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]

Daejeon, Republic of Korea, May 2019

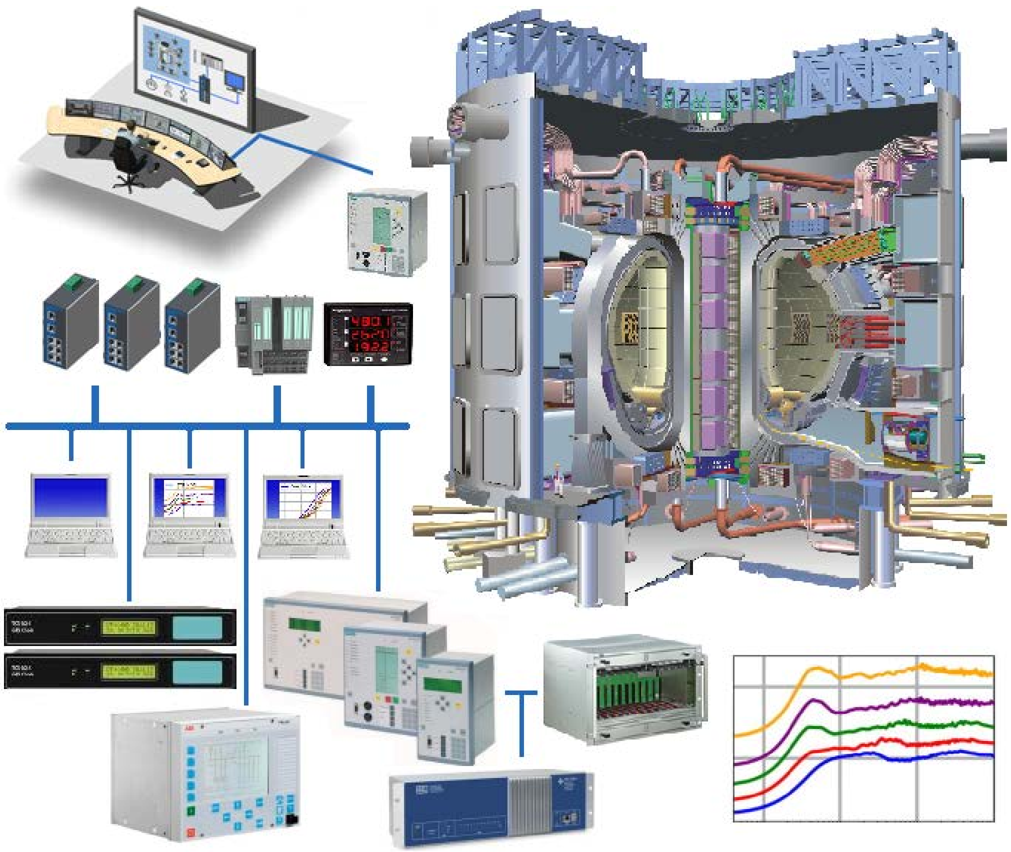
1 PONTIFICIA UNIVERSIDAD CATOLICA DE VALPARAISO
2 UNED
3 GOBIERNO DE ESPAÑA MINISTERIO DE ECONOMÍA Y COMPETITIVIDAD
Ciemat Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas

Outline

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

Introduction

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

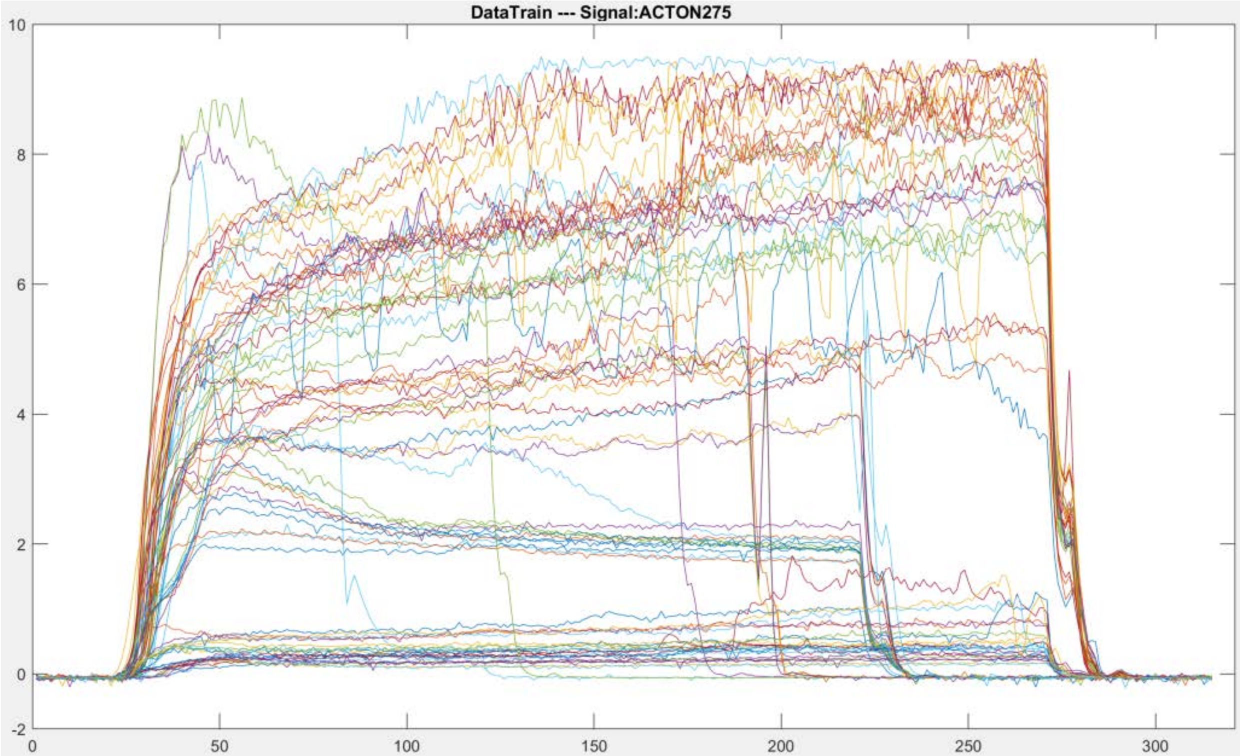


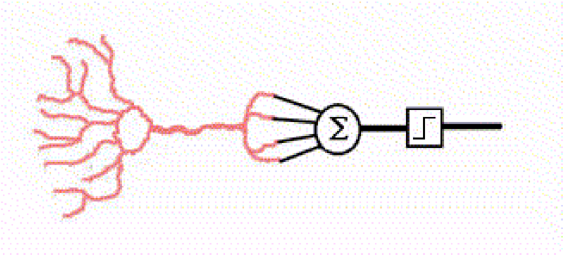
A shot of few seconds can generate huge quantity of data:

- **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.

Background

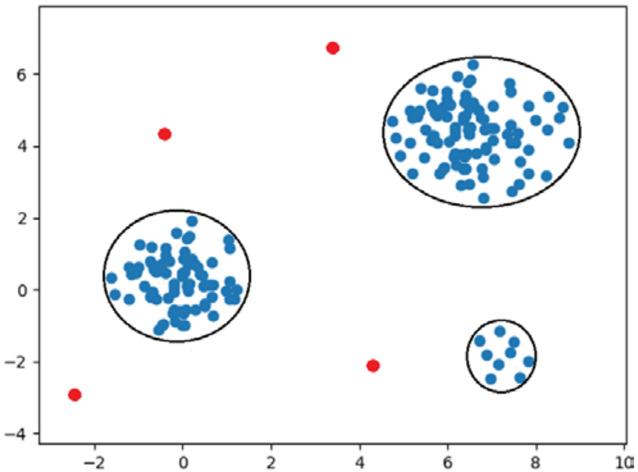
- The idea is to use Artificial Intelligence to deal with this data.
- Create systems that allow specialists to analyze and interpret data more quickly and efficiently than manually.



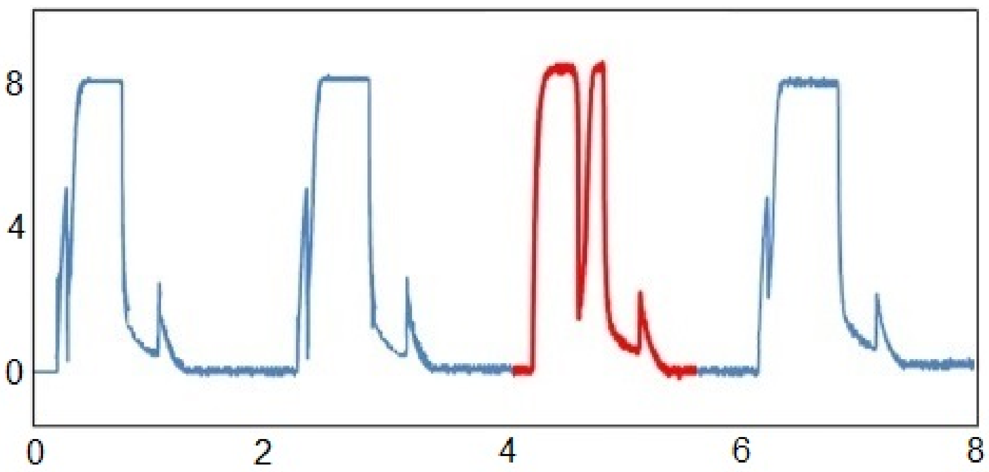


Background – Anomalies

- **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.
- The other type of anomaly is an anomalous behavior, which is a **periodic collapsing phenomenon in time series**.



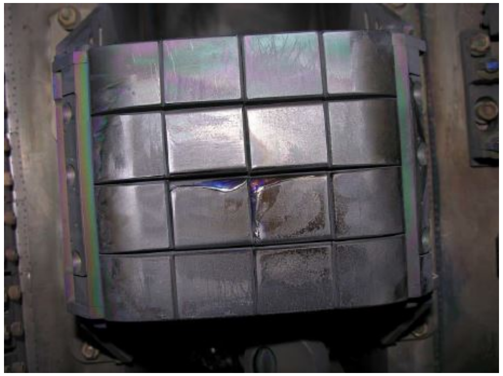
Clustering

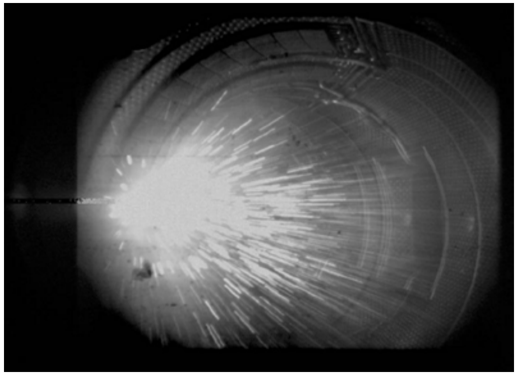


Pattern recognition

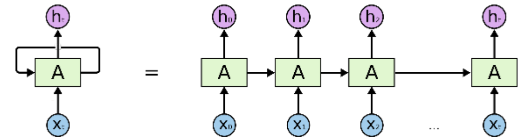
Background – Anomalies

- We try to find anomalies in signals (known and unknown).
- **Unknown** (plasma behavior).
- **Known**: disruptions or L-H and H-L transitions.

