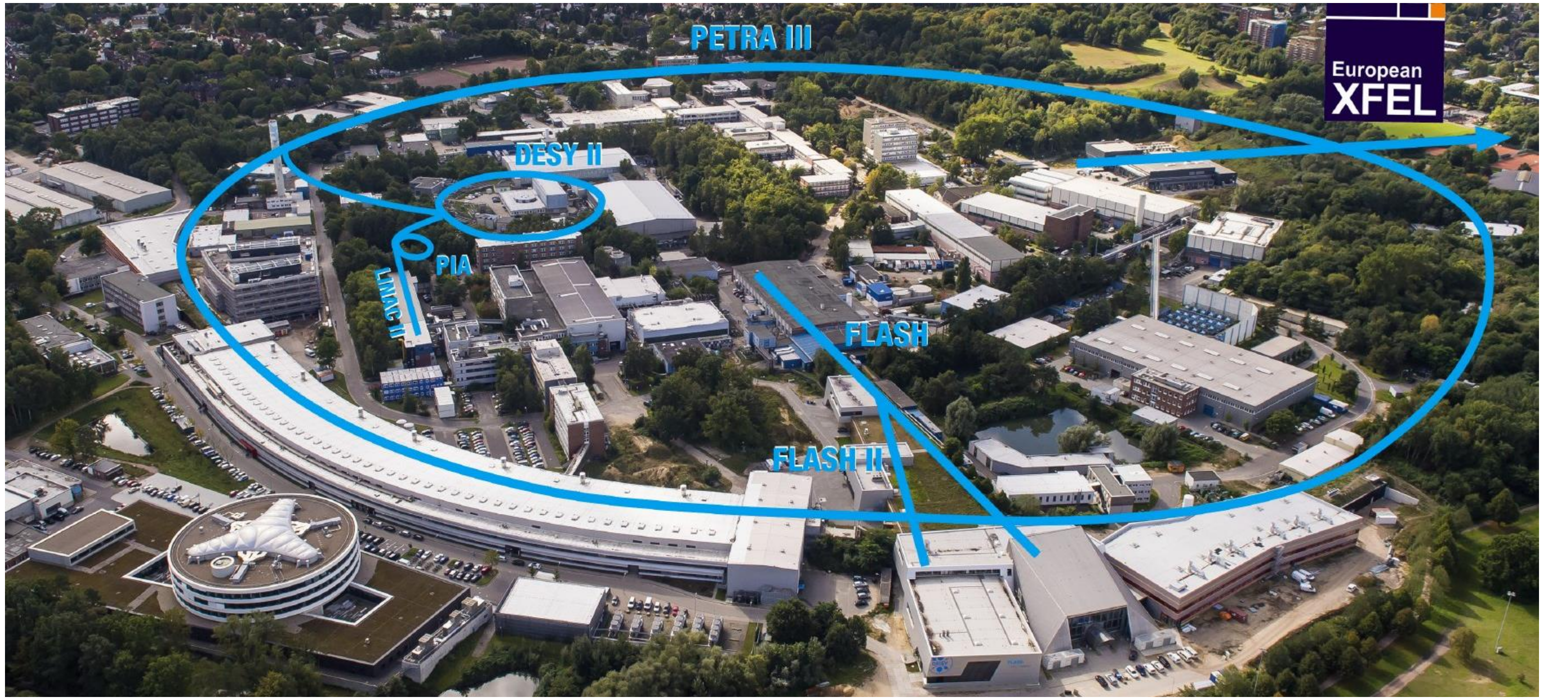
Neural Networks for ID Gap Orbit Distortion Compensation in PETRA III

ACCLAIM Meeting, July 3rd 2023, Jena

Bianca Veglia, DESY MPY

PETRA III.

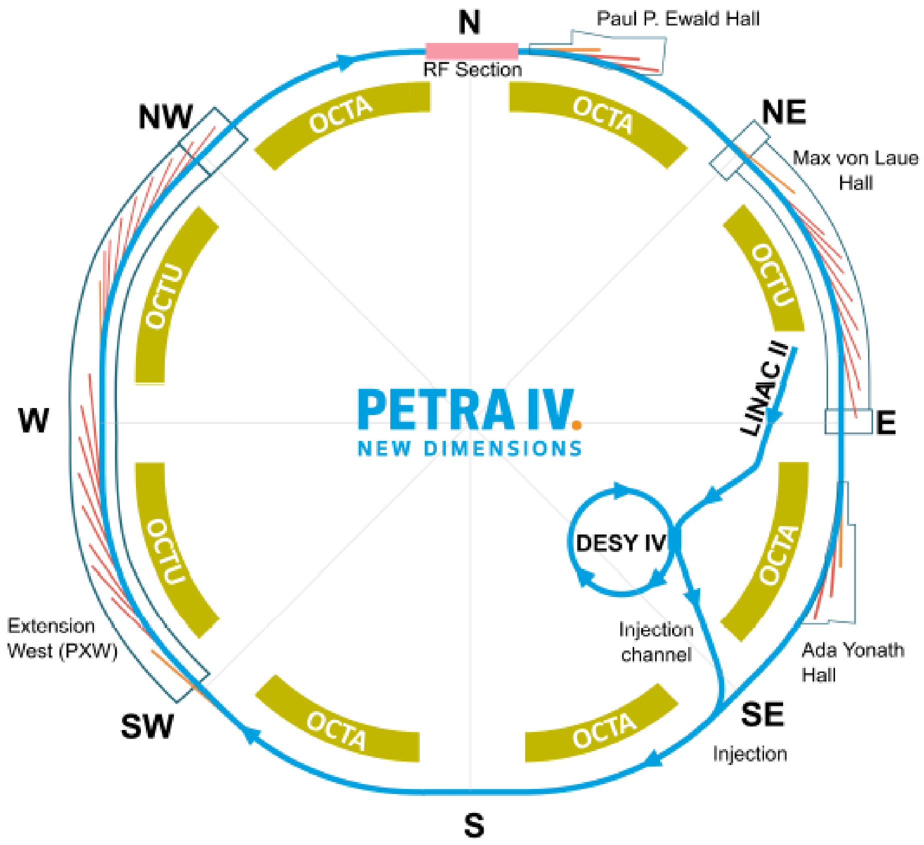


Each year, more than 2000 users are performing measurements at the PETRA III beamlines.

- High brilliance 3rd Generation Synchrotron Radiation Source.
- Extremely low emittances.
- 25 beamlines.
- Hybrid lattice with FODO and DBA (Double Bend Achromat) cells.

Parameter	PETRA III
Energy /GeV	6
Circumference /m	2304
Emittance (hor. / vert.) /nm	1.3 / 0.012
Total current / mA	100

Parameter	H6BA
Tunes ν_x, ν_y	135.8, 86.27
Natural chromaticity ξ_x, ξ_y	-233, -156
Momentum compaction α_c	3.3 10^{-5}
U_0 /MeV	100
Standard ID section /m	4.7 -4.9
Hor. Emittance w/o IDs (zero current) /pm	20
Hor. Emittance with IDs (zero current) /pm	20
Rel. energy spread with IDs (zero current)	0.9 10^{-3}
Beta at ID /m	$\beta_x = 2.2$ $\beta_y = 2.2$
RF Voltage 1st, 3rd /MV	8, 2.4

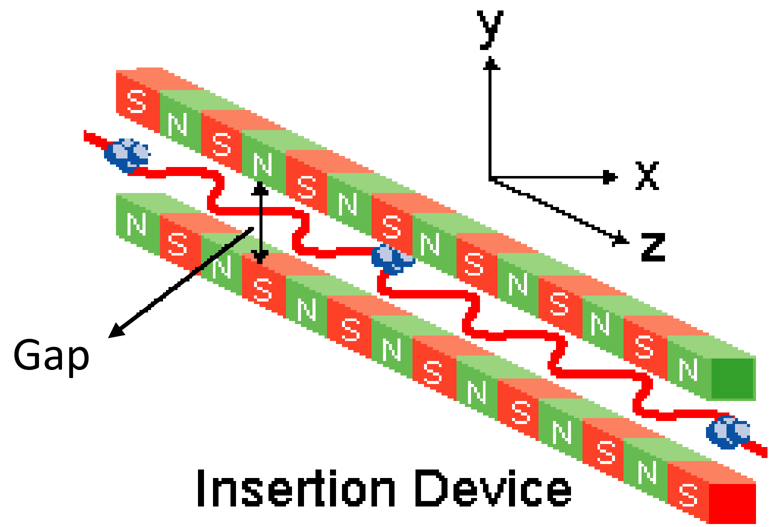


Up to **30 undulator** insertions (photon beam can be further split to allow more experimental stations).

