How will AI enable autonomous particle accelerators?

V. Kain
Data Science for Beam Operation
Beams Department, CERN

Why autonomous accelerators? Examples...

Summary talk *CERN Injector and Experimental Facility Workshop (IEF) '21*

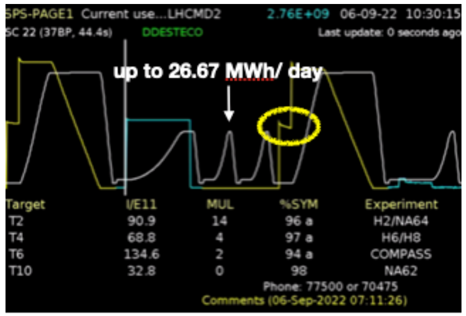
2. Address reproducibility and availability

- Availability OK, under control of Groups. **Reproducibility** is critical concern with increasing flexibility and multi-destination operation
- Transmission problems and instability in beam delivery in many locations. "Need more time in 2022"→have to ensure this is there (add in schedule?) #A
- Addressing reproducibility relies on many factors including equipment, accelerator modelling and high-level controls approach

Input from CERN Joint Accelerator Performance Workshop'22

→ **Hysteresis** is severe limitation for efficiency and flexibility in most machines, current mitigation methods wasting energy

- ~ 15 % of yearly cost of SPS fixed target cycle for "waste" cycles and quasi-degauss Cycle MD1



Efficiency Think Tank (ETT): Oct ’22 - Mar ‘23

ETT = Body for brainstorming for strategy definition for more efficient CERN accelerator exploitation

Response to *Injector and Experimental Facility Workshop 2021* concerns with efficiency and reproducibility

- large extended team for community discussions, small core team to synthesise directions

Wide range of efficiency topics touched

- shorter turn-around, more flexibility, energy efficiency,…

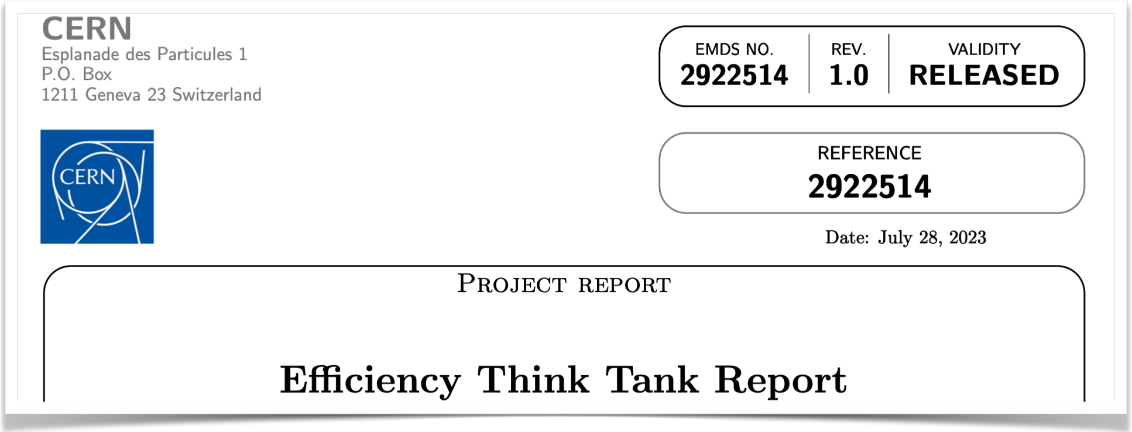
Key target areas identified

- 7 high priority recommendations

→ CERN accelerator sectorwide project to implement recommendations

- **Efficient Particle Accelerators (EPA) project:** 5 year project → improvements ready for HL-LHC

