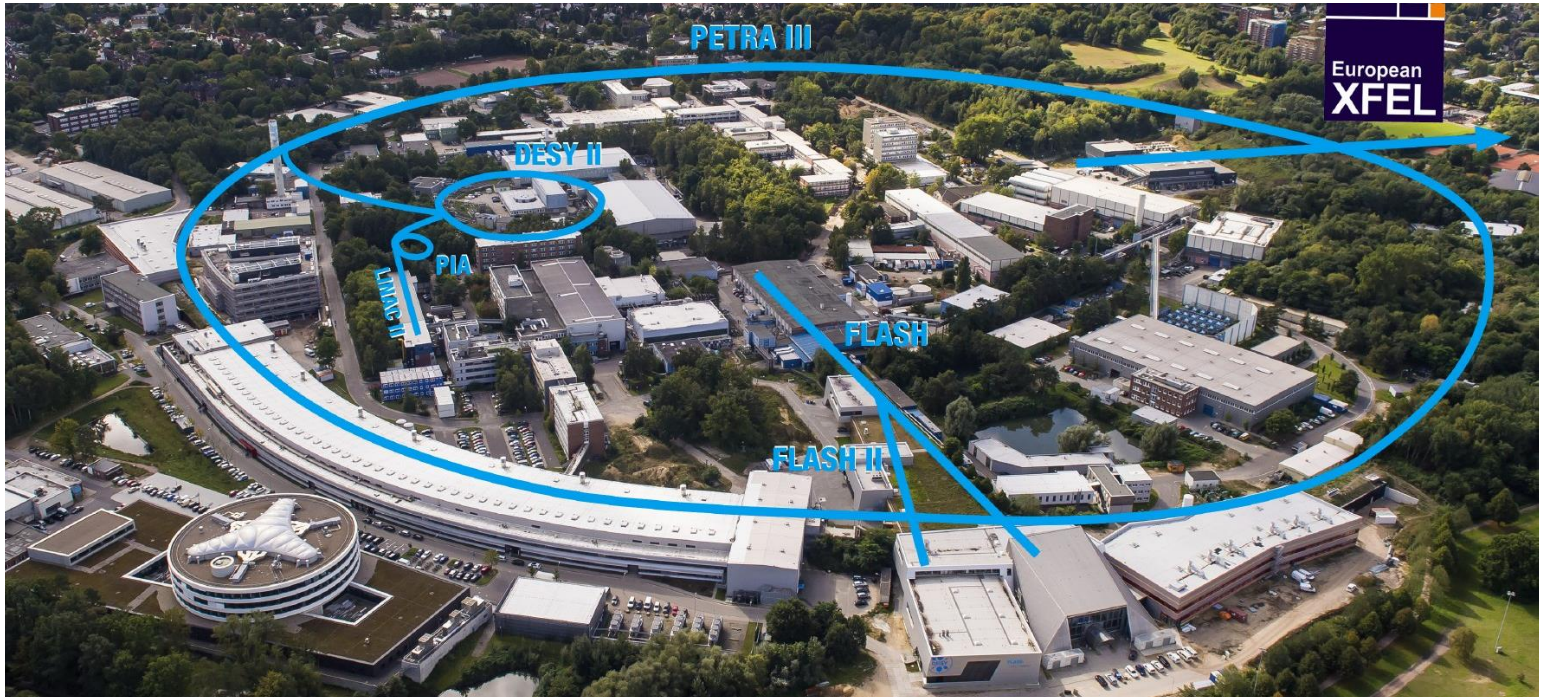
Machine Learning-based ID gap compensation scheme for PETRA III

3rd ICFA Beam Dynamics Mini-Workshop on
Machine Learning Applications for Particle Accelerators

HELMHOLTZ RESEARCH FOR GRAND CHALLENGES

DESY.

PETRA III.



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

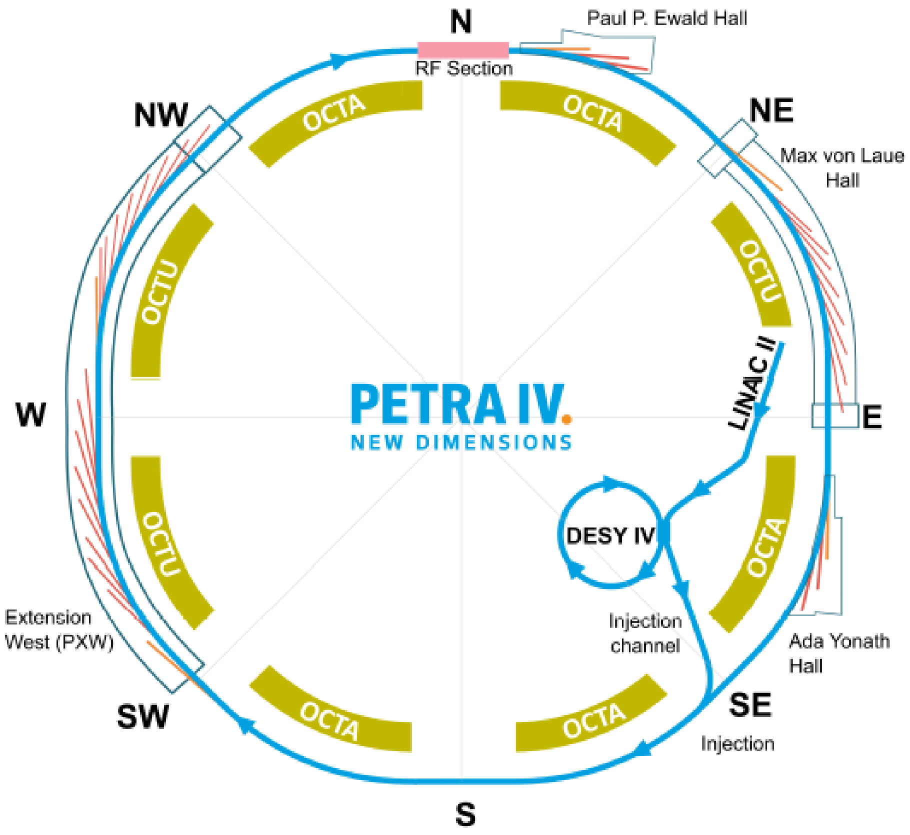
Each year, more than 2000 users are performing measurements at the PETRA III beamlines.

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

PETRA IV.
NEW DIMENSIONS

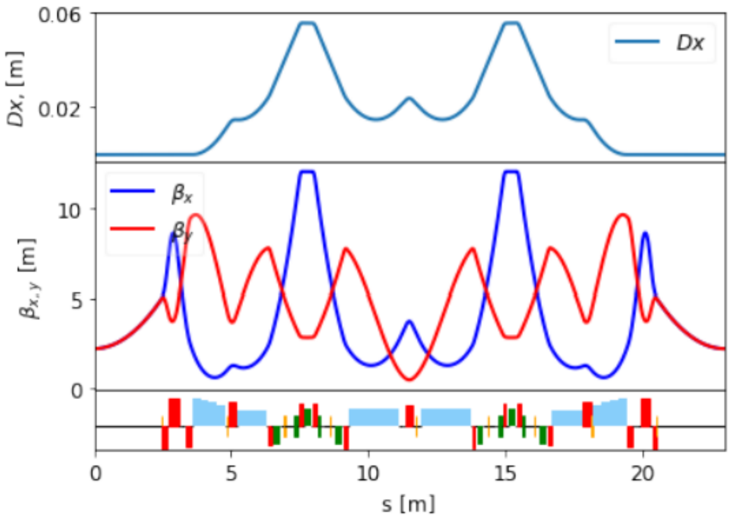
- New high-resolution 3D X-ray microscope for chemical and physical processes.
- Construction within the existing PETRA ring tunnel.
- Nanometre scale for the first time.
- Ultra low emittances in the region of 10 pm.
- Each of the eight arcs is composed of nine hybrid six-bend achromat (H6BA) cells.

PETRA IV. NEW DIMENSIONS



Layout of the PETRA IV facility and the H6BA cell.
Courtesy of I. Agapov

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

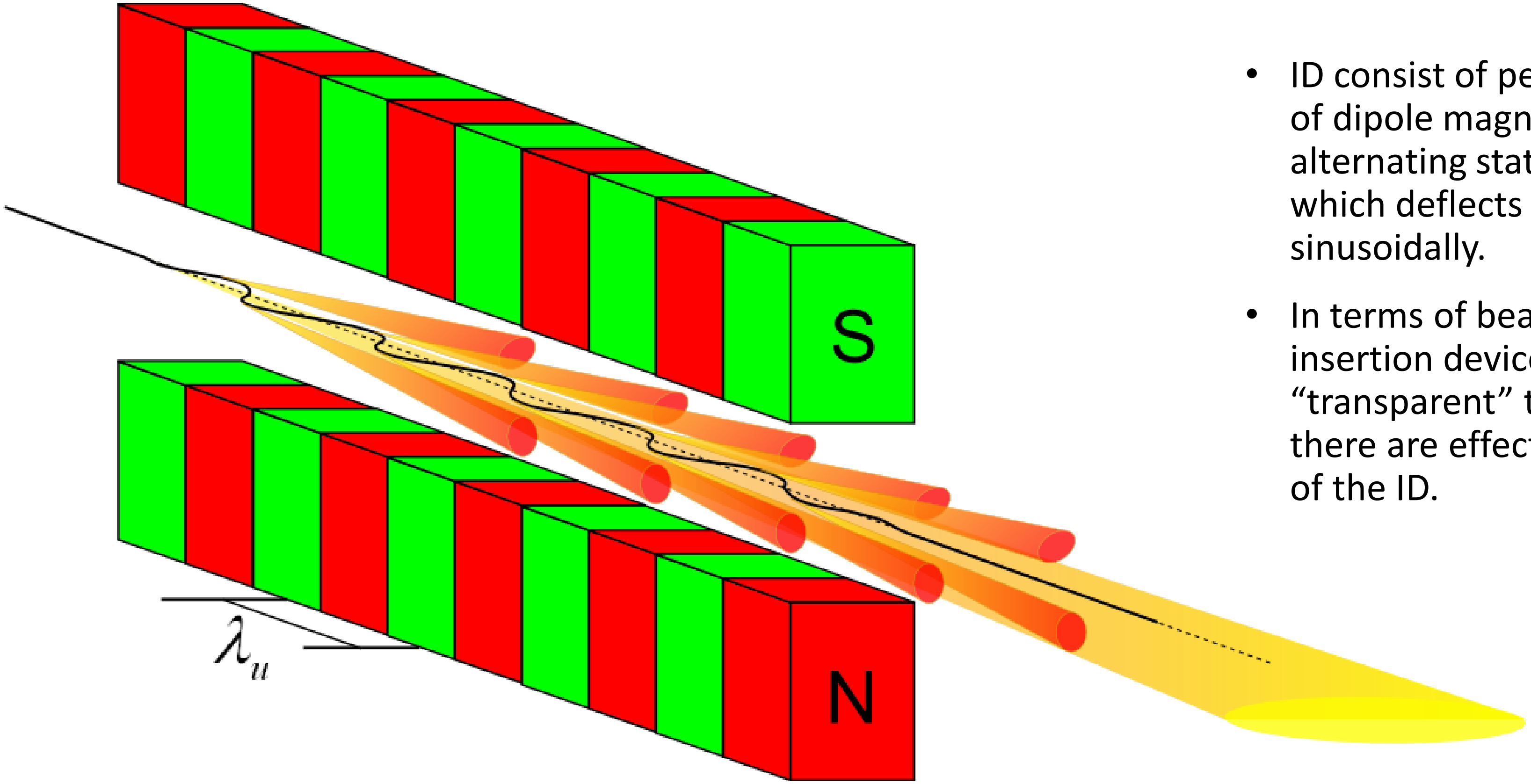


The storage ring feeds up to 30 undulator insertions (photon beam can be further split to allow more experimental stations). The storage ring will operate in two modes: brightness mode with 1920 stored bunches with the total current of 200 mA and the timing mode with 80 bunches and total current of 80 mA.

Insertion Devices

Synchrotron sources provide especially brilliant light that can be used to examine a vast variety of probes and samples. To produce this radiation, insertion devices are deployed.

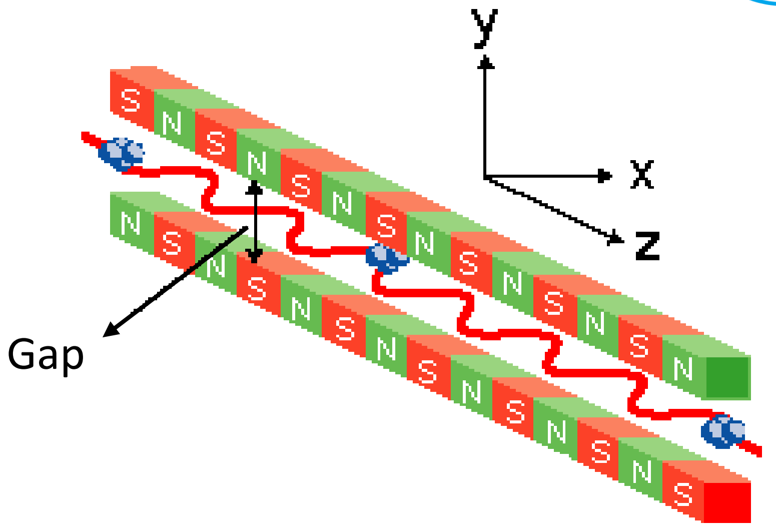
An insertion device is a special magnetic apparatus with periodic magnetic field designed to make the electron trajectory wiggle and generate intense synchrotron radiation. So-called „Undulators“ or „Wigglers“ are often „inserted“ in straight sections of storage rings → ID.



- ID consist of periodic arrangements of dipole magnets generating an alternating static magnetic field which deflects the electron beam sinusoidally.
- In terms of beam dynamics, an insertion device should be “transparent” to the machine. But there are effects due to field errors of the ID.

IDs induce an orbit distortion which varies with the gap size

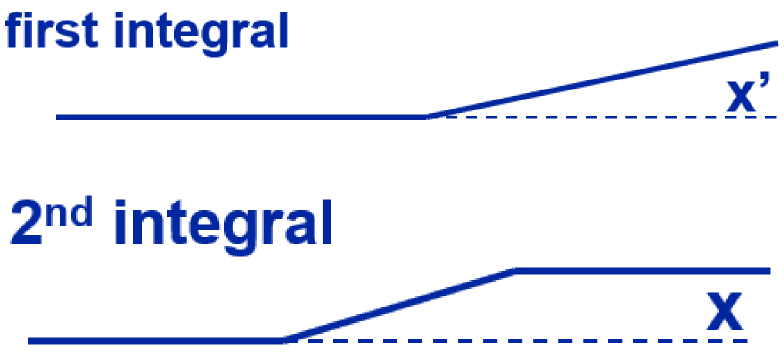
- The magnetic fields of IDs introduce perturbations to the circulating electron beam and hence affect the linear and nonlinear beam dynamics of the electron beam in the storage ring.
- Often users adjust the spectrum from undulators by changing undulator gap size. It's important to keep the orbit constant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



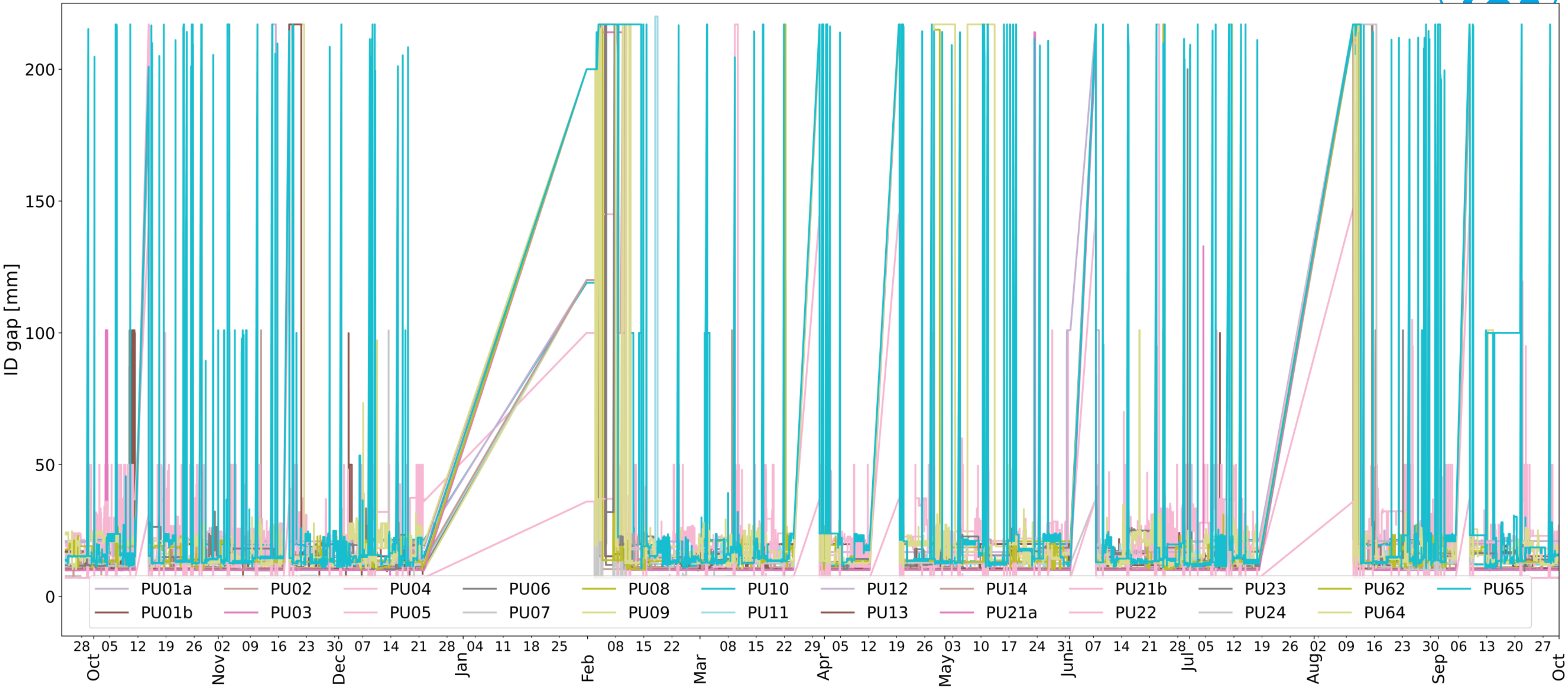
The field integrals determine the overall effect of the undulator on the electron beam orbit.

$$
\begin{aligned}
I_{1x} &\equiv \int_{z_0}^{z_0+L} B_x(z_1)\, dz_1 \\
I_{1y} &\equiv \int_{z_0}^{z_0+L} B_y(z_1)\, dz_1 \\
I_{2x} &\equiv \int_{z_0}^{z_0+L} \int_{z_0}^{z_2} B_x(z_1)\, dz_1\, dz_2 \\
I_{2y} &\equiv \int_{z_0}^{z_0+L} \int_{z_0}^{z_2} B_y(z_1)\, dz_1\, dz_2
\end{aligned}
\quad \Rightarrow \quad
\begin{aligned}
x'_{exit} &= -\frac{q}{\gamma m v_z} I_{1y} \\
y'_{exit} &= \frac{q}{\gamma m v_z} I_{1x} \\
x_{exit} &= -\frac{q}{\gamma m v_z} I_{2y} \\
y_{exit} &= \frac{q}{\gamma m v_z} I_{2x}
\end{aligned}
$$

Small closed orbit distortions but very sensitive experiments required sub-µrad corrections



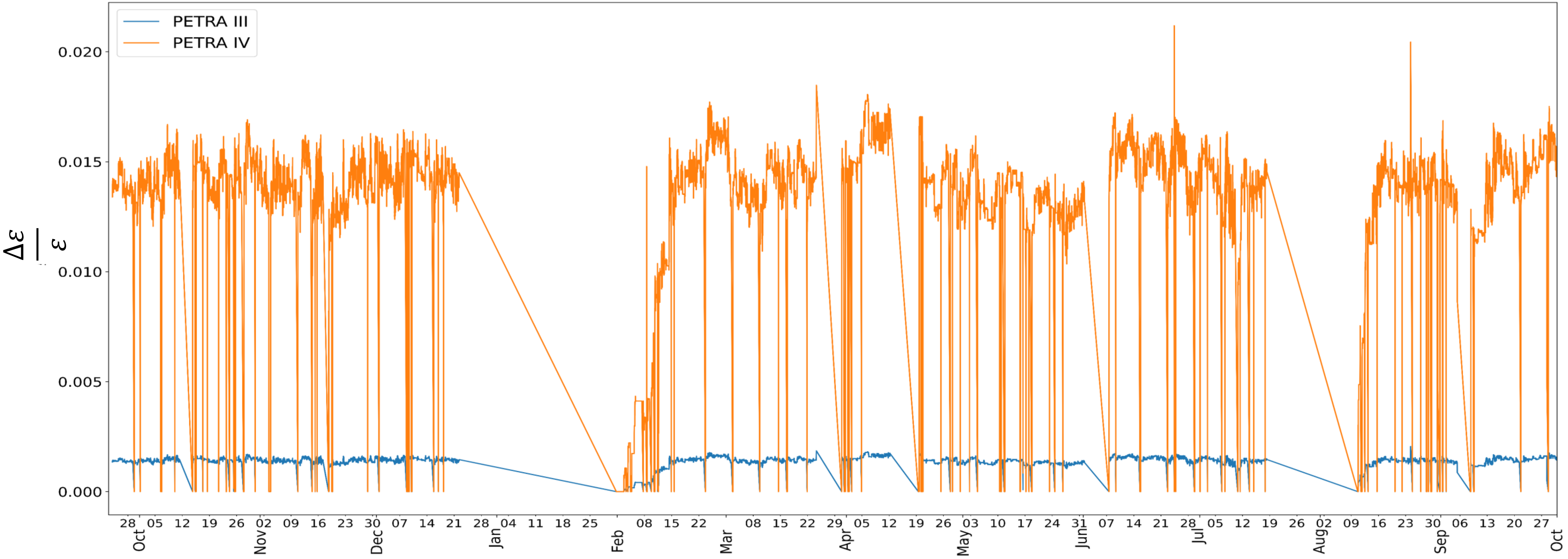
IDs affect the beam dynamics of the stored electron beam



The data was filtered to represent only user operations (while the Fast Orbit Feedback system is on).
The IDs have different maximum and minimum gap sizes.

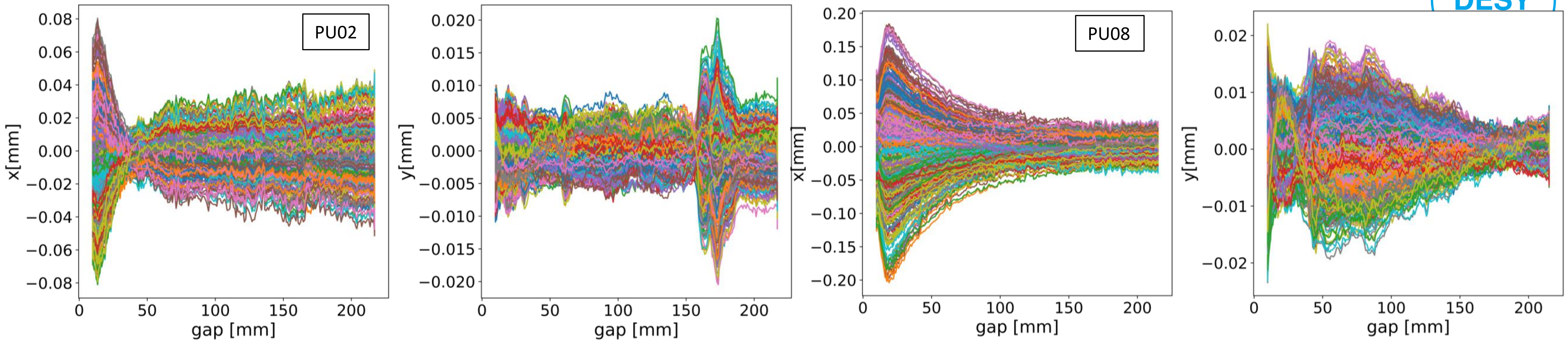
IDs affect the beam dynamics of the stored electron beam. The intensity of the effect depend on the ID gap.

Impact on emittance, projections for PETRA IV.

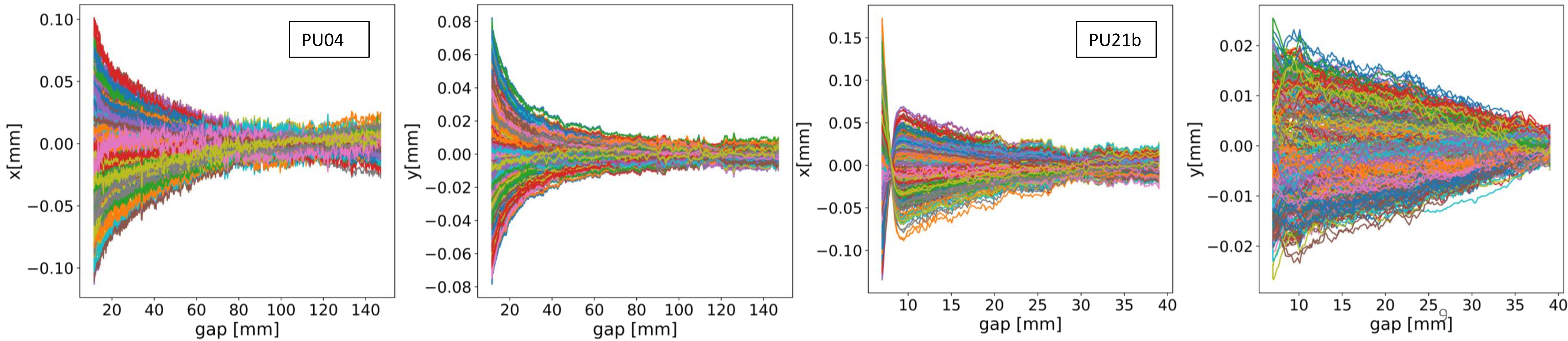


$$\varepsilon = C_q \gamma^2 \frac{I_5}{I_2 - I_4} \qquad \frac{\Delta\varepsilon}{\varepsilon} \approx \frac{1}{1+\frac{I_2^{ID}}{I_2}} \qquad I_2^{ID} = \frac{L_{ID}}{2\rho_{ID}^2} \qquad \rho_{ID} \propto \exp\left(\frac{3 * gap}{\lambda_p}\right)$$

Closed orbit distortion measurements



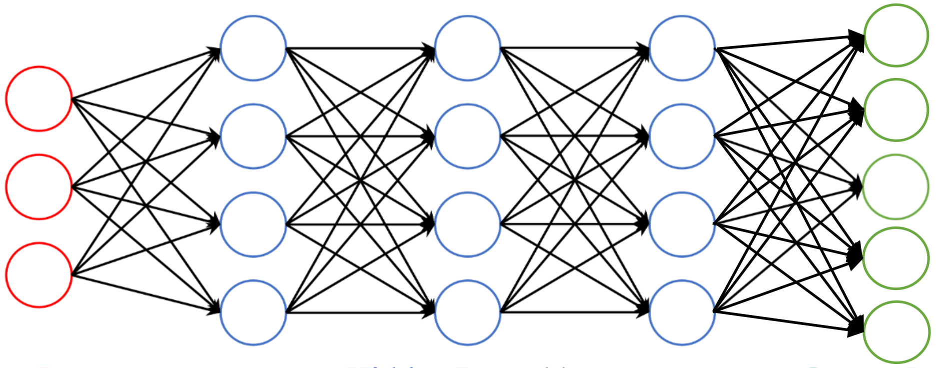
Measurements of the horizontal and vertical orbit were taken varying the gap size in their operation range for 18 IDs in PETRA III. Each colour represents a BPM along the ring.



Building the NN

- 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%

ID gaps



Closed Orbit

Input Layer:
ID gaps
(18)

Hidden Layers:
number of
neurons in each layer

Output Layer:
Closed Orbit
(2x239=Number of BPMs)

Keras

```
model = keras.Sequential()
…
model.compile(loss="mse",
optimizer="adam", metrics=["mae"])
```

Hyperparmeter sweeps performed with

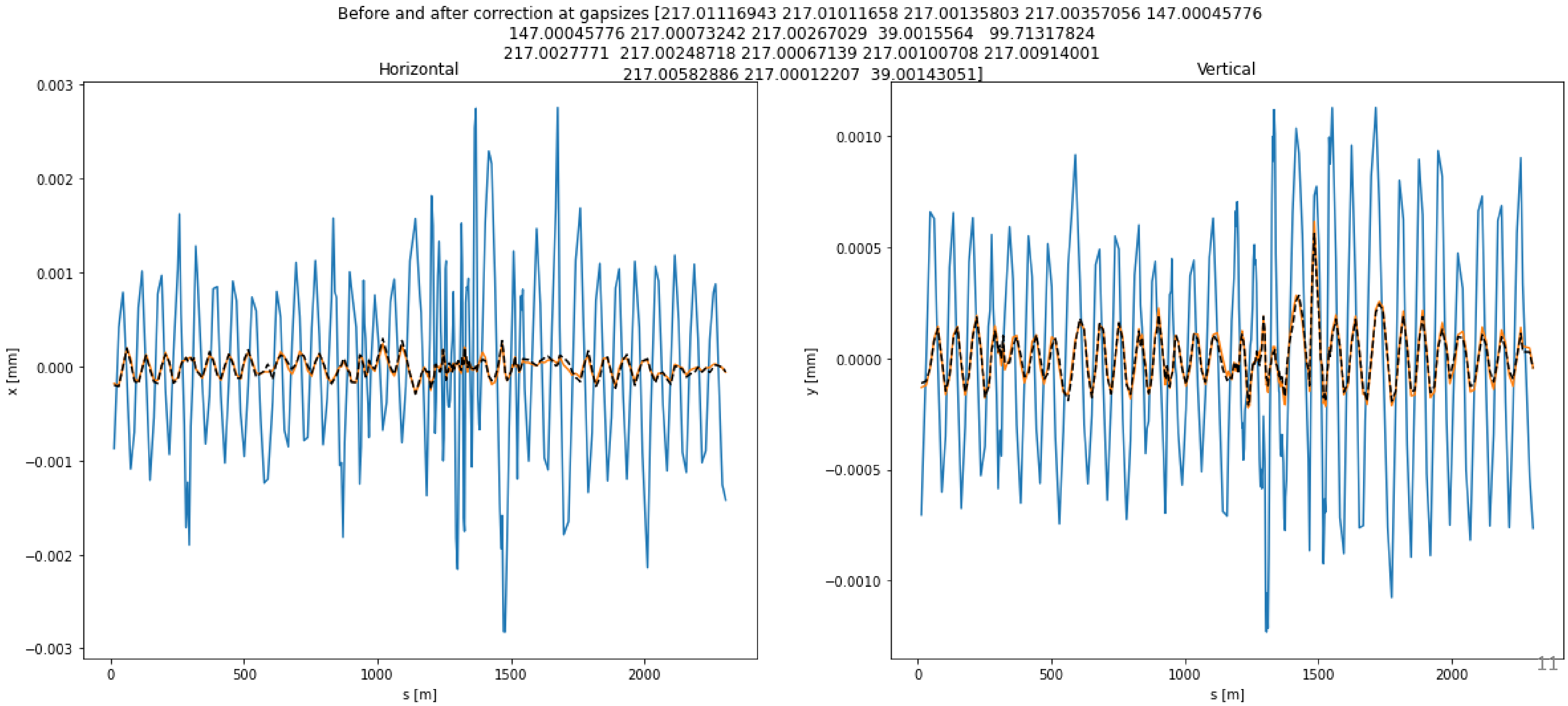
Weights & Biases

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.

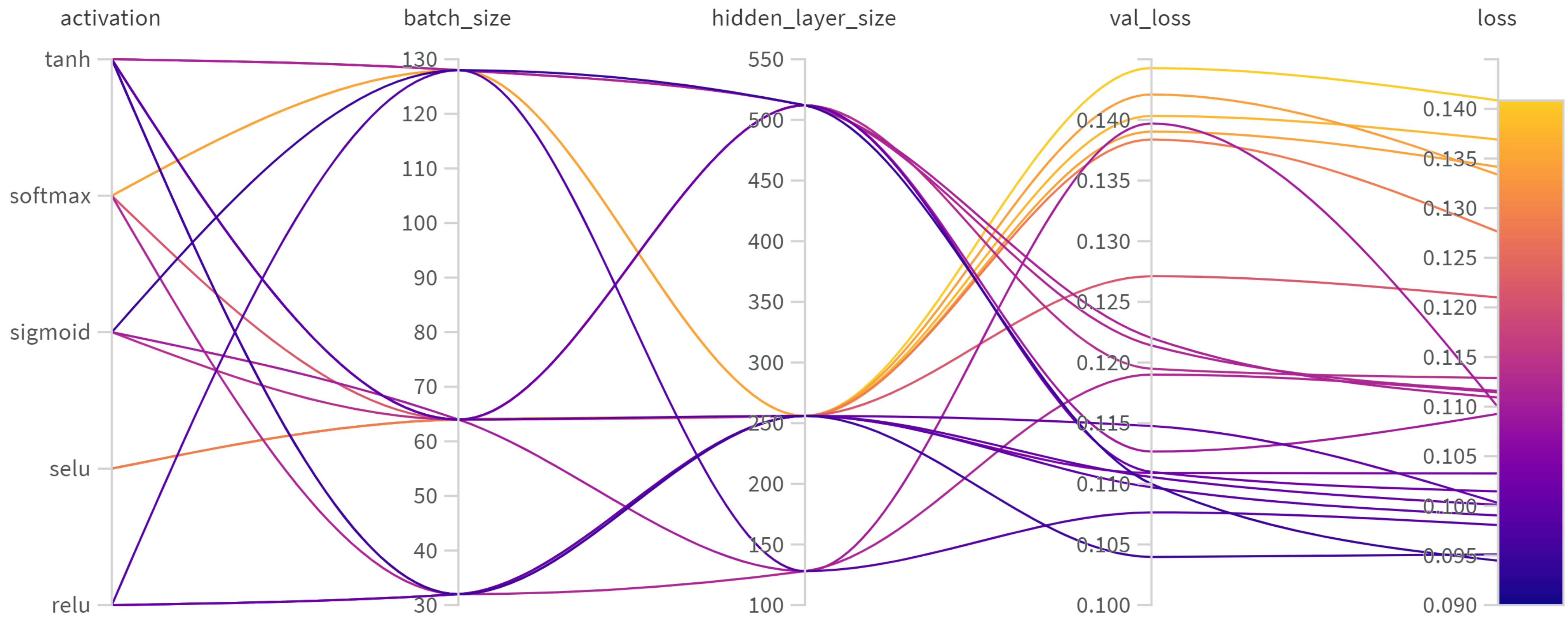
From prediction to correction:

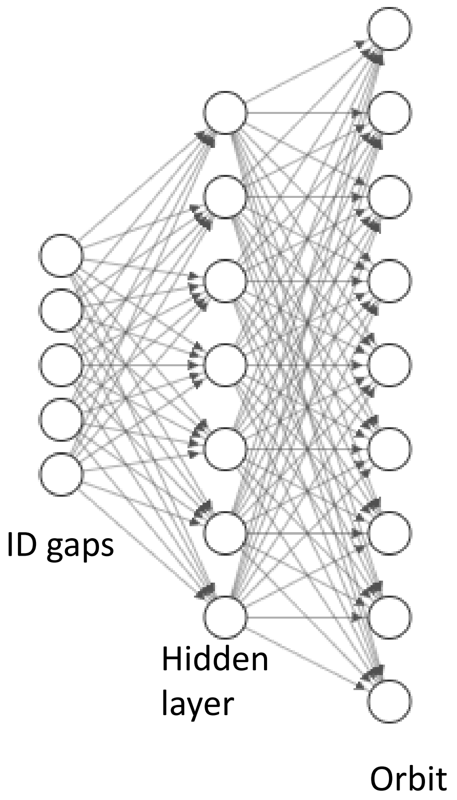
- Random configuration of ID gaps.
- Once the orbit is predicted, the strengths of the correctors are computed through SVD.



Shallow feed-forward fully connected NN

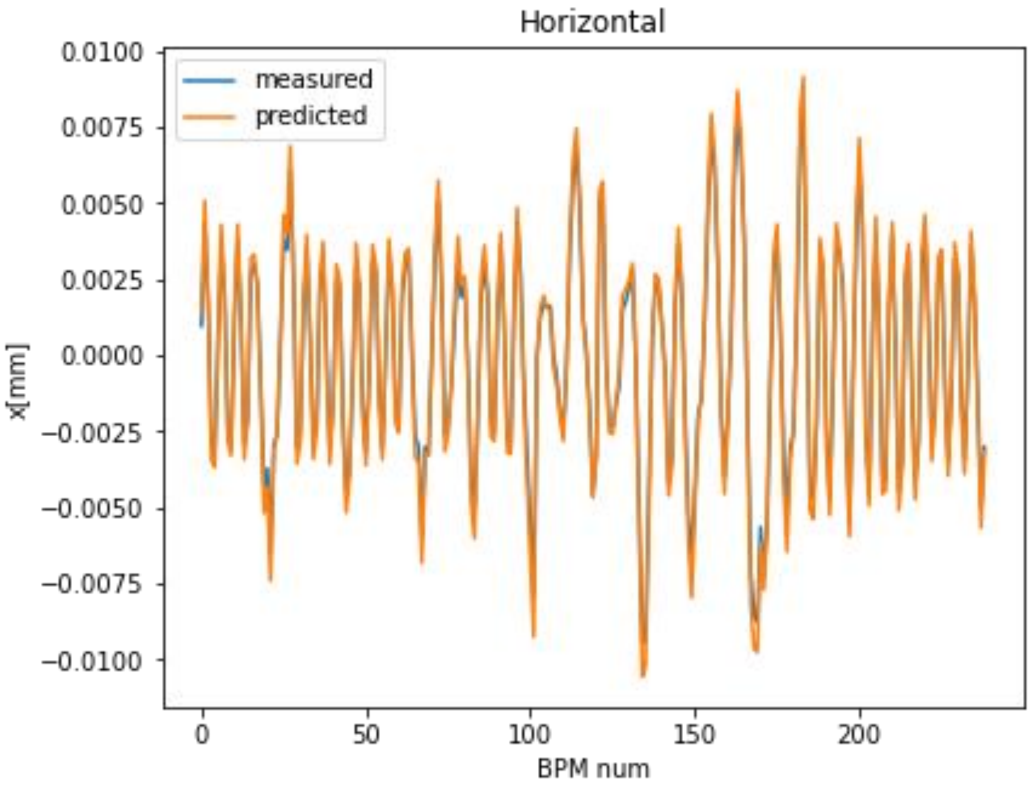
One hidden layer: exploring the impact of different activation functions and batch size.

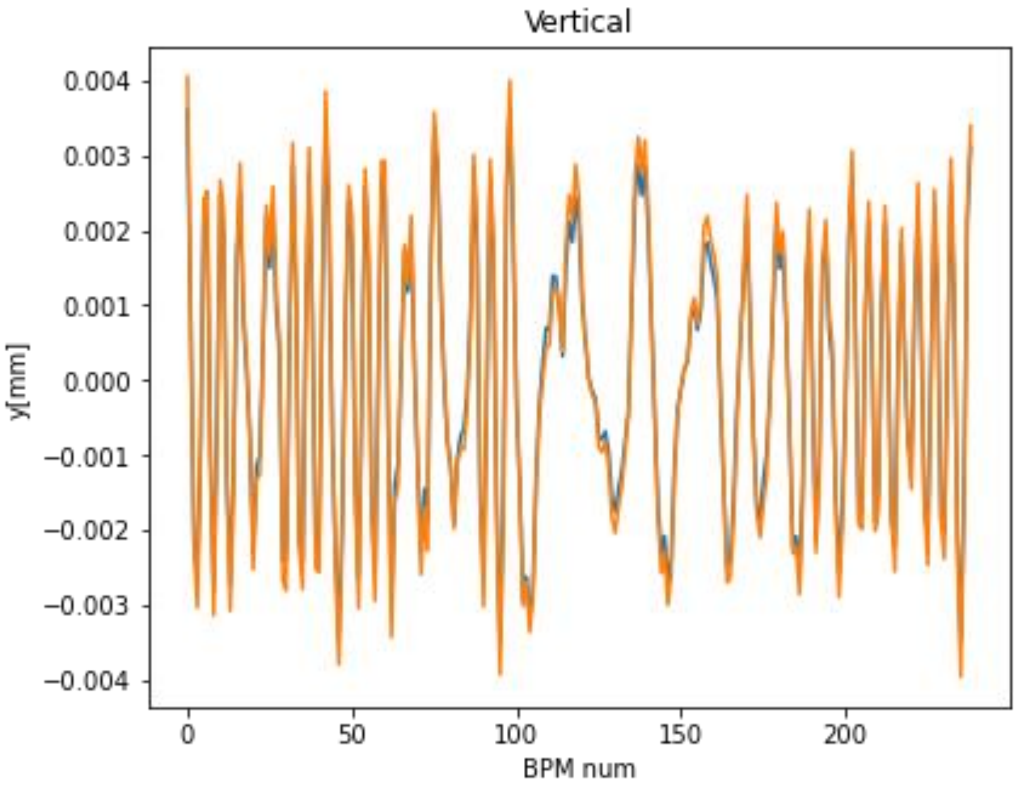


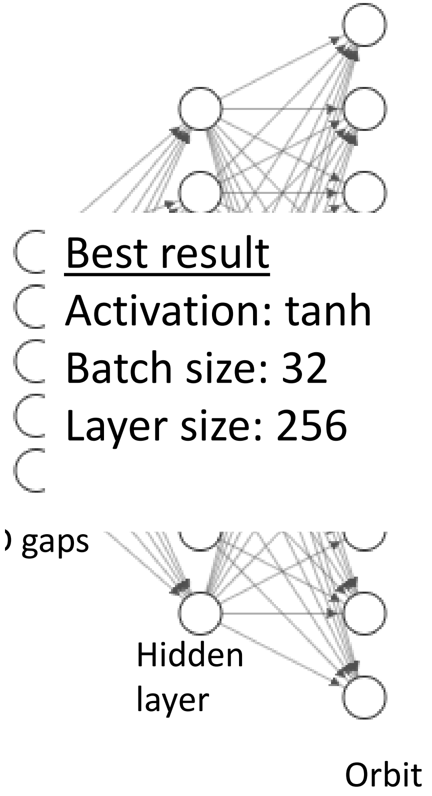


Shallow feed-forward fully connected NN

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







<u>Best result</u>

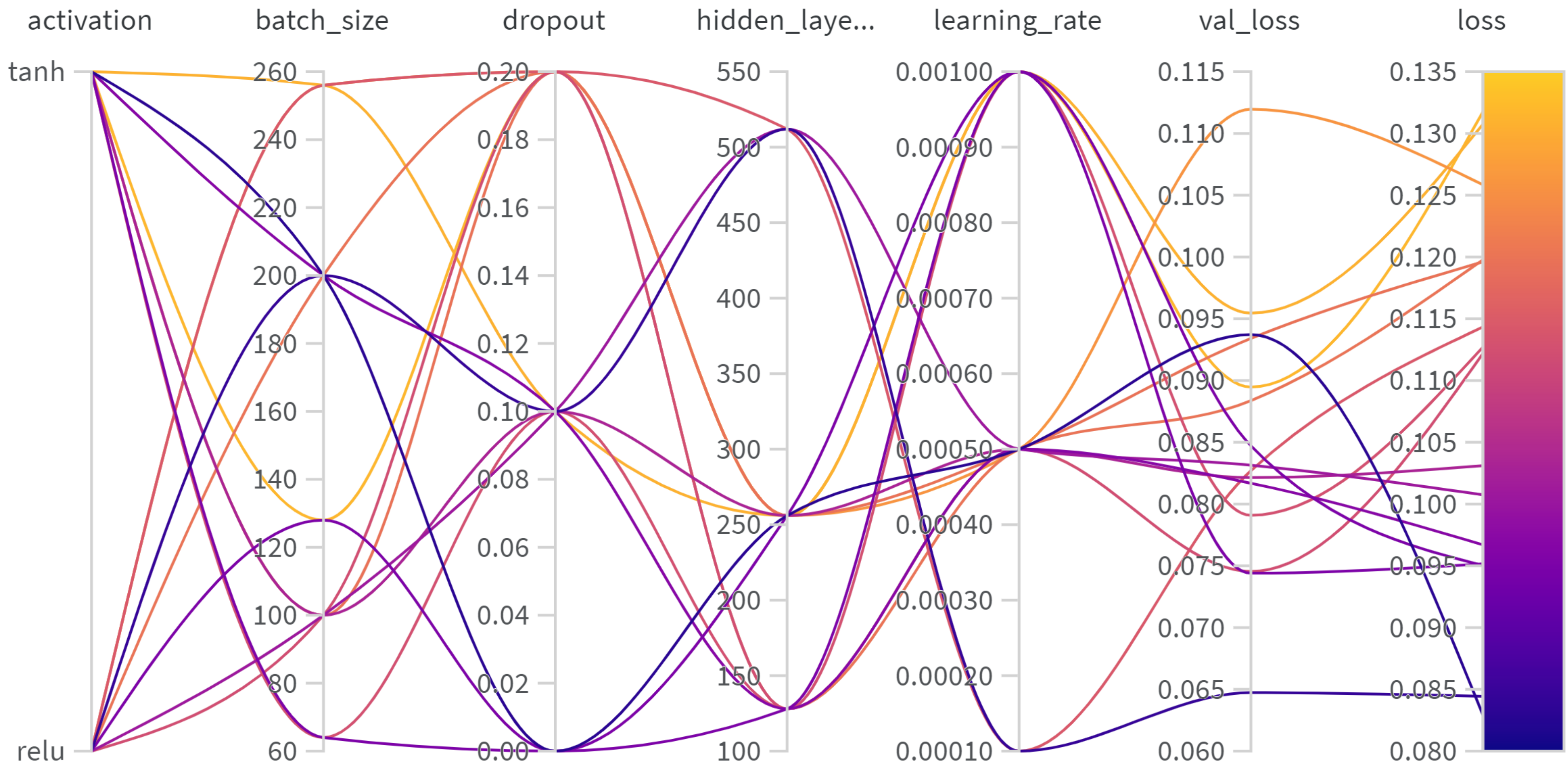
Activation: tanh

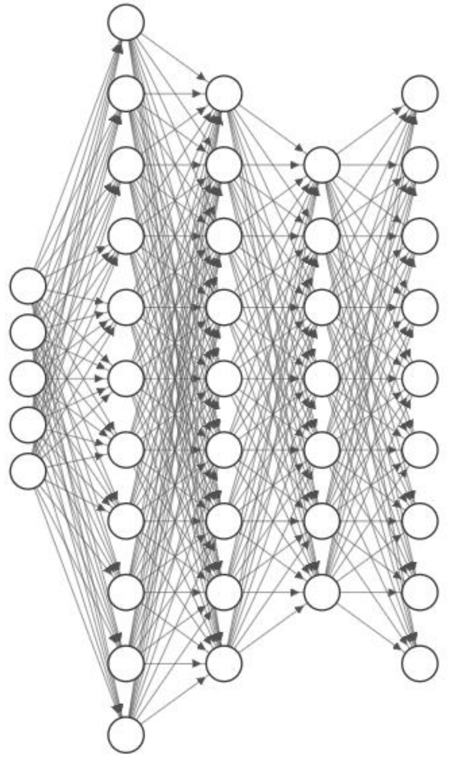
Batch size: 32

Layer size: 256

Deep feed-forward fully connected NN

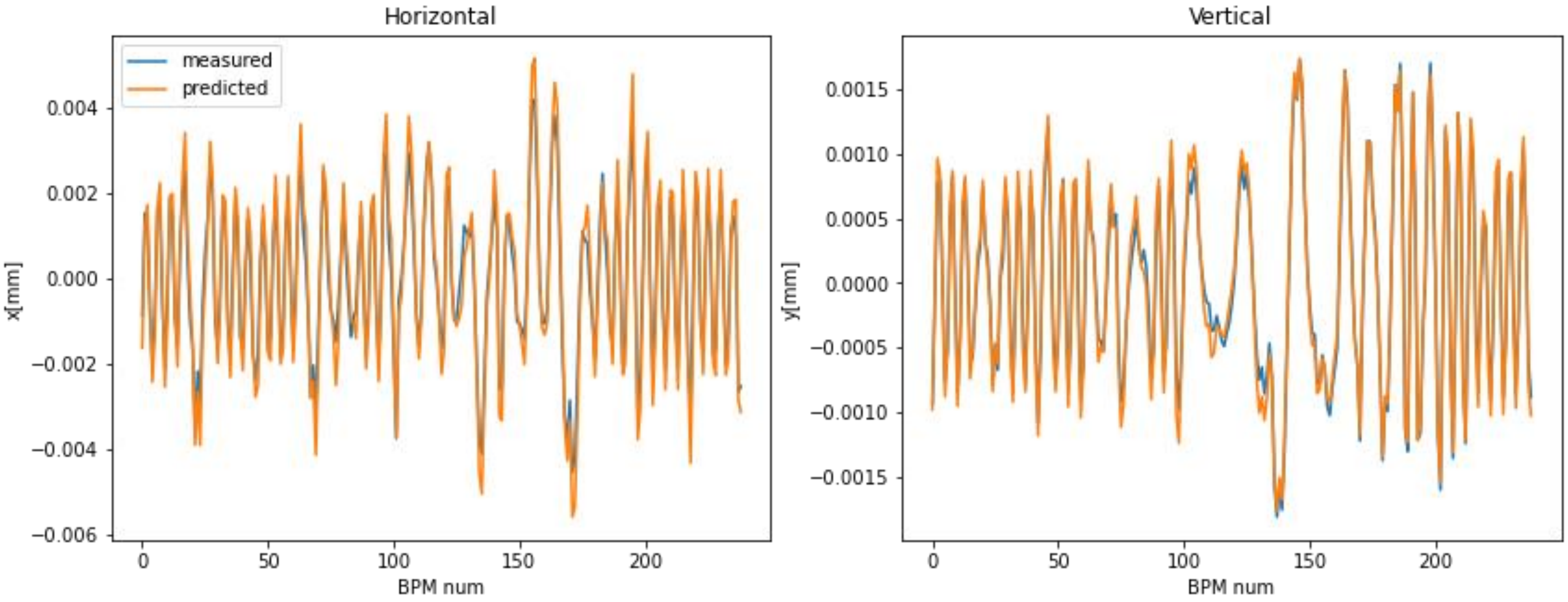
Multiple hidden layers: exploring the impact of dropout, batch and layer size and learning rate.

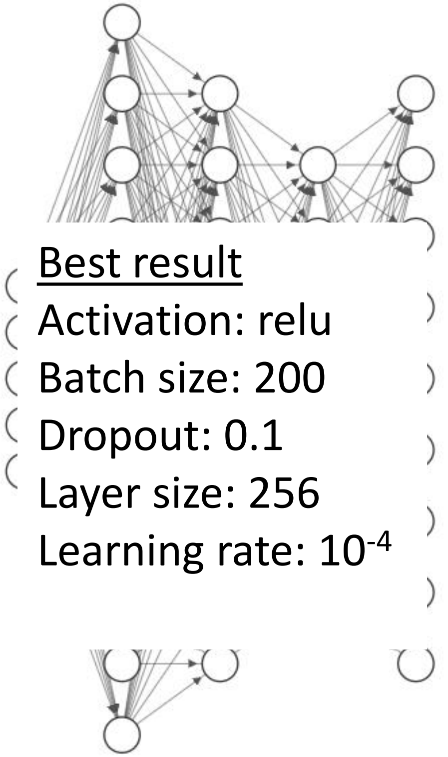




Deep feed-forward fully connected NN

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





Best result

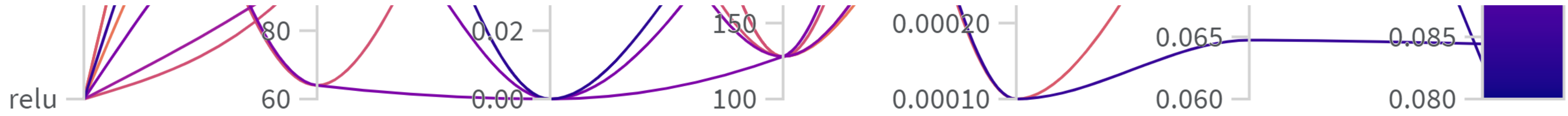
Activation: relu

Batch size: 200

Dropout: 0.1

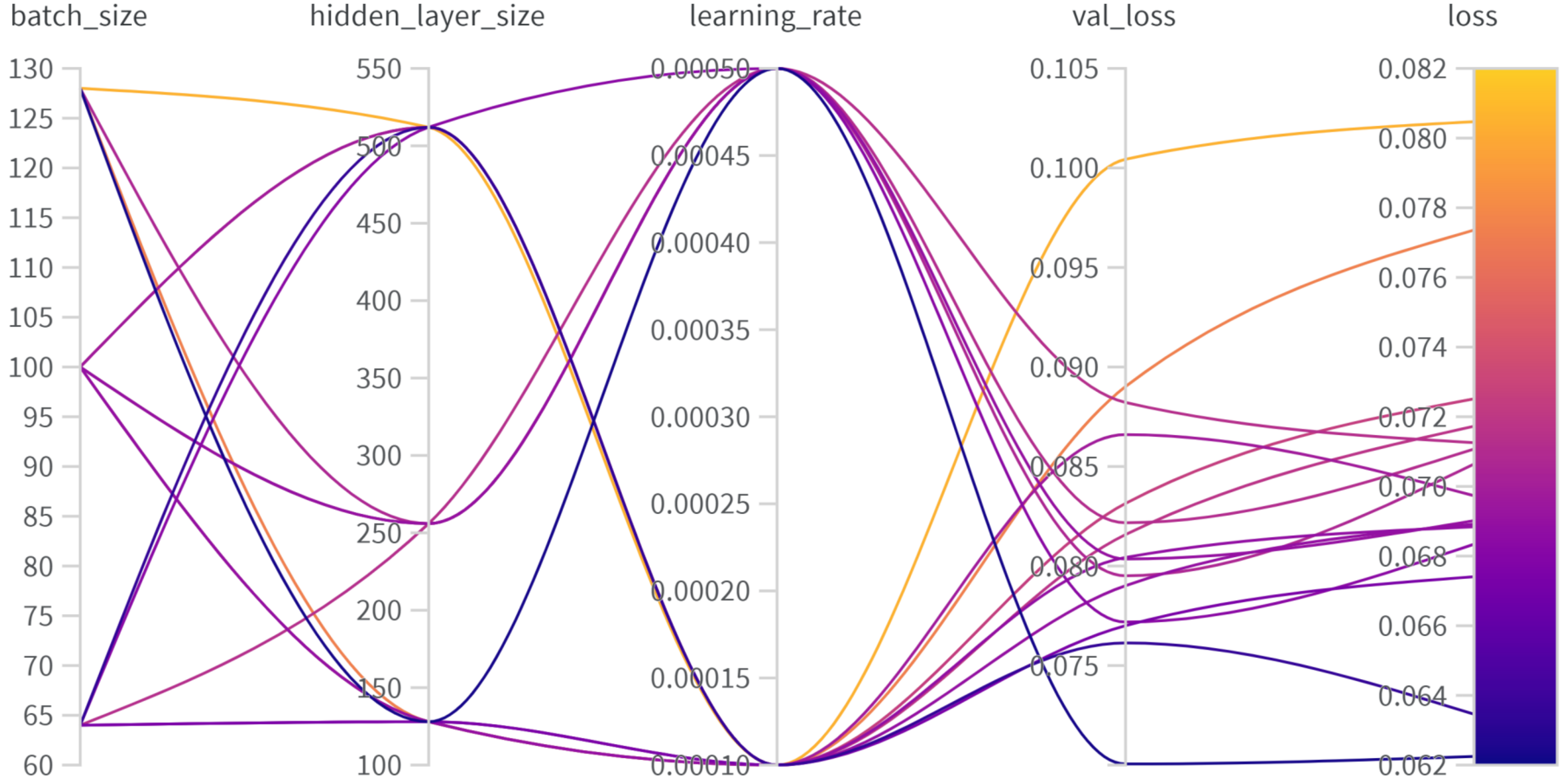
Layer size: 256

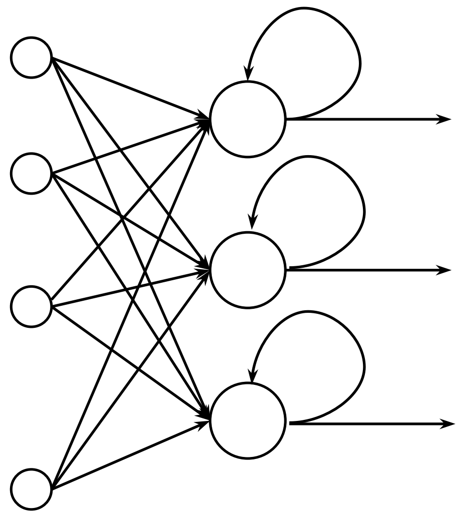
Learning rate: 10^{-4}



Recurrent NN

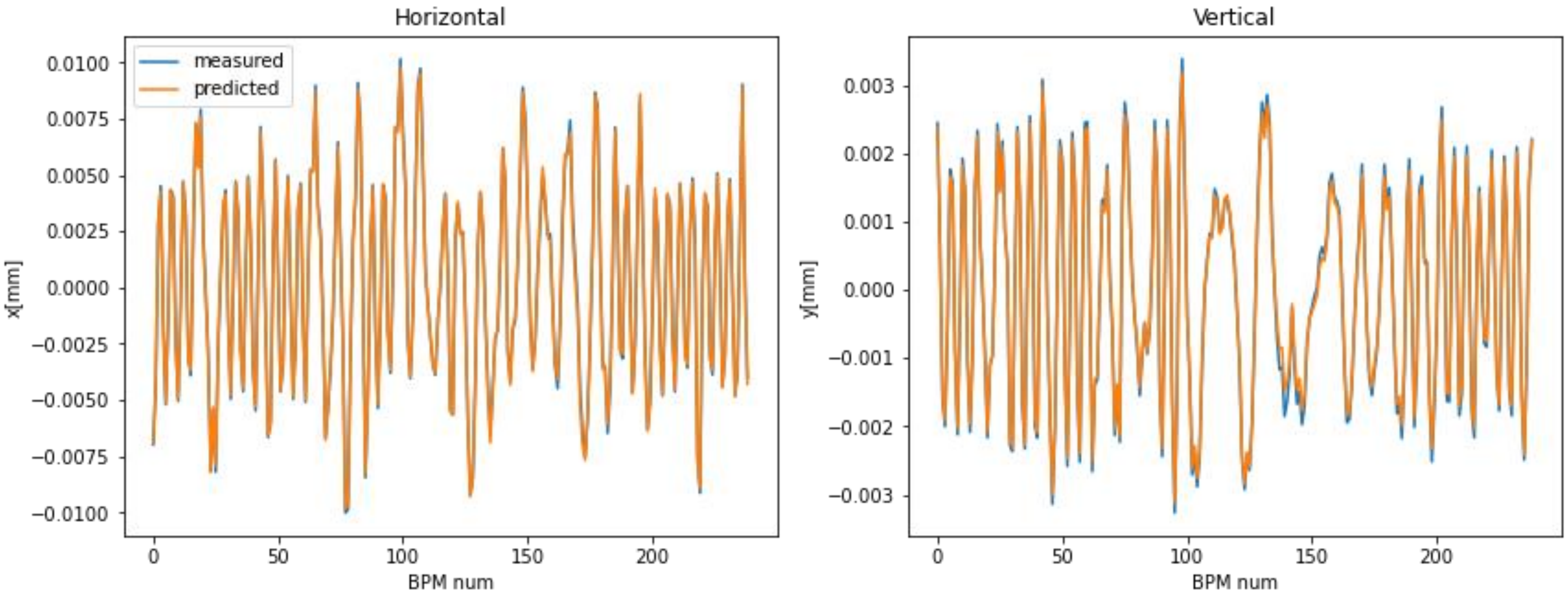
In a RNN the information cycles through a loop. When it makes a decision, it considers the current input and also what it has learned from the inputs it received previously.





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

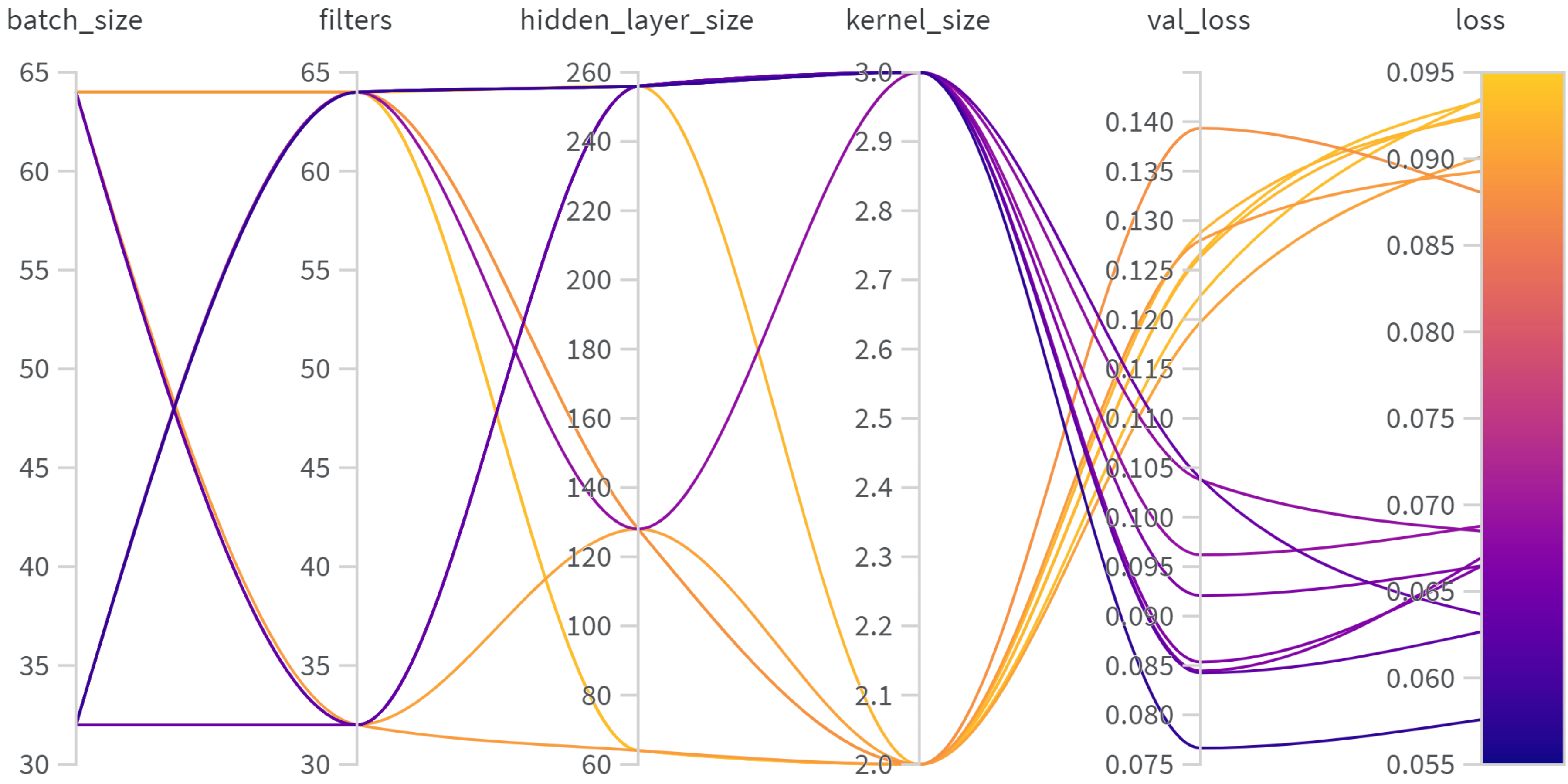
Layer size: 128

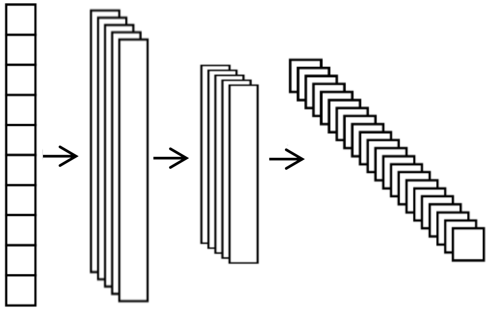
Learning rate: $5x10^{-4}$

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.

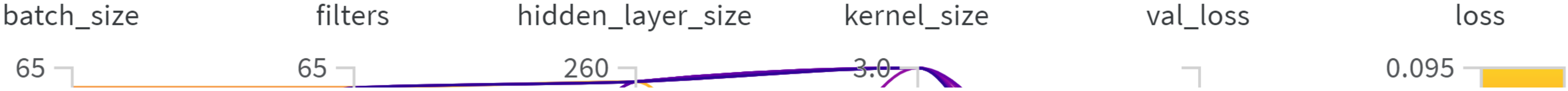


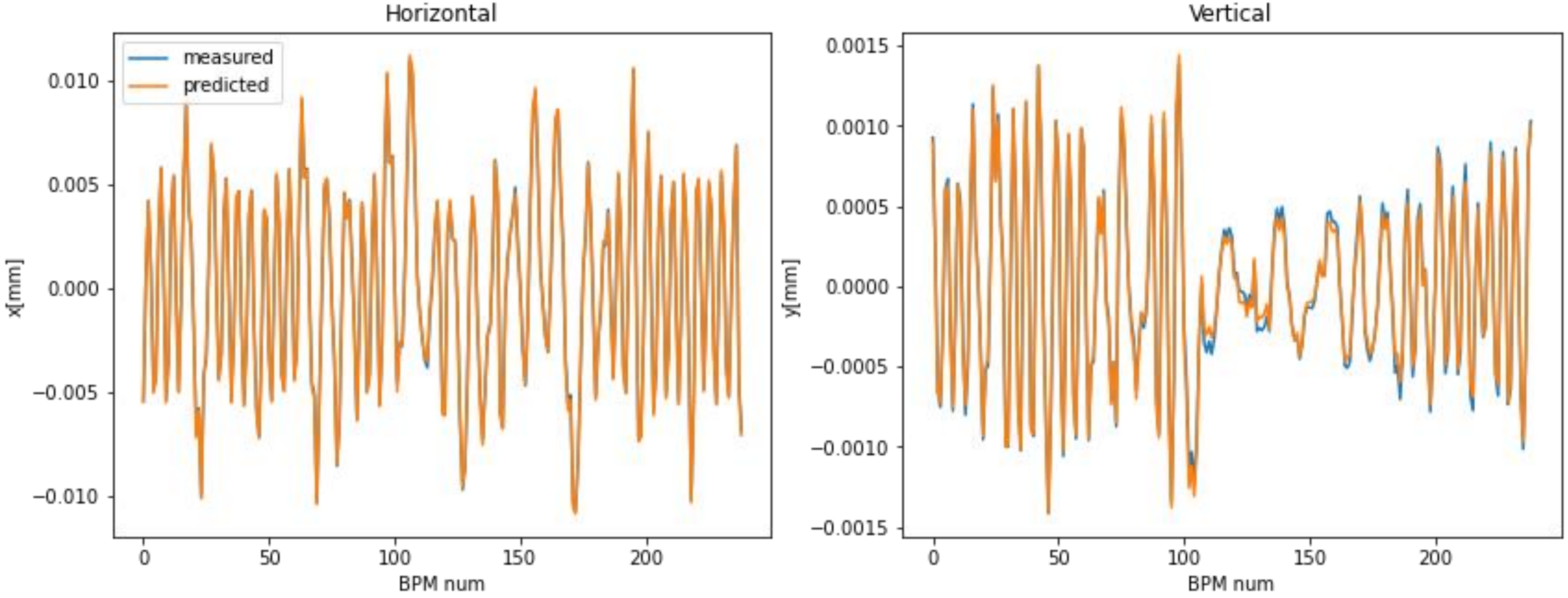


1D Convolutional NN

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>

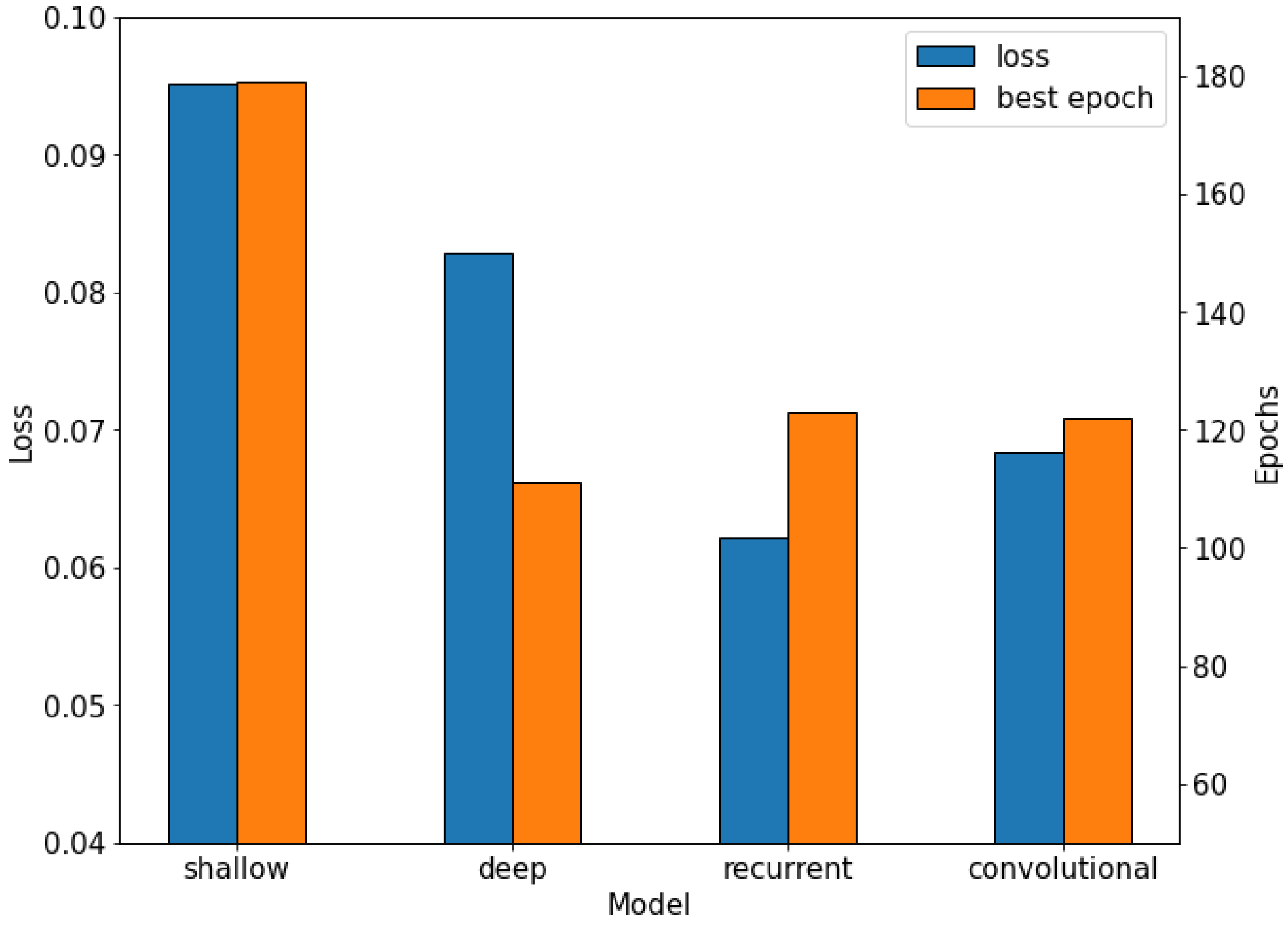
Batch size: 32

Filters: 32

Layer size: 256

Kernel size: 3

Comparing the architectures



The convolutionl and recurrent structure outperform the fully connected NN in a reasonable amount of epochs.

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.
- The prediction can be used to calculate the corrector magnet strength.
- The same scheme could be applied to PETRA IV considering as well the expected significant impact on the emittance.

Thank you

Contact

DESY. Deutsches Elektronen-Synchrotron

www.desy.de

Bianca Veglia

MPY

bianca.veglia@desy.de