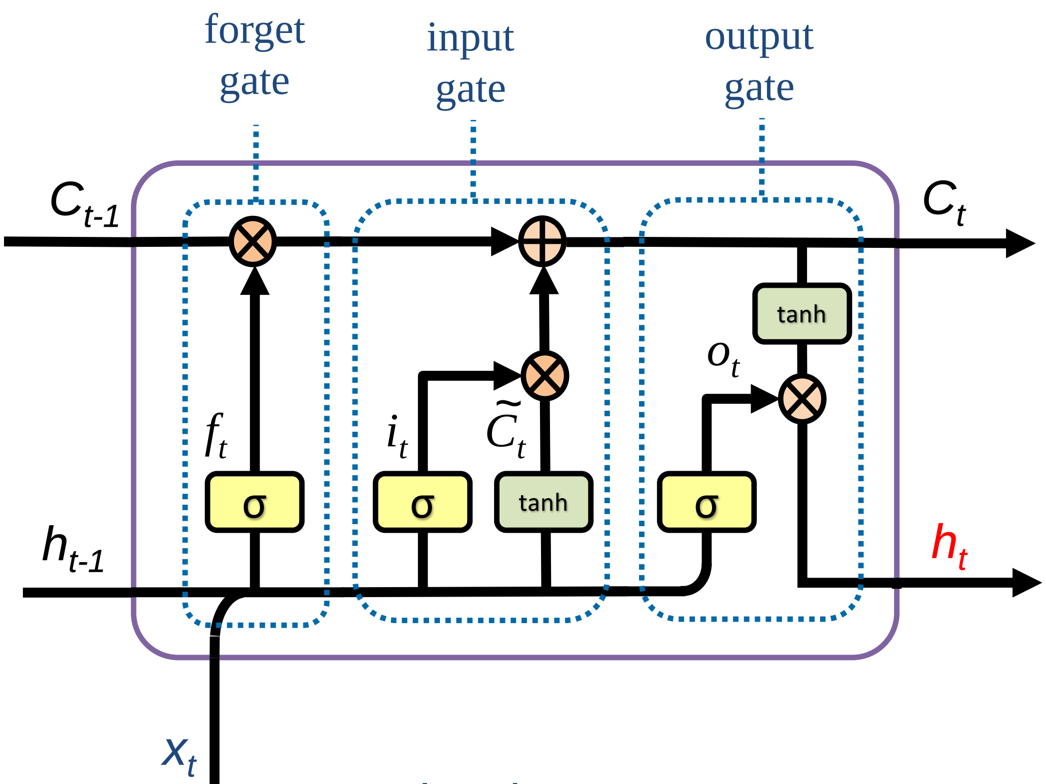




Background – LSTM



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



*They have memory

Forget gate layer

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

Input gate layer

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

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

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

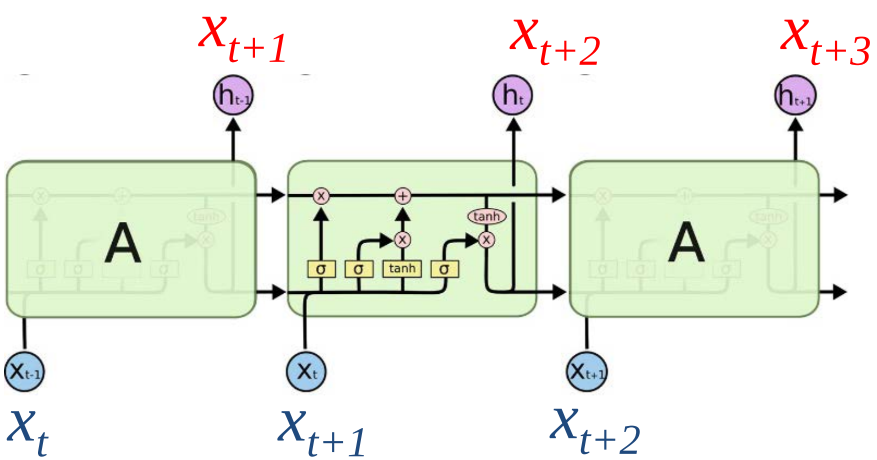
Output gate layer

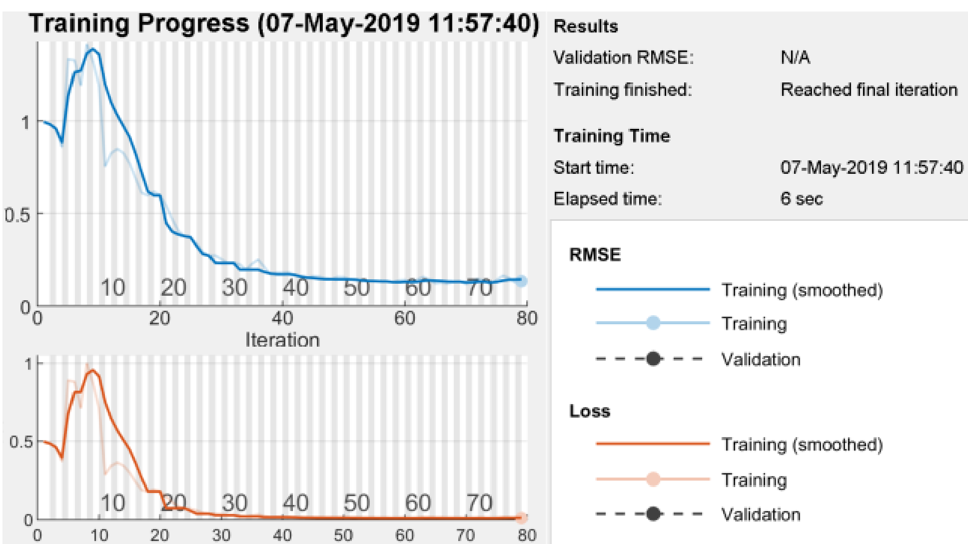
$$o_t = \sigma\,(\boldsymbol{W_o}\,[h_{t-1}, x_t] + \boldsymbol{b_o})$$