7 recommendations → Automating exploitation


CERN
Esplanade des Particules 1
P.O. Box
1211 Geneva 23 Switzerland

EMDS NO.	REV.	VALIDITY
2922514	1.0	RELEASED

REFERENCE
2922514

Date: July 28, 2023

PROJECT REPORT

Efficiency Think Tank Report

1. **Hysteresis compensation**
2. **Automatic and dynamic beam scheduling**
3. **Automatic LHC filling**
4. **Auto-pilots**
5. **Automatic fault analysis, recovery and prevention**
6. **Automatic testing and sequencing**
7. **Automatic parameter optimisation**

7 recommendations → Automating exploitation

CERN
Esplanade des Particules 1
P.O. Box
1211 Geneva 23 Switzerland

EMDS NO.	REV.	VALIDITY
2922514	1.0	RELEASED

REFERENCE
2922514

Date: July 28, 2023

PROJECT REPORT

Efficiency Think Tank Report

1. **Hysteresis compensation**
2. **Automatic and dynamic beam scheduling**
3. **Automatic LHC filling**
4. **Auto-pilots**
5. **Automatic fault analysis, recovery and prevention**

→ **Fully automated standard physics operation**

6. **Automatic testing and sequencing**
7. **Automatic parameter optimisation**

→ **Goal: reduce commissioning time by 50 %**

7 recommendations → Automating exploitation

CERN
Esplanade des Particules 1
P.O. Box
1211 Geneva 23 Switzerland

EMDS NO.	REV.	VALIDITY
2922514	1.0	RELEASED

REFERENCE
2922514

Date: July 28, 2023

PROJECT REPORT

Efficiency Think Tank Report

- **AI** Hysteresis compensation
- 2. Automatic and dynamic beam scheduling
- **AI** Automatic LHC filling
- **AI** Auto-pilots
- **AI** Automatic fault analysis, recovery and prevention

→ Fully automated standard physics operation

- 6. Automatic testing and sequencing
- **AI** Automatic parameter optimisation

→ Goal: reduce commissioning time by 50 %

Predicting magnetic hysteresis and eddy current effects

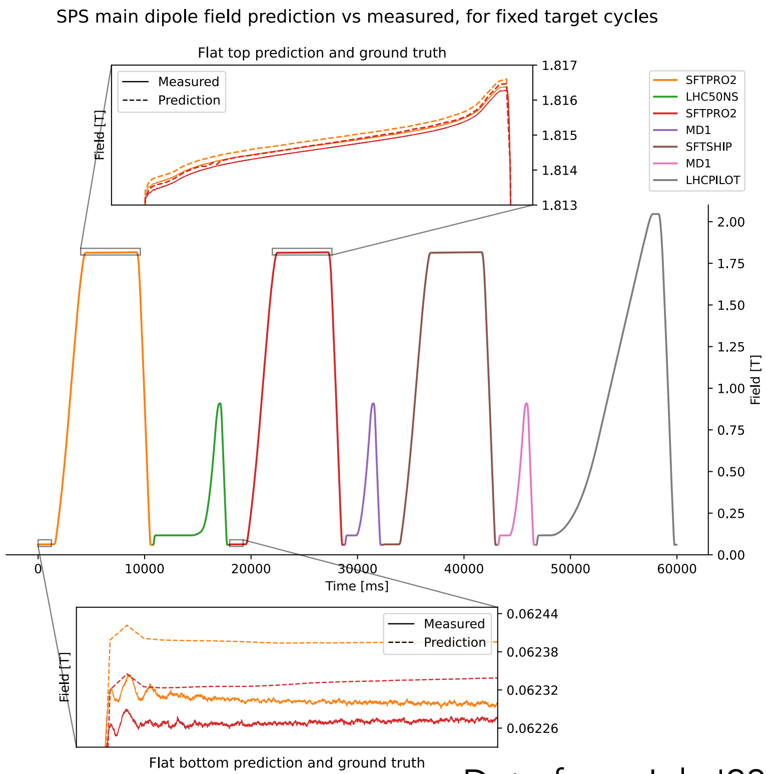
Potentially game-changing!

Time-series forecasting problem: need magnets to be measured on test bench
$[B_t, B_{t+1}, \ldots, B_{t+n-1}], [I_t, I_{t+1}, \ldots, I_{t+n+N}] \rightarrow [B_{t+n}, B_{t+n+1}, \ldots, B_{t+n+N}]$

First operational experience:

- **feedforward correction** triggered before every cycle
- accuracy not sufficient yet

First results PhyLSTM for SPS main dipoles assuming $\ddot{B} + g(B, \dot{B}) = \Gamma I(t)$, next: Transformers

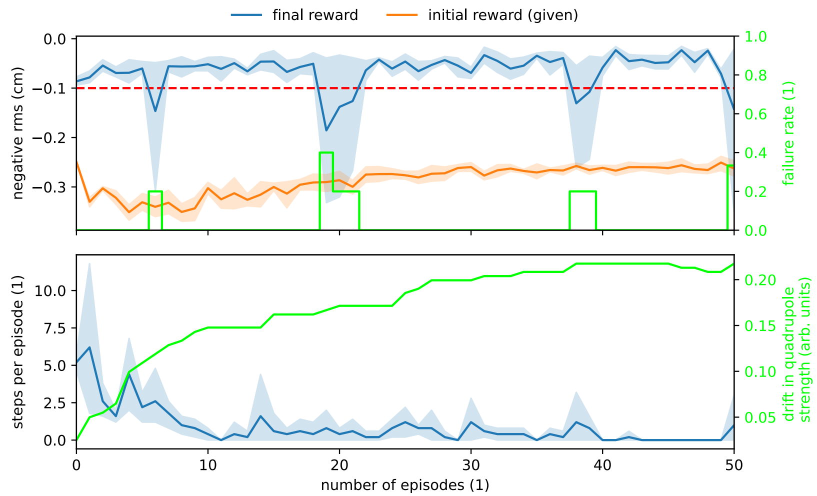


Data from July '23

Auto-pilots: Adaptive control for time-varying systems

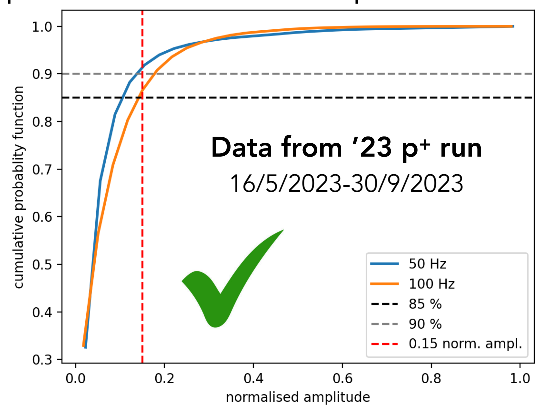
Time-varying response: limitation for using data-driven surrogate models for control so far.

- ◉ Optimal control with Gaussian Processes
  - ✳ Model-based RL with GPs for dynamics model: very sample-efficient
  - ✳ Deal with time-varying systems → add time t in state space
  - ✳ Example: with 10 DOF for trajectory steering problem and time-varying lattice at AWAKE electron line



- ◉ Continuous control with Adaptive Bayesian Optimisation
  - ✳ Bayesian Optimisation for time-varying systems
  - ✳ Model objective function f as $f(x, t)$: separate kernels for x, t
  - ✳ Example: successful $n \times 50$ Hz intensity ripple control for slow extracted spill for CERN North Area fixed target experiments

Requirement: norm. ampl. < 0.15 for 85 %



Plans and remaining challenges

5-year project to automate CERN's accelerator fleet

- Resources to be approved middle of next year, GPUs to come as soon as available

Scalable solutions for many tasks available

- Data processing framework (UCAP) with GPUs and frameworks to define RL problems and other controllers
- CERN Machine Learning Platform as model repository

Many challenges still ahead

- Active learning for online systems, dealing with uncertainties,..

Need to organise transition to new exploitation model