Data-Driven Tokamak Plasma Magnetic Response Modeling

Jiyuan Li [1], Zongyu Yang [1], Tailin Wu [2], Yihang Chen [1], Junzhao Zhang [1], Da Li [1], Jingyue Yuan [1], Ruoshu Qiu [1], Wulu Zhong [1], Bo Li [1]

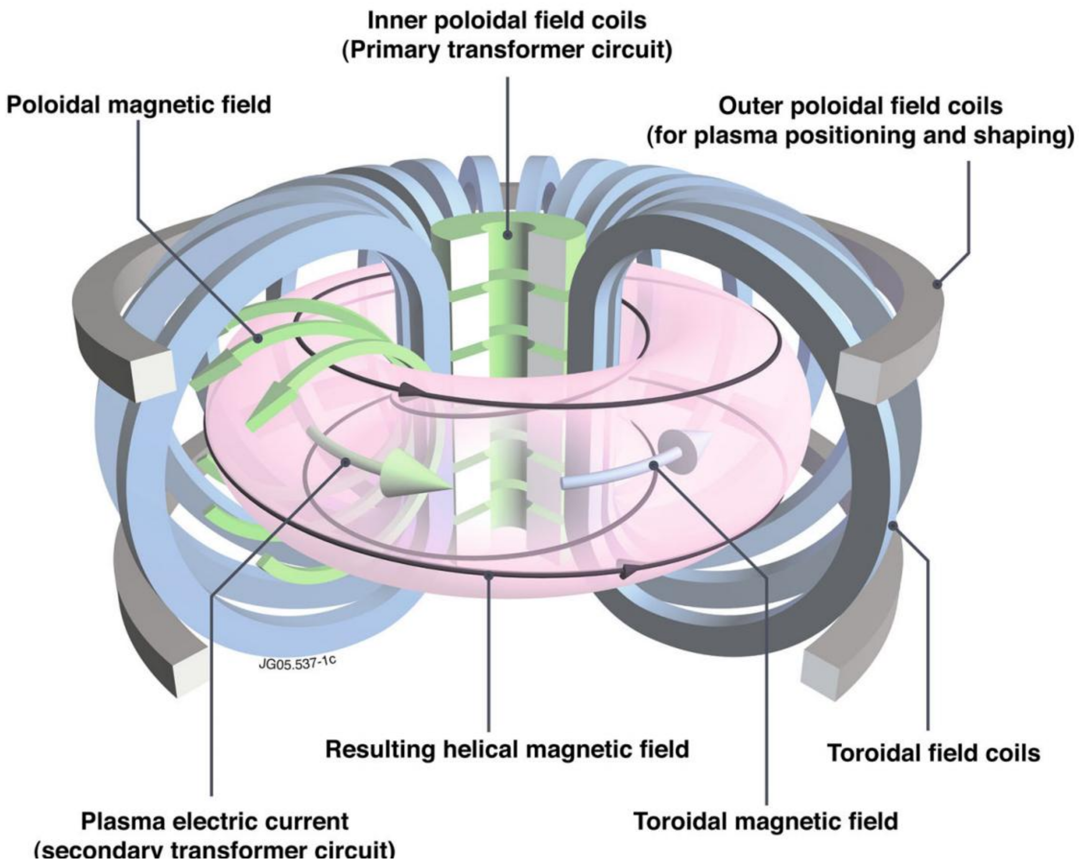
[1]Southwestern Institute of Physics

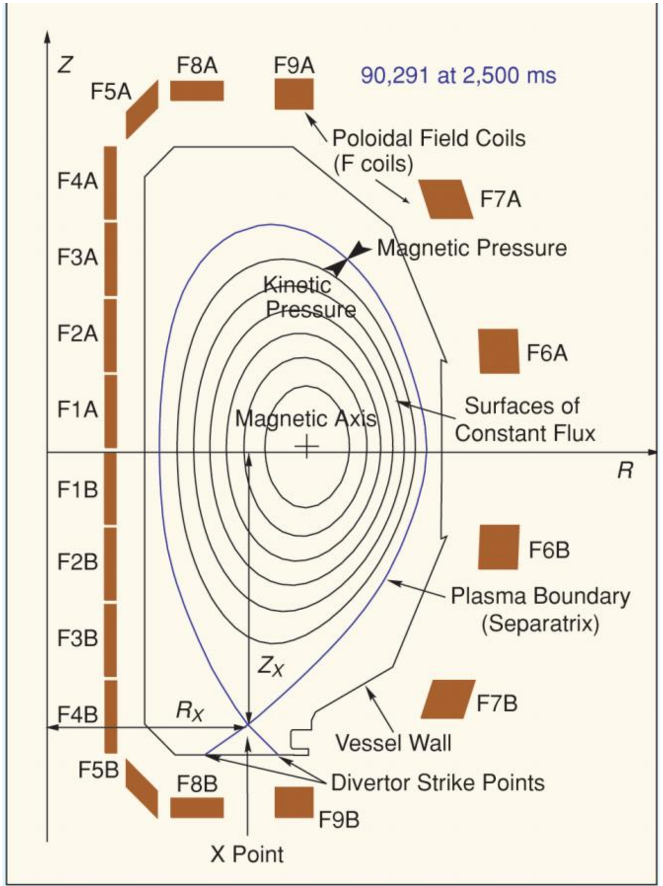
[2]Westlake University

Contents

Research Background

- Tokamaks use magnetic fields to confine hot plasma for fusion.
- Magnetic control precisely adjusts field interactions to maintain plasma position and stability, preventing wall contact or disruptions - essential for safe fusion operation.





