Optimization of Fuel Consumption in a Hybrid Electric Vehicle using Reinforcement Learning and Control Vector Parameterization

Vishnu Masampally, Abhishek Agarwal, Aditya Pareek, Venkataramana Runkana*

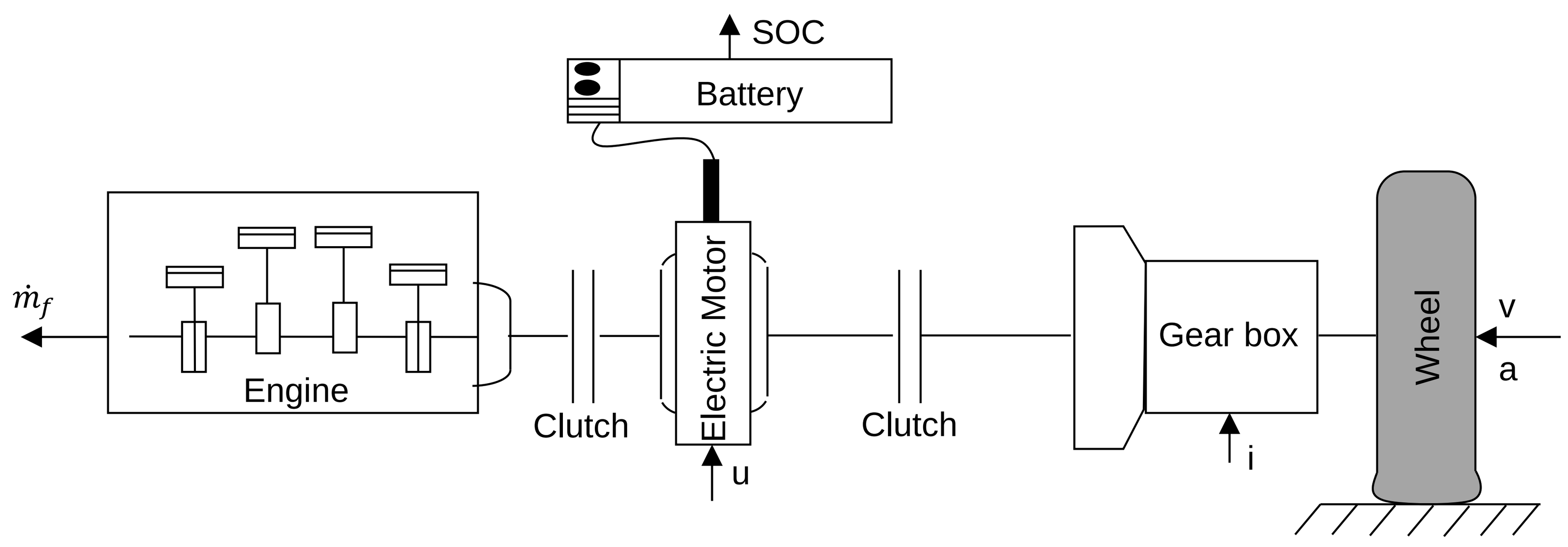
TCS Research, Tata Research Development and Design Centre,
Tata Consultancy Services, 54B, Hadapsar Industrial Estate, Pune - 411013, India
***e-mail: venkat.runkana@tcs.com**

Introduction

- Rise in fuel prices, pollutant emissions and global warming are driving forces for the development of alternative drive vehicles like hybrid electric vehicles (HEVs).
- A hybrid electric vehicle (HEV) uses an electric motor (EM) driven by a battery alongside the conventional internal combustion engine (ICE) to provide the required energy to run the vehicle.
- Finding the optimal torque distribution between the two sources of energy while minimizing fuel consumption of the vehicle and maintaining the state of charge (SOC) of the battery is an optimal control problem.
- The vehicle energy management strategies (EMS) are broadly classified into two classes, namely,:
  1. Heuristics:
     - Example: Fuzzy logic[1]
  2. Optimal control:
     - Example: Pontryagin's minimum principle (PMP), Control Vector Parameterization (CVP)[2] approach using Sequential Quadratic Programming (SQP).

