First Studies on Machine Learning for the ALS Storage Ring

Simon C. Leemann & Shuai Liu
ALS Accelerator Physics, ATAP Division, Lawrence Berkeley National Laboratory
March 28, 2019

In collaboration with: Alex Hexemer (co-PI), Hiroshi Nishimura, Matthew A. Marcus, David Shapiro, Changchun Sun, Dani Ushizima, Nathan Melton, Greg Penn, Thorsten Hellert

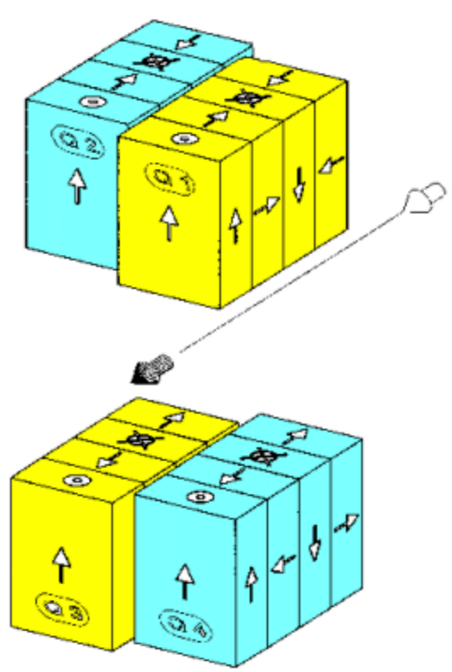
BERKELEY LAB

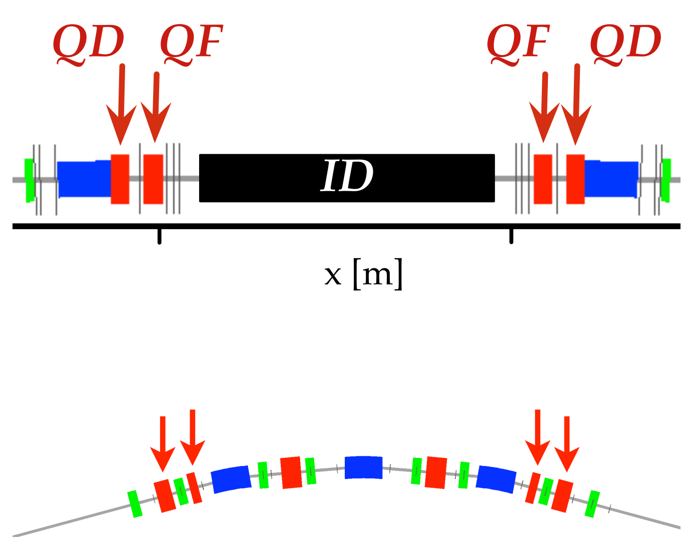
U.S. DEPARTMENT OF ENERGY | Office of Science

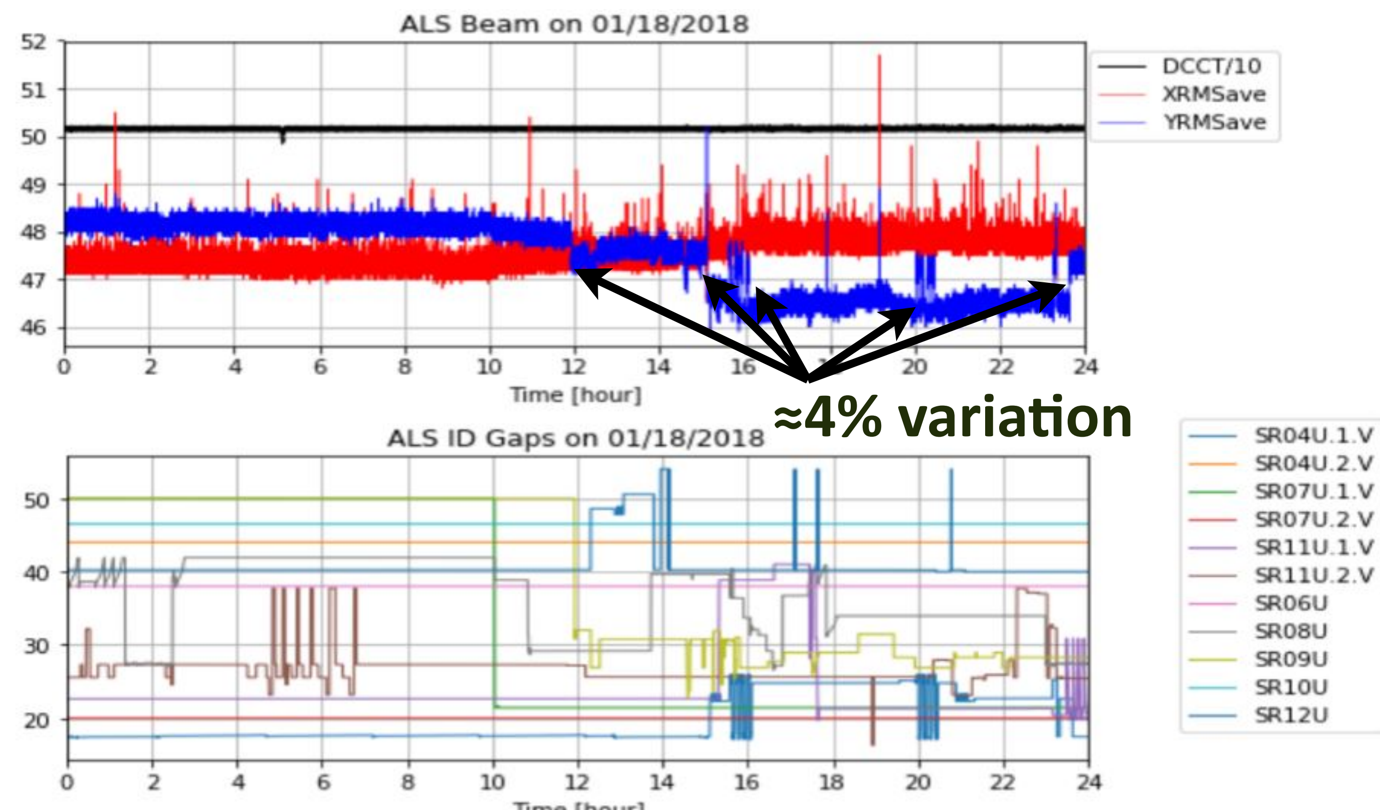
ALS ADVANCED LIGHT SOURCE

Background

- ALS Storage Ring (SR) orbit is very well controlled up to many tens of Hz
- ALS beam size control is an entirely different matter: depends strongly on plane, local optics tuning, global optics corrections, eg. ID feed-forwards (FFs), tune feedback (FB)
- ID FFs work well, but require look-up tables (≈12 hrs AP shift per EPU to generate two 3D tables) & correction susceptible to machine drift ➔ constantly re-record tables



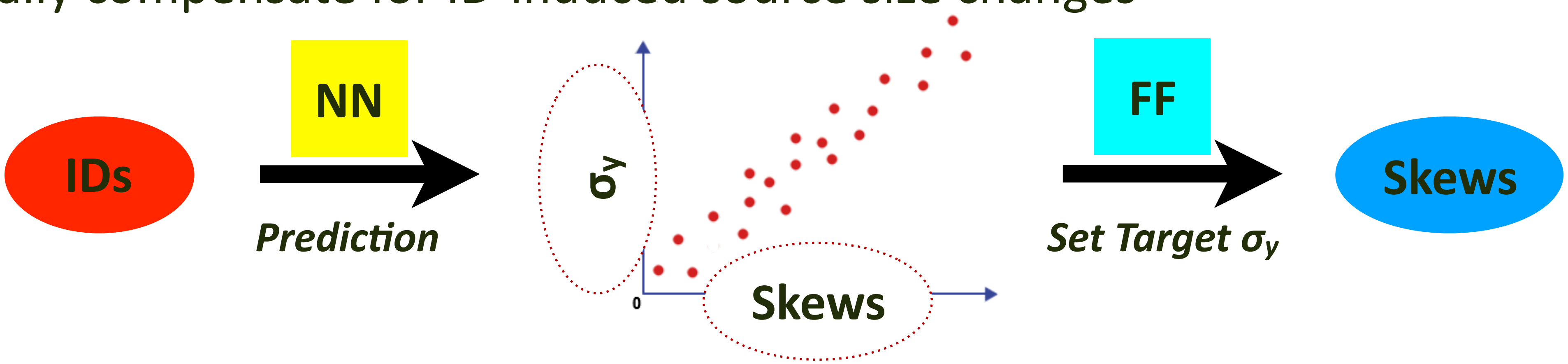




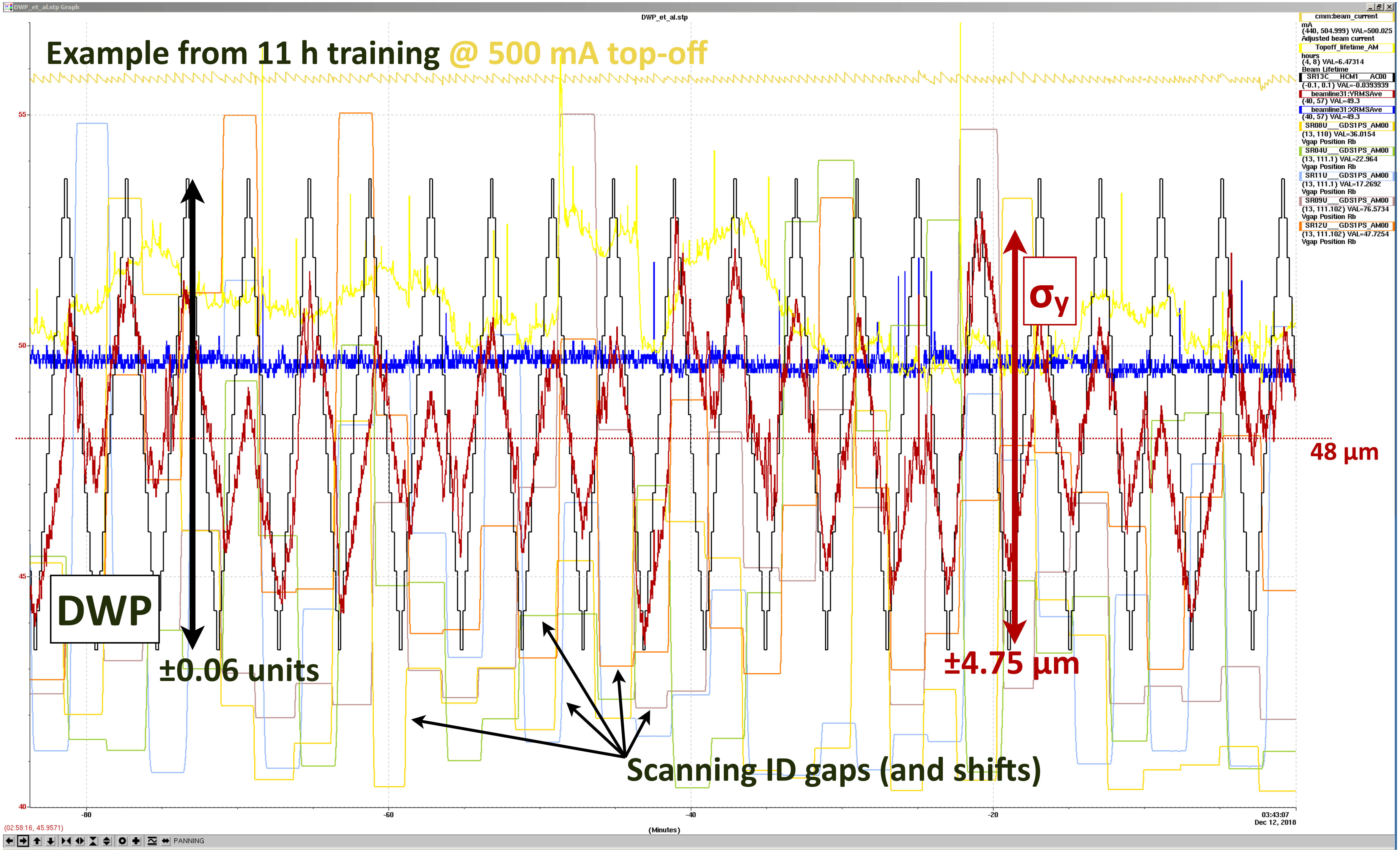
Courtesy: H. Nishimura

Background (cont.)