Research Background

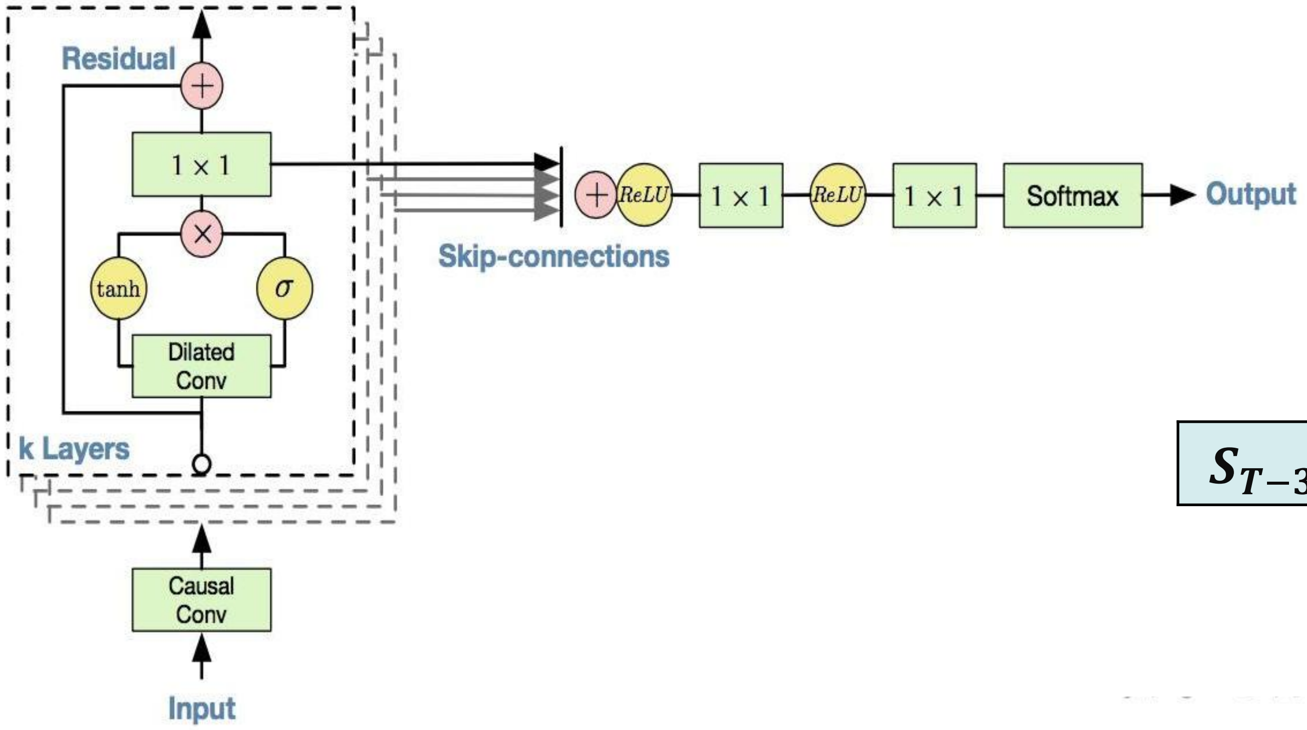
- **Importance of Tokamak Plasma Magnetic Response Simulation Models:**
  - Provide safe and efficient environment for developing and testing new control strategies, avoiding risks of direct experimentation on actual devices.

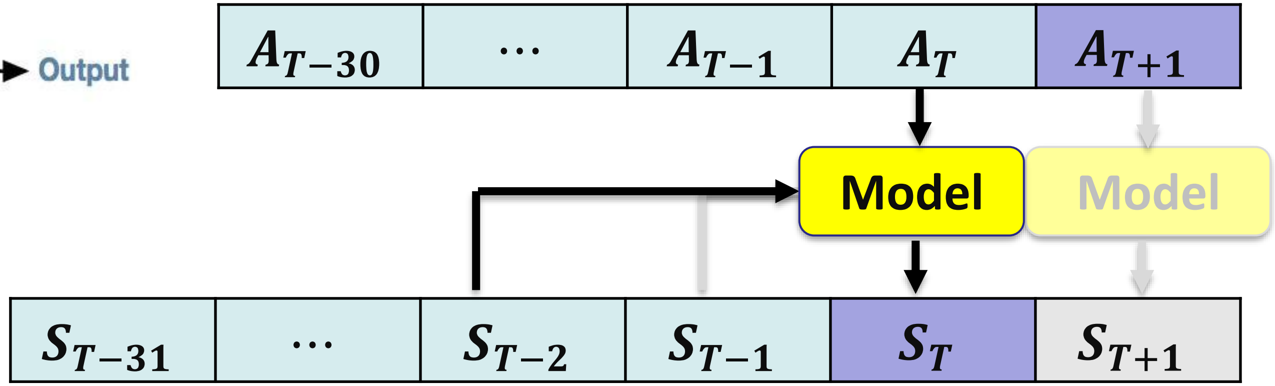
- **Traditional Plasma Magnetic Response Models (such as RZIP, CREATE-L) have limitations:**
  - Require simplified assumptions about plasma properties.
  - High computational complexity.
  - Although linear approximation can improve computational speed, it reduces model accuracy, especially in long-term experiments where errors accumulate continuously.

- **Data-driven methods show clear advantages:**
  - No need for simplified assumptions about plasma properties.
  - Low computational cost, high-speed real-time applications.
  - Provide high-precision simulation environment for control strategy development and fault detection.

Model Framework

- **Wavenet: A deep generative neural network model proposed by DeepMind team in 2016**
  - Causal dilated convolution structure.
  - Residual and skip connections.
- **Trained WaveNet Model Autoregressive Usage**
  - A: PF+CS coil currents
  - S: Plasma current and configuration parameters



