Operational Integration of Machine Learning for Beam Size Control in the Advanced Light Source

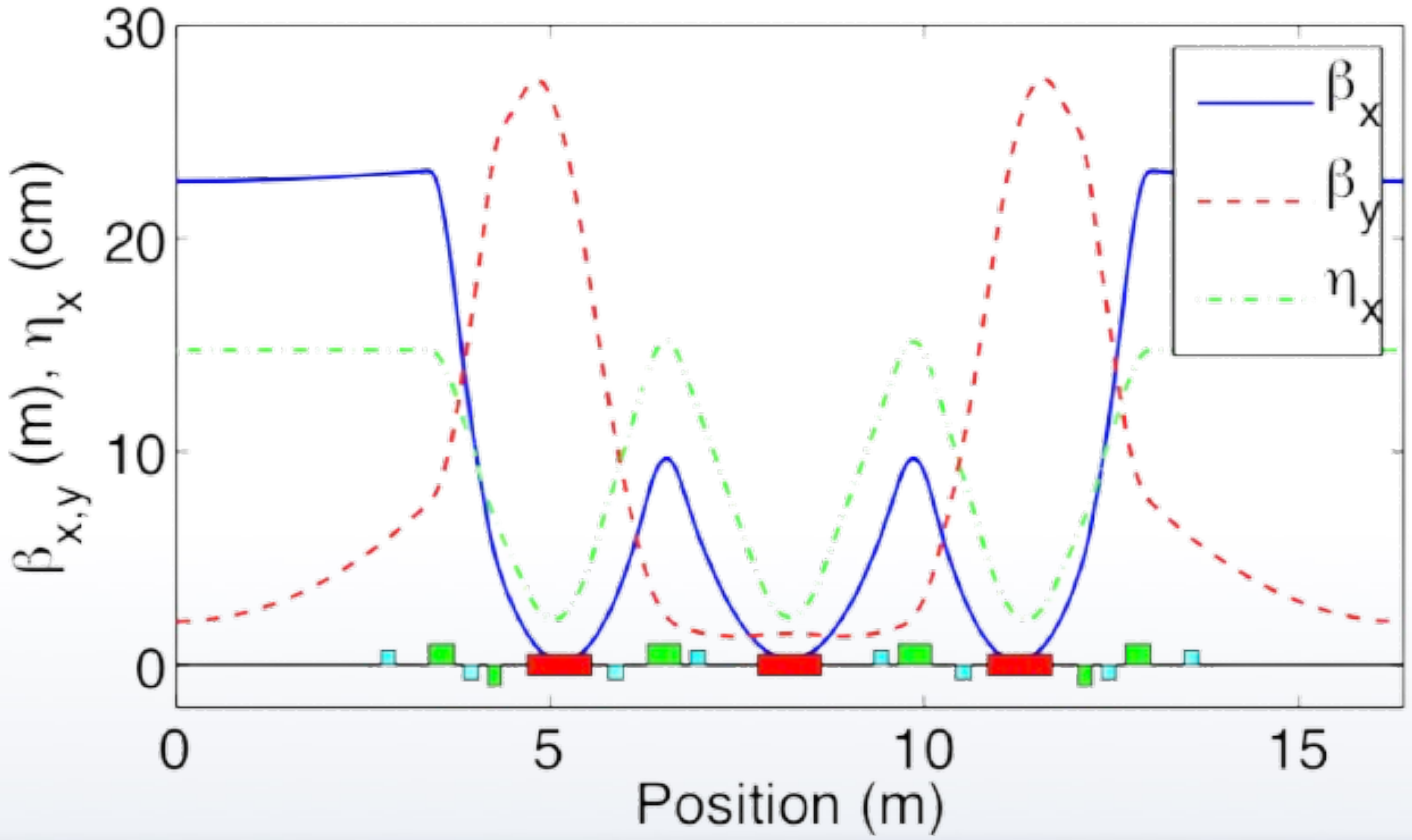
Thorsten Hellert, Tynan Ford, Simon C. Leemann, Hiroshi Nishimura, Andrea Pollastro, Marco Venturini

ALS Accelerator Physics Group

06.03.24

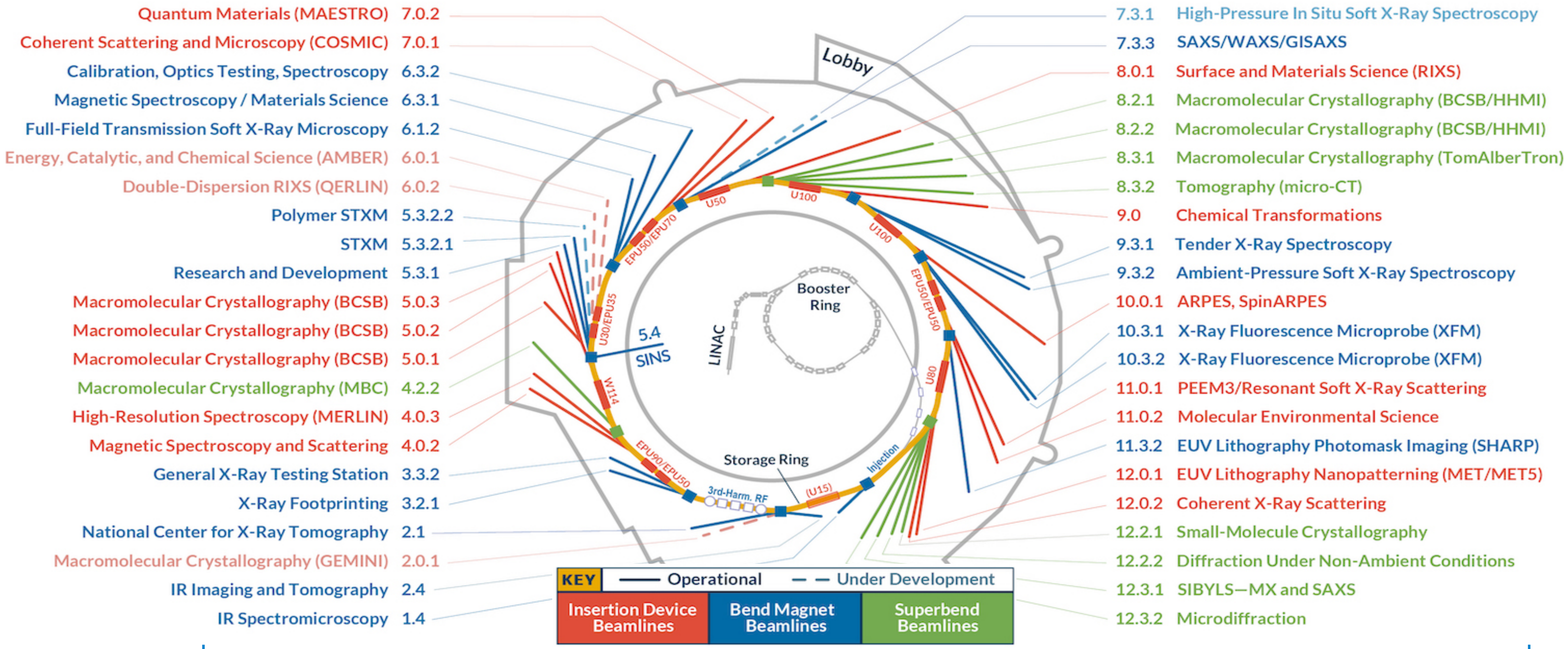
PARAMETER	VALUE
Beam energy	1,9 GeV
Circumference	196.8 m
Beam current	500 mA
Horizontal emittance	2 nm·rad
Vertical emittance	0.04 nm·rad

ALS Triple-Bend Achromat Lattice



Insertion Devices at the Advanced Light Source

14 vertical gaps (1 always fixed) + 7 EPUs with 2x horizontal offsets = 27 free parameters



Electron Beam Stability at the Advanced Light Source

- Beam Current:
  - Top-off operation keeps current variations below 1mA
- Beam Position:
  - Orbit feedback and ID feed-forwards stabilize source positions to sub-micron level
- Beam Size:
  - ID skew quadrupole feed-forwards stabilize source size
  - Requires lookup tables

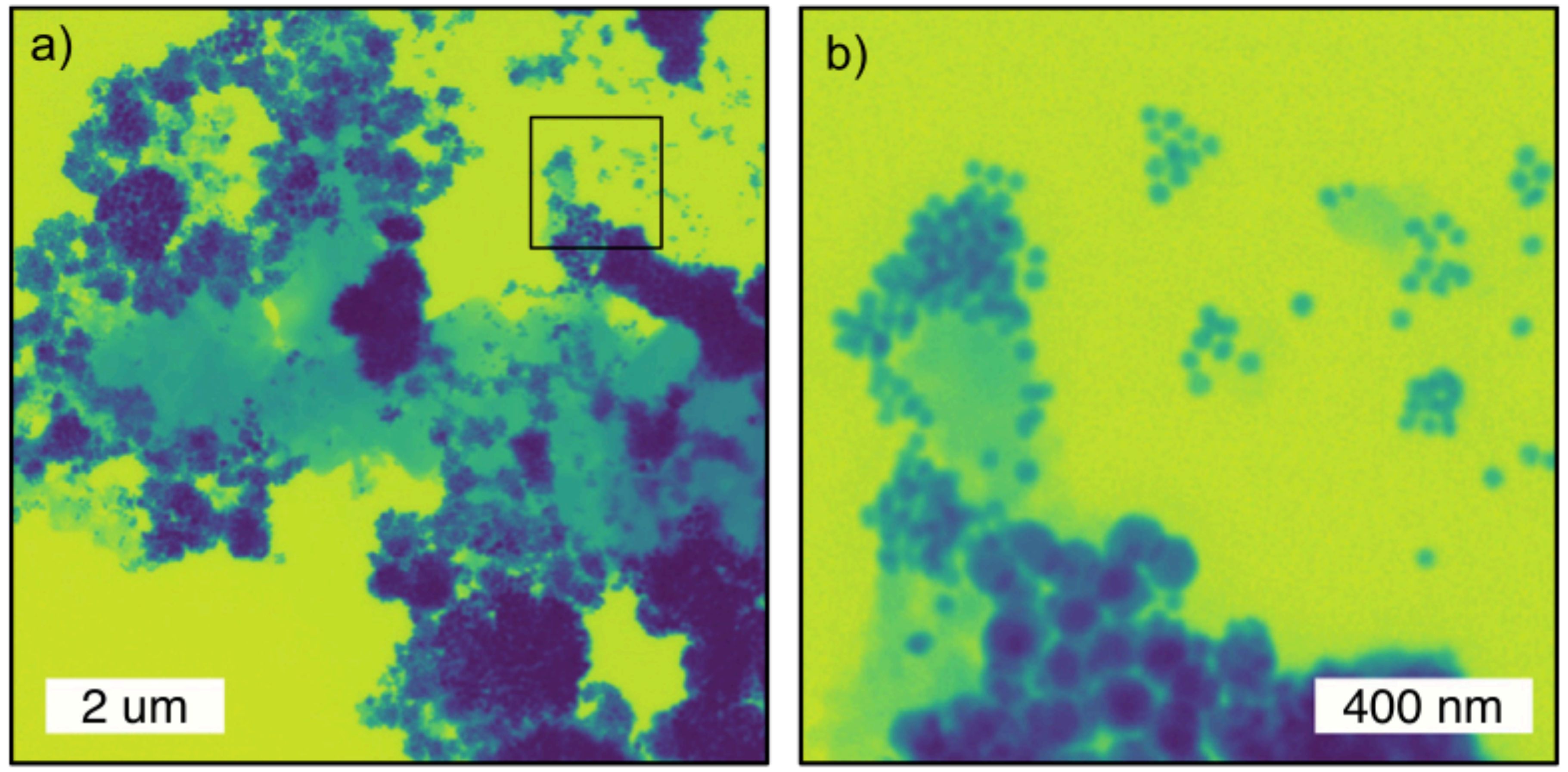
Scanning Transmission X-ray Microscopy at the Advanced Light Source

- STXM Beamlines:
  - Widely used for nanoscale studies
  - Fast raster scanning
  - No averaging
  - ≈ 1 ms/pixel, 1 s/line, 6 min/scan

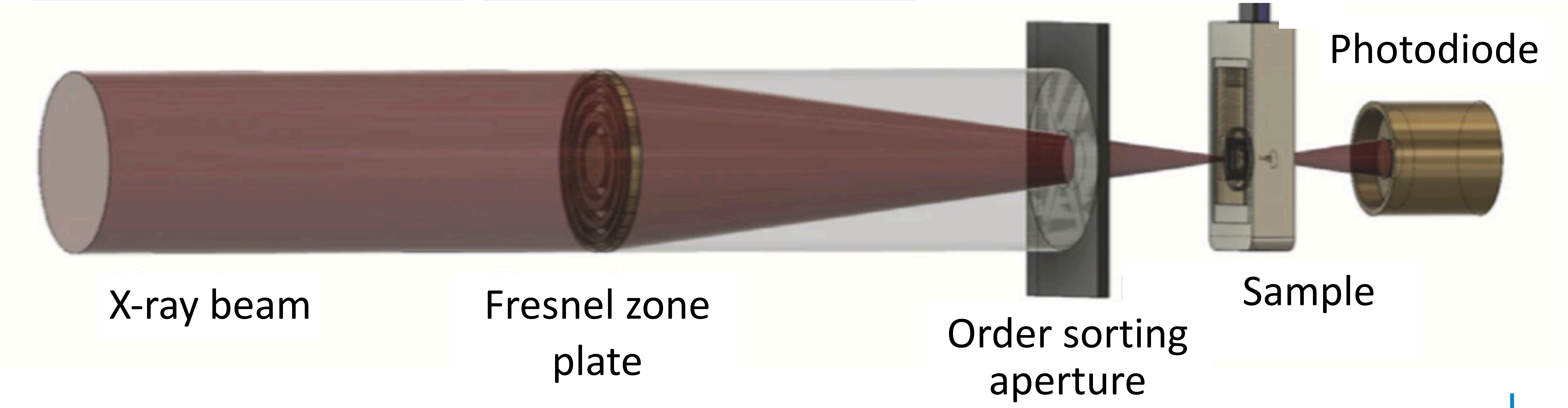
Contents lists available at ScienceDirect
ELSEVIER
Journal of Electron Spectroscopy and Related Phenomena
journal homepage: www.elsevier.com/locate/elspec

Scanning transmission X-ray microscopy at the Advanced Light Source

Thomas Feggeler [a,b,*], Abraham Levitan [c,b], Matthew A. Marcus [b], Hendrik Ohldag [b,d,e], David A. Shapiro [b]

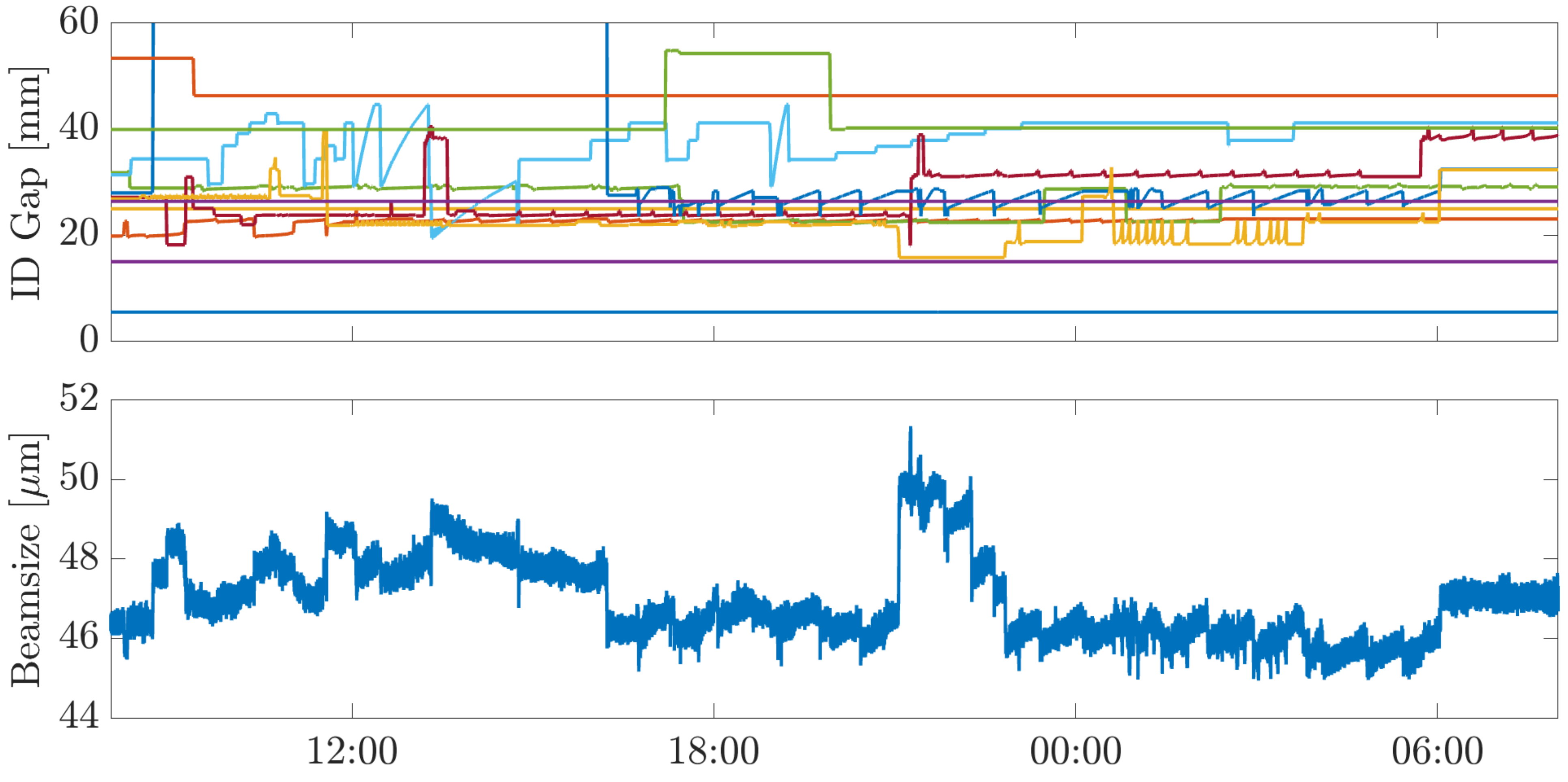


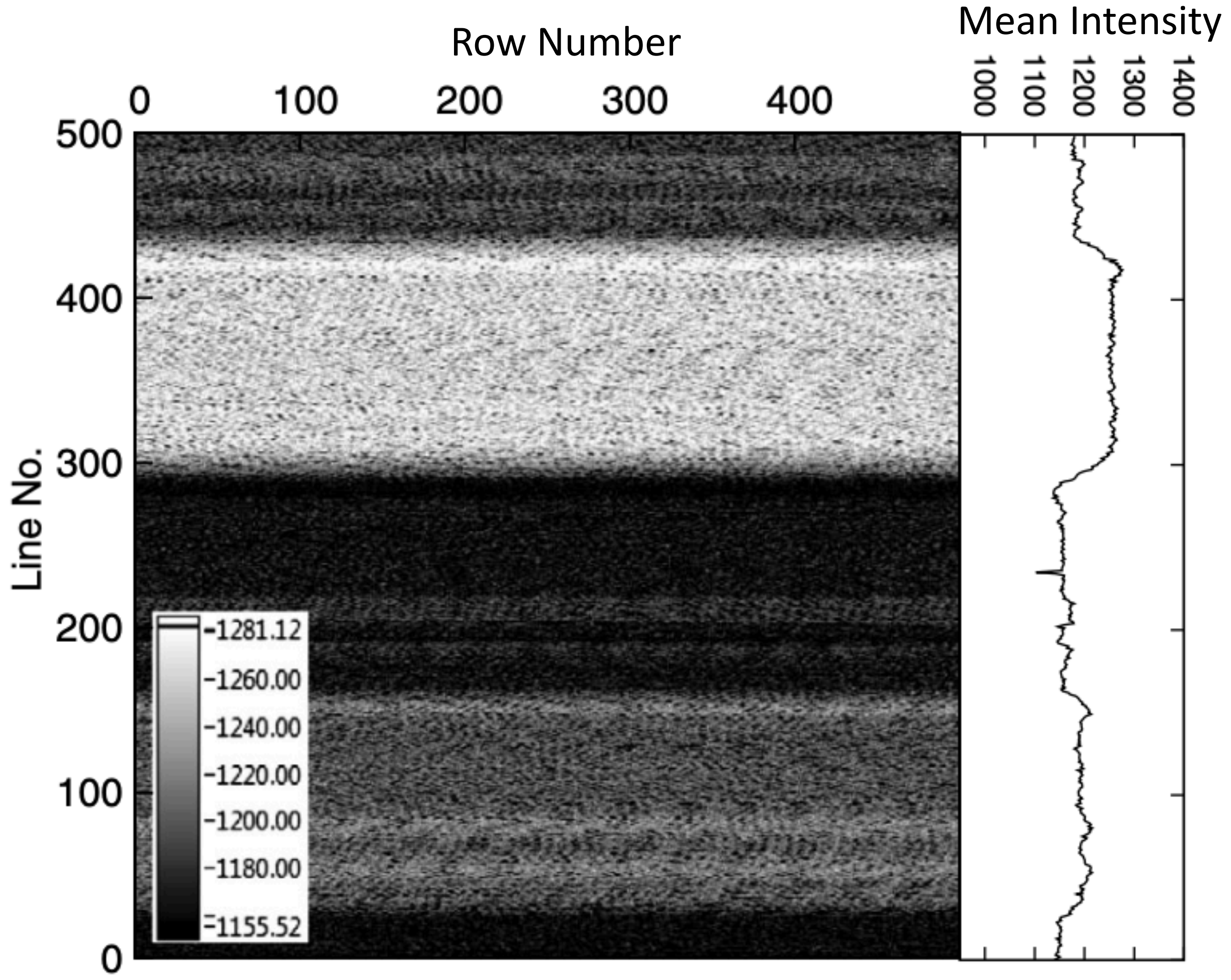
“Section of a ptychography reconstruction of 40nm and 100nm gold nanoparticles on a silicon nitride membrane“