- Early 2018: DOE BES ADRP suggests we form collaboration with SSRL on ***"Beam Based Optimization and Machine Learning for Synchrotrons"***
- DOE BES & ASCR: granted funding Aug 2018 for two years (≈$0.7M @ ALS & similar at SSRL) → S.C. Leemann & A. Hexemer (PIs @ ALS), X. Huang & J. Safranek (PIs @ SSRL)
- Presently in the process of hiring postdoc to work on this full time; in the meantime collaborated with other accelerator (H. Nishimura) and ML experts (N. Melton) on preparing **first ML application for ALS SR**
  - Idea is *not* to replace any existing FBs or FFs, but to use an ***ML-based FF to remove residual fluctuations of vertical source size in ALS SR*** (sub-micron, ALS-U, etc.)
  - Employ NN to predict SR beam size as function of arbitrary ID gap/shift configurations → adjust skew quadrupole excitation ("vertical dispersion wave") to globally compensate for ID-induced source size changes



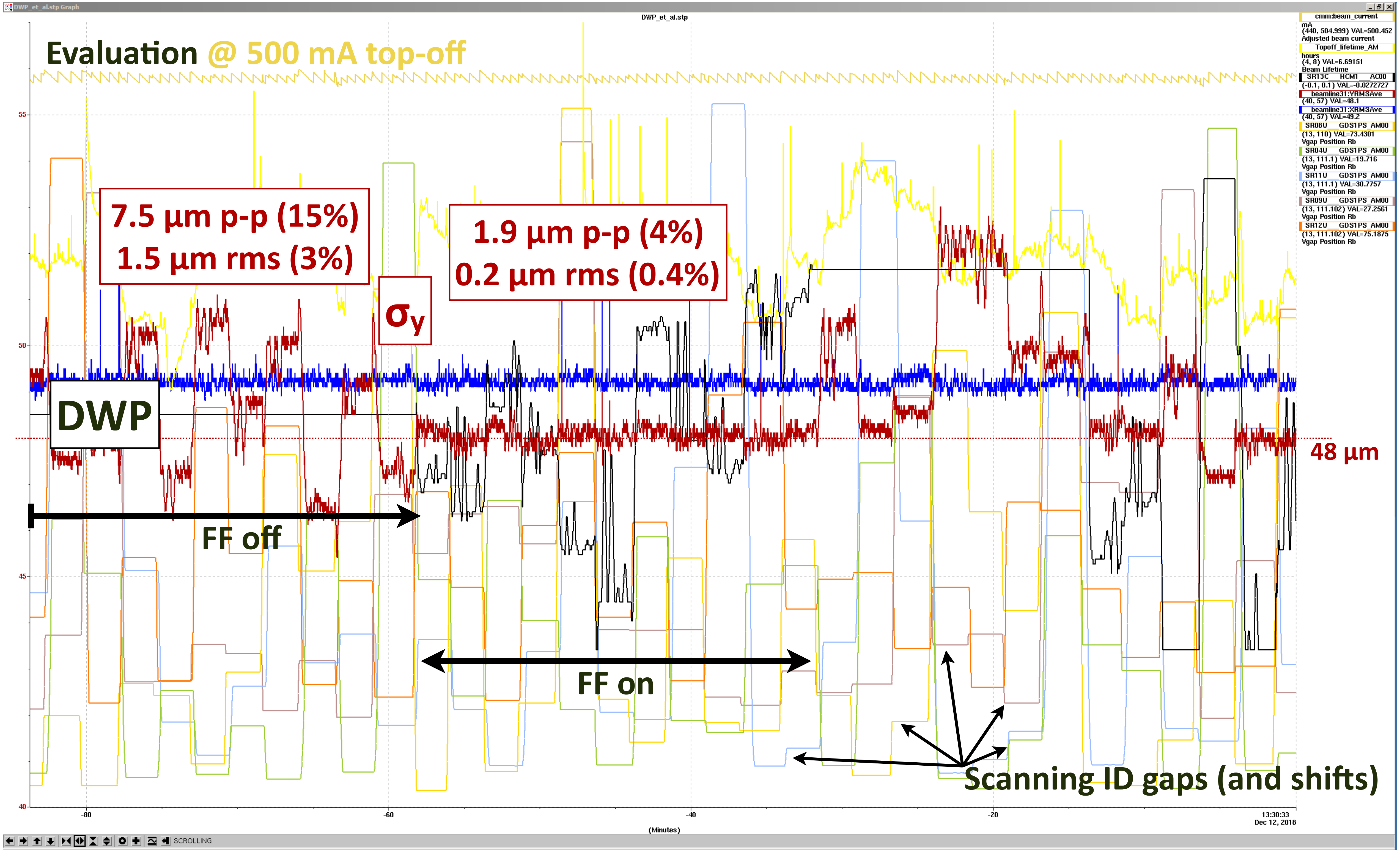
Training a Neural Network During AP Shift