$$h_t = o_t * tanh\,(C_t)$$

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

Background – 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

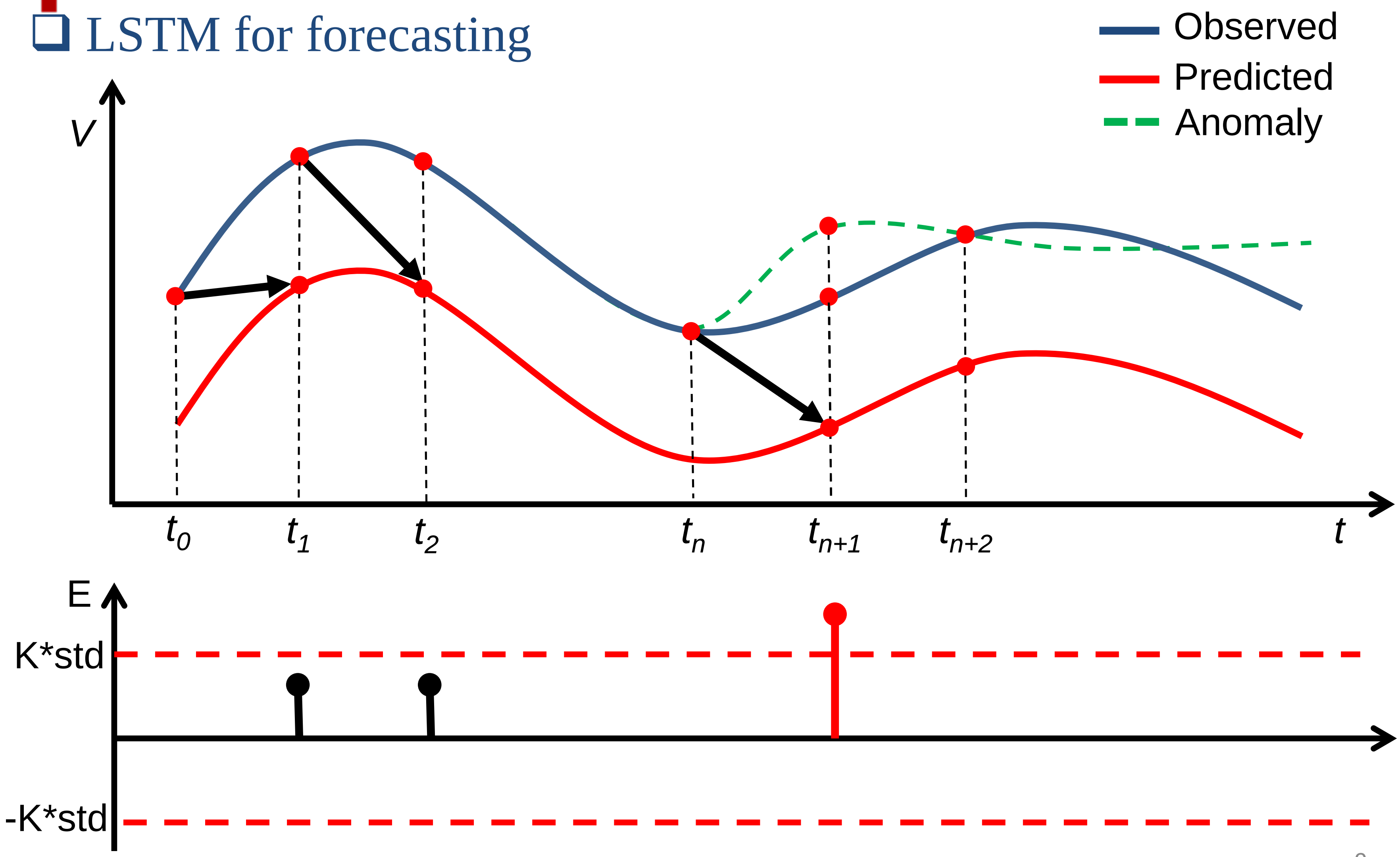
$$o_t = \sigma\left(\boldsymbol{W_o}[h_{t-1}, x_t] + \boldsymbol{b_o}\right)$$

$$h_t = o_t * tanh(C_t)$$

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

Background – Forecasting (test stage)

- LSTM for forecasting



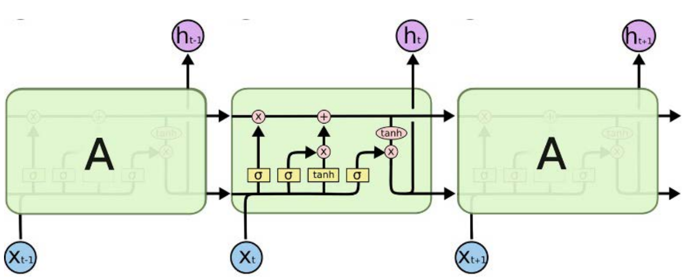
Goals

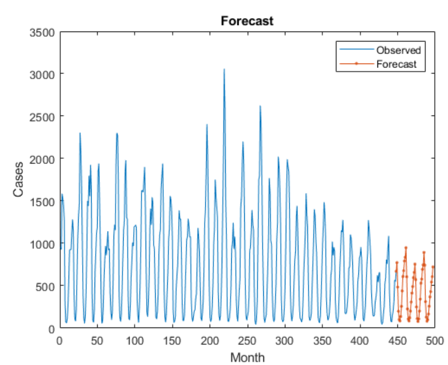
❑ General Goal

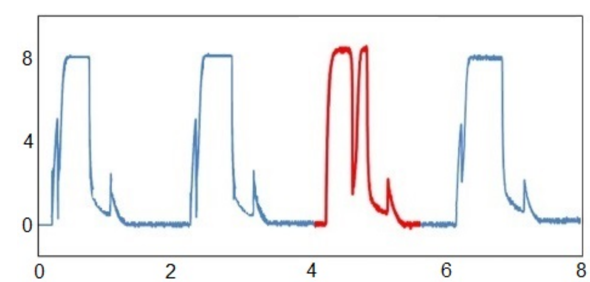
- ❑ Anomaly detection using Recurrent Neural Network (LSTM - Long Short Term Memory).