- New high-resolution 3D X-ray microscope for chemical and physical processes.
- Construction within the existing PETRA ring tunnel.
- Nanometre scale for the first time.
- Ultra low emittances in the region of **10 pm**.
- Each of the eight arcs is composed of nine hybrid six-bend achromat (H6BA) cells.

IDs induce an orbit distortion which varies with the gap size

- The magnetic fields of IDs introduce perturbations to the circulating electron beam and hence affect the linear and nonlinear beam dynamics of the electron beam in the storage ring.
- Often users adjust the spectrum from undulators by changing undulator gap size. It's important to keep the orbit constant during these field changes to not disrupt other users.

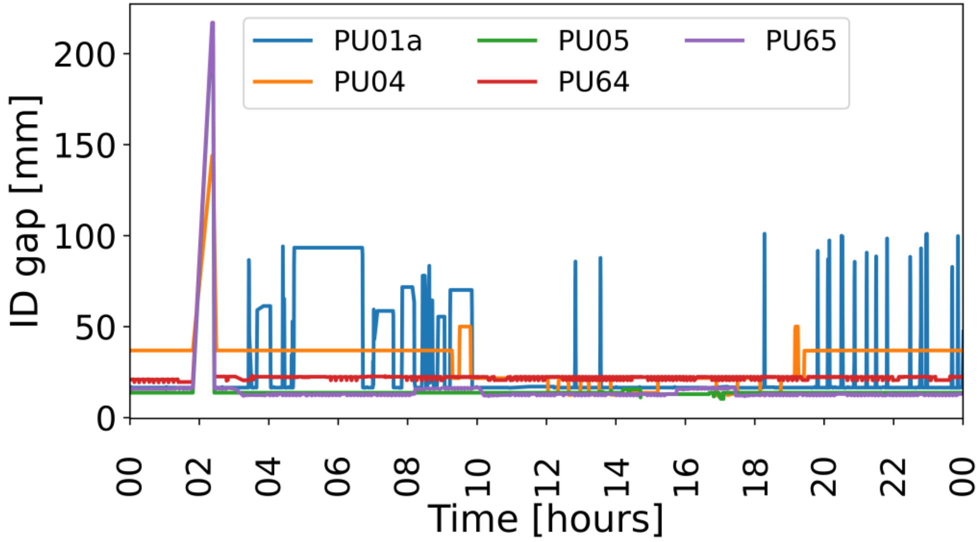


The field integrals determine the primary effect of the undulator on the electron beam orbit.

$$I_{1x} \equiv \int_{z_0}^{z_0+L} B_x(z_1)\, dz_1$$
$$I_{1y} \equiv \int_{z_0}^{z_0+L} B_y(z_1)\, dz_1$$
$$I_{2x} \equiv \int_{z_0}^{z_0+L} \int_{z_0}^{z_2} B_x(z_1)\, dz_1\, dz_2$$
$$I_{2y} \equiv \int_{z_0}^{z_0+L} \int_{z_0}^{z_2} B_y(z_1)\, dz_1\, dz_2$$

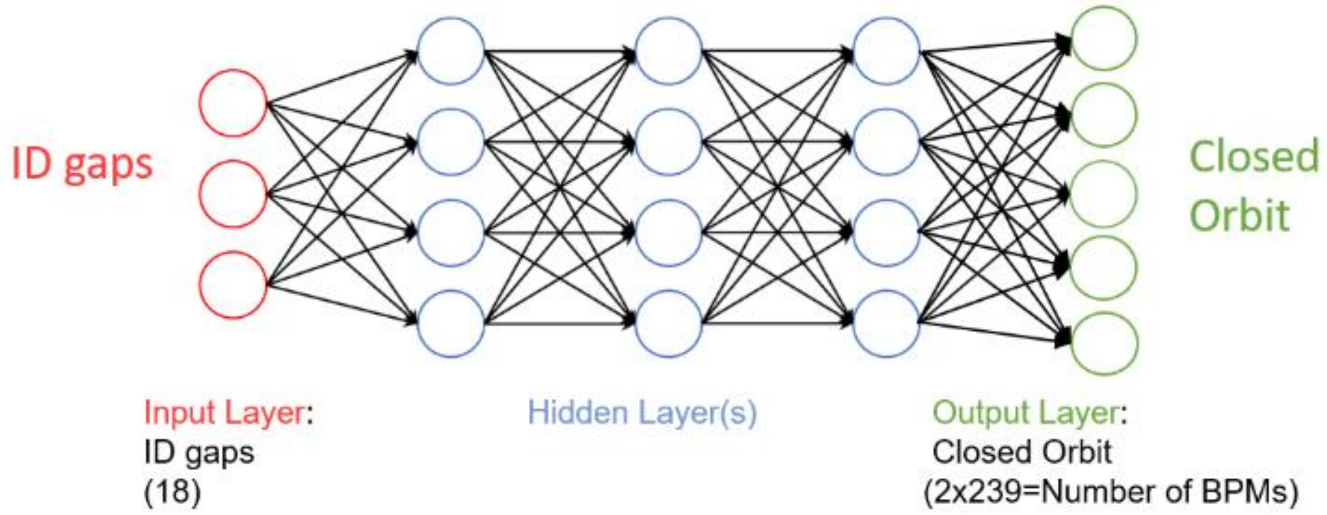
$$\Rightarrow$$

$$x'_{exit} = -\frac{q}{\gamma m v_z} I_{1y}$$
$$y'_{exit} = \frac{q}{\gamma m v_z} I_{1x}$$
$$x_{exit} = -\frac{q}{\gamma m v_z} I_{2y}$$
$$y_{exit} = \frac{q}{\gamma m v_z} I_{2x}$$



Building the NNs

- The NN takes as input a vector containing the ID gap sizes and gives as output the predicted orbit at the location of each BPM.
- The different model are trained on 80% of the measurements took in July and validated vs the remaining 20%