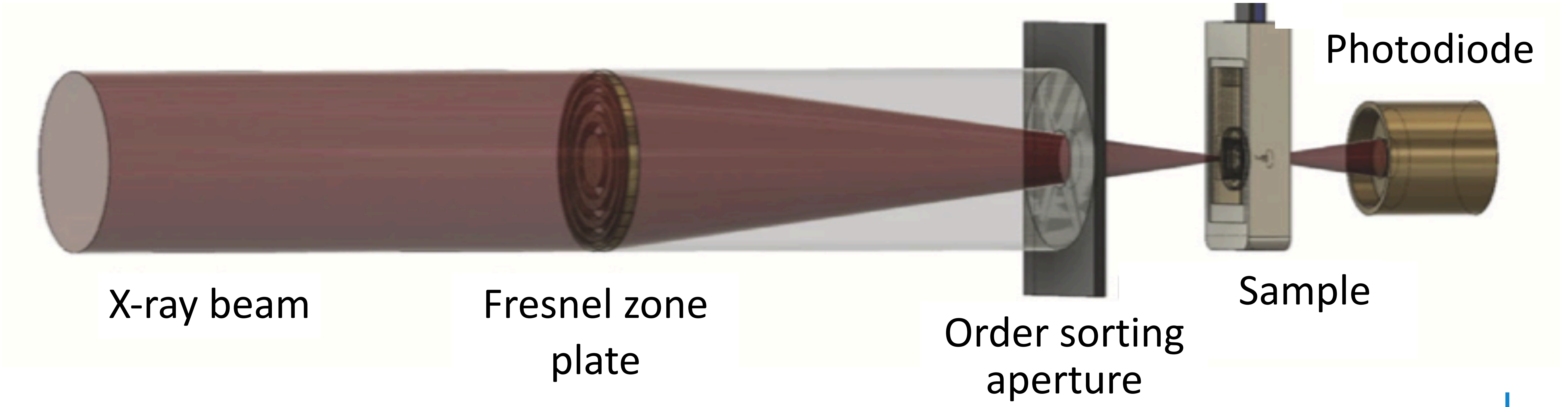


Scanning Transmission X-ray Microscopy at the Advanced Light Source

- STXM Beamlines:
  - Widely used for nanoscale studies
  - Fast raster scanning
  - No averaging
  - ≈ 1 ms/pixel, 1 s/line, 6 min/scan
- Current ID Feed Forward:
  - LOCO at various gaps to obtain (local) skew quadrupole corrector settings
  - Interpolation table used in FF correction
  - Takes one night to measure
  - Aging of tables results in large variation of vertical beam size

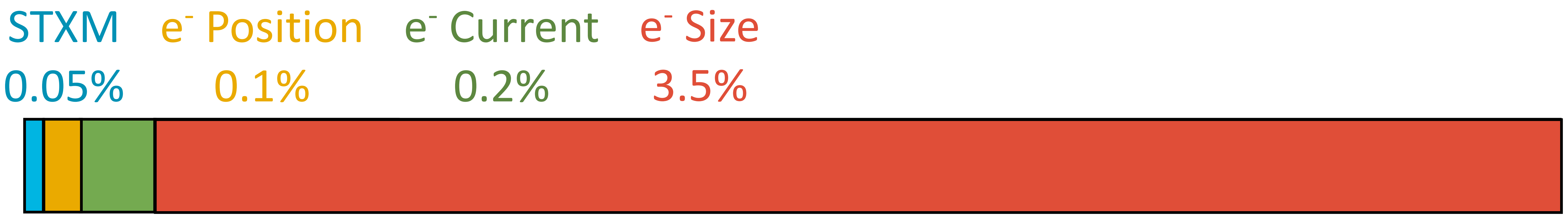


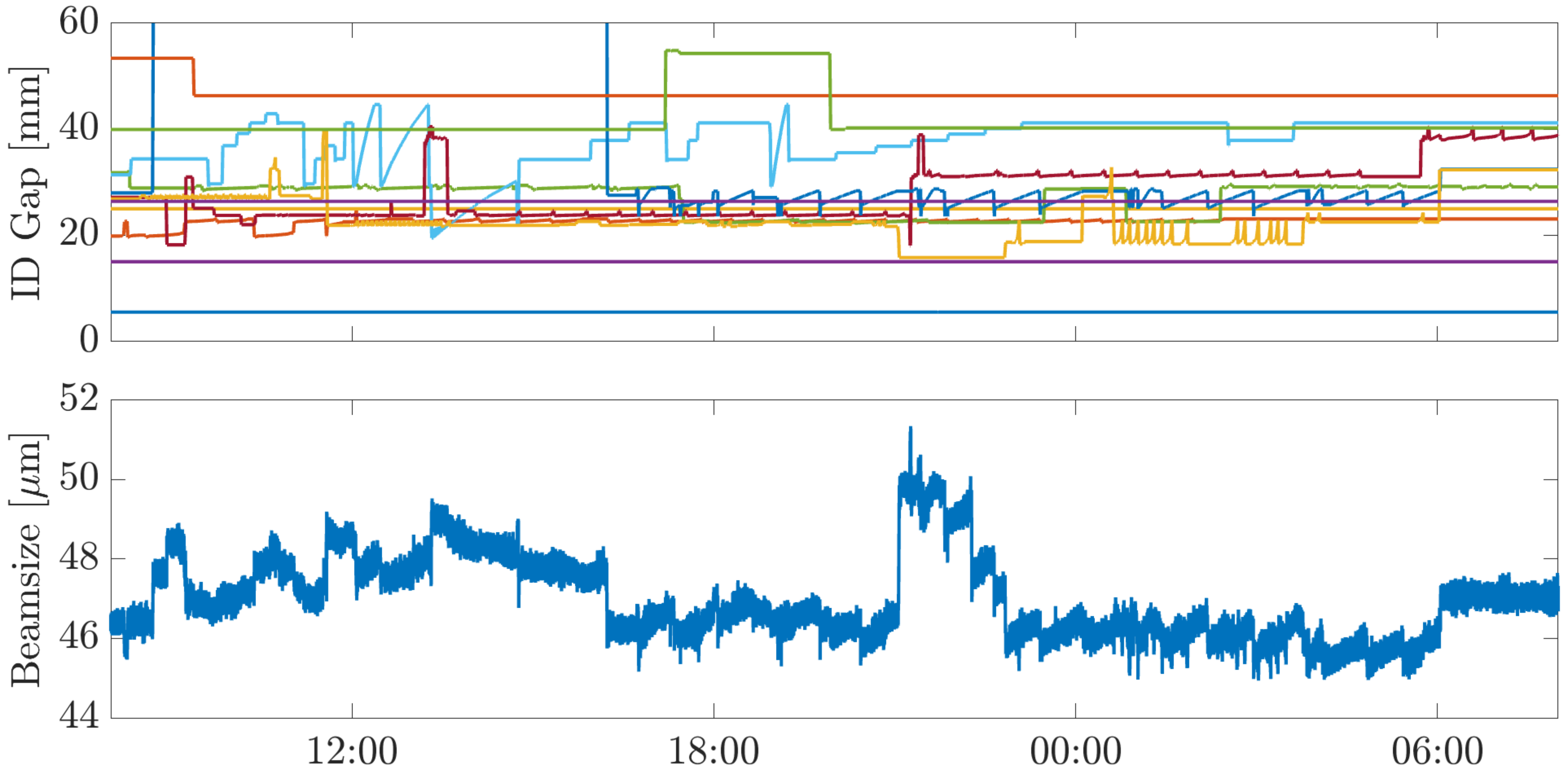


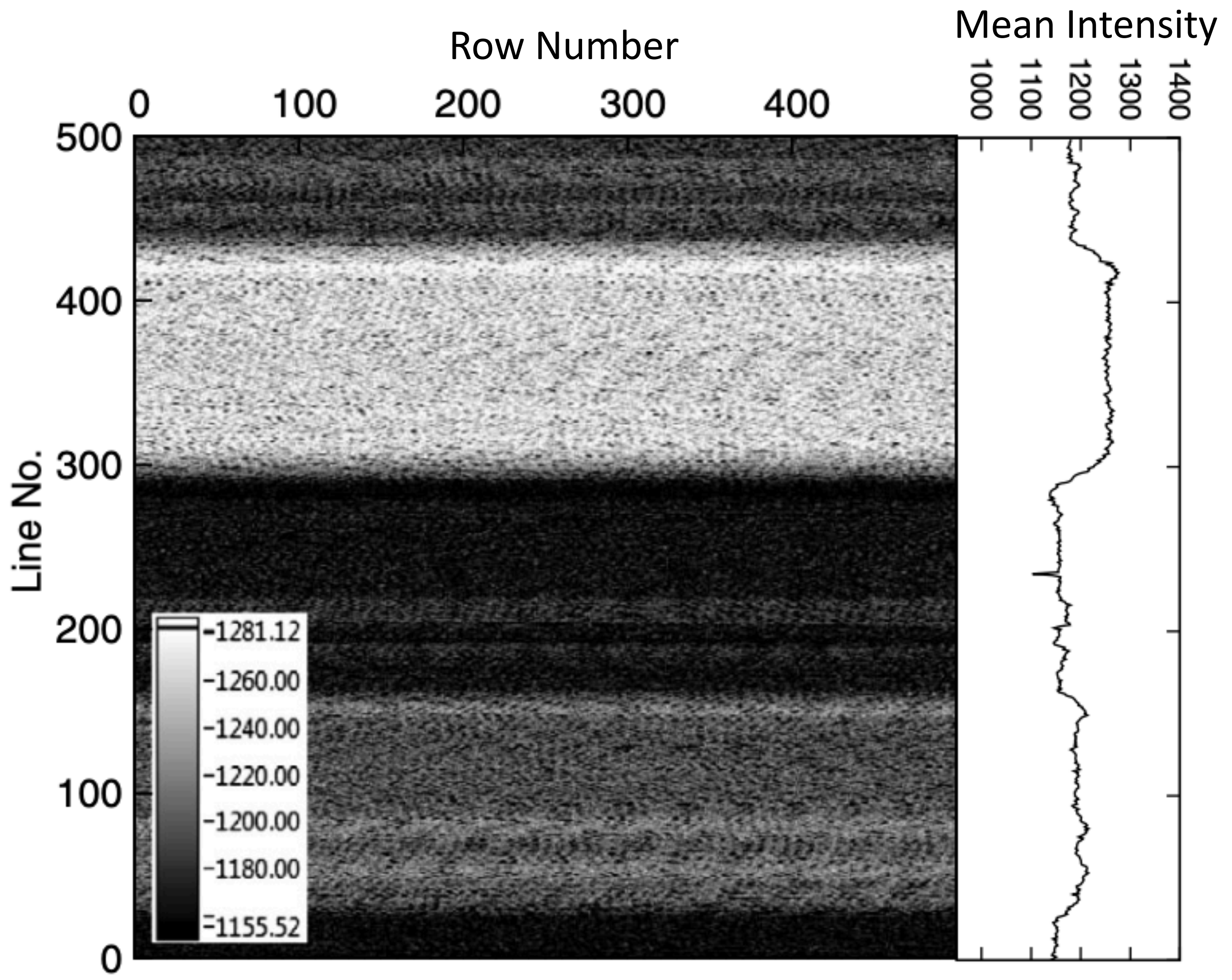


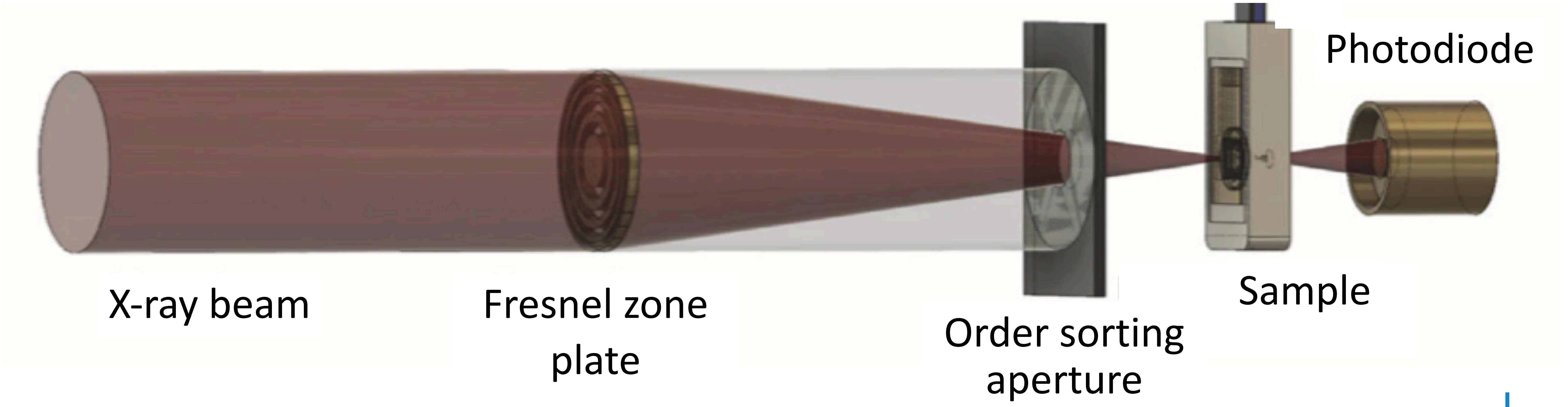
Scanning Transmission X-ray Microscopy at the Advanced Light Source

- STXM Beamlines:
  - Widely used for nanoscale studies
  - Fast raster scanning
  - No averaging
  - ≈ 1 ms/pixel, 1 s/line, 6 min/scan
- Current ID Feed Forward:
  - LOCO at various gaps to obtain (local) skew quadrupole corrector settings
  - Interpolation table used in FF correction
  - Takes one night to measure
  - Aging of tables results in large variation of vertical beam size
- STXM Intensity Fluctuations:
  - Residual ID induced vertical beamsize variations biggest contributor
  - POC: Leemann et al., PRL 123,194801



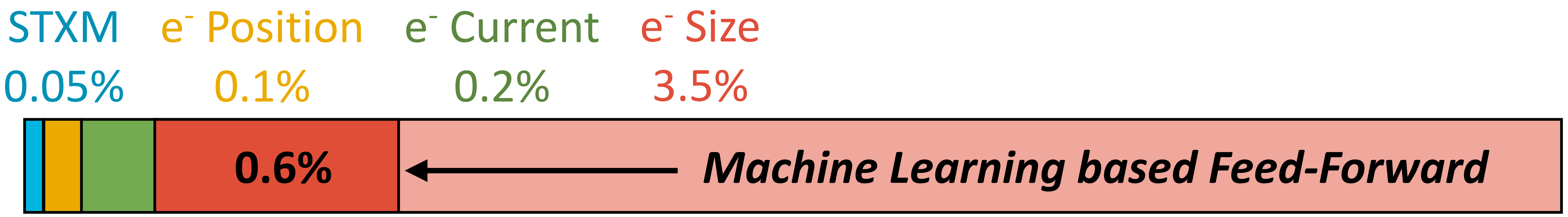


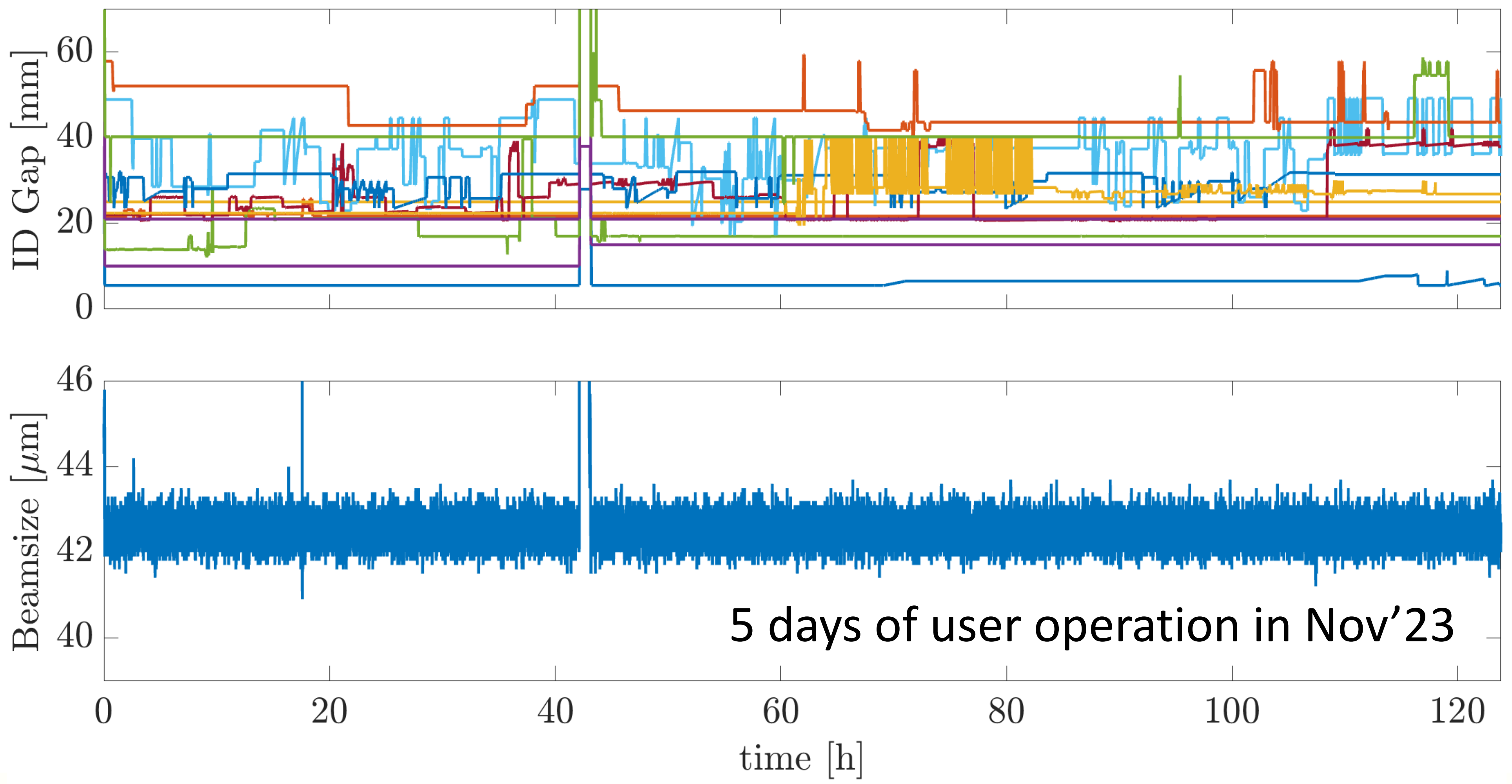


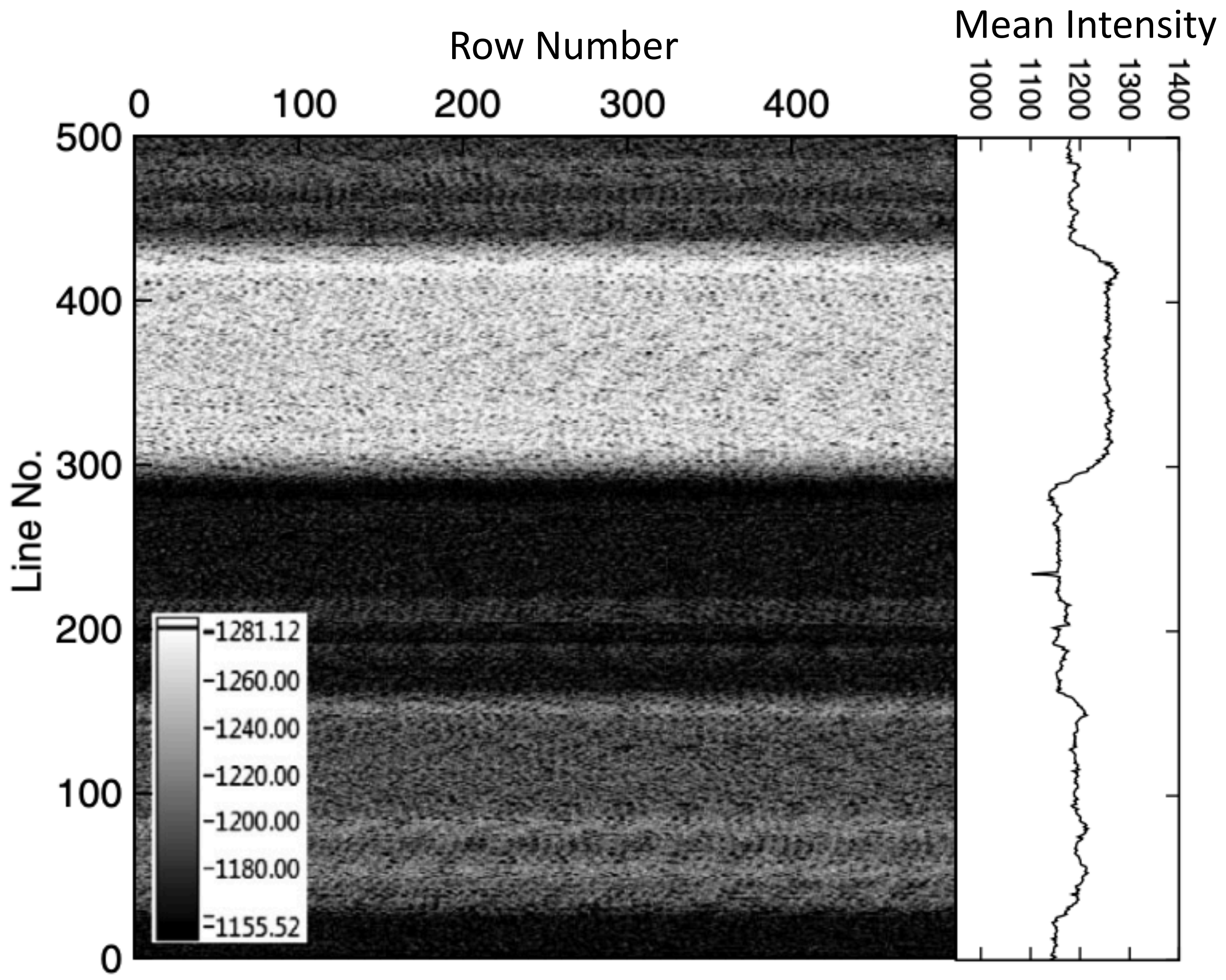


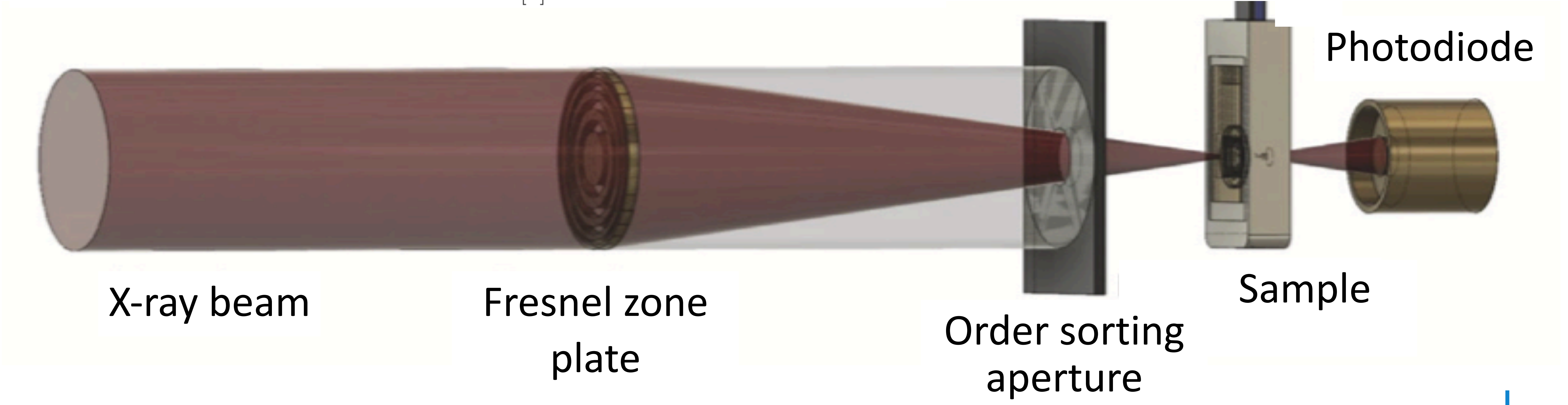
Scanning Transmission X-ray Microscopy at the Advanced Light Source