❑ Specific Goal

- ❑ The LSTM Neural Network learns the waveform to detect anomalies through forecasting.





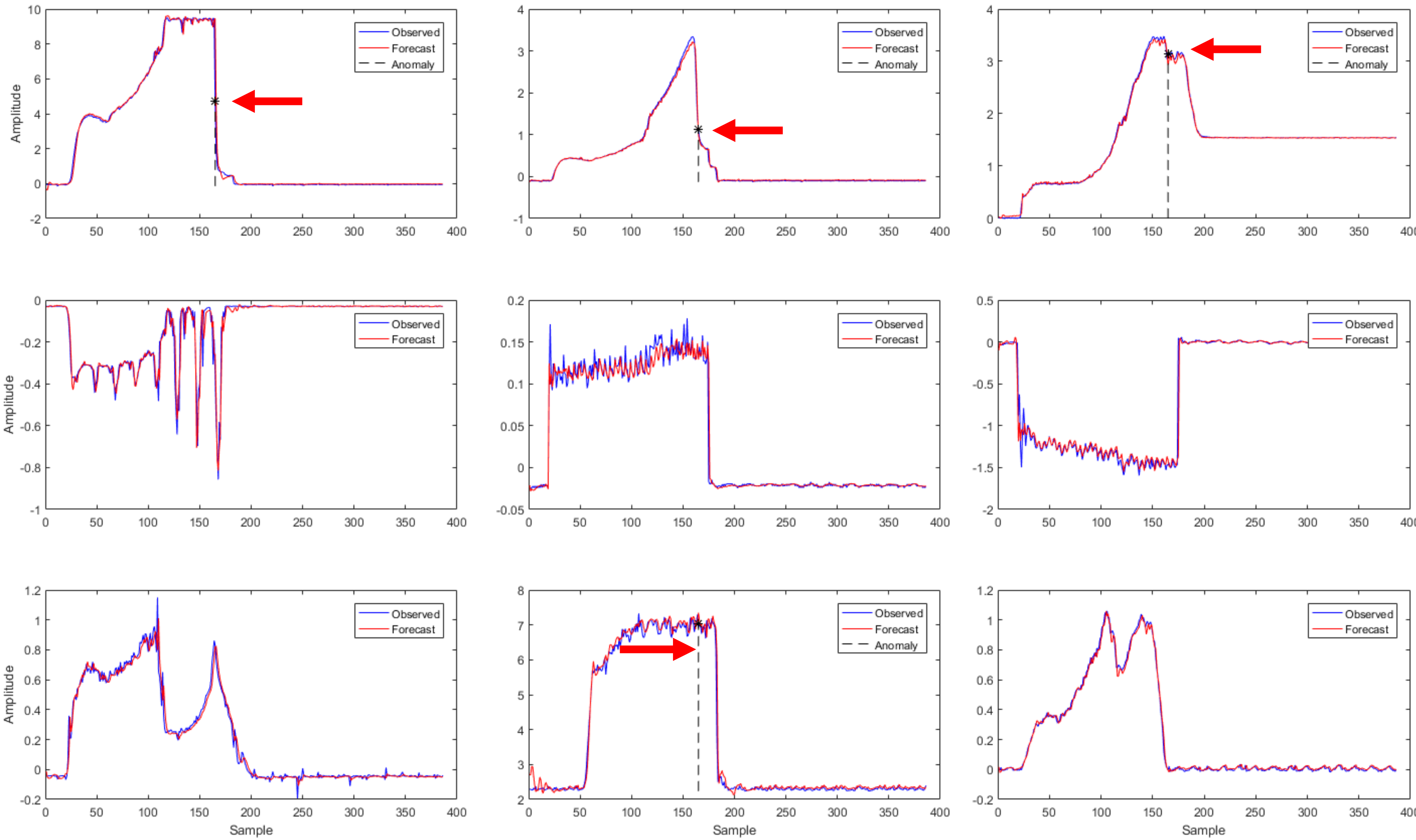


* https://colah.github.io/posts/2015-08-Understanding-LSTMs/