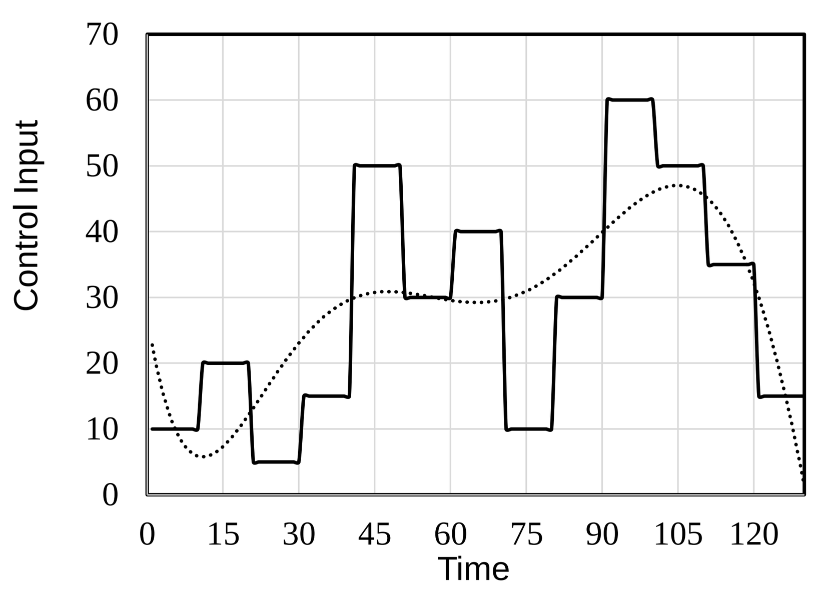
Parallel Hybrid Electric Vehicle

- Powertrain model of a HEV consists of chassis, gearbox, torque coupler, ICE, EM and battery models.
- A quasi-static backward approach is used to model this vehicle in which the direction of calculation is from the wheel to the engine in this approach[3].
- The fuel consumption rate and the SOC of the battery are the output variables obtained from the HEV powertrain model.



CVP-SQP METHODOLOGY FOR OPTIMAL CONTROL

- CVP-SQP methodology is direct method of optimal control
- The control variable is discretized in the form of a polynomial function during each time interval.
- The problem reduces to estimation of coefficient of these polynomials such that the objective function defined over the entire time period is minimized.
- In this work, piecewise constant inputs are considered.
- For SQP, the objective function is approximated to a quadratic model subject to the linearization of the constraints.



Objective function:

- Objective: minimize the fuel consumption over a drive cycle.
- The amount of fuel over the entire drive cycle: $\min \sum_i \dot{m}_f$

Constraints and bounds:

- Inequality constraints:
  - Torques of both the EM and the ICE
  - Speeds of both the EM and the ICE
  - SOC of the battery (0.4 to 0.7)
- Equality constraint:
  - SOC at the end of a drive cycle is equal to 0.5.

Constraints on ICE

$$T_{ICE,i} \geq 0$$
$$T_{ICE,\max}(\omega_{ICE}) - T_{ICE,i} \geq 0$$
$$\omega_{ICE} - \omega_{ICE,idle} \geq 0$$
$$\omega_{ICE,\max} - \omega_{ICE} \geq 0$$