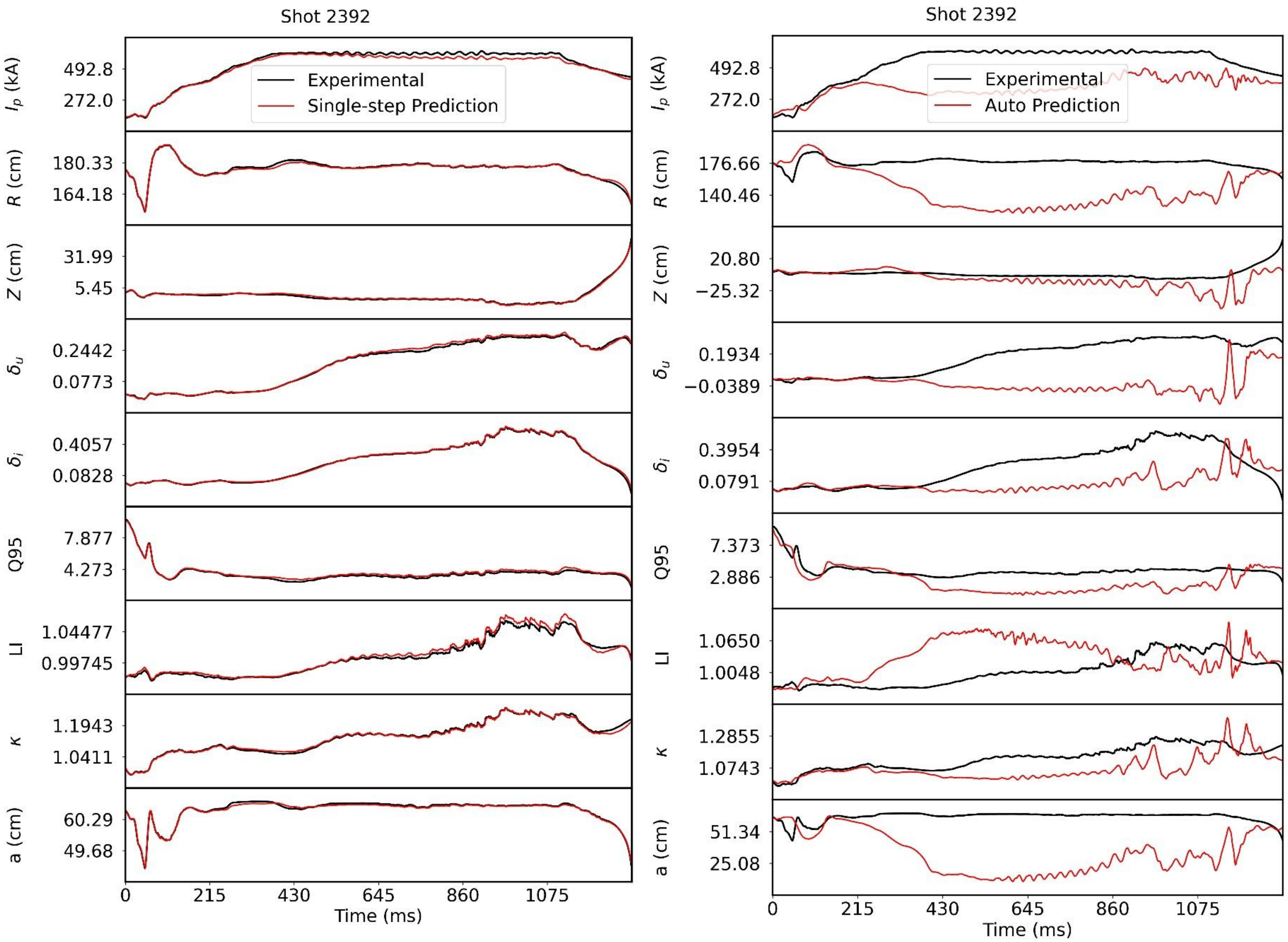


Model Framework

- Dataset:100 shots for validation
- Evaluation Metric： $R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \tilde{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$ 。
  - Single-step prediction

    $R^2 \approx 0.9$
  - Autoregressive prediction

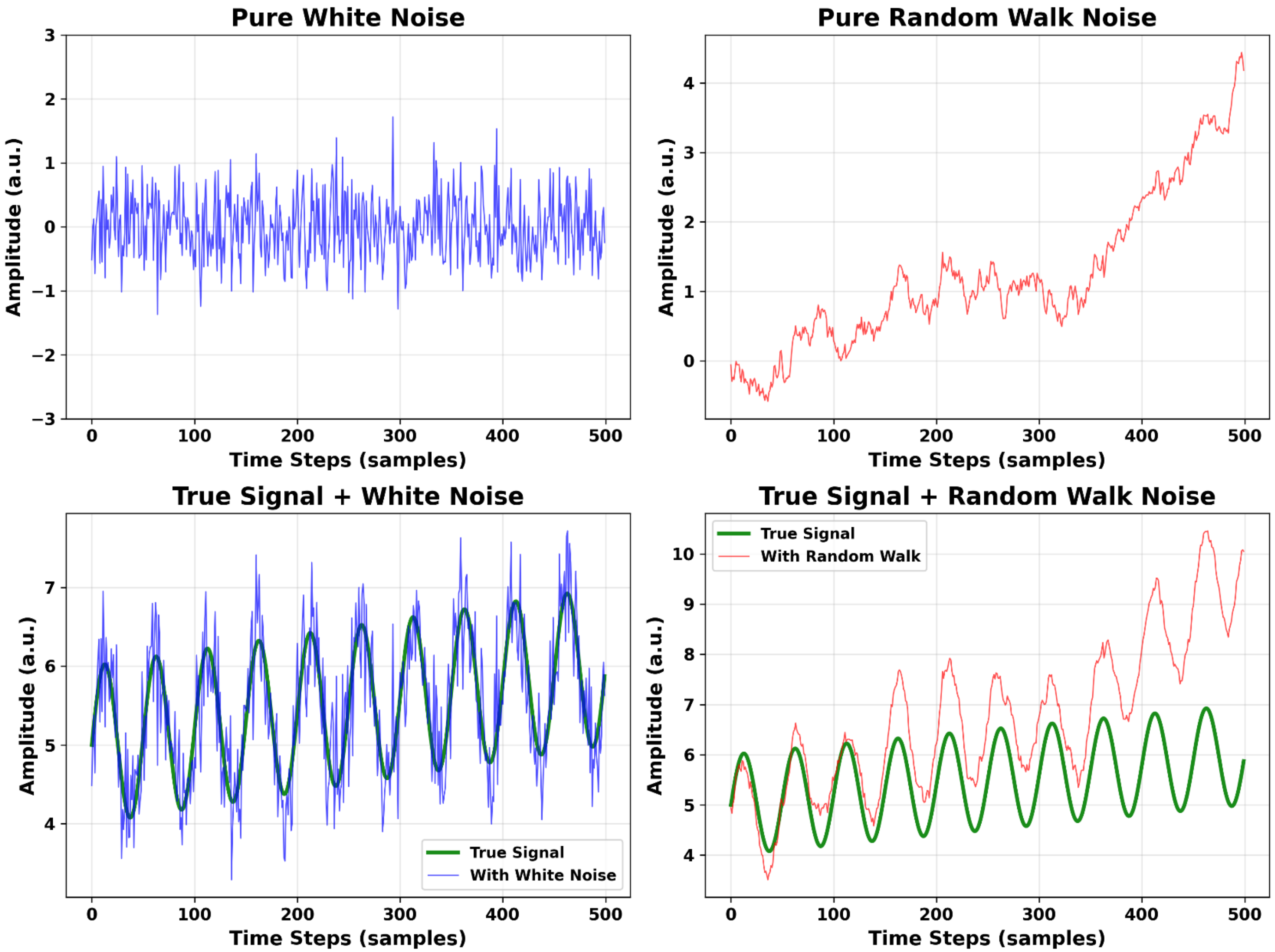
    $R^2 \approx 0.6$
- Need optimization for autoregressive prediction
  - Error Accumulation
  - Curve fluctuation fitting accuracy



