Journal of Zhejiang University-SCIENCE A (Applied Physics & Engineering)

Real-time energy management controller design for a hybrid excavator using reinforcement learning

Qian Zhu

Cite this as: Qian Zhu, Qing-feng Wang, 2017. Real-time energy management controller design for a hybrid excavator using reinforcement learning. *Journal of Zhejiang University-SCIENCE A (Applied Physics & Engineering)*, 18(11):855-870. http://dx.doi.org/10.1631/jzus.A1600650

Optimal energy management problem

Find an optimal assist motor power

$$P_{assist} : \left[t_0, t_f \right] \rightarrow \Omega_{P_{assist}} \left(t \right)$$

such that the constraints

$$\omega_{engine} \in \left[\omega_{emin}, \omega_{emax} \right]$$

$$P_{engine} \in \left[P_{emin} \left(\omega_{engine} \right), P_{emax} \left(\omega_{engine} \right) \right]$$

$$\omega_{motor} \in \left[\omega_{mmin}, \omega_{mmax} \right]$$

$$P_{motor} \in \left[P_{mmin} \left(\omega_{motor} \right), P_{mmax} \left(\omega_{motor} \right) \right]$$

$$SOC_{min} \le SOC\left(t \right) \le SOC_{max}$$

$$SOC\left(t_0 \right) = SOC\left(t_f \right) = SOC_{target}$$

are satisfied and such that the overall fuel

$$J = \int_{t_0}^{t_f} \dot{m}_f \left(P_{assist} \left(t \right), SOC\left(t \right), t \right) dt$$

is minimized

Optimal control problem

Find an optimal control

$$u : \left[t_0, t_f \right] \rightarrow \Omega_u \left(t \right)$$

such that the constraints

$$x\left(t_0 \right) = x_0$$

$$x\left(t_f \right) = x_f$$

$$\dot{x}\left(t \right) = f\left(x\left(t \right), u\left(t \right), t \right)$$

$$x\left(t \right) \in \Omega_x \left(t \right) \quad \text{for all} \quad t \in \left[t_0, t_f \right]$$

are satisfied and such that the cost function

$$J\left(u \right) = \int_{t_0}^{t_f} L\left(x\left(t \right), u\left(t \right), t \right) dt$$

is minimized

Optimal energy management problem

Find an optimal assist motor power

$$P_{assist}:\left[t_0,t_f\right]\rightarrow\Omega_{P_{assist}}(t)$$

such that the constraints

$$\omega_{engine}\in\left[\omega_{emin},\omega_{emax}\right]$$

$$P_{engine}\in\left[P_{emin}\left(\omega_{engine}\right),P_{emax}\left(\omega_{engine}\right)\right]$$

$$\omega_{motor}\in\left[\omega_{mmin},\omega_{mmax}\right]$$

$$P_{motor}\in\left[P_{mmin}\left(\omega_{motor}\right),P_{mmax}\left(\omega_{motor}\right)\right]$$

