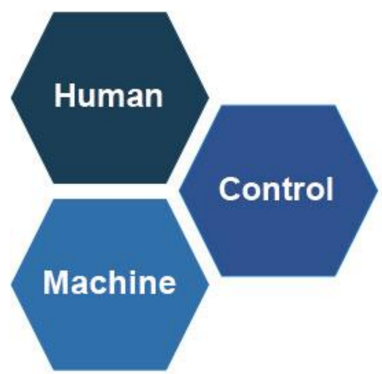
RL-based Adaptive Admittance Control for Optimal Landing of a Single-Leg Robot

Artificial Intelligence, M.S. Student, Jangho Kim

2025.09.16 Tue



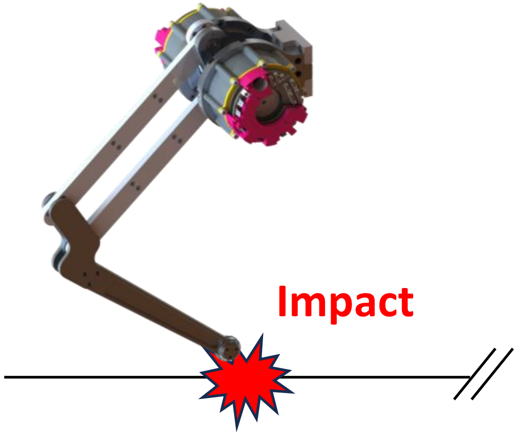
We are dedicated to helping people find better ability in motions through various technologies of robotics and control engineering.
Four keywords of our research are: Control, Actuator, Mobility and Muscle.

Knowledge and technologies on human,machine and control are converged to achieve our goals

Contents

Ⅰ. Introduction

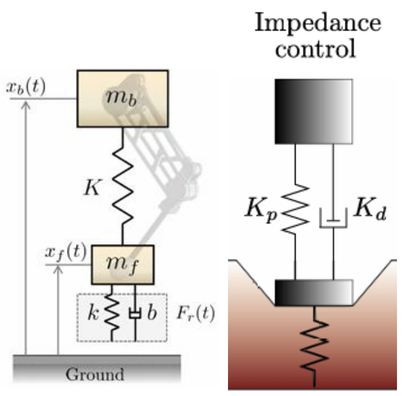
- Motivation



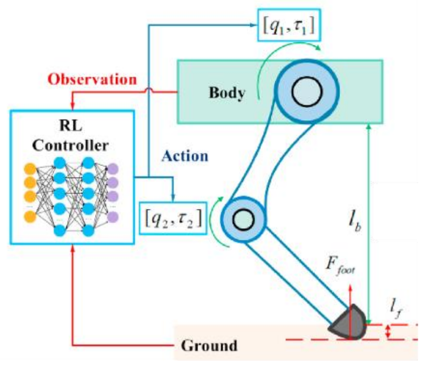
Extensive research has been conducted on legged robots, yet studies on impact mitigation remain insufficient.

- Conventional Methods

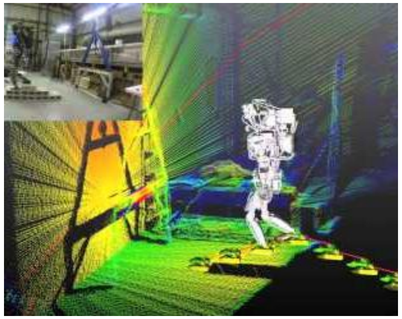
Model based



RL



SLAM



Model-based control struggles with modeling errors and irregular terrain, while end-to-end RL is time-consuming. Alternatively, SLAM relies on many sensors, leading to noise issues and high cost.

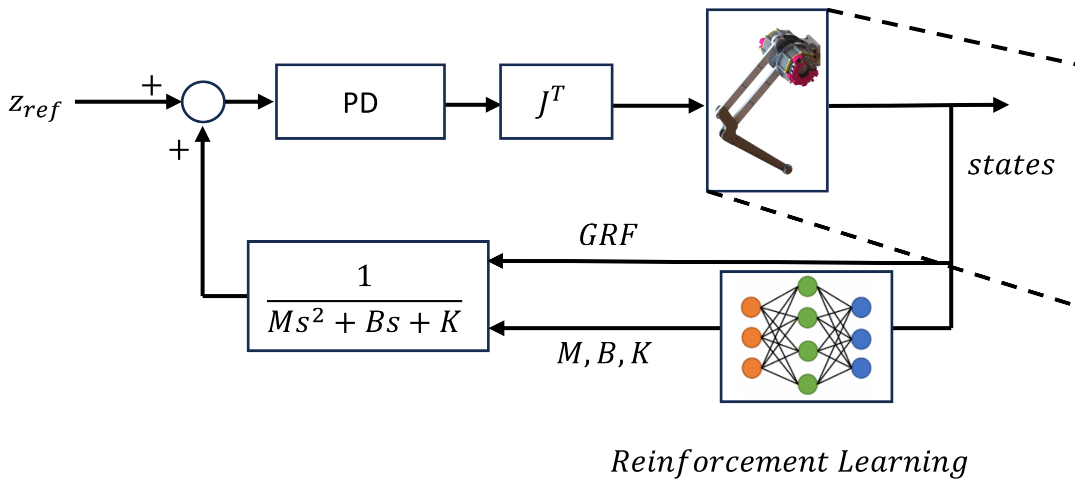
Ⅰ. Introduction

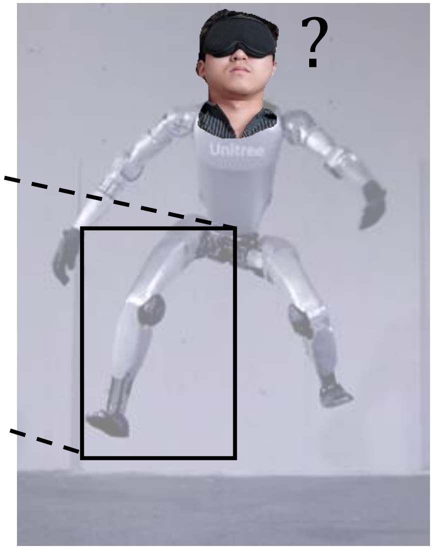
- Conventional Methods

Types	Pros	Cons
Model based	Interpretability	Model mismatch
Reinforcement Learning	High adaptability	Reward design complexity, Very Low interpretability
SLAM	Environment awareness	High computation cost
RL based Adaptive Admittance	Adaptability, Interpretability	Low interpretability

Thus, I propose RL-based Adaptive Admittance Control that combines the strengths of existing methods.