Hyperparmeter sweeps performed with

Weights & Biases

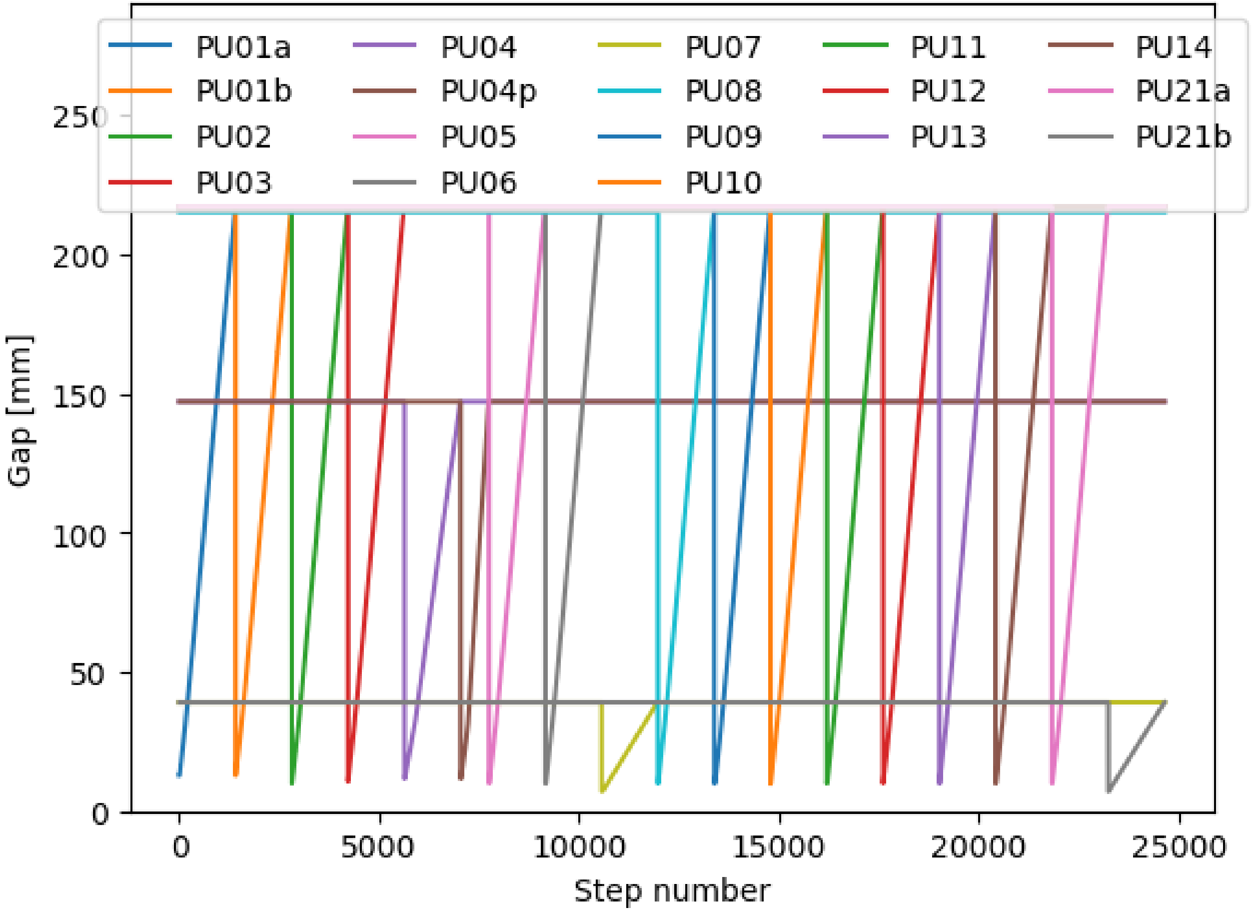
Keras

The models are trained using the back propagation method employing the Adam optimizer for 200 epochs (10-50 minutes).

Why using machine learning on this specific problem?

Because of its flexibility and ability to model also highly nonlinear processes. Measurements for feed forward systems are time consuming and need to be updated regularly.

Measurements



Horizontal and vertical orbit were taken varying the gap size in their operation range for 17 IDs in PETRA III.

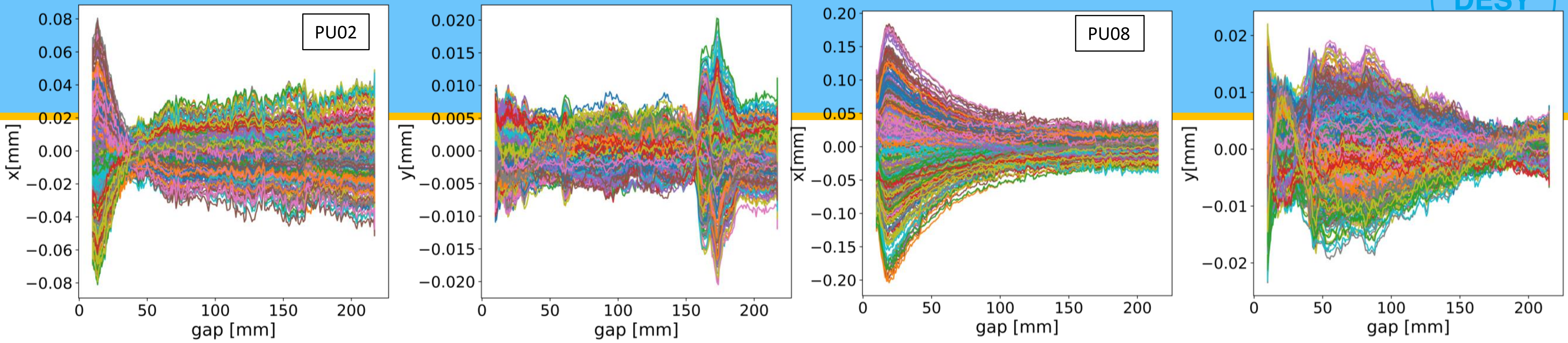
2 in-vacuum undulators (limited open gap).

During the movement of each one IDs the other were kept to the full opening gap (no impact on the orbit).

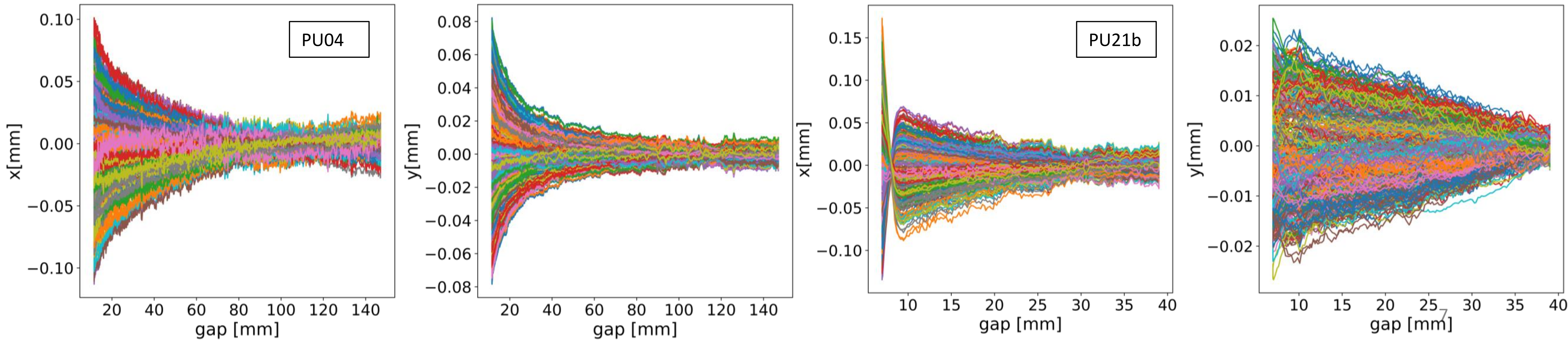
The correction coils and FOFB were disabled.

PU04 is an Apple-II type undulator, providing light with different polarizations.

Closed orbit distortion measurements

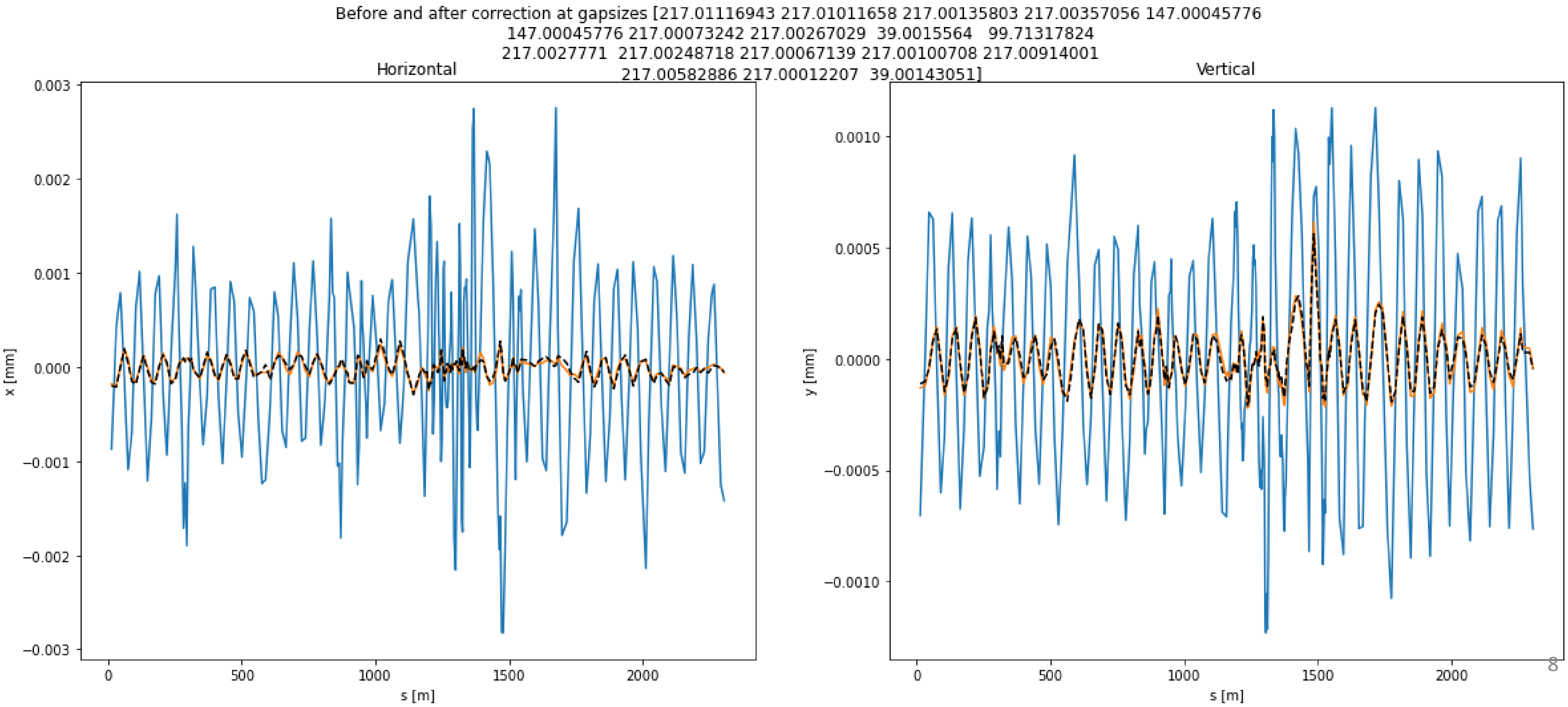


Measured orbit distortion while closing 4 undulators. Each colour represents a BPM along the ring.



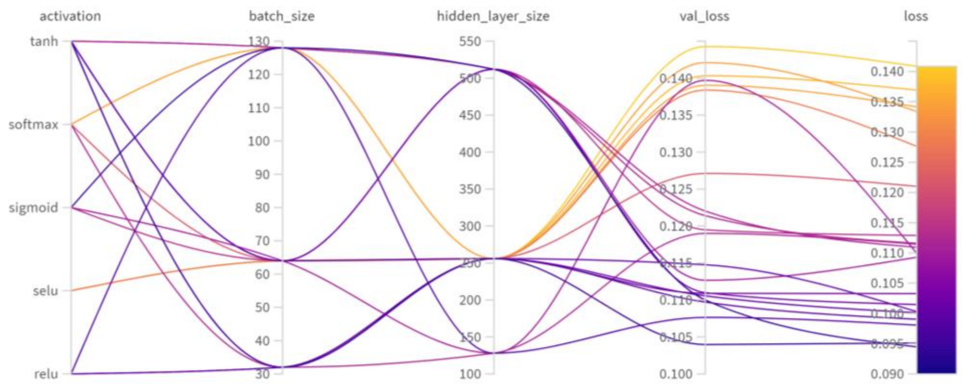
From prediction to correction

- Once the orbit is predicted, the strengths of the correctors are computed through SVD of ORM in pyAT.

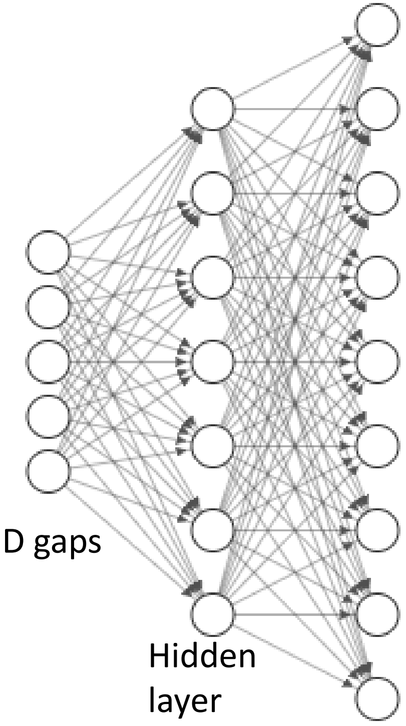


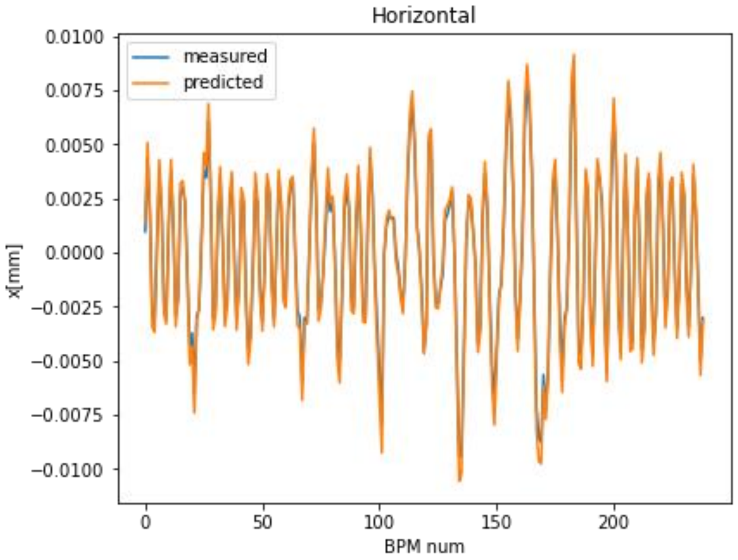
Shallow feed-forward fully connected NN

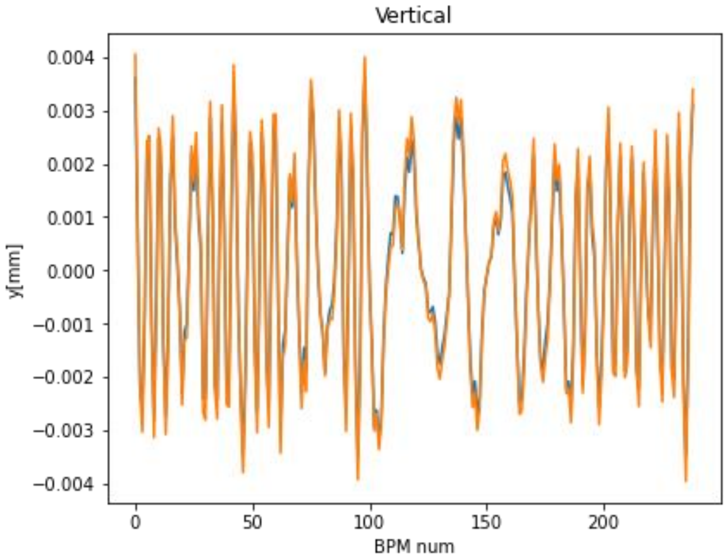
One hidden layer: exploring the impact of different activation functions and batch size.



Best result
Activation: tanh
Batch size: 32
Layer size: 256

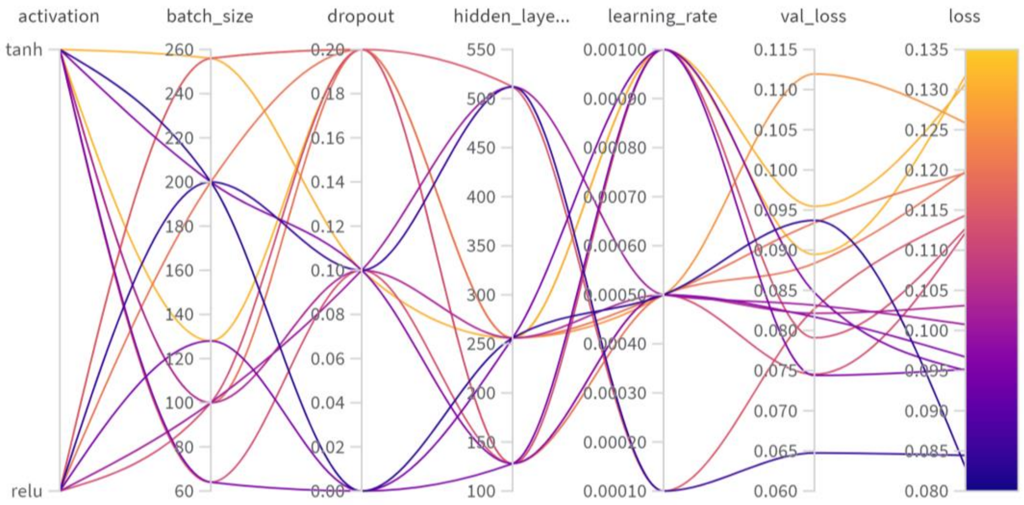






Deep feed-forward fully connected NN

Multiple hidden layers: exploring the impact of dropout, batch and layer size and learning rate.



<u>Best result</u>

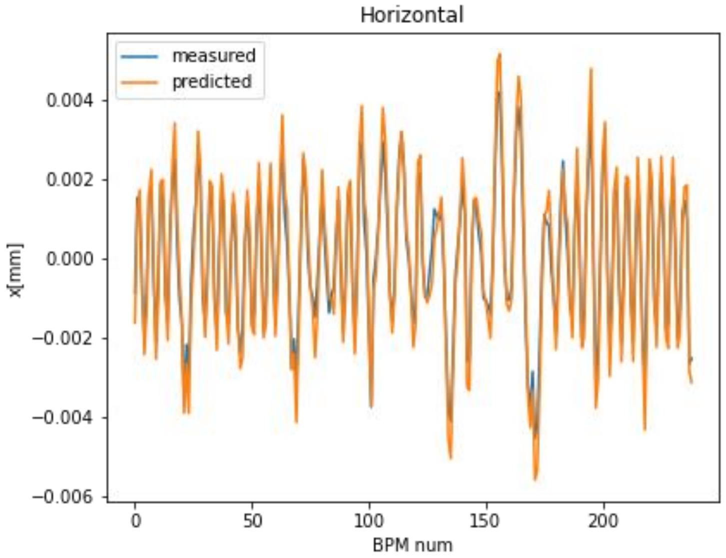
Activation: relu

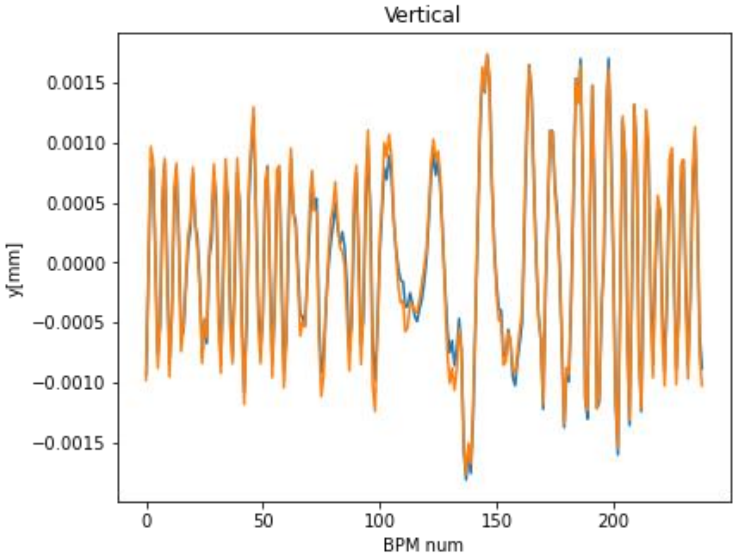
Batch size: 200

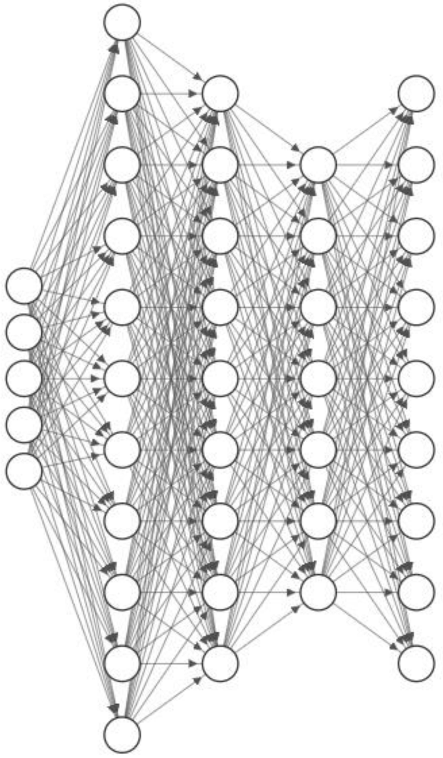
Dropout: 0.1

Layer size: 256

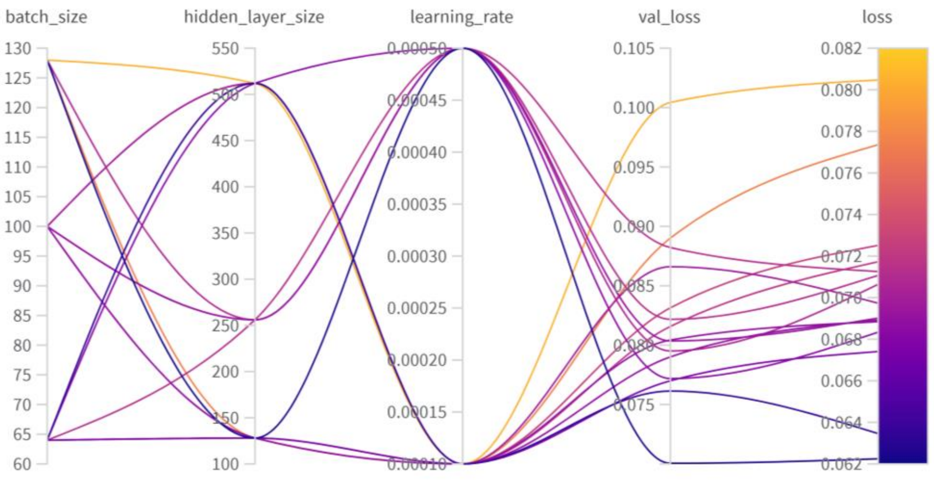
Learning rate: 10^{-4}







Recurrent NN



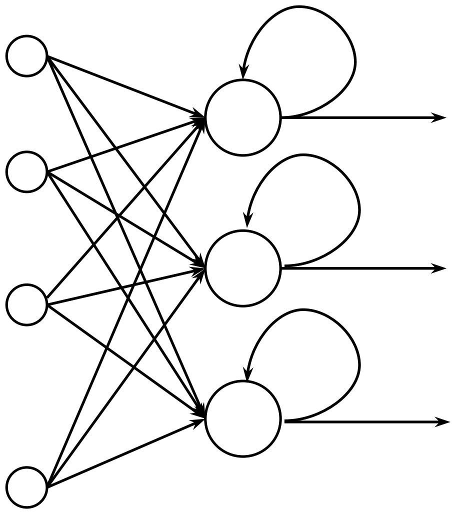
In a RNN the information cycles through a loop. When it makes a decision, it considers the current input and also what it has learned from the inputs it received previously.

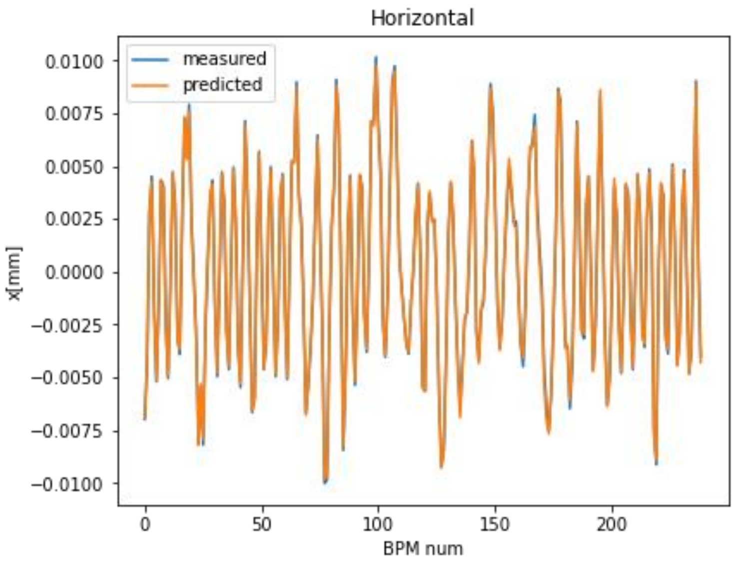
<u>Best result</u>

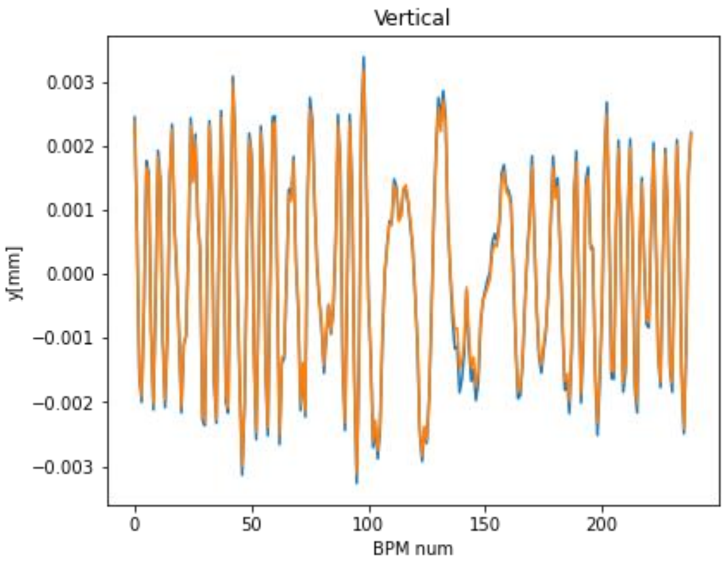
Batch size: 128

Layer size: 128

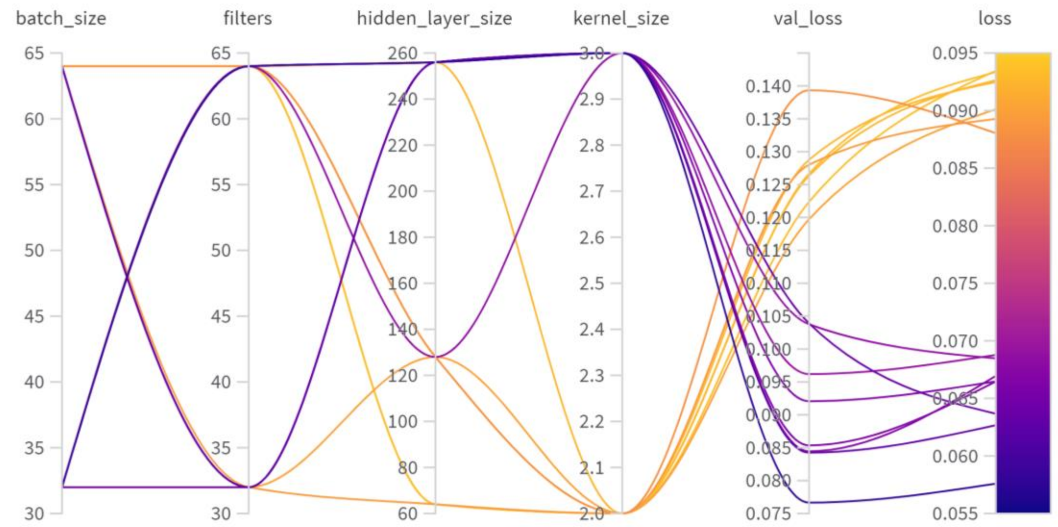
Learning rate: 5×10^{-4}





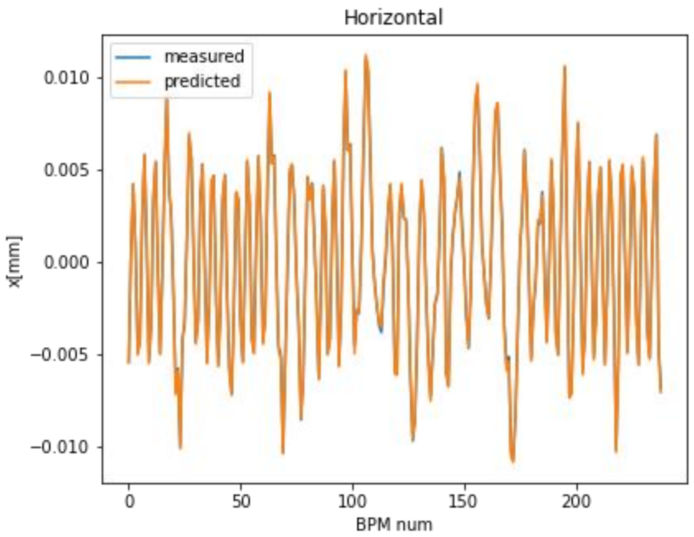


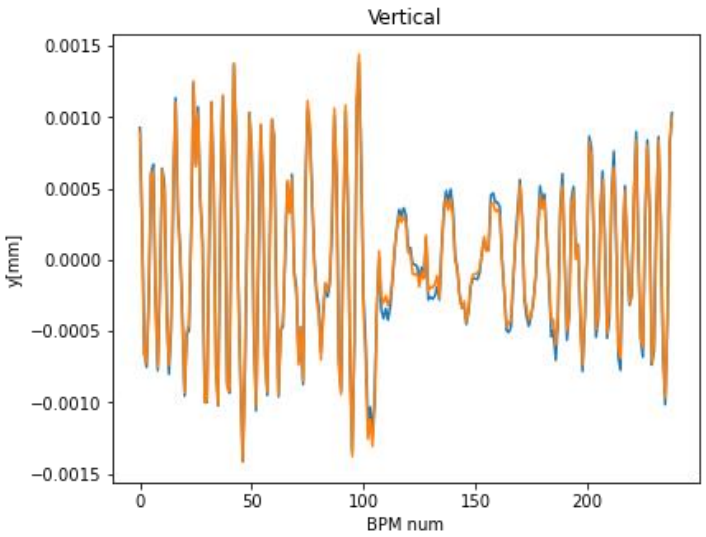
1D Convolutional NN



A convolution layer systematically apply learned filters to input in order to extract features.

The kernel is a matrix (in this case 1D) of weights which are multiplied with the input to extract relevant features.





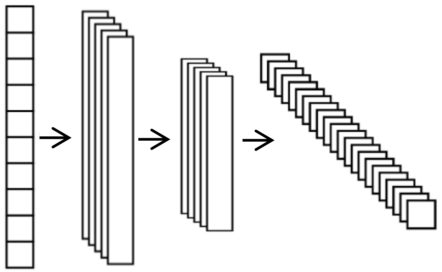
<u>Best result</u>

Batch size: 32

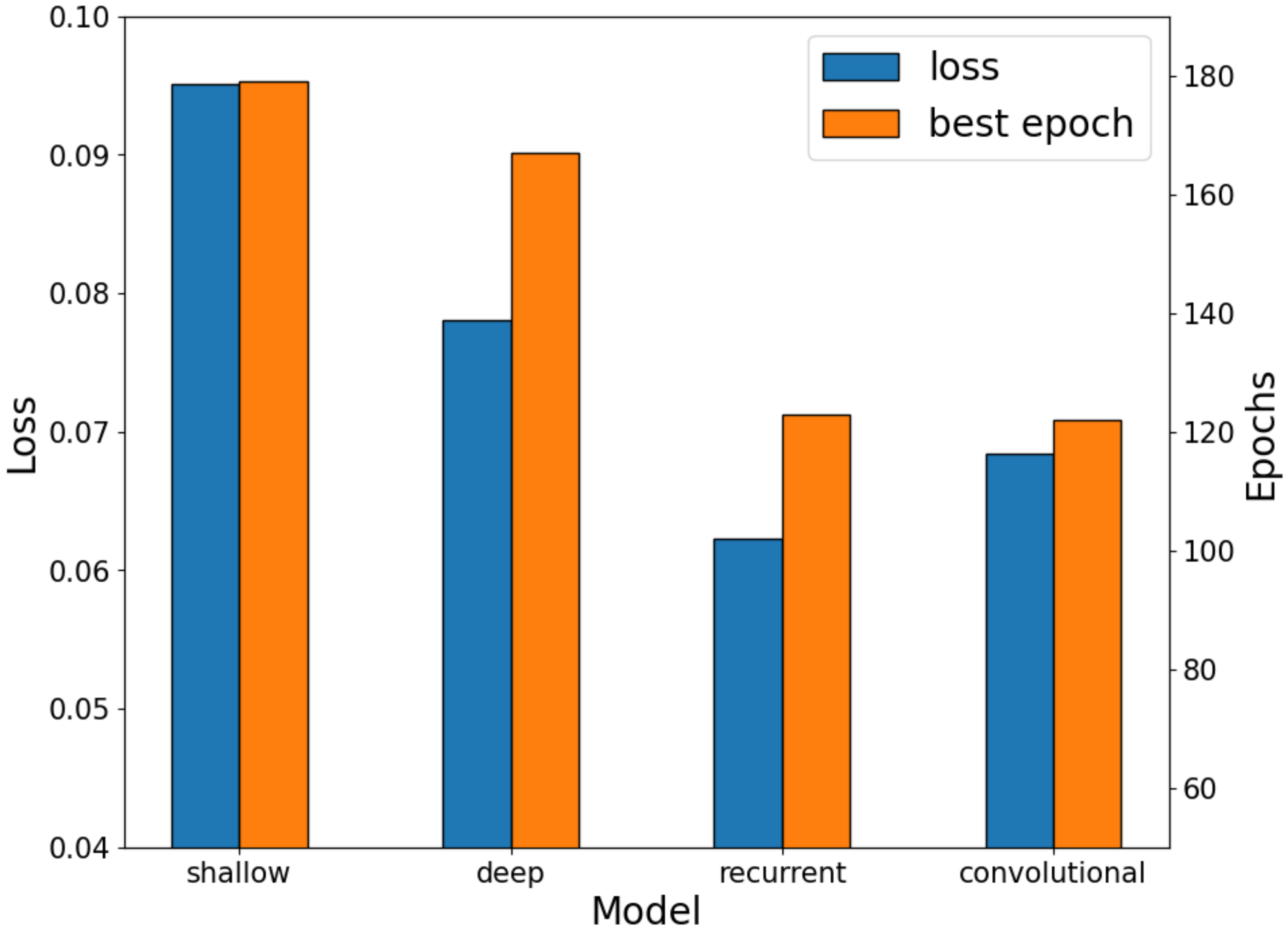
Filters: 32

Layer size: 256

Kernel size: 3



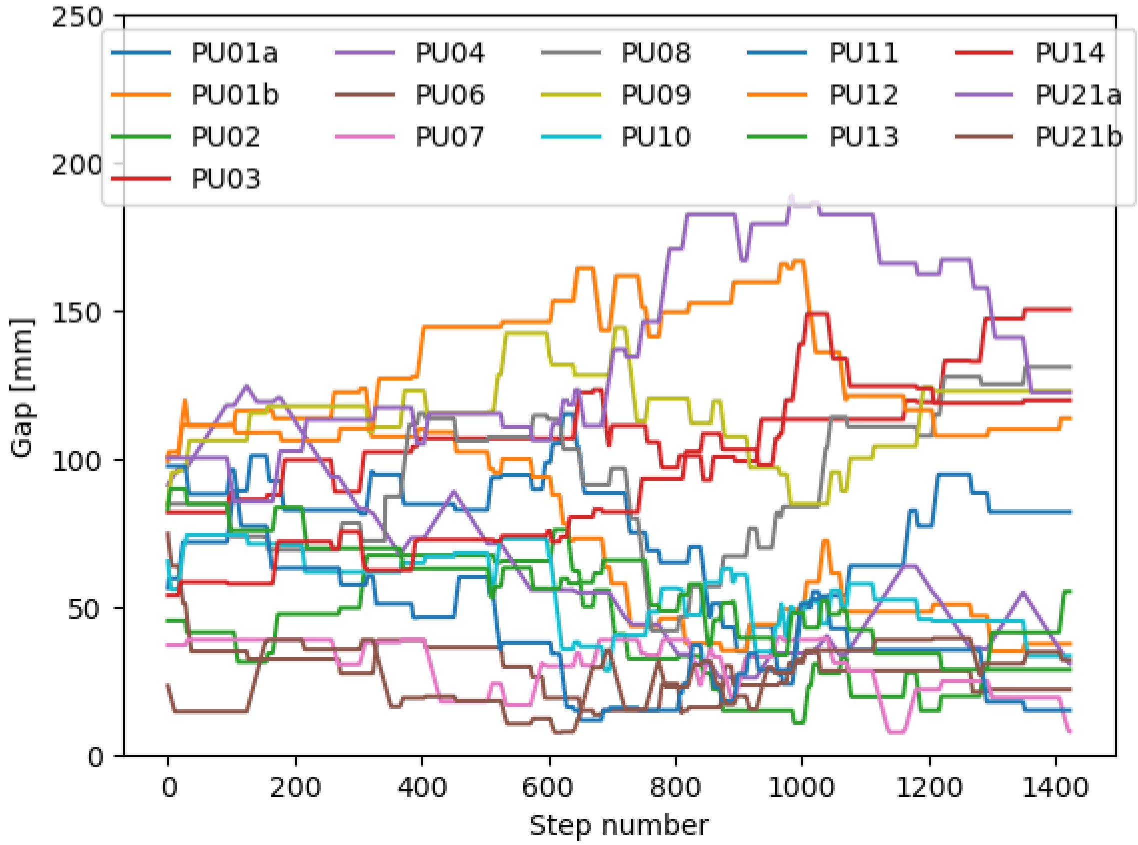
Comparing the architectures



The convolutional and recurrent structures outperform the fully connected NN in a reasonable amount of epochs.

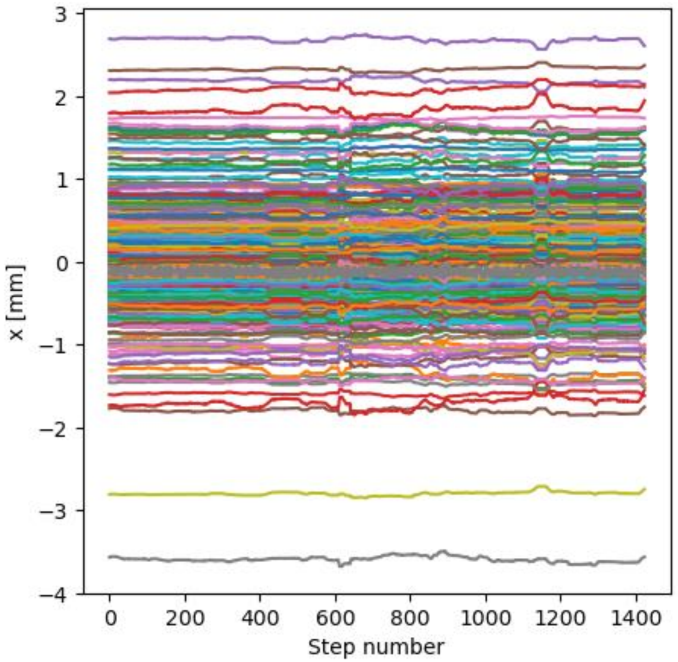
This is due to the format of the measurements used for training, where the gap size was recursively increased and decreased for each undulators. The input data exhibited a serial structure that is better modelled by NNs containing recurrent or convolutional features.

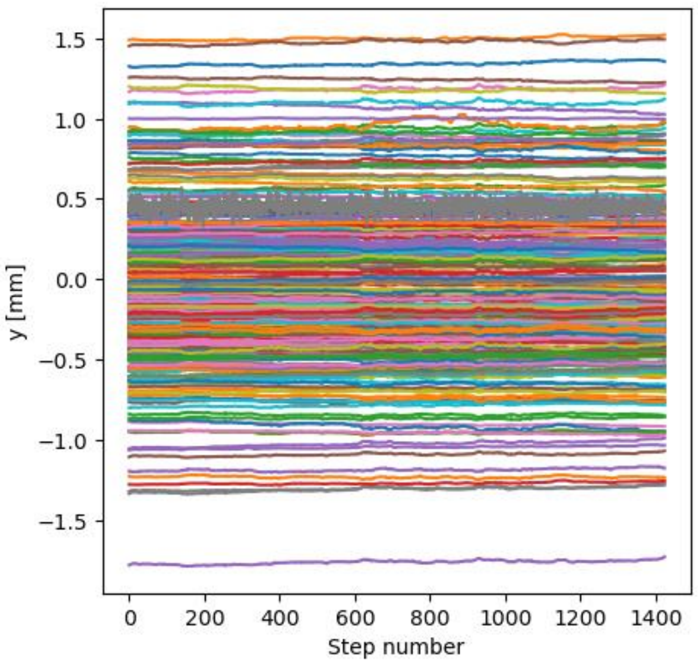
Additional measurements



Gap variations limited by the orbit interlock system.

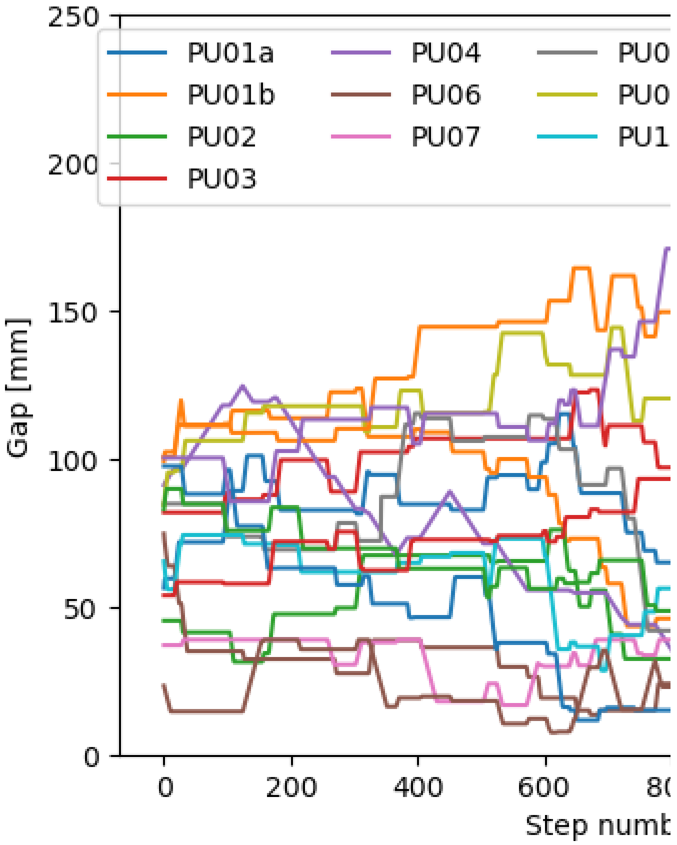
More realistic IDs movement patterns (many closing at the same time.

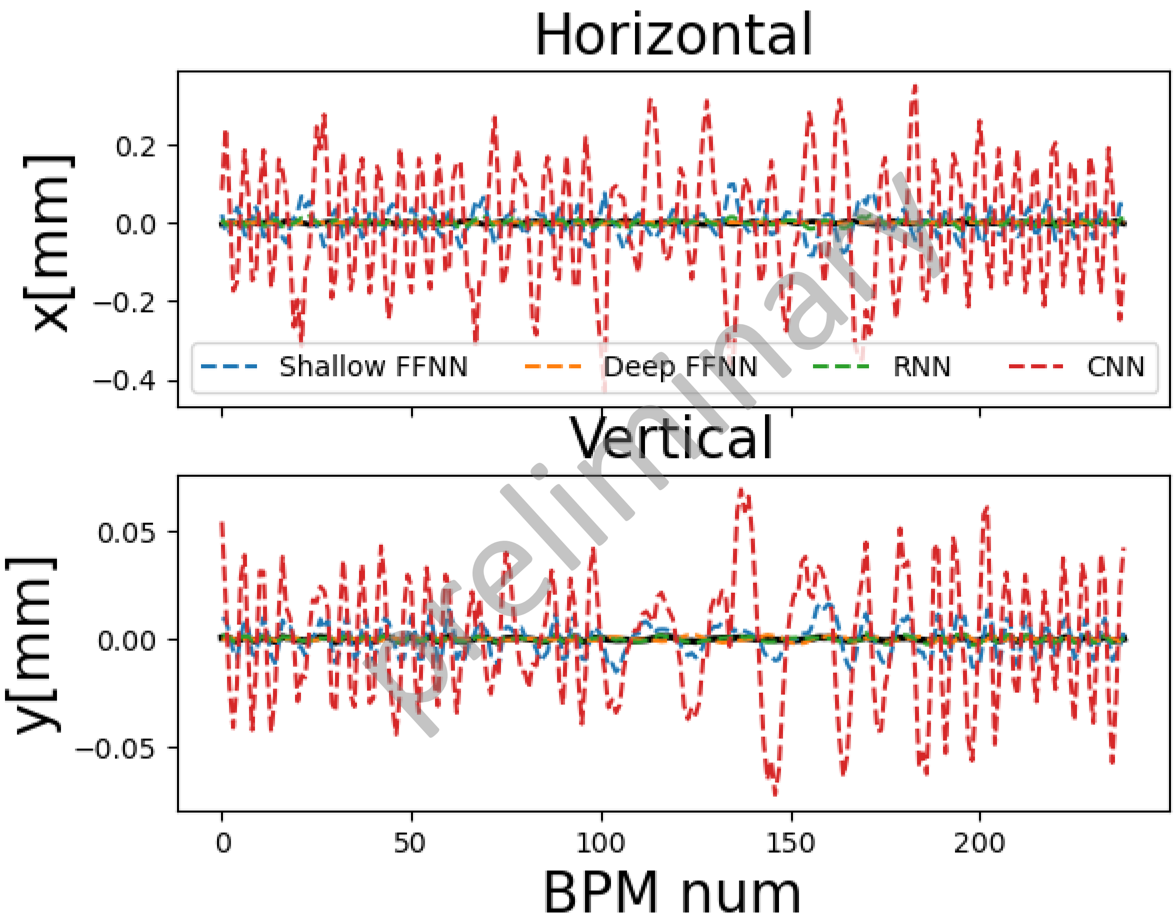




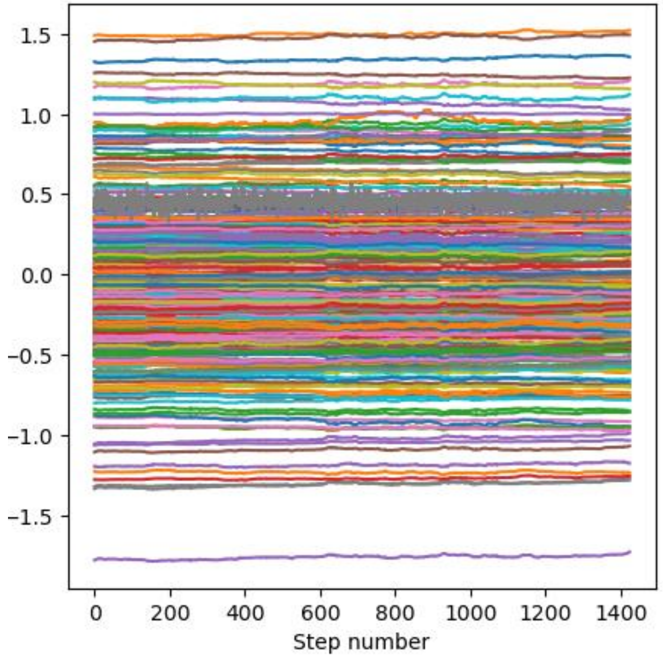
- Can the NNs make accurate predictions?
- 17 IDs (vs 20) suffice to train the networks?

Additional measurements





terns (many closing at



te predictions?

- 17 IDs (vs 20) suffice to train the networks?

Gap variations limited by the orbit interlock system.

Summary and conclusions

- The varying gap size of the IDs impact the circulating beam dynamics. One major effect is orbit distortions that need to be compensated.
- Neural networks were trained on PETRA III measurements to learn the correlation between arbitrary ID configurations and the orbit.
- Different NN architecture models were tested and compared. The Recurrent and Convolutional NN structure showed better predictivity.
- Additional measurements taken with multiple IDs closing at the same time.
- <u>Global correction</u>: through the ORM it is possible to compute the kick at the correctors along the ring necessary to counteract the distortion.
- <u>Local correction</u>: use the trained NN to map the ID movements to the current of the compensation coils at each end of the undulators (through field integrals)

> Tests of the implementation of the related compensation schemes planned in the future.

- A similar approach could also be considered to counteract the perturbation introduced by ID gap variations to the betatron coupling and the vertical dispersion that could impact the extremely low emittances of PETRA IV .

Thank you

Contact

DESY. Deutsches Elektronen-Synchrotron

www.desy.de

Bianca Veglia

MPY

bianca.veglia@desy.de