- STXM Beamlines:
  - Widely used for nanoscale studies
  - Fast raster scanning
  - No averaging
  - ≈ 1 ms/pixel, 1 s/line, 6 min/scan
- Current ID Feed Forward:
  - LOCO at various gaps to obtain (local) skew quadrupole corrector settings
  - Interpolation table used in FF correction
  - Takes one night to measure
  - Aging of tables results in large variation of vertical beam size
- STXM Intensity Fluctuations:
  - Residual ID induced vertical beamsize variations biggest contributor
  - POC: Leemann et al., PRL 123,194801





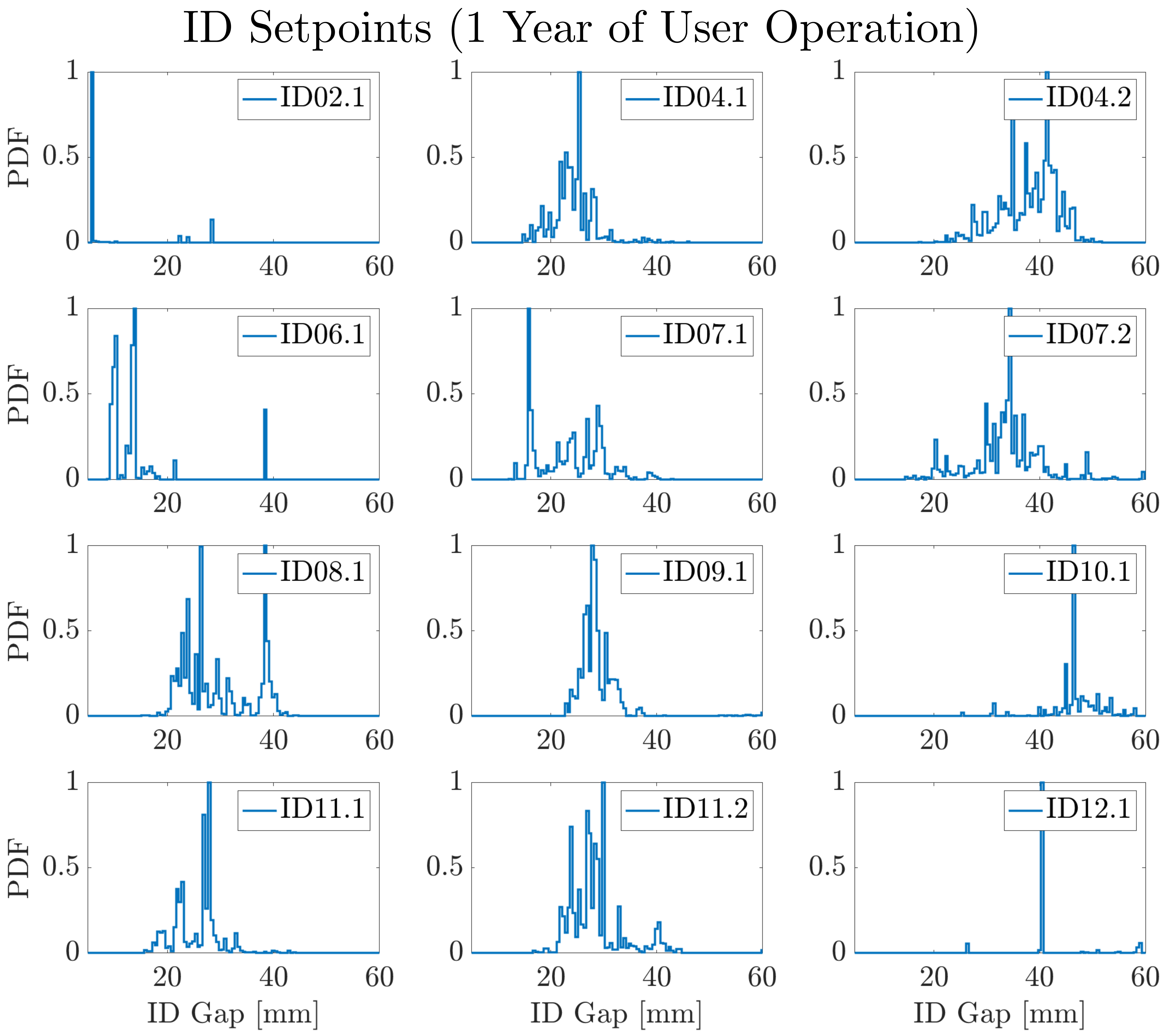




Model Development

Acquiring Training Data

- Data Sampling:
  - Derived from two years of user operation data
  - Ensures representative operational conditions

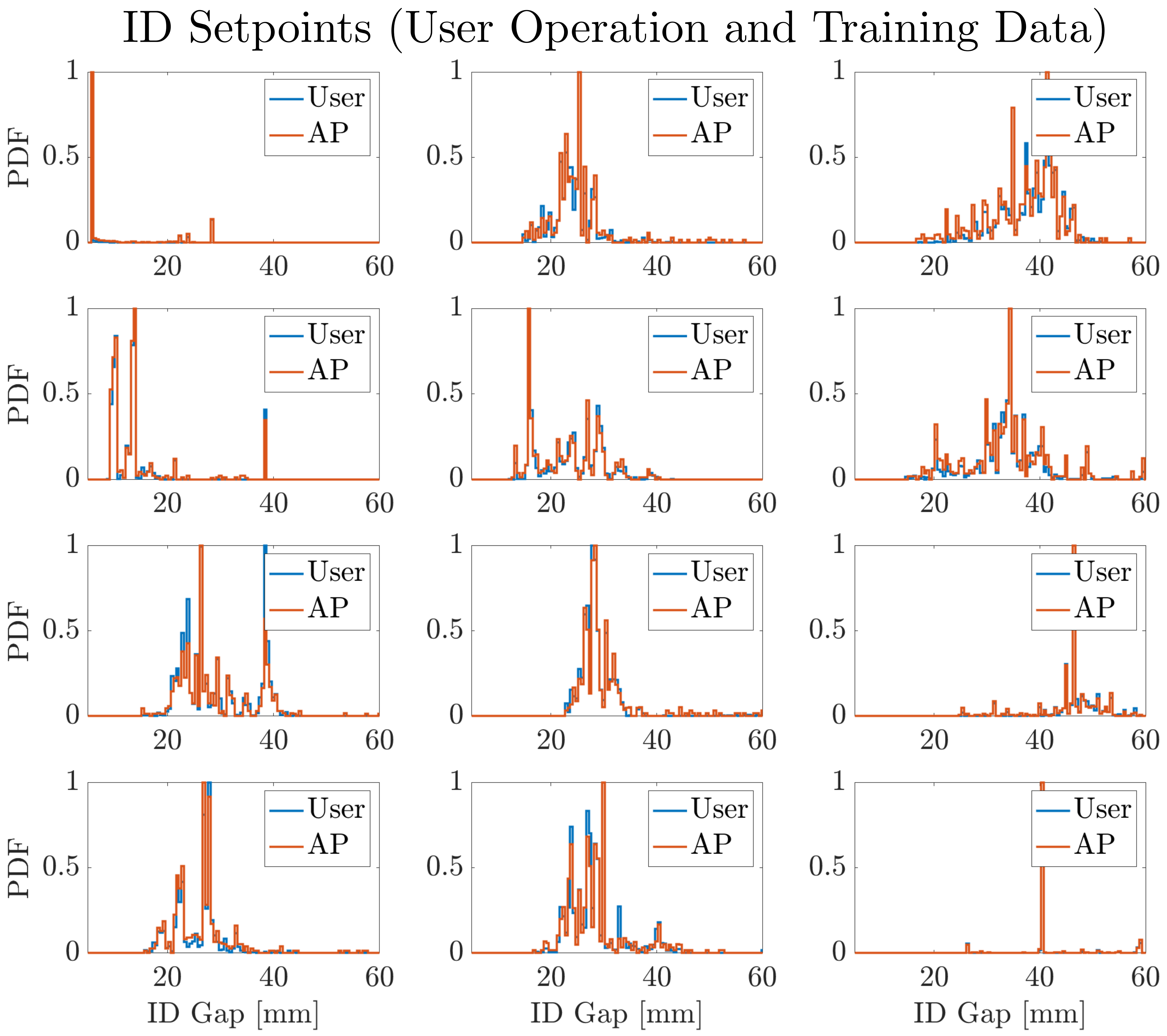


Acquiring Training Data

- Data Sampling:
  - Derived from two years of user operation data
  - Ensures representative operational conditions
- Data Acquisition and Recording:
  - Gathered during accelerator physics shifts
  - Independent exercise of each insertion device
  - All ID read-backs and beam size recorded at 10Hz
  - EPICS based archiver system
  - 12-hour, 27 ID parameters (466k x 27 samples)
- Operational Challenges:
  - High value of AP time leads to nighttime shifts
  - ID setup not optimized for fast ramping (ID amplifier trips, local ID FF trips)
  - Implementation of watchdog with for operational oversight very important

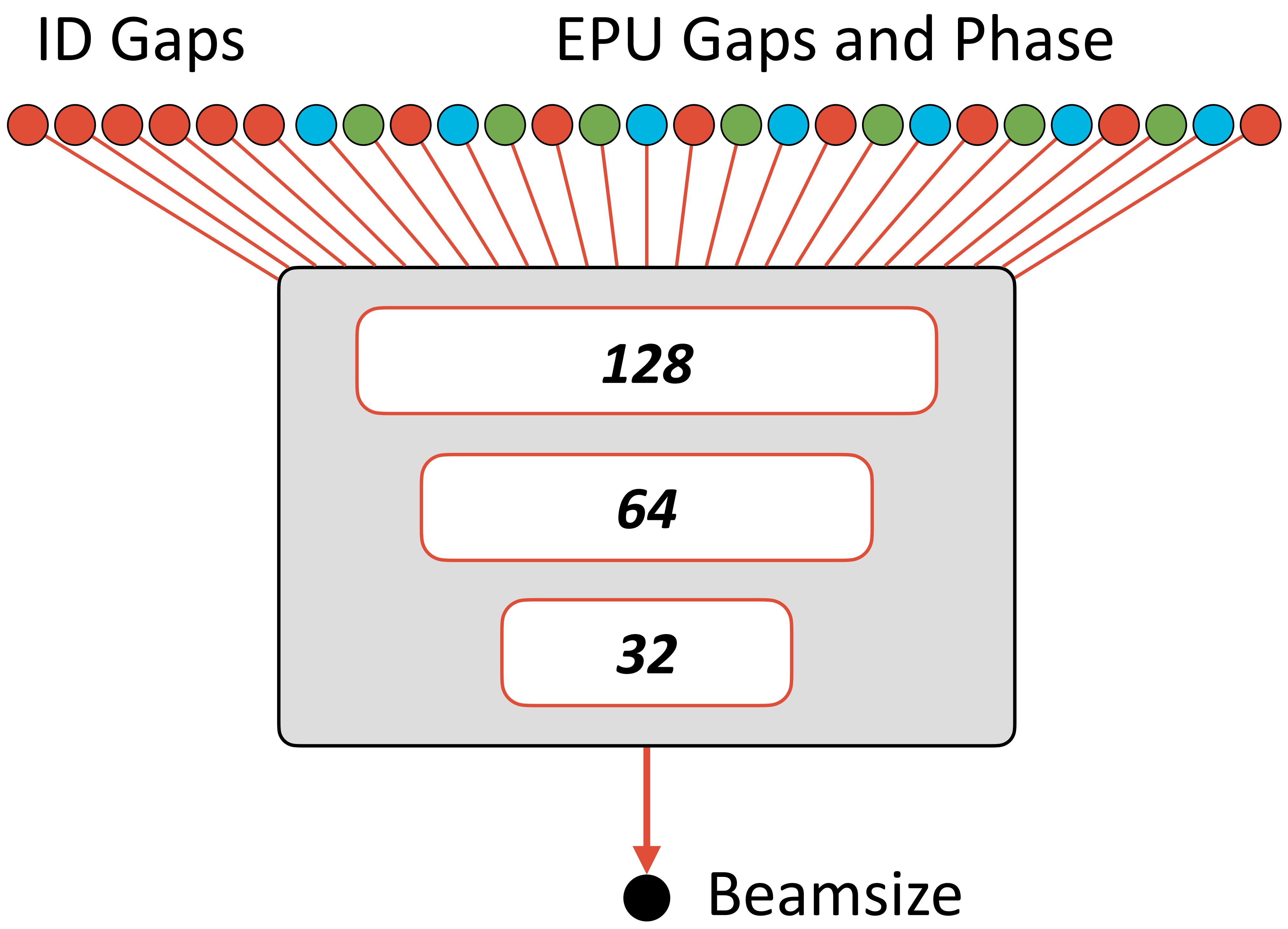
Acquiring Training Data

- Data Sampling:
  - Derived from two years of user operation data
  - Ensures representative operational conditions
- Data Acquisition and Recording:
  - Gathered during accelerator physics shifts
  - Independent exercise of each insertion device
  - All ID read-backs and beam size recorded at 10Hz
  - EPICS based archiver system
  - 12-hour, 27 ID parameters (466k x 27 samples)
- Operational Challenges:
  - High value of AP time leads to nighttime shifts
  - ID setup not optimized for fast ramping (ID amplifier trips, local ID FF trips)
  - Implementation of watchdog with for operational oversight very important



Neural Network Architecture

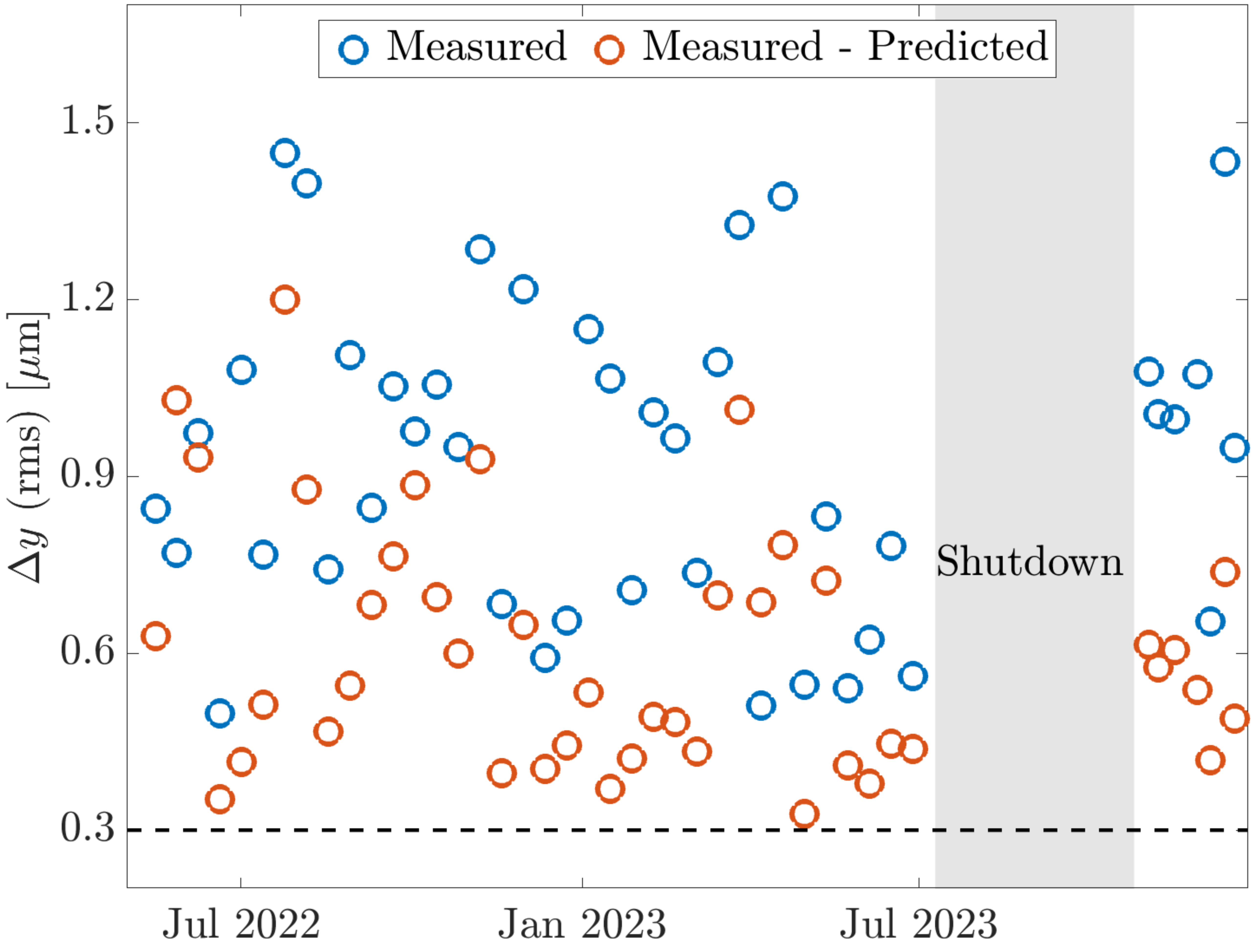
- Model Input/Output:
  - 27 ID input parameters
  - 1 beam size prediction output
  - Dispersion wave used to correct beamsize
- Studied Neural Network Types:
  - RNN, CNN, ***MLP***
- MLP Hyperparameter Search:
  - Number of hidden layers: 3
  - Neurons per Layer: 128/64/32
  - Activation Function: Tanh
- Final Hyperparameter Search:
  - Weight decay: 1E-3
  - Dropout rate: 0.2
- Takes about 15min on RTX2060 GPU



Hyperparameter	Search Space
Number of Hidden Layers	$\{1, 2, 3\}$
Number of Neurons per Layer	$\{2^n\}, 1 \leq n \leq 9$
Activation Function	{ReLU, Tanh, Sigmoid}
Weight decay	$\{10^{-n}\}, 1 \leq n \leq 5$
Dropout rate	$\{0.2, 0.4, 0.6, 0.8\}$

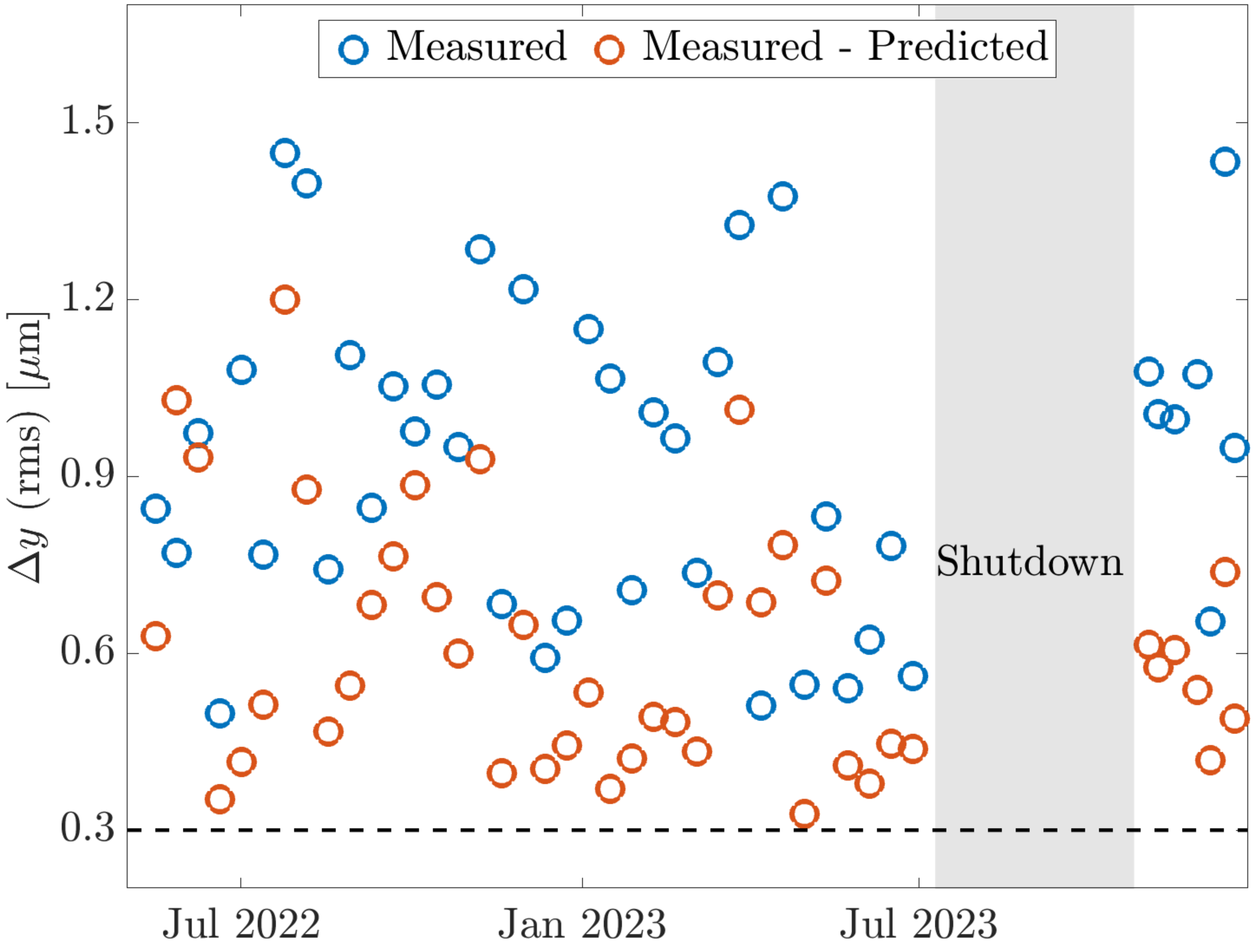
Evaluation on Historical User Operation Data

- Archive of Operational Data
  - 18 months of user ops data available
  - Subject to asynchronous downsampling before shutdown
  - Training on old data not possible
- Observations:
  - Prediction accuracy varies significantly between weeks
  - No significant long term drift apparent
  - Average performance of 0.6µm (noise floor: 0.3µm)



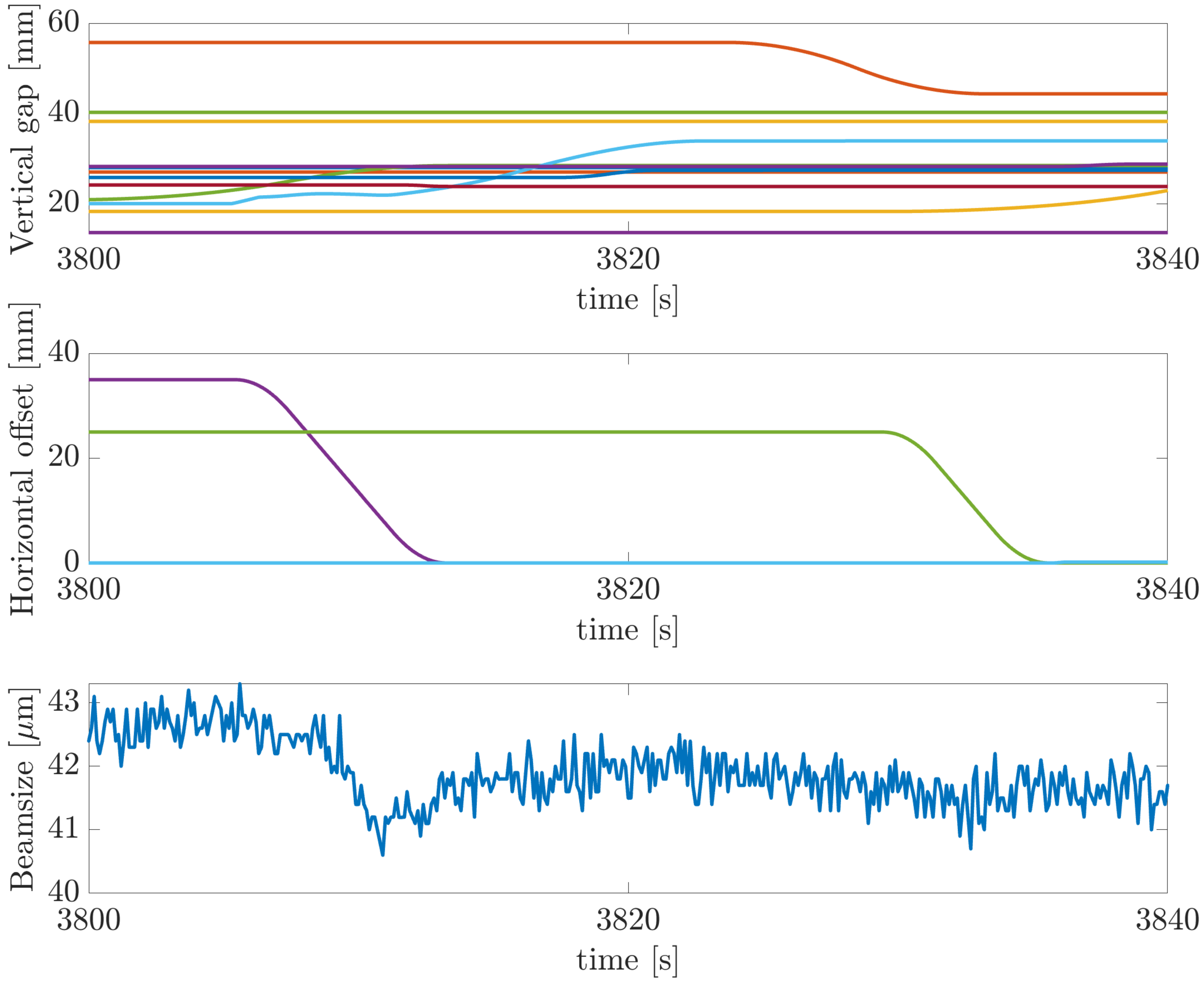
Impact of Training Data Size on Model Performance

- How much training time is required for *perfect* model?