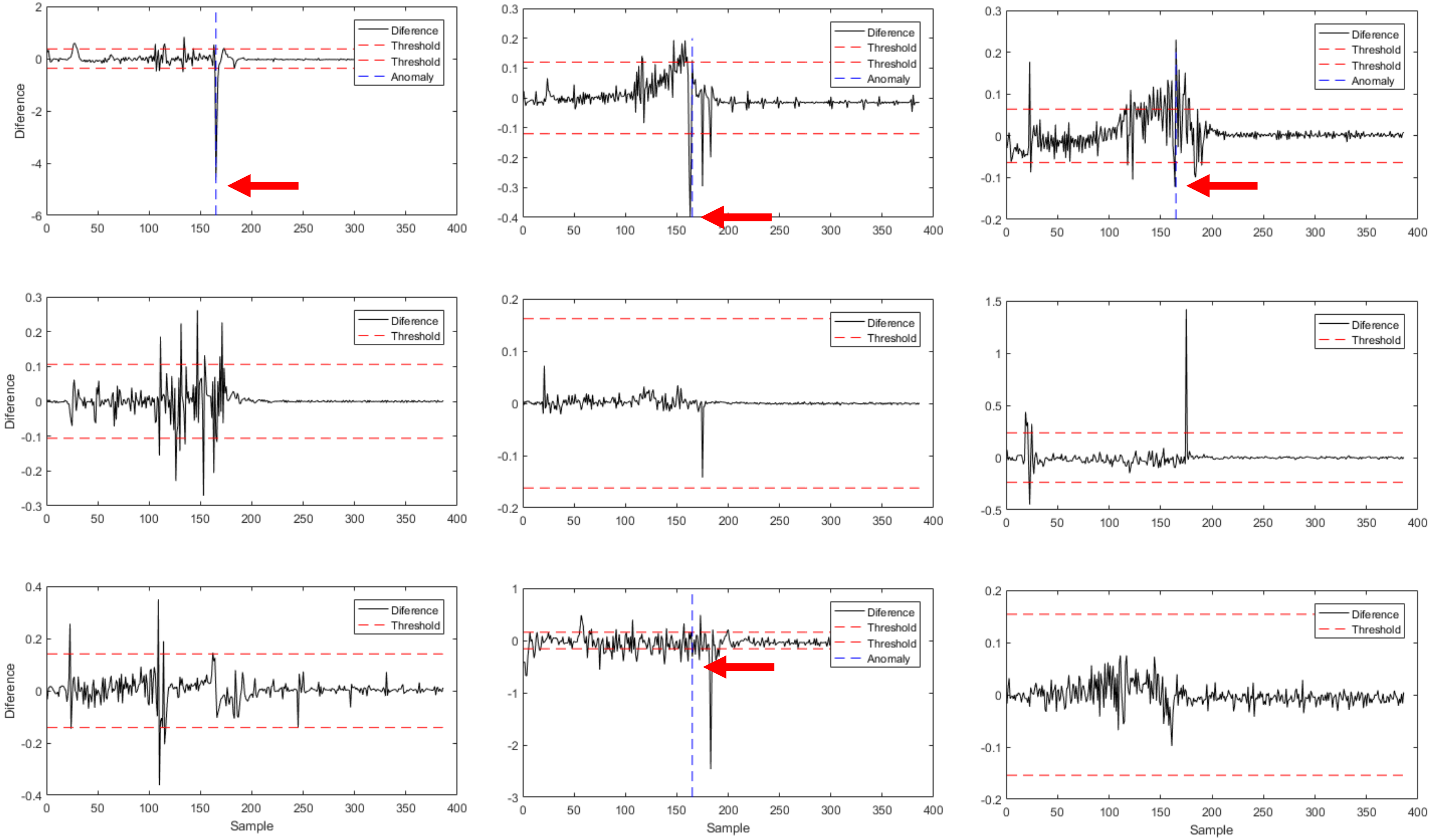
Anomaly Detection – Threshold (th=k*std)

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

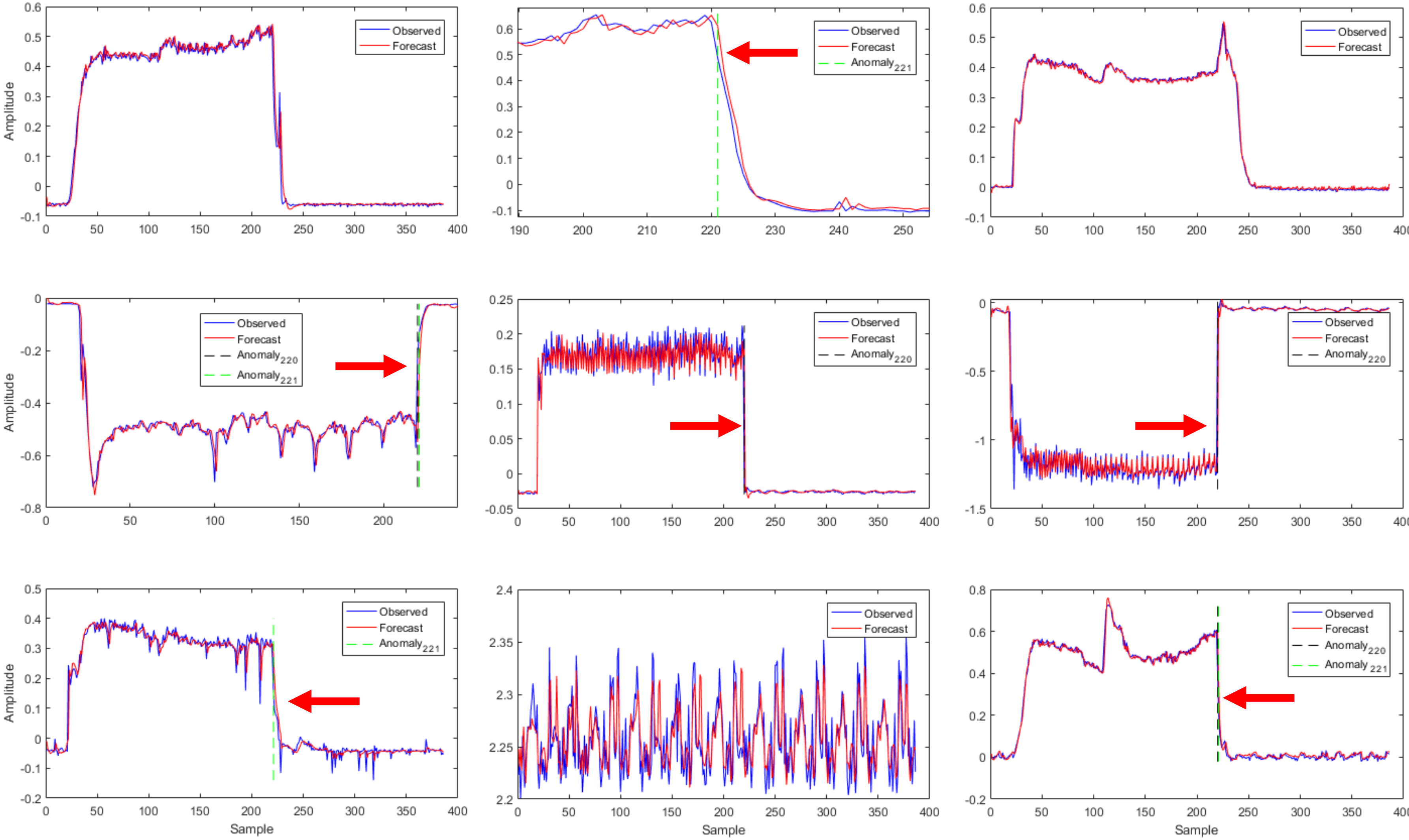
Anomaly Detection – Simultaneous (t=165)



