

Real-time realization of Dynamic Programming using machine learning methods for IC engine waste heat recovery system power optimization

Bin Xu, Dhruvang Rathod*, Adamu Yebi, Zoran Filipi

Clemson University, Department of Automotive Engineering, 4 Research Dr., Greenville, SC 29607, USA

HIGHLIGHTS

- Dynamic Programming optimizes the power of organic Rankine Cycle.
- Proper Orthogonal Decomposition Galerkin Projection reduced model is used.
- 11 machine learning models extract rules from offline Dynamic Programming.
- Contribution of inputs to rule extraction is compared.
- Transient optimization is proved to be different from steady state optimization.

ARTICLE INFO

Keywords:

Waste heat recovery
Organic Rankine Cycle
Dynamic Programming
Transient optimization
Real-time implementation
Machine learning

ABSTRACT

This study aims to present a method for real-time realization of Dynamic Programming algorithm for power optimization in an organic Rankine Cycle waste heat recovery system. Different from existing studies, for the first time machine learning algorithms are utilized to extract the rules from offline Dynamic Programming results for optimal power generation. In addition, for the first time a single state Proper Orthogonal Decomposition and Galerkin Projection based reduced order model is combined with Dynamic Programming for its high accuracy and low computation cost. For a transient driving cycle, Dynamic Programming algorithm is utilized to generate the optimal working fluid pump speed. A total of eleven state-of-art machine learning algorithms are screened to predict this pump speed. Random Forest algorithm is then selected for its best pump speed prediction accuracy. A rule-based method is added to the Random Forest model to improve energy recovery. As one of the main discoveries in this study, in the rule extraction process, the Random Forest model reveals that the time delayed exhaust gas mass flow rate and exhaust temperature improve the rule extraction accuracy. This observation points out the difference between steady state and transient optimization and that the steady state optimization results do not necessarily hold true in transient conditions. Another key observation is that Random Forest – rule-based method retrieves 97.2% of the energy recovered by offline Dynamic Programming in a validation driving cycle. In addition, the inclusion of rule-based method significantly increases the Random Forest model's energy recovery from 66.5% to 97.2%. This high accuracy means that the machine learning models can be used to extract Dynamic Programming rules for real-time application.

1. Introduction

In an internal combustion engine, 30–60% of fuel energy is wasted as heat [1]. Given such large amount of waste heat, waste heat recovery (WHR) techniques have the potential to reduce fuel consumption and CO₂ emissions. Among the WHR technologies, thermoelectric generators, turbo-compounding and Organic Rankine Cycle (ORC) are popular in literature. The thermoelectric generator generates electricity using temperature difference between waste heat and thermoelectric

material coolant. The structure of thermoelectric generator is compact and simple, but the thermal efficiency is low and limited by the materials [2]. Turbo-compounding integrates the turbocharger with electrical generator or with engine crankshaft [3]. Even though it can generate electricity or improve crankshaft power output, its heat source is limited to tail pipe exhaust gas. In addition, for a heavy-duty diesel engine, turbo-compounding turbine is installed downstream of main turbocharger, where thermal energy is in the form of low availability for the turbo-compounding turbine. ORC utilizes heat sources to

* Corresponding author.

E-mail address: dhruvar@clemson.edu (D. Rathod).

| Nomenclature | |
|------------------------------------|--|
| A | area [m^2] |
| b | number of simulation time steps |
| c_p | specific heat capacity [J/kg K] |
| d | diameter [m] |
| h | enthalpy [J/kg] |
| k | thermal conductivity [W/m K] |
| l, L | length [m] |
| m | mass [kg] |
| \dot{m} | mass flow rate [kg/s] |
| n | finite volume model number of cells |
| N | revolution speed [rpm] |
| O | valve opening [%] |
| p | pressure [Pa] |
| q | dimension of state |
| Q | heat [J] |
| t | time [s] |
| T | temperature [K] |
| u | velocity [m/s], internal energy [J/kg] |
| U | heat transfer coefficient [W/m ² K] |
| v | dynamic viscosity [m^2/s] |
| V | volume [m^3] |
| W | work/power [W] |
| x | state |
| z | space coordinate [m] |
| r | Ratio |
| γ | specific heat ratio |
| ρ | density [kg/m ³] |
| ∂ | partial derivative operator |
| Φ | basis function |
| <i>Abbreviations</i> | |
| ORC | organic Rankine Cycle |
| RC | Rankine Cycle |
| WHR | waste heat recovery |
| RB | rule-based |
| RF | random forest |
| HT | high temperature |
| LT | low temperature |
| MPC | Model-based Predictive Control |
| NMPC | Nonlinear Model-based Predictive Control |
| PDEs | partial differential equations |
| POD | Proper Orthogonal Decomposition |
| ROM | reduced order model |
| SVD | singular value decomposition |
| MBM | moving boundary method |
| FVM | finite volume method |
| EGR | exhaust gas recirculation |
| FTP | Federal Test Procedure |
| CSVL | Constant Speed Variable Load |
| TP | tail pipe |
| sim | simulation |
| exp | experiment |
| <i>Subscripts and superscripts</i> | |
| f | working fluid |
| w | wall |
| e | exhaust gas |
| v | vapor |
| l | liquid |
| in | inlet/ upstream |
| out | outlet/ downstream |
| p | pressure |
| U | heat transfer coefficient |
| is | isentropic |
| d | discharge |
| Op | valve opening |
| vlv | valve |
| HPP | high pressure pump |
| TP | tail pipe |
| turb | turbine |

evaporate working fluid and the generated vapor then spins the turbine to produce power. Even though ORC system is more complex and heavier than thermoelectric generator or turbo-compounding, it has the highest thermal efficiency [1]. Thus, this study focuses on ORC-WHR technology.

For the power optimization of the ORC-WHR system, the optimal results can be achieved by either optimizing actuators positions (e.g. pump speed, valve opening, and expander speed [4]) or state variables trajectories (e.g. evaporation temperature, vapor temperature, and evaporation pressure [5]) for the given heat source conditions. These optimal actuator positions or state variable trajectories can then be stored as a function of heat source and working fluid conditions (e.g. exhaust gas mass flow rate, exhaust gas temperature, and working fluid mass flow rate [6]) for the real-time implementation.

The ORC-WHR system development can be divided into concept phase, components selection phase, system integration phase, control development phase and experimental validation phase [1]. In literature, most of power optimization studies were conducted in concept phase or components selection and development phase, where details of components were not given and only steady state engine operating conditions were considered (or heat exchanger models were static). Nazari et al. optimized Rankine Cycle (RC) WHR exergy efficiency and system cost at the same time [7]. Heat exchanger types were not given and two turbines were chosen as expanders for the two loops. Steam turbine inlet pressure, organic turbine inlet pressure and organic

preheater pinch temperature were considered as the optimization variables. Heat source was from a SGT-500 Siemens gas turbine and only a steady state condition was considered. The exhaust gas temperature is fixed at 400 °C. Genetic algorithm was considered in the optimization. The results showed that the optimization improved exergy efficiencies by 2.3–4.2% depending on working fluid types. Dai et al. optimized exergy efficiency at steady state waste heat condition [8]. Two heat exchangers were considered in ORC loop and condenser was configured as a recuperator downstream of the working fluid pump. Heat source temperature is fixed at 145 °C. Turbine was chosen as expander and the turbine inlet temperature and pressure were the optimization variables. Ten working fluids were compared in the optimization. The results showed that R236EA, isobutene, butane, R245CA and R123 were the top five working fluids in terms of net power production. Sadeghi et al. optimized net power and turbine expander size in steady state geothermal conditions [9]. Details of the components were not decided. Four variables were chosen as optimization parameters including evaporation pressure, superheat temperature, pinch point temperature difference of evaporator one, and evaporation pressure of evaporator two. Temperature of geothermal heat source was fixed as 100 °C. Among multiple working fluid types, zeotropic working fluid like outperformed pure fluids by 20–30% net power production. The optimal superheat temperature was 14.83 °C. Yang et al. optimized ORC-WHR net power and turbine expander size in steady state compressed natural gas engines [10]. Components like

control valves, and reservoir were ignored. Geometry details of fin-and-tube evaporator and plate type heat exchanger were provided and different heat transfer coefficients were given in different working fluid phases. Pressures and temperatures related to evaporation and condensation were chosen as the optimization variables. Results showed that optimal superheat temperature was among the range of 0.43–12.35 °C and optimal evaporation pressure was among 2.5–2.9 MPa.