- RL-based Adaptive Admittance Control

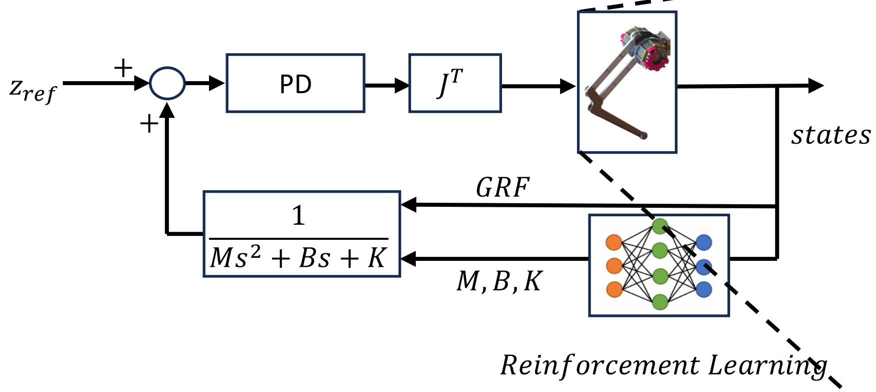


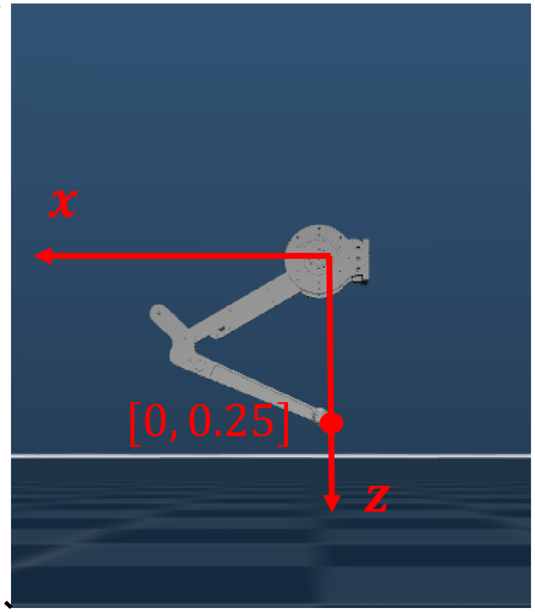


RL-based Adaptive Admittance Control learns time-varying admittance gains in real time using Reinforcement Learning.

The ultimate goal is to achieve stable landing even without prior knowledge of ground height.

- System Specification





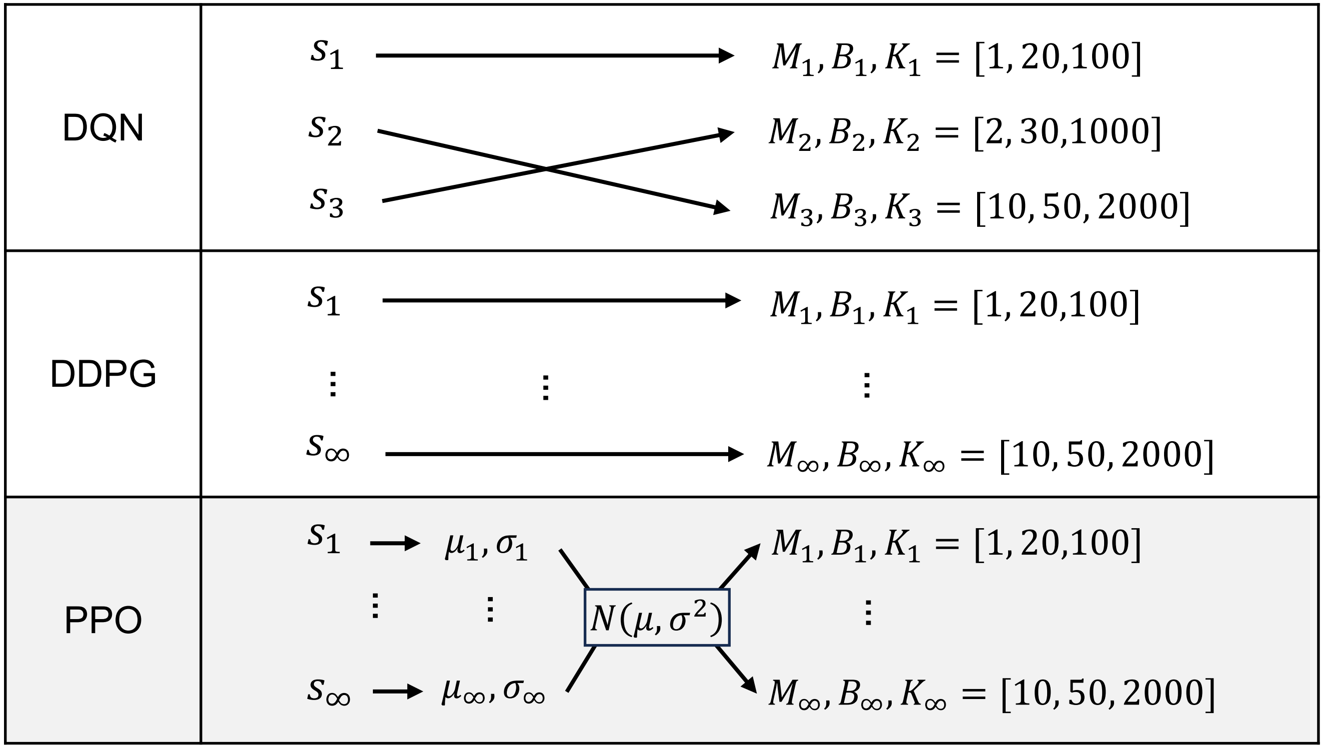
$$^{B}P_F^{ref} = [0, 0.25]$$

$$K_p, K_d = [2000,100]$$

$$M_d, B_d, K_d \rightarrow \textit{Determined by RL}$$

The simulation system is simplified to a 2D single leg (body is fixed along z-direction) with fixed reference and fixed PD gains, while M, B, K gains are optimized by RL.

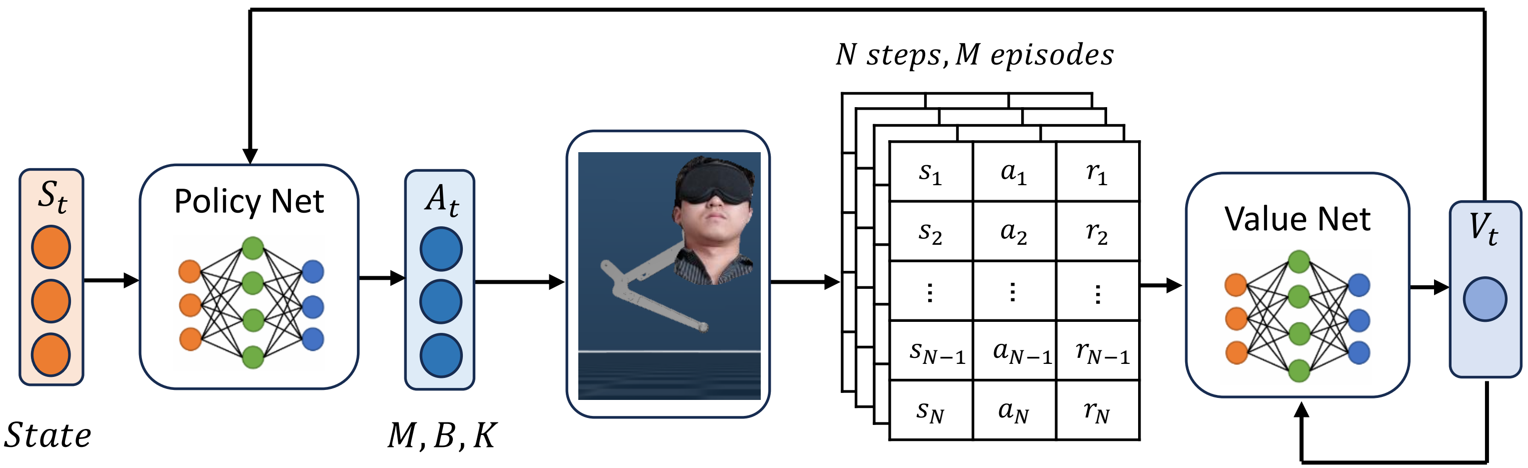
- Types of Reinforcement Learning



Since the goal is to estimate continuous M, B, K, DQN was excluded.

PPO was selected for its stochastic policy, which enables more diverse exploration.

- PPO Training Structure



$$State = \left[\theta, \dot{\theta}, {}^{B}\boldsymbol{p}_F, {}^{B}\boldsymbol{v}_F, GRF_z, Touch\ Flag\right]$$

To reflect the condition that the robot does not know its global position, the state includes the foot's relative position and velocity with respect to the body.

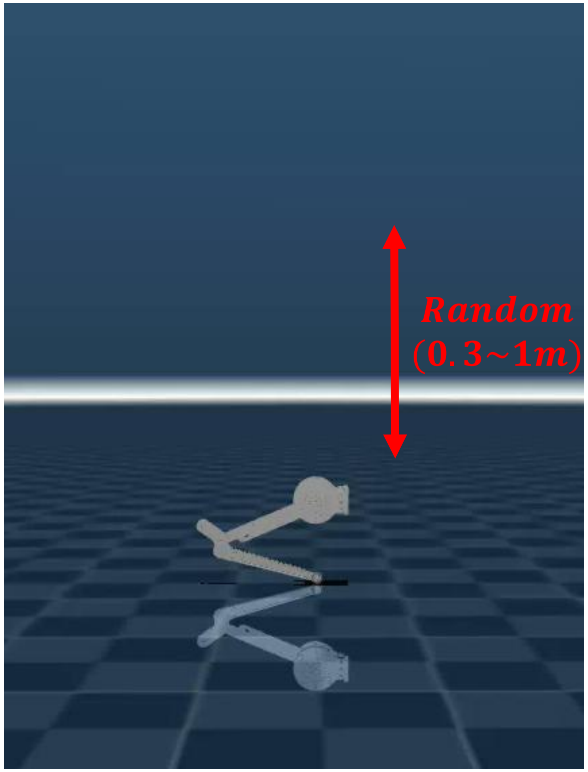
Ⅱ. Proposed Method

- Reward Function

Reward	Expression
Ground Reaction Force	$C_F(-\|F_z - W\|^2)$
Peak Ground Reaction Force	$C_{peak}(-\left\|\max\left(0, F_z - F_z^{peak}\right)\right\|^2)$
Foot Position	$C_p\left(-\left\|p^{des} - p\right\|^2\right)$
Foot Velocity	$C_v(-\|v\|^2)$
Angular Velocity	$C_\omega(-\|\omega\|^2)$
Joint Torque	$C_\tau(-\|\tau\|^2)$

The reward function was designed with a primary focus on impact mitigation, while also incorporating terms for energy efficiency and tracking performance.

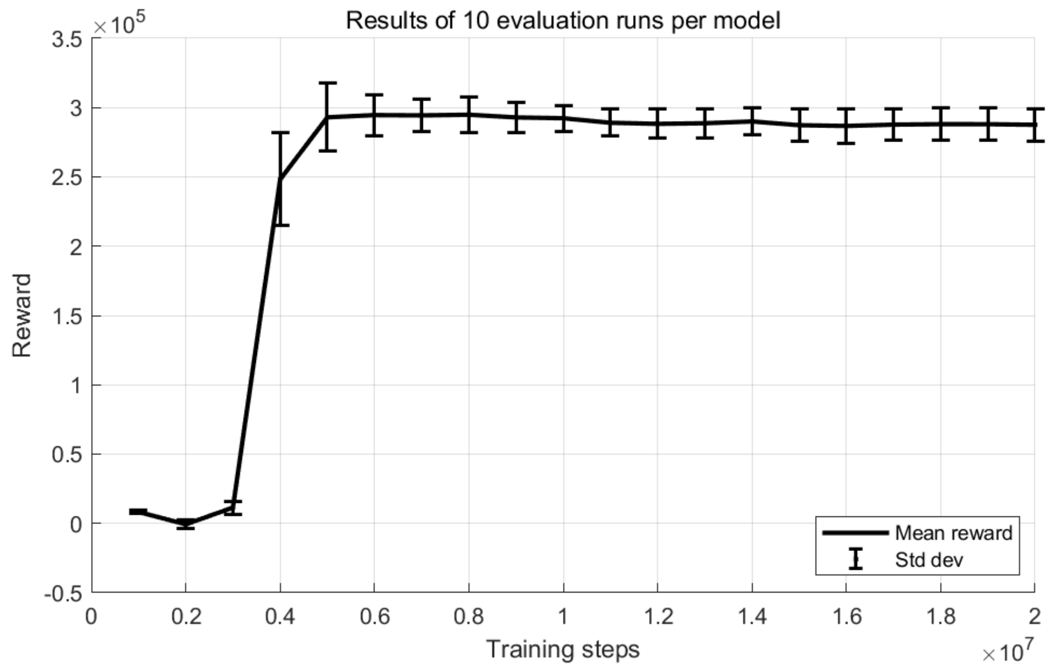
- Training Setup



Training was performed with random drop heights from 0.3 to 1.0 m, over a total of 20 million steps.

Ⅲ. Simulation

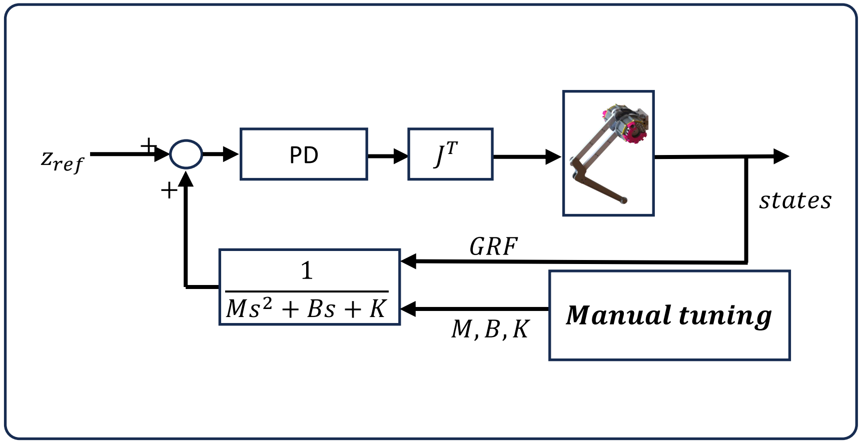
- Reward Curve



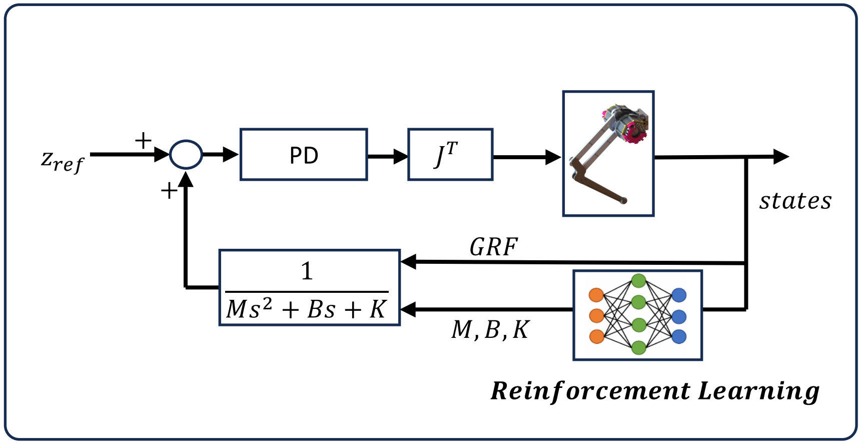
The reward curve shows stable convergence as training progresses.

- Comparative Analysis

Conventional Method



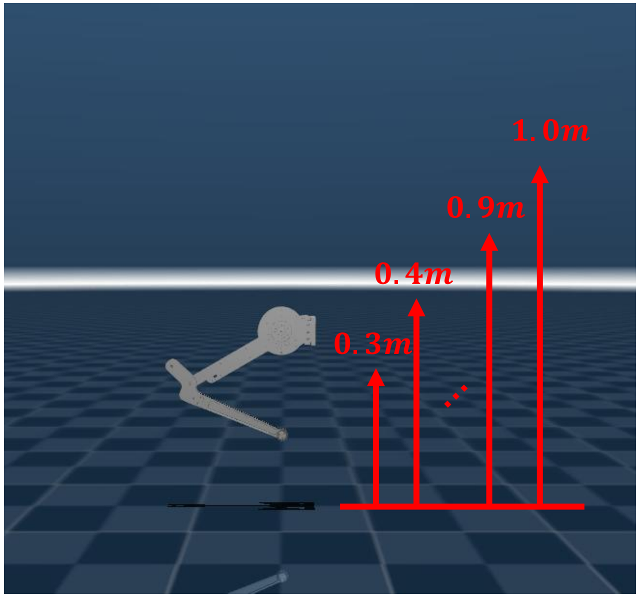
Proposed Method



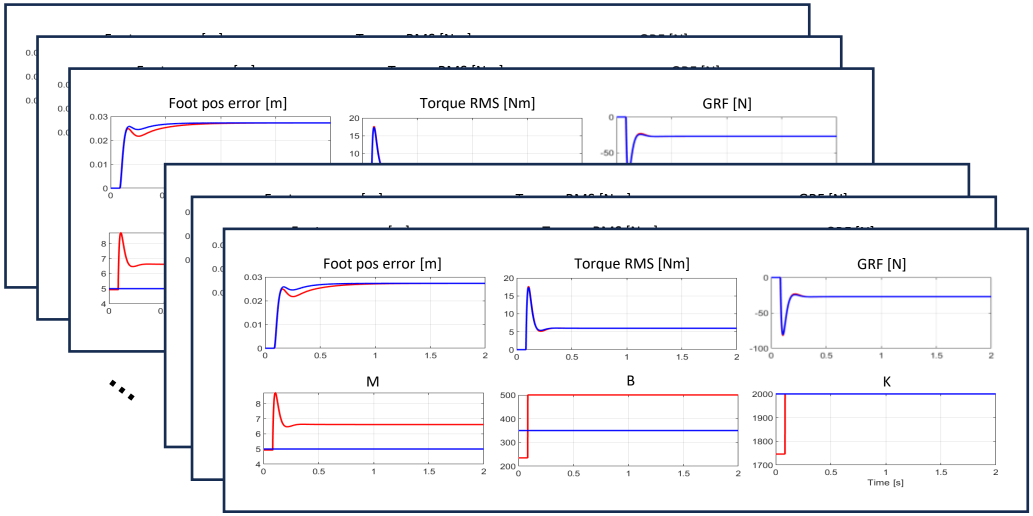
Performance Metrics	Peak GRF	Total Energy Usage	Torque RMS	Foot Position RMSE

To evaluate the algorithm's performance, I compared the proposed method with manual model-based control ($\underline{\text{Manually Tuned for 0.65m case } Fixed\ M, B, K = [5, 350, 2000]}$) using key performance metrics.

- Test Method



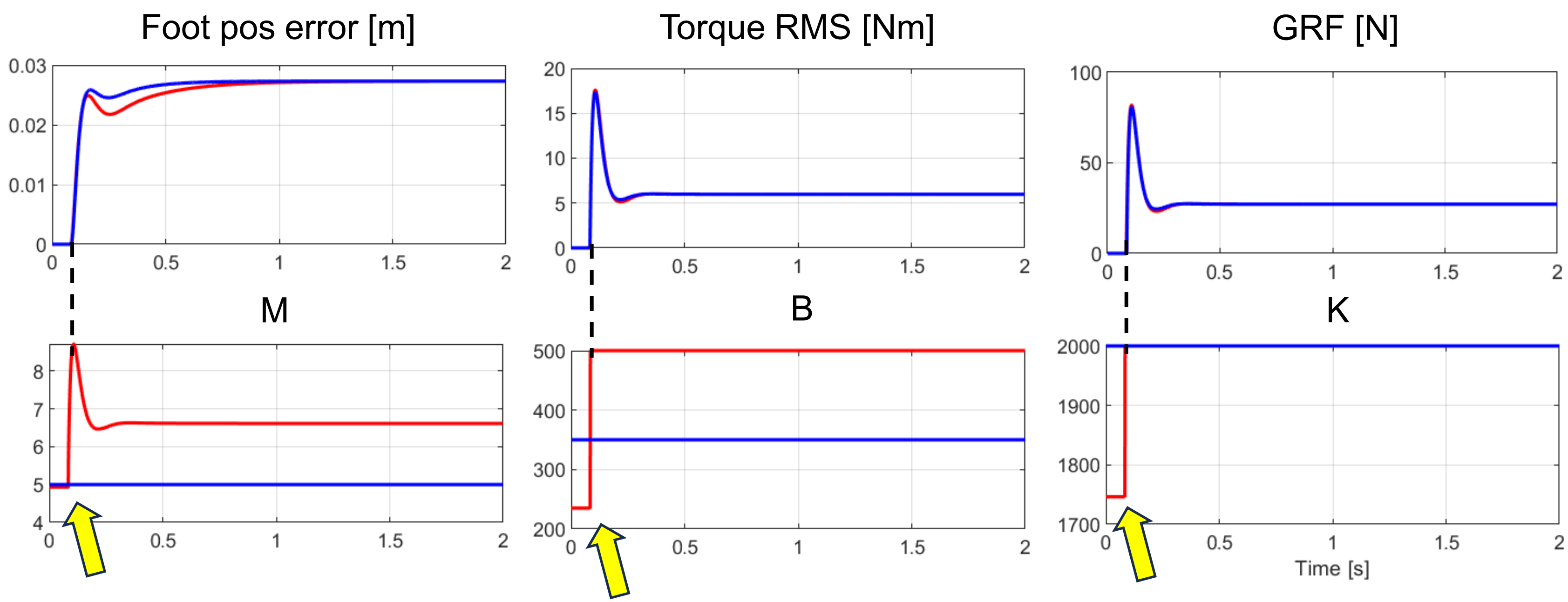
Conventional vs Proposed



Both conventional and proposed methods were evaluated from 0.3 m to 1.0 m in 0.1 m increments.

- Test Result

<0.3m-high landing>



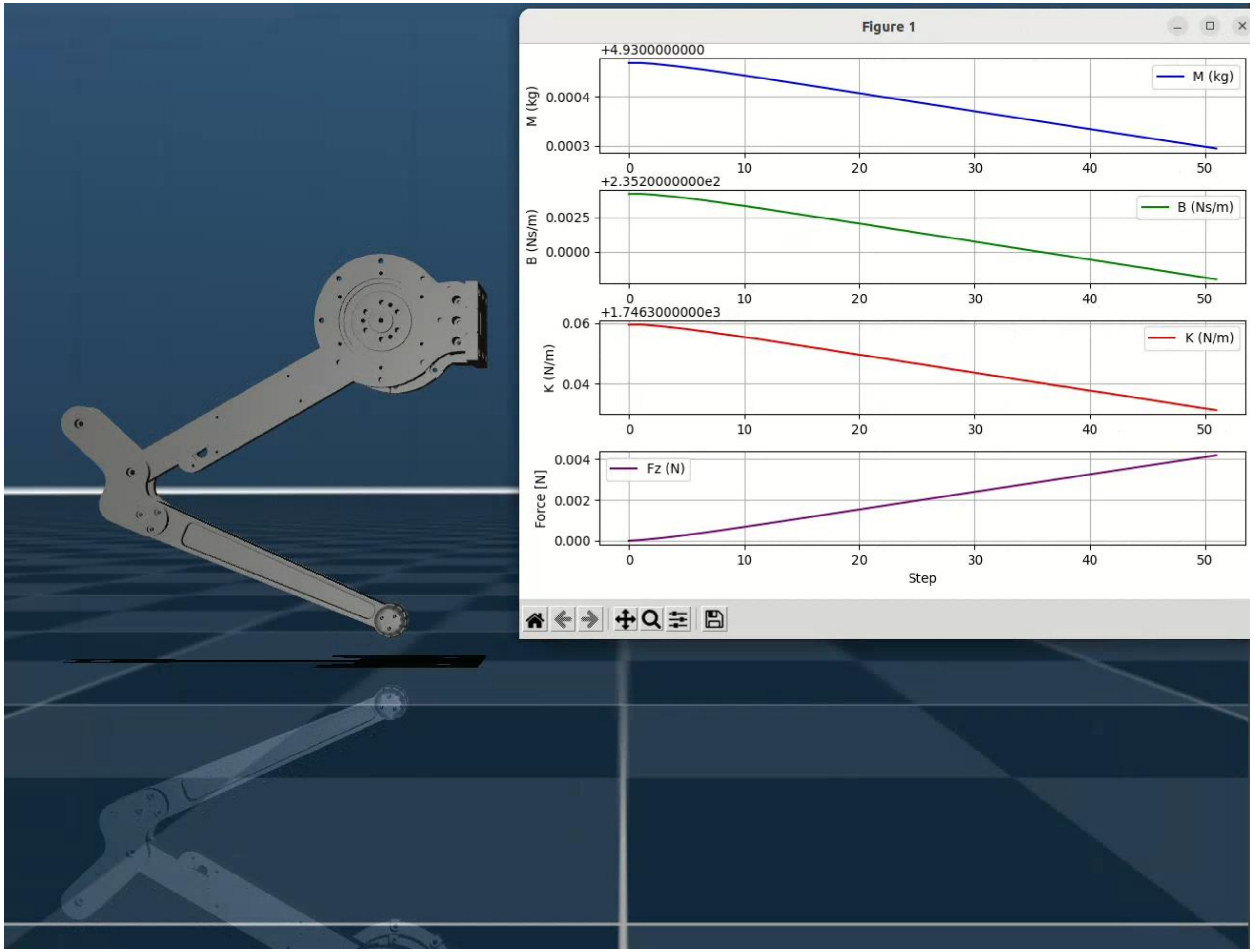
$$M \in [1, 10]$$
$$B \in [0, 500]$$
$$K \in [0, 2000]$$

RL: Proposed Method
Manual: Conventional Method

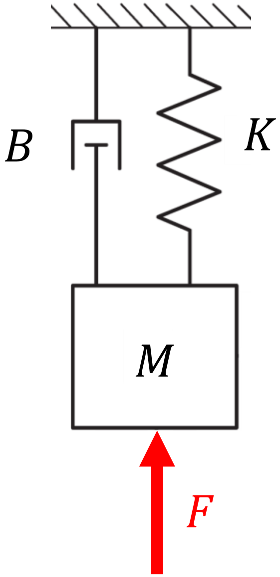
Temporal variations of foot position error, torque RMS, GRF, and varying gains for the 0.3m-high case over time are shown above.

- Test Result

<0.3m-high landing>



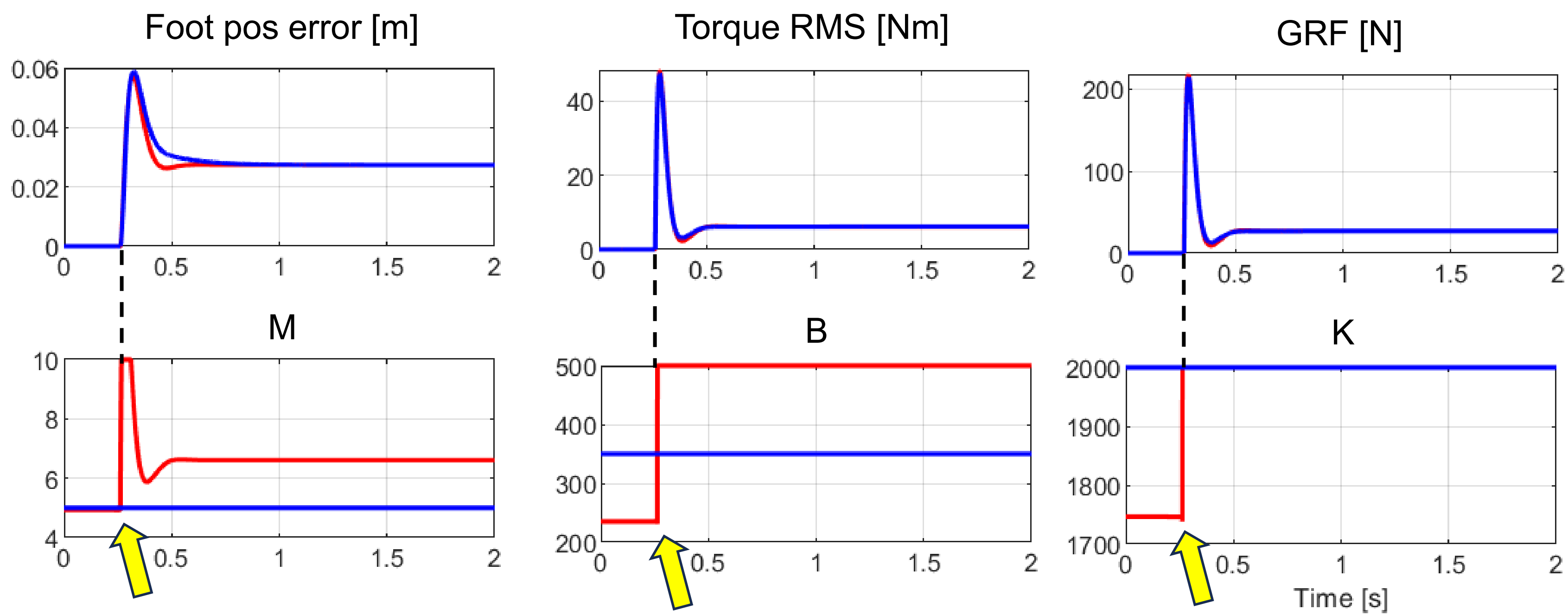
$M \in [1, 10]$
$B \in [0, 500]$
$K \in [0, 2000]$



Ⅲ. Simulation

- Test Result

<0.6m-high landing>



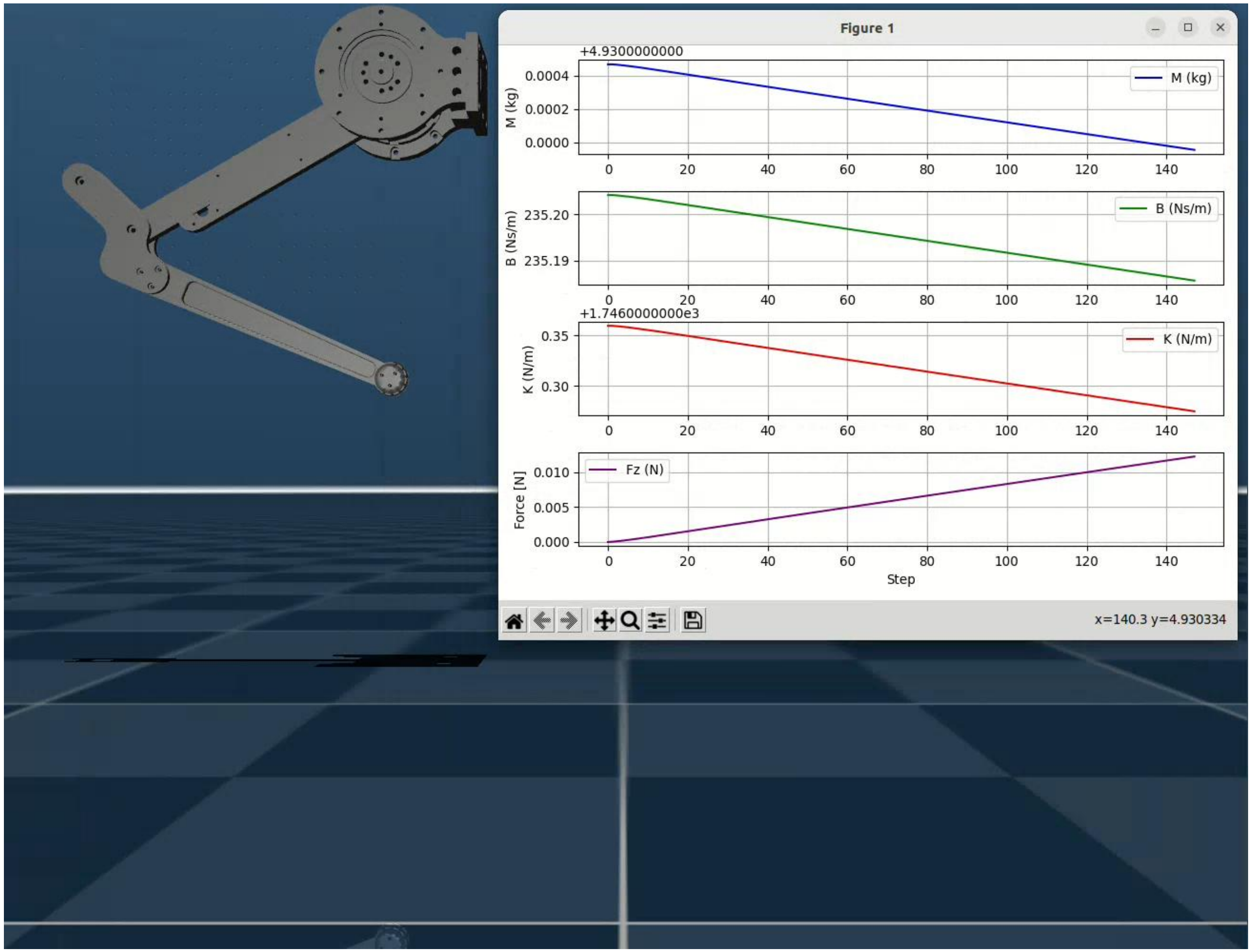
$$M \in [1, 10]$$
$$B \in [0, 500]$$
$$K \in [0, 2000]$$

RL: Proposed Method
Manual: Conventional Method

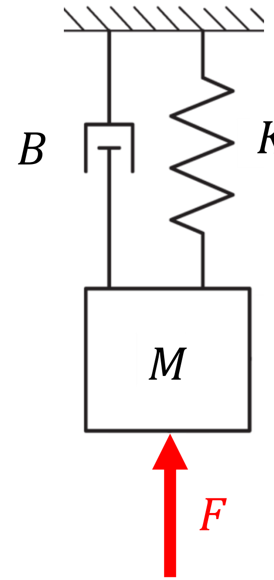
Temporal variations of foot position error, torque RMS, GRF, and varying gains for the 0.6m-high case over time are shown above.

- Test Result

<0.6m-high landing>

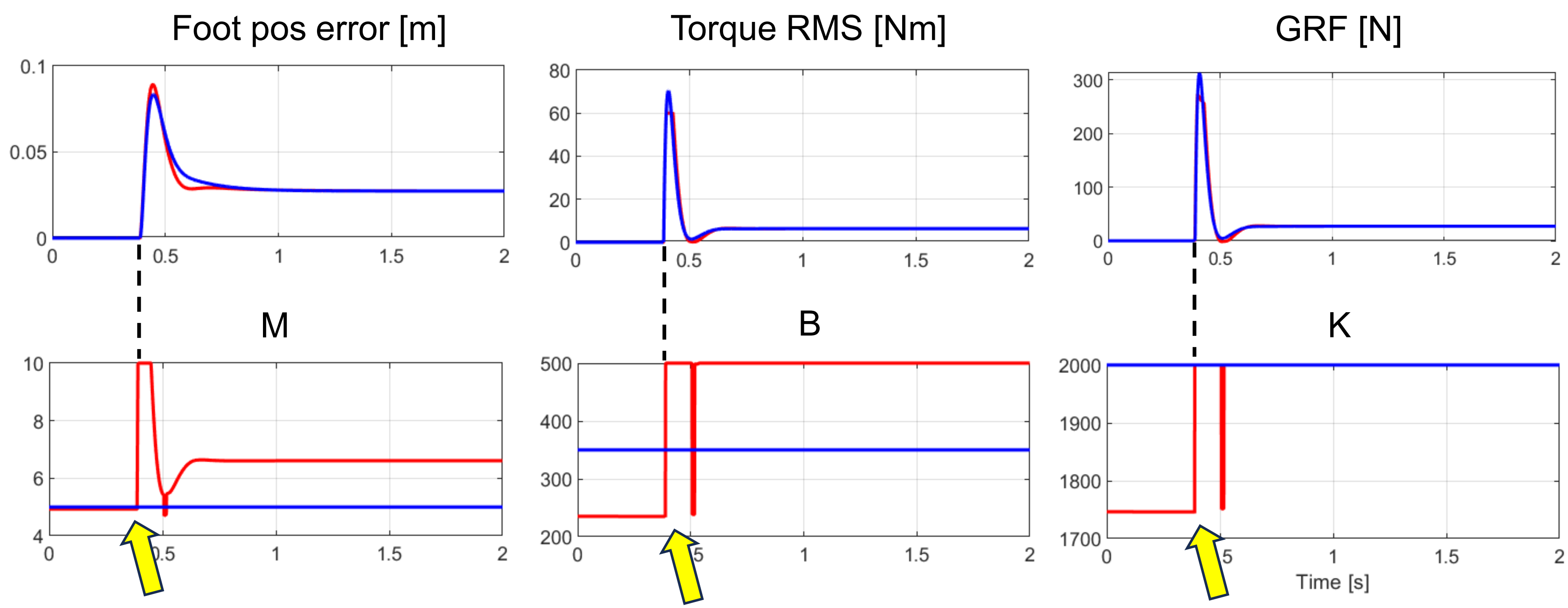


$$M \in [1, 10]$$
$$B \in [0, 500]$$
$$K \in [0, 2000]$$



- Test Result

<1.0m-high landing>



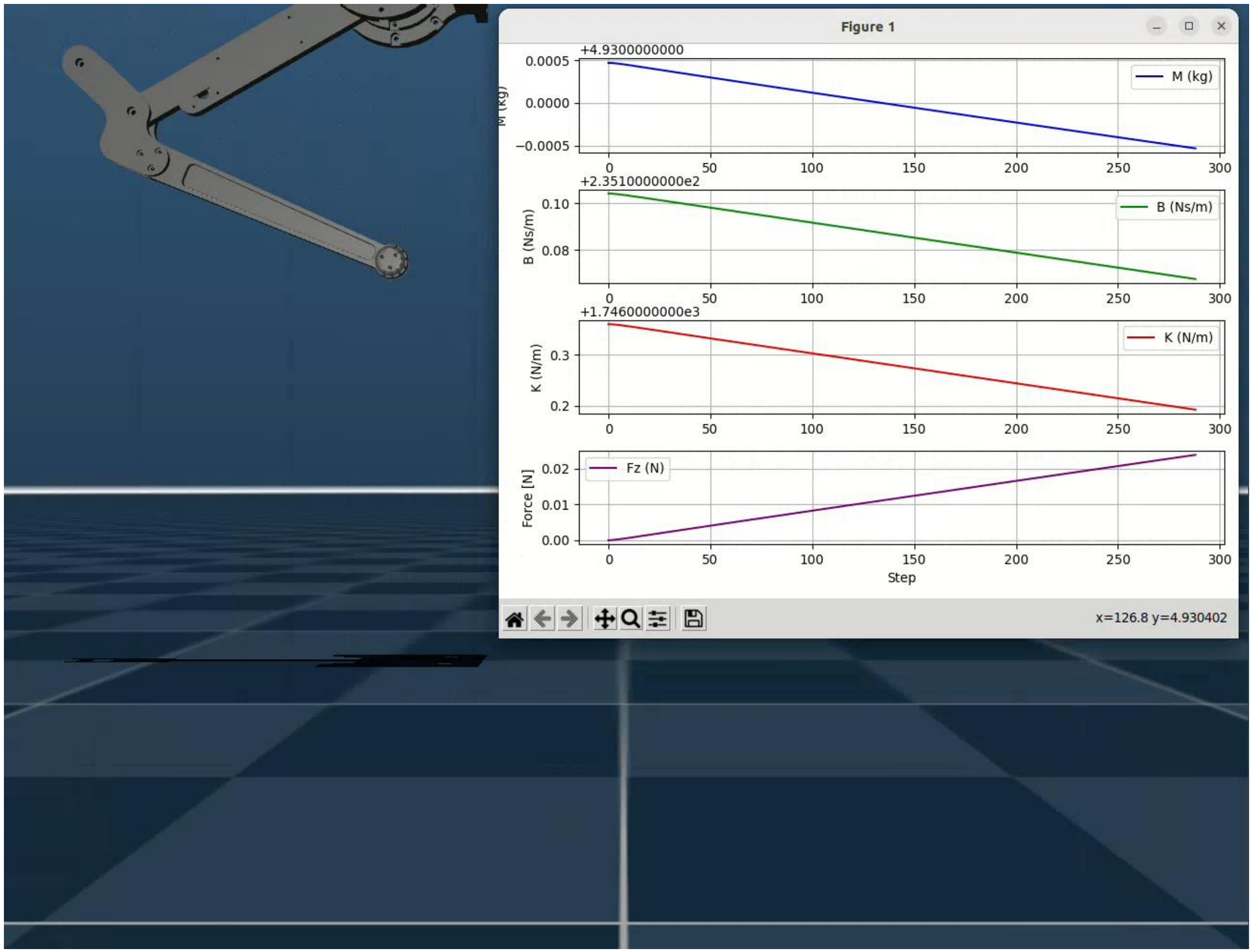
$M \in [1, 10]$
$B \in [0, 500]$
$K \in [0, 2000]$

RL: Proposed Method
Manual: Conventional Method

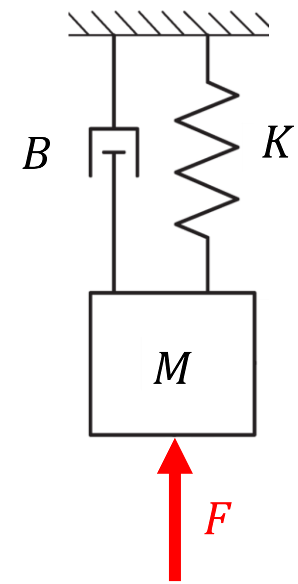
Temporal variations of foot position error, torque RMS, GRF, and varying gains for the 1.0m-high case over time are shown above.

- Test Result

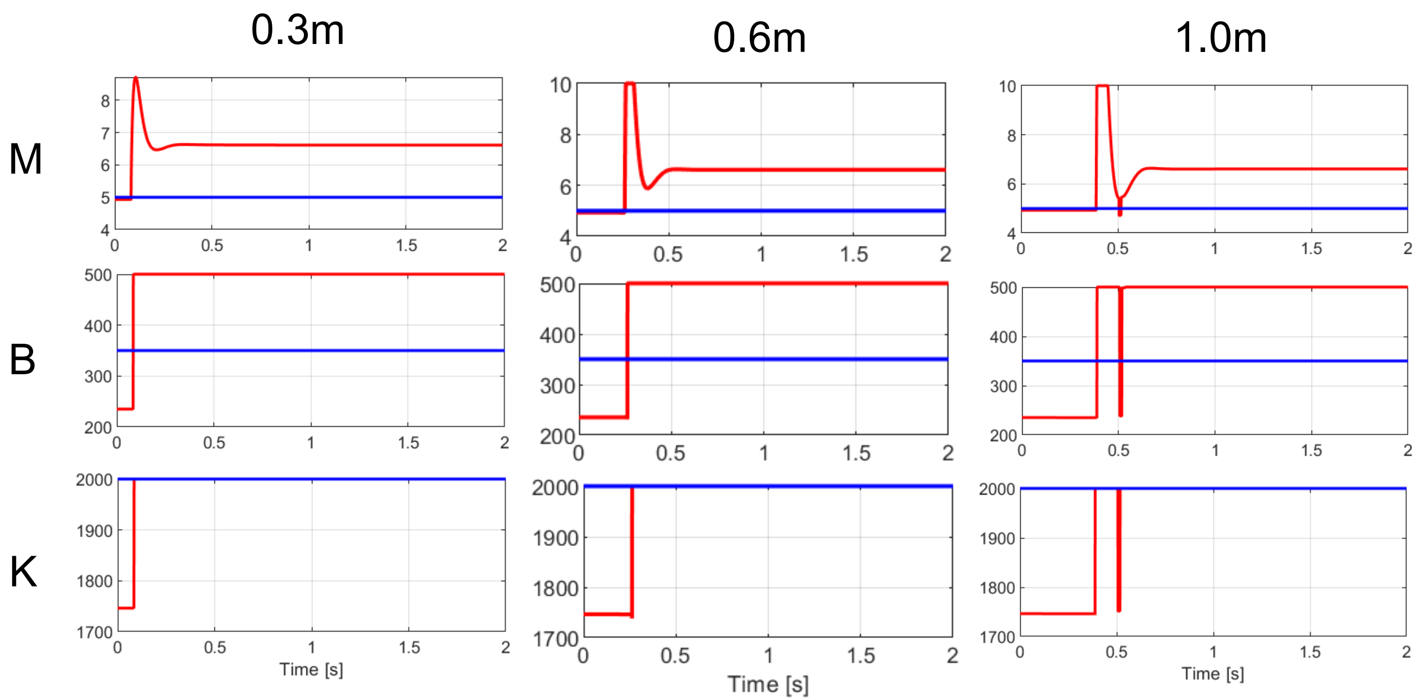
<1.0m-high landing>



$M \in [1, 10]$
$B \in [0, 500]$
$K \in [0, 2000]$



- Temporal Variations of gains

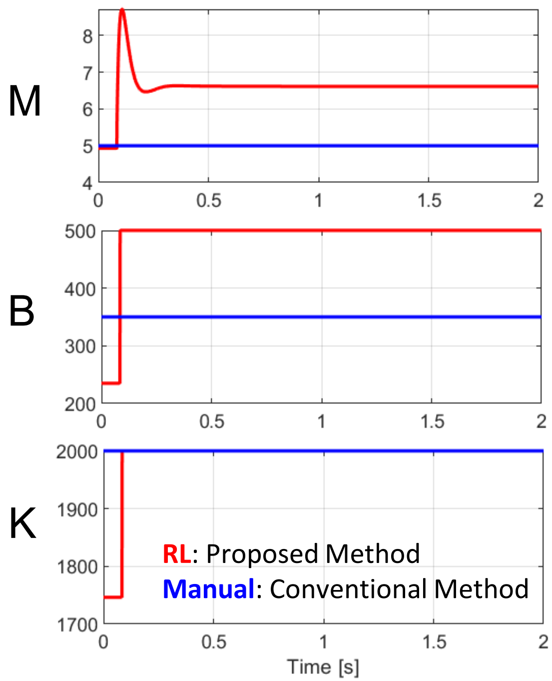


RL: Proposed Method
Manual: Conventional Method

Across all cases, M, B, and K exhibited consistent variation patterns.

Ⅲ. Simulation

- Reward Design Reflection



Reward	Expression	Weight
Ground Reaction Force	$C_F(-\|F_z - W\|^2)$	1.0
Peak Ground Reaction Force	$C_{peak}(-\left\|\max\left(0, F_z - F_z^{peak}\right)\right\|^2)$	3.0
Foot Position	$C_p\left(-\left\|p^{des} - p\right\|^2\right)$	2.0
Foot Velocity	$C_v(-\|v\|^2)$	0.05
Angular Velocity	$C_\omega(-\|\omega\|^2)$	0.05
Joint Torque	$C_\tau(-\|\tau\|^2)$	0.02

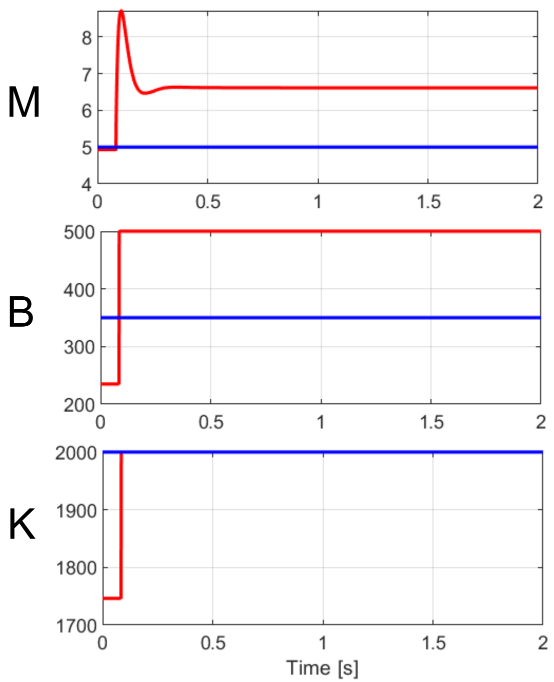
The increase of M, B, and K under impact is attributed to the foot position reward design and weight tuning.

- Performance Comparison

Height [m]	Types	Peak GRF [N]	Energy Usage [J]	Torque RMS [Nm]	Foot Pos RMSE [m]
0.3	Conventional	**81.02**	**1.8384**	**6.2734**	0.0261
	Proposed	81.79	1.9155	6.2837	**0.0256**
0.6	Conventional	**214.82**	11.1979	**8.4189**	0.0279
	Proposed	216.96	**11.0058**	8.5024	**0.0271**
1.0	Conventional	312.79	**22.9018**	**10.7480**	0.0301
	Proposed	**273.82**	23.1577	10.7543	**0.0298**

Evaluation of performance metrics shows that proposed method mainly focused on minimizing foot position tracking error.

- Discussion



Reward	Expression	Weig ht
Ground Reaction Force	$C_F(-\|F_z - W\|^2)$	1.0
Peak Ground Reaction Force	$C_{peak}(-\|\max(0, F_z - F_z^{peak})\|^2)$	3.0
Foot Position	$C_p\left(-\|p^{des} - p\|^2\right)$	2.0
Foot Velocity	$C_v(-\|v\|^2)$	0.05
Angular Velocity	$C_\omega(-\|\omega\|^2)$	0.05
Joint Torque	$C_\tau(-\|\tau\|^2)$	0.02

The simulation results differed from my expectations, but the time-varying M, B, and K provided intuitive insights into the causes.

With further tuning toward impact mitigation under this system, I believe more ideal outcomes can be achieved.

IV. Conclusion

- Future Work

Future work involves verifying online applicability on real robots and training and testing in more diverse simulation environment for adaptability.