ETH zürich

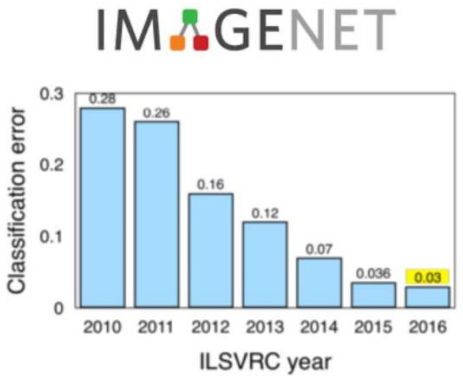
Prescriptive Maintenance with Deep Reinforcement Learning

Yuan Tian

Chair of Intelligent Maintenance Systems

Success of Data-driven Methods for General Tasks

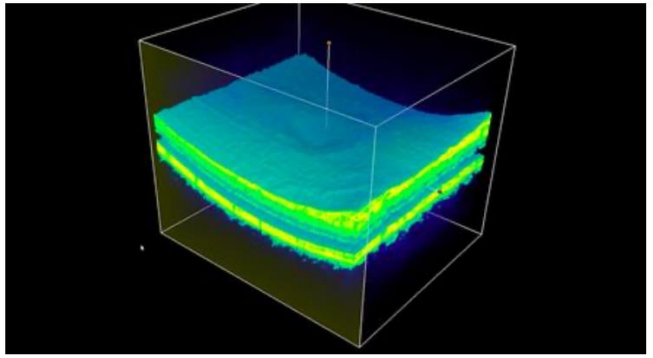
Image classification



Detection and segmentation

Machine Translation

Eye disease diagnosis

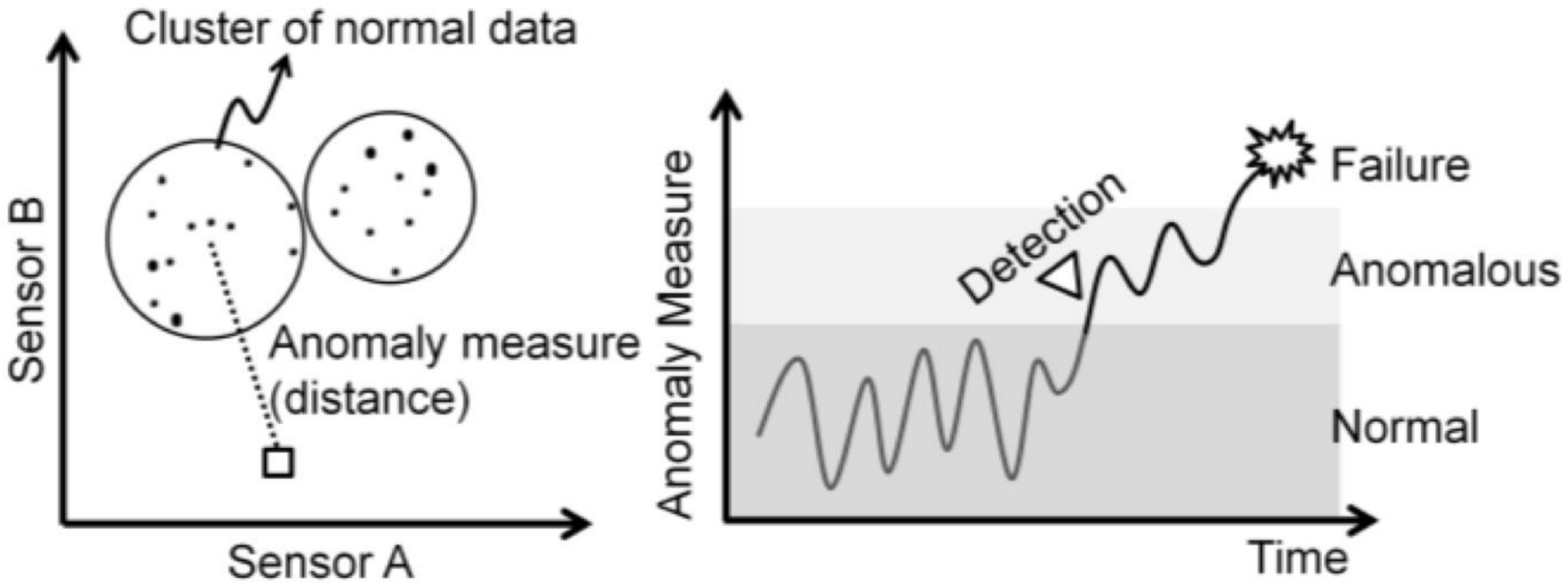


Success of Data-driven Methods on Maintenance Applications

- Detection, Diagnostics, RUL Prediction, etc.

Success of Data-driven Methods on Maintenance Applications

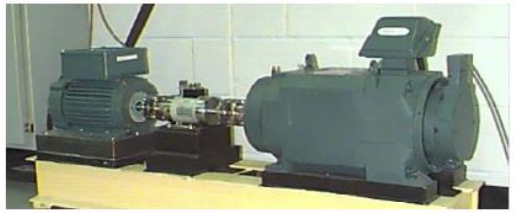
- Detection, Diagnostics, RUL Prediction, etc.



Success of Data-driven Methods on Maintenance Applications

- Detection, Diagnostics, RUL Prediction, etc.

- From one operating condition to another

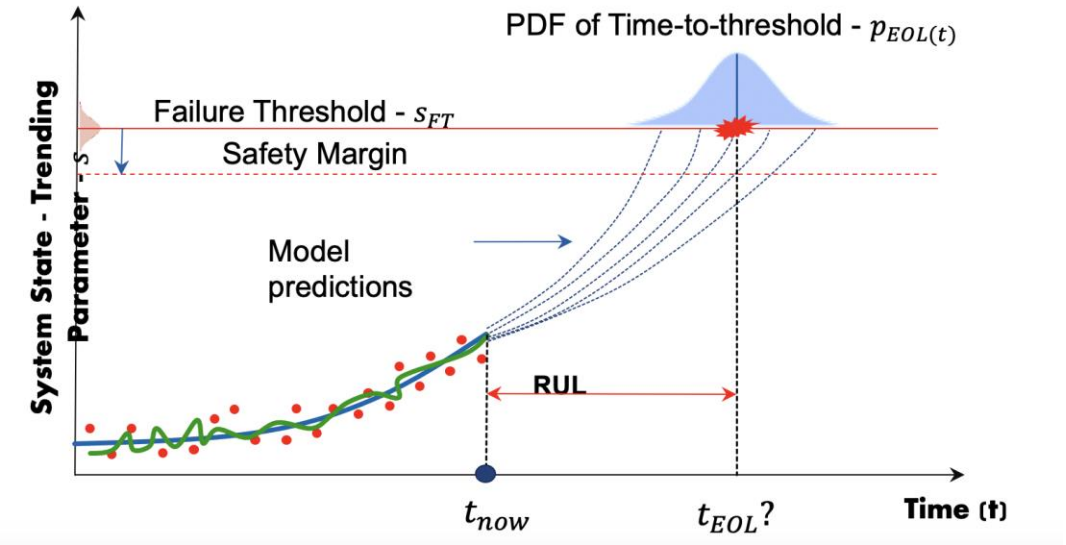


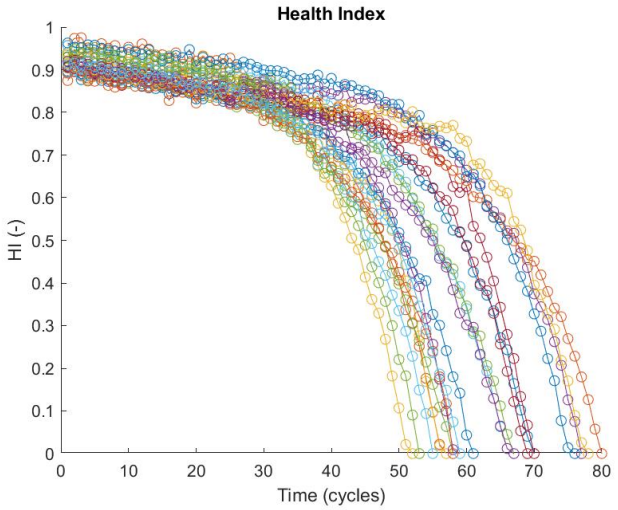
- From one machine to a fleet

- From one manufacturer to another

Success of Data-driven Methods on Maintenance Applications

- Detection, Diagnostics, RUL Prediction, etc.





Type of Problem Addressed

Predict the remaining useful life

Anticipate the failure

Reduce the impact of the failure

Determine the optimal point in time for maintenance intervention

Type of Problem Addressed

Predict the remaining useful life

Anticipate the failure

Reduce the impact of the failure

Determine the optimal point in time for maintenance intervention

What can we do to prolong the remaining useful life?

How can we proactively adjust the operating conditions?

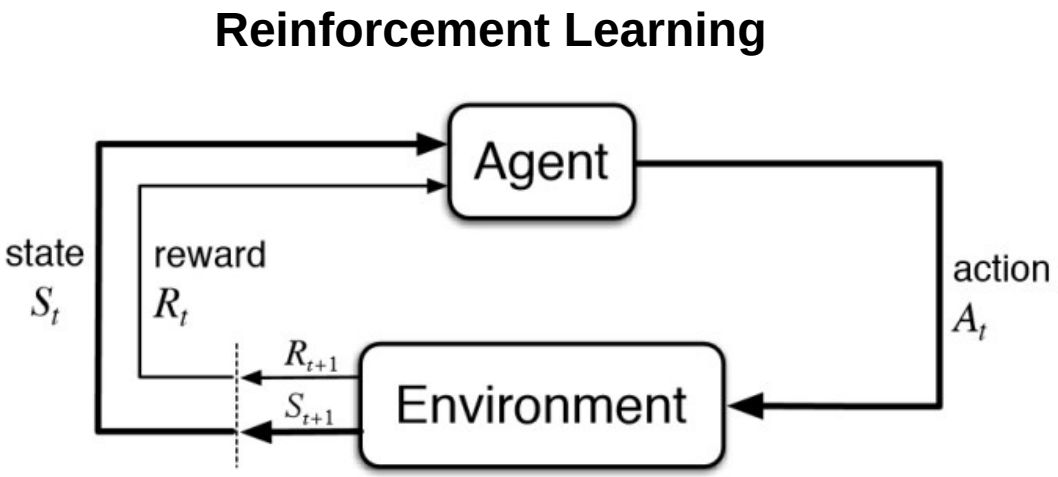
How can we control the process parameters?