All the aforementioned optimization are in initial phases of the ORC-WHR system development and provide sensitivity study of different parameters, which are helpful for the component design and system development. However, in these optimization literatures, only major components are considered, many components geometries are not decided and some components are ignored. Thus, the optimization results are far from the ORC-WHR system operation with finalized components design.

Some researchers investigated optimal operating condition in control development and power optimization phases. In these phases, components of the ORC-WHR system are already selected or designed. Different from the goal of initial phases, the goal of the later phases aims to maximize power production in system operation with fixed components and working fluids.

Quoilin et al. considered 31 steady state points to find the optimal evaporation temperatures [6]. R245fa was chosen as working fluid and an oil-free scroll expander was utilized. Only a single evaporator was considered and the geometry details of expander and heat exchangers were given. The heat source were between 120 and 300 °C. After the optimization, the optimal temperatures were correlated with working fluid mass flow rate, working fluid condensation temperature and heat source temperature. The correlation then served as reference in real-time control. Xu et al. utilized two steady state points to find optimal pump speed, turbine speed, condenser coolant pump speed and mass flow distribution of parallel evaporators [4]. There were two heat sources including TP exhaust gas and exhaust gas recirculation (EGR), which were supplied by a heavy-duty diesel engine. Turbine expander and ethanol were selected as expander and working fluid, respectively. Details of components modeling were given. Operational maps were built for respective steady state engine conditions, which gave guidance of working fluid pump speed and turbine speed setup in order to achieve optimal power generation.

Even though Quoilin et al. and Xu et al. generated correlations or maps using optimization results, the optimizations were conducted in steady state engine operating conditions. Most of engine operating conditions were transient rather than steady-state. During the highly transient engine conditions, the thermal inertia in the ORC-WHR system plays a key role in the temperature response. According to an evaporator thermal inertial sensitivity study in [11], working fluid vapor temperature settling time could triple as thermal inertia increases in the evaporator. This response time difference can not be captured by steady state optimization as thermal inertia only affects transient performance rather than steady state performance. When ORC-WHR system is operated based on the map generated from the steady state optimization, the thermal inertia leads to different vapor temperature, pressure, thus net power production. Therefore, there are gaps between steady state optimization results and real-time transient optimization results. Transient optimization is needed to fill the gaps.

Xu et al. optimized working fluid vapor temperature reference over a transient heavy-duty diesel engine driving cycle [11]. Three working fluid vapor temperature strategies were compared, namely: (1) constant vapor temperature, (2) constant superheat temperature, and (3) fuzzy logic superheat temperature based on waste heat power levels. Vapor temperature was controlled by a PID control, which followed the three different vapor temperature trajectories. The evaporator models were finite volume models and all the component models were identified and validated with experimental data. The results showed that the fuzzy logic superheat temperature produced the highest net power across the

driving cycle. However, the PID controller experienced large variation and left much room for improvement. To address the PID temperature tracking issue, many Model-based Predictive Control (MPC) temperature control publications can be found in the areas of ORC-WHR. Yebi et al. designed a Nonlinear MPC (NMPC) for a Heavy-Duty Diesel Engine WHR system [12]. NMPC showed shorter settling time and less overshoot. Compared with PID control, the NMPC recovered 9% more energy over a 1200 s driving cycle and increased the turbine expander operation time duration by 12%. Hernandez et al. also concluded that the MPC produced better vapor temperature tracking performance [13]. Similar conclusion can also be found in [14] in Heavy-Duty ORC-WHR application. Even though MPC is quite efficient in vapor temperature tracking, this method does not directly optimize net power in ORC-WHR system and thus its performance highly depends on the optimality of the vapor temperature trajectory. Direct net power optimization could address this concern.

One solution is to formulate MPC for net power optimization. Esposito et al. defined net power production in the cost function of NMPC [15]. The NMPC outperformed a feedforward control with regard to working fluid pressure and temperature constraints violation. The total number of states in the ORC-WHR system was 16 by implementing moving boundary model in evaporator and condenser, which resulted in intensive computation in the NMPC solver and compromised real-time implementation potential. Merino et al. designed an MPC for a Heavy-Duty Truck ORC-WHR system [16]. A 6-state moving boundary model was used as the evaporator control-oriented model. During a 1200 s exhaust gas mass flow rate three-step change transient condition, the MPC recovered 2% more energy than the PID did. However, the optimality of MPC results depends on the horizon it looks forward, which is not the global optimum for the entire driving cycle, but the optimum in a tiny time window. The time window is 5 s in both literature [15].

Another solution is Dynamic Programming (DP) net power optimization. In this case, the results are global optimum. DP, as one of the state-of-art transient optimization algorithms, was utilized in the ORC-WHR system application. Paralez et al. optimized working fluid evaporation and condensation pressure over transient engine conditions using DP algorithm [17]. Evaporator was modelled with 0-D model and components models were not given. The results revealed that DP was capable of handling transient operating condition and maximized net power production. However, DP algorithm is implemented offline due to its high computational cost, which is known as the curse of dimensionality [18]. The computational cost grows exponentially as number of states increases. Due to the high computation cost, DP is run offline and the number of states is restricted to finish the DP simulation in a reasonable time duration. This limitation on the number of states diminishes the accuracy of the reduced order model utilized in DP, which is another drawback brought out by DP algorithm in the ORC-WHR application. Even though the DP is implemented offline, the off-line DP rule extraction and its real-time implementation are not well studied. In ORC-WHR field, DP work was only limited to offline optimization [17] and real-time DP is not researched, which is the main motivation of this study.

One possible approach to implement the DP algorithm in real-time is extracting the DP rules from its offline generated results. Different from the correlation method studied in literature, this study proposes a method to extract the DP rules which can be implemented in real-time, and is the main contribution of this study. More specifically, different machine learning algorithms are evaluated to fit the relation between working fluid pump speed and the exhaust gas conditions. The trained machine learning models are then evaluated to mimic the DP algorithm and predict the optimal pump speed over a driving cycle. During the DP optimization, a high accuracy and low cost evaporator reduced model is utilized, which is derived using Proper Orthogonal Decomposition and Galerkin Projection method.

In the past decade, machine learning has been soaring in many

fields for its excellent performance in complex systems and nonlinear problems, such as pattern recognition, image processing, voice recognition, and natural language processing [19]. ORC-WHR system is a complex system, which includes high nonlinearity in working fluid phase change. Therefore, machine learning algorithms could have the potential to extract rules from DP optimization results for a highly nonlinear system.

This study is organized as follows: **Section 2** introduces architecture and modeling of the ORC-WHR system. Optimization problem formulation is described in **Section 3**, followed by DP optimization in **Section 4**. In **Section 5**, the DP rule extraction is conducted with Random Forest (RF) machine learning model. In **Section 6**, the performance of RF model is then validated and evaluated over a drive cycle. Finally, the study ends with conclusion section.

2. Modeling

DP algorithm is a model-based method, which requires a dynamic system model to calculate power output at different time steps in a given transient operating condition. This section first presents schematic of the ORC-WHR system, followed by details of components models.