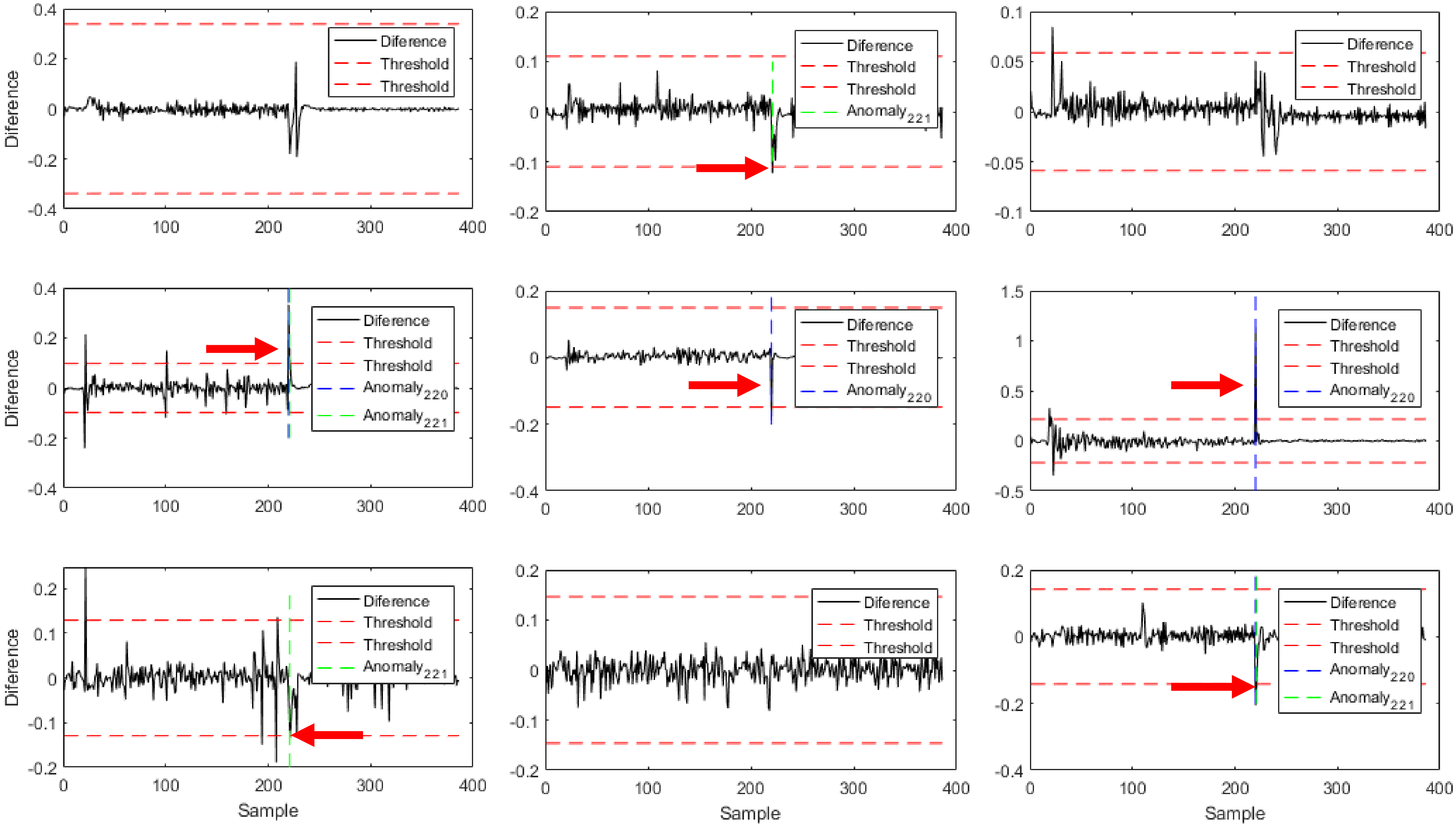
Anomaly Detection – Simultaneous (t=165)



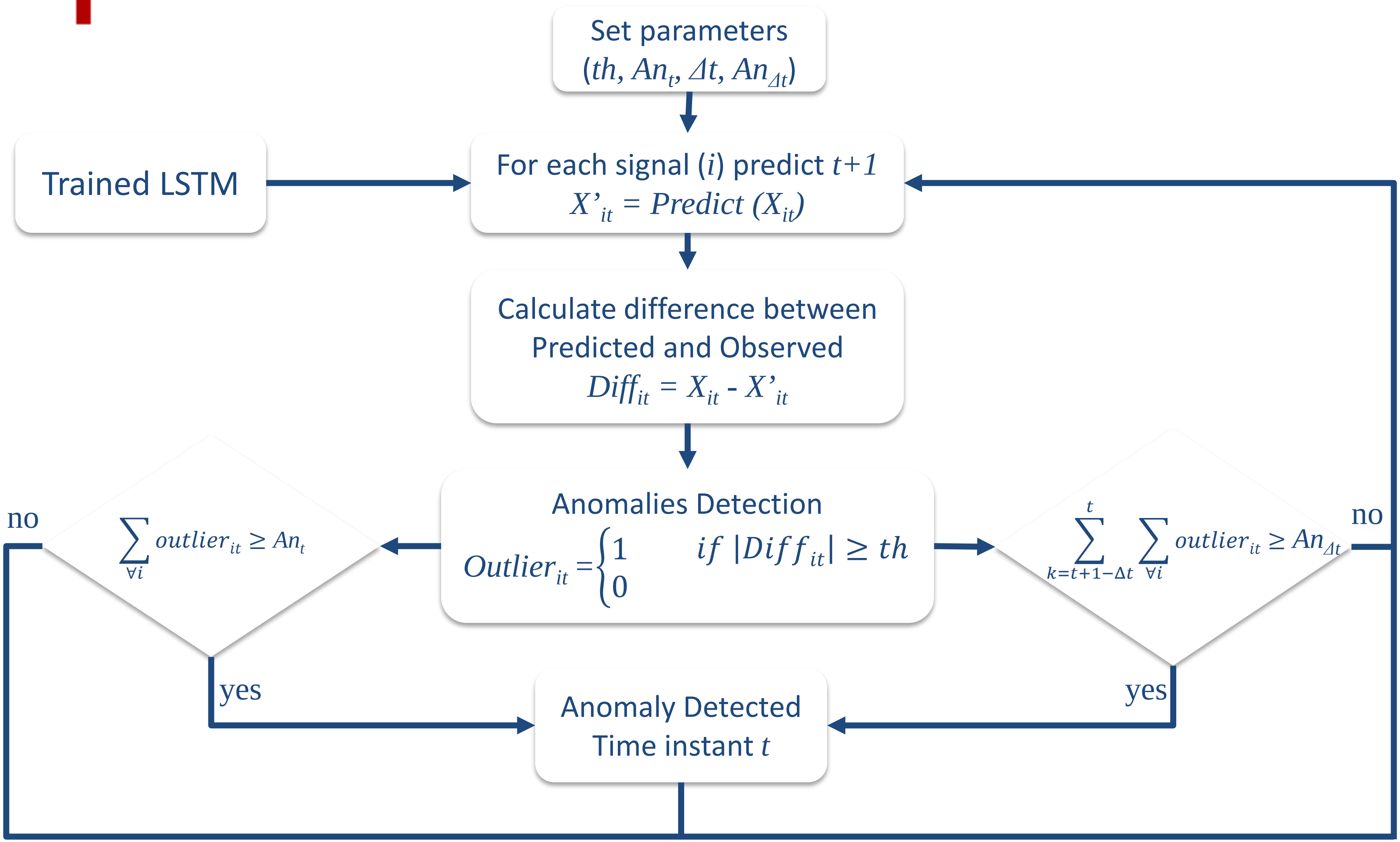
Anomaly Detection – Simultaneous (Δt)



Anomaly Detection – Simultaneous (Δt)



Anomaly Detection – Algorithm



Results

- Database from TJ-II Fusion Device
- 430 Shots with 9 signals each (80% for training, 20% for testing randomly selected)
- Training time (5 sec/signal and shot, GPU)
- Testing time (forecasting, ~10 sec/signal and shot, GPU)

originalData | data

430x10 cell

	1	2	3	4	5	6	7	8	9	10
1	10104	65536x2 double	65536x2 double	6552x2 double	61440x2 double	65536x2 double	65536x2 double	65536x2 do...	65536x2 double	30208x2 double
2	10107	65536x2 double	65536x2 double	32766x2 double	61440x2 double	65536x2 double	65536x2 double	65536x2 do...	65536x2 double	30208x2 double
3	10108	65536x2 double	65536x2 double	32766x2 double	61440x2 double	65536x2 double	65536x2 double	65536x2 do...	65536x2 double	30208x2 double
4	10109	65536x2 double	65536x2 double	10921x2 double	61440x2 double	65536x2 double	65536x2 double	65536x2 do...	65536x2 double	30208x2 double
5	10110	65536x2 double	65536x2 double	9361x2 double	61440x2 double	65536x2 double	65536x2 double	65536x2 do...	65536x2 double	30208x2 double
6	10112	65536x2 double	65536x2 double	8190x2 double	61440x2 double	65536x2 double	65536x2 double	65536x2 do...	65536x2 double	30208x2 double
7	10114	65536x2 double	65536x2 double	9361x2 double	61440x2 double	65536x2 double	65536x2 double	65536x2 do...	65536x2 double	30208x2 double
8	10115	65536x2 double	65536x2 double	9361x2 double	61440x2 double	65536x2 double	65536x2 double	65536x2 do...	65536x2 double	30208x2 double
9	10116	65536x2 double	65536x2 double	9361x2 double	61440x2 double	65536x2 double	65536x2 double	65536x2 do...	65536x2 double	30208x2 double
10	10119	65536x2 double	65536x2 double	9361x2 double	61440x2 double	65536x2 double	65536x2 double	65536x2 do...	65536x2 double	30208x2 double
11	10120	65536x2 double	65536x2 double	9361x2 double	61440x2 double	65536x2 double	65536x2 double	65536x2 do...	65536x2 double	30208x2 double

Results

- Simultaneous Anomalies Detection in a Shot (Δt=1)

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_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

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

Conclusions

- 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 allows the detection of previous detected/studied anomalies.

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]

Daejeon, Republic of Korea, May 2019

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