Algorithm Optimization（Error Accumulation）

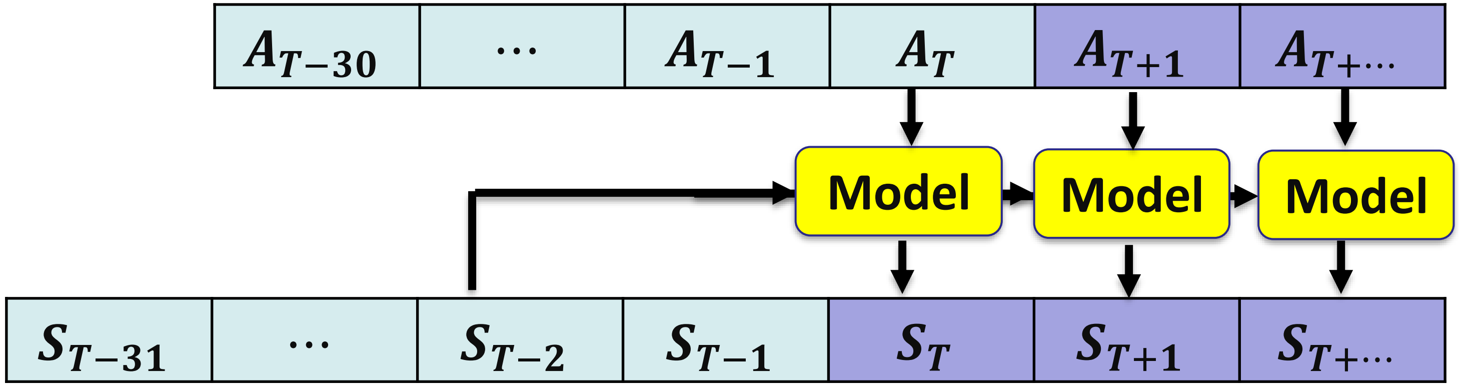
- Optimization 1：Input with Random Walk Noise
  - Random Walk Noise is a special type of time-correlated noise, characterized by current noise value equals previous time step noise value plus a Gaussian random perturbation.
  - It can simulate error accumulation problems in autoregressive situations, improving model robustness and generalization ability.

$$N(t) = N(t-1) + \varepsilon(t) \qquad N(0) = \varepsilon(t)$$



Algorithm Optimization（Error Accumulation）

- Optimization 2： Autoregressive Multi-step Loss Function Calculation.
  - Consider cumulative errors of multiple prediction steps during training. Not only calculate single-step prediction loss, but also consider cumulative errors when current prediction results are used as input for subsequent predictions.
  - Different weights assigned to each output step during loss calculation in training.
  - Reduce the problem of accumulated errors; Improve the stability of long sequence prediction; Make the model more robust to its own prediction errors.



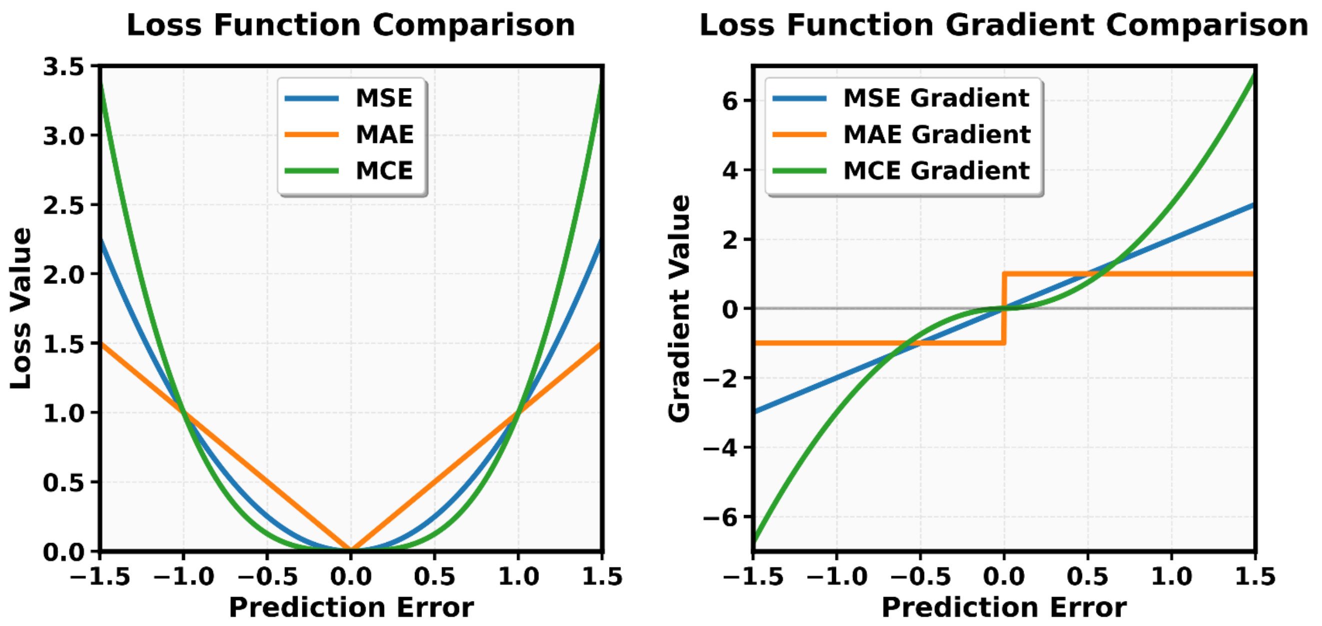
Algorithm Optimization (Accelerate Convergence, Improve Accuracy)