As shown in **Fig. 1**, the ORC-WHR system utilized in this study only considers TP exhaust as heat source. For this study, a 13L heavy-duty diesel engine is modeled in GT-POWER. For a detailed description of this model, readers are advised to refer [20]. The ORC contains high pressure region and low pressure region. Working fluid pump creates the high pressure region and circulates the working fluid through the ORC system. The feed pump supplies working fluid for the high pressure pump to avoid cavitation effect and protect the high pressure pump. Evaporator extracts heat from exhaust gas and evaporates working fluid from liquid to vapor phase. The high pressure vapor passes through the turbine expander and generates electricity from turbine shaft using a generator. Working fluid then flows to the low pressure region, releases heat in condenser and turns back to liquid phase. Expansion tank downstream of condenser acts as a working fluid buffer during transient and start/ stop operating conditions. Ethanol is chosen as working fluid. The working fluid selection requires a systematic process, but it is not the focus on this study.

The rest of this section mainly describes the component modeling of the ORC system, including high pressure pump model, TP evaporator model, turbine bypass valve, turbine valve model, and turbine model. Three components in the lower pressure region (i.e. condenser, expansion tank and feed pump) are not considered in this modeling section. Working fluid temperature at high pressure pump inlet is assumed to be fixed at 30 °C. The models are built in Mathworks Simulink®. Details of modeling work are given below:

2.1. Evaporator

Governing equations of the evaporator can be represented by partial differential equations (PDEs). Following assumptions are made in the evaporator model: (i) The TP exhaust gas mass flow rate across the heat exchanger is assumed to be steady state (i.e., the mass flow rate at the evaporator inlet is the same as the outlet). The TP exhaust gas mass flow rate varies with time. (ii) The radial and axial heat conduction is neglected inside all three mediums (TP exhaust gas, working fluid and the wall separating exhaust gas/working fluid). (iii) Pressure drop across the heat exchanger is neglected. The governing equations of the evaporator include working fluid mass balance, energy balance, wall energy balance, and exhaust gas energy balance. These four PDEs are shown in Eqs. (1)–(4) respectively. In Eq. (1), the term $\frac{\partial(A_f \rho_f)}{\partial t}$ represents working fluid mass change in the volume with unit length in axial direction and unit time. The term $\frac{\partial m_f}{\partial z}$ represents mass flow rate difference at each end of the unit length in axial direction. In Eq. (2),

the first term on the left side ($\frac{\partial(\rho_f A_f h_f)}{\partial t}$) represents heat change in the unit volume at unit time. The term $\frac{\partial(m_f h_f)}{\partial z}$ represents heat flow rate difference at each end of the unit length in axial direction. The term $\pi d_{f,w} U_{f,w} (T_w - T_f)$ represents heat flow from wall to working fluid in the unit volume at unit time. Three terms in Eq. (3) have similar meanings corresponding to the terms in Eq. (2) and the only difference is the medium as Eq. (2) applied to working fluid and Eq. (3) applied to exhaust gas. In Eq. (4), the term $\frac{\partial(\rho_w C_{p_w} A_w T_w)}{\partial t}$ represents the heat change in the unit wall volume at unit time between the exhaust gas and working fluid. The term $\frac{\partial(k_w A_w)}{\partial z}$ represents heat conduction rate difference at each end of the unit length in axial direction in the wall. Two terms on the right hand side represent heat transfer between wall and the other two mediums, which are the same as the two terms on the right hand side of Eqs. (2) and (3).

$$\frac{\partial(A_f \rho_f)}{\partial t} + \frac{\partial m_f}{\partial z} = 0 \quad (1)$$

$$\frac{\partial(\rho_f A_f h_f)}{\partial t} + \frac{\partial(m_f h_f)}{\partial z} = \pi d_{f,w} U_{f,w} (T_w - T_f) \quad (2)$$

$$\frac{\partial(\rho_e C_{p_e} A_e T_e)}{\partial t} + \frac{\partial(m_e C_{p_e} T_e)}{\partial z} = \pi d_{e,w} U_{e,w} (T_w - T_e) \quad (3)$$

$$\begin{aligned} \frac{\partial(\rho_w C_{p_w} A_w T_w)}{\partial t} + \frac{\partial(k_w A_w)}{\partial z} \\ = \pi d_{f,w} U_{f,w} (T_f - T_w) + \pi d_{e,w} U_{e,w} (T_e - T_w) \end{aligned} \quad (4)$$

Finite volume approach is a popular approach to solve evaporator PDEs thanks to its high accuracy [21]. As shown in **Fig. 2**, finite volume model (FVM) composes of n volumetric uniform cells. Exhaust gas releases heat to wall and heat evaporates working fluid from liquid to vapor phase. In each cell, Eqs. (1)–(4) are solved. In general, n is greater than 10 to ensure model accuracy [21]. DP has the curse of dimensionality issue that computation cost increases exponentially with number of states. n cells include $4n$ equations, which makes FVM computationally inefficient for online applications. Thus, FVM cannot be directly utilized in DP algorithm. Alternative evaporator models are moving boundary model (MBM) and 0D model. The main concept of MBM is to calculate locations of working fluid phase boundaries, based on which the evaporator is separated into three sections (i.e. liquid phase, mixed phase, and vapor phase). Its number of states generally varies from 6 to 9, depending on assumptions. MBM has comparable vapor temperature prediction accuracy with FVM and its computation cost is much lower [22]. Thus, MBM is generally implemented in real-time model-based control, such as model predictive controls [23]. Even

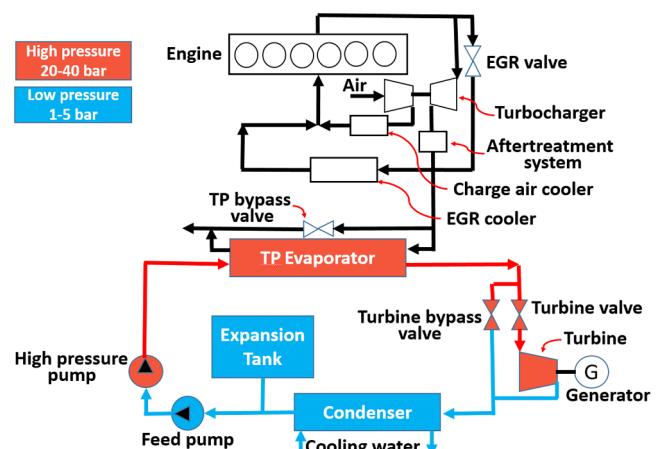


Fig. 1. Schematic of ORC-WHR system. TP evaporator locates downstream of aftertreatment system.

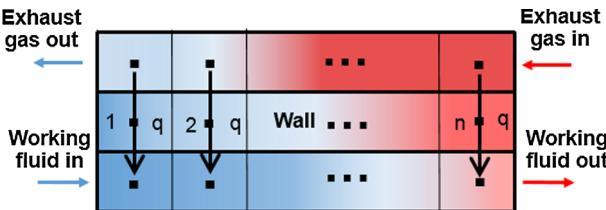


Fig. 2. Schematic of finite volume method. The working fluid flow and exhaust gas flow are in reverse direction.

though computation cost of MBM is low compared with FVM, it is still far from DP application and no MBM is used in DP optimization in literature. In terms of computation cost, 0D model has extreme low number of states (2 or 3 states) and thus low computation cost. 0D model can be found in DP application [17] and ORC-WHR analysis [24]. The main concept of 0D model is considering each medium as a single cell in the evaporator and there is no discretization in axial direction. It is the same as single cell FVM. However, its model accuracy is compromised in comparison to computation cost. In this study, TP evaporator model is built via the Proper Orthogonal Decomposition (POD) Galerkin projection method [25]. It is derived from the FVM, and thus takes advantage of high accuracy characteristic from FVM. In addition, its number of states can be as low as one, which reduces its computation cost.

The POD Galerkin reduced order model (ROM) derivation procedures are shown in Fig. 3. The POD ROM derivation requires snapshots, which are state parameters value over a period of time. Finite volume heat exchanger model is utilized to generate the snapshots via simulation. Assume there are b time steps in simulation, the snapshot for working fluid enthalpy has the dimension of $\mathbb{R}^{n \times b}$ and the wall temperature snapshot shares the same dimension. POD analysis then extracts basis functions from each type of state snapshots (working fluid enthalpy type and wall temperature type). Singular value decomposition (SVD) is considered as the POD analysis method. Finally, Galerkin projection derives the reduced order model with help of basis functions. More information about the POD analysis and Galerkin projection can be found in [25]. Temporal dynamics in exhaust gas $\partial(A_{e,cross}\rho_e c_p e T_e)/\partial t$ in Eq. (3) are neglected due to their fast transient characteristics. After the derivation, final form of the reduced order model are shown in Eqs. (5)–(7):

$$m_f \dot{x}_{f,k}(t) = -\dot{m}_f(0)L \left[\left(\sum_{i=1}^{q_f} x_{f,i}(t) \phi_{f,i}(L) \right) \phi_{f,k}(L) - x_{f,in} \phi_{h_f,k}(0) \right. \\ \left. - \sum_{i=1}^{q_f} x_{f,i}(t) \left(\sum_{j=1}^n \phi_{f,i,j} \dot{\phi}_{f,k,j} \right) \right] + A_{f,w} \sum_{j=1}^n U_{f,j} \phi_{f,k,j} \left(\sum_{i=1}^{q_w} x_{w,i}(t) \phi_{w,i,j} - T_{f,j}(t) \right) \quad (5)$$

$$m_w C_{pw} \dot{x}_{w,k}(t) = -A_f \left(\sum_{i=1}^{q_w} x_{w,i}(t) \left(\sum_{j=1}^n \phi_{T_{w,i},j} U_{f,j} \phi_{w,k,j} \right) \right. \\ \left. - \sum_{j=1}^n U_{f,j} T_{f,j}(t) \phi_{w,k,j} \right) - A_{e,w} U_e \left(x_{w,k}(t) - \sum_{j=1}^n T_{e,j}(t) \phi_{w,k,j} \right) \quad (6)$$

$$\frac{\partial(\dot{m}_e c_p e T_e)}{\partial z} - \pi d_e U_{e,w} (T_w - T_e) = 0 \quad (7)$$

where x_f and x_w are POD ROM states relating to working fluid and wall, respectively, q_f , q_w represent dimension of working fluid enthalpy and wall temperature, respectively, ϕ_f , ϕ_w represent basis functions of working fluid enthalpy states and wall temperature states, respectively. $\Phi_f = (\phi_{f,1}, \dots, \phi_{f,q_f}) \in \mathbb{R}^{n \times q_f}$ and $\Phi_w = (\phi_{w,1}, \dots, \phi_{w,q_w}) \in \mathbb{R}^{n \times q_w}$. Basis functions are calculated from finite volume model simulation using SVD method. An example of the basis functions is shown in Fig. 4, where the dominating basis functions from working fluid states and wall states are plotted (i.e. first column of matrix Φ_f and matrix Φ_w).

The POD ROM is validated with experimental data under transient

conditions. In the experimental setup, the ORC system is integrated downstream the aftertreatment system of a 13L heavy-duty diesel engine. The dyno is AVL 480 kW double end A/C type. The dyno and engine are monitored by an AVL PUMA system. ORC actuators controls are achieved via a dSPACE MicroAutobox. The signal communication diagram in the test cell is shown in Fig. 5. In POD ROM validation, the stand-alone POD ROM is validated given measured boundary conditions, which include working fluid inlet temperature, inlet mass flow rate, inlet/outlet pressure, exhaust gas inlet temperature, inlet mass flow rate. Outputs of the simulation and experiment include working fluid outlet temperature and exhaust gas outlet temperature. The POD ROM is solved with fixed time step differential method and the time step is 1.0 s.

Validation case 1: In this validation case, working fluid vapor temperature experiences step changes and engine operates at steady state condition (1575 rpm, 1540 Nm). Measured inputs are shown in Fig. 6(a-c). Exhaust gas mass flow rate and temperature at the evaporator inlet are constant due to steady state engine operating condition. Working fluid mass flow experiences multiple step changes to achieve working fluid vapor temperature 20 °C step changes shown in Fig. 6(b) and Fig. 6(c). Working fluid temperature at evaporator inlet is constant around 44 °C, which is not largely impacted by vapor temperature step changes thanks to excellent cooling effect in condenser. Fig. 6(c) shows average working fluid pressure of inlet and outlet. In the experiment, the pressure is controlled by turbine bypass valve and turbine inlet valve is fully closed. The pressure is closely related to the working fluid mass flow rate variation. In Fig. 6(d), the working fluid vapor temperature at evaporator outlet between simulation and experiment show well alignment with limited discrepancy. In addition, exhaust gas temperature exhibit small discrepancy.

Validation case 2: In this validation case, engine speed and torque experience a step change as shown in Fig. 7, while the working fluid vapor temperature stays constant. As shown in Fig. 8(a), the measured engine exhaust gas mass flow rate decreases nearly 50% after the engine speed and torque step change. Due to the working fluid evaporator locates downstream the heavy-duty diesel engine aftertreatment system, which has large thermal inertia, exhaust gas temperature variation at the evaporator inlet is slow compared to mass flow rate. In order to maintain the working fluid temperature during the engine step change, the working fluid mass flow rate drops significantly at 150 s in Fig. 8(b). In Fig. 8(d), simulation and experiment results only show small difference.

2.2. Turbine

The turbine is designed for high expansion ratio, in which condition the turbine operates in choke flow. Thus, the mass flow rate has linear relation with the inlet pressure as follows:

$$\dot{m}_{turb} = a_{turb} p_{turb,in} + b_{turb} \quad (8)$$

where the $a_{turb} = 2.43 \times 10^{-8}$, and $b_{turb} = -3.3 \times 10^{-3}$. The coefficients a_{turb} and b_{turb} are derived from experimental data. Working fluid enthalpy and temperature at turbine outlet are calculated as follows:

$$h_{turb,out} = h_{turb,in} - \eta_{turb,is} (h_{turb,in} - h_{turb,is,out}) \quad (9)$$

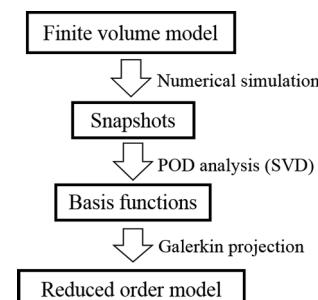


Fig. 3. POD Galerkin reduced order model derivation procedure.

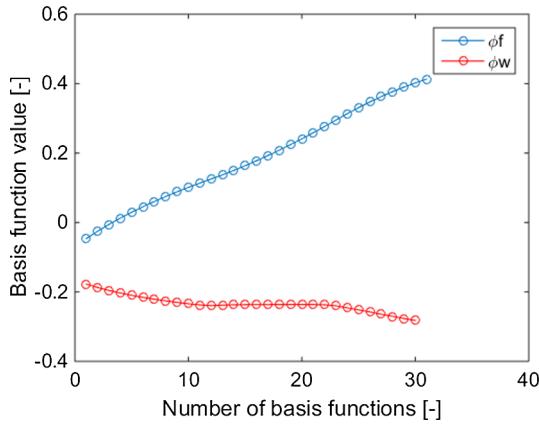


Fig. 4. Basis functions example when $q_f = 1$ and $q_w = 1$.

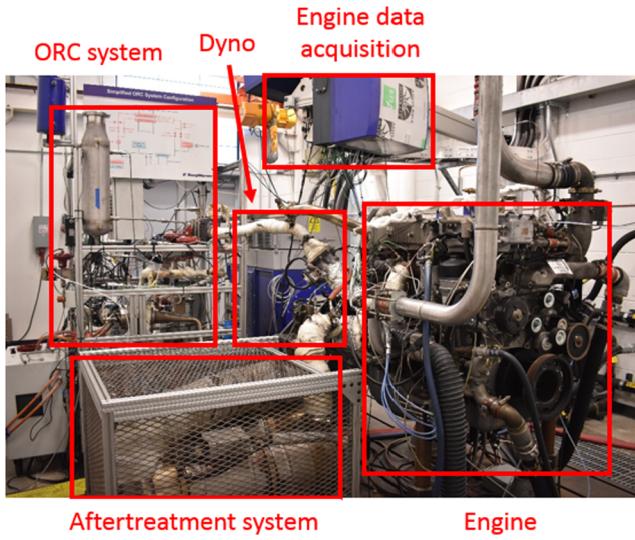


Fig. 5. Experiments configurations for ORC-WHR testing.

$$T_{turb,out} = f(h_{turb,out}, P_{turb,out}) \quad (10)$$

where the right hand side of Eq. (10) is the Ethanol thermodynamic map shown in Fig. 9. The isentropic efficiency $\eta_{turb,is}$ is interpolated via a map provided by turbine manufacturer. The inputs to the map include turbine speed, expansion ratio and inlet temperature. The efficiency map is not shown due to the confidential content. The turbine outlet pressure $P_{turb,out}$ is assumed constant at 1.4 bar.

2.3. Pump

Both the high pressure pump and the feed pump in Fig. 1 are positive displacement pumps. The functionality and construction of these two pumps are the same. Thus, only one pump model is given. Mass flow rate is linearly correlated to the pump speed in Eq. (11). Coefficients, $b_{pump} = 1.75 * 10^{-3}$, which are fitted from manufacturer pump data. Pump outlet temperature and power consumption are calculated in Eqs. (12) and (13). Isentropic efficiency is given in Eq. (14) [6].

$$\dot{m}_{pump} = a_{pump} N_{pump} + b_{pump} \quad (11)$$

$$T_{pump,out} = T_{pump,in} + \frac{(1 - \eta_{pump,is}) P_{pump}}{\dot{m}_{pump} c_{p,pump}} \quad (12)$$

$$W_{pump} = \frac{\dot{m}_{pump}}{\rho} \frac{(p_{pump,out} - p_{pump,in})}{\eta_{pump,is}} \quad (13)$$

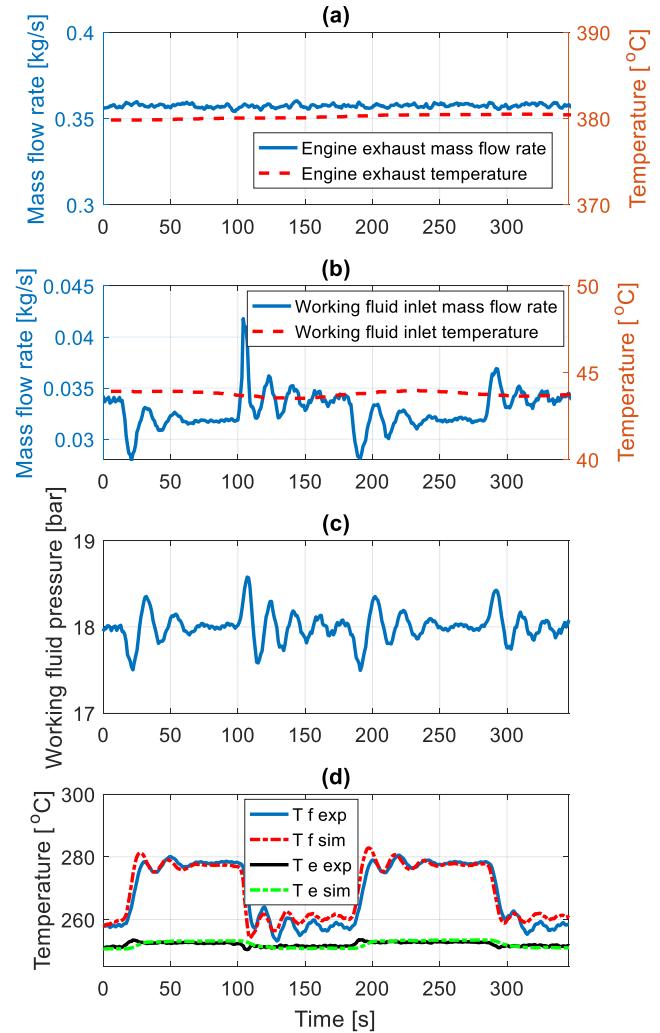


Fig. 6. Validation case 1: measured inputs from experiments and outputs: (a) exhaust gas mass flow rate and temperature, (b) working fluid mass flow rate and temperature at heat exchanger inlet, (c) working fluid pressure in heat exchanger, and (d) working fluid vapor temperature. (sim – simulation, exp – experiment).

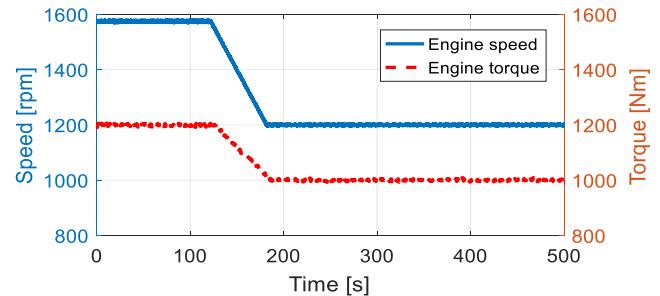


Fig. 7. Validation case 2: measured engine speed and torque step changes in experiments.

$$\eta_{pump,is} = 0.93 - 0.11 \log \left(\frac{\dot{m}_{pump}}{\dot{m}_{pump,max}} \right) - 0.2 \log \left(\frac{\dot{m}_{pump}}{\dot{m}_{pump,max}} \right)^2 - 0.06 \log \left(\frac{\dot{m}_{pump}}{\dot{m}_{pump,max}} \right)^3 \quad (14)$$

2.4. Valve

Turbine inlet valve and bypass valve ensure safe and efficiency operation of turbine expander. During ORC system warmup, there is working fluid liquid at turbine inlet. Turbine bypass valve is fully open

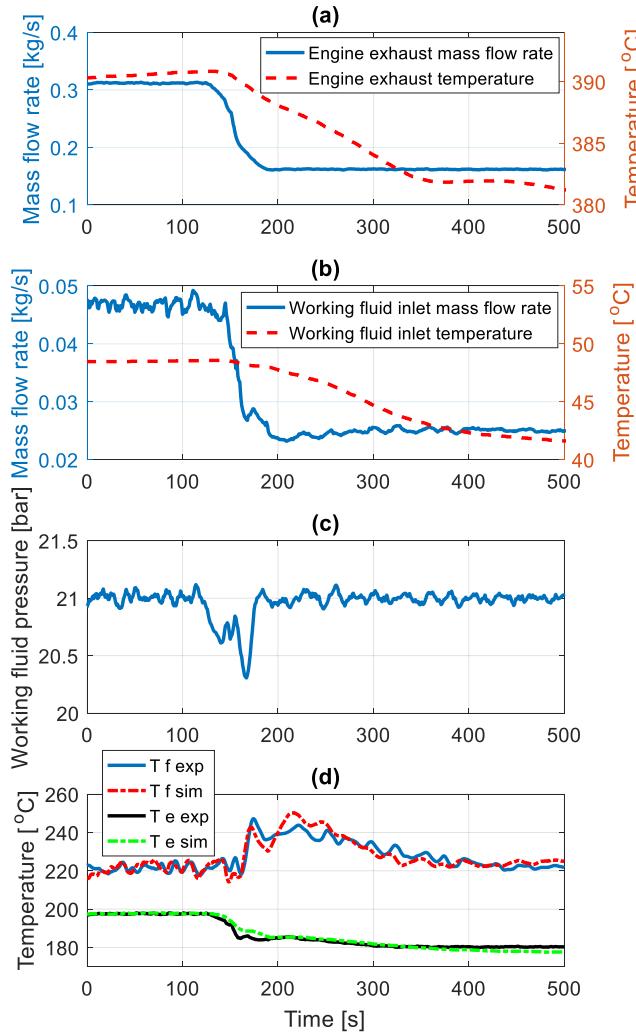


Fig. 8. Validation case 2: measured input from experiments and output: (a) exhaust gas mass flow rate and temperature, (b) working fluid mass flow rate and temperature at heat exchanger inlet, (c) working fluid pressure in heat exchanger, and (d) working fluid vapor temperature.

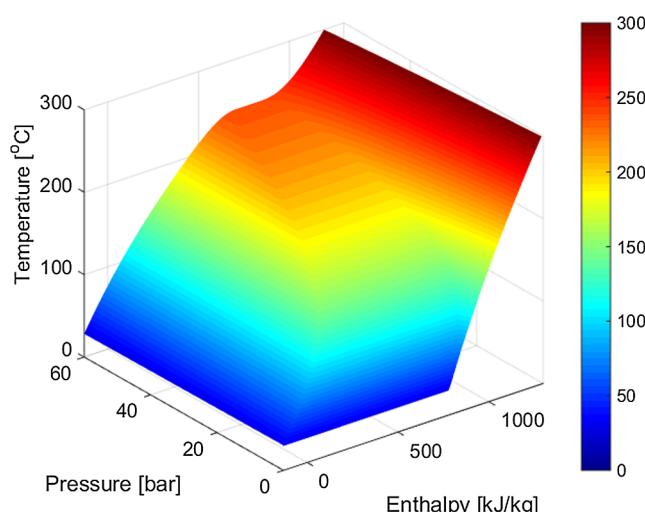


Fig. 9. Thermodynamic map of Ethanol temperature as a function of pressure and enthalpy.

and turbine inlet valve is fully close. This operation avoids turbine spinning at wet working fluid condition and protects turbine blades. When ORC system is fully warm and working fluid reaches certain superheat at turbine inlet, turbine inlet valve turns to open and turbine bypass valve turns to close. Turbine then starts producing power. During normal operation, when working fluid vapor quality drops below 1.0, turbine inlet valve and bypass valve switch open and close positions for turbine protection. In addition, turbine bypass valve opens when turbine inlet pressure is above system safety pressure. Turbine inlet valve and bypass valve is modelled based on pressure ratio as follows:

$$\left\{ \begin{array}{l} \text{If } \left(\frac{2}{\gamma+1} \right)^{\frac{\gamma}{\gamma-1}} \leq \frac{p_{out}}{p_{in}} \leq 1 \text{ (subsonic):} \\ \dot{m} = C_d A_{Op} \sqrt{\frac{2\gamma}{\gamma-1} p_{in} \rho_{in} \left[\left(\frac{p_{out}}{p_{in}} \right)^{\frac{2}{\gamma}} - \left(\frac{p_{out}}{p_{in}} \right)^{\frac{\gamma+1}{\gamma}} \right]} \quad (a) \\ \text{If } 0 \leq \frac{p_{out}}{p_{in}} \leq \left(\frac{2}{\gamma+1} \right)^{\frac{\gamma}{\gamma-1}} \text{ (supersonic):} \\ \dot{m} = C_d A_{Op} \left(\frac{2}{\gamma+1} \right)^{\frac{\gamma+1}{2(\gamma-1)}} \sqrt{\gamma p_{in} \rho_{in}} \quad (b) \end{array} \right. \quad (15)$$

$$C_d = (a_1 O^2 + a_2 O + a_3) \cdot \left(\frac{p_{in}}{a_4} \right) \quad (15c)$$

$$\left\{ \begin{array}{l} a_1 = -1.109e^{-5} \\ a_2 = 1.397e^{-5} \\ a_3 = 3.376e^{-6} \\ a_4 = 2.1e6 \end{array} \right. \quad (15d)$$

The discharge coefficients are calibrated with experimental data and the calibrated parameters are shown in Eq. (15d). If valve upstream pressure is in the 5–60 bar range and valve downstream pressure is around 1.4 bar, the pressure ratio $\frac{p_{out}}{p_{in}}$ will vary only from 0.047 to 0.28,

whereas the value of $\left(\frac{2}{\gamma+1} \right)^{\frac{\gamma}{\gamma-1}}$ is around 0.5–0.6 in normal operating conditions. Thus, Eq. (15b) dominates.

3. Optimization problem formulation

In ORC-WHR system, a generator is connected to turbine expander to produce electricity. High pressure pump consumes power. Condenser cooling circuit power consumption and turbine inlet/bypass valve actuation power are ignored in this study. The input of the DP problem is pump speed, which is the only independent variable in this DP optimization. The cost function of the DP is the net power production by the ORC-WHR system, which is the difference between turbine generated power and power consumed by the working fluid pump. The goal of the optimization is to maximize the power output from the system. The cost function is given as follows:

$$J = \int_{T_0}^{T_{end}} -(W_{turb} - W_{pump}) d\tau \quad (16)$$

To minimize the cost function, the optimization problem is formulated in Eq. (17) and the whole system dynamics are expressed in Eqs. (5)–(15):

$$\min J(x(t), u(\cdot)) \quad (17)$$

Table 1
Lower and upper boundaries of states, outputs, and inputs.

| Parameters | Value | Parameters | Value |
|-----------------|--|-----------------|--------------------------------|
| x^{lb} | 2e5 | x^{ub} | 3.3e5 |
| y^{lb} | $T_{sat} + 5 \text{ }^{\circ}\text{C}$ | y^{ub} | 330 $\text{ }^{\circ}\text{C}$ |
| u^{lb} | 100 rpm | u^{ub} | 1500 rpm |
| δu^{lb} | -500 rpm/s | δu^{ub} | 500 rpm/s |

$$\left. \begin{array}{l} \dot{x} = f(x, u), \text{ Eq. (5)} \\ y = g(x, u), \text{ Eqs. (7 - 15)} \\ x^{lb} \leq x(t) \leq x^{ub} \\ y^{lb} \leq y(t) \leq y^{ub} \\ u^{lb} \leq u(t) \leq u^{ub} \\ \delta u^{lb} \leq \dot{u}(t) \leq \delta u^{ub} \\ x = h_{f,k} \\ u = N_{pump} \end{array} \right\} s, t:$$

where $\dot{x} = f(x, u)$ represents system dynamics, x represents $x_{f,k}(t)$ in Eq. (5). The dynamics of $x_{w,k}(t)$ in Eq. (6) is neglected, which results in 1-state evaporator model. The details of equation derivation about Eqs. (5)–(7) can be found in [25]. $y = g(x, u)$ represents the system output, which is the working fluid temperature. All the rest of equations are utilized to calculate the net power generation. Superscripts lb/ub represent lower/upper bounds, u is input, δu is input change rate (see Table 1). x^{lb} is determined by ambient condition (25 $\text{ }^{\circ}\text{C}$, 1 bar) and x^{ub} is determined by high pressure high temperature condition (330 $\text{ }^{\circ}\text{C}$, 40 bar). y^{lb} is determined by turbine operating vapor quality limits and y^{ub} is determined by decomposition temperature of the working fluid (ethanol in this study). u^{lb}/u^{ub} are determined by waste power at engine idle and waste power at rated engine operating condition. $\delta u^{lb}/\delta u^{ub}$ are determined by measured step response speed of the working fluid pump. The cost function J is generally the net power output from the ORC-WHR system. The only parameter that DP algorithm can make decision is the working fluid pump speed u . As the pump speed changes, the working fluid mass flow rate changes accordingly, which then affects working fluid enthalpy and temperature inside the evaporator.

4. Optimization utilizing Dynamic Programming

4.1. Dynamic Programming

Dynamic Programming (DP) is mostly used as an offline optimization algorithm to generate a benchmark for control performance evaluation or generate a reference trajectory for feedback controls. DP is a solution method for complex/ nonlinear problem by breaking down the complex problem into simpler problems. The advantage of DP is its global optimal characteristic. However, the computation time increases exponentially with number of states, which is the main reason why DP is mostly implemented offline. In this study, the POD ROM control model only has one state.

The time step of the ORC WHR model is set at 0.4 s. Therefore, a 1200 s Federal Test Procedure (FTP) driving cycle produces 3000 pairs of inputs and outputs. In the model, there is only one state x_f and the initial condition is given. The final condition is constrained by lower and upper bounds. The state is discretized into four values and the time is discretized by 0.4 s in this study.

4.2. Optimization

The FTP driving cycle engine condition is shown in Fig. 10. From Fig. 10(a), this driving cycle contains large range of heavy-duty diesel engine speed and torque, which covers majority of heavy-duty diesel

engine operating conditions. Fig. 10(b) and (c) show TP exhaust gas mass flow rate and temperature downstream of aftertreatment system from the GT-POWER engine model. The mass flow rate is very dynamic in the range of 0.1 kg/s to 0.4 kg/s. The temperature shows relatively slow response due to the large thermal inertial in the aftertreatment system. The exhaust gas mass flow and temperature serve as exogenous inputs to the TP evaporator and DP algorithm knows the exhaust conditions in advance.

The DP optimal working fluid vapor temperature, working fluid pump speed, normalized net power production and normalized accumulative energy are shown in Fig. 11.

The net power is mainly affected by the turbine generated power, which is proportional to the working fluid mass flow rate (i.e. working fluid pump speed for the displacement pump in this study). From Fig. 7(b) it can be seen that the DP algorithm maintains the vapor temperature very close to the saturation temperature for maximizing turbine power. Therefore, as the exhaust gas mass flow increases, the working fluid pump speed increases and vice-versa. This can also be confirmed from Fig. 11(a) and Fig. 11(c), where the working fluid pump speed and net power production have similar trend with respect to the exhaust gas mass flow rate shown in Fig. 10(b). This is because the heat transfer between the exhaust gas and working fluid changes in accordance with exhaust gas mass flow.

The accumulative energy is calculated through Eq. (18). In Fig. 11(d), accumulative energy is normalized by its maximum value. The subplot shows drastic increase between 600 s and 900 s. During this period, the exhaust gas mass flow rate is large and exhaust gas temperature is high in Fig. 10.

$$E_{acc} = \sum_{k=0}^{k=N} (W_{turb}(k) - W_{pump}(k)) \Delta t \quad (18)$$

PID control is implemented in the same ORC-WHR model and driving cycle as the reference of DP performance. Working fluid temperature of PID control is shown in Fig. 12(a). PID vapor temperature shows larger variation compared with DP vapor temperature as shown in Fig. 11(b). The large variation results in working fluid saturation in some parts of the driving cycle, which leads to turbine shutdown and no power production as shown in Fig. 12(b). In comparison, DP keeps producing power over the entire driving cycle and the net power is always higher than the PID control thanks to the higher mass flow rate and lower vapor temperature. Over the driving cycle, DP produces 19.7% more energy than PID does.

In summary, DP algorithm shows excellent performance in terms of turbine operation duration and power production. With the DP algorithm, the turbine operation is not disturbed in the middle of driving cycle as working fluid vapor temperature does not touch saturation temperature. In comparison, PID control leads to several turbine shutdowns due to the working fluid temperature saturation at the turbine inlet. The net power shows huge gap during those turbine shutdowns in Fig. 12(b). Even though the PID produces power most of time, DP net power still outperforms the PID net power. The DP results in Fig. 11 is not real-time implementable. To address this real-time implementation problem, DP rule extraction is conducted in the following section. The extracted rules can then be implemented in real-time with low computation cost in comparison to implementing the DP algorithm itself.

5. Dynamic Programming rule extraction

This section describes the proposed DP rule extraction method. In literature, if DP rules were implemented in feedback control, the rules were mostly saved in table or correlated with nonlinear functions. If the rules were saved in table, the table would be complicated with high input dimension, especially when input dimension is greater than 3. Whereas, if the rules were correlated with nonlinear functions, the nonlinear function selection processes are time-consuming and there is no guidance for the selection. In addition, ORC-WHR system is a highly

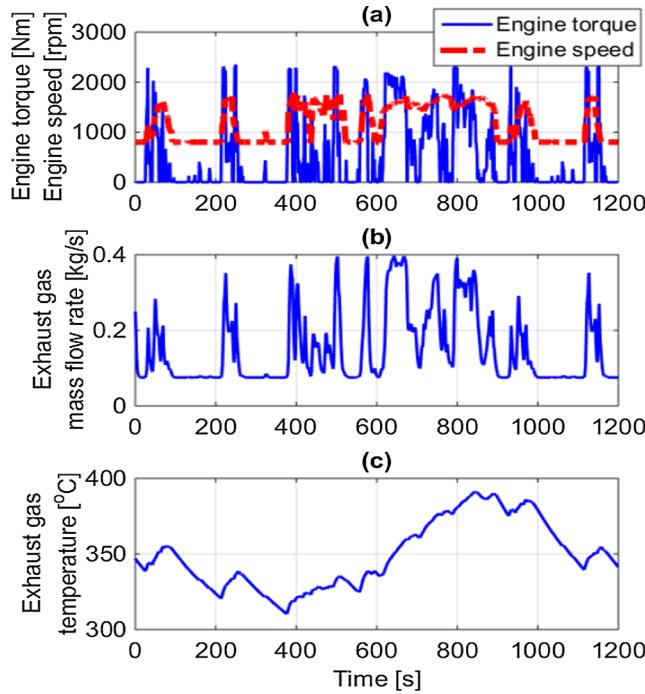


Fig. 10. Exhaust gas conditions from FTP heavy-duty diesel engine driving cycle: (a) engine speed and torque, (b) exhaust gas mass flow rate, and (c) exhaust gas temperature.

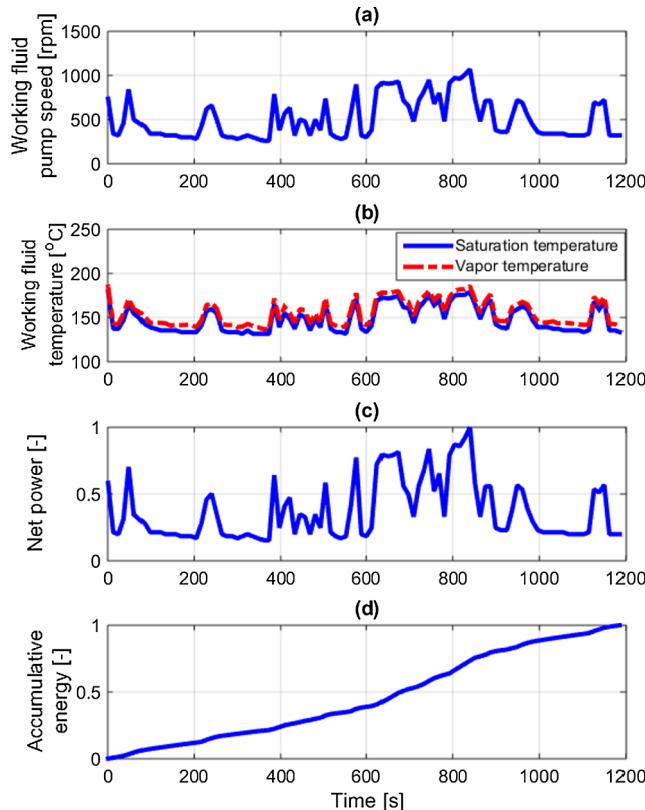


Fig. 11. DP results over FTP driving cycle: (a) working fluid pump speed, (b) working fluid vapor temperature at heat exchanger outlet, (c) normalized net power generation from the ORC-WHR system, and (d) normalized accumulative energy.