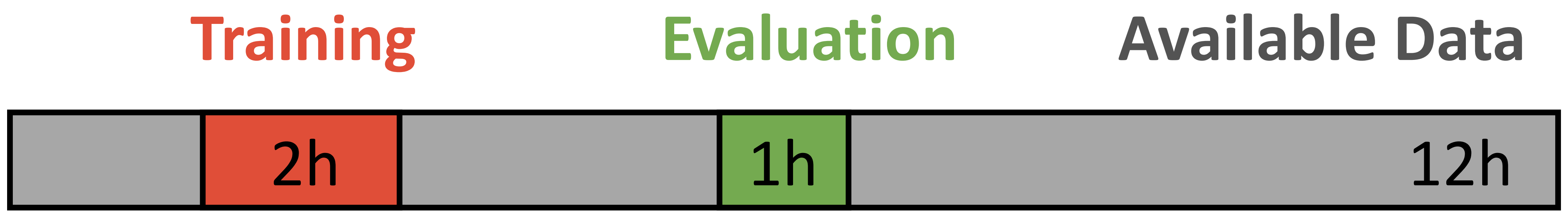
Impact of Training Data Size on Model Performance

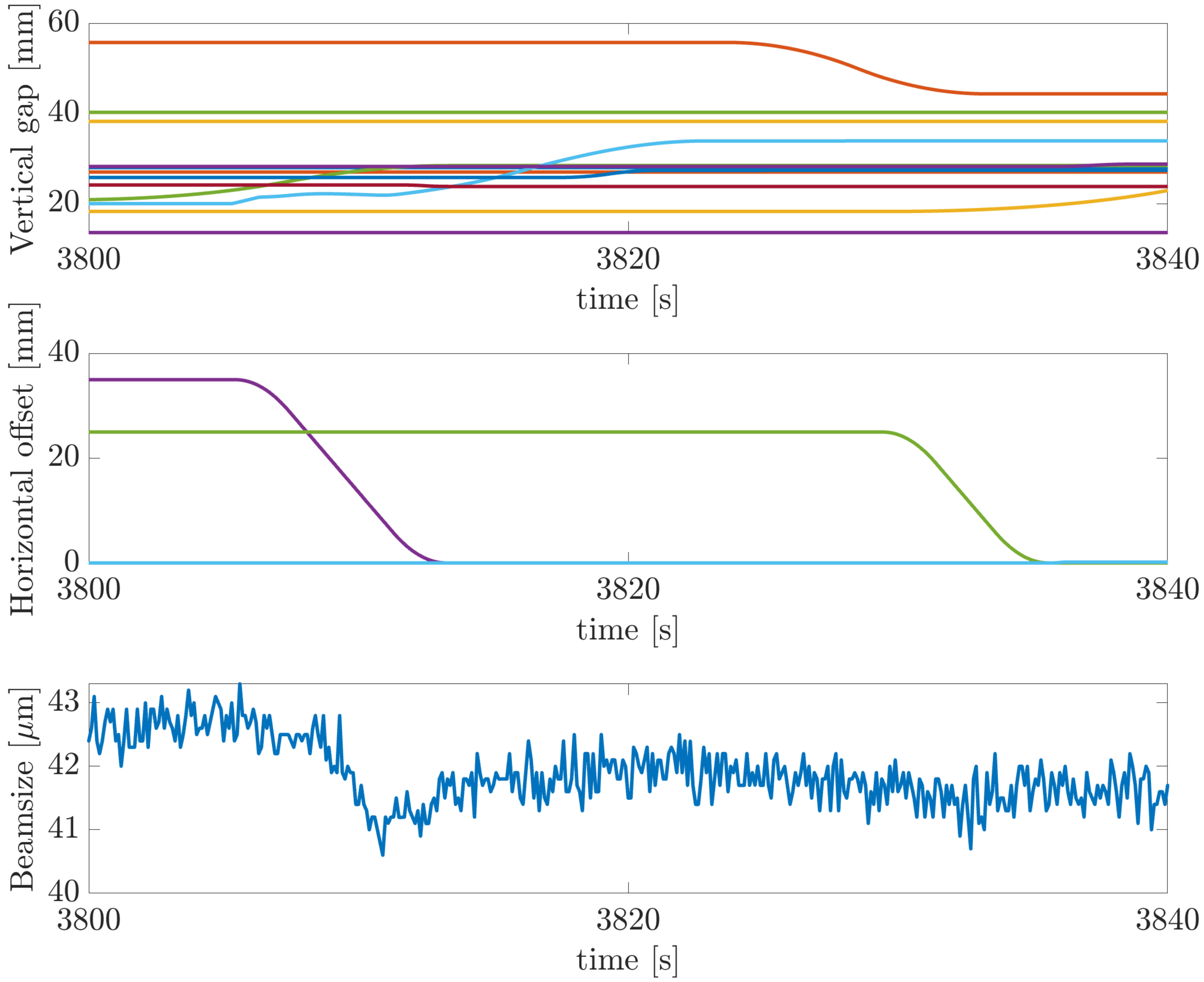
- How much training time is required for *perfect* model?
- Chronologically Split Data:
  - Can not randomly select datapoints for evaluation (oversampling at 10Hz)



Impact of Training Data Size on Model Performance

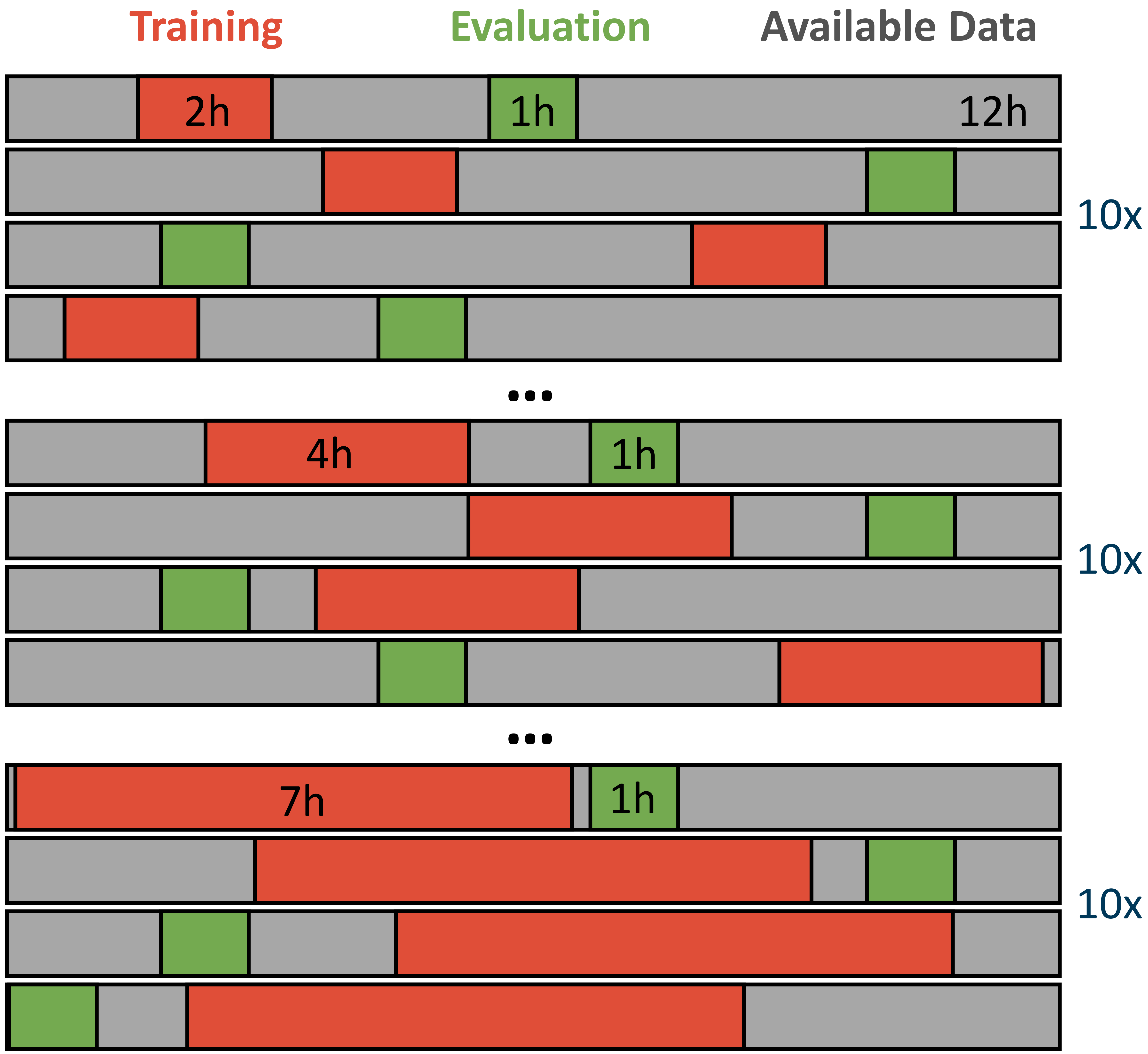
- How much training time is required for *perfect* model?
- Chronologically Split Data:
  - Can not randomly select datapoints for evaluation (oversampling at 10Hz)
- Evaluation Procedure:
  - Remove 1h randomly from the data set for evaluation
  - Choose [1,…11]h for training
  - 10 seed for each configuration
  - Evaluate RMSE on evaluation data





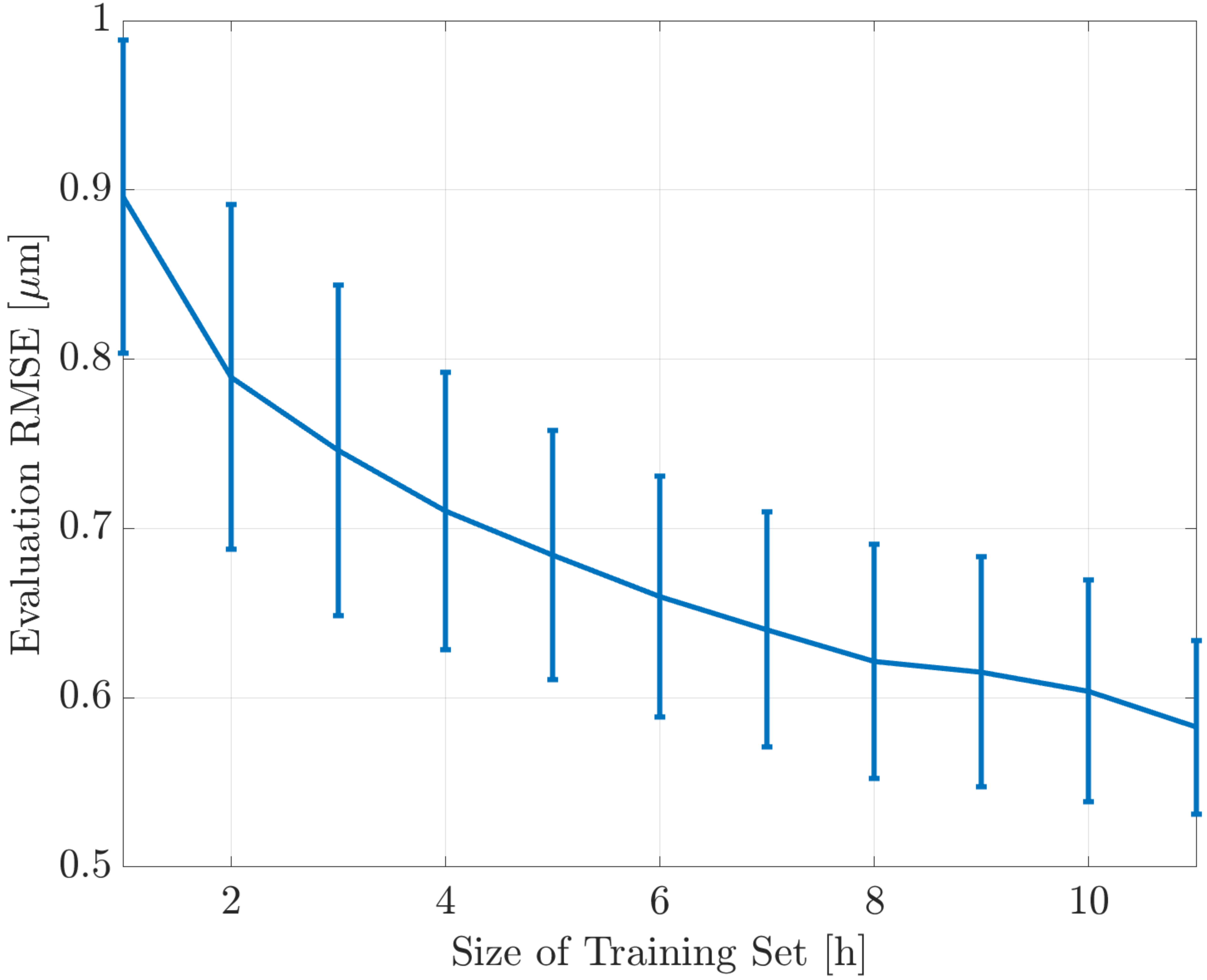
Impact of Training Data Size on Model Performance

- How much training time is required for *perfect* model?
- Chronologically Split Data:
  - Can not randomly select datapoints for evaluation (oversampling at 10Hz)
- Evaluation Procedure:
  - Remove 1h randomly from the data set for evaluation
  - Choose [1,…11]h for training
  - 10 seed for each configuration
  - Evaluate RMSE on evaluation data



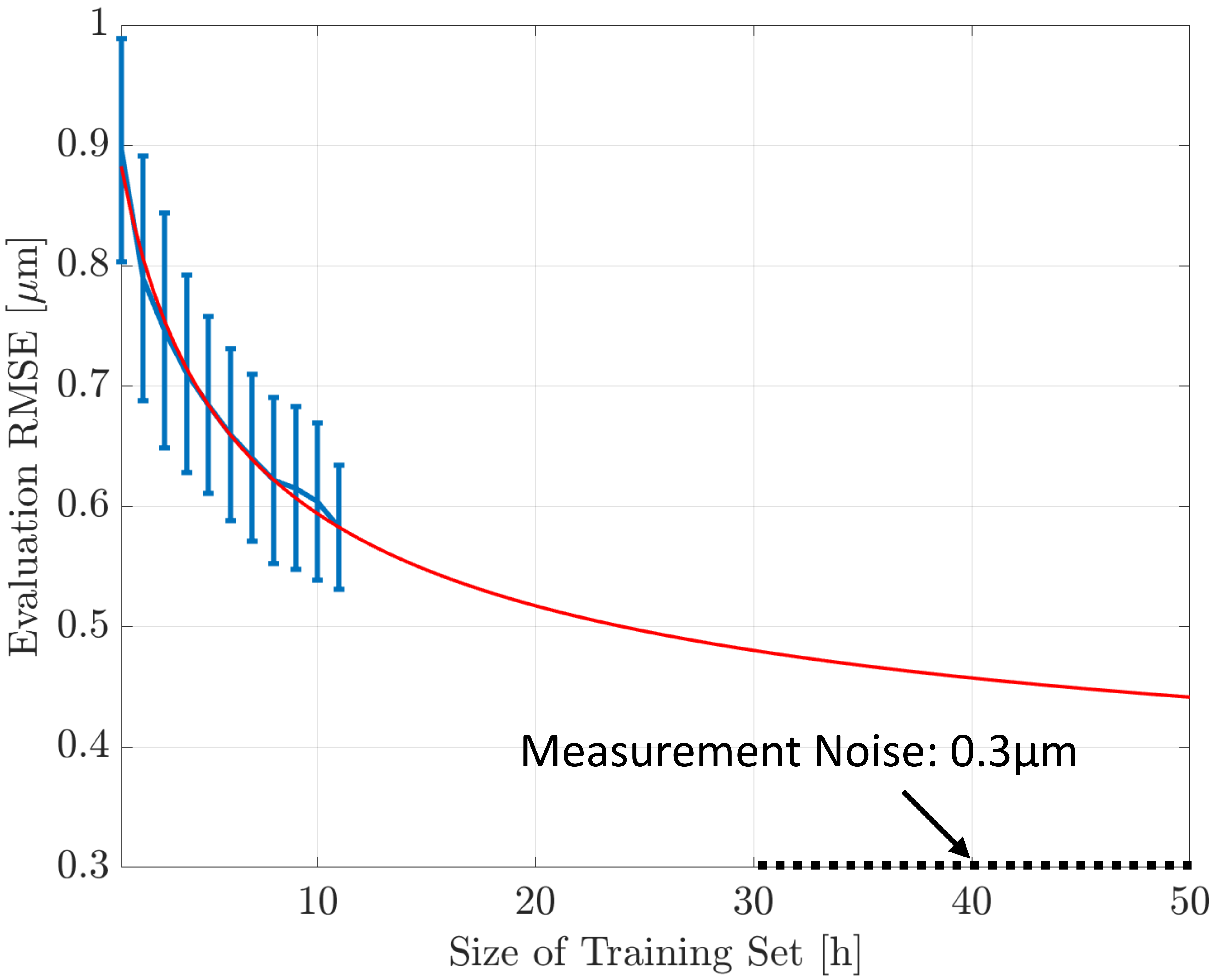
Impact of Training Data Size on Model Performance

- How much training time is required for *perfect* model?
- Chronologically Split Data:
  - Can not randomly select datapoints for evaluation (oversampling at 10Hz)
- Evaluation Procedure:
  - Remove 1h randomly from the data set for evaluation
  - Choose [1,…11]h for training
  - 10 seed for each configuration
  - Evaluate RMSE on evaluation data
- Observed Convergence:
  - Reasonable convergence at first
  - Trend suggests infeasible amount of data required to reach noise level

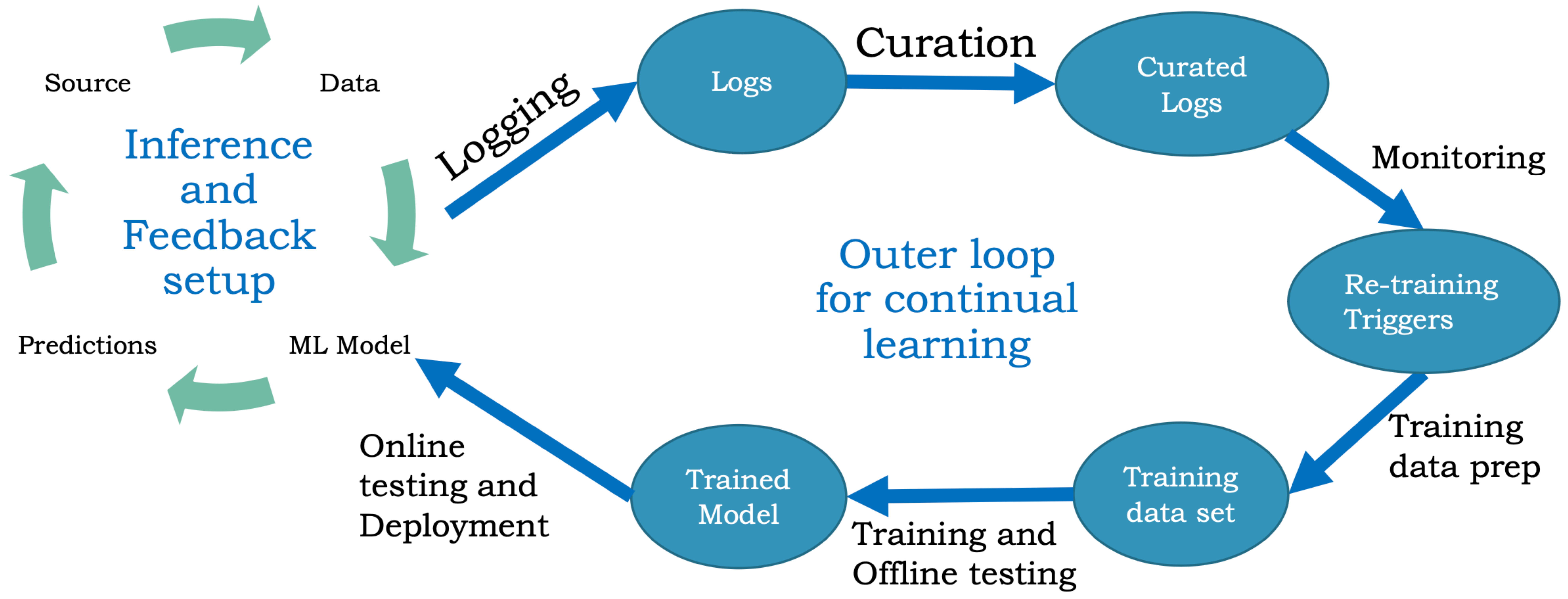


Impact of Training Data Size on Model Performance

- How much training time is required for *perfect* model?
- Chronologically Split Data:
  - Can not randomly select datapoints for evaluation (oversampling at 10Hz)
- Evaluation Procedure:
  - Remove 1h randomly from the data set for evaluation
  - Choose [1,…11]h for training
  - 10 seed for each configuration
  - Evaluate RMSE on evaluation data
- Observed Convergence:
  - Reasonable convergence at first
  - Trend suggests infeasible amount of data required to reach noise level