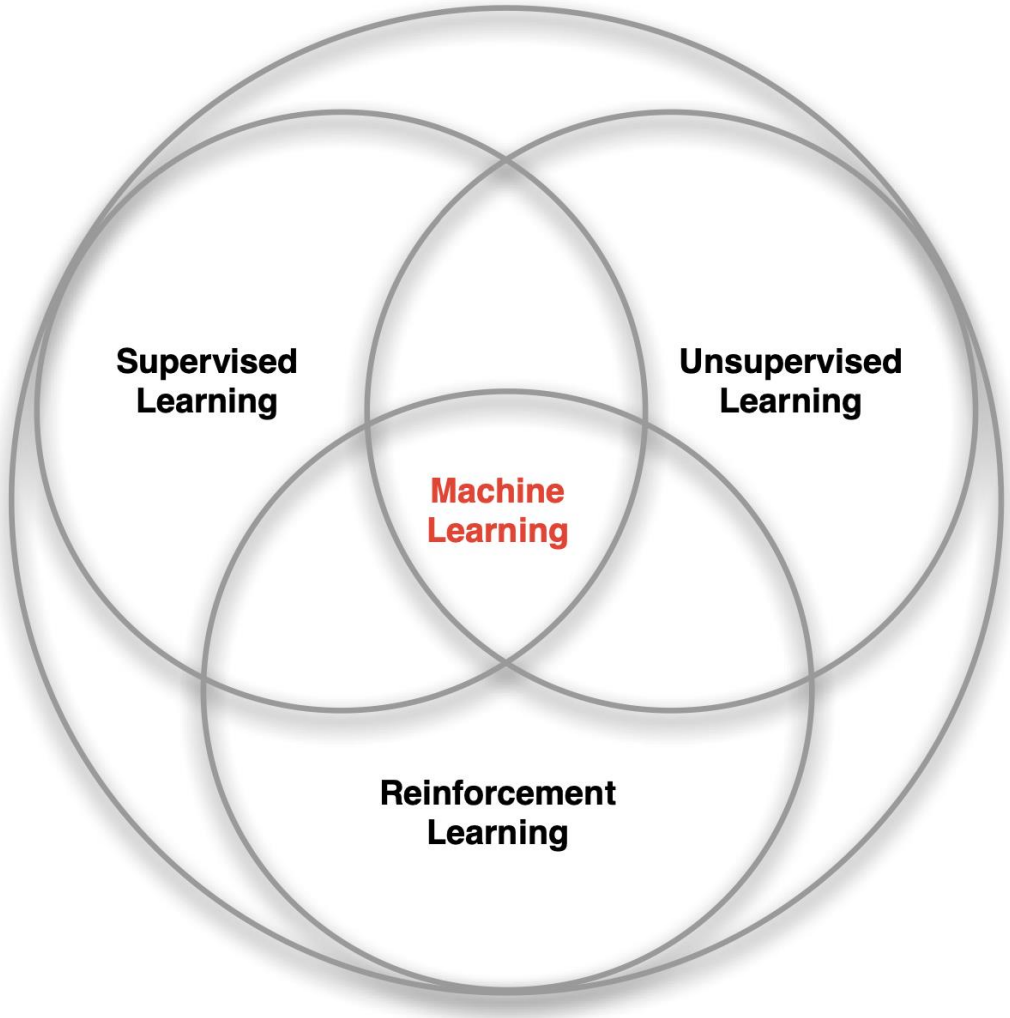
Type of Problem Addressed

Type of Problem Addressed

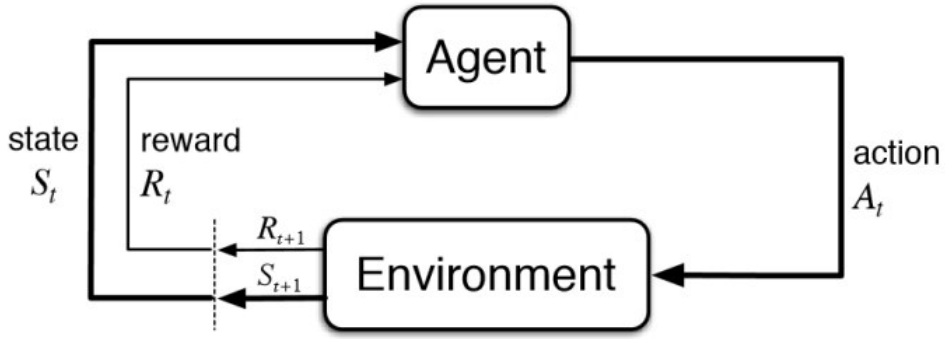
- Sequential decision-making

Reinforcement Learning





Reinforcement Learning



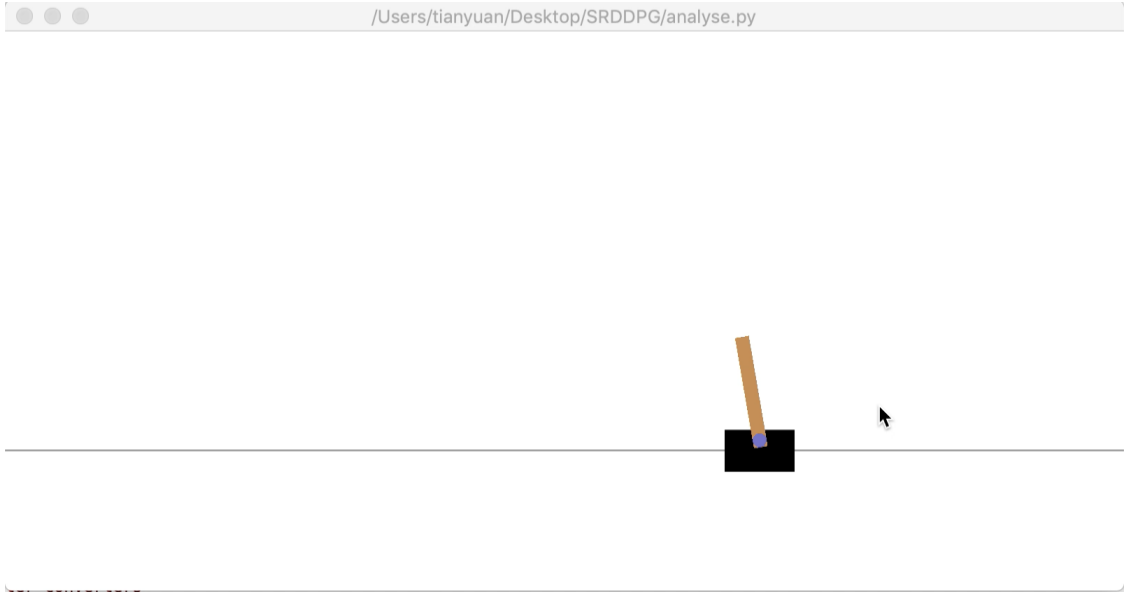
What makes reinforcement learning different from other machine learning paradigms?

- There is no supervisor, only a reward signal
- Feedback may be delayed, not instantaneous
- Time really matters (sequential, non i.i.d data)
- Agent's actions affect the subsequent data it receives

Reinforcement Learning

What makes reinforcement learning different from other machine learning paradigms?

- There is no supervisor, only a reward signal
- Feedback may be delayed, not instantaneous
- Time really matters (sequential, non i.i.d data)
- Agent's actions affect the subsequent data it receives

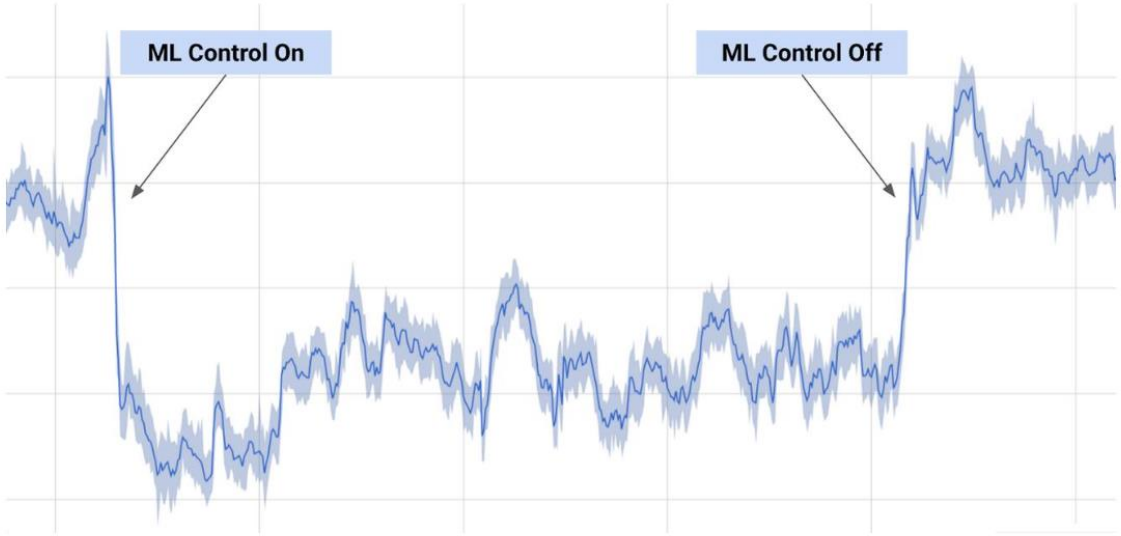


Current Success

- **Deepmind AI Reduces Google Data Centre Cooling Bill by 40%**
  - Successful system management
- **OpenAI Dota 2 AI defeated the world's top professional team**
  - Capturing the messiness and continuous nature of the real world
  - Ability to solve long time horizons, partially-observed state, high-dimensional continuous action/state space problems
  - End-to-End
- **ETH Anymal legged robot**
  - Real-world control task
- **Deepmind Alpha Zero**
  - Without human knowledge
- **AutoML**
  - Automated optimization

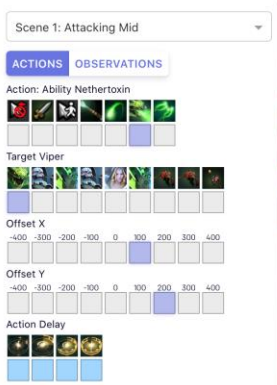
Current Success

- **Deepmind AI Reduces Google Data Centre Cooling Bill by 40%**
  - Successful system management
- **OpenAI Dota 2 AI defeated the world's top professional team**
  - Capturing the messiness and continuous nature of the real world
  - Ability to solve long time horizons, partially-observed state, high-dimensional continuous action/state space problems
  - End-to-End
- **ETH Anymal legged robot**
  - Real-world control task
- **Deepmind Alpha Zero**
  - Without human knowledge
- **AutoML**



Current Success

- **Deepmind AI Reduces Google Data Centre Cooling Bill by 40%**
  - Successful system management
- **OpenAI Dota 2 AI defeated the world's top professional team**
  - Capturing the messiness and continuous nature of the real world
  - Ability to solve long time horizons, partially-observed state, high-dimensional continuous action/state space problems
  - End-to-End
- **ETH Anymal legged robot**
  - Real-world control task
- **Deepmind Alpha Zero**
  - Without human knowledge
- **AutoML**

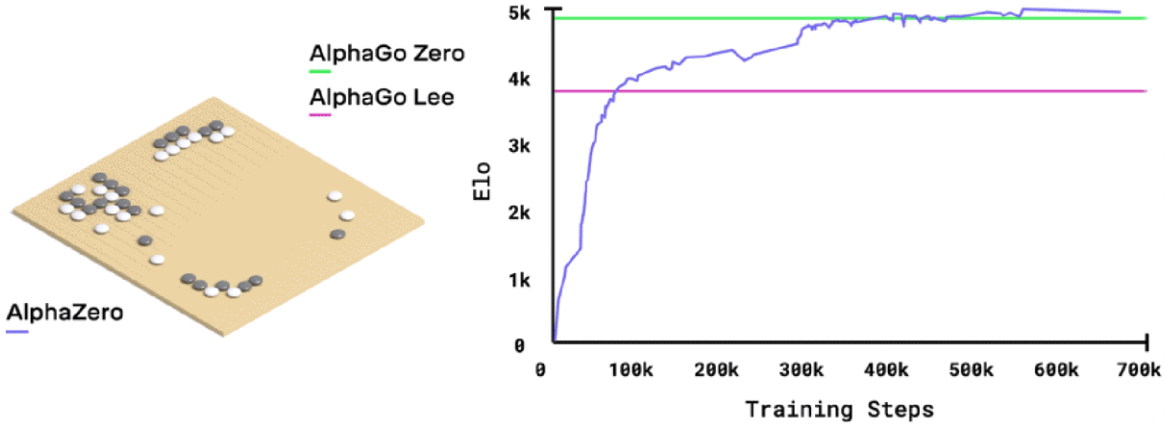


Current Success

- **Deepmind AI Reduces Google Data Centre Cooling Bill by 40%**
  - Successful system management
- **OpenAI Dota 2 AI defeated the world's top professional team**
  - Capturing the messiness and continuous nature of the real world
  - Ability to solve long time horizons, partially-observed state, high-dimensional continuous action/state space problems
  - End-to-End
- **ETH Anymal legged robot**
  - Real-world control task
- **Deepmind Alpha Zero**
  - Without human knowledge
- **AutoML**

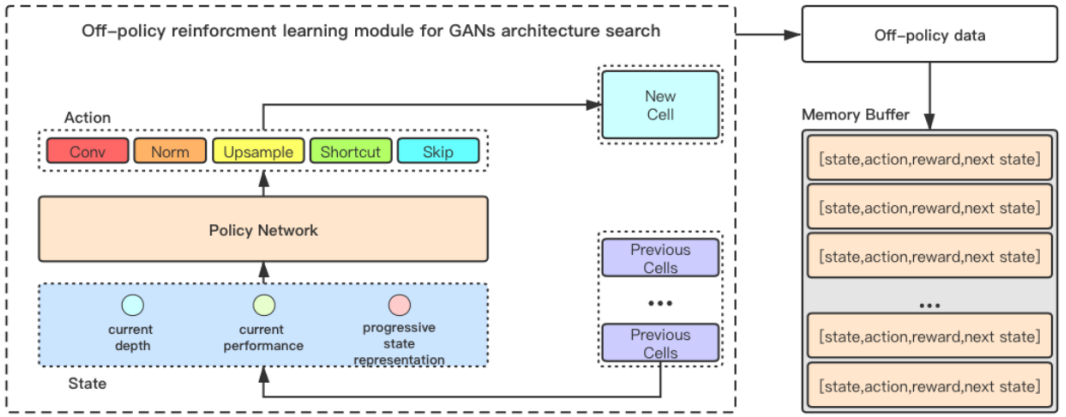
Current Success

- **Deepmind AI Reduces Google Data Centre Cooling Bill by 40%**
  - Successful system management
- **OpenAI Dota 2 AI defeated the world's top professional team**
  - Capturing the messiness and continuous nature of the real world
  - Ability to solve long time horizons, partially-observed state, high-dimensional continuous action/state space problems
  - End-to-End
- **ETH Anymal legged robot**
  - Real-world control task
- **Deepmind Alpha Zero**
  - Without human knowledge
- **AutoML**

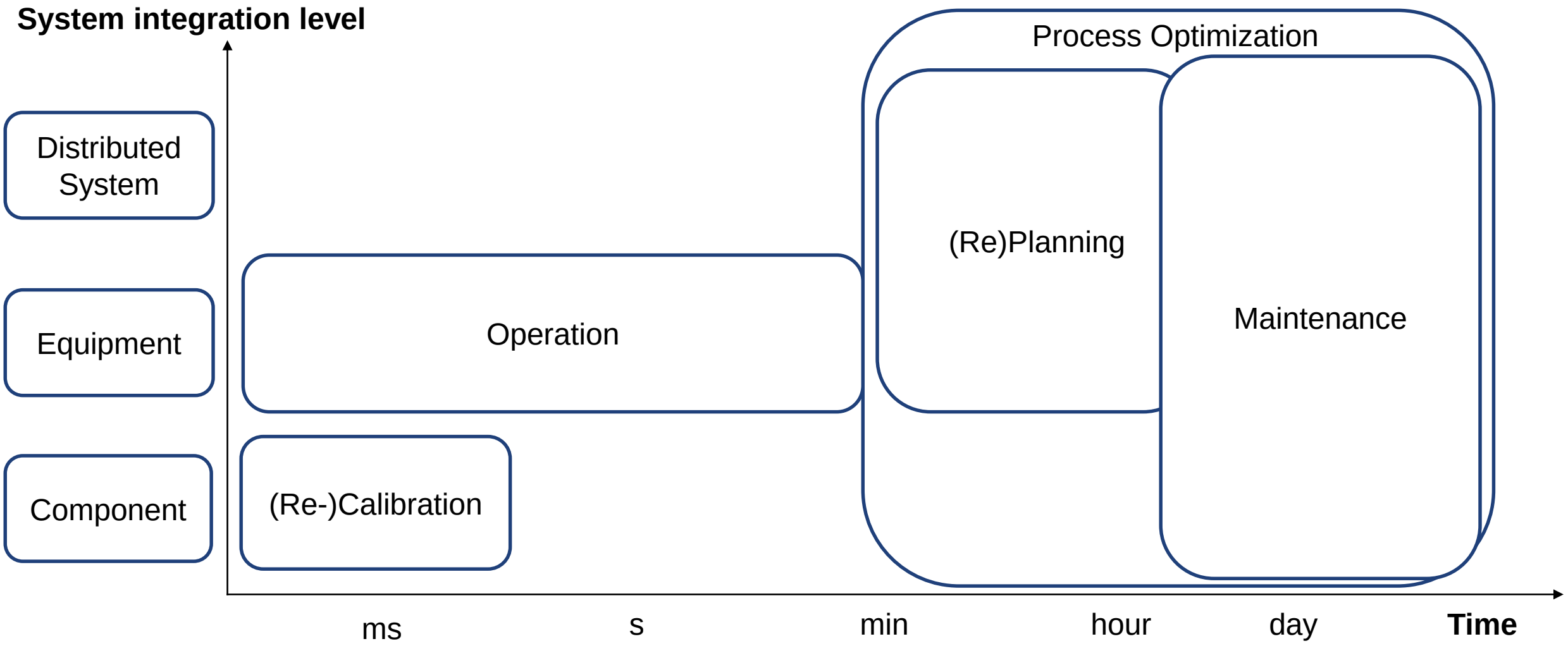


Current Success

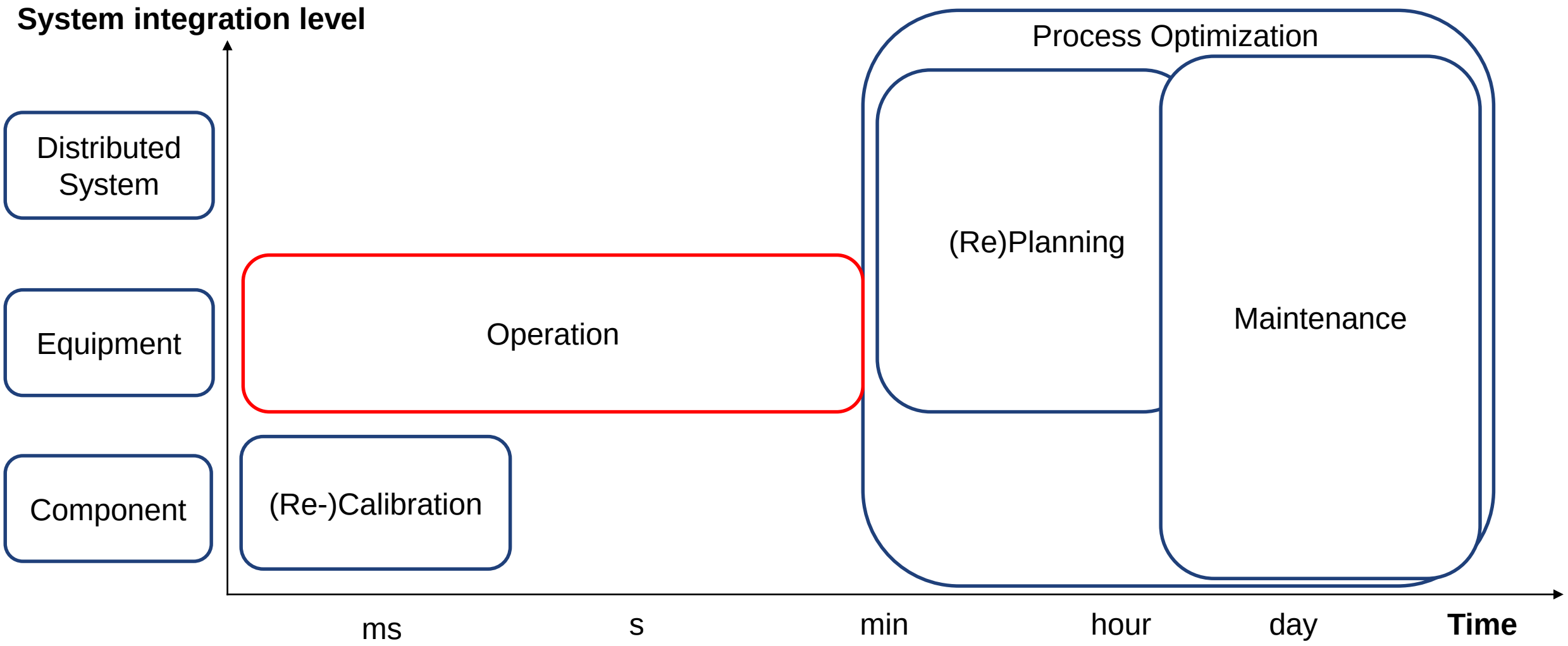
- **Deepmind AI Reduces Google Data Centre Cooling Bill by 40%**
  - Successful system management
- **OpenAI Dota 2 AI defeated the world's top professional team**
  - Capturing the messiness and continuous nature of the real world
  - Ability to solve long time horizons, partially-observed state, high-dimensional continuous action/state space problems
  - End-to-End
- **ETH Anymal legged robot**
  - Real-world control task
- **Deepmind Alpha Zero**
  - Without human knowledge
- **AutoML**



Prescriptive Maintenance



Prescriptive Maintenance

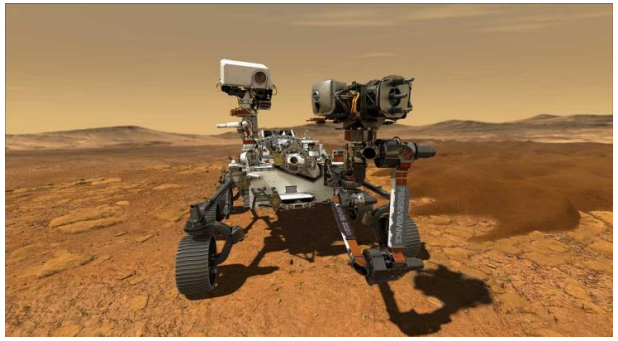


Case study I : Power Allocation

Electric Vehicles

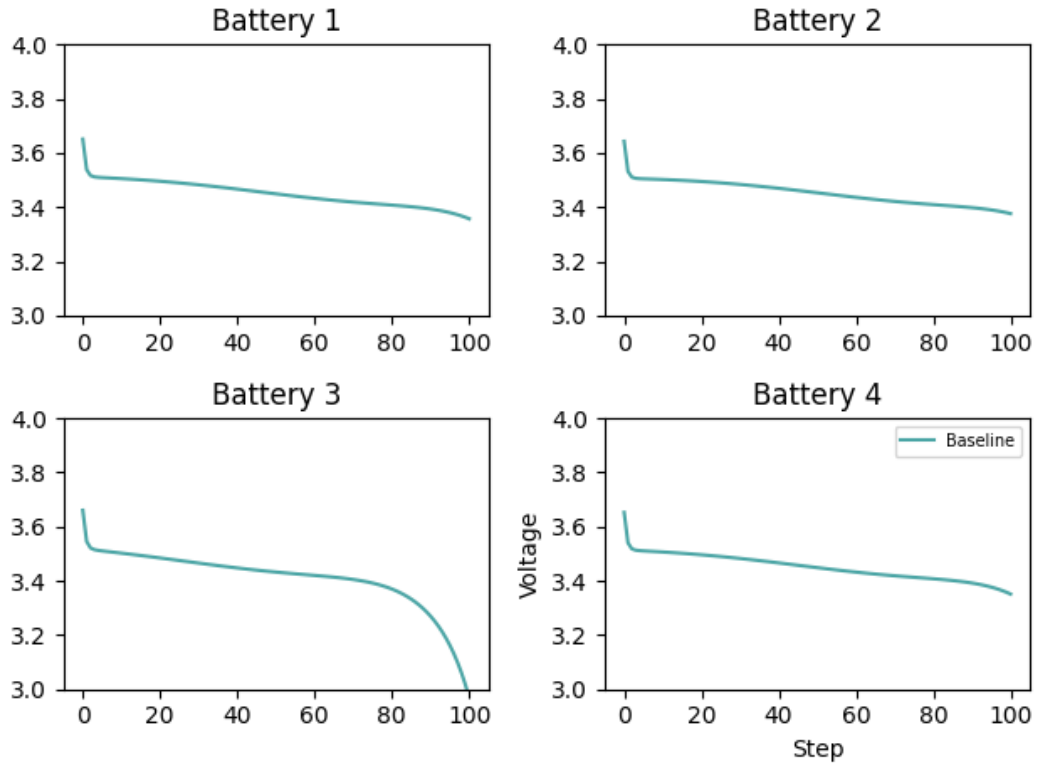
Personal Mobility

Space Exploration



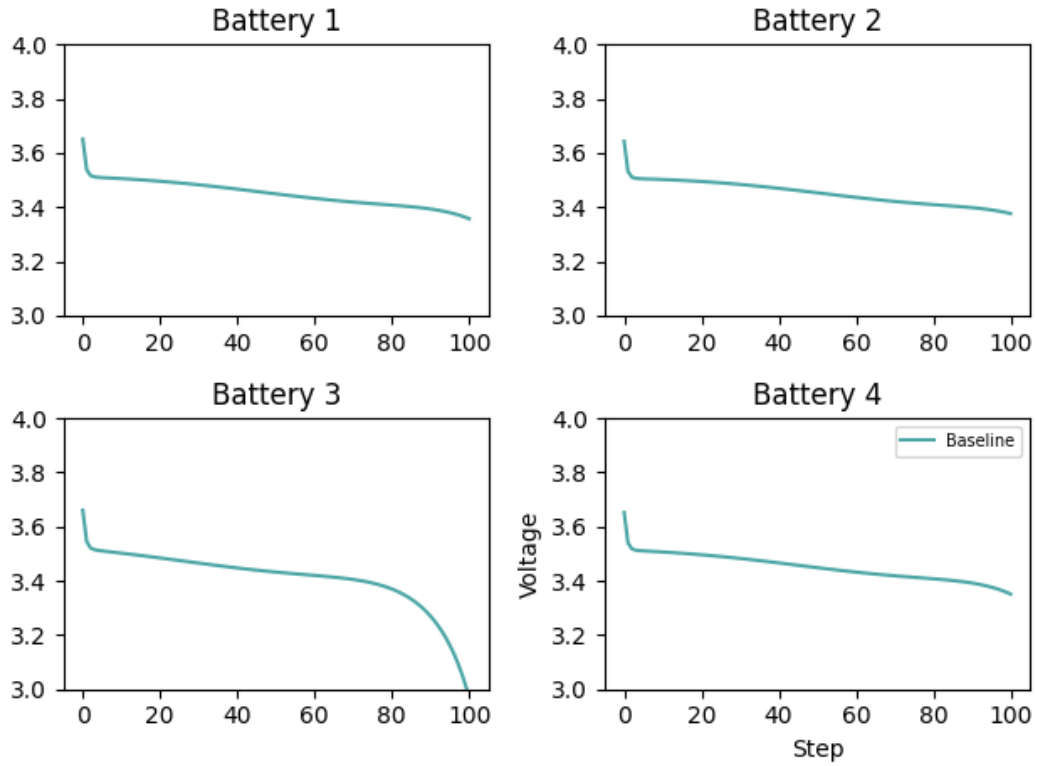
Electronic Devices

Challenge

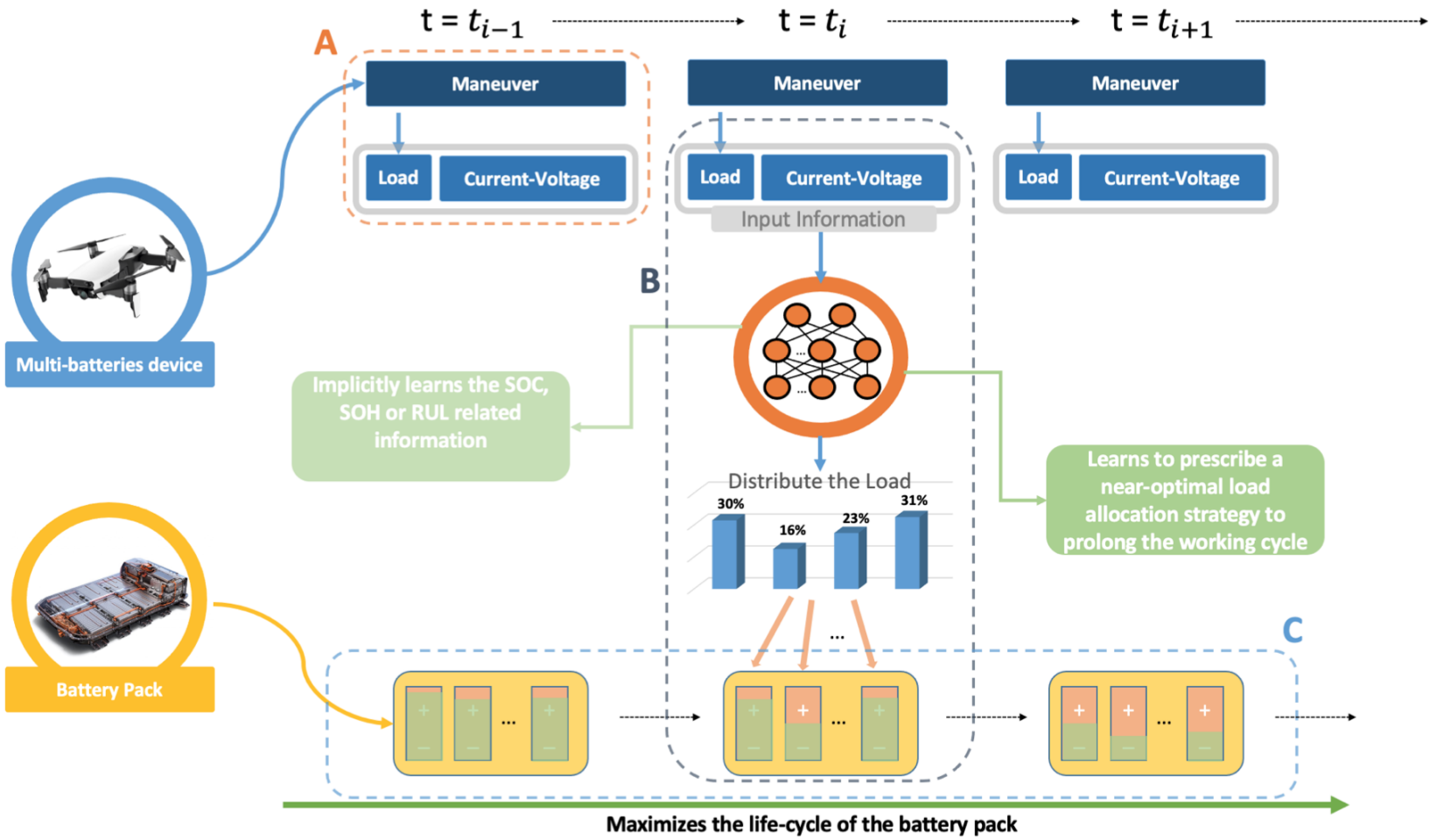

Battery 1
Battery 2
Battery 3
Battery 4
Baseline
Voltage
Step
4.0
3.8
3.6
3.4
3.2
3.0
0
20
40
60
80
100

Challenge

- **How to smartly distribute the load demand in order to prolong the working cycle?**

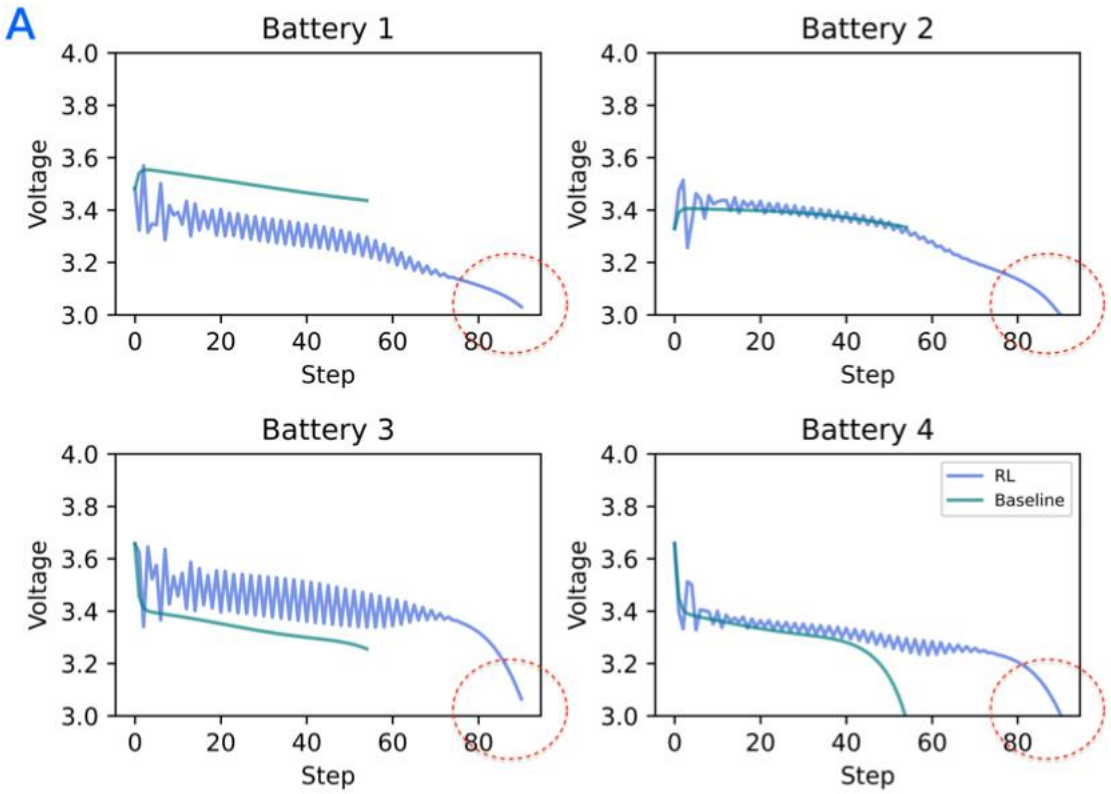


Proposed Framework



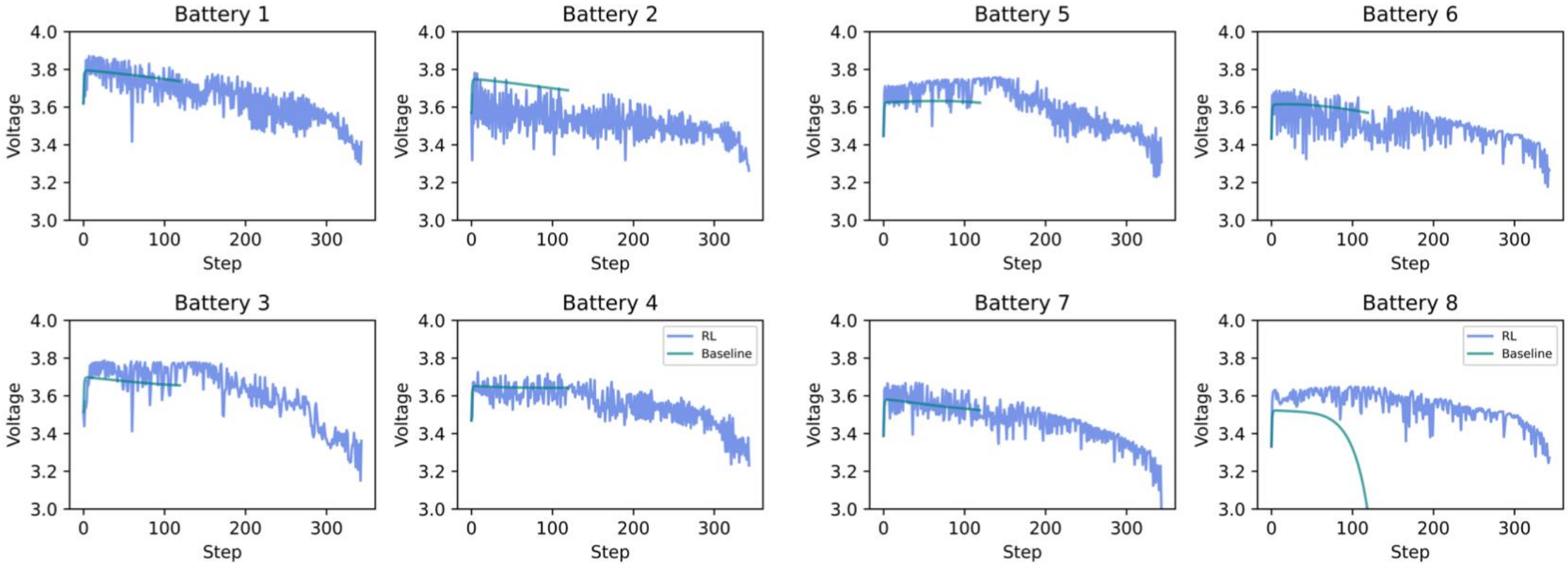
Results

On Average, the working cycle of the deployed system prolongs **15.2%**



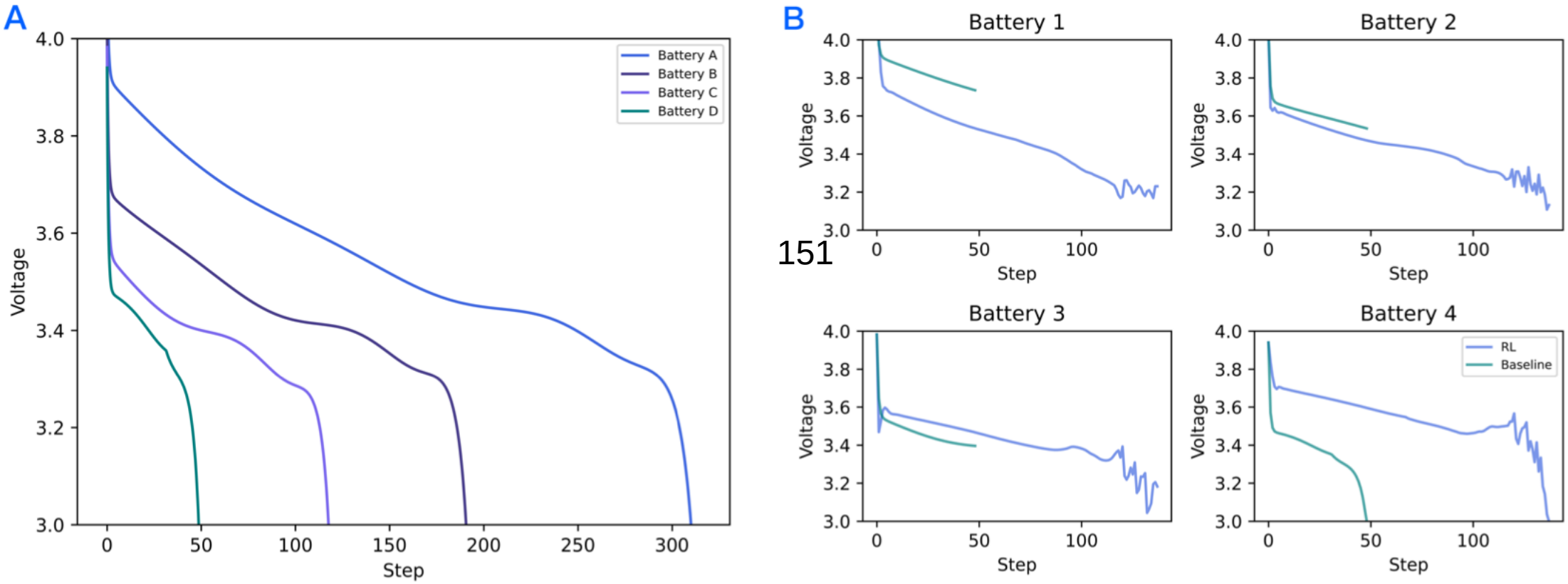
Results

On Average, the working cycle of the deployed system prolongs **31.9%**



Results

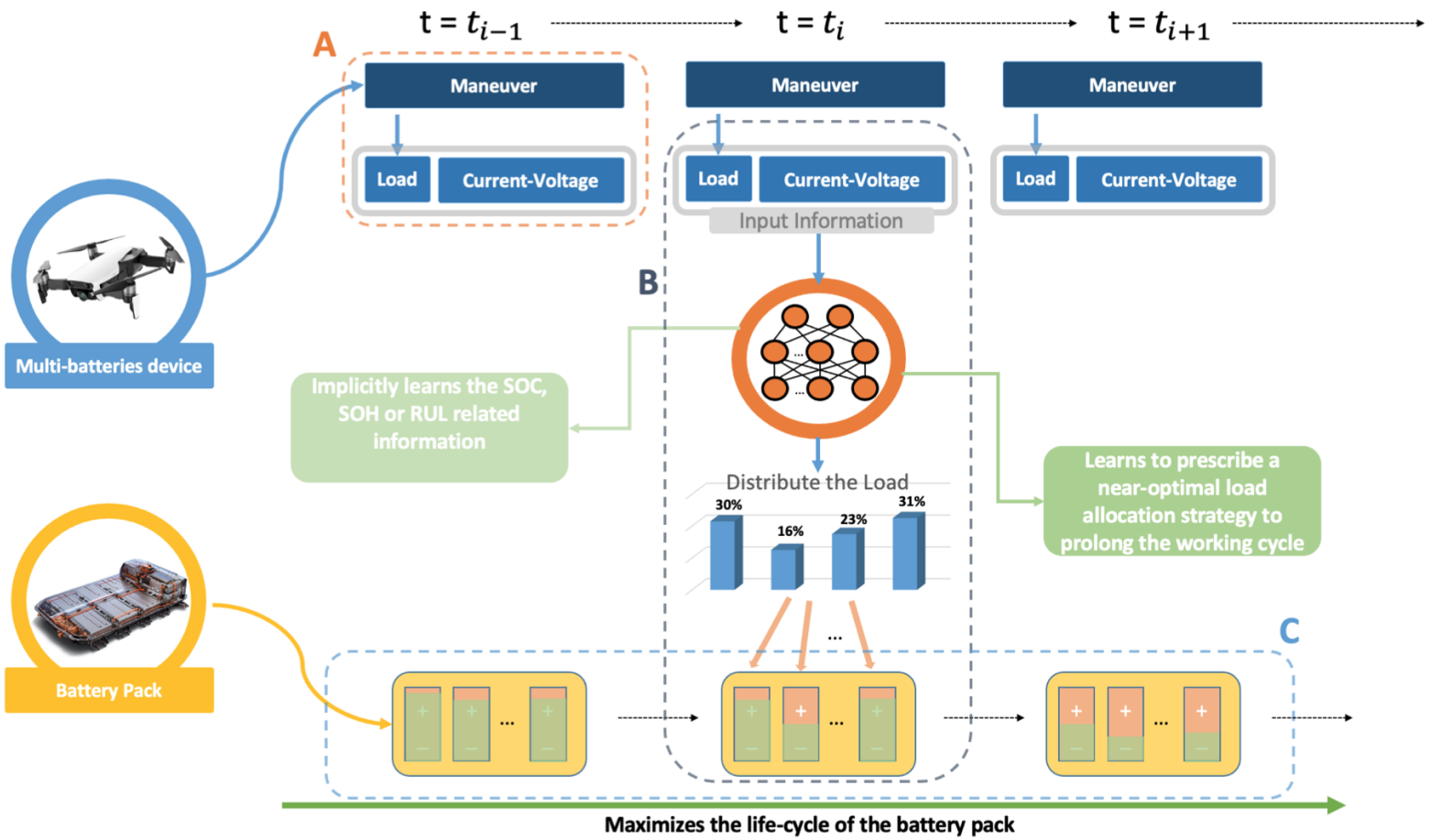
On Average, the working cycle of the deployed system prolongs **151.9%**



151

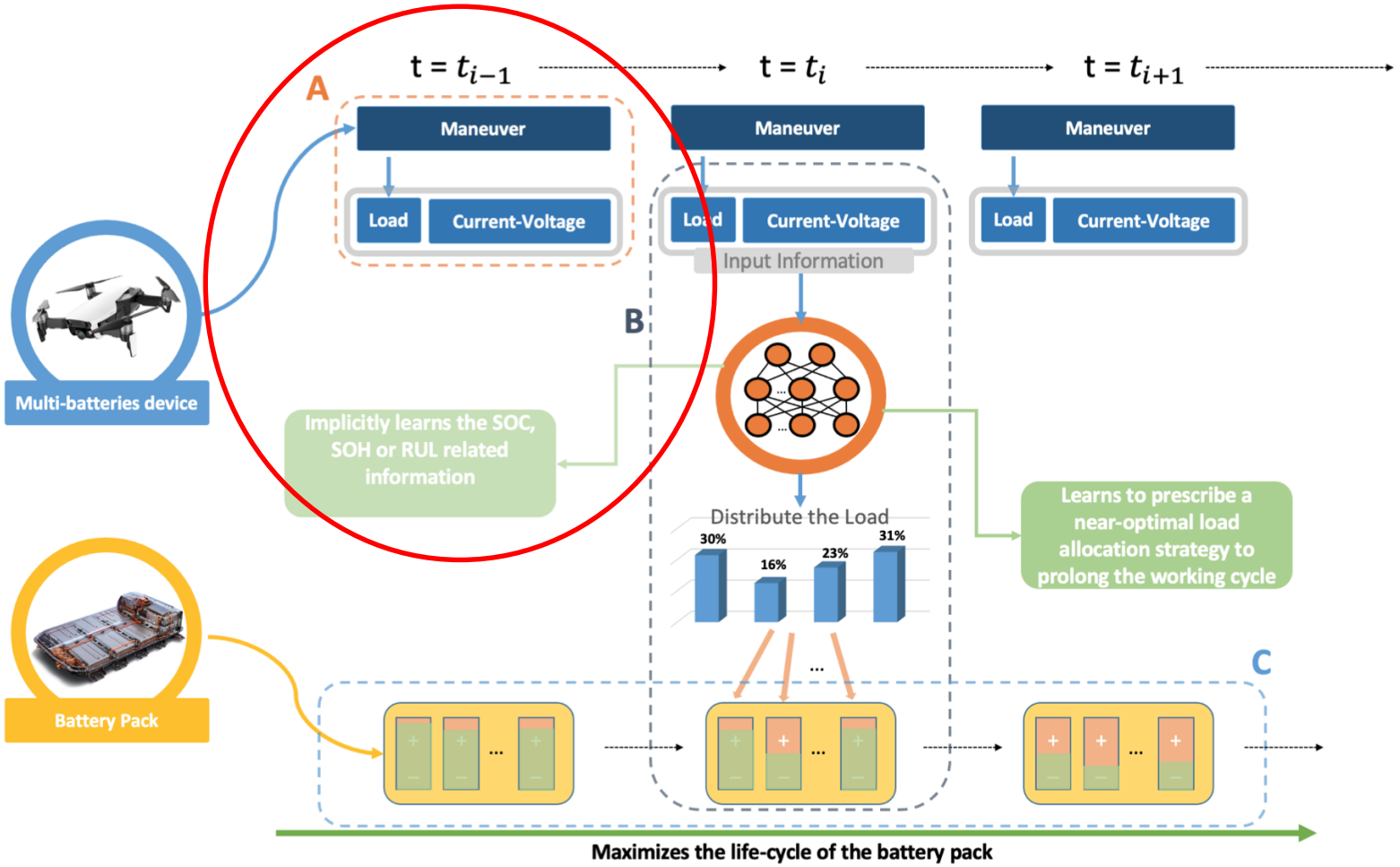
Advantages

- End-to-End
- Scalable
- Global optimization
- Without Prior Knowledge



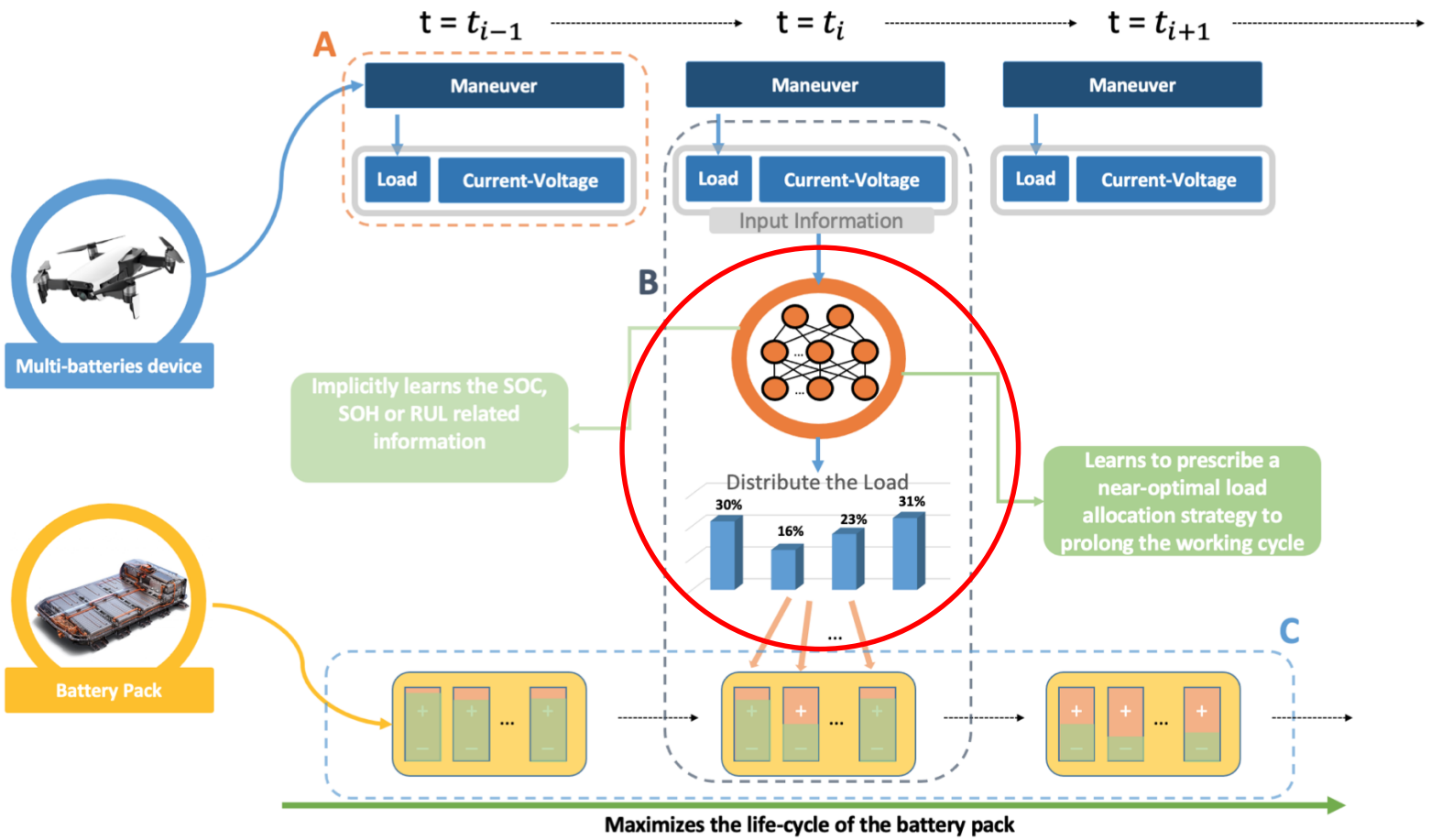
Advantages

- End-to-End
- Scalable
- Global optimization
- Without Prior Knowledge



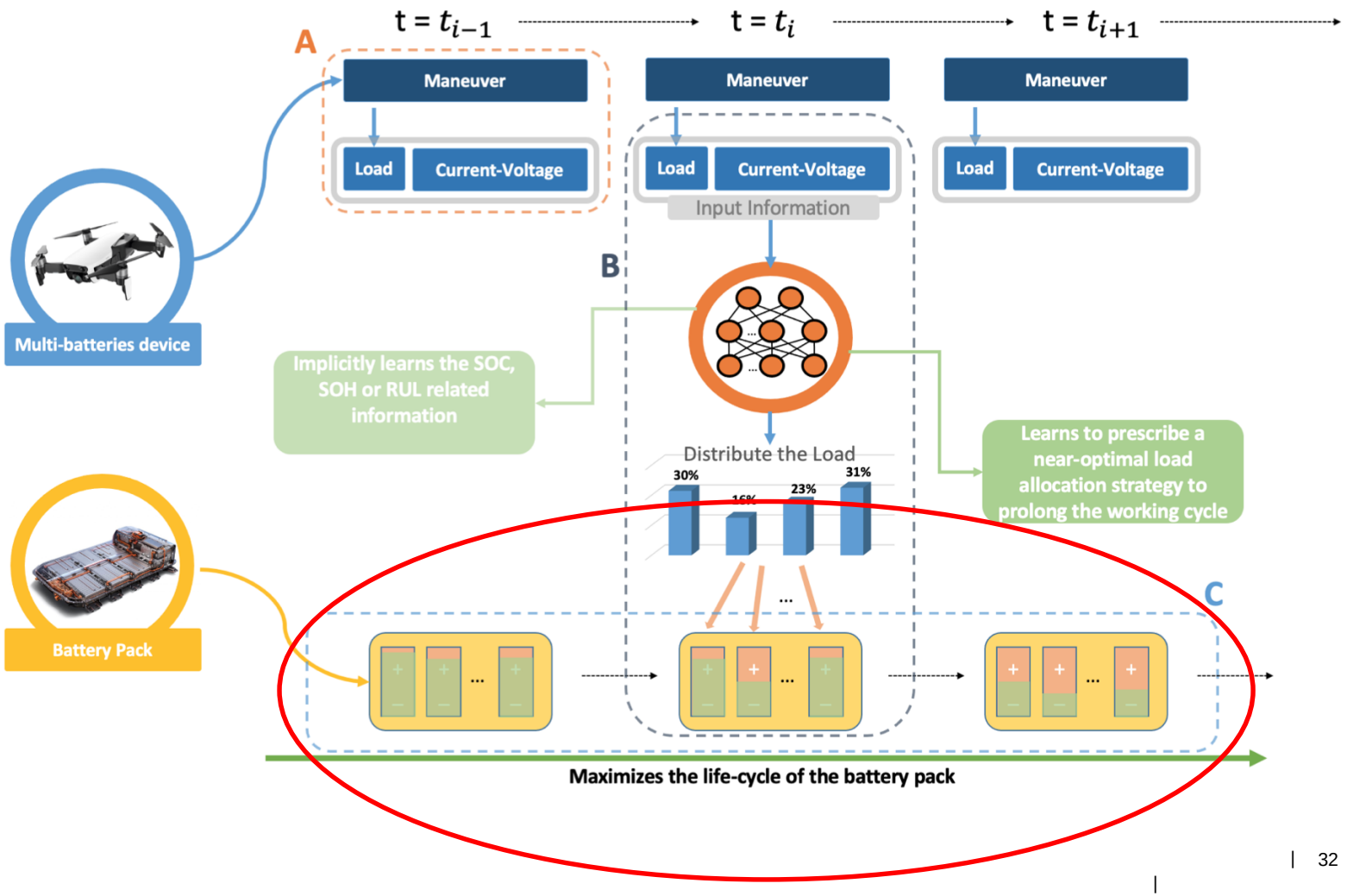
Advantages

- End-to-End
- Scalable
- Global optimization
- Without Prior Knowledge

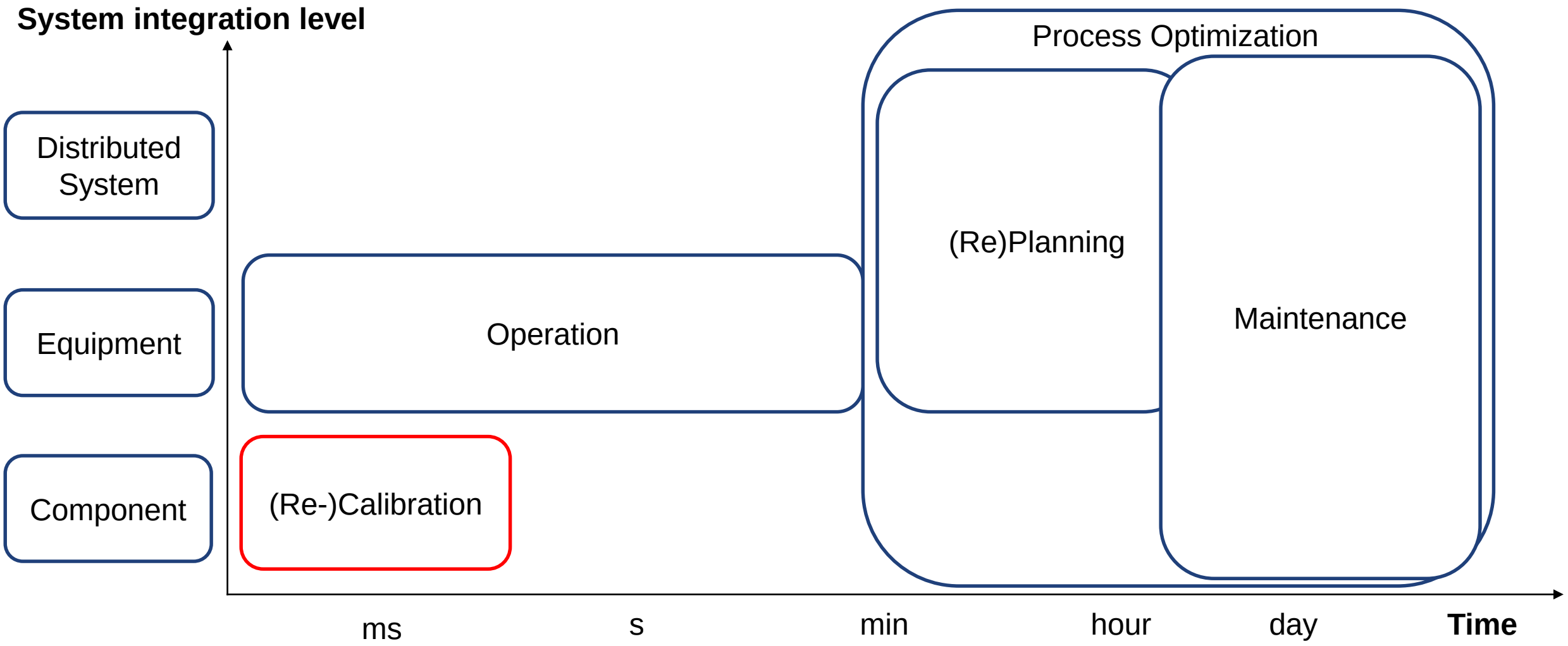


Advantages

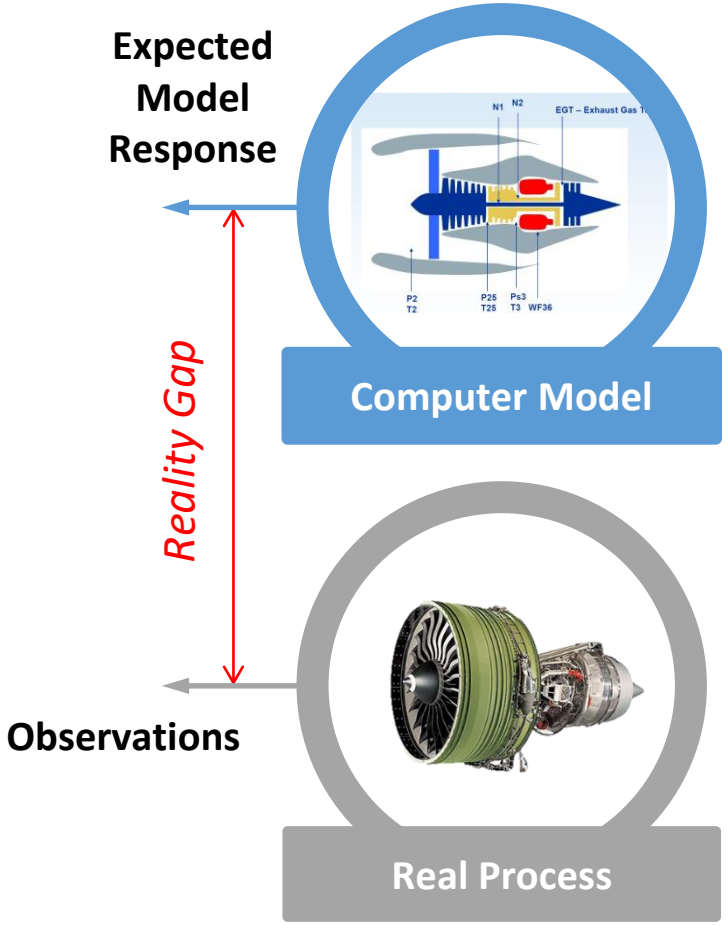
- End-to-End
- Scalable
- Global optimization
- Without Prior Knowledge

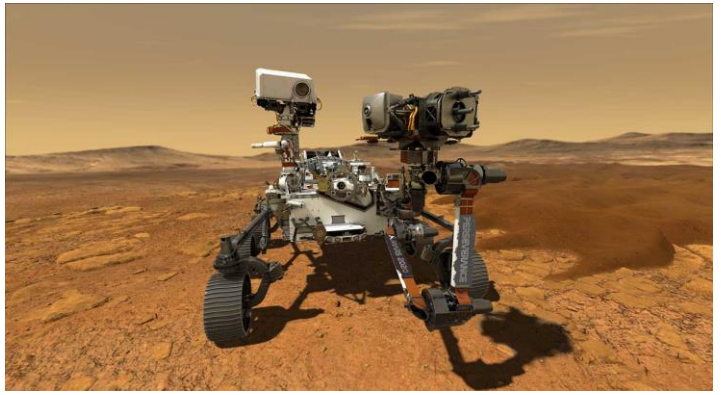


Prescriptive Maintenance



Case study II : Real-time Model Calibration



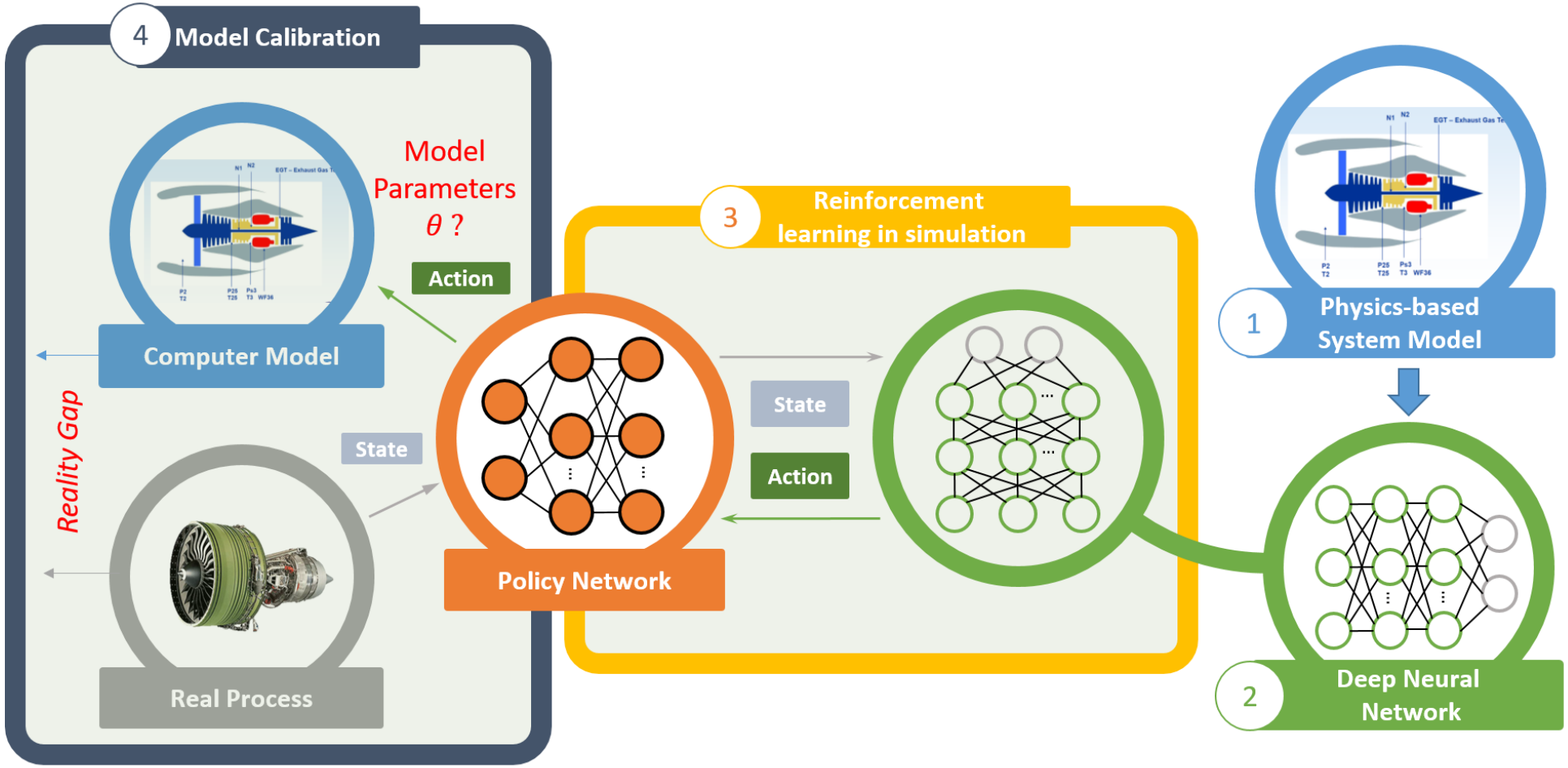


… bridging the reality gap is a general (and very important) problem in science

Challenge

- **Real-time**

Proposed Framework

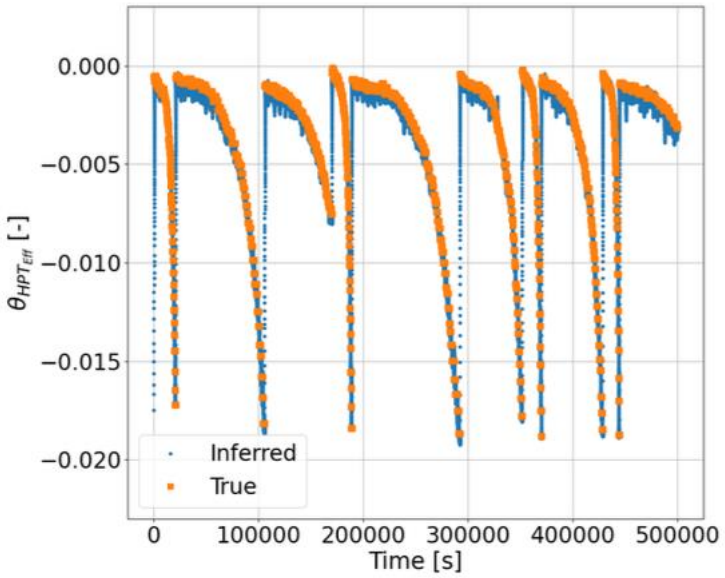


Tracking Problem

- Observe the trajectory or move of the front car (real assets output)
- Controlling the steering wheel (degradation parameters)
- Try to follow the front car (fill the reality gap)
- Once tracked (Calibrated)
- Inverse problem

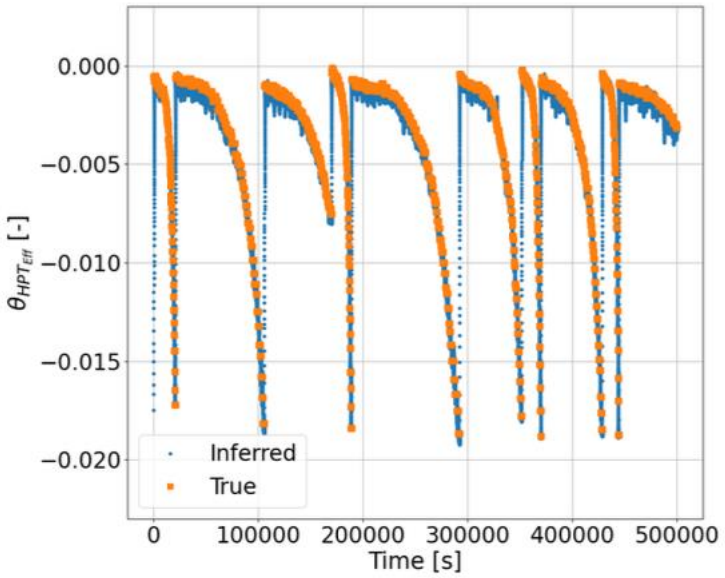
Results

Unscented Kalman Filter

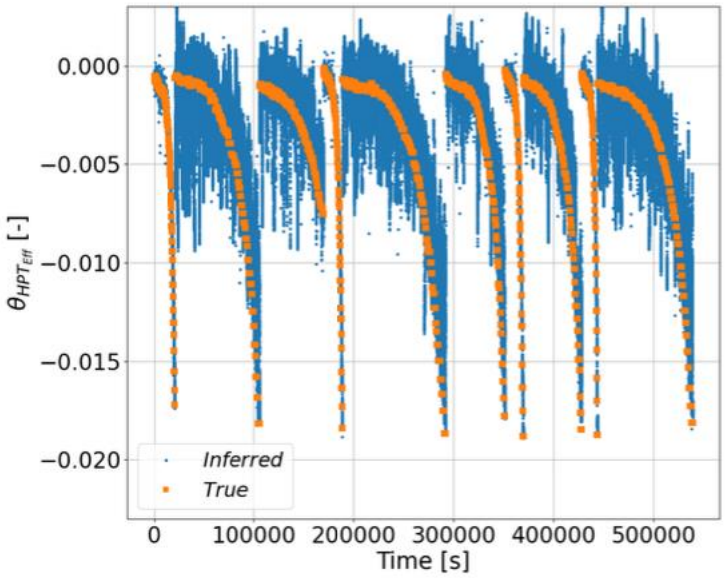


Results

Unscented Kalman Filter

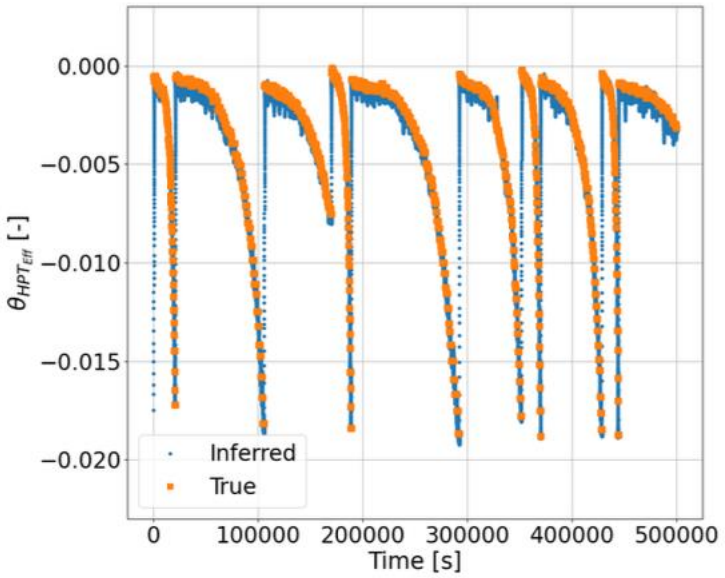


Regression

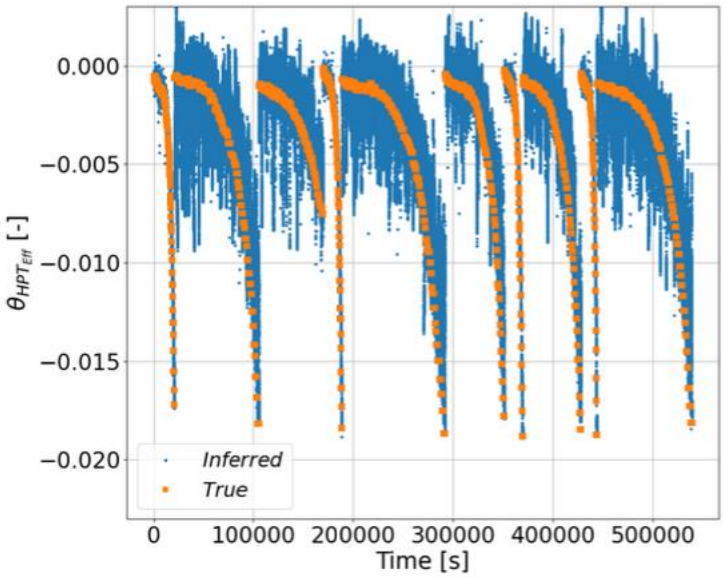


Results

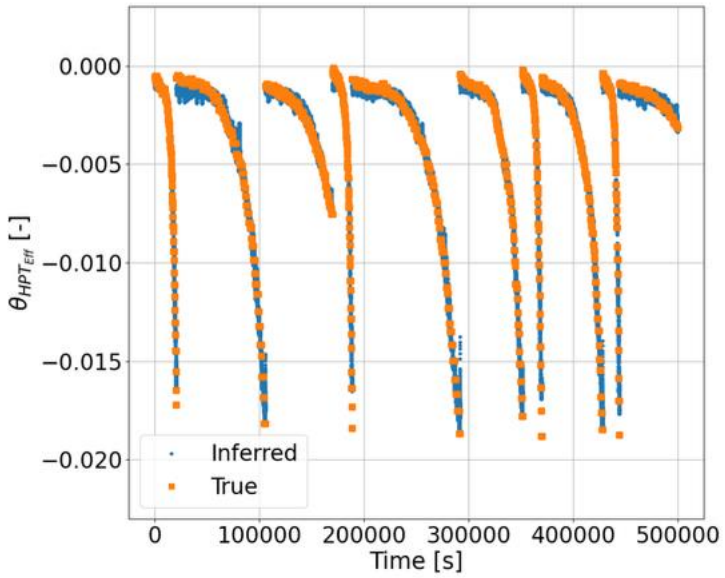
Unscented Kalman Filter



Regression



RL

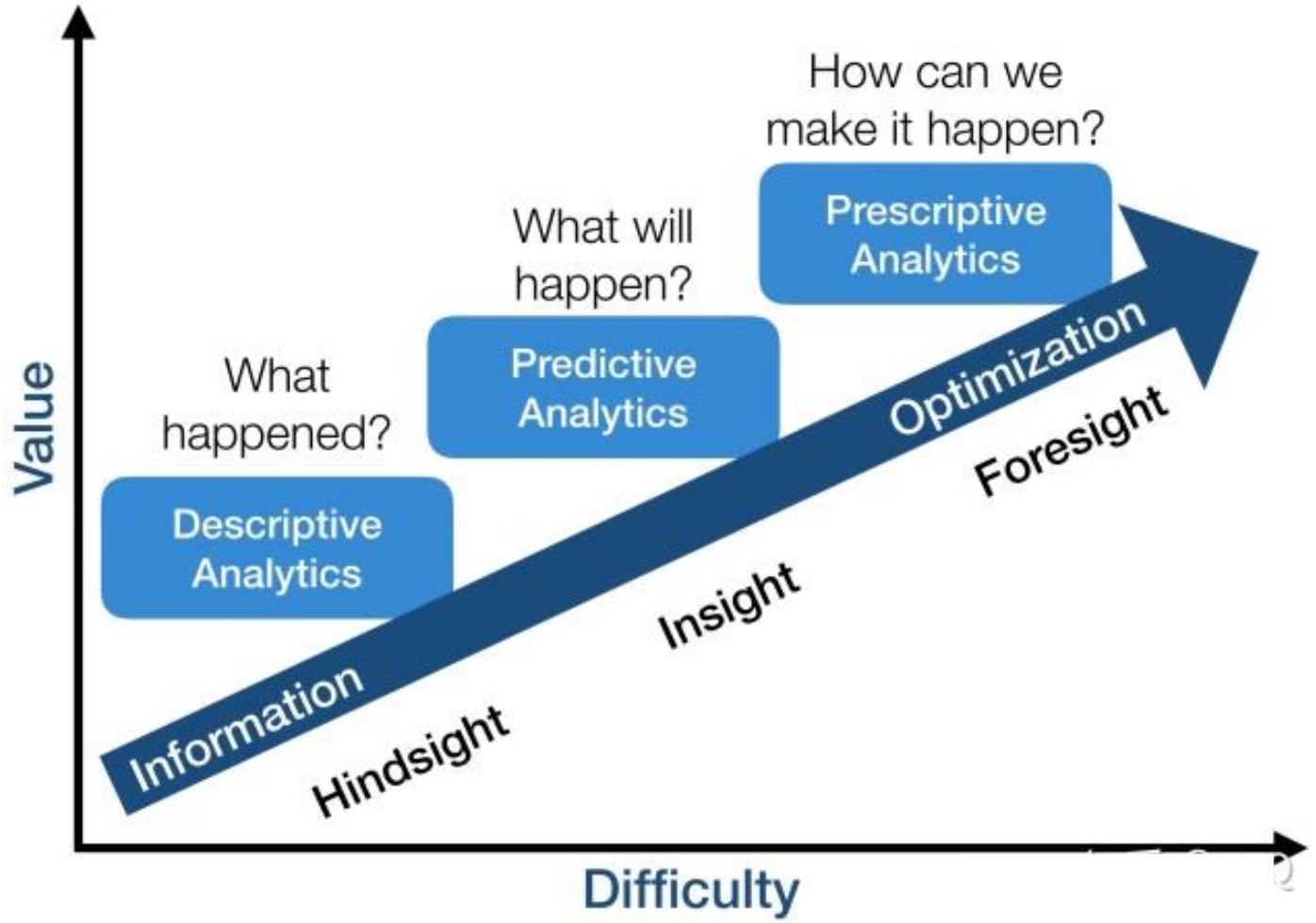


Method	UKF	Reg	RL
Deployment Time [s]	6	4.2e-02	**4.0e-02**

Advantages

- End-to-End
- Scalable
- Without labeled data
- Robust to sensor noise and model bias

Take-home message



Take-home message

Preventive Maintenance

Take-home message

Preventive Maintenance

Predictive Maintenance

Take-home message

Preventive Maintenance

Predictive Maintenance

Prescriptive Maintenance

Take-home message

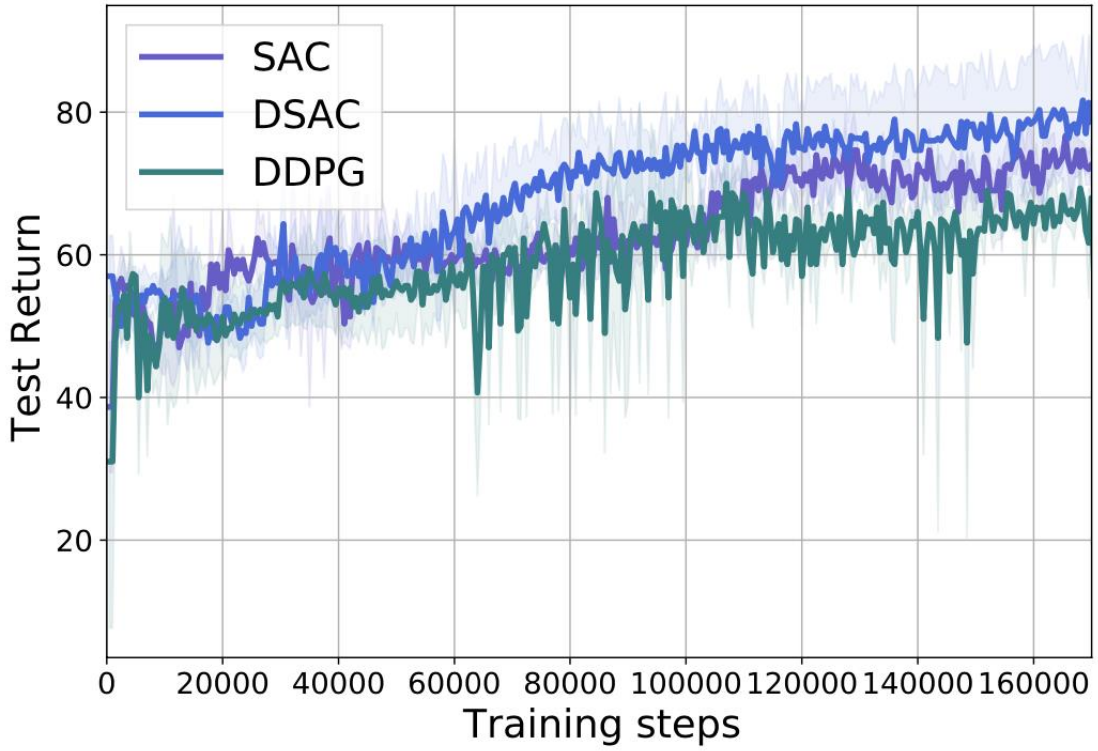
- Reinforcement learning is a promising alternative in maintenance domain, especially for prescriptive operation.

Q&A

Thanks for your attention

Case study I : Power Allocation

- **Dirichlet Soft Actor-Critic (DSAC) compared with the SOTA**
  - **Soft Actor-Critic(SAC)**
  - **Deep Deterministic Policy Gradient (DDPG)**



Method	Case Study #1	Case Study #2
UKF	3.42e–04	3.51e–03
E2E	1.36e–03	–
CLAC	**3.30 ± 0.38e–04**	**2.50e–03**

[2]

Case study II : Real-time Model Calibration