➢ Optimization 3: Composite Loss Function

- Combine three different error measures: MSE (Mean Squared Error); MAE (Mean Absolute Error) and MCE (Mean Cubic Error). Formula shown below, where λ_1 and λ_2 are weight coefficients controlling contribution ratio of different error measures.

$$Loss = MSE + \lambda_1 * MAE + \lambda_2 * MCE$$

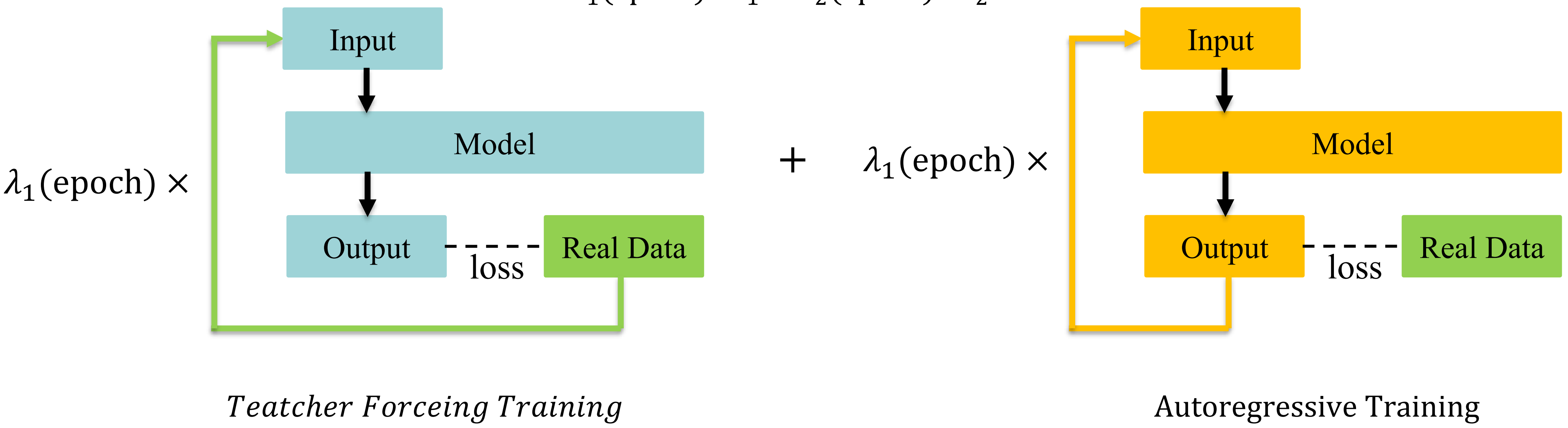
- MSE can achieve fast convergence, MAE provides constant gradient for continued optimization near optimal solution, MCE further amplifies punishment for large errors.



Algorithm Optimization (Accelerate Convergence, Improve Accuracy)

- Optimization 4: Teacher Forcing (Combined with Autoregressive Multi-step Loss Function)
  - Use real values instead of predicted values as input for subsequent steps during autoregressive multi-step loss function training
  - Accelerate training convergence, avoid instability caused by incorrect predictions in early training stages
  - Weights of both methods change with increasing epochs

$$L = \lambda_1(\text{epoch}) \cdot L_1 + \lambda_2(\text{epoch}) \cdot L_2$$



Results

➢ Dataset:100 shots for validation

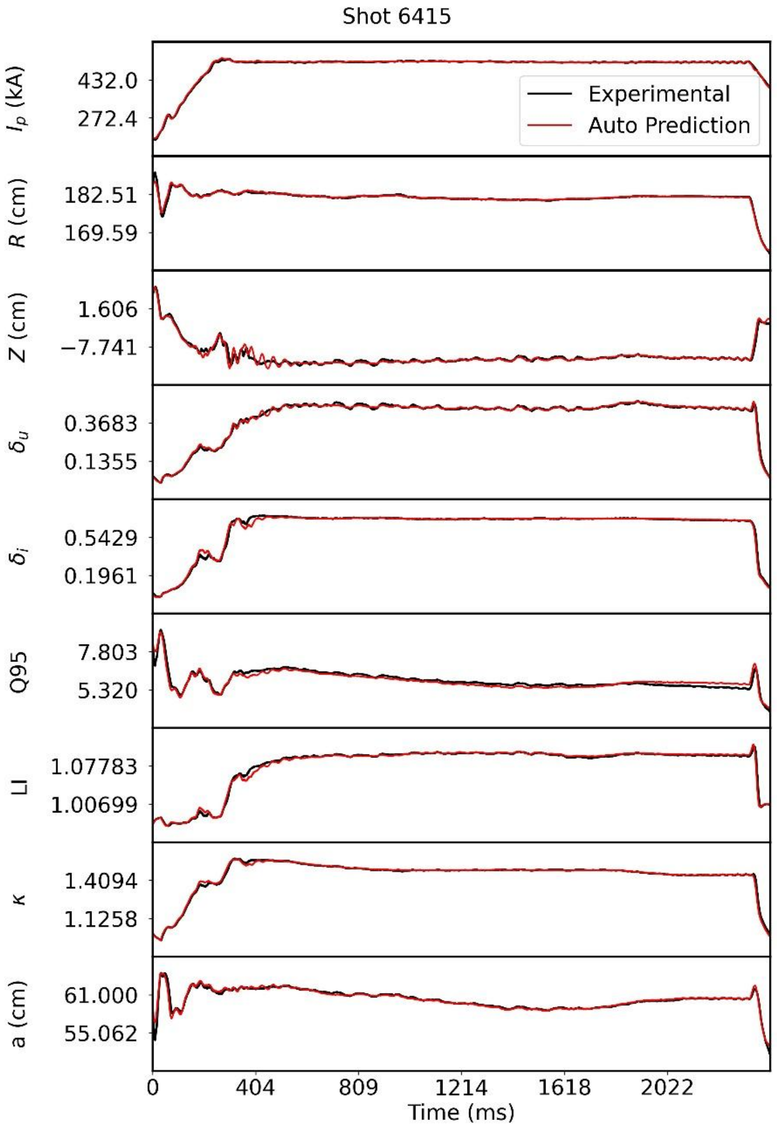
➢ Evaluation Metric： $R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \widetilde{y_i})^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$

- Optimization 1: Input with random walk noise
- Optimization 2: Autoregressive multi-step loss function calculation
- Optimization 3: Composite loss function
- Optimization 4: Teacher forcing (combined with autoregressive multi-step loss function)

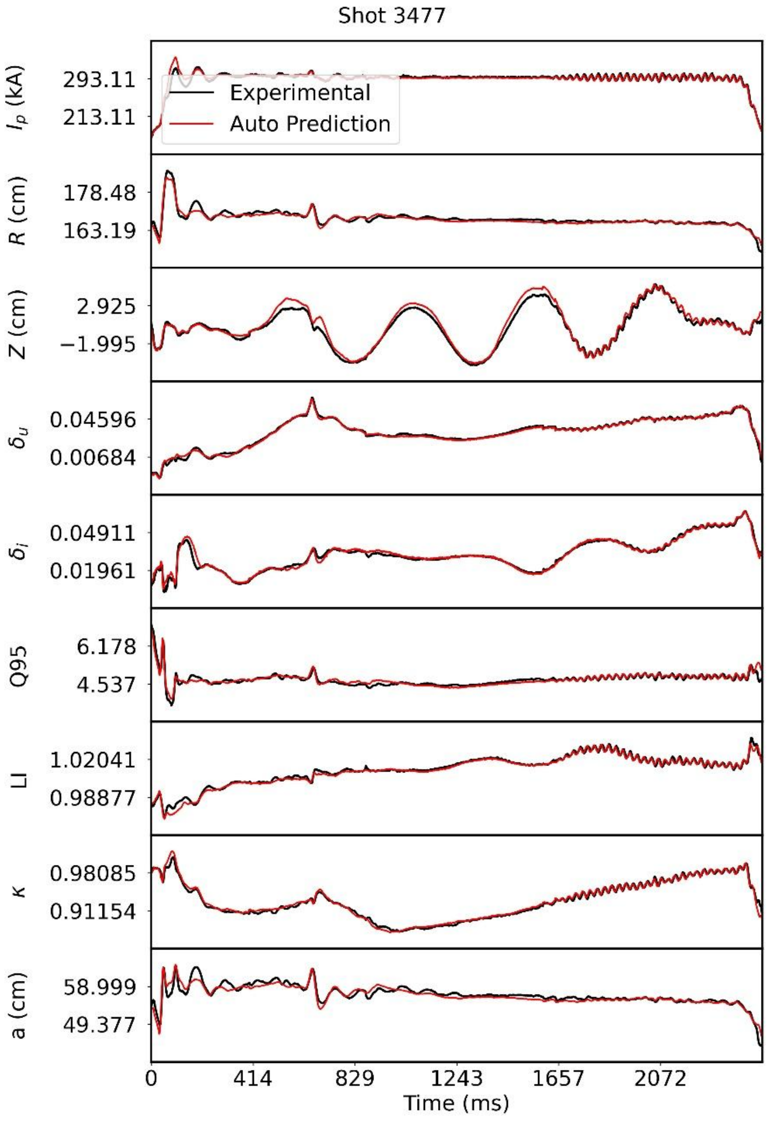
R^2	Baseline (Opt 1+2)	Opt 3	Opt 4	ALL Opt
IP	0.9477	0.9402	0.9451	0.9476
R	0.7397	0.7634	0.7716	0.7823
Z	0.8503	0.8572	0.8210	0.8632
Triangle_Up	0.9339	0.9361	0.9217	0.9425
Triangle_Down	0.9255	0.9335	0.9469	0.9526
Q95	0.6970	0.7304	0.7404	0.7512
LI	0.8432	0.8576	0.8740	0.8814
Elongation	0.9383	0.9432	0.9576	0.9637
a	0.7421	0.7398	0.7636	0.7728
Average	0.8464	0.8557	0.8602	0.8734

Results

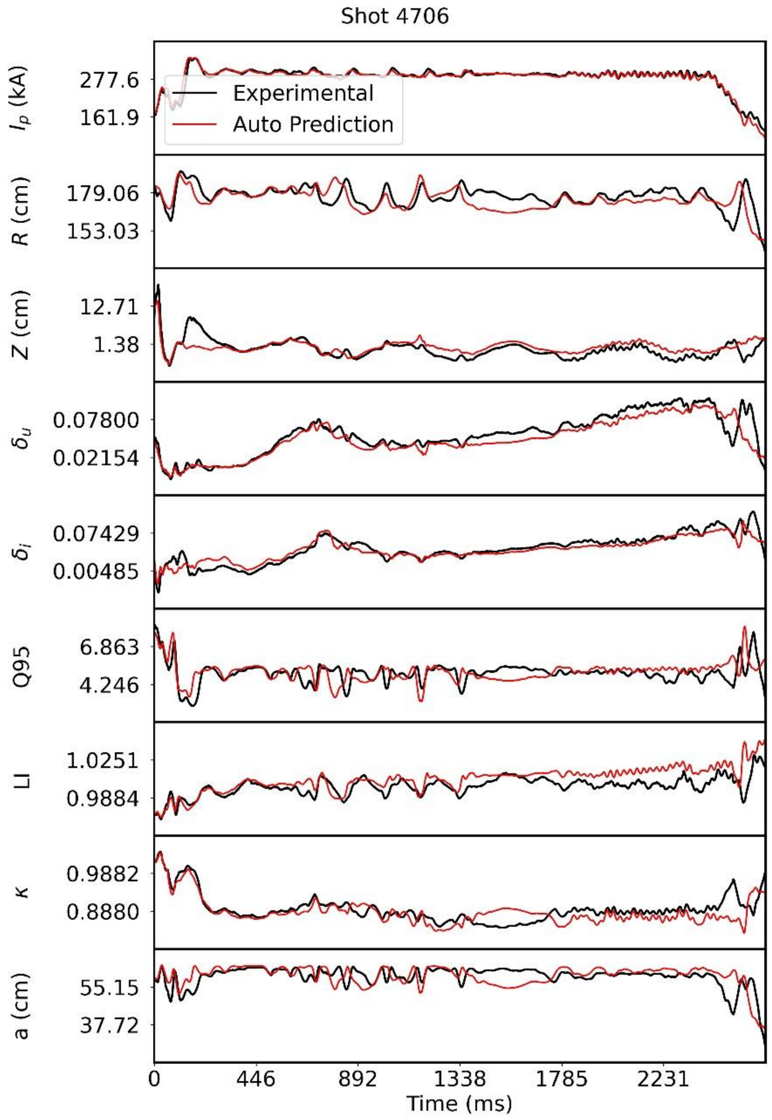
HL-3 SHOT 6415



HL-3 SHOT 3477

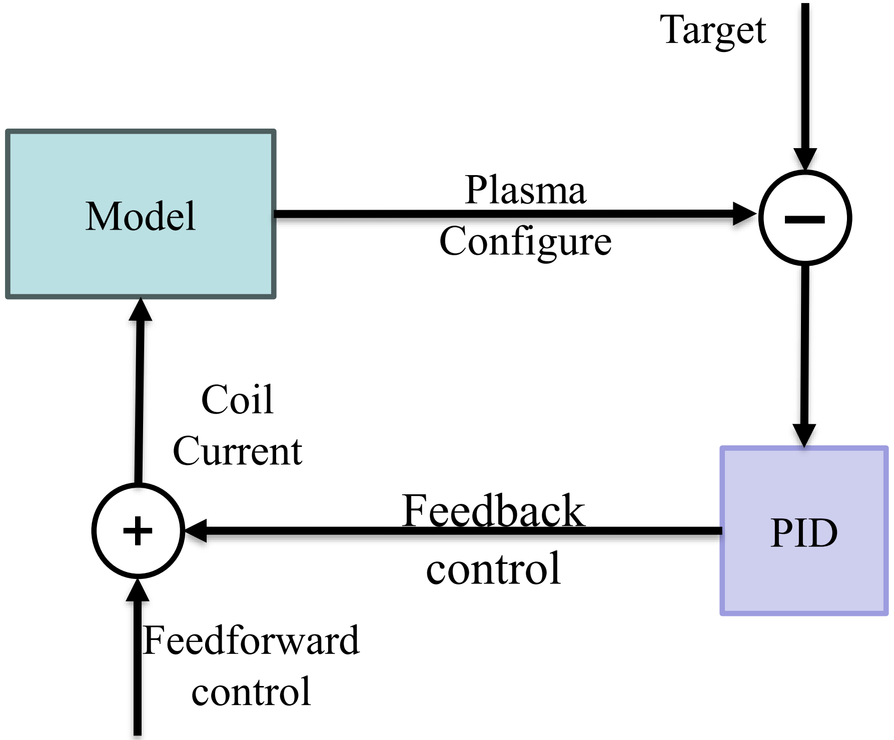


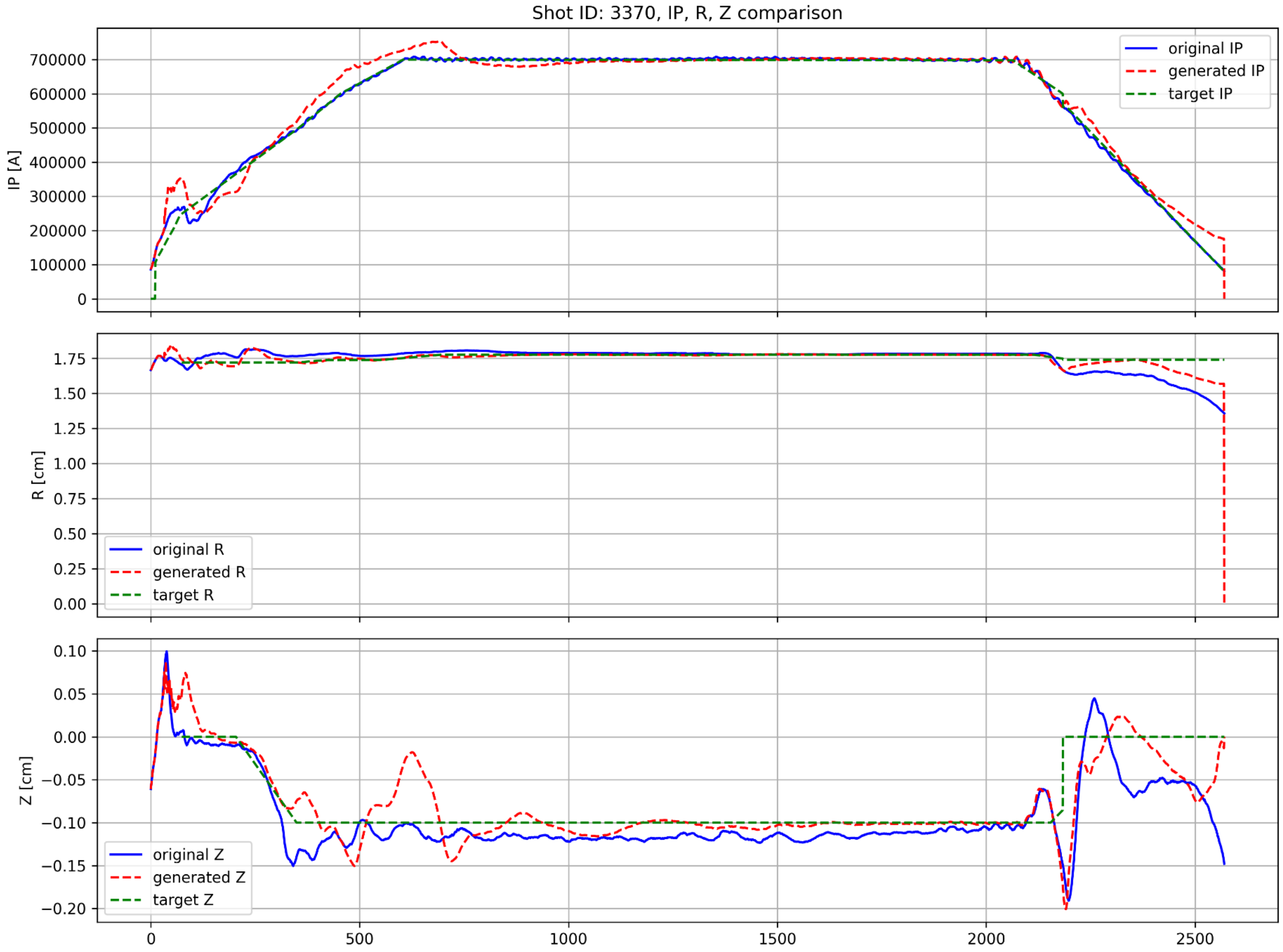
HL-3 SHOT 4706



Results

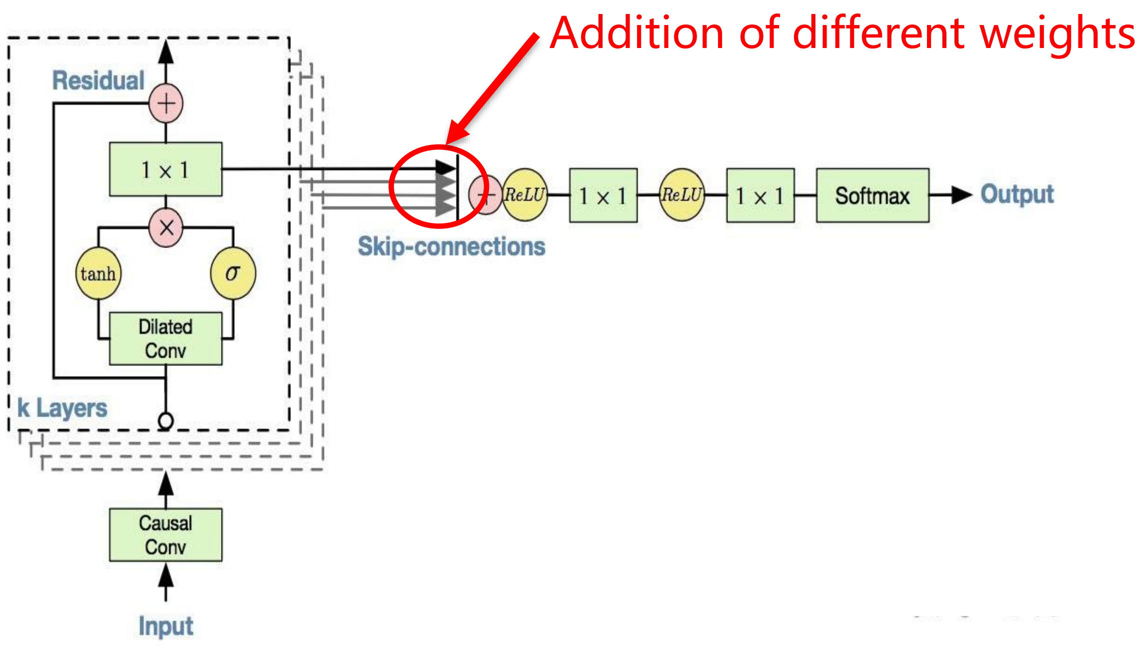
- Successfully integrated with traditional PID control algorithm

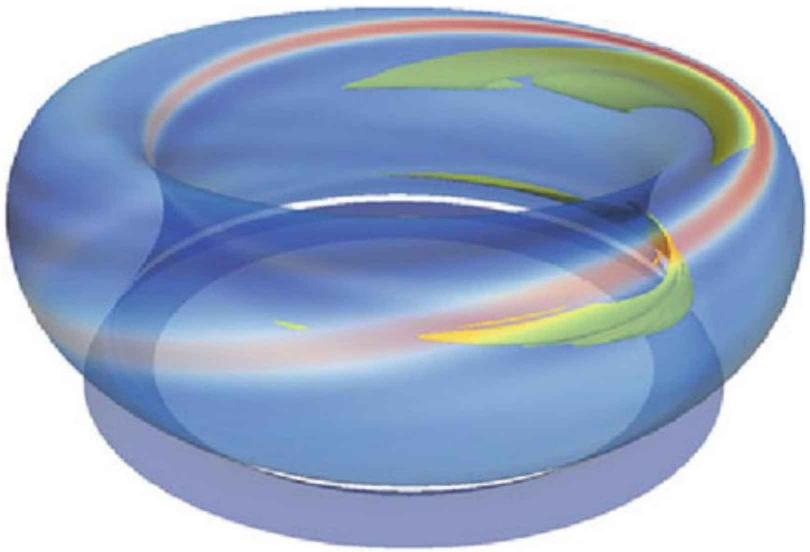




Discussion

- Applications: Offline simulation environment, MPC, Fault detection systems
- Model expansion: Extend from 1D time series prediction to 2D and 3D, enabling plasma profile and MHD instability autoregressive evolution.
- Architecture improvements: Modify skip connection weights in different WaveNet layers to focus on user-desired frequencies





Thank you!

Jiyuan Li
lijiyuan@swip.ac.cn
2025.9 ·**ShangHai**