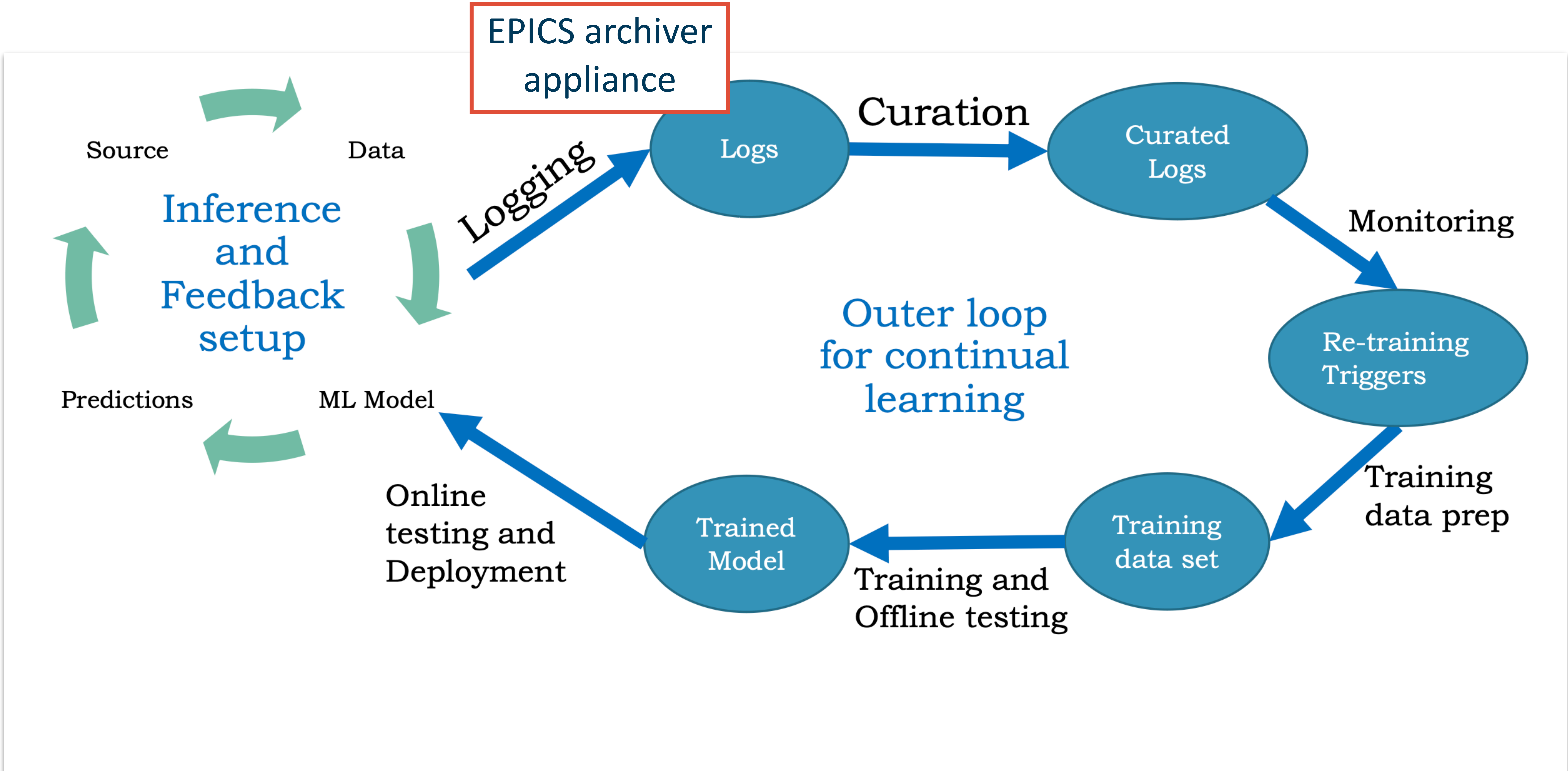


Continual Online Fine-Tuning



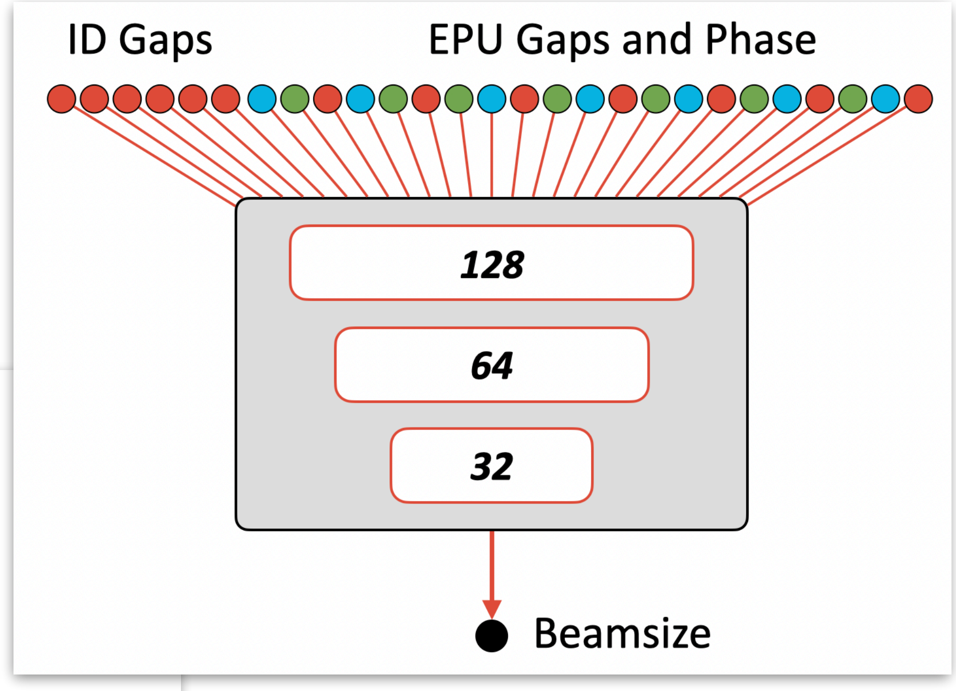
K. Rajput, "Model up-keep with continual learning", MaLAPA'24

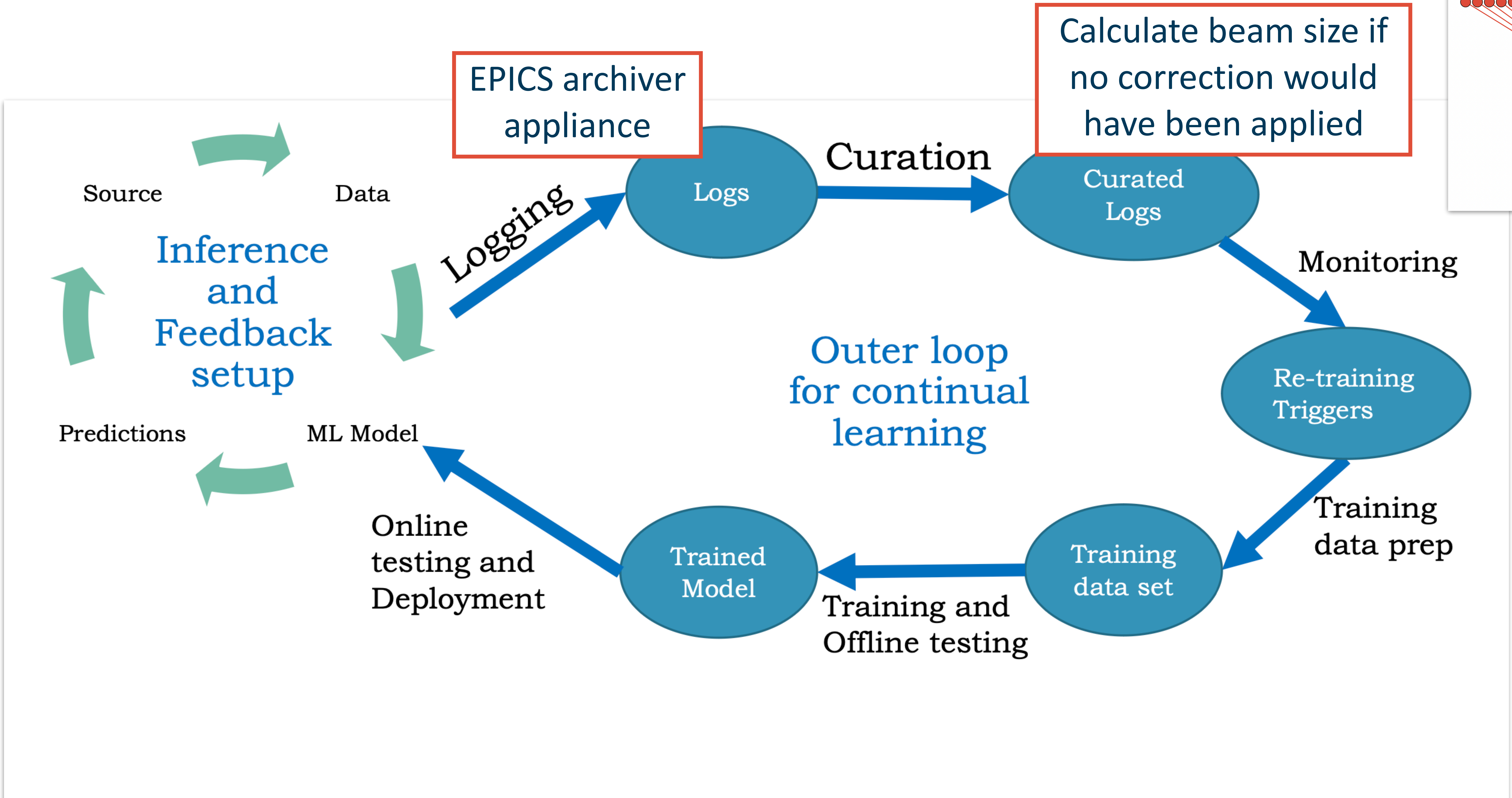
Continual Online Fine-Tuning



K. Rajput, "Model up-keep with continual learning", MaLAPA'24

Continual Online Fine-Tuning



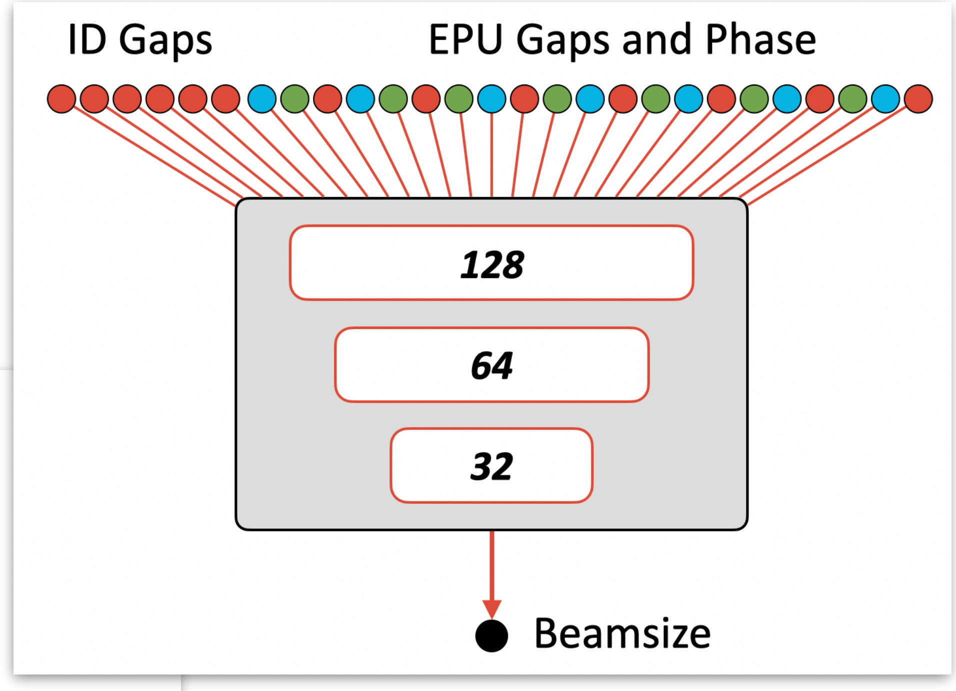


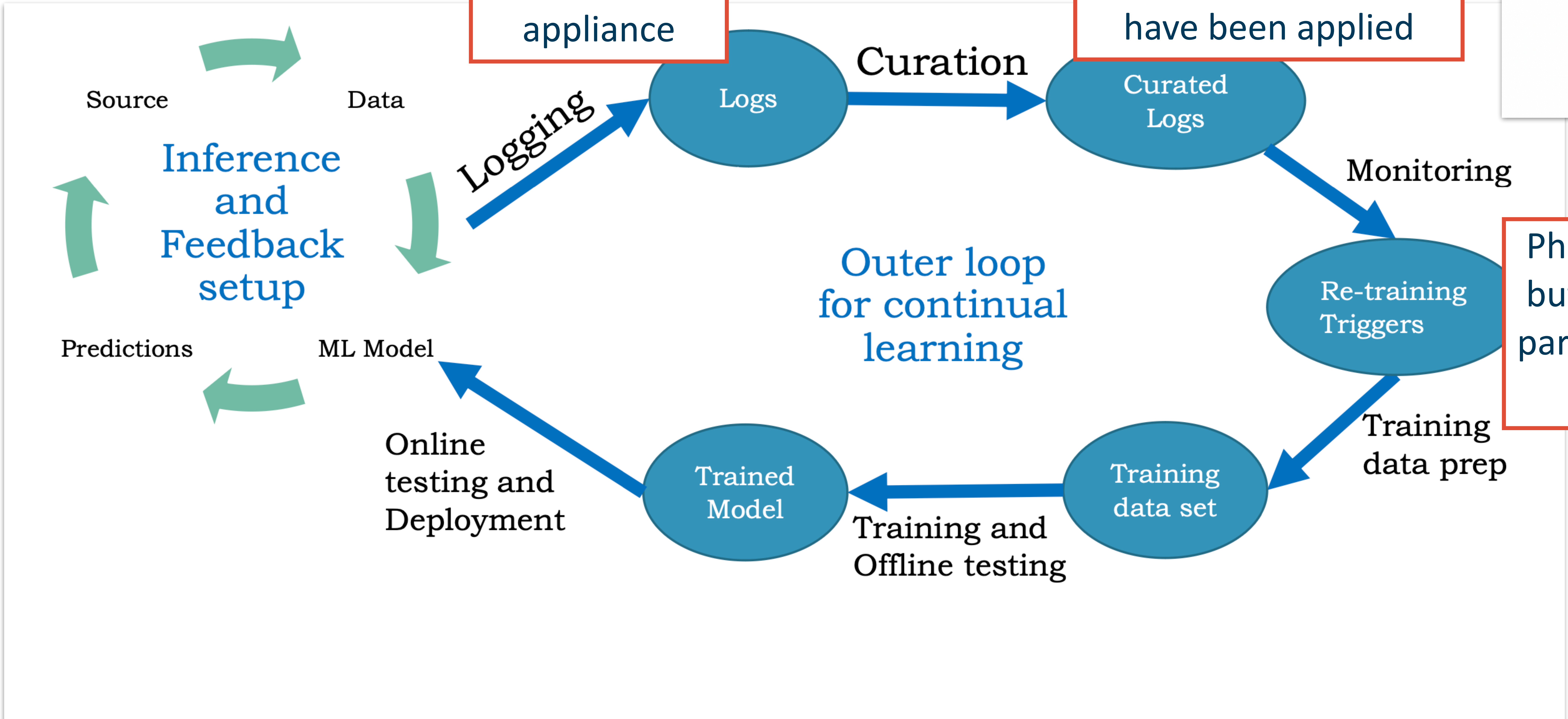
K. Rajput, "Model up-keep with continual learning", MaLAPA'24

Continual Online Fine-Tuning

EPICS archiver appliance

Calculate beam size if no correction would have been applied

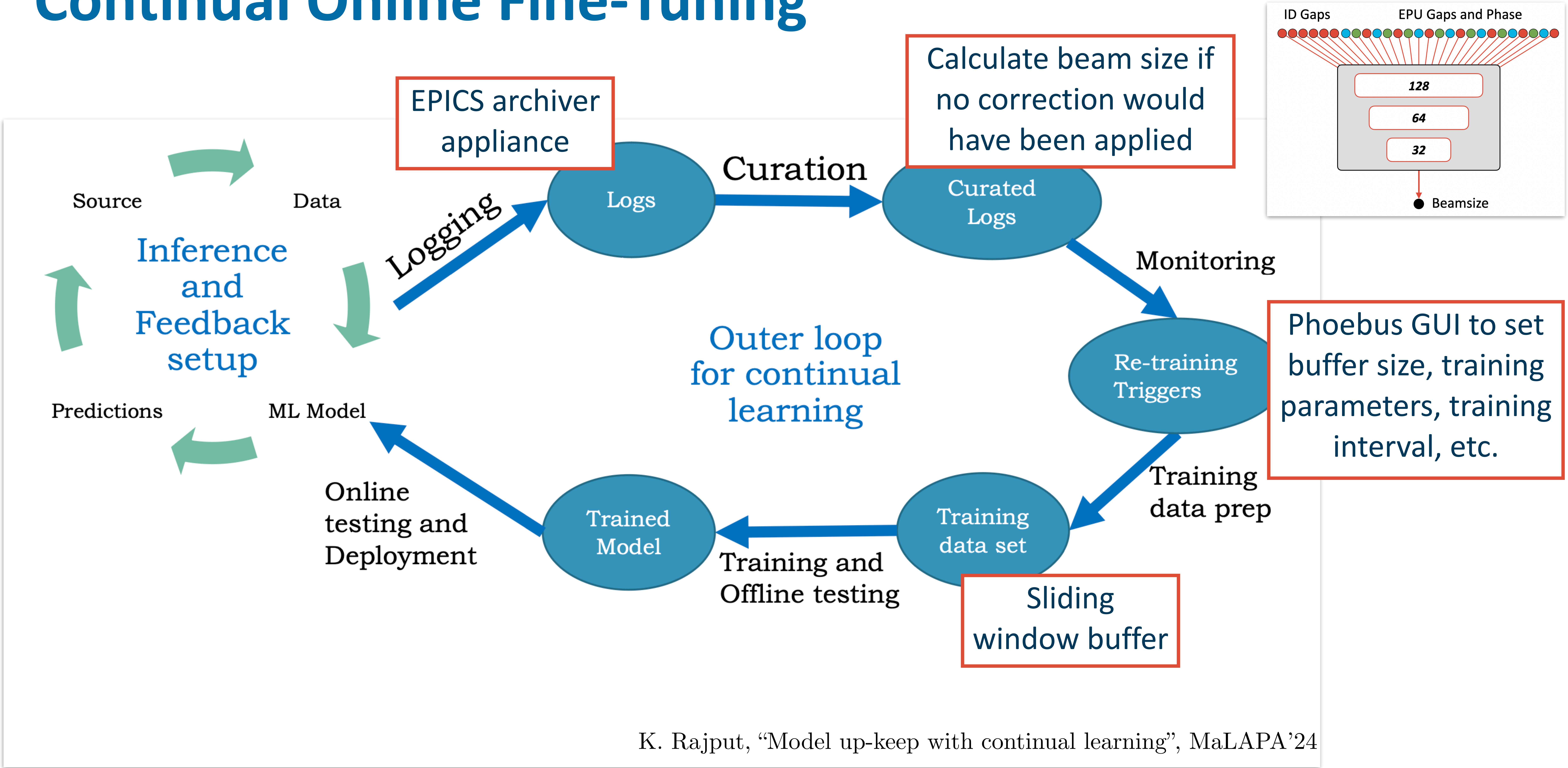




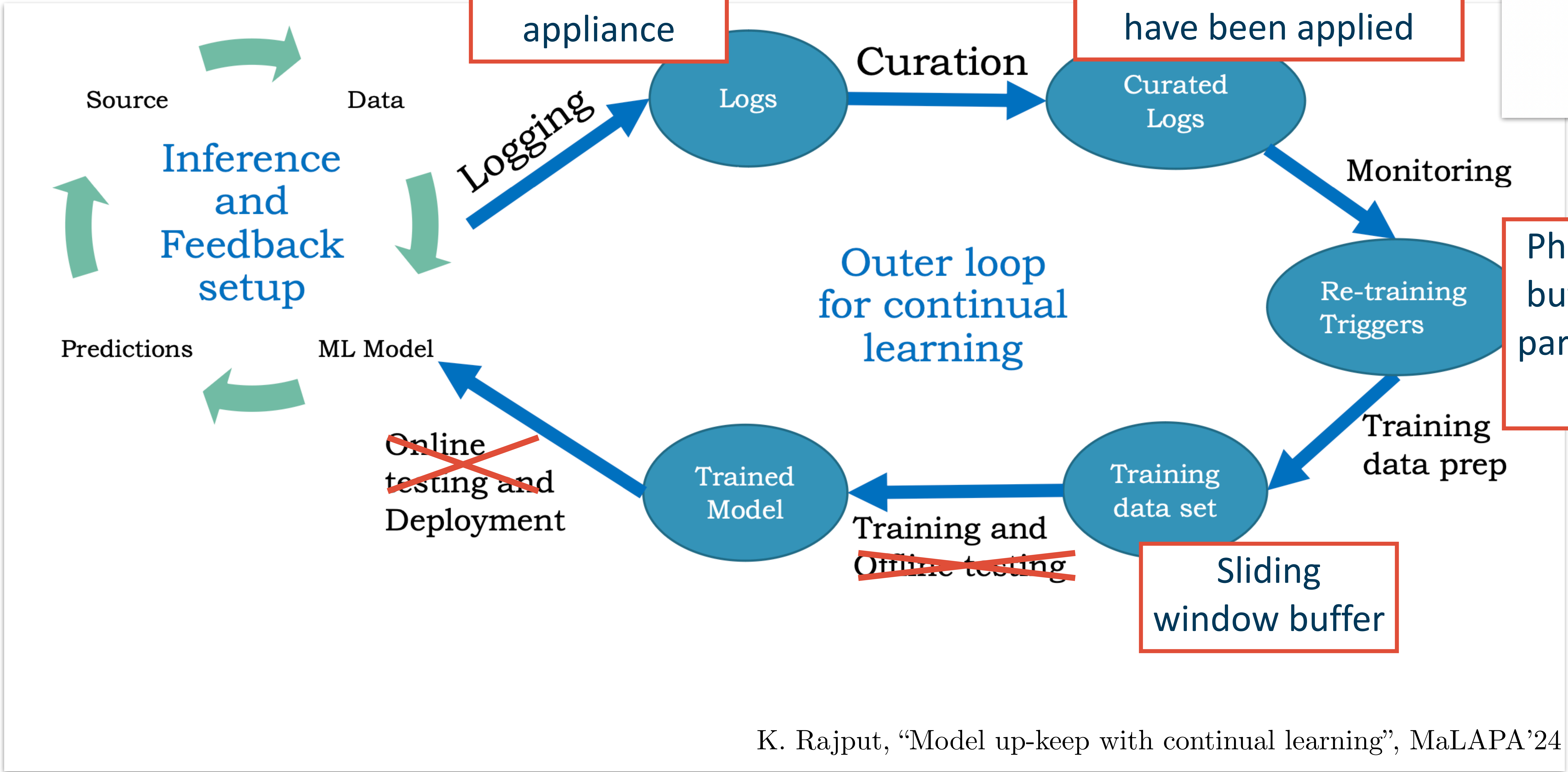
K. Rajput, "Model up-keep with continual learning", MaLAPA'24

Phoebus GUI to set buffer size, training parameters, training interval, etc.