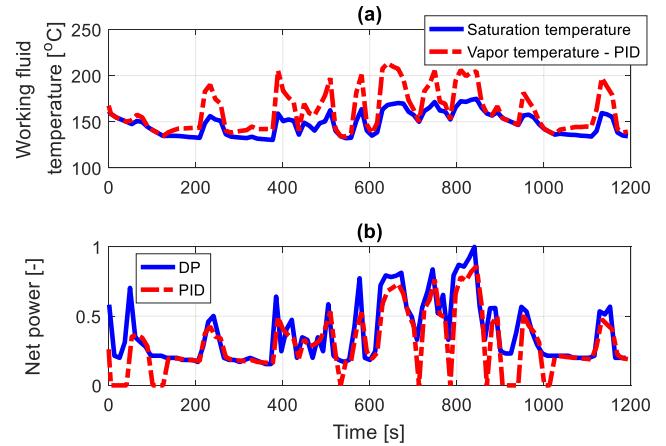


Fig. 12. DP and PID comparison for FTP driving cycle: (a) working fluid vapor temperature, and (b) net power from ORC system.

nonlinear system, which adds challenge to the nonlinear fitting. This study proposes a new method, which utilizes nonlinear machine learning algorithm to learn the rule from DP results. The advantage of the machine learning models is their high nonlinearity, which fits complex system problems. Machine learning algorithms were implemented in regression problems from nonlinear system, such as lithium-ion battery capacity prediction [26], solar thermal energy prediction [27], wheat biomass prediction [28], etc.

In ORC-WHR power optimization, machine learning algorithms have not been utilized for rule extraction in existing literature. Rule extraction is a regression problem owing to the continuous output of the optimal working fluid pump speed. There are more than 20 types of regression models and this study chooses the models based on their popularity in literatures over the past decade. In this study, eleven state-of-art machine learning regression models are compared, which include Linear Regression (LR) [29], Ridge Regression (RR), Kernel Ridge Regression (KRR), K-Nearest Neighbor (KNN), Gaussian Process Regression (GPR), Decision Tree (DT), Random Forest (RF), Bagged Trees (BT), Adaboost (ADB), Xgboost (XGB), and Artificial Neural Networks (ANN). A brief description of these algorithms is given in Appendix A.

The datasets utilized in the algorithm selection are the DP results for FTP driving cycle shown in Fig. 11. For all machine learning models, there are two inputs, which are TP exhaust gas mass flow rate and TP exhaust gas temperature. The output is the working fluid pump speed. The model predicts the optimal working fluid pump speed based on two inputs. For each algorithm, the parameters are swept in the 5-fold cross validation [30] and the parameters corresponding to the lowest mean validation error are kept.

The results are shown in Fig. 13. The algorithms are ranked based on the relative error in the cross validation. There is not significant error difference among the top 7 models. Interestingly, 5 out of top 7 algorithms are tree based methods, except for KRR and GPR, which are kernel based methods. Ensemble trees (RF-top1, BT-top4) and boost trees (ADB-top3, XGB-top5) algorithms exhibit less error than the single tree (DT-top7) algorithm thanks to the anti-overfitting features of the ensemble trees. Linear algorithms (LR and RR) exhibit large error because this ORC-WHR problem is a nonlinear problem. Even though Neural Network (ANN) shows slightly better performance than the linear algorithms (LR, RR), all these three algorithms show obvious error gap ($> 2\%$) with the top 7 algorithms. Therefore, even though the most recent artificial intelligence revolution is reportedly brought out by the Neural Networks, its application in the current WHR-ORC system

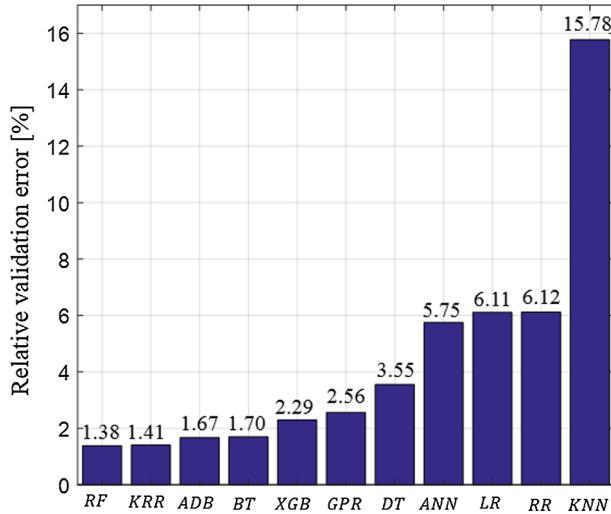


Fig. 13. Mean relative error of working fluid pump speed prediction from 5-fold cross validation.

does not outperform some other machine learning algorithms such as RF and KRR. One possible reason is that the size of the dataset in traditional engineering fields is much smaller than other fields like image processing, voice recognition, and translation (e.g. the Stanford and Princeton image processing database “imageNET” contains 14 millions of images [31]). KNN has the worst performance among the eleven algorithms. The reason could be the large magnitude difference between features (mass flow rate and temperature). When the distance is calculated, the temperature (K) often has much larger impact on the distance than the mass flow rate (kg/s). Based on the performance of different algorithms, RF is selected in this study.

RF regressor has low overfitting and high robustness characteristics [32], which reduce its training effort compared with Neural Networks regressor. The RF method considers multiple decision trees, which are trained with randomly selected inputs and sub-datasets. Fig. 14 shows a simple RF tree example used in this study. In this example, there are N decision trees. In each tree, there are multiple nodes, which separate the path to the next nodes level. The inputs enter the tree from the top. The inputs in this example is the exhaust gas mass flow rate and exhaust gas temperature. The path which the inputs follow is determined by the judgement condition in each node. The judgment condition threshold is identified in the model training process. The target is the optimal working fluid pump speed.

Besides the TP exhaust gas mass flow rate and temperature, the inputs can be expanded based on the existing inputs. A time delayed process is implemented to mass flow rate and temperature. The TP exhaust gas mass flow rate is delayed by one time step, two time steps and three steps to create three new inputs. Similarly, three new temperature inputs are created. Overall, six new inputs are added on top of existing two inputs, which forms eight inputs to the RF model. The added inputs are like the memories of the mass flow and temperature history to give the model the moving direction and slope information. For the thermal system, the system response could be slow. The exhaust gas release the heat to the wall and then the wall heat up the working fluid. As shown in Fig. 15, the working fluid vapor temperature takes nearly 50 s to settle down at the TP exhaust gas mass flow rate 5% step change. This feature engineering tries to inject some history values of TP exhaust gas mass flow rate and temperature so that the machine learning models can detect the change direction and rate of these two types of inputs. The number of time step delays are optimized by grid

search and the three time delays are selected based on the optimization results.

The RF model utilized in this study is from the open-source sklearn library written in Python [33]. The main hyper-parameters of RF model include number of trees, maximum tree depth, and minimum samples to be a leaf node. These hyper-parameters should be optimized to improve the RF model prediction accuracy. In literature, grid search and random search are two popular methods utilized in the hyper-parameter optimization [34]. Random search method is utilized in this study for its better performance than grid search given limited amount of computation power. As the comparison, the inputs with and without time delays are considered in the hyper-parameter optimization. The lower and upper boundaries setup and results of different inputs comparison are given in Table 2. Iterations of both optimization are set to be 100. The lowest validation error percentage from the two-inputs RF model is 1.08%, while the eight-inputs RF model is 0.85%. The validation error decreases 22% by adding the six time delayed mass flow rate and temperature inputs to the RF model, which reveals that the delayed inputs do assist the RF model prediction. The iteration process is shown in Fig. 16. The validation error percentage of the best visited results of the two optimization cases are shown in Fig. 16. In both optimization cases, the validation error varying range is as large as 5–7%, which reveals that hyper-parameter optimization could substantially improve the RF model prediction accuracy.