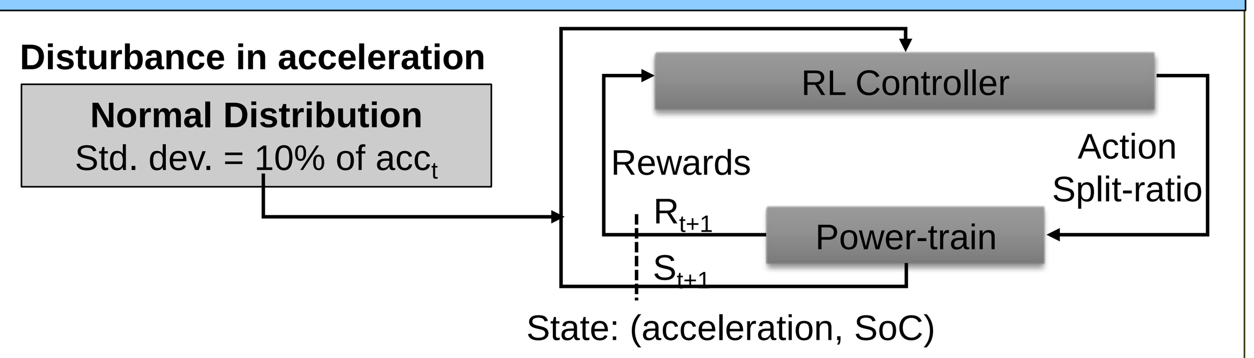
Constraints on EM

$$T_{EM,i} - T_{EM,\min}(\omega_{EM}) \geq 0$$
$$T_{EM,\max}(\omega_{EM}) - T_{EM,i} \geq 0$$
$$\omega_{EM} \geq 0$$
$$\omega_{EM,\max} - \omega_{EM} \geq 0$$

Constraints on SOC

$$0.7 - SOC_i \geq 0$$
$$SOC_i - 0.4 \geq 0$$
$$SOC_{final} - 0.5 = 0$$

Reinforcement Learning (Q-learning) for Optimal Control

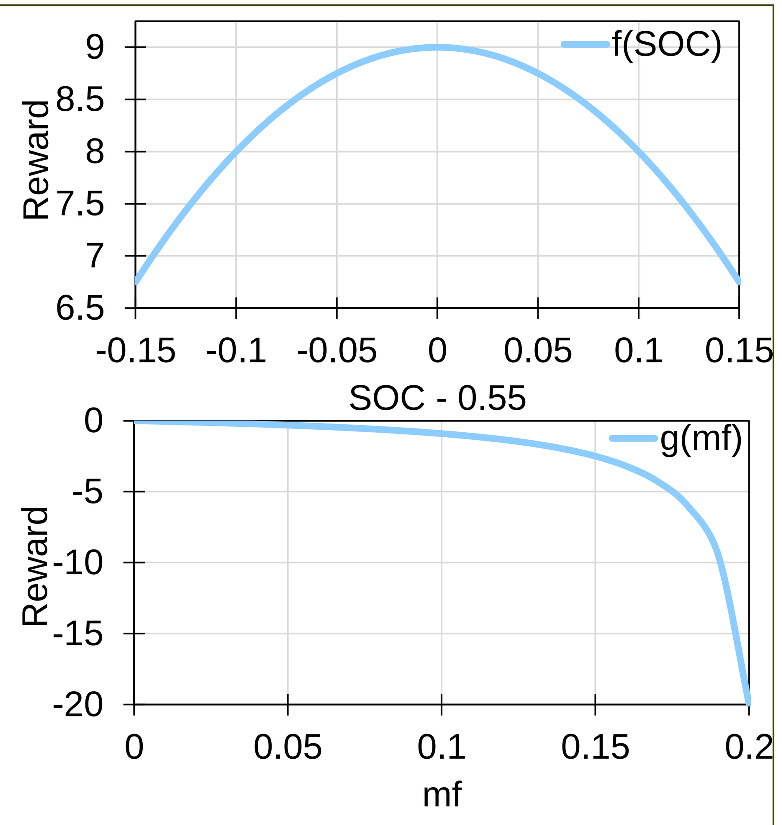


- RL agent: controller
- Environment: Powertrain model
- The RL agent learns the torque-split ratio (action) that minimizes fuel consumption of the ICE based on reinforcements (rewards) for its actions.