Continual Online Fine-Tuning

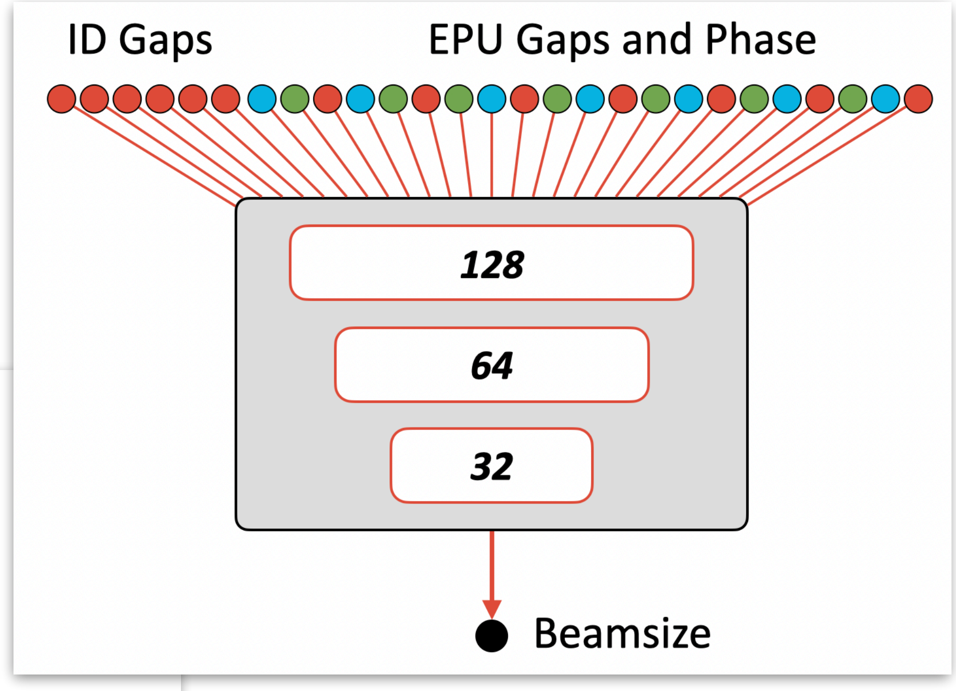


Continual Online Fine-Tuning



EPICS archiver appliance

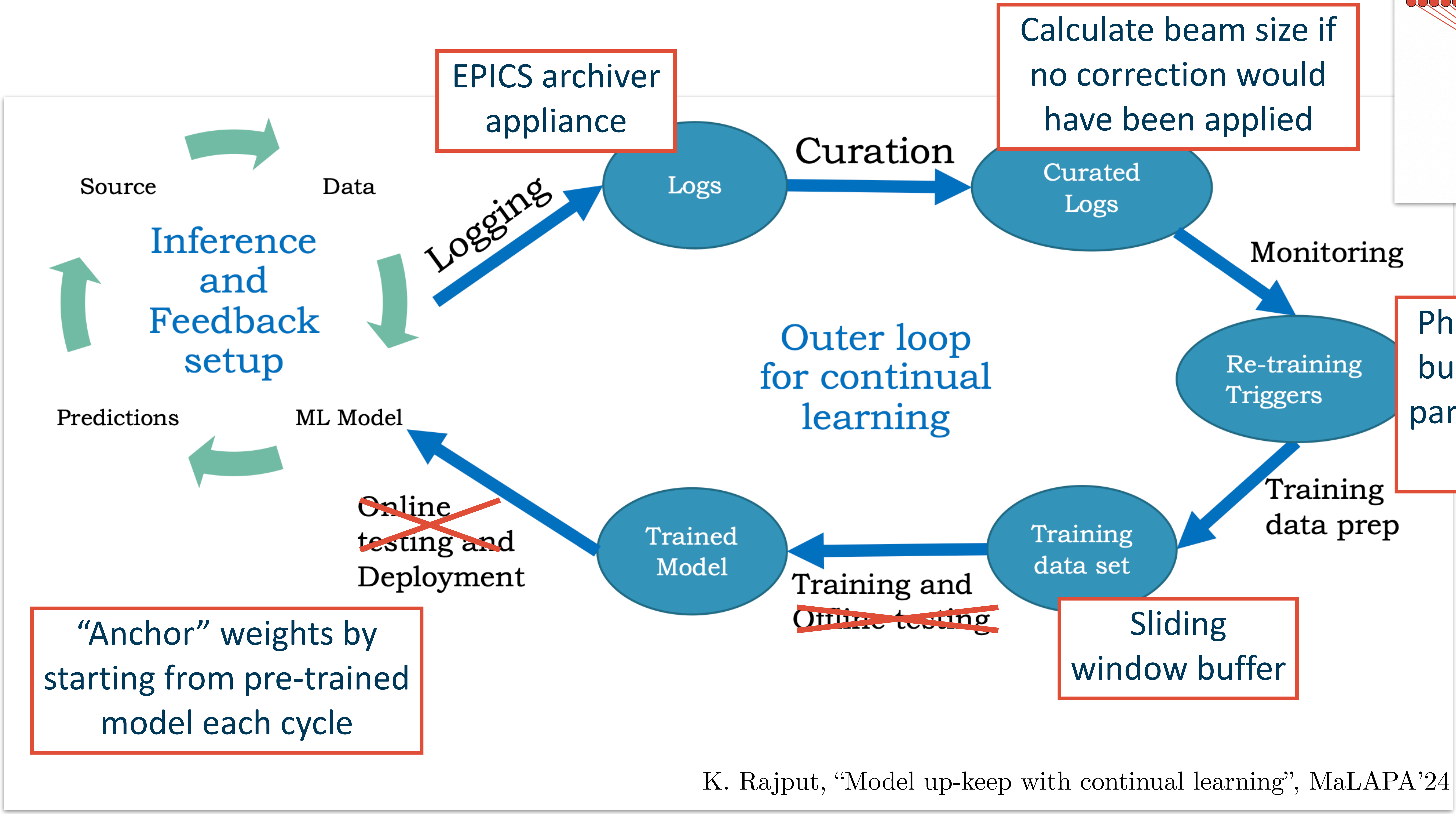
Calculate beam size if no correction would have been applied

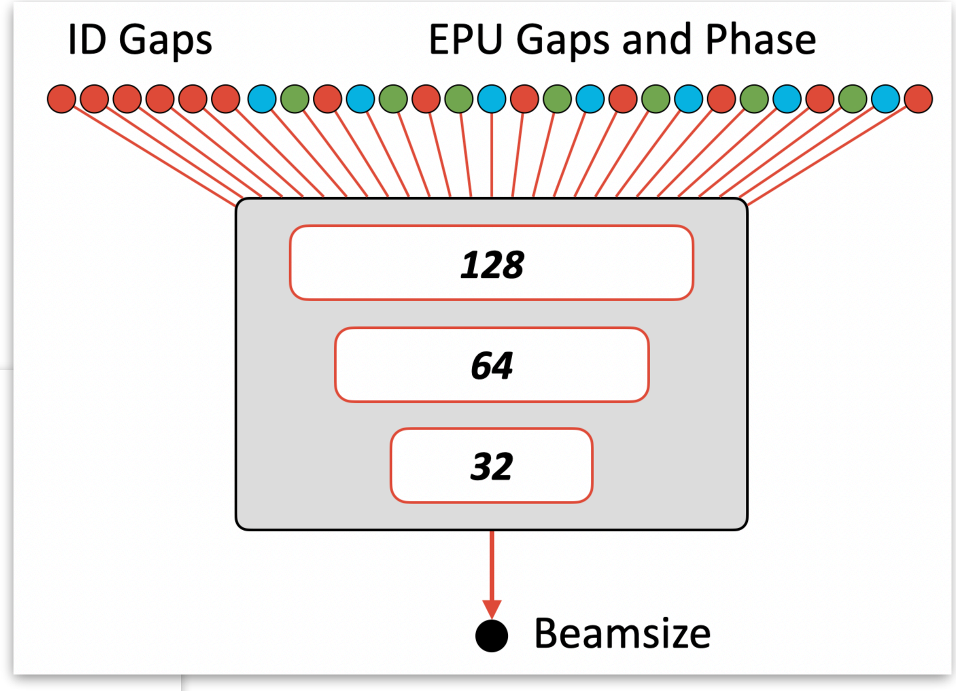


Phoebus GUI to set buffer size, training parameters, training interval, etc.

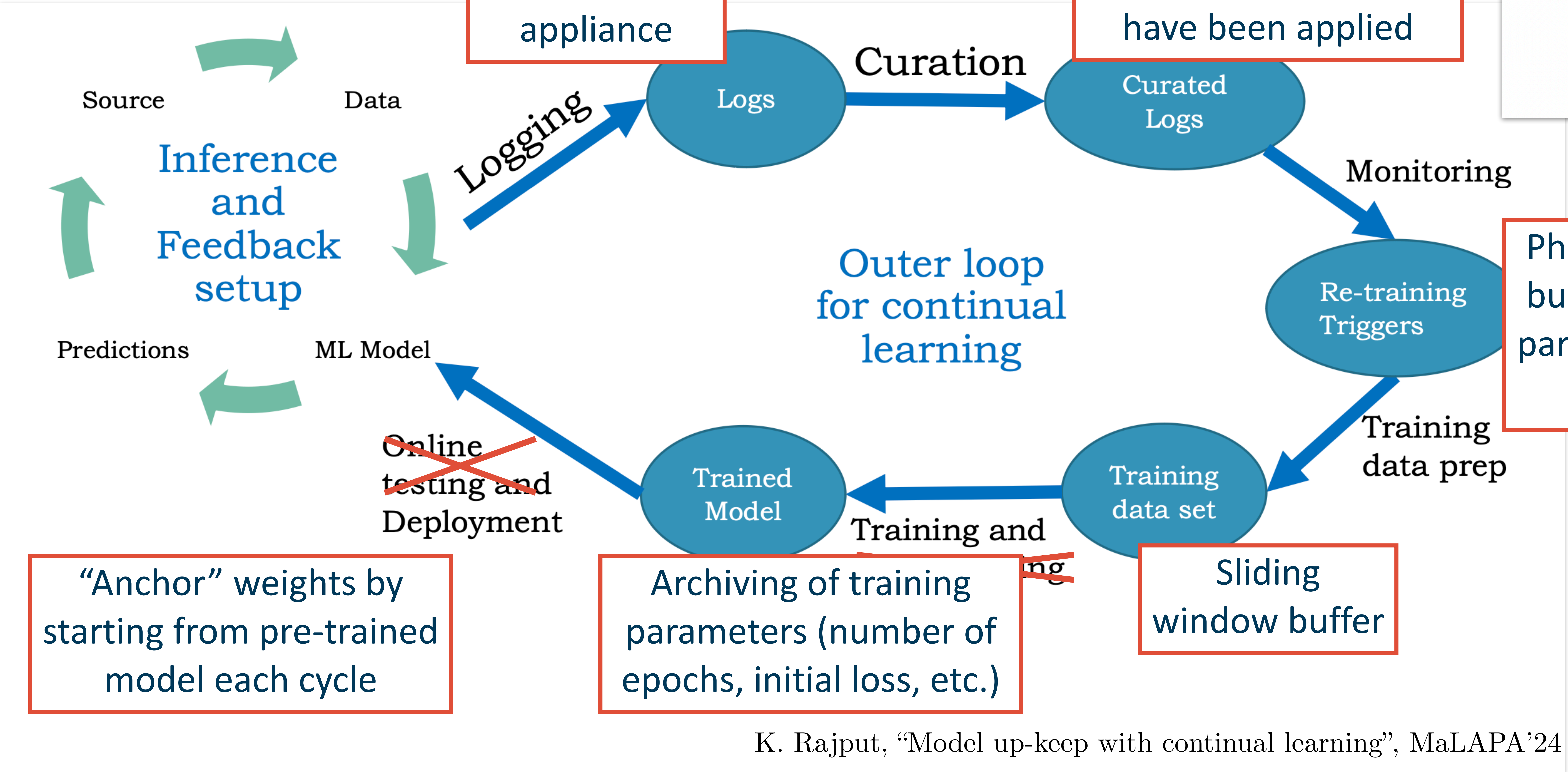
Sliding window buffer

Continual Online Fine-Tuning





Continual Online Fine-Tuning



EPICS archiver appliance

Calculate beam size if no correction would have been applied

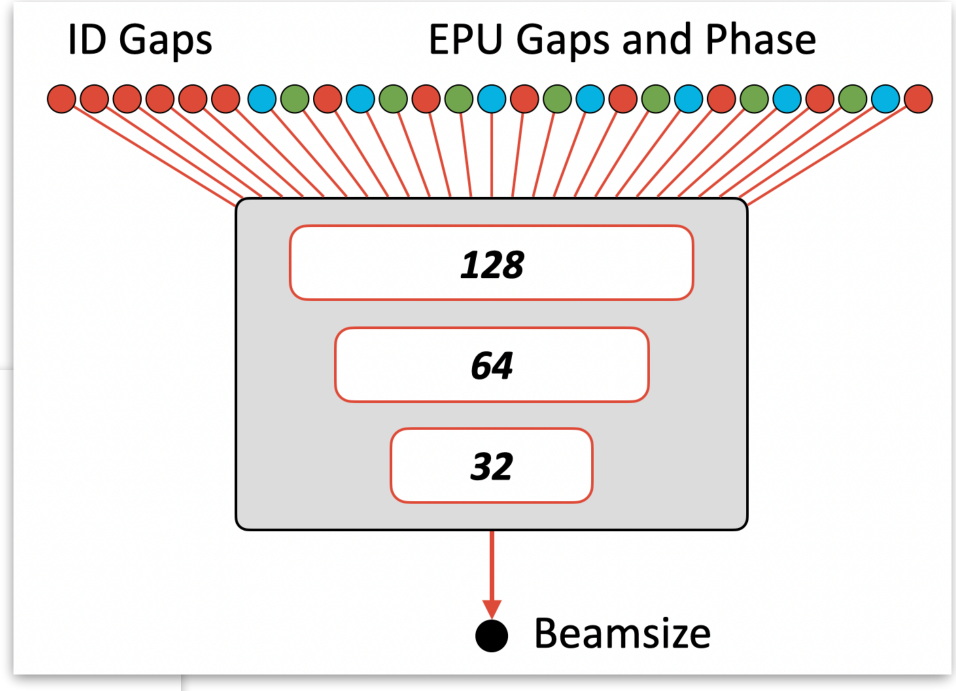
Phoebus GUI to set buffer size, training parameters, training interval, etc.

Sliding window buffer

Archiving of training parameters (number of epochs, initial loss, etc.)

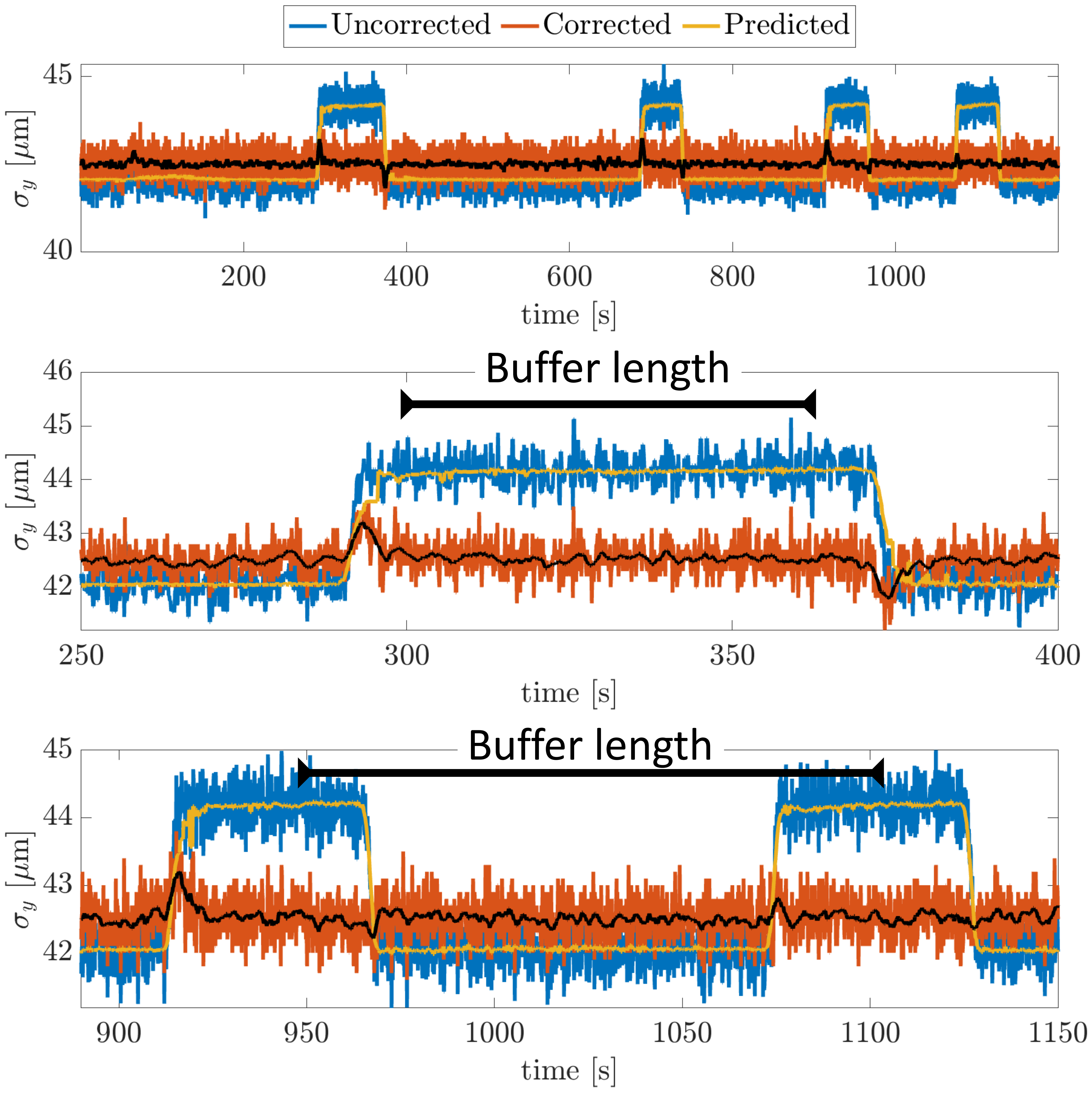
"Anchor" weights by starting from pre-trained model each cycle

K. Rajput, "Model up-keep with continual learning", MaLAPA'24



Continual Online Fine-Tuning

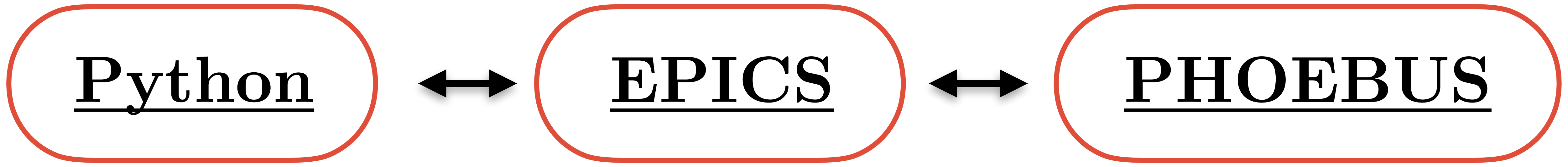
- Online Fine-Tuning:
  - Circular buffer to record model input
  - Train base model on data in buffer only
  - Start from base model each cycle to avoid runaway
  - Uncorrected beamsize calculated with DWP
- Parameters:
  - Typically 1k samples in buffer
  - Takes less then 100 epochs and about 1s
- Feedback vs Feedforward:
  - Online retraining acts as feedback
  - Buffer size controls impact of FB vs. FF