$$SOC_{min}\le SOC(t)\le SOC_{max}$$

$$SOC(t_0)=SOC(t_f)=SOC_{target}$$

$$J=\int_{t_0}^{t_f}\dot{m}_f\left(P_{assist}(t),SOC(t),t\right)dt$$

is minimized

Energy management problem

↓

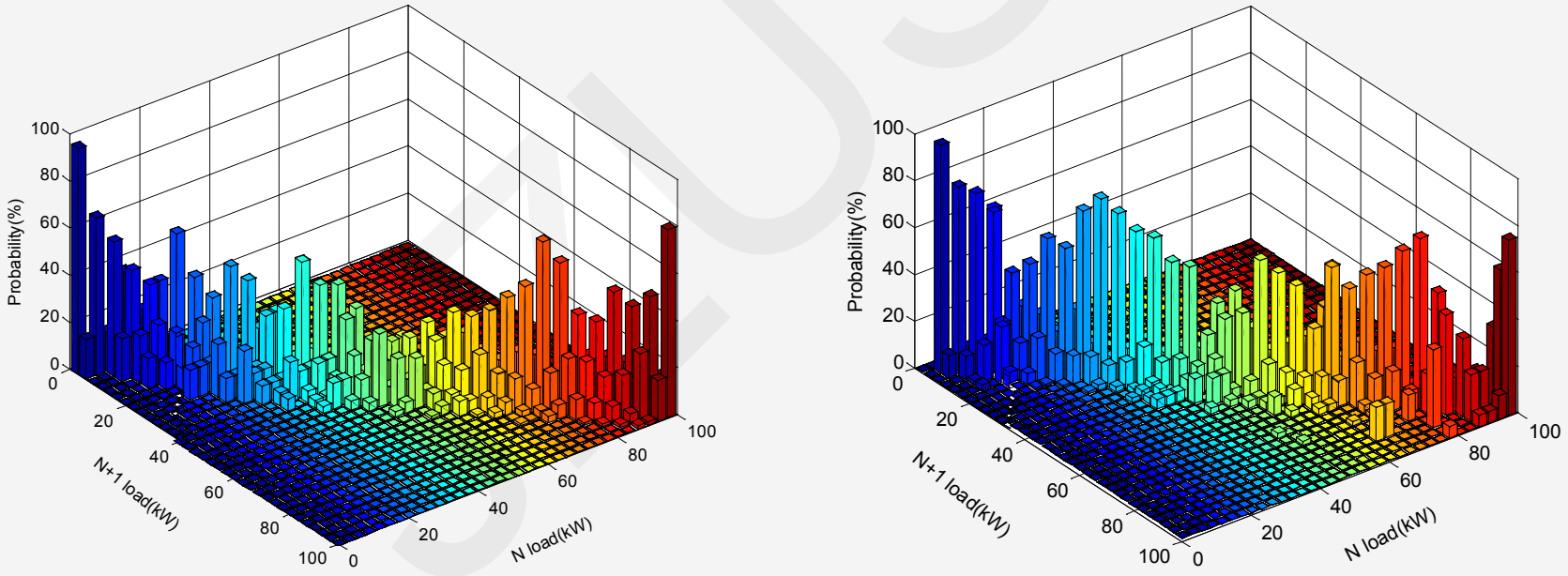
Optimal control problem

Optimal control algorithm	Traditional control algorithm
Global optimal	Sub-optimal
Overall system efficiency	Component efficiency
Long term fuel economy	Instantaneous fuel economy

- **An optimal control problem (Nonlinear, state constraint, control constrain) .**
- **Real-time energy management controller based on reinforcement learning is proposed to deal with the optimal control problem**

Implementation of RL algorithm

- For hybrid excavators, the load cycle can be modeled as a stochastic process with the load power being seen as a stochastic state. Thus the energy management problem in hybrid excavators is essentially a Markov Decision Process (MDP).

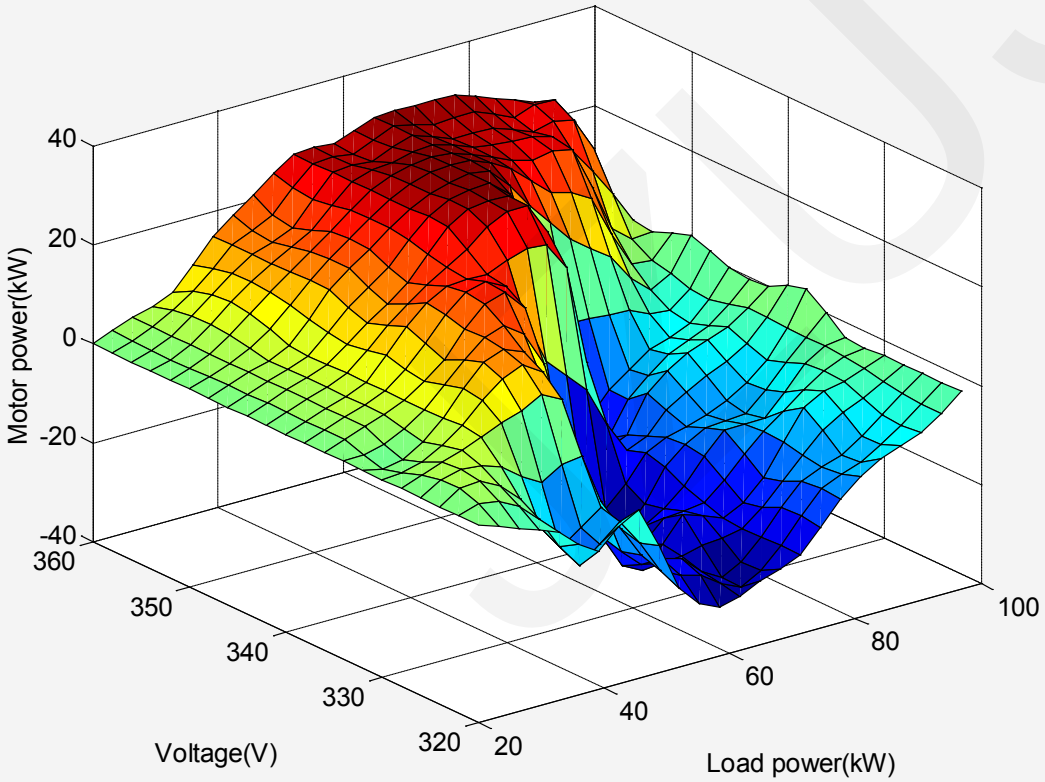


Markov chain model of load cycles: digging(left), lifting(right)

Implementation of RL algorithm

- In RL, energy management problem is time-independent compared with DP algorithm which computes only time-dependent solution for a give load cycle. Thus the optimal solution obtained from RL algorithm can be easily implemented online. The energy management problem for hybrid excavators is formulated as follows using a discounted reward over an infinite horizon:

$$\psi_i = \lim_{k\to\infty} \mathrm{E}\left[\sum_{s=1}^{k} \gamma^{s-1} r\left(x_s, \pi\left(x_s\right), x_{s+1}\right)\right]$$

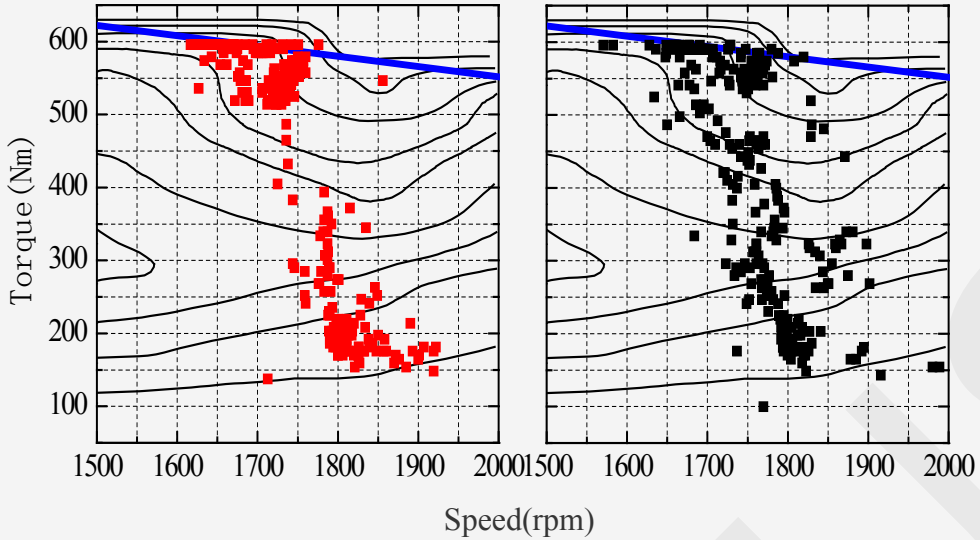


Optimal solution of RL

$$u = f\left(x, l\right)$$

Optimal policy of RL algorithm

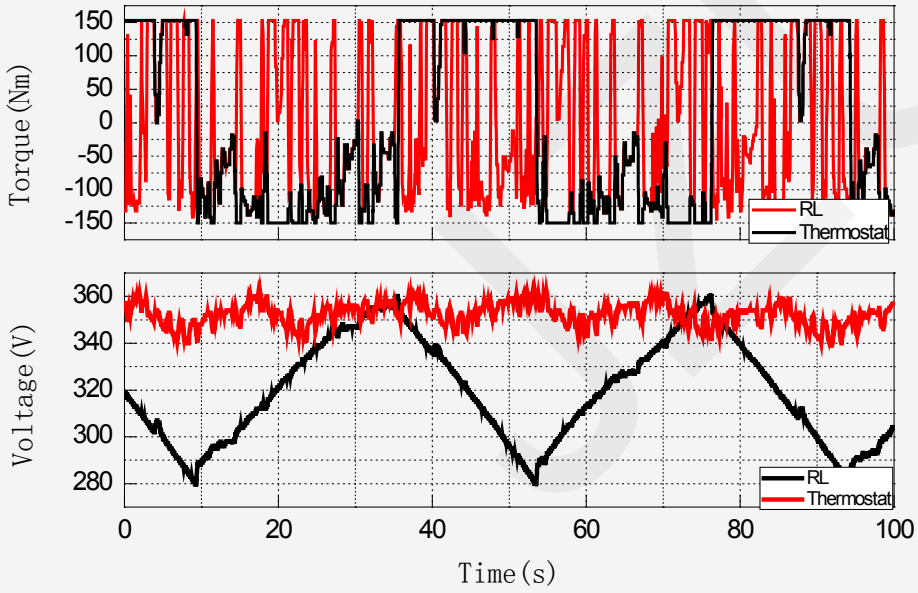
Experiment results



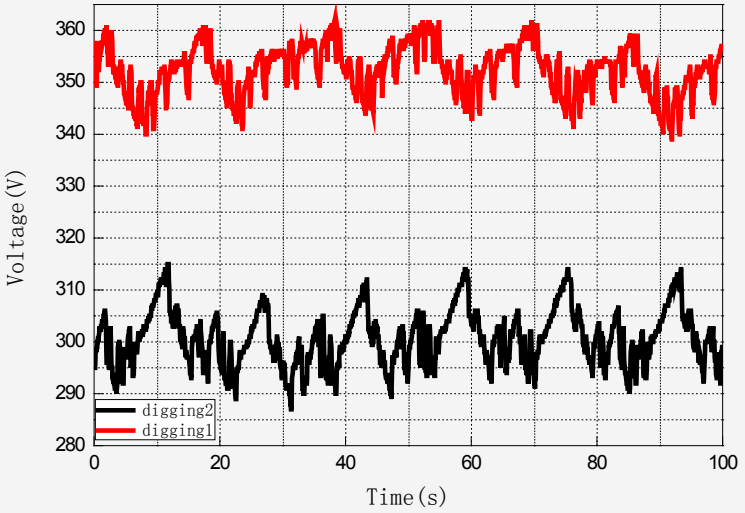
■Experimental results

Advantages of RL:

- ✓ The engine works more efficiently
- ✓ Small SOC fluctuation
- ✓ Applicable to different conditions



RL algorithm VS Thermostat algorithm



SOC for different load cycles