Reward functions:

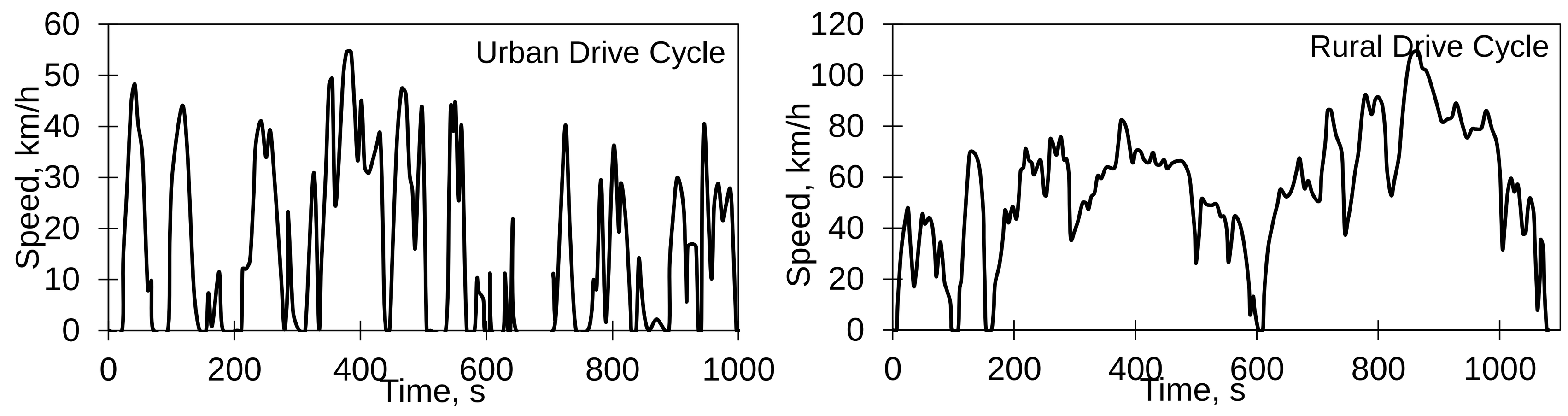
- *f(SOC)*: maintains SOC close to 0.55
- *g(m_f)*: reduces the fuel consumption

Element	Detail
Agent	Controller
Environment	Powertrain model
State Vector	Acceleration (a) and SOC
Action	Torque-Split ratio (u $=T_m/T_g$)
State Space	Acceleration (m/s^2): [-5.0, -4.9, …., 5.0] SOC: [0.40, 0.41, …, 0.70]
Action Space	[-3.00, -2.99, …, 1.10]
Rewards	$\alpha f(SOC) + \beta g(m_f)$

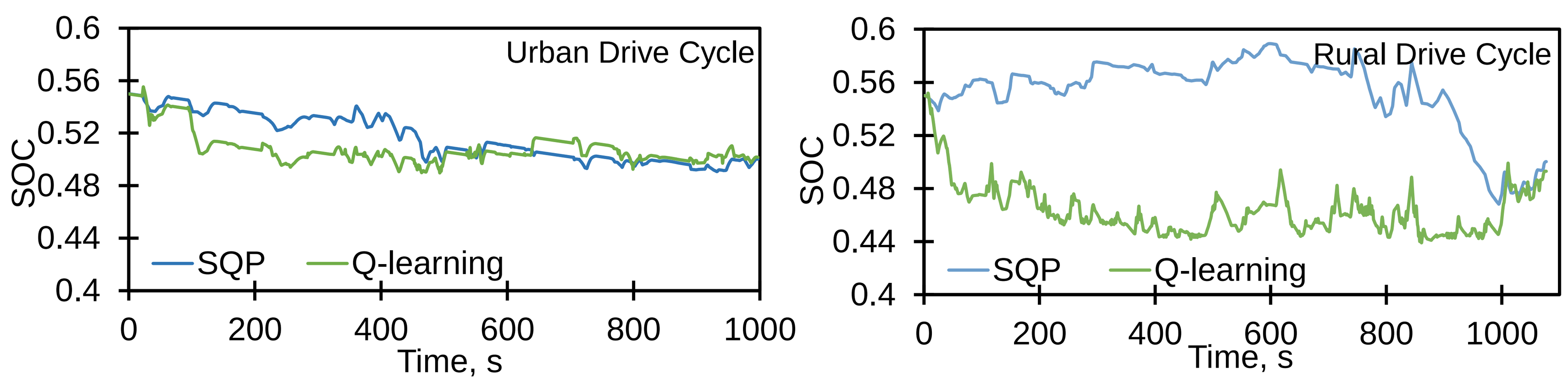


Results

- Results from Q-learning are compared with results from the CVP-SQP approach.
- Common Artemis Driving Cycles (CADC):
  - Urban drive cycle
  - Rural drive cycle



SOC at the optimized settings of the torque-split ratio without any external disturbance:

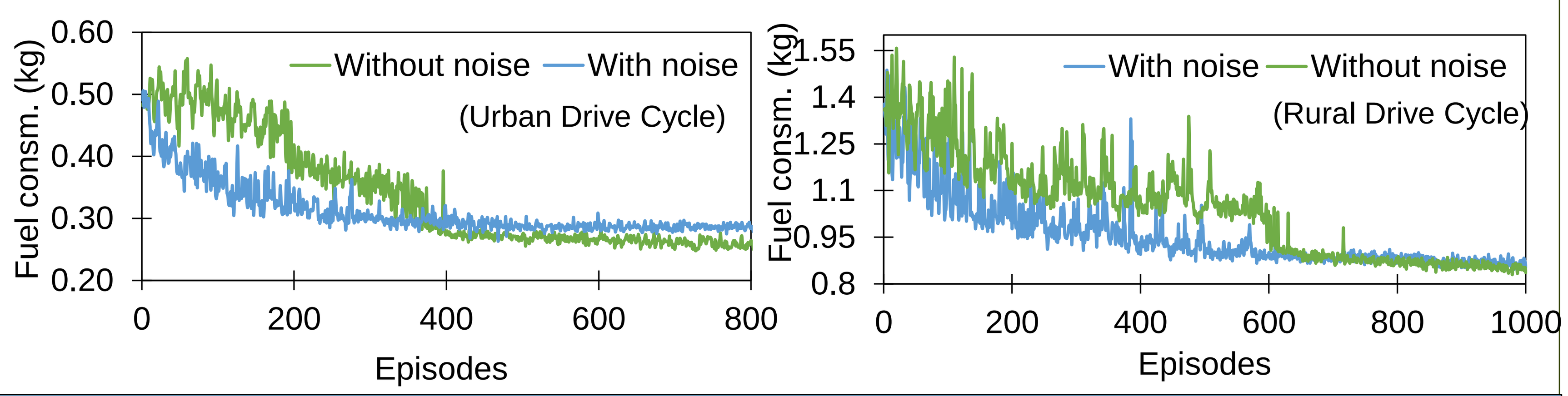


Fuel consumption using CVP-SQP and RL:

Strategy	Total minimal fuel consumed (kg)			
	Without disturbance		With disturbance	
	Urban	Rural	Urban	Rural
SQP-CVP	0.236	0.797	[-]	[-]
RL	0.259	0.844	0.287	0.854

The capability of Q-learning in handling the noise:

- The variations in acceleration were assumed to be in the form of Gaussian noise with standard deviation of 10% about the mean acceleration.



Conclusion

- The true optimal fuel consumption can only be achieved if the entire driving cycle is known beforehand, hence CVP-SQP approach is used as a benchmark to analyse the performance of the Q-learning algorithm.
- Q-learning algorithm could minimize the fuel consumption and the optimal fuel consumption was comparable to that obtained from SQP.
- The controller based on RL algorithms learned quite well to cater to the power demands as can be seen by comparing the SOC profile with that of the SQP.
- In the presence of external disturbances, the Q-learning algorithm managed to converge to the optimal fuel consumption.

References

1. Schouten, Niels J., Mutasim A. Salman, and Naim A. Kheir. "Energy management strategies for parallel hybrid vehicles using fuzzy logic." *IFAC Proceedings Volumes* 33.26 (2000): 83-88.
2. Goh, C.J. and Teo, K.L., 1988. Control parametrization: a unified approach to optimal control problems with general constraints. *Automatica*, 24(1), pp.3-18.
3. Sundström, O., 2009. Optimal control and design of hybrid-electric vehicles, *Ph. D. thesis, ETH Zurich, Switzerland*.

TATA
TATA CONSULTANCY SERVICES

Acknowledgement
Authors would like to thank Mr. K Ananth Krishnan, CTO, Tata Consultancy Services for his constant encouragement and support during this project.