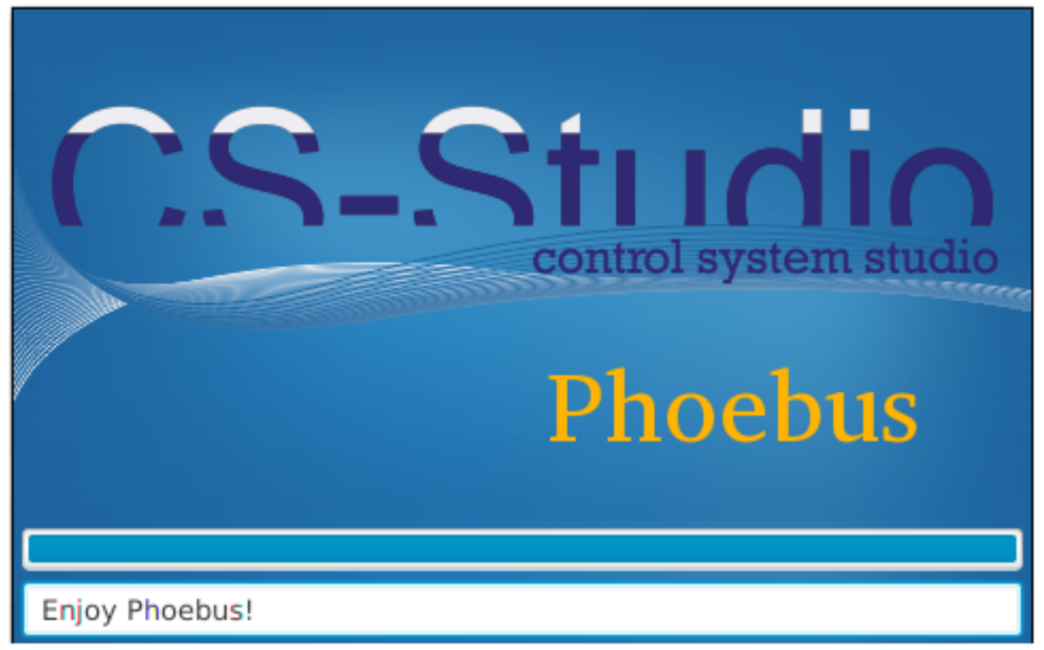


Model Deployment

Beam Size Control Backend Layout

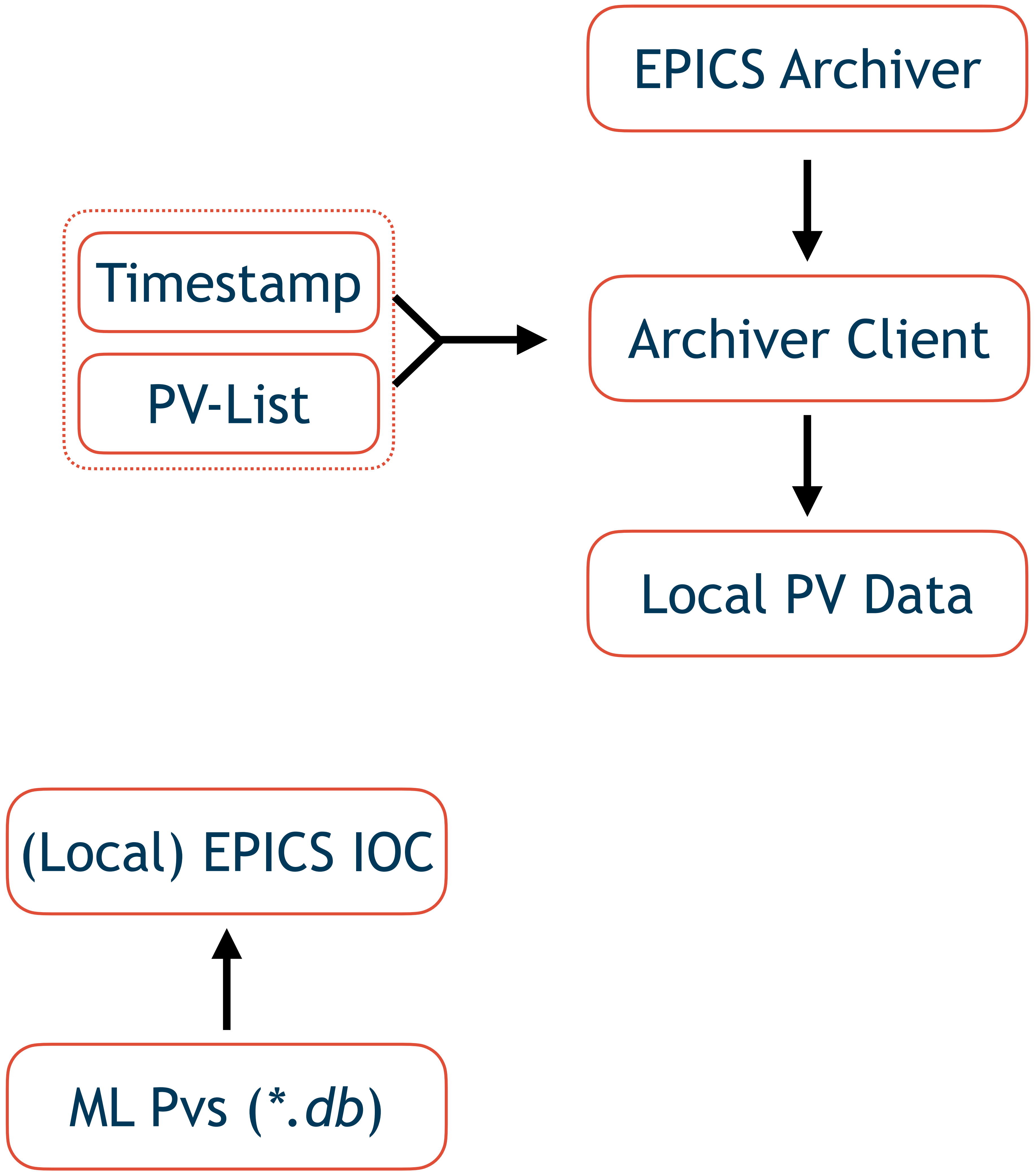
- Python Backend:
  - PyTorch
  - Currently on control room VM
  - Plan to implement on IOC this year
- Dedicated EPICS IOC:
  - 600 PVs required
  - Goal: concentrate all logic on EPICS
- PHOEBUS GUI:
  - State of the art control system GUI
  - Easy integration with EPICS
  - Expert/Operator Panel





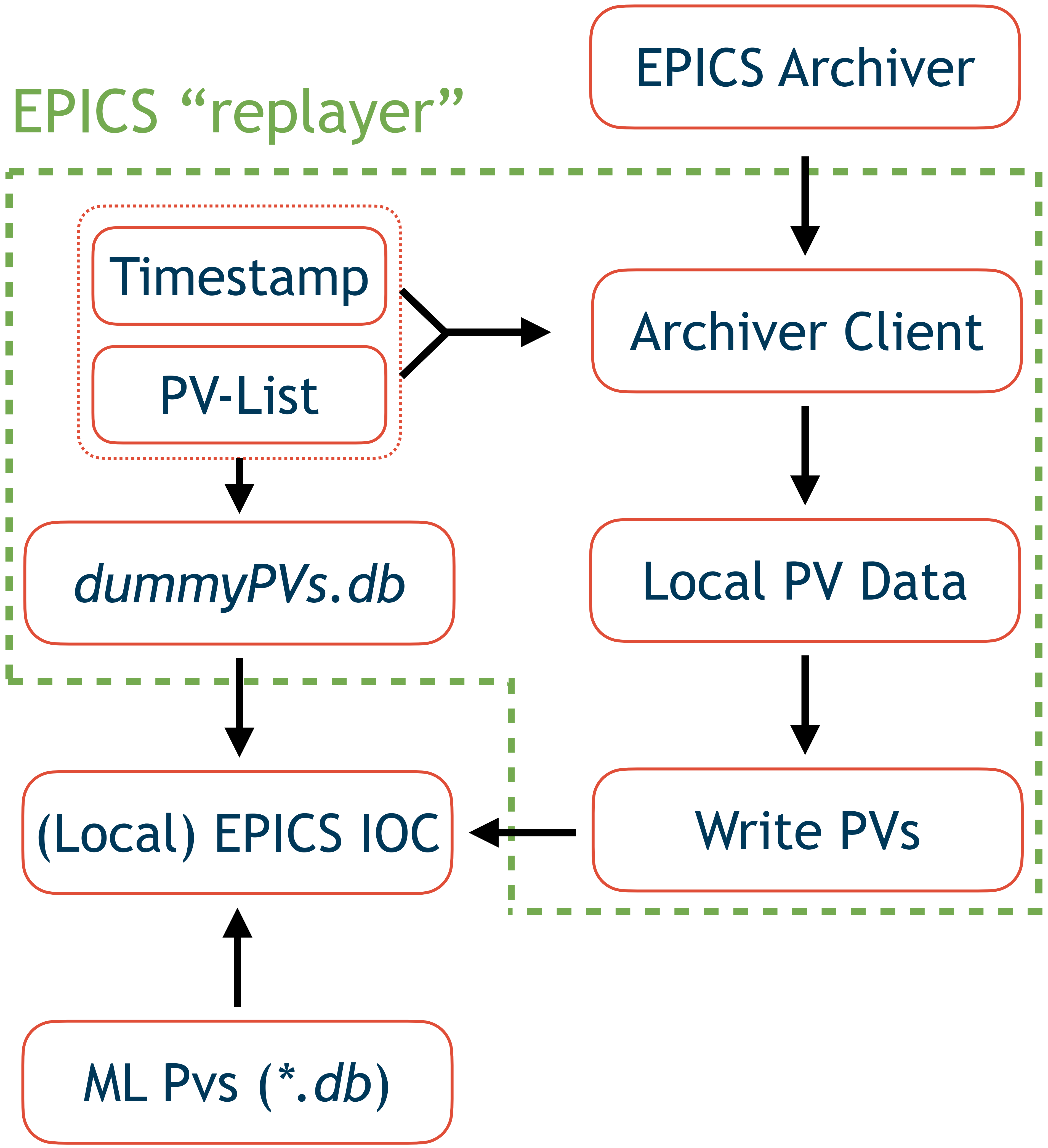
Archiver Replayer: Invaluable R&D Tool

- Local IOC Sandbox:
  - Beam size control IOC acting as a development sandbox
- Archived Data Utilization:
  - Archiver client to download historical data



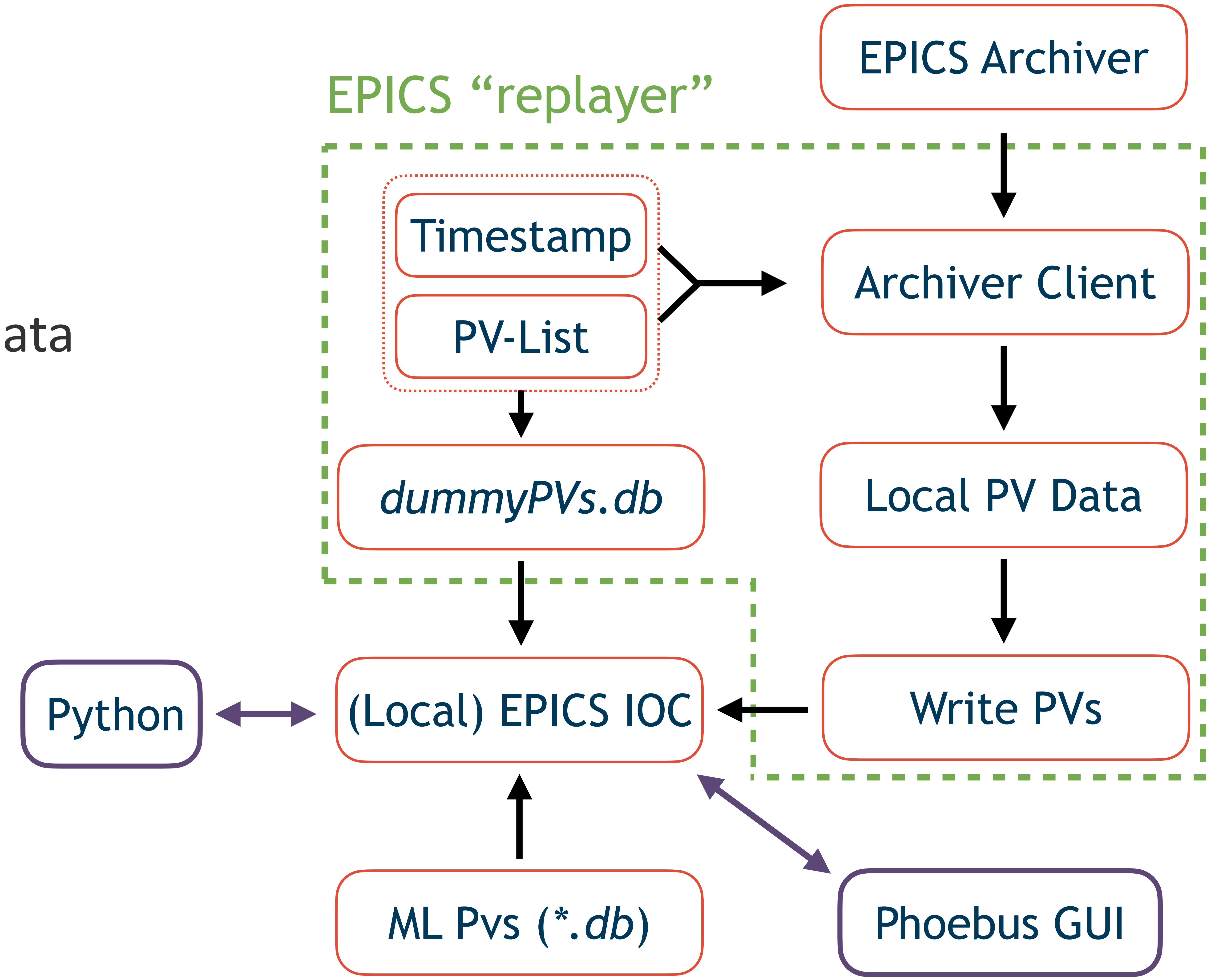
Archiver Replayer: Invaluable R&D Tool

- Local IOC Sandbox:
  - Beam size control IOC acting as a development sandbox
- Archived Data Utilization:
  - Archiver client to download historical data
  - dummyPVs.db automatically generated from archiver client
  - Python loop over PVs in **.csv* file



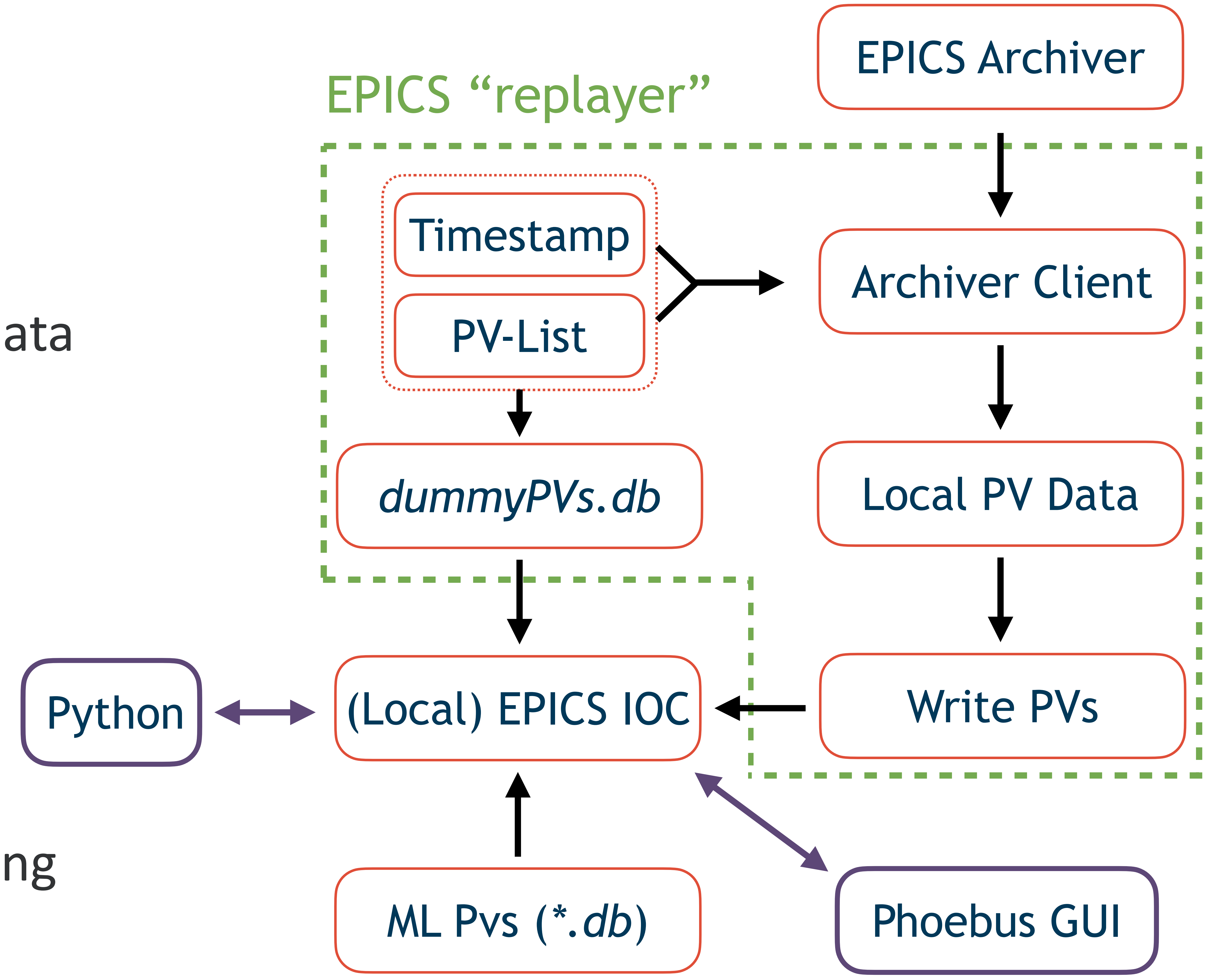
Archiver Replayer: Invaluable R&D Tool

- Local IOC Sandbox:
  - Beam size control IOC acting as a development sandbox
- Archived Data Utilization:
  - Archiver client to download historical data
  - dummyPVs.db automatically generated from archiver client
  - Python loop over PVs in **.csv* file
- Phoebus GUI Integration
  - Phoebus running locally

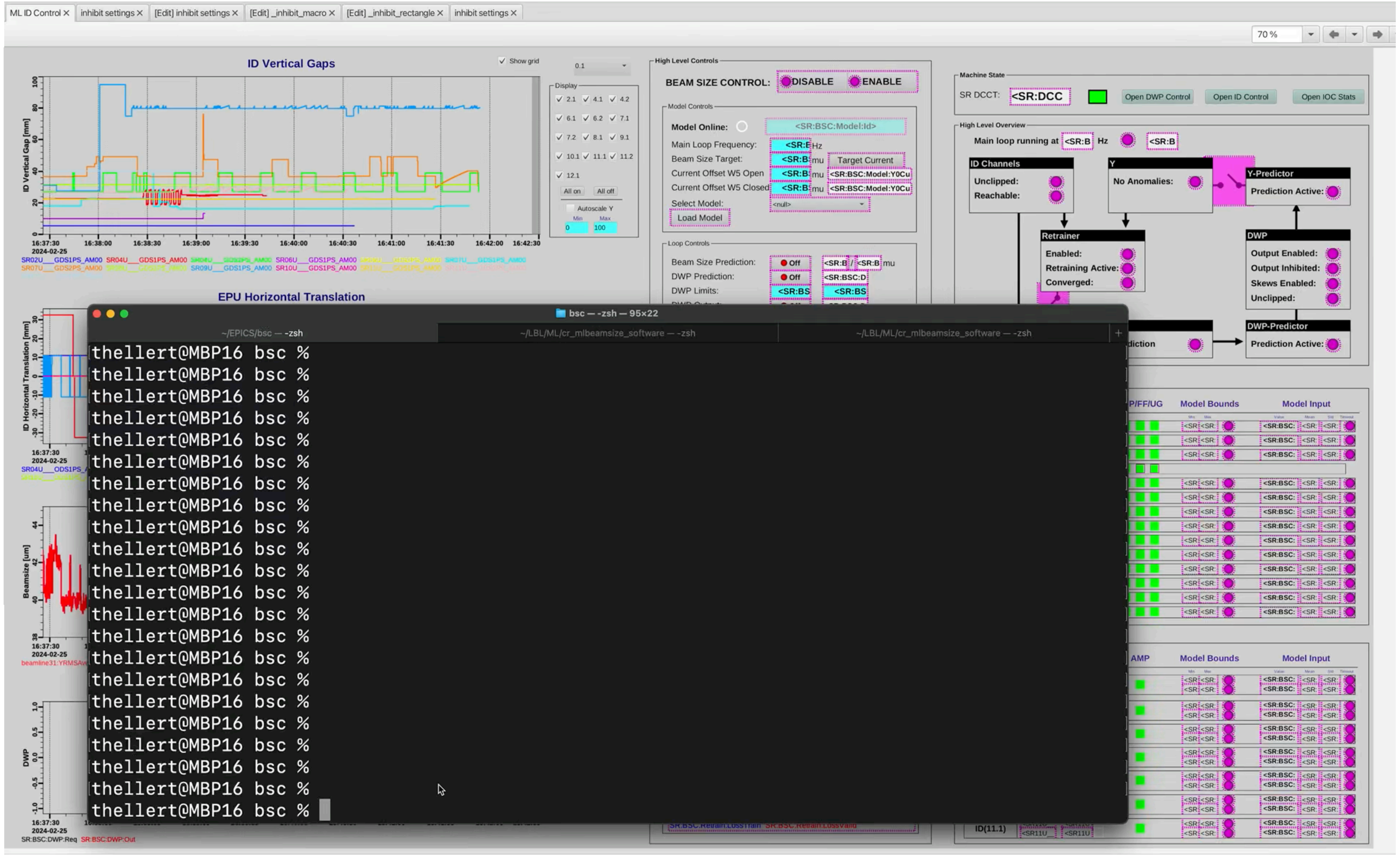


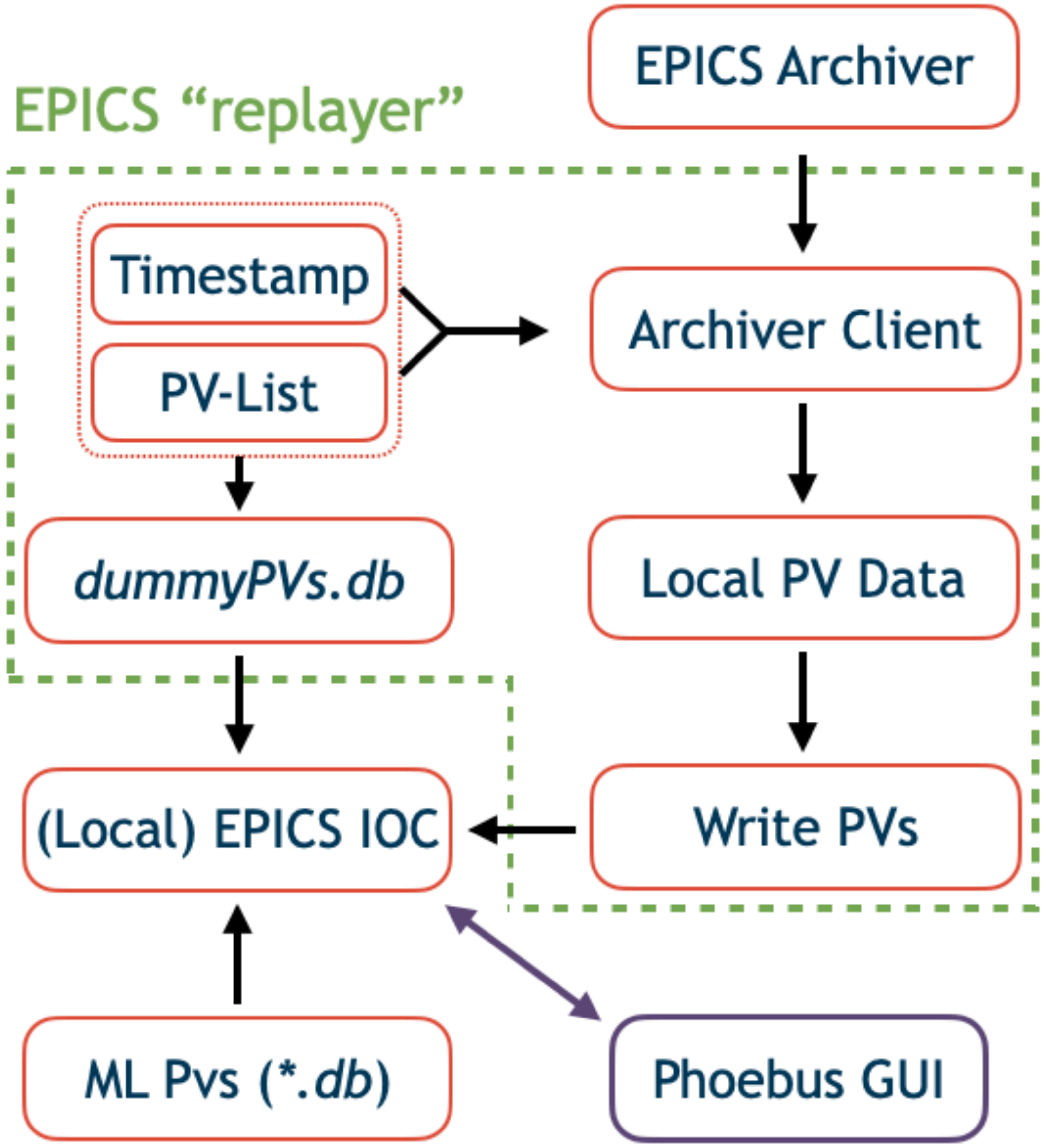
Archiver Replayer: Invaluable R&D Tool

- Local IOC Sandbox:
  - Beam size control IOC acting as a development sandbox
- Archived Data Utilization:
  - Archiver client to download historical data
  - dummyPVs.db automatically generated from archiver client
  - Python loop over PVs in **.csv* file
- Phoebus GUI Integration
  - Phoebus running locally
- Efficient R&D and Deployment:
  - Enabled rapid prototyping and debugging of the complete pipeline



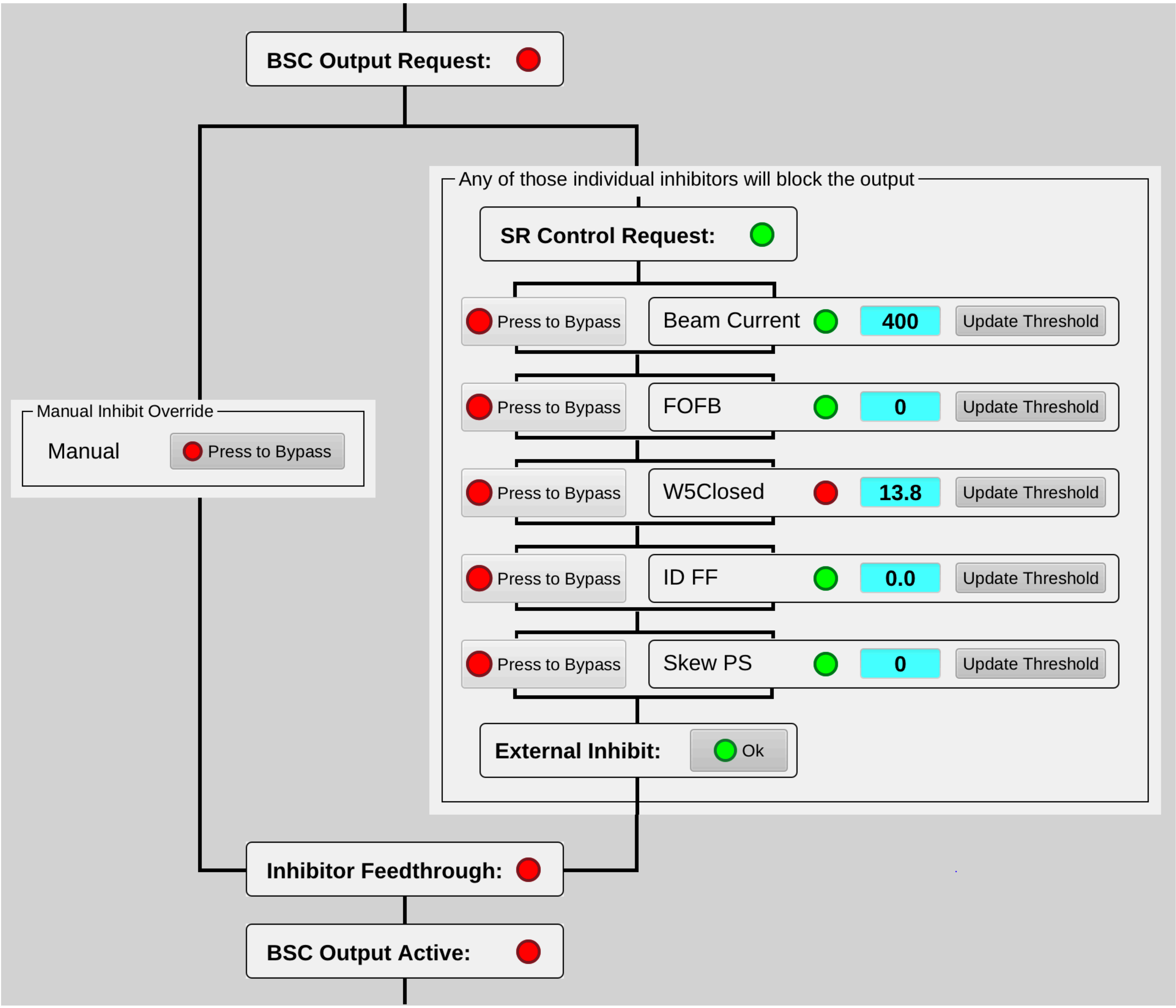
Archiver Replayer: Invaluable R&D Tool





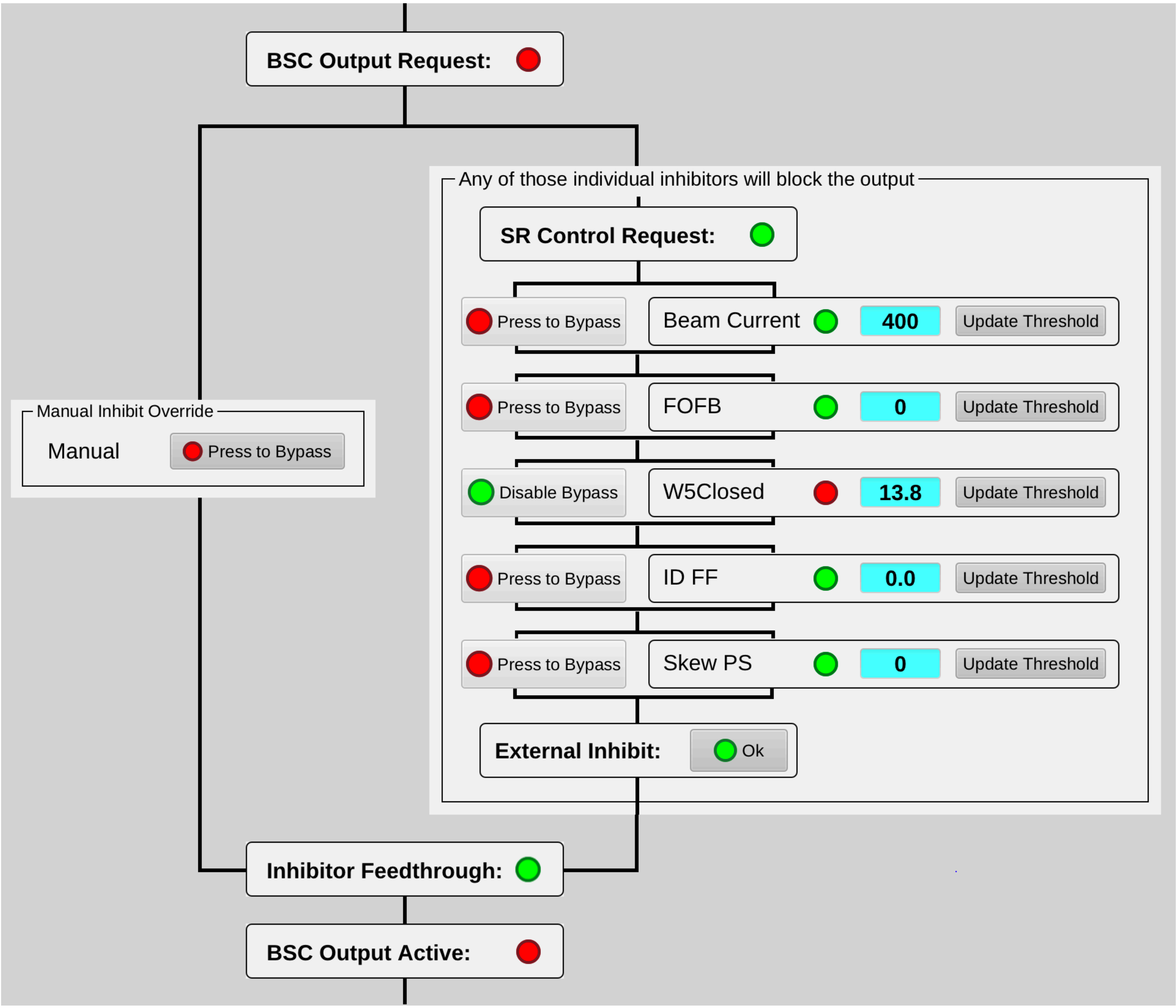
Inhibitor Chain

- Inhibitor Chain:
  - Can not activate beam size control
  - Can only pause operation
  - Includes manual override options
- Critical Conditions:
  - Beam current, FOFB, local ID FF
  - Skew quads ok, W5 closed
  - Control room request
- Crucial for Reliable Operation



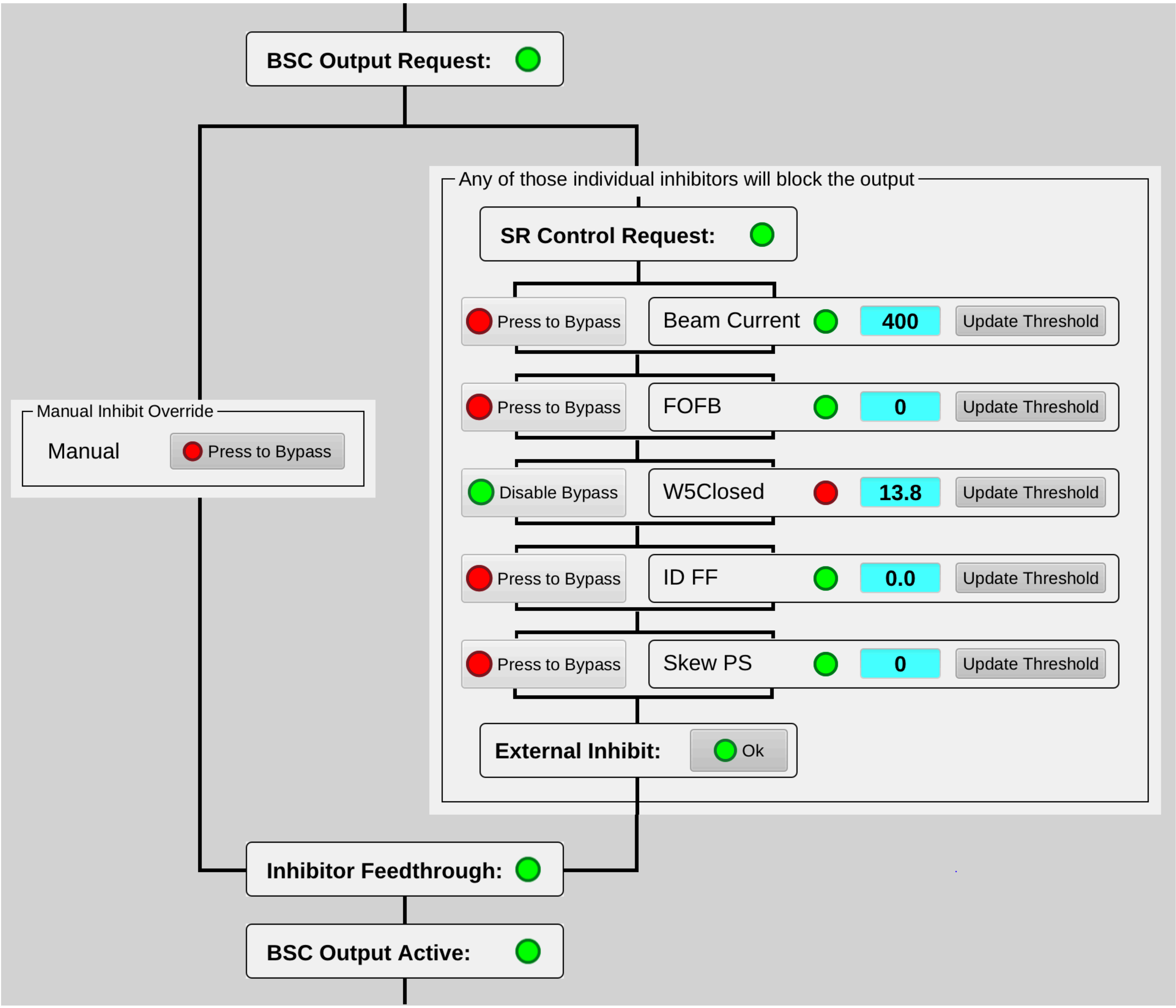
Inhibitor Chain

- Inhibitor Chain:
  - Can not activate beam size control
  - Can only pause operation
  - Includes manual override options
- Critical Conditions:
  - Beam current, FOFB, local ID FF
  - Skew quads ok, W5 closed
  - Control room request
- Crucial for Reliable Operation



Inhibitor Chain

- Inhibitor Chain:
  - Can not activate beam size control
  - Can only pause operation
  - Includes manual override options
- Critical Conditions:
  - Beam current, FOFB, local ID FF
  - Skew quads ok, W5 closed
  - Control room request
- Crucial for Reliable Operation

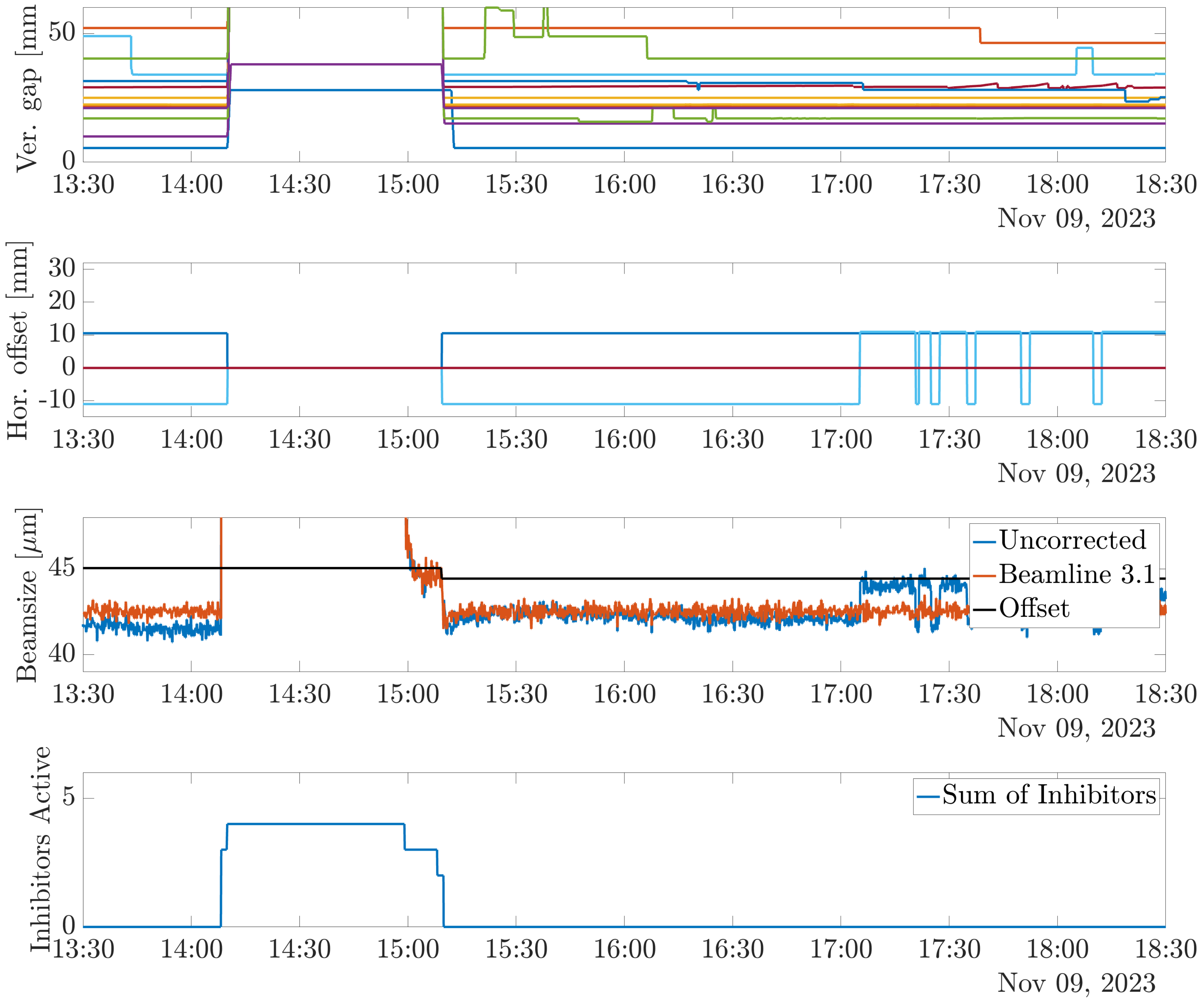


Hands-Off Beam Outage Recovery

- Beam Outage and Recovery Events
  - 12 beam outages recorded in 2 months
  - Autonomous reactivation of BSC algorithm
  - No manual intervention required
  - Automated model update

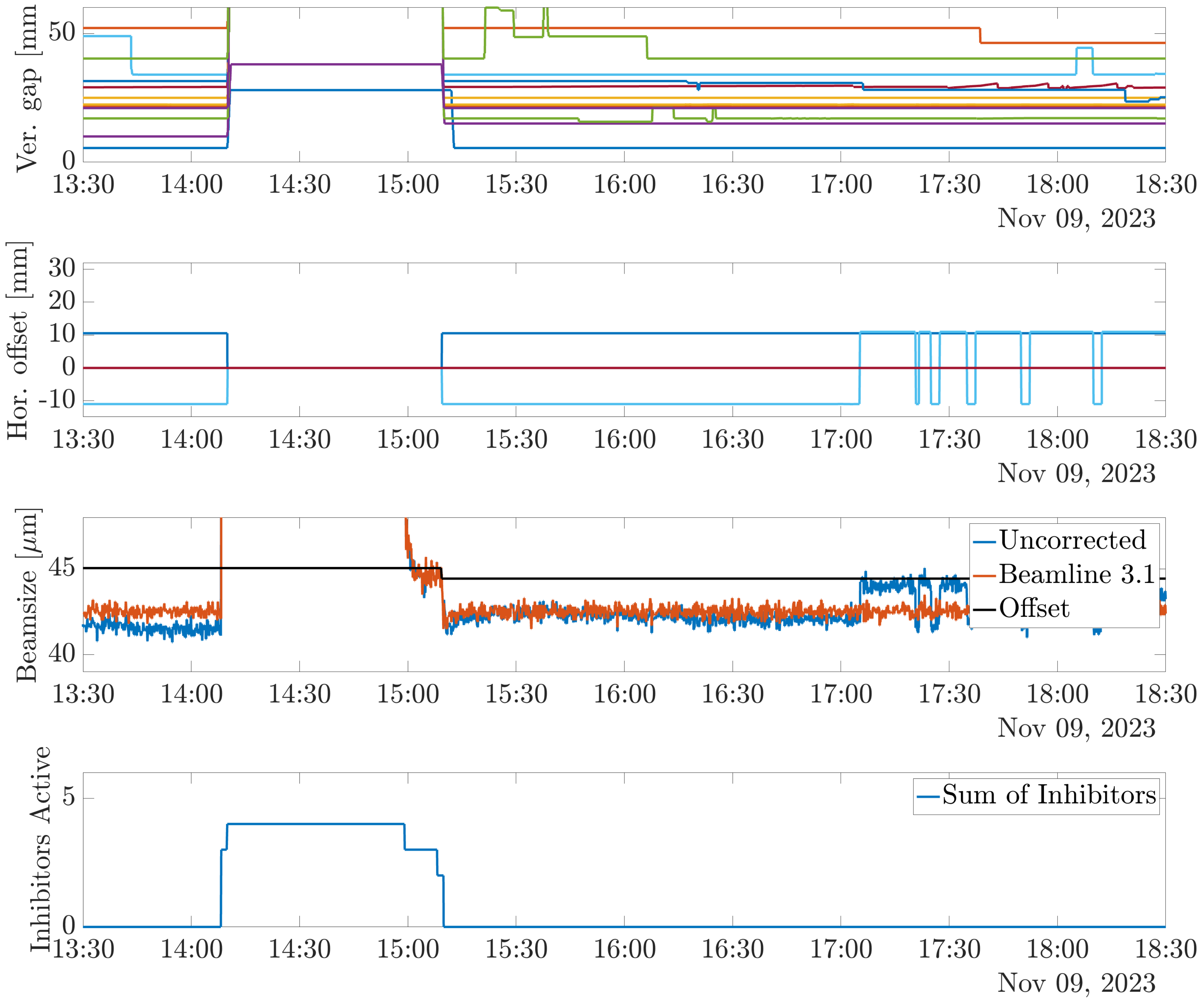
Hands-Off Beam Outage Recovery

- Beam Outage and Recovery Events
  - 12 beam outages recorded in 2 months
  - Autonomous reactivation of BSC algorithm
  - No manual intervention required
  - Automated model update
- Outage Example: RF Power Trip Incident:
  - RF power trip at 14:07
  - Triggered inhibitor PVs
  - BSC disengaged
  - IDs opened for refilling the machine



Hands-Off Beam Outage Recovery

- Beam Outage and Recovery Events
  - 12 beam outages recorded in 2 months
  - Autonomous reactivation of BSC algorithm
  - No manual intervention required
  - Automated model update
- Outage Example: RF Power Trip Incident:
  - RF power trip at 14:07
  - Triggered inhibitor PVs
  - BSC disengaged
  - IDs opened for refilling the machine
- Restoration Process:
  - Injection process from 14.55 am to 15.10
  - Conditions met for FF algorithm from 15.12
  - Vertical beam size back to target 42.5um



Summary

- Vertical Beam Size Variation at ALS
  - Dominating source of variation at STXM beamlines
  - Conventional correction techniques insufficient
- Model Development:
  - Comprehensive model- and hyper parameter search
  - Evaluation on historical user operation data
- Online Finetuning:
  - Continuous fine tuning of base model during operation
  - Outperforms conventional feedback correction
- Routine Deployment:
  - Utilization of EPICS backend and Phoebus Frontend
  - ML FF in routine operation since October’23