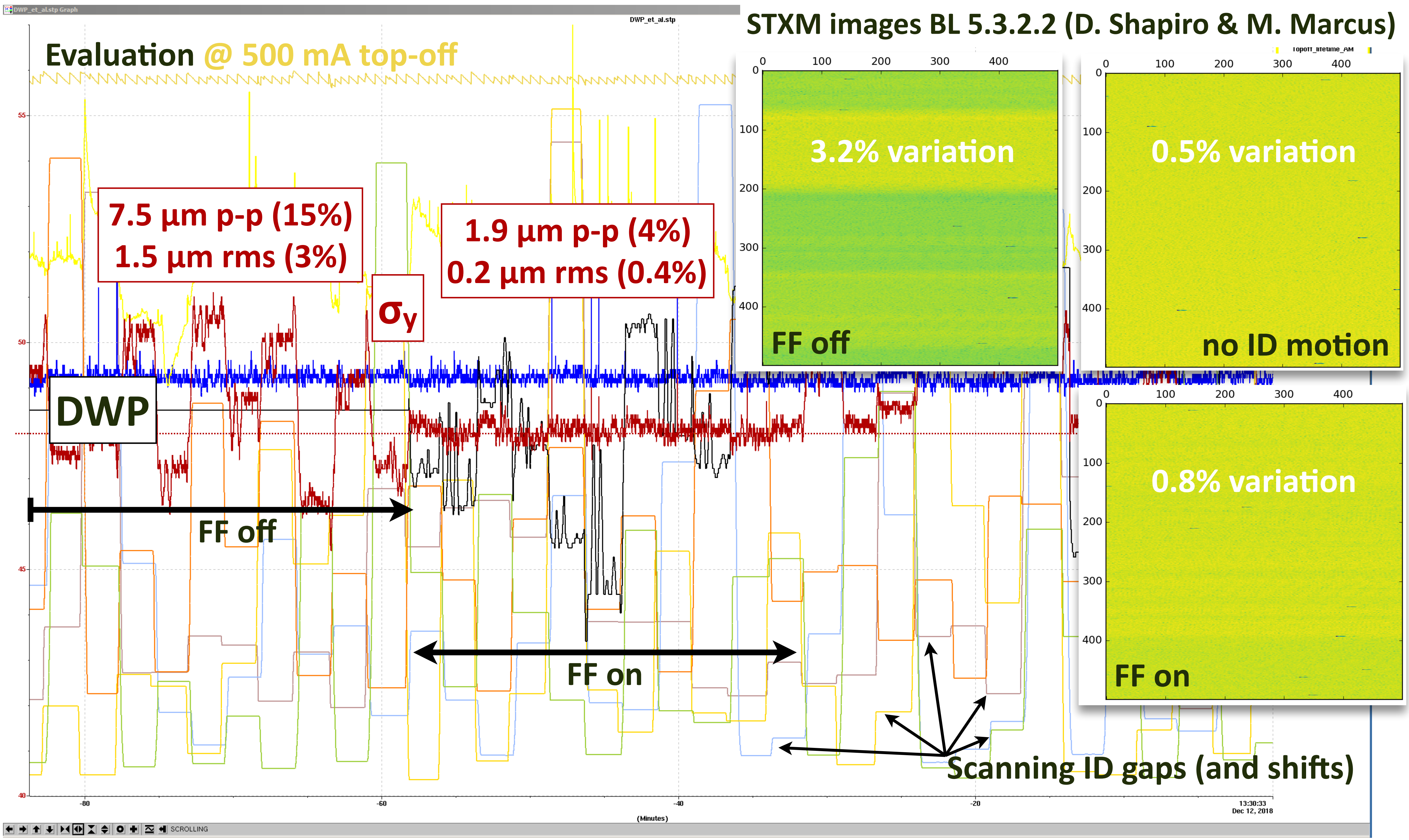
BERKELEY LAB

ALS

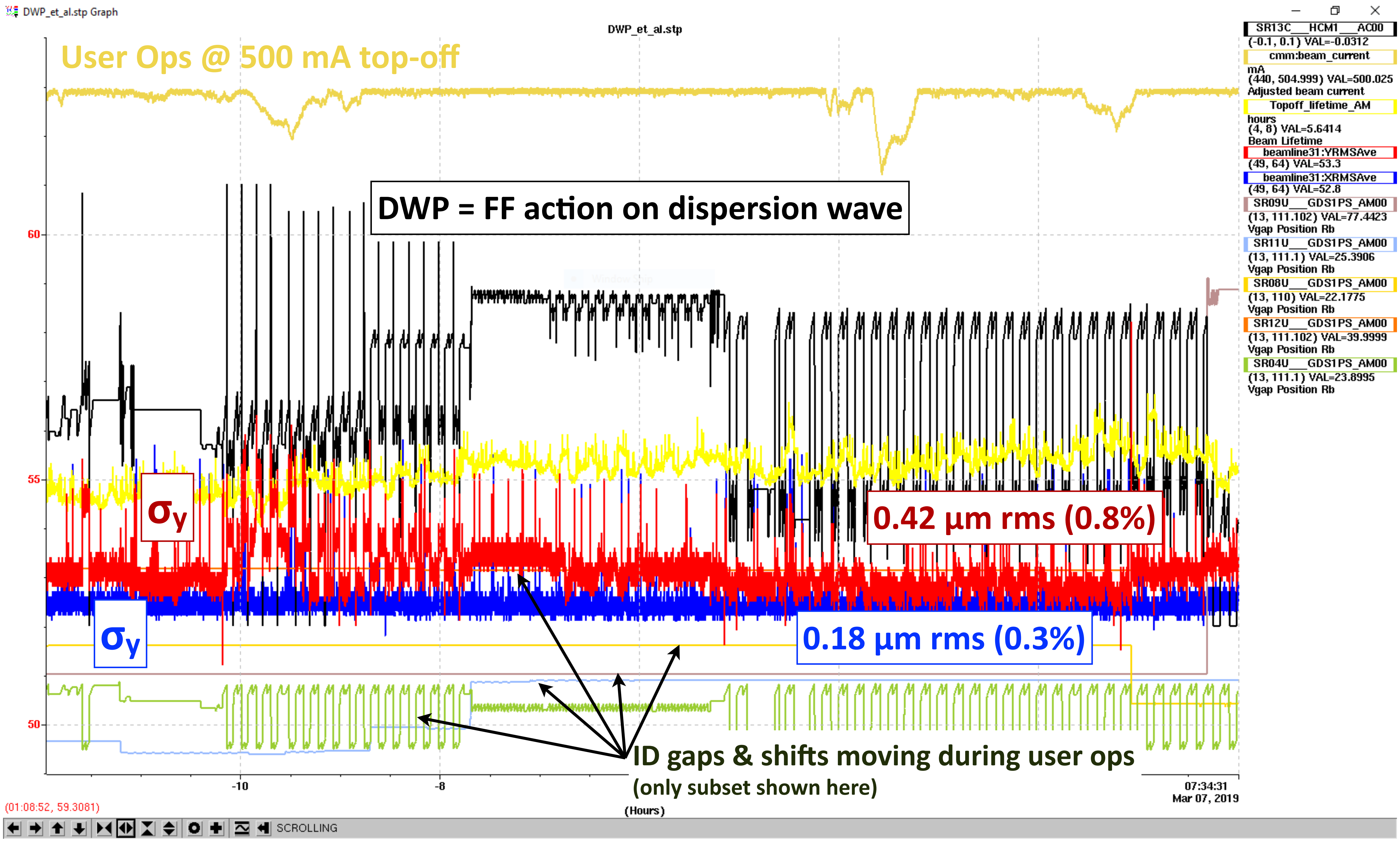
Evaluating NN-Based FF During AP Shift



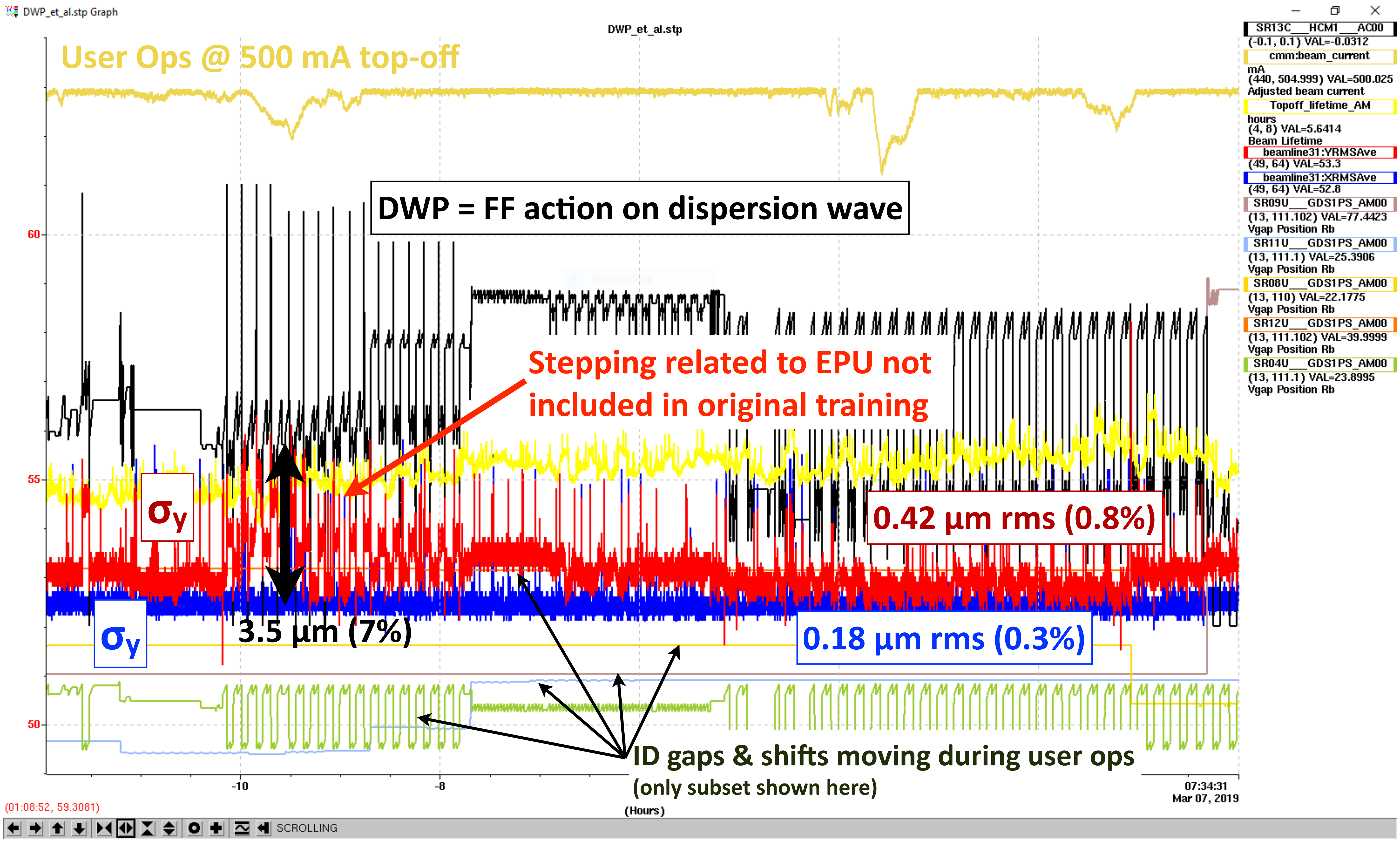
Evaluating NN-Based FF During AP Shift (cont.)



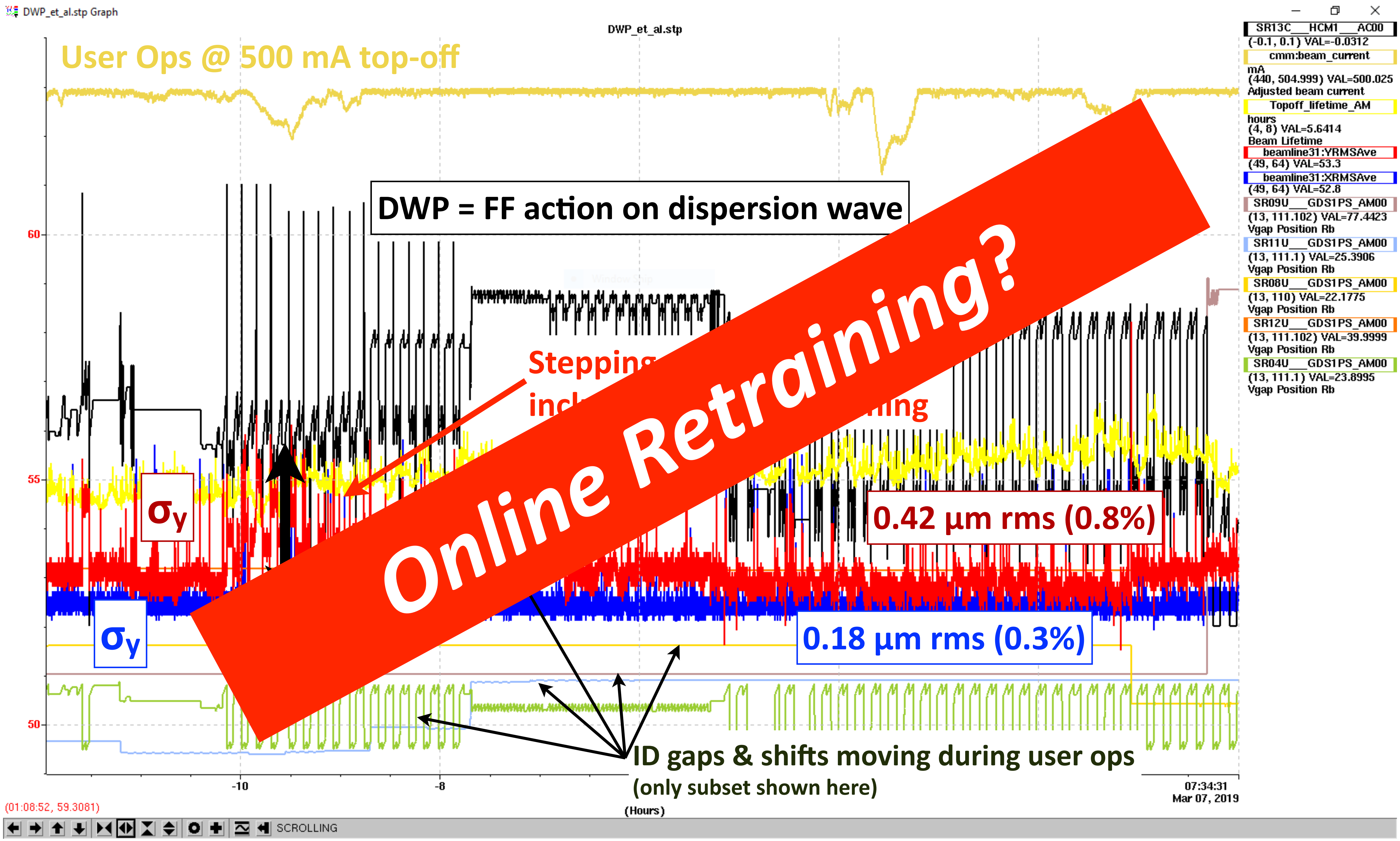
During User Ops: Stabilization Confirmed



During User Ops: Stabilization Confirmed, but...

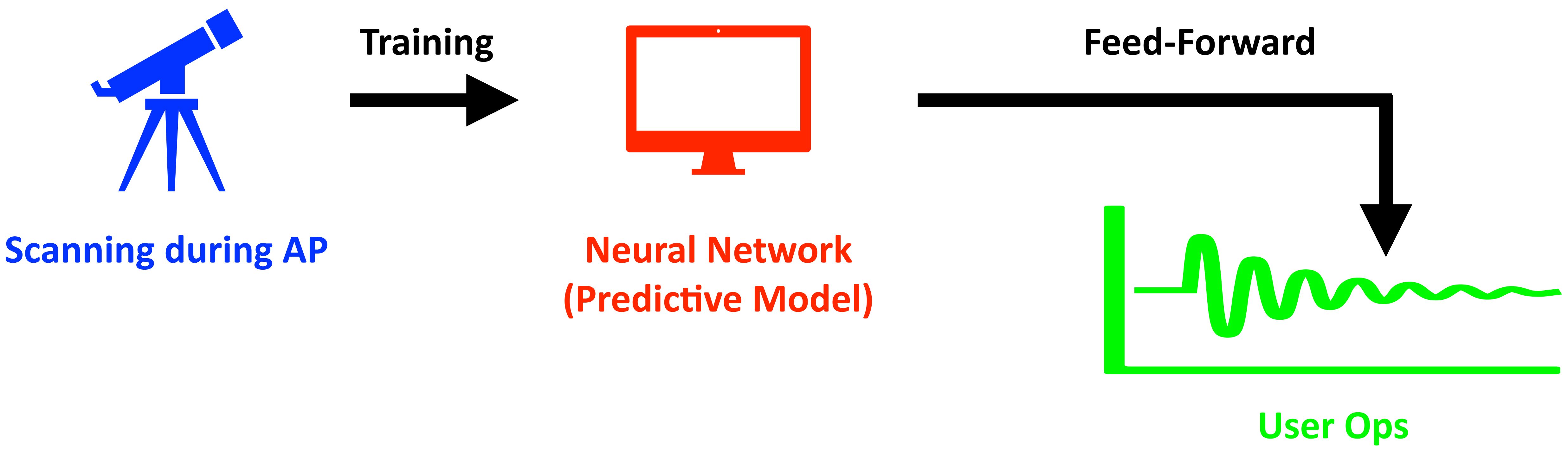


During User Ops: Stabilization Confirmed, but…



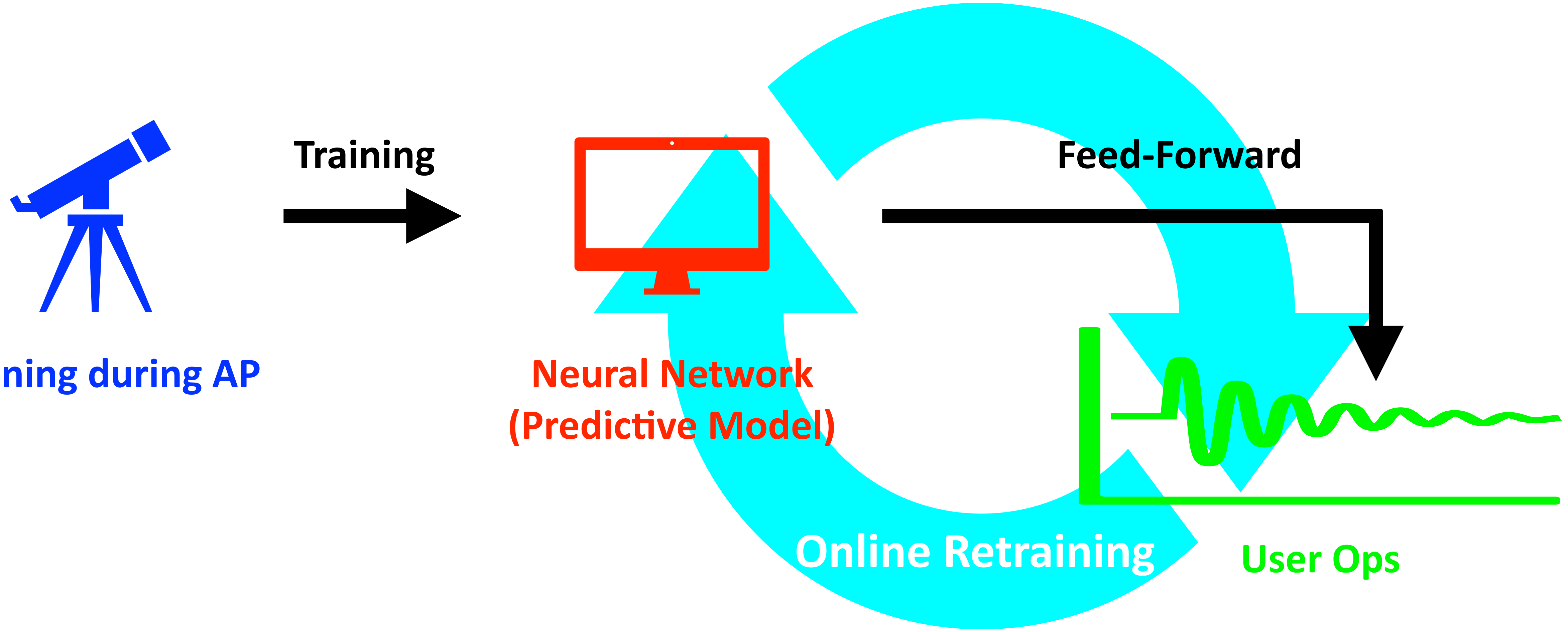
So what's different with online retraining?

"Conventional" Machine Learning



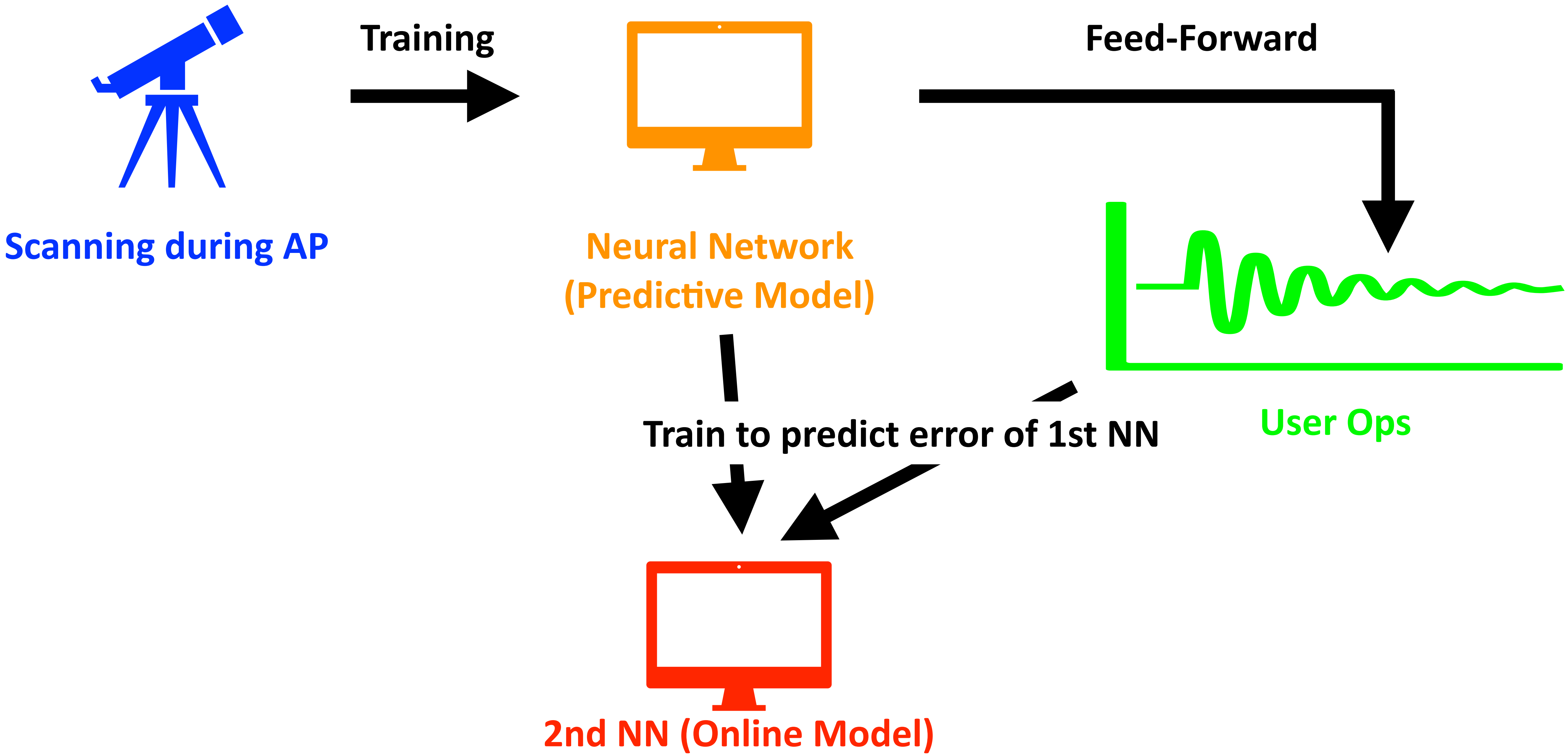
So what's different with online retraining? (cont.)

Online Retraining (Simple Approach)



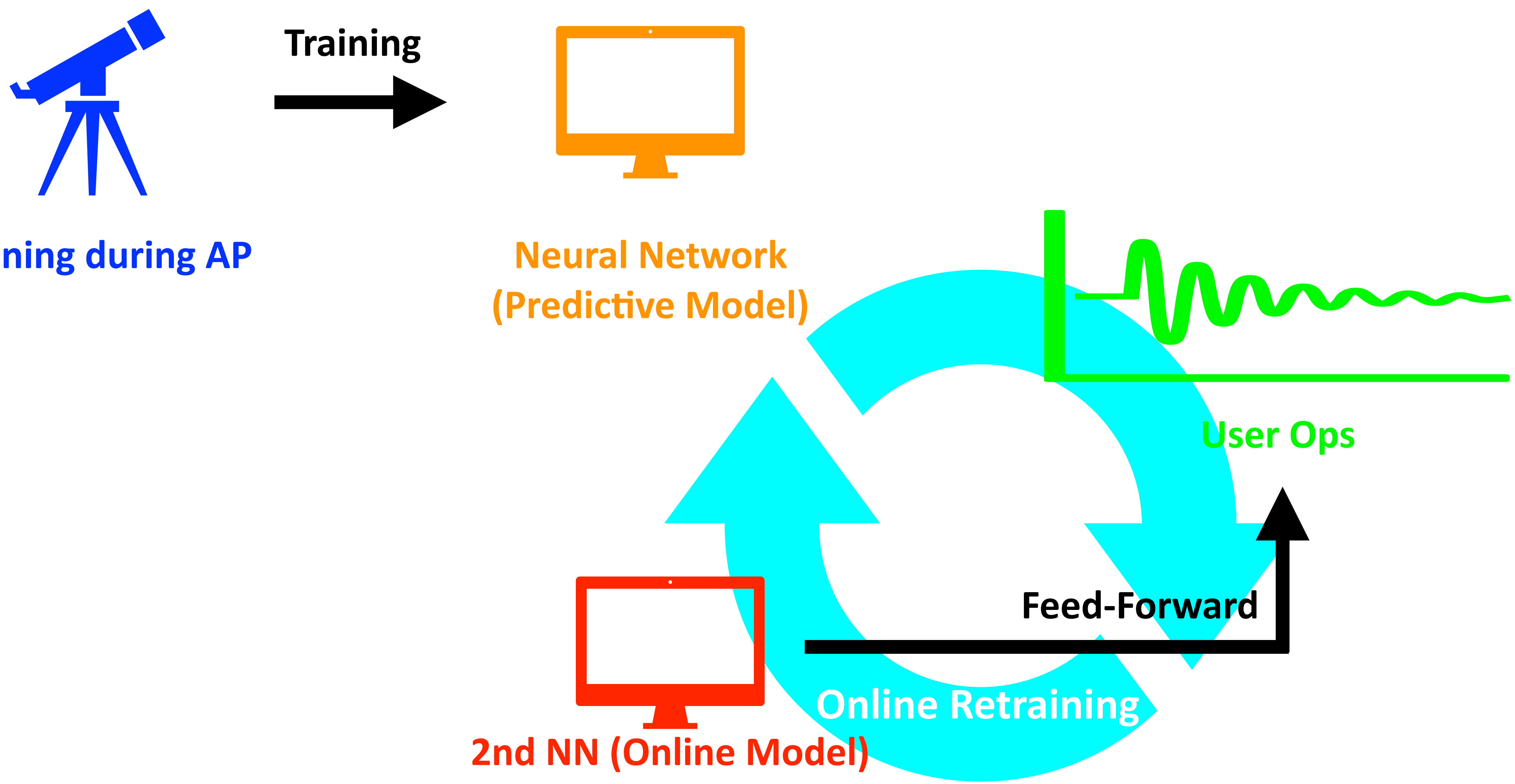
So what's different with online retraining? (cont.)

Online Retraining (Smarter Approach), Step 1



So what's different with online retraining? (cont.)

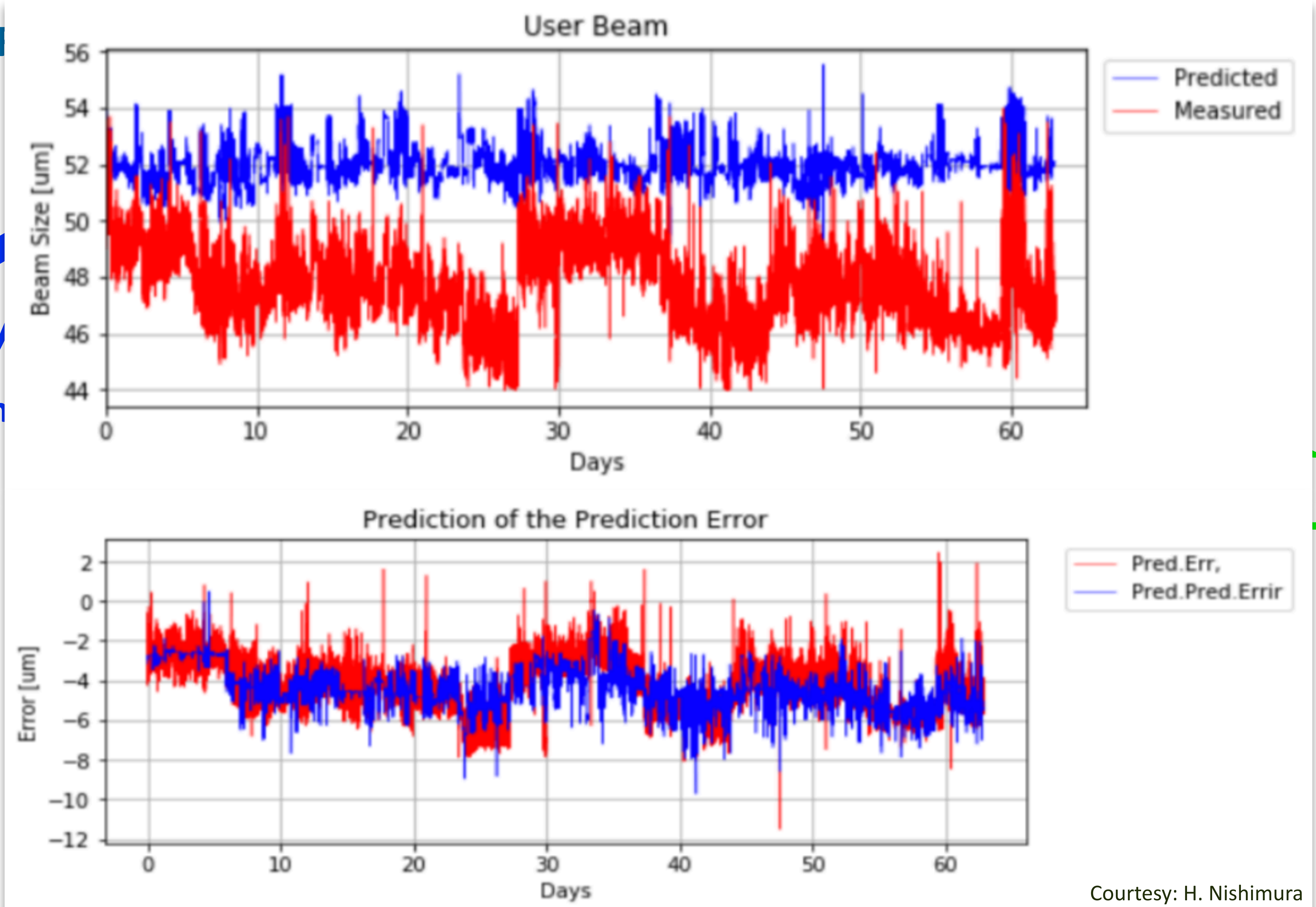
Online Retraining (Smarter Approach), Step 2



So what's different with online retraining? (cont.)

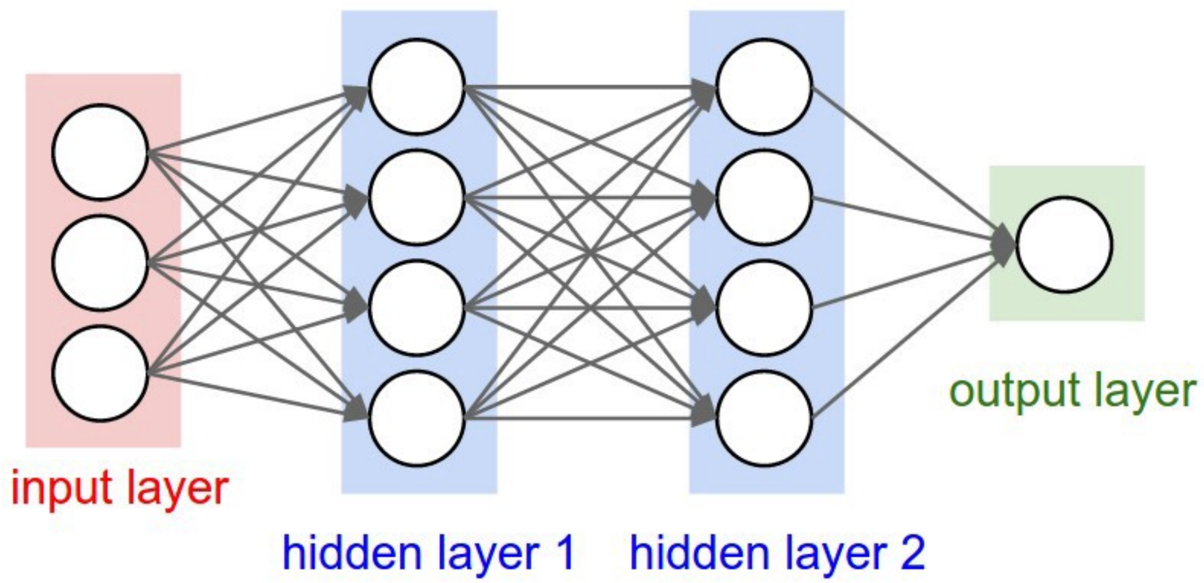
Online I

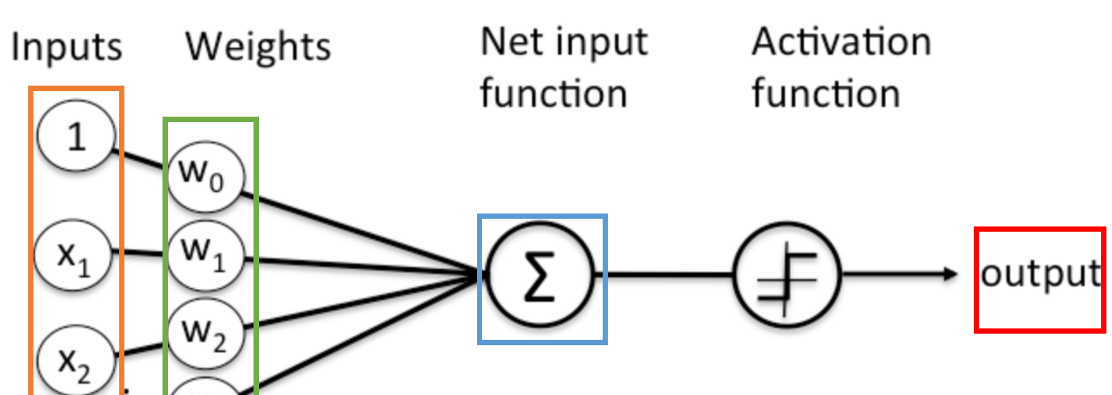
Scannin



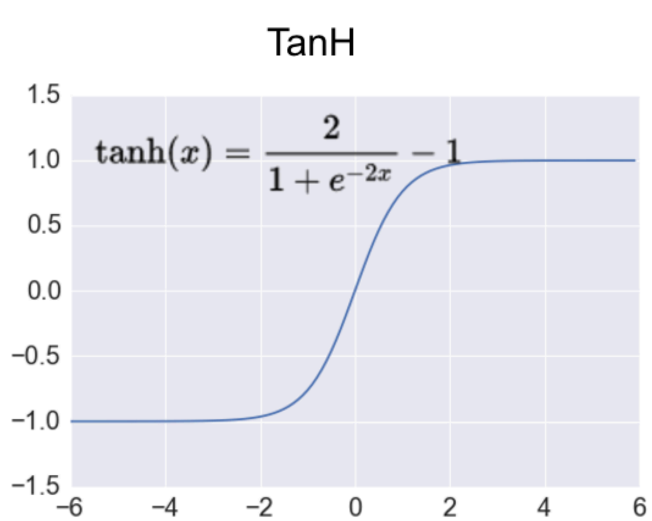
Courtesy: H. Nishimura

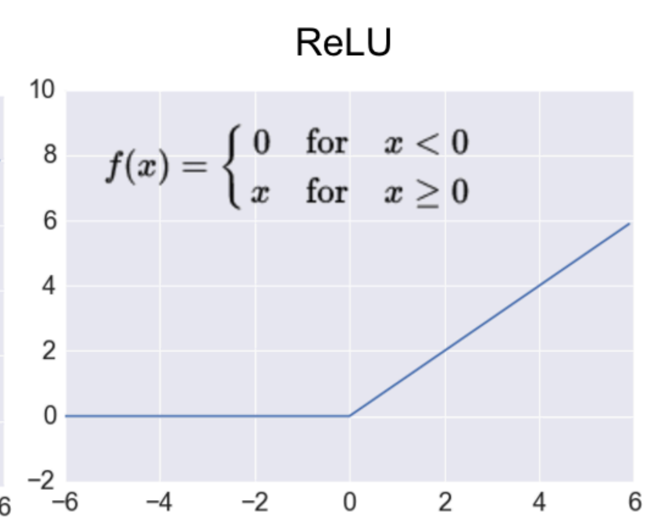
So what does this NN look like in detail?



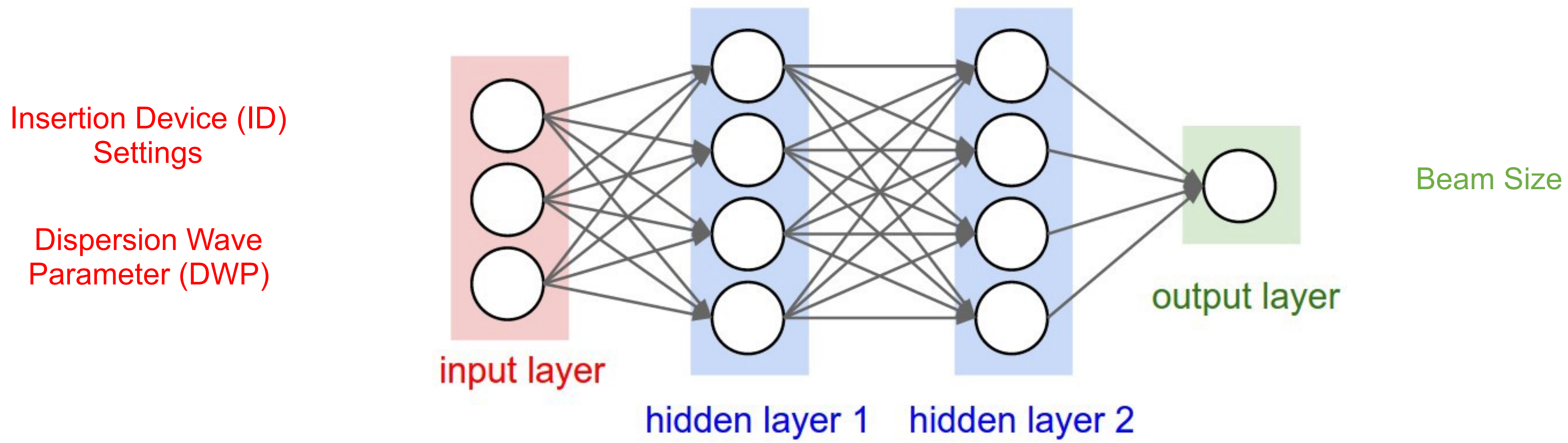


$$x_j^{l+1} = f\left(\sum_j w_{ij} x_i^l\right)$$





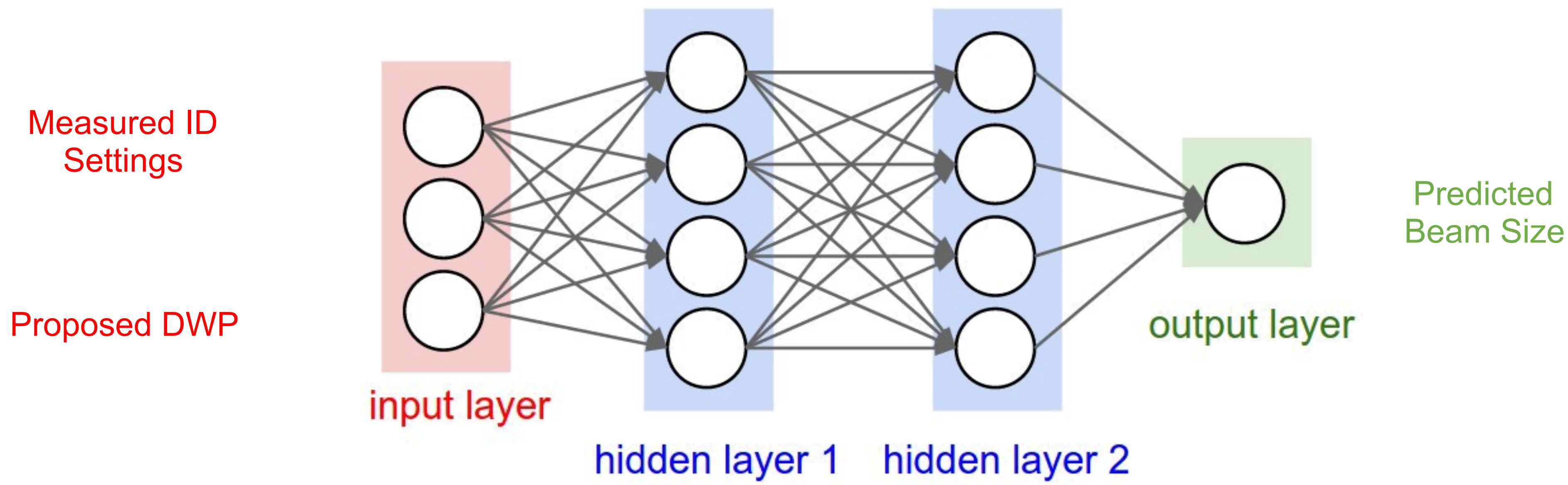
And how is it trained?



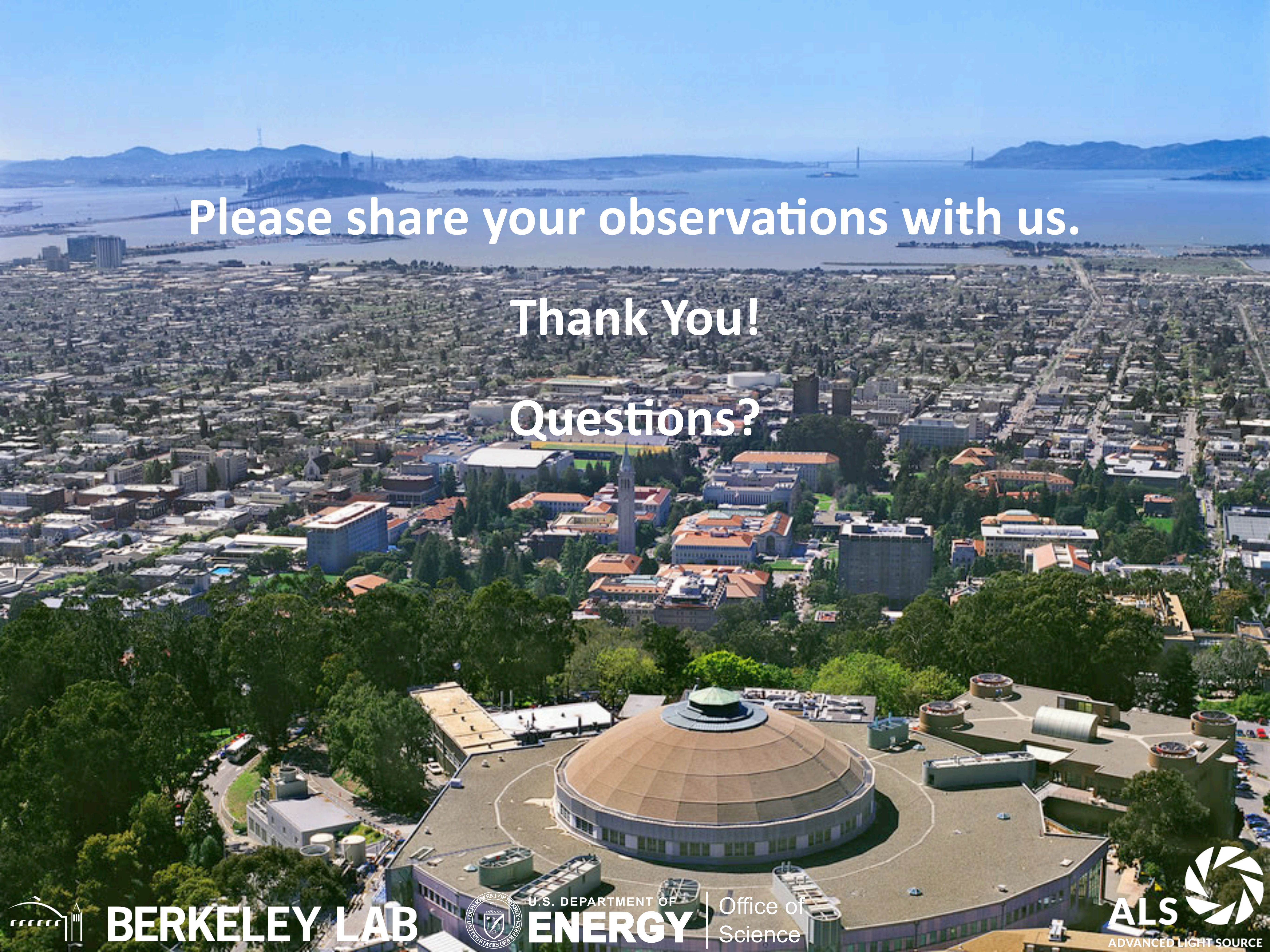
Input Layer: ID settings (22 Dimension) and DWP (1 Dimension)
Three Hidden Fully Connected Layers: 128, 64, 32 neurons in each layer
Output Layer: Vertical Beam Size (1 Dimension)

Regularization: L_2 regularizer with $\lambda = 10^{-4}$
Optimization: Adam Optimizer with learning rate $\alpha = 10^{-3}$

So then how do we use it?



Proposed DWPs		Predicted Beam Sizes		
-0.06		50.3		
….		….		
0	Measured ID Settings → Neural Network	52.1	Compare with Target → Beam Size	Choose proper DWP
…		…		
0.06		54.0		



Please share your observations with us.

Thank You!

Questions?

BERKELEY LAB

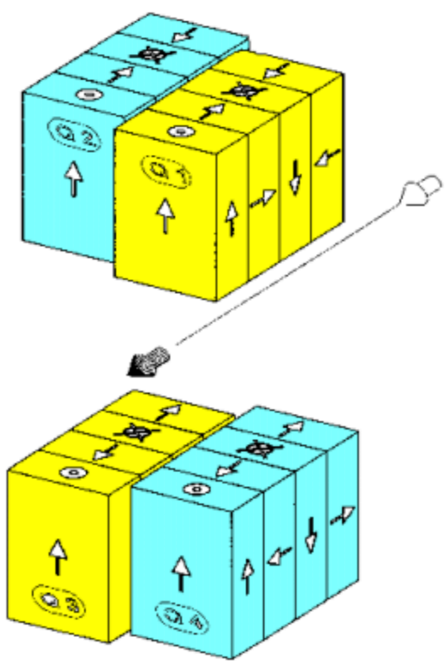
U.S. DEPARTMENT OF ENERGY | Office of Science

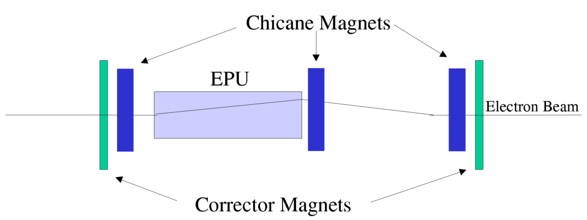
ALS ADVANCED LIGHT SOURCE

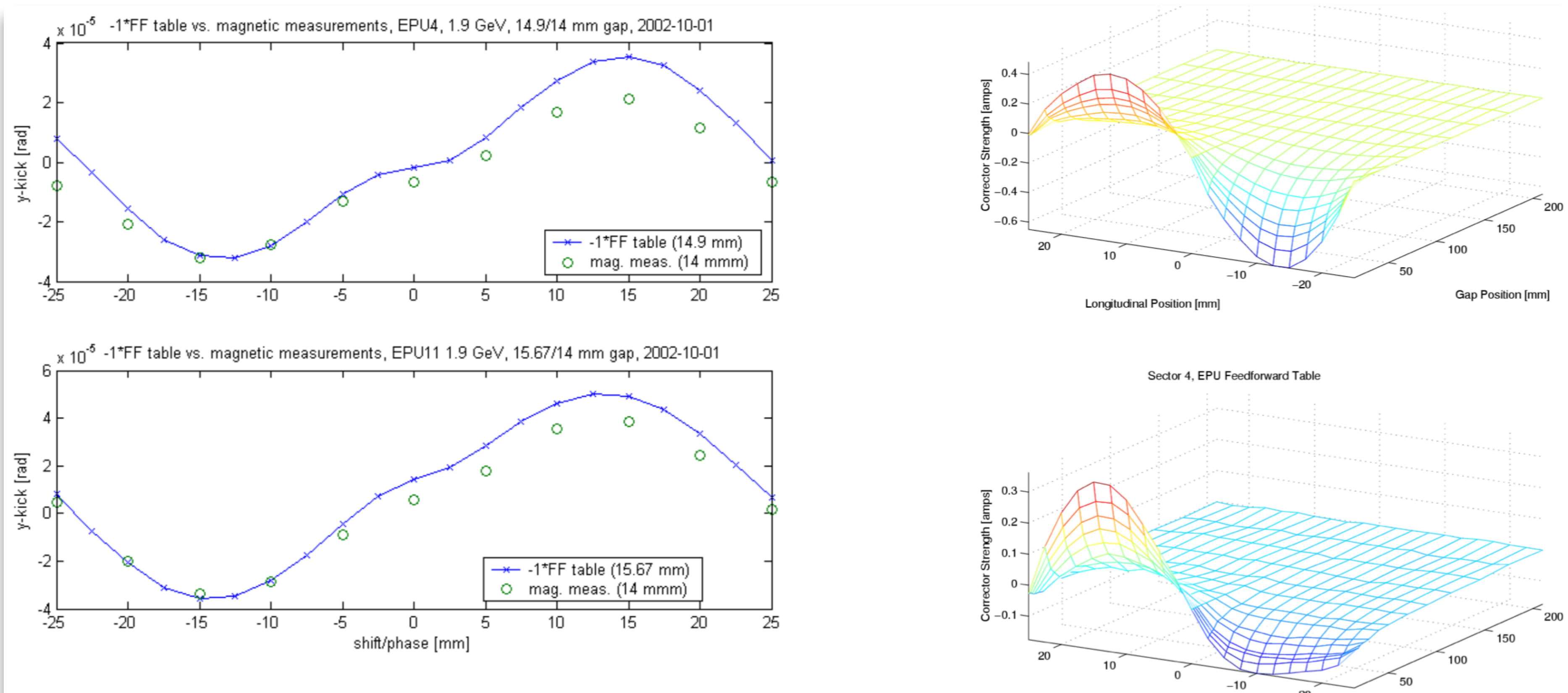
Backup Slides

Errors Caused by IDs & Required Corrections

- Orbit distortions
  - caused by on-axis variation of field integrals (with gap or EPU phase)
    - corrected by shims (magic fingers) & local orbit correctors (FF, 200 Hz)
    - corrected by ring corrector magnets (FB, ≈1 Hz SOFB & 1.1 kHz FOFB)



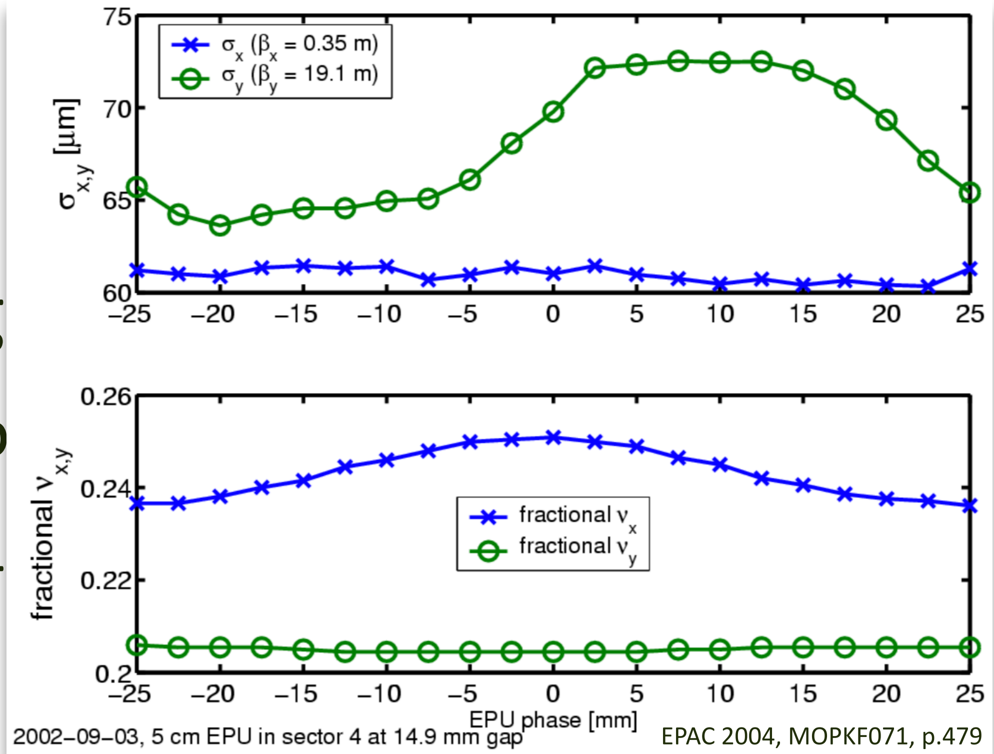




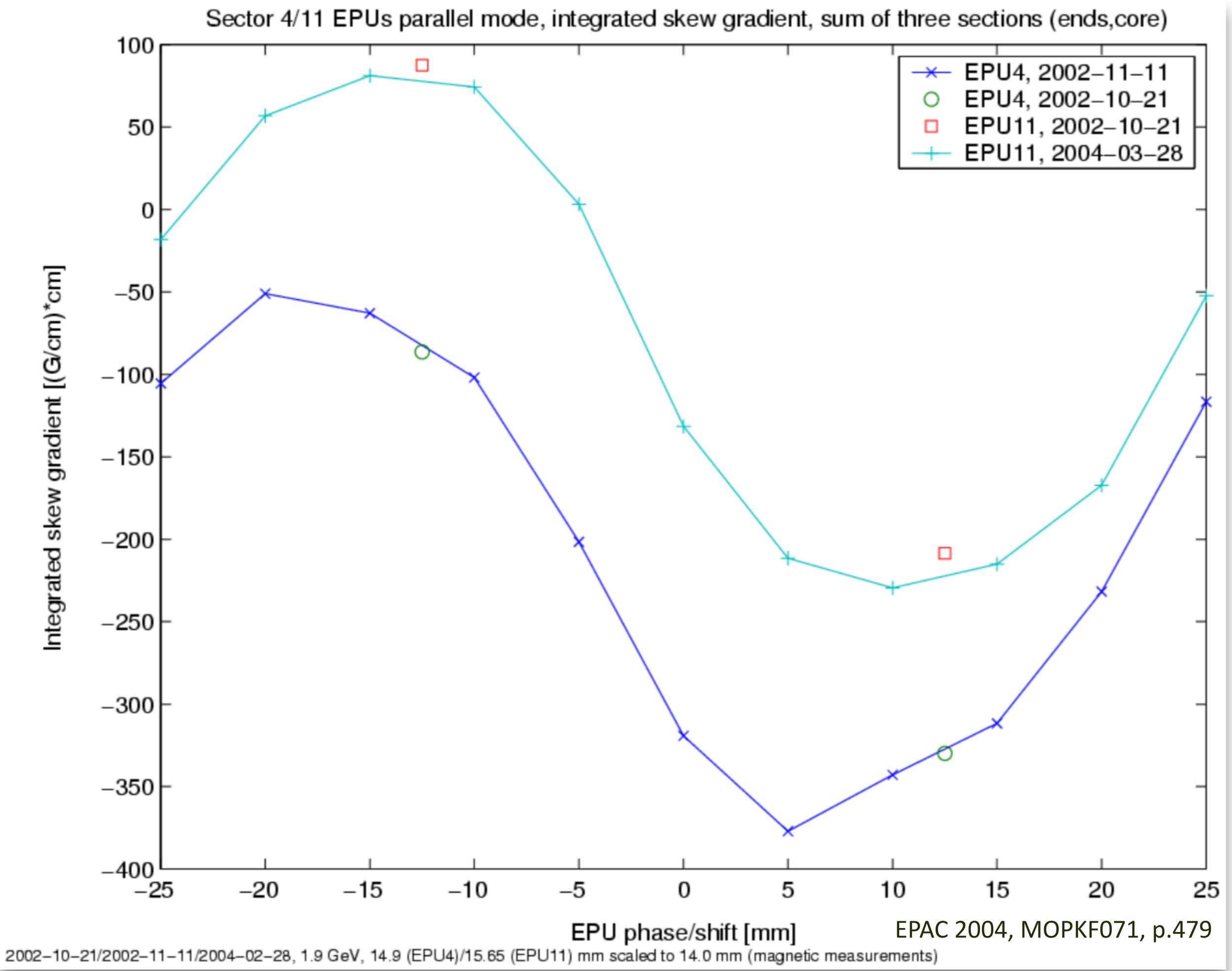
EPAC 2004, MOPKF071, p.479

Errors Caused by IDs & Required Corrections

- Orbit distortions
  - caused by on-axis variation of field integ
    - corrected by shims (magic fingers) & local o
    - corrected by ring corrector magnets (FB, ≈1
- Beam size (primarily vertical)
  - caused by variation of ID focusing terms (with gap or EPU phase)
    - corrected by local quad trims and global quad adjustment (FF & FB)



Errors Caused by IDs & Required Corrections



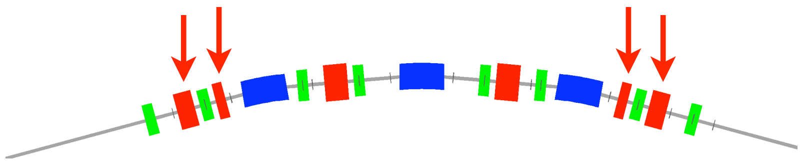
- Orbit distortions
  - caused by on-axis variation of fie
    - corrected by shims (magic fingers)
    - corrected by ring corrector magnet
- Beam size (primarily vertical)
  - caused by variation of ID focusing
    - corrected by local quad trims and g
  - caused by variation of ID-induced coupling (usually with EPU phase)
    - corrected by local skew quad coils (FF)

Errors Caused by IDs & Required Corrections

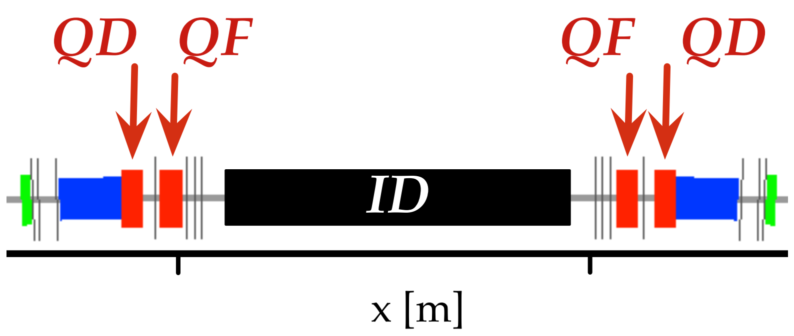
- Orbit distortions
  - caused by on-axis variation of field integrals (with gap or EPU phase)
    - corrected by shims (magic fingers) & local orbit correctors (FF, 200 Hz)
    - corrected by ring corrector magnets (FB, ≈1 Hz SOFB & 1.1 kHz FOFB)
- Beam size (primarily vertical)
  - caused by variation of ID focusing terms (with gap or EPU phase)
    - corrected by local quad trims and global quad adjustment (FF & FB)
  - caused by variation of ID-induced coupling (usually with EPU phase)
    - corrected by local skew quad coils (FF)
- Reduced injection efficiency & lifetime (nonlinear beam dynamics)
  - caused by higher-order ID effects (eg. field roll-off) → sets requirements for ID design and machine optics

ID Focusing Corrections Implemented in ALS ID FF

- Global Corrections
  - tunes (using lattice quads: 24 QF & 24 QD)
  - in addition: **tune FB** using same quads



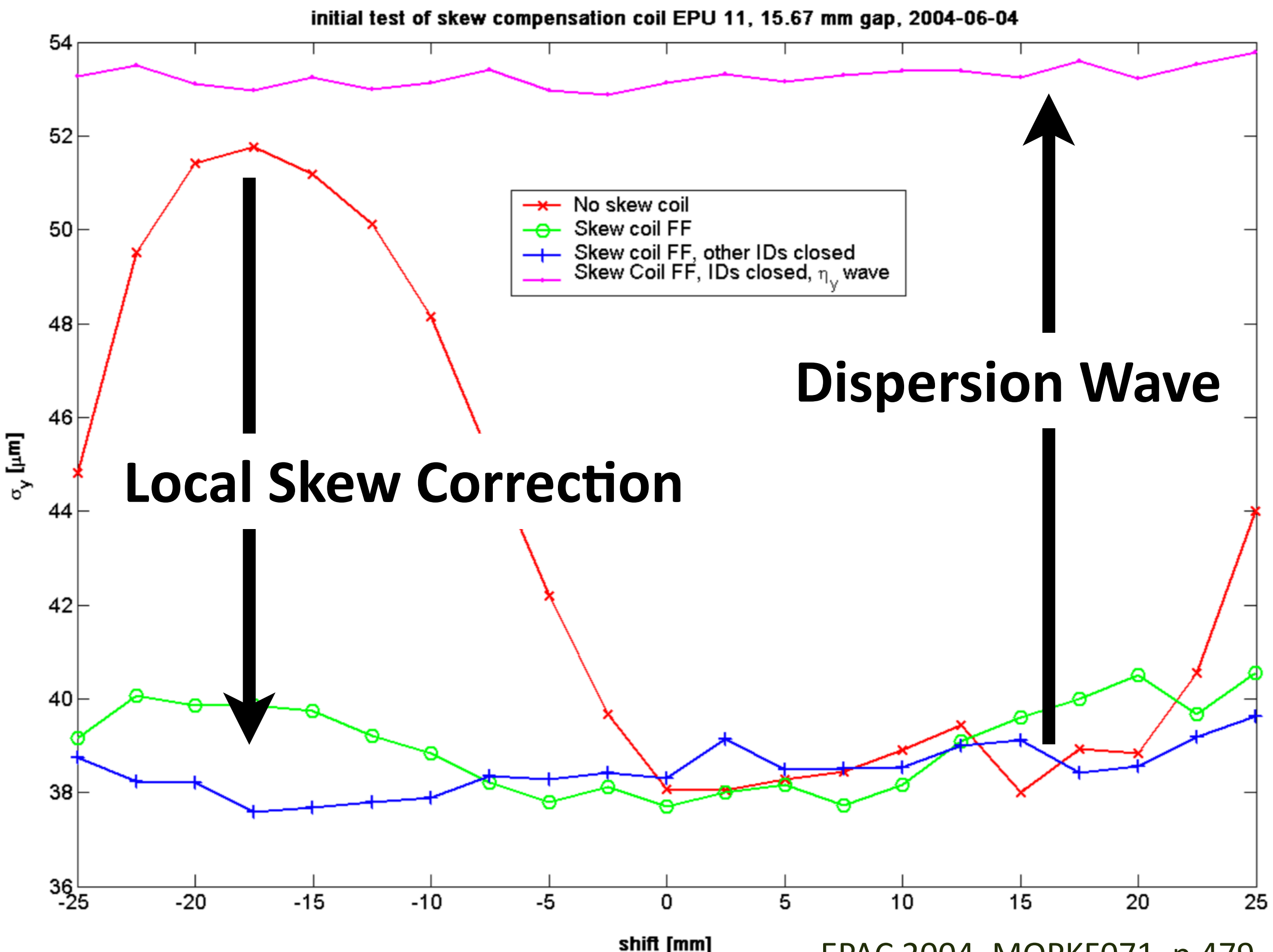
- Local Corrections for all IDs
  - β_y beat (using 2 QF & 2 QD locally) ➔ slightly increases $\Delta\nu_y$ ➔ can be removed by global tune correction



- Local Corrections for EPUs only
  - β_x beat (using 2 QF & 2 QD locally) ➔ locally also corrects $\Delta\nu_x$ since $\beta_x \approx 21$ m

Vertical Dispersion Wave Determines Effective ε_y

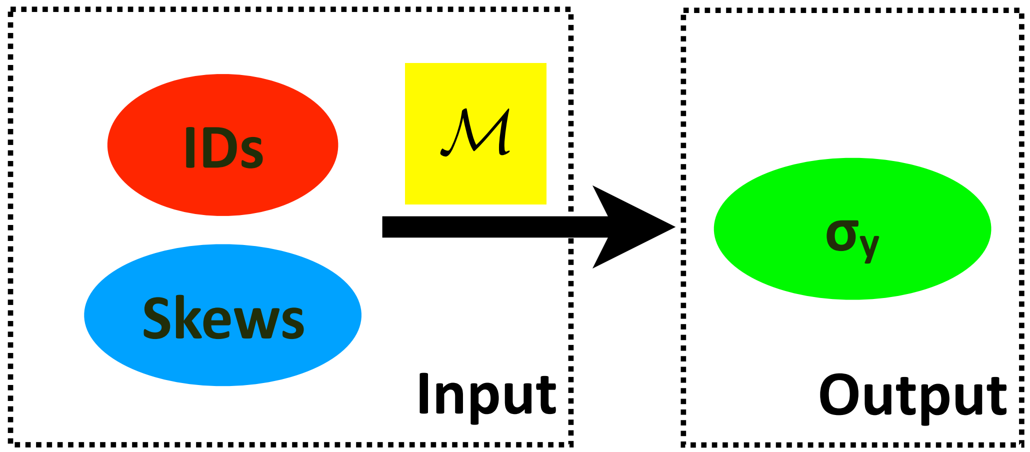
- Vertical source size is determined by
  - optics and coupling (local)
  - vertical emittance (global) consisting of
    - natural contribution (emission of SR is quantum process)
    - imperfections (unavoidable in real machines)
    - systematic η_y contributions (Dispersion Wave)



EPAC 2004, MOPKF071, p.479

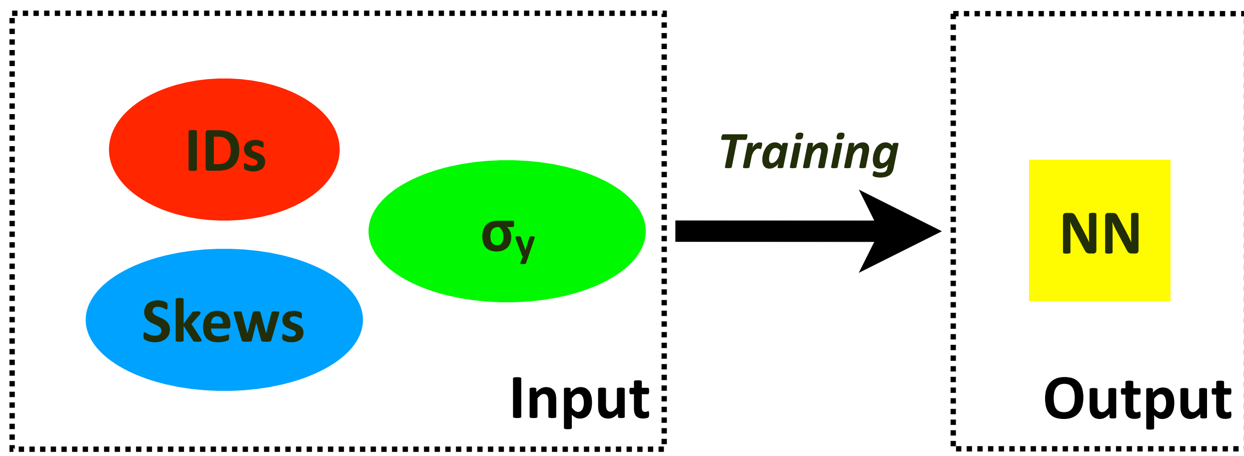
For Accelerators Deep Learning is a Paradigm Shift

Traditional Approach



- *Requires physics/model knowledge*
- *Tends to rely on many (sometimes poor) assumptions*

Deep Learning



- *Requires only large amounts of data & reproducibility*

Application during ops: