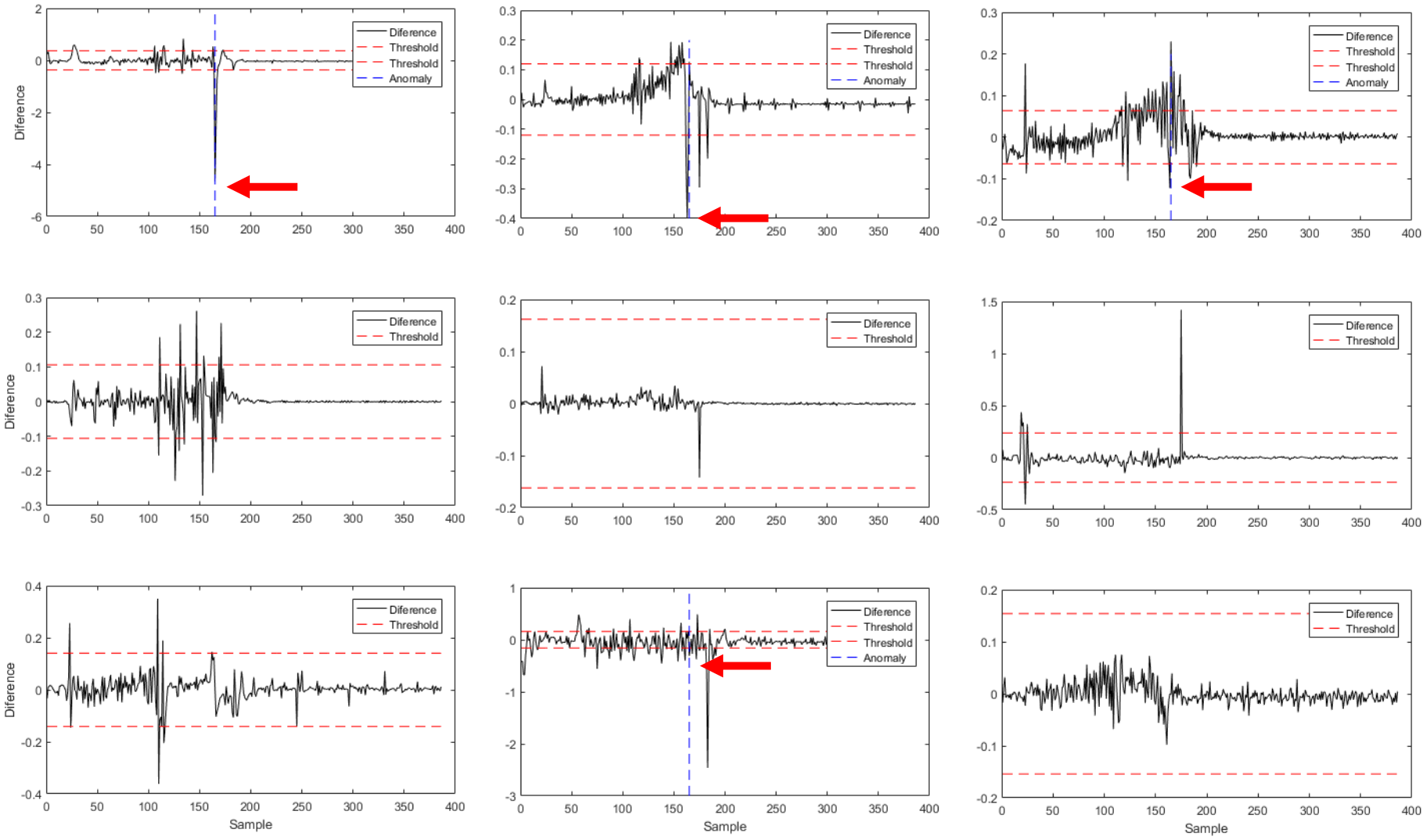
- **How the Anomaly is detected?**
  - We fix a **threshold** proportional to the **Standard Deviation** of the **Error**.



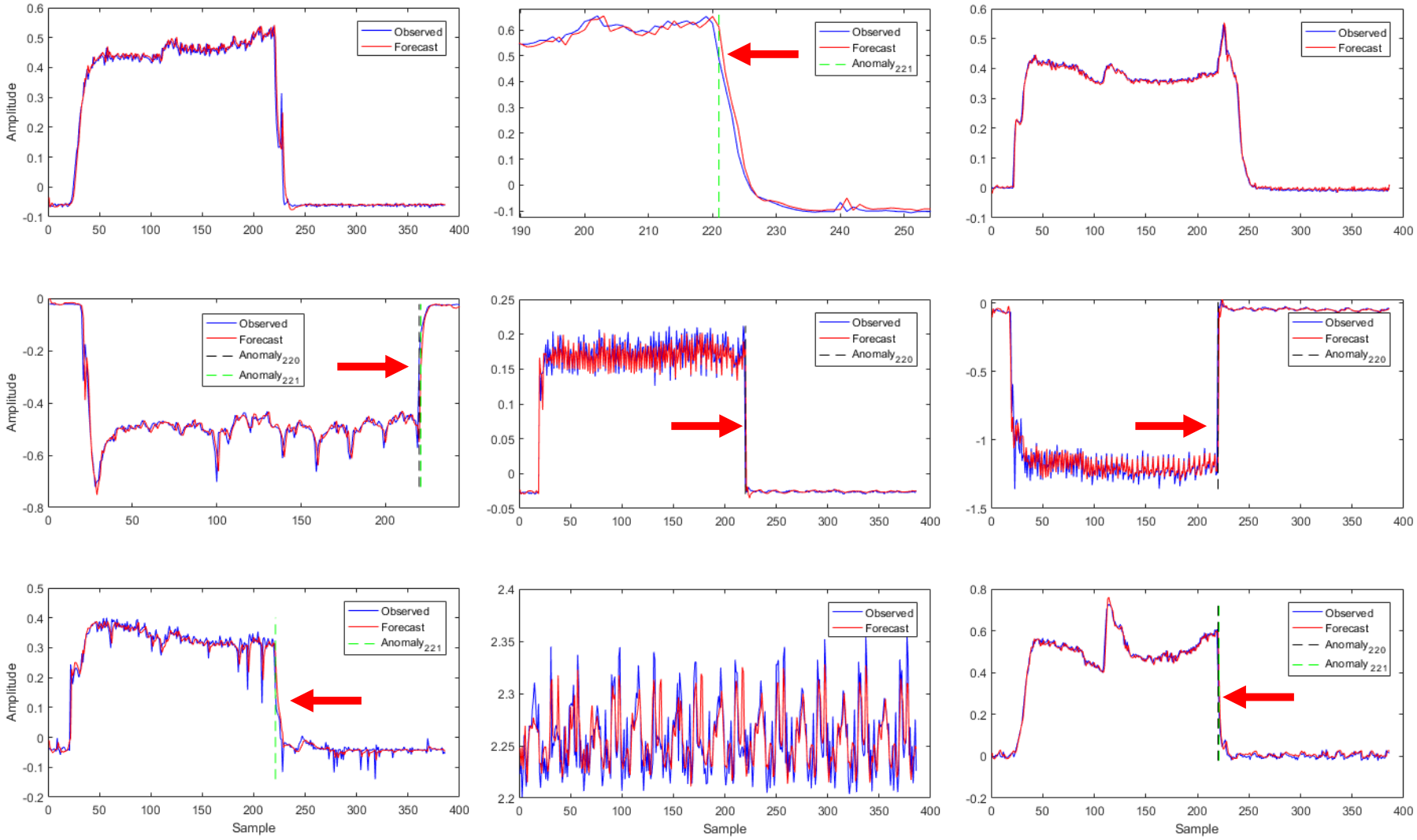
Anomaly Detection – Simultaneous (t=165)



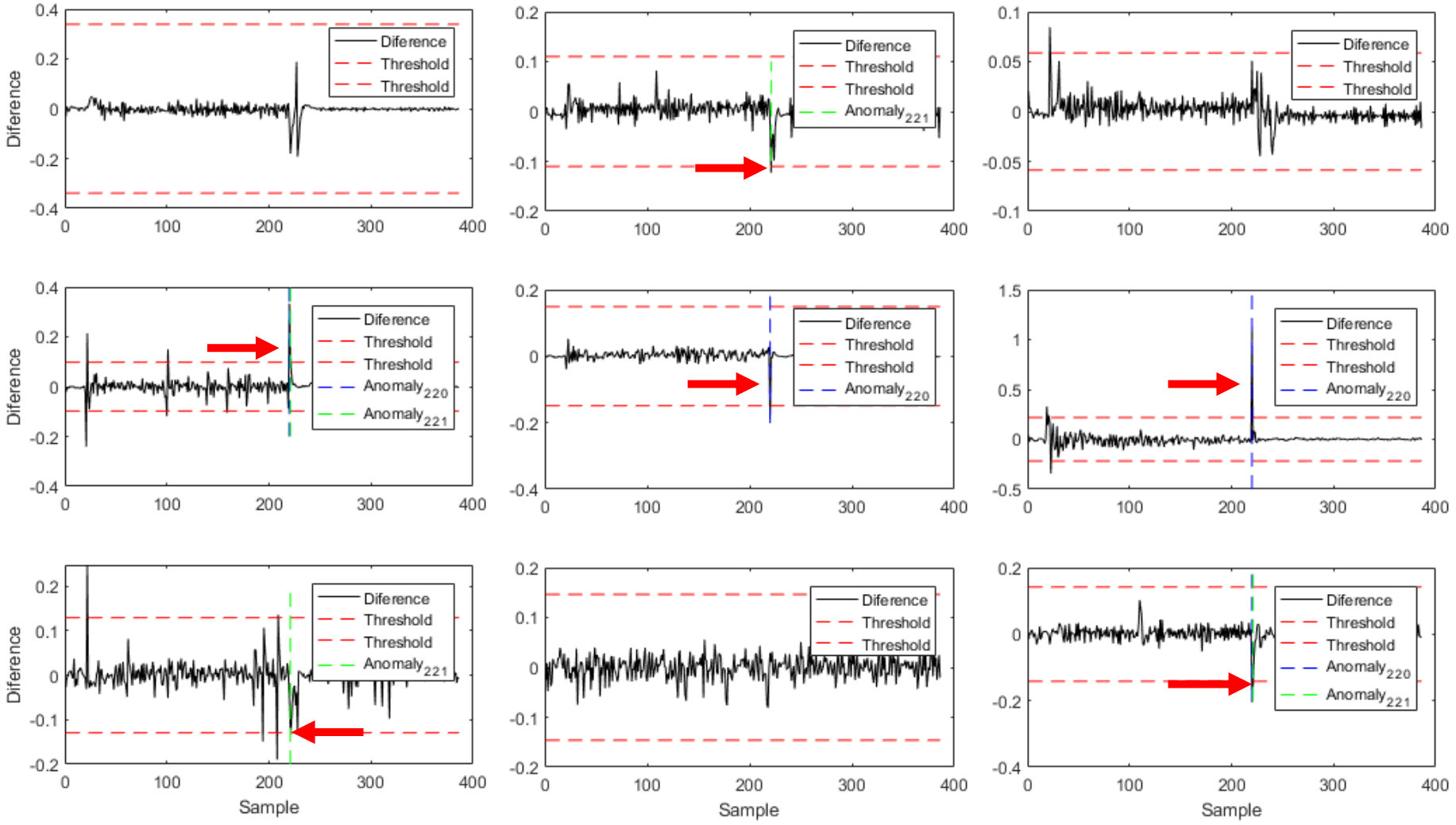
Anomaly Detection – Simultaneous (t=165)



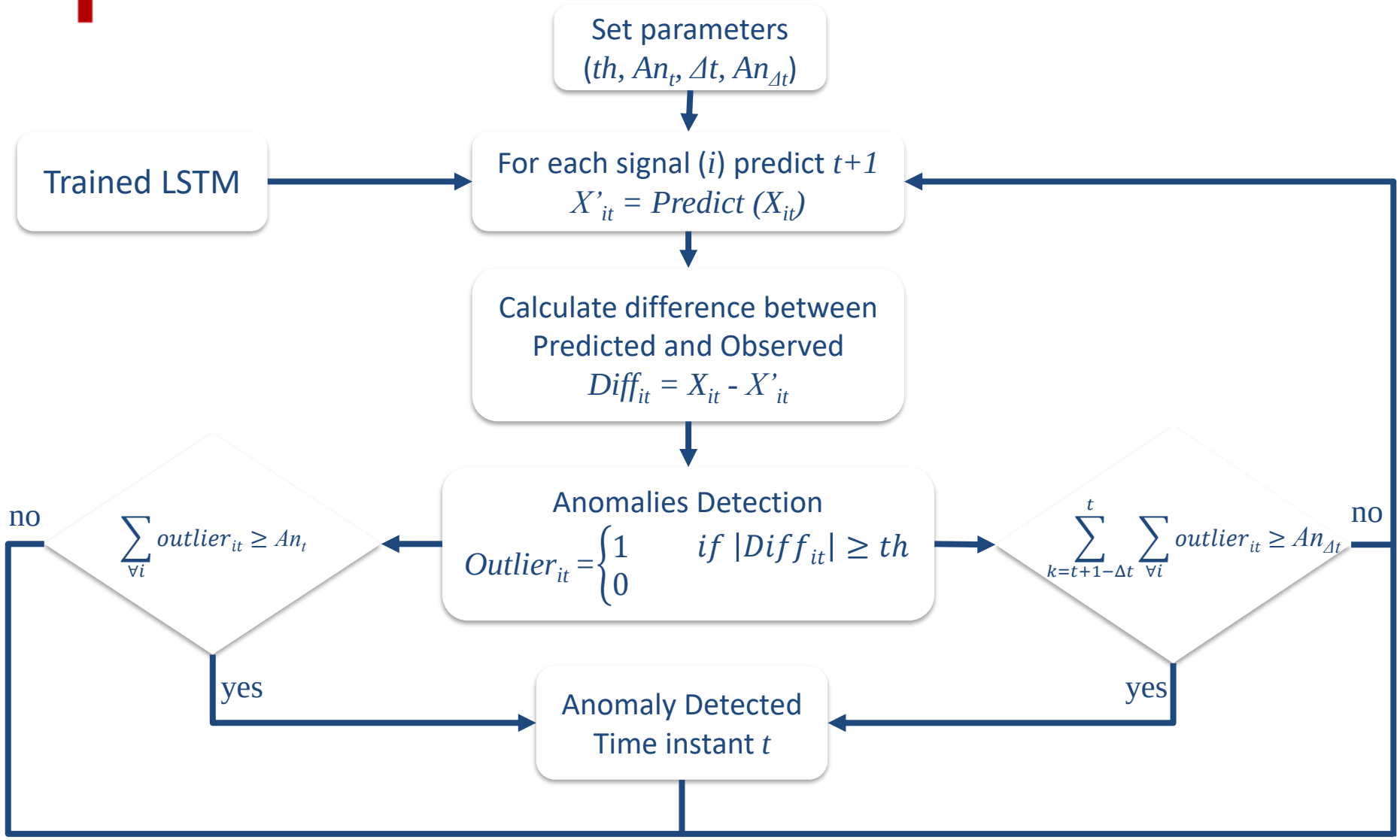
Anomaly Detection – Simultaneous (Δt)



Anomaly Detection – Simultaneous (Δt)



Anomaly Detection – Algorithm



Outline

- Introduction
- Background
  - Anomalies
- Proposed Solution
  - Recurrent Neural Networks (LSTM)
  - Anomaly Detection
- **Results**
- Summary

Results

❑ Simultaneous Anomalies Detection in a Shot (t)

The wider is the band the less anomalies are detected

	K (th = K*STD)							
An_t	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
1	190	109	67	40	21	11	8	6
2	96	34	8	3	2	0	0	0
3	49	11	4	0	0	0	0	0
4	21	1	0	0	0	0	0	0
5	4	0	0	0	0	0	0	0
6	1	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0

The more simultaneity is required, the less anomalies are detected.

1 simultaneous anomaly in 4 signals for k=2 at given time (t)

*100 shots randomly selected

Results

❑ Simultaneous Anomalies Detection in Time Windows (Δt=5)

The wider is the band the less anomalies are detected

The more simultaneity is required, the less anomalies are detected.

	K (th = K*STD)							
$An_{\Delta t}$	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
1	266	204	153	110	62	33	30	25
2	212	98	50	35	21	5	3	0
3	146	54	25	3	2	0	0	0
4	92	35	8	0	0	0	0	0
5	64	5	0	0	0	0	0	0
6	30	3	0	0	0	0	0	0
7	15	0	0	0	0	0	0	0
8	4	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0

4 simultaneous anomalies in 8 signals for k=1 with Δt=5

*100 shots randomly selected

Summary

- LSTM networks can learn the shape of a waveform (one model for signal).
- LSTM networks can be used for anomaly detection in signals.
- The specialists have to define the parameters to distinguish the noise from the real anomalies.
- It is possible to design supervised systems that allow the detection of previous detected/studied anomalies.
- In the paper ID. 484 you can find other anomaly detection methods.

Automatic recognition of anomalous patterns in discharges by recurrent neural networks

Gonzalo Farias[1]
Ernesto Fabregas[2]
Sebastián Dormido-Canto[2]
Jesús Vega[3]
Sebastián Vergara[1]

May 13-17 2019 / Daejeon, KOREA

1 PONTIFICIA UNIVERSIDAD CATOLICA DE VALPARAISO
2 UNED
3 GOBIERNO DE ESPAÑA MINISTERIO DE ECONOMÍA Y COMPETITIVIDAD
Ciemat Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas