SPS ongoing studies: towards reinforcement learning for electrostatic septum alignment

M. Schenk[□], M. Fraser[*], B. Goddard[*], S. Hirlaender[*□], V. Kain[*], T. Pieloni[□]

[*]CERN & [□]EPFL, Switzerland

[□]also at University of Malta

Overview

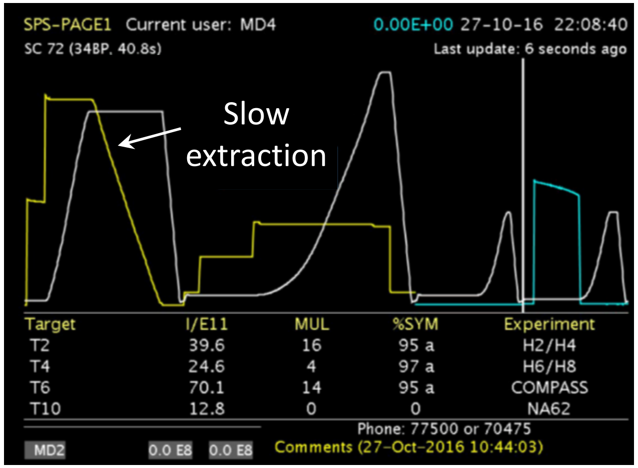
- **Introduction**
  - SPS slow-extraction scheme
  - Electrostatic septum: an optimisation problem
- **Motivation:** reinforcement learning
- **Data-driven model**
  - Overview
  - Results: training and validation
- **Final remarks**

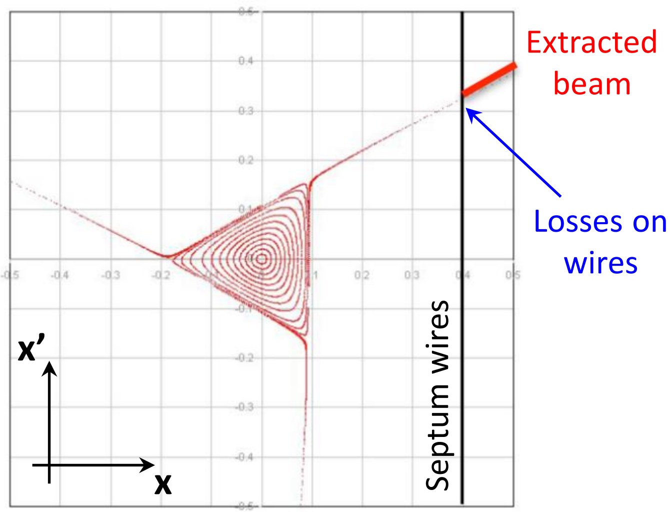
Introduction

SPS slow-extraction scheme

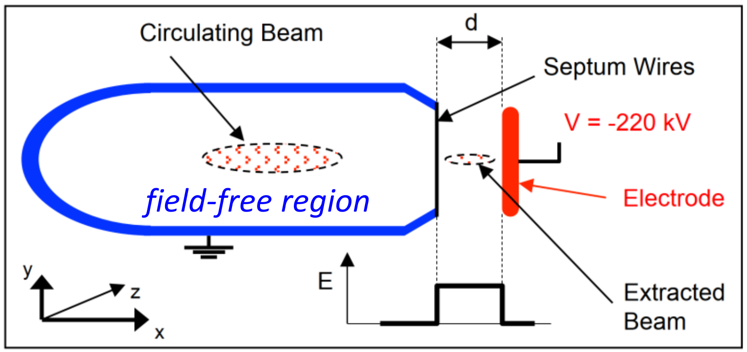
Super Proton Synchrotron (SPS) delivers protons to Fixed Target experiments (FT)

- FT request **constant particle flux** for several seconds (= spill)
- **Solution:** *multi-turn resonant extraction scheme*
  1. Excite third-order resonance
  2. Extract beamlets by means of electrostatic septum (ZS)
- **Challenge:** *inherently a lossy process*
  Optimisation problem: align ZS anode wires to reduce particle losses





B. Goddard (TE-ABT)

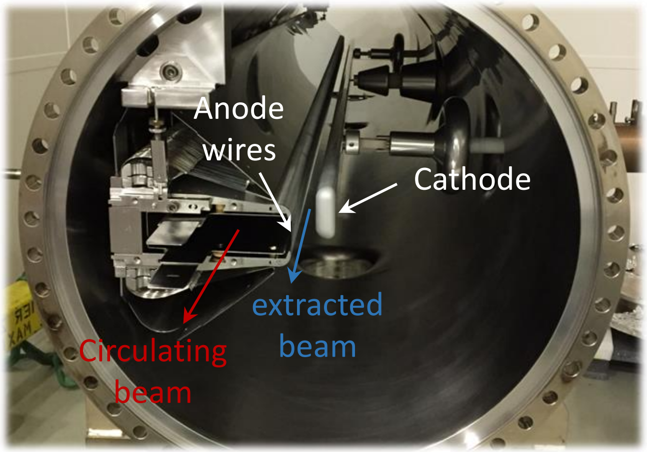


M. Fraser (TE-ABT)

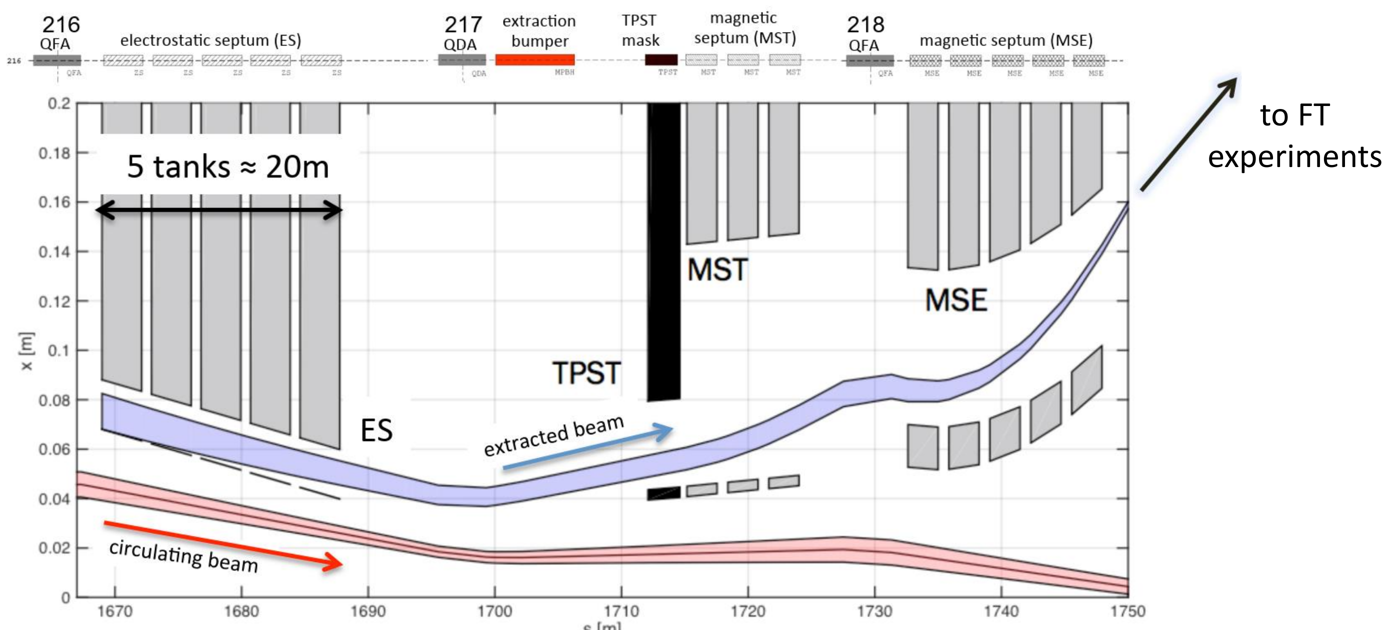
Introduction
SPS ZS

SPS ZS composed of 5 tanks

- System with 12 degrees of freedom (dof)
  - 10 dof: adjustable positions for anode wires up- and downstream for every tank
  - 2 dof: girder positions up- and downstream
- Loss monitoring: > 20 beam loss monitors (BLMs)
- ***Entangled dynamics:* modeling and optimisation of system not straightforward**



M. Fraser (TE-ABT)

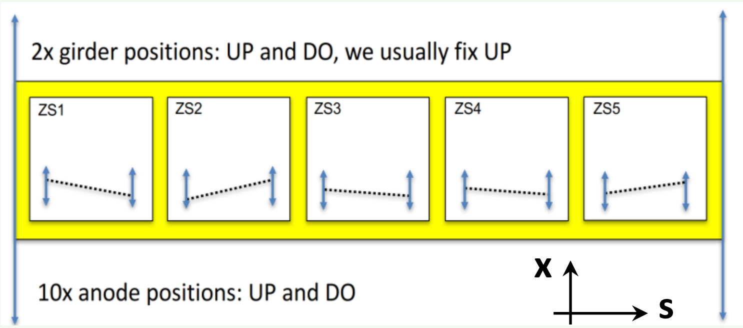


M. Fraser (TE-ABT)

Introduction

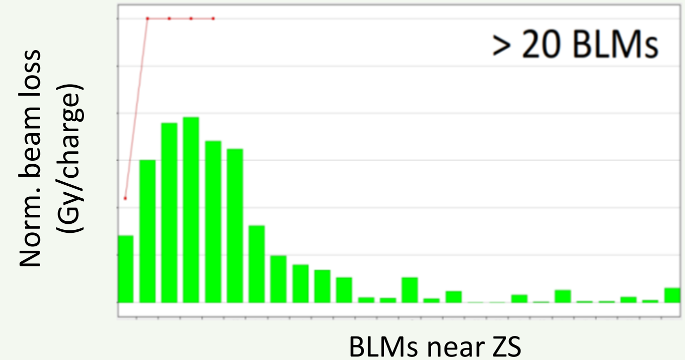
SPS ZS alignment procedures

Controls *(features)*

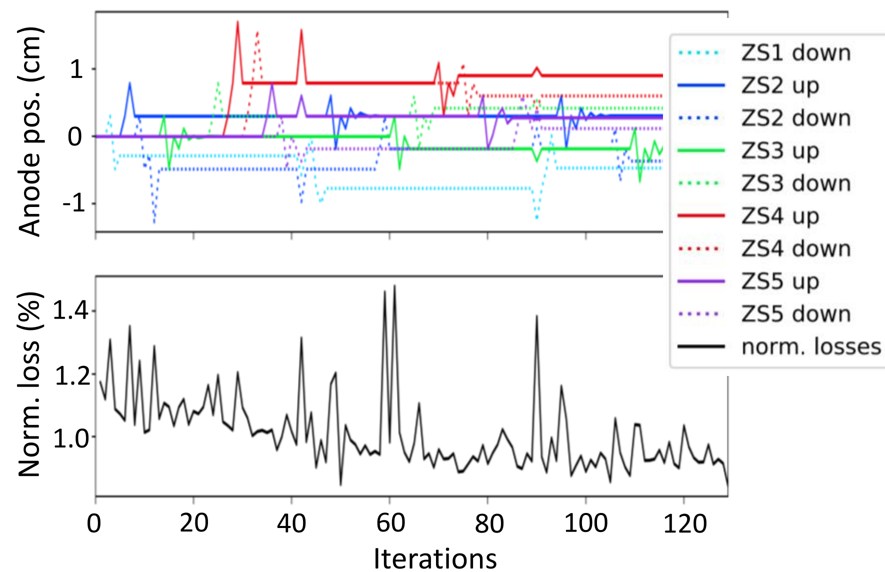


M. Fraser (TE-ABT)

Output *(observables)*



- **Until 2018:**
  - Manual anode-by-anode optimisation
  - Time-consuming (≈ 8 h) and tedious
- **Test in Nov. 2018:**
  - **Automatic alignment:** proof-of-principle with *modified* Powell optimiser
  - Does first line search along every direction, then simultaneous parameter adjustment
  - **Big impact:** time for full adjustment reduced to 40’ with same improvement on total loss



S. Hirlaender, V. Kain (BE-OP) & TE-ABT

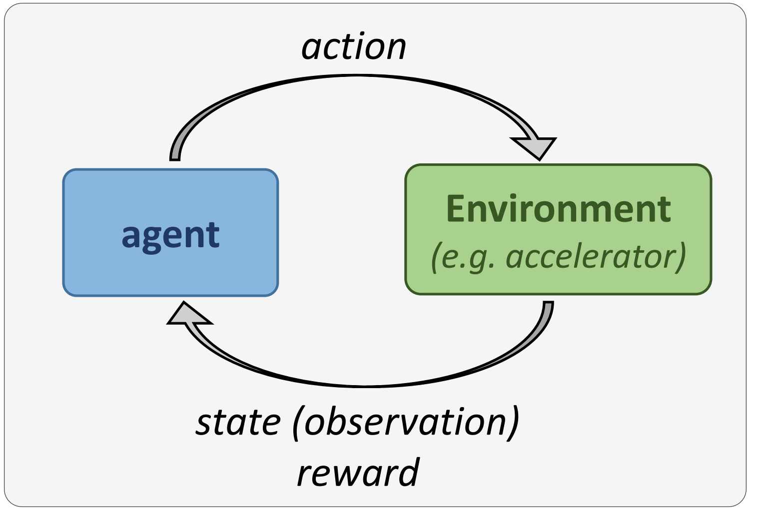
Motivation
Reinforcement learning

Can we do even better?

- Powell optimiser has *no memory*:
  ZS alignment from scratch every time
- *Reinforcement learning (RL)*
  - Agent interacts with environment and learns dynamics of the system
  - State *not* restricted to action space
  - Agent strategy / knowledge typically represented by neural network (Deep RL)

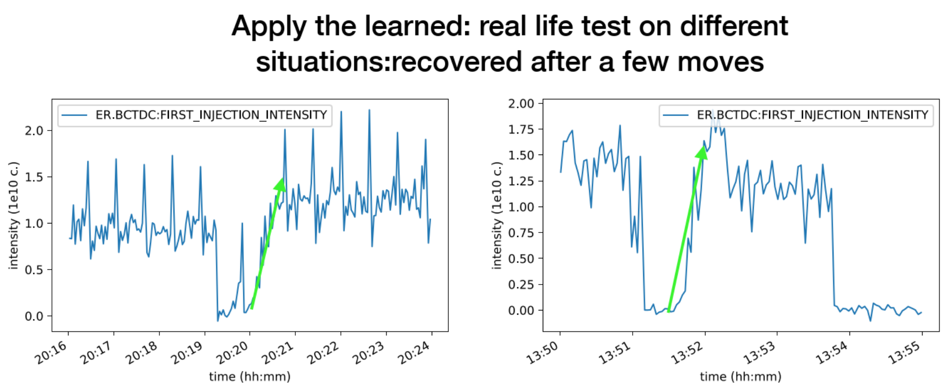
  => once trained, agent finds optimum in a few steps

Reinforcement learning

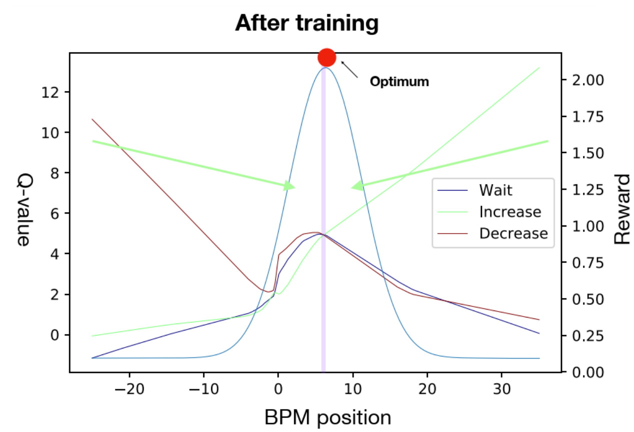


Motivation
Reinforcement learning

- **RL defines state-of-the-art performance** in robotics, playing games (e.g. Atari, AlphaGo), … and artificial intelligence problems in general
- **Applied to particle accelerator domain**
  - Proof-of-principle at CERN Low Energy Ion Ring (LEIR)
  - 1 control parameter: optimise injected intensity by tuning dipole magnet strength
  - Solved as discrete control problem
  - Model-free RL agent **successfully learned the dynamics** of the system



S. Hirlaender, V. Kain, M. Schenk



- Implemented and studied many **continuous control algorithms using [OpenAI gym](#) environment templates** *(S. Hirlaender, V. Kain)*
- Experience shows: **sample efficiency is of major importance for our problems**

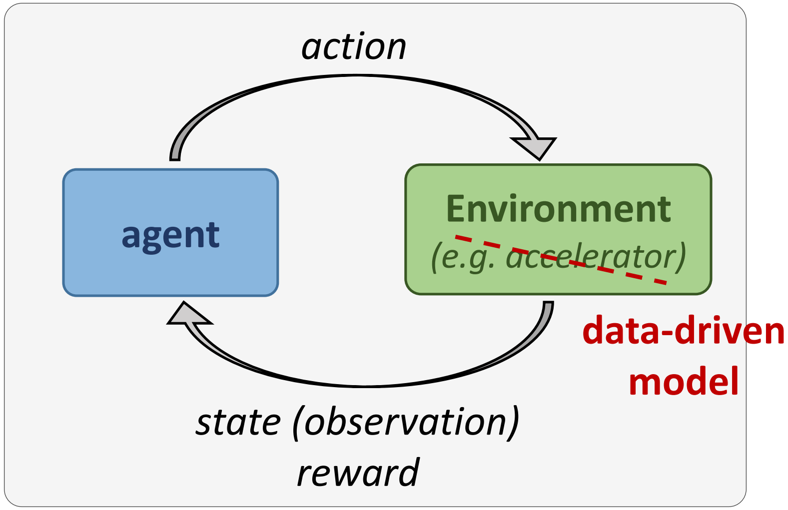
Motivation
Reinforcement learning

Can we do even better?

- Powell optimiser has *no memory*:
  ZS alignment from scratch every time
- *Reinforcement learning (RL)*
  - Agent interacts with environment and learns dynamics of the system
  - State *not* restricted to action space
  - Agent strategy / knowledge typically represented by neural network (Deep RL)

  => once trained, agent finds optimum in a few steps

Reinforcement learning

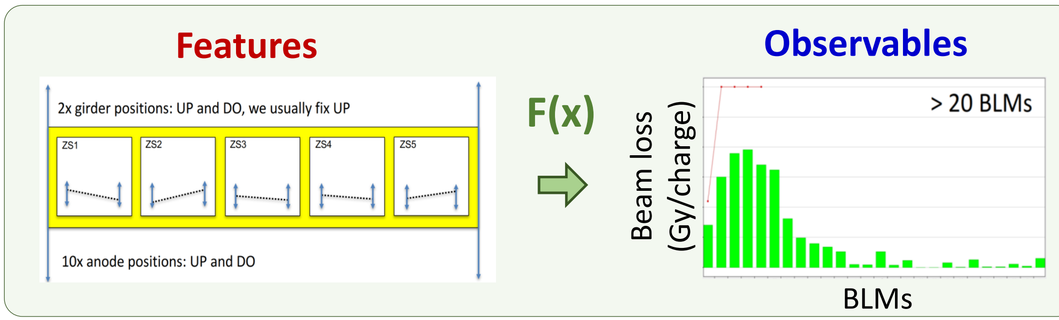


Reinforcement learning (RL)

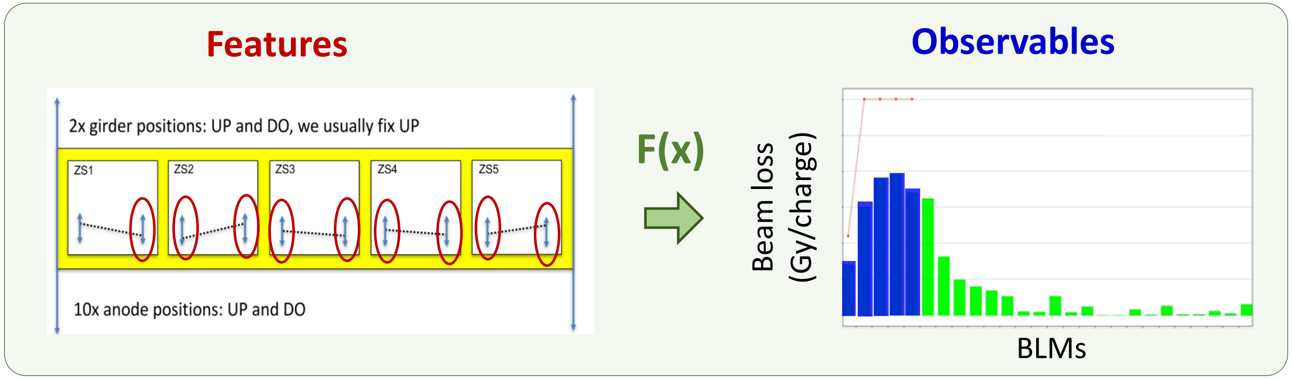
- Sample efficiency is key as *machine time is expensive*
- Idea: pre-train RL agent offline for 'warm start' in the accelerator

Offline training requires a model F(x) of the system

- Tracking simulation, data-driven model, …
- Fast, cheap evaluation needed: Pre-training may require few thousand iterations



Data-driven model
Overview



Start with simple model

- **9 features:** All anode positions *(except ZS1 UP)*, no girder positions
- **5 observables:** BLMs ZS1, ZS2, .., ZS5
- **Neural network:** feed-forward, dense, 1 hidden layer, 7 / 15 nodes, leaky ReLU activation
- Adam optimizer
- *Missing features:* girders, orbit, cathode voltage, ...

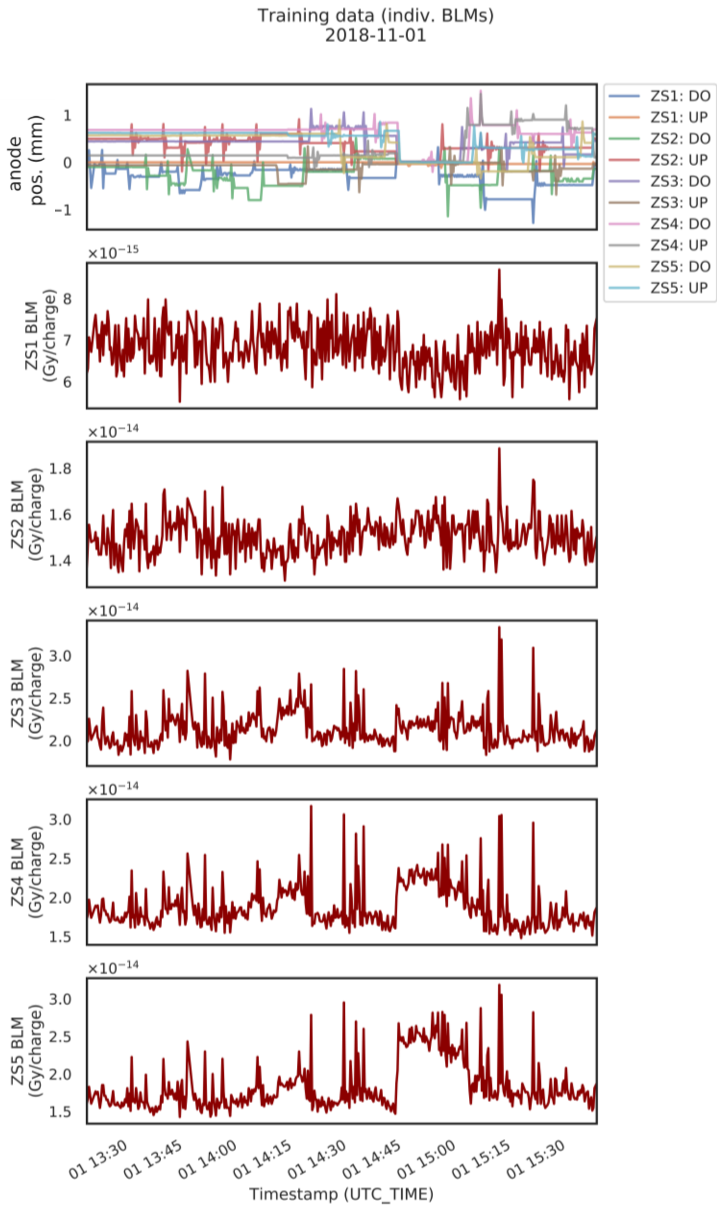
Training and test sets from existing data

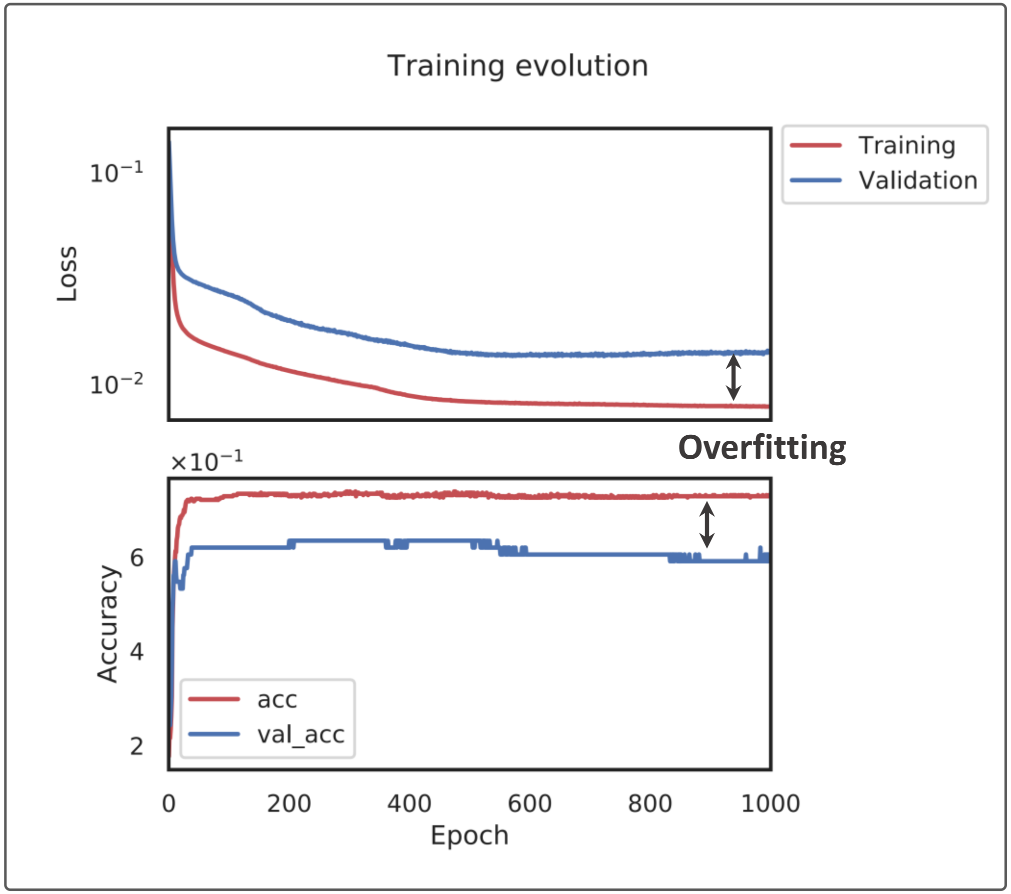
I. *Training on data from Powell scan (01.11.18, 467 samples)*
   => Predict manual scan 27.03.19 (test set)
   => Fake scans for all anode positions

II. *Training on data from Powell scan and manual scan (27.03.19, 1308 samples)*
   => Predict manual scan 30.03.19 (test set)
   => Perform anode scans as above

Data-driven model

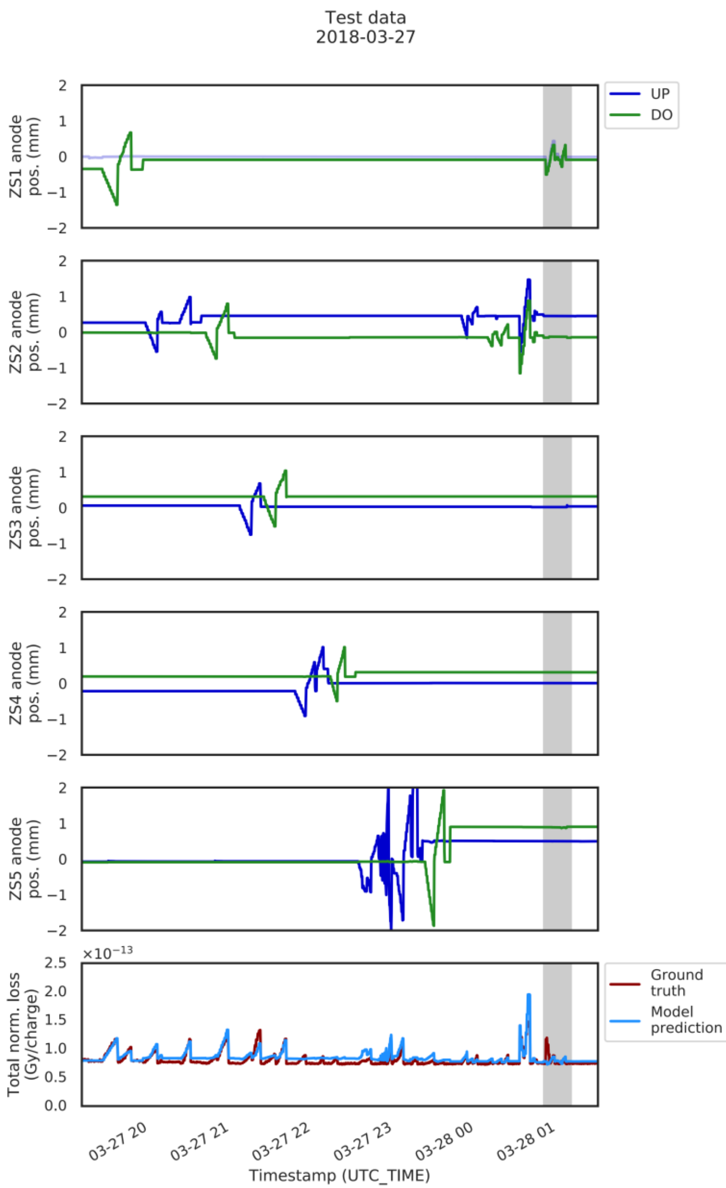
Case I: Training data and evolution (7-node neural network)





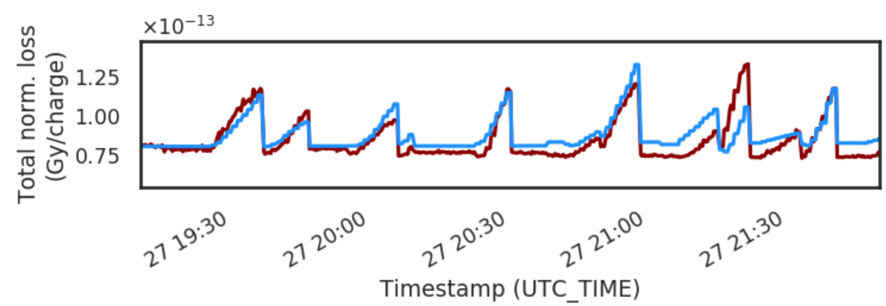
Data-driven model

Case I: Test set predictions (total loss)



- **Good quantitative agreement,** given simplicity of model and low number of samples
- Training and test set separated by **7 months**
- Offset in 'baseline' & some peaks only **qualitative agreement**
  - Missing feature (girder, orbit)?
  - See individual loss on BLM ZS5 *(backup)*
- Another source of discrepancy is **range in anode positions**
  - Training set: ≈ ± 0.5 - 1 mm
  - Test set: up to ≈ ± 2 mm

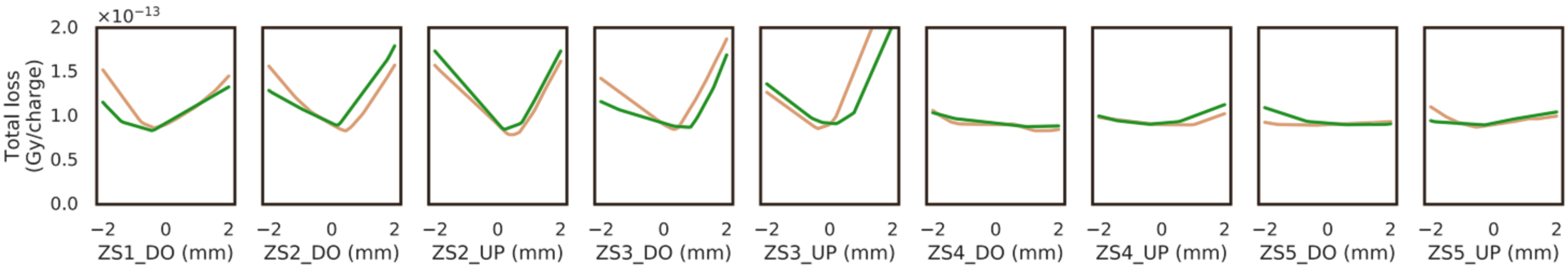
Zoomed view of total loss



Data-driven model

Validation: orthogonal anode scans

Scan individual anode positions and predict total loss



Case I: Trained on Powell data (7-node NN)

Case II: Trained on Powell and manual scan data (15-node NN)

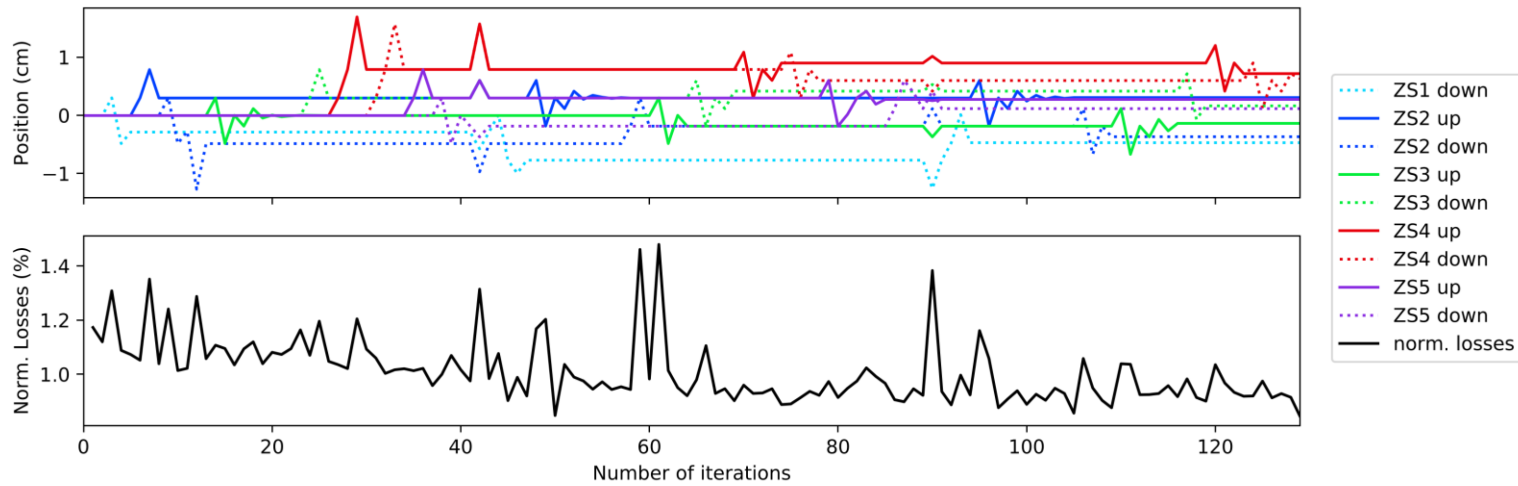
- **Loss response is convex** – *as expected*
- **Small impact** on total loss from **ZS4 and ZS5** – as *observed in the machine*
- **'Piecewise linear'** functions *due to simplicity of network*
- Case II: A second ZS NN model with 15 nodes trained on Powell *and* manual scan data performs even better on a test set *(backup)*

Good news: both models predict similar loss response

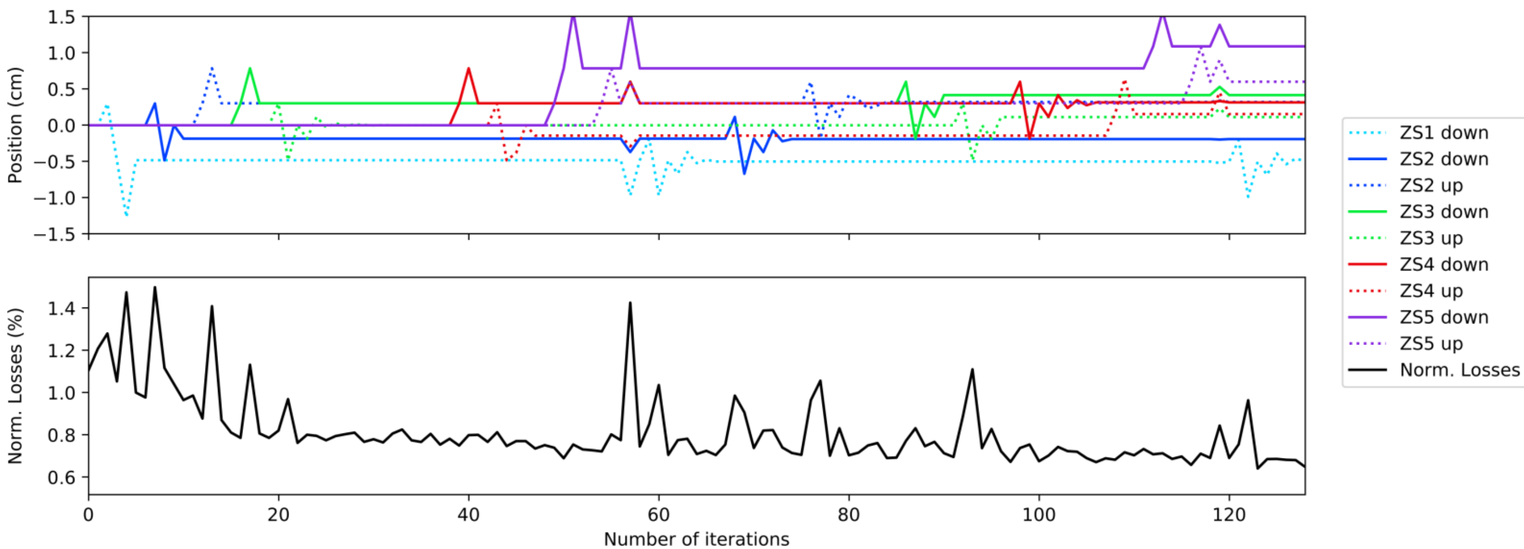
Data-driven model

Validation: Powell optimisation on SPS machine and trained model

Measurement



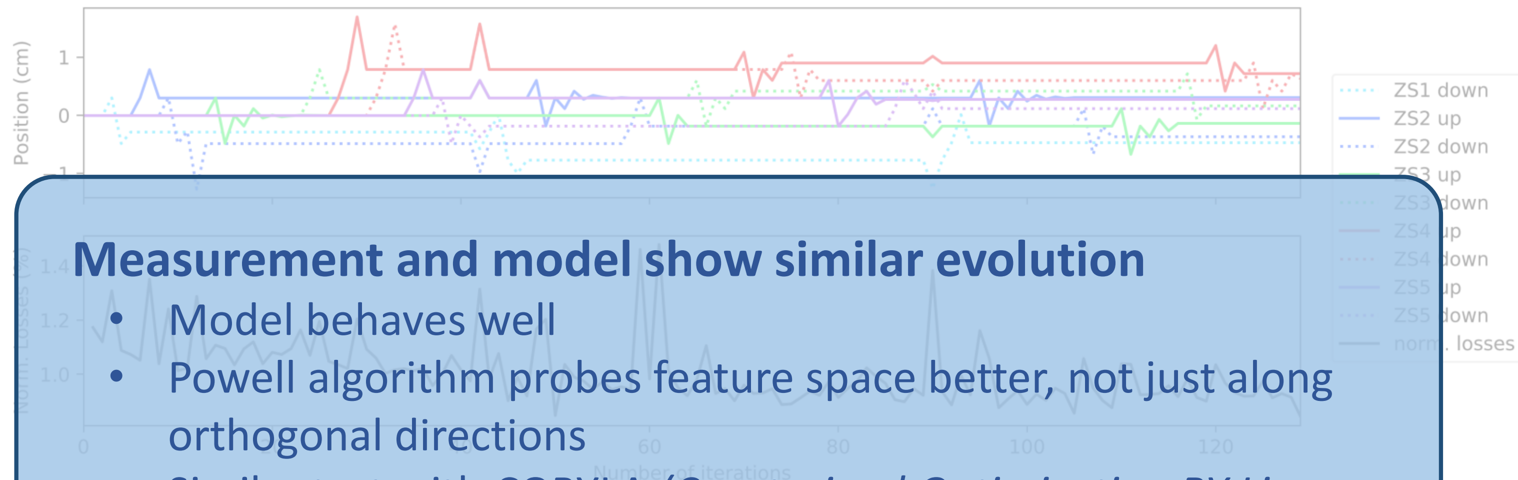
Model



Data-driven model

Validation: Powell optimisation on SPS machine and trained model

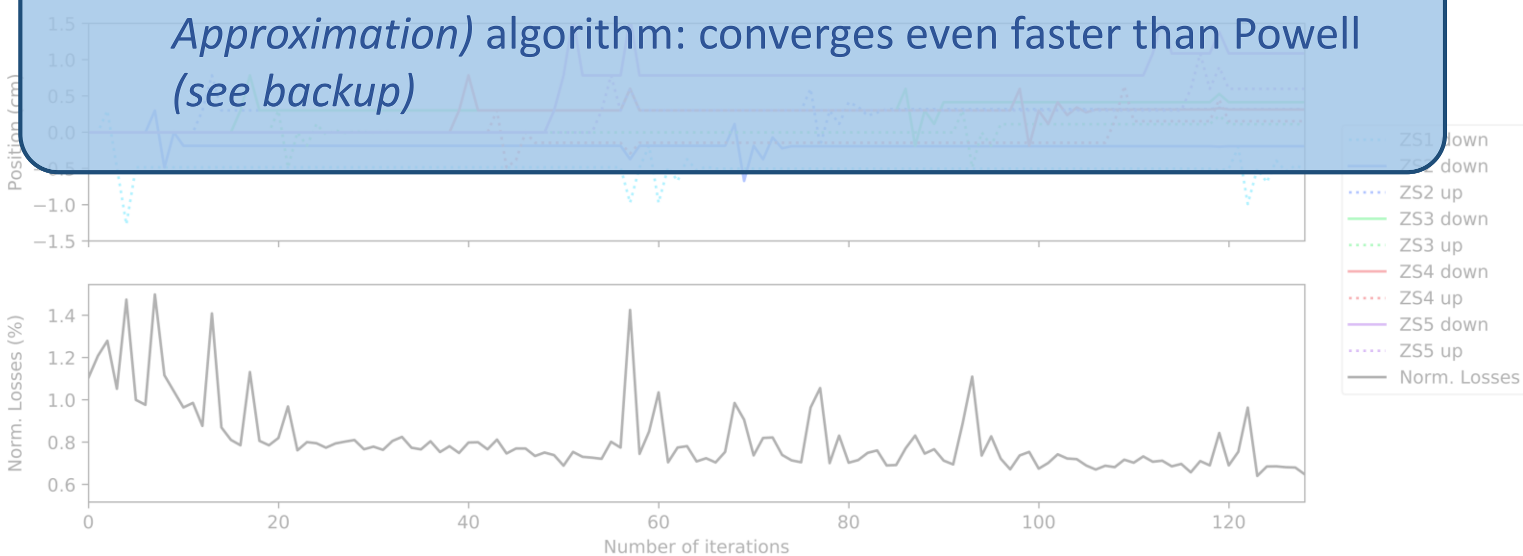
Measurement



Measurement and model show similar evolution

- Model behaves well
- Powell algorithm probes feature space better, not just along orthogonal directions
- Similar test with COBYLA *(Constrained Optimisation BY Linear Approximation)* algorithm: converges even faster than Powell *(see backup)*

Model



Final remarks

- Even with little training data: **ZS model performs well and passes validation tests**
- **It is yet incomplete and hence improved further:** adding more features / observables

- Trained ZS model is now **embedded in *OpenAI gym* environment** and used for **RL benchmarks**
- **Excellent testbed for multidimensional RL optimisation problems** to study various algorithms for **continuous control**
- **Sample efficiency is a key player** *(for accelerator domain)*
- **Data-driven or surrogate models can be game changers**

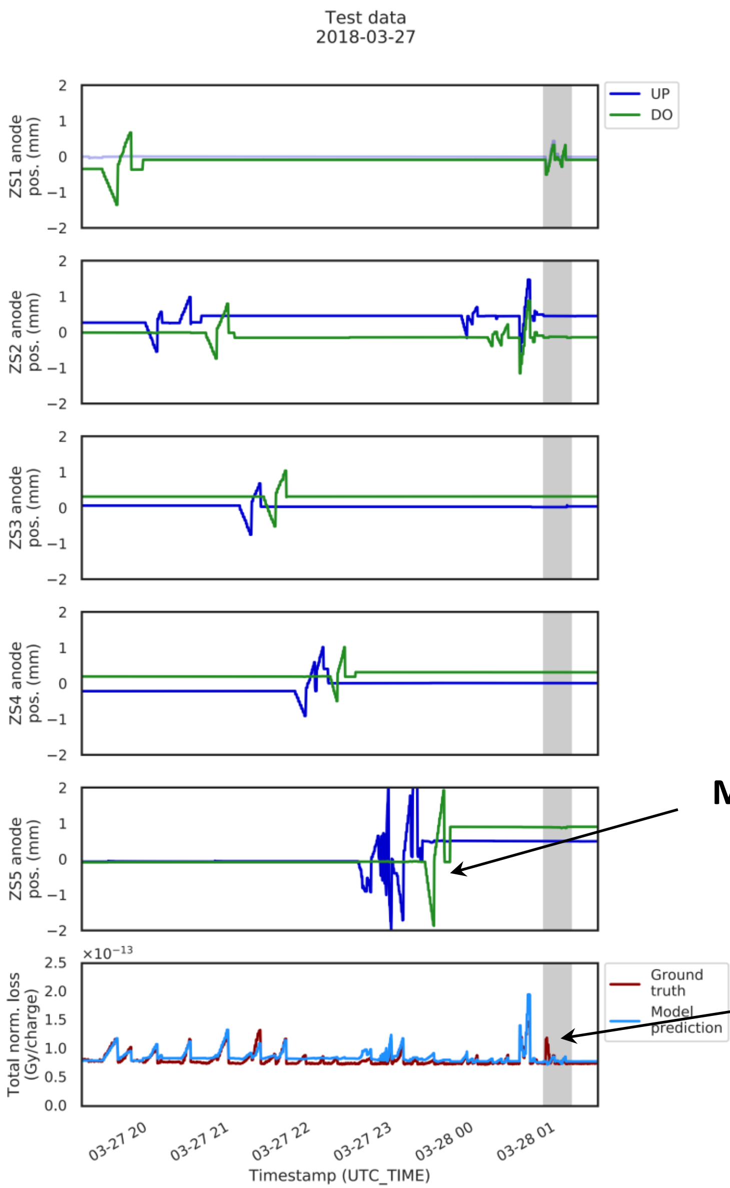
Backup

Other ongoing and upcoming projects

- **Gather knowledge and expertise in machine learning, reinforcement learning, and advanced numerical optimisers**
  - Bayesian optimisation
  - Reinforcement learning with NAF, Gaussian processes, and other algorithms
  - Explore for best sample efficiency, robustness, etc.
  - Use self-made OpenAI gym environments (target steering model, ZS model)
- **Electron cooling in LEIR**
  - Build surrogate model based on simulation code
  - Analysis of Schottky spectra: convolutional neural networks, autoencoders?
  - Provide operational tool
- **SPS slow extraction**
  - Model for hysteresis of main magnets in SPS
  - Reinforcement learning for spill optimisation?
- **Optimisation of transition crossing in the SPS**
- **LINAC4 and AWAKE: beam matching**

Data-driven model

Case I: Test set predictions (total loss)



- **Good quantitative agreement,** given simplicity of model and low number of samples
- Training and test set separated by **7 months**
- Offset in 'baseline' & some peaks only **qualitative agreement**
  - Missing feature (girder, orbit)?
  - See individual loss on BLM ZS5 *(backup)*
- Another source of discrepancy is **range in anode positions**
  - Training set: ≈ ± 0.5 - 1 mm
  - Test set: up to ≈ ± 2 mm

Moving in range [-2, 2] mm

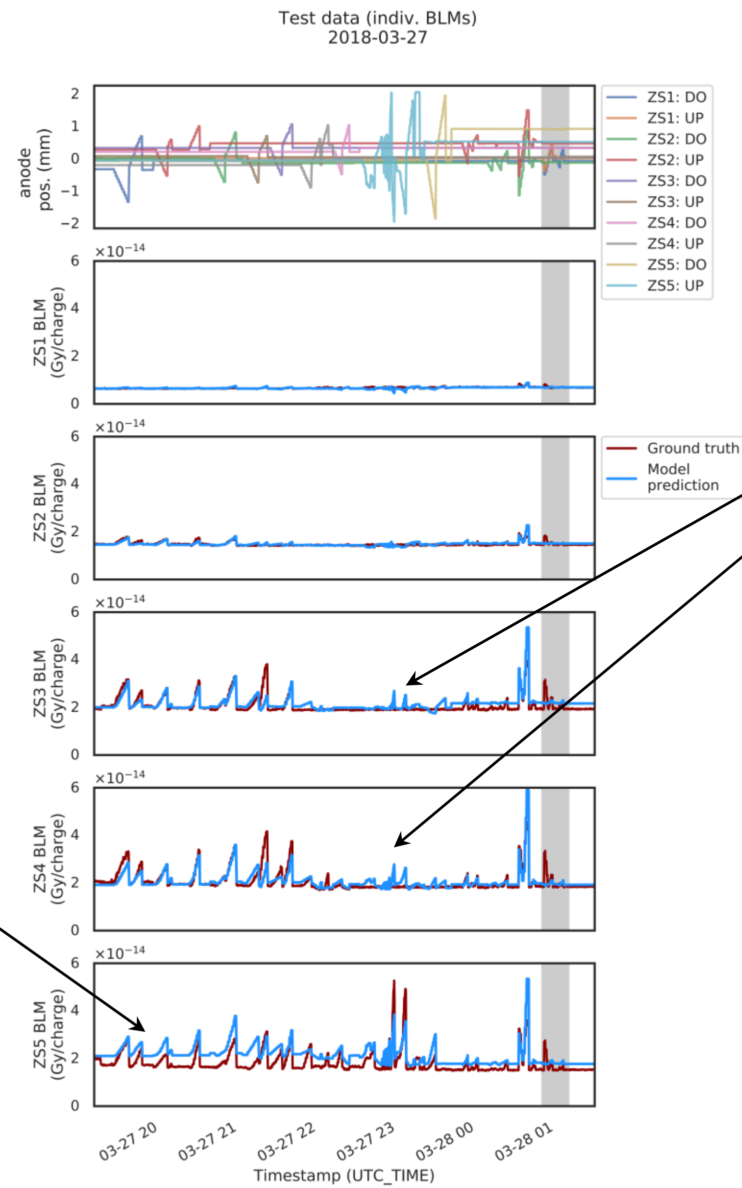
- Not trained for that (extrapolation)
- But: ZS5 no big impact on total loss, hence prediction still OK

Peak not predicted

- Moving ZS1 UP
- Not trained for that

Data-driven model

Case I: Test set predictions (individual BLMs)



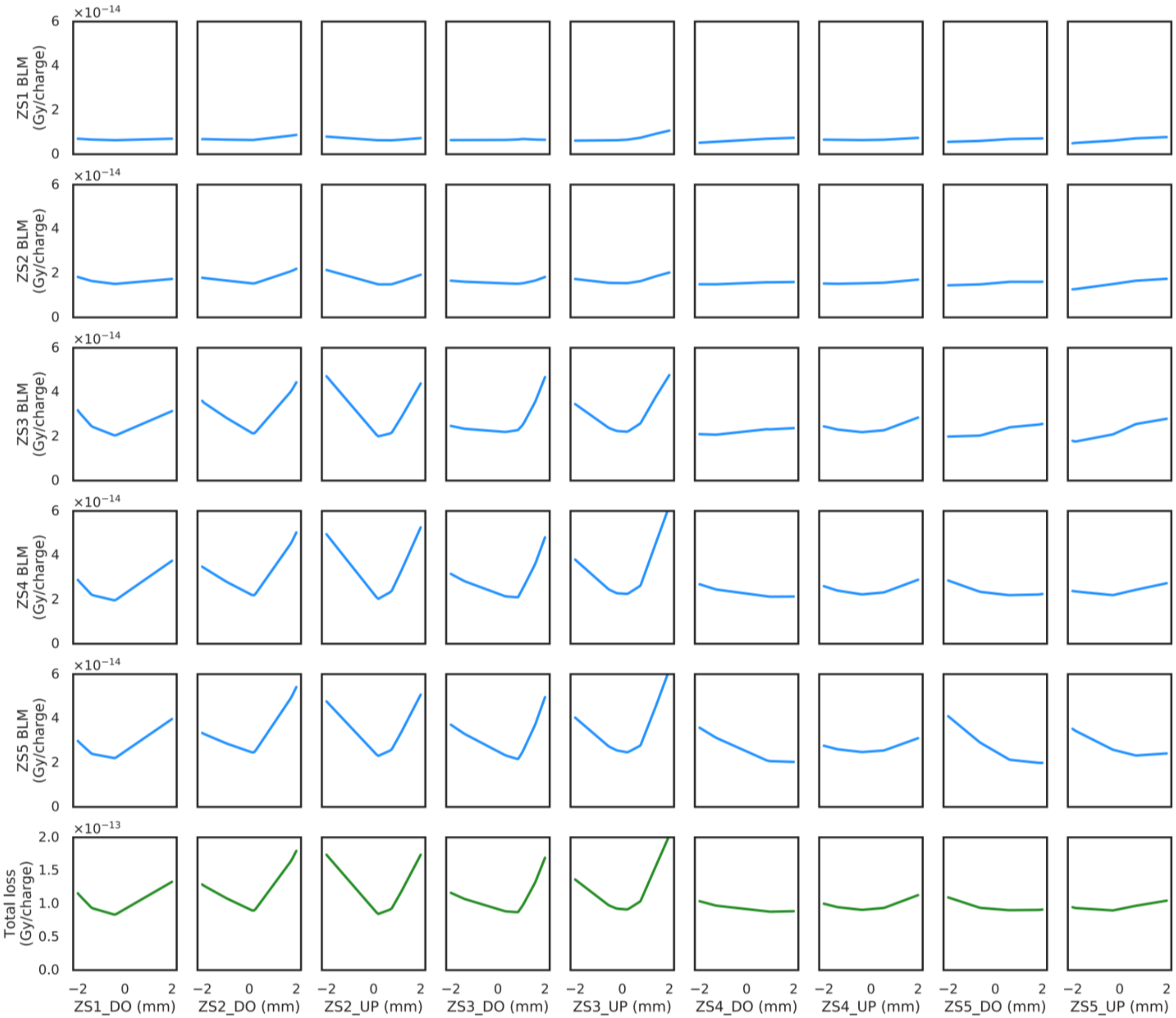
Overfitting artefacts?

Offsets on BLM ZS5

Possibly due to changed settings on features that are not included in the model (orbit, girder, cathode voltage, etc.)

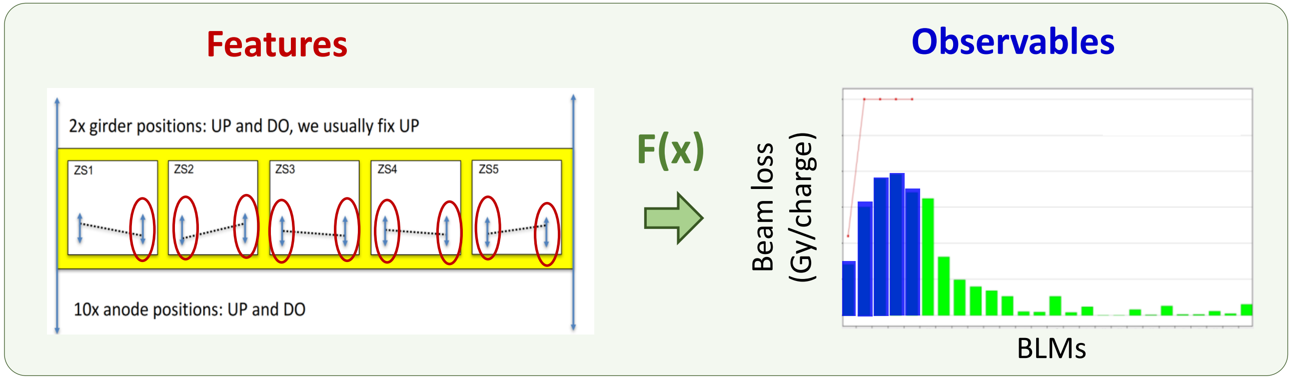
Data-driven model

Case I: Tests with 'fake' anode scans



- Orthogonal scans reasonable: *convex shapes*
- No strong impact on total loss from ZS4 and ZS5, *as observed in the machine*
- Piecewise linear functions *due to simplicity of network*

Data-driven model
Overview



Start with simple model

- **9 features:** All anode positions *(except ZS1 UP)*, no girder positions
- **5 observables:** BLMs ZS1, ZS2, .., ZS5
- **Neural network:** feed-forward, dense, 1 hidden layer, 7 – 15 nodes, leaky ReLU activation
- Adam optimizer
- *Missing features:* girders, orbit, cathode voltage, ...

Training and test sets from existing data

I. *Training on data from Powell scan (01.11.18, 467 samples)*

=> Predict manual scan 27.03.19 (test set)

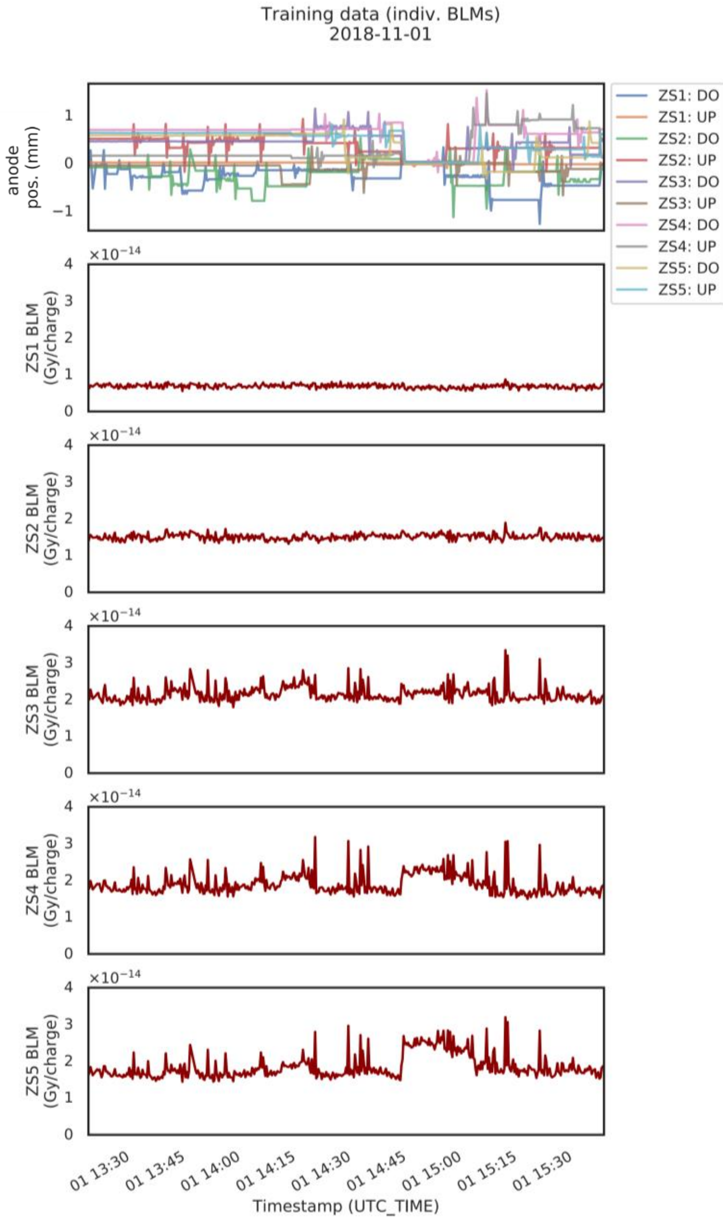
=> Perform 'fake' scans for all anode positions

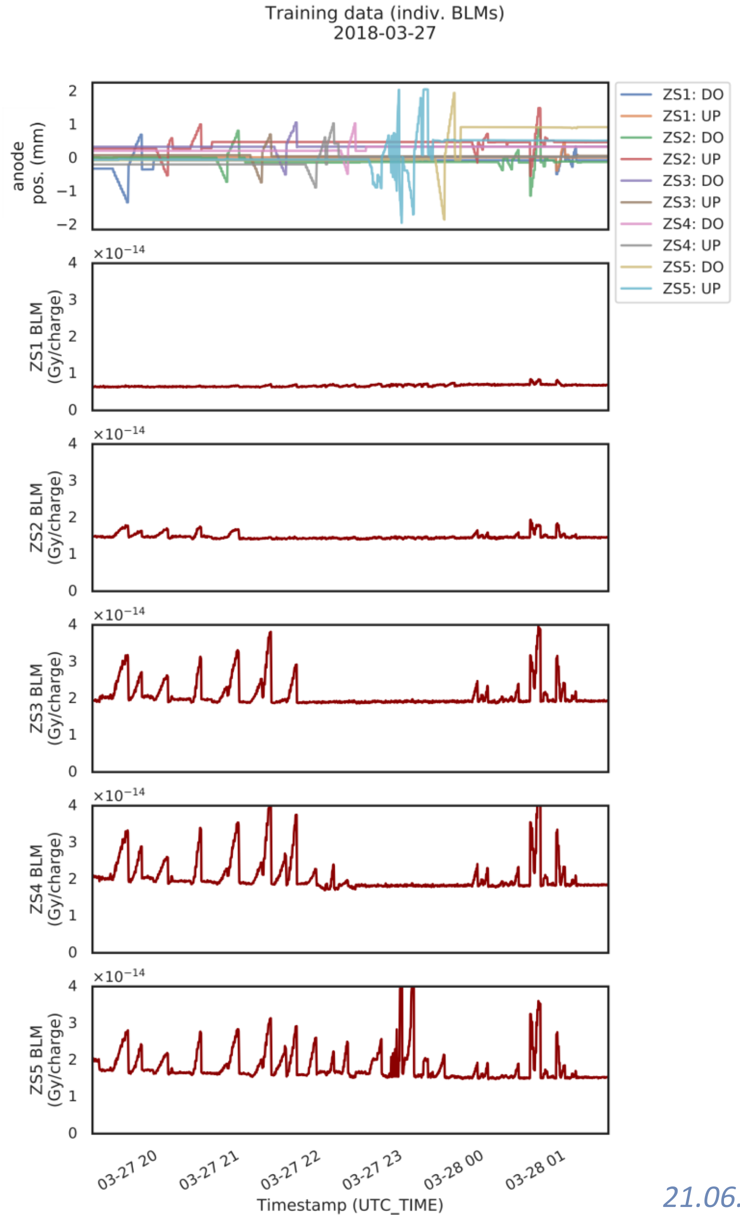
II. *Training on data from Powell scan and manual scan (27.03.19, 1308 samples)*

=> Predict manual scan 30.03.19 (test set)

=> Perform 'fake' anode scans as above

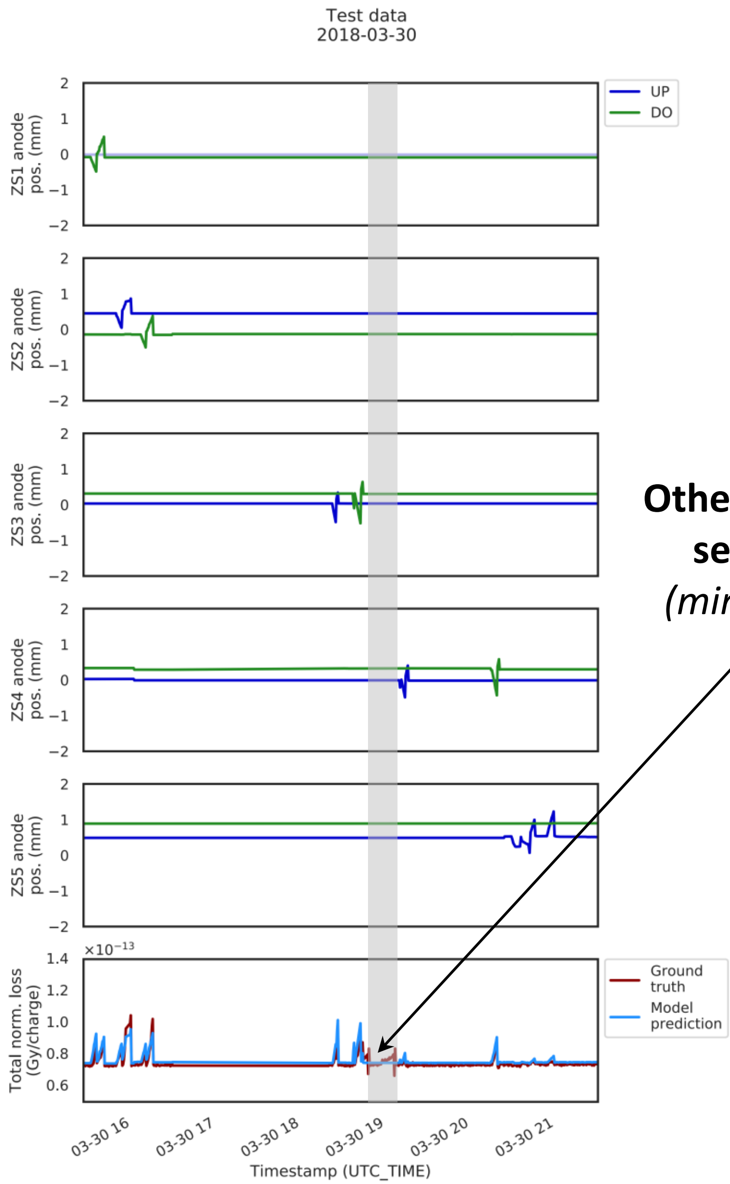
Data-driven model
Case II: Training data (15-node neural network)



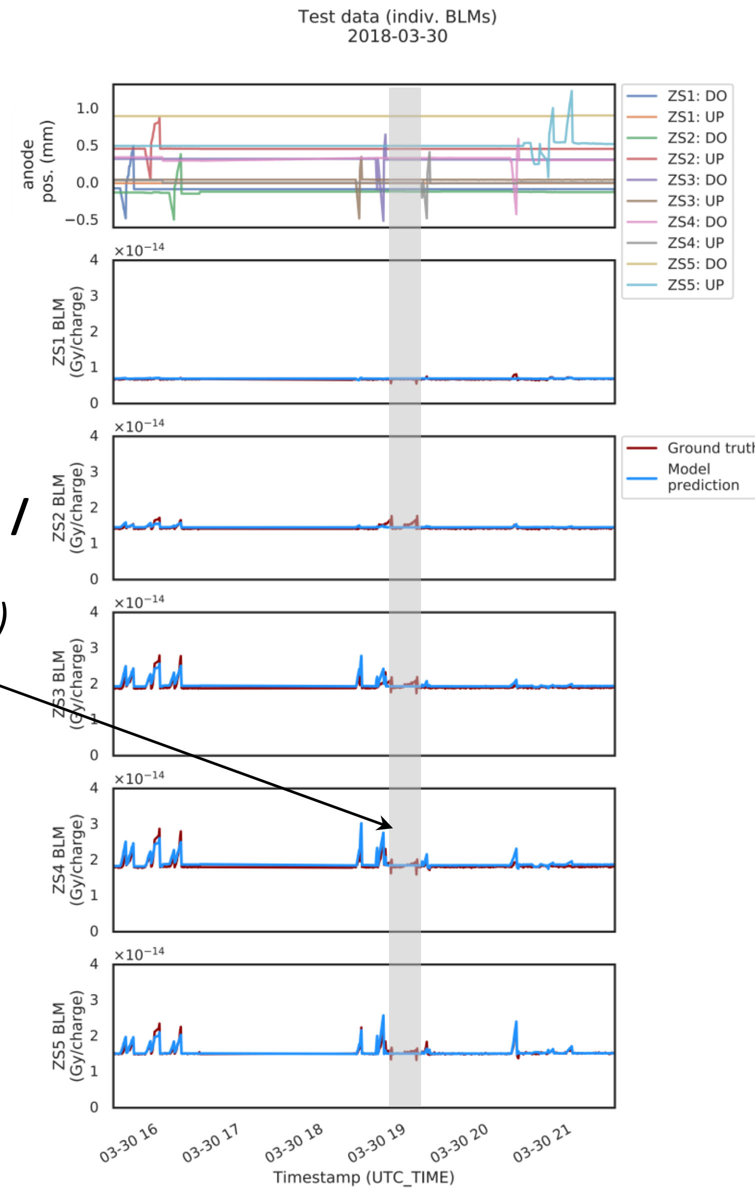


Data-driven model

Case II: Test set predictions (on 'independent' data set)

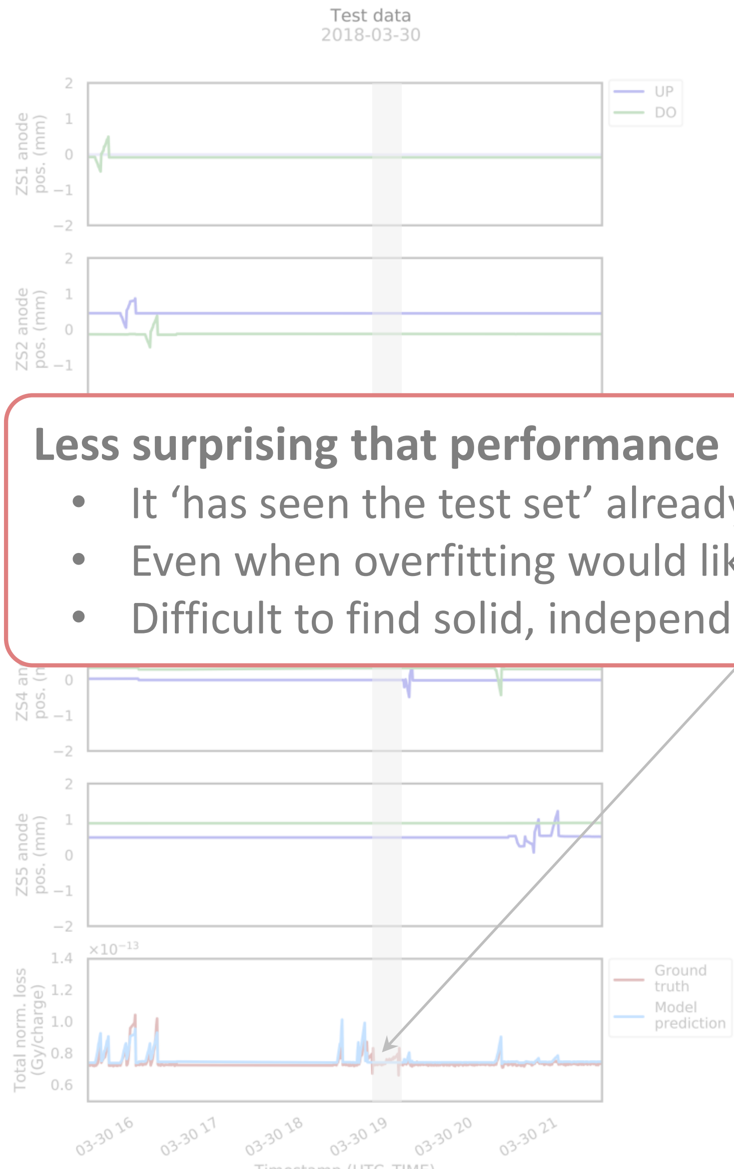


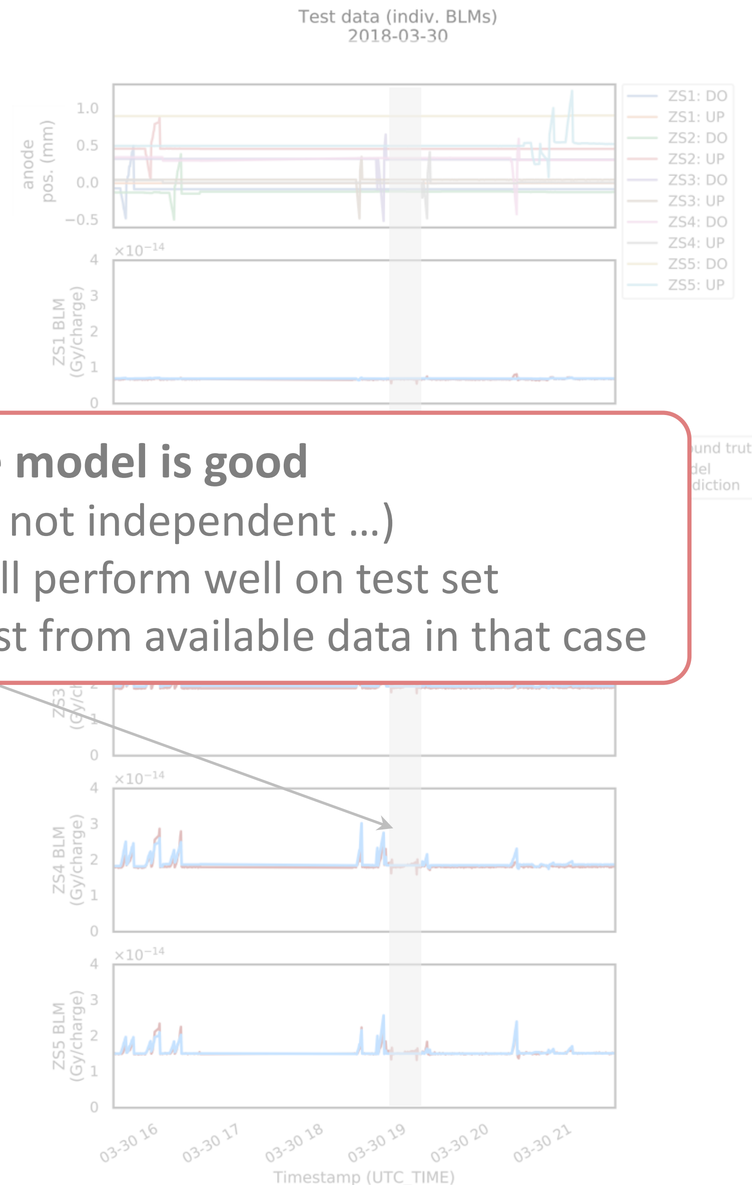
Other scans / settings
(miniscans)



Data-driven model

Case II: Test set predictions (on 'independent' data set)





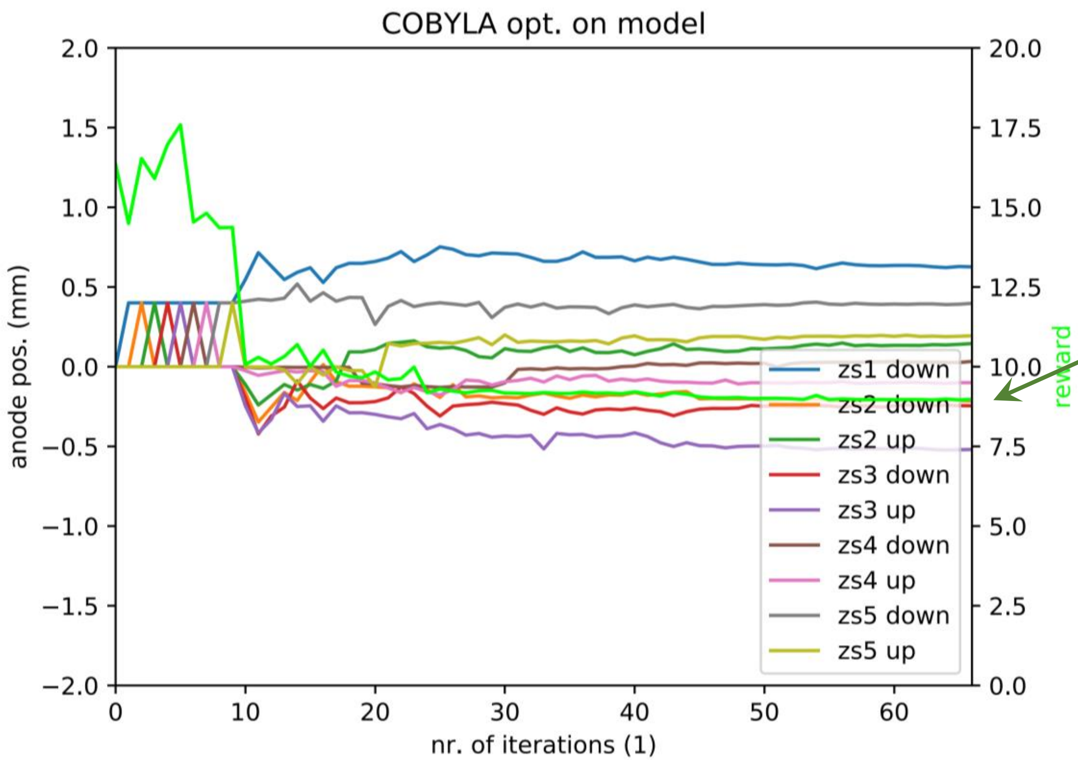
Less surprising that performance of the model is good

- It 'has seen the test set' already (sets not independent …)
- Even when overfitting would likely still perform well on test set
- Difficult to find solid, independent test from available data in that case

Data-driven model

Validation: COBYLA performance on trained model (7-node NN)

- **COBYLA with constraints on anode positions to ± 2 mm:**
  Optimisation on NN model to check if it produces sensible output
- Not clear yet how NN model behaves outside 'trained range' …
- COBYLA or Powell algorithms probe feature space better, not just along orthogonal directions



- Reward = total loss x 10^{14}
- Optimum expected at 8×10^{-14} Gy/charge *(from data and NN model)*
- Even faster convergence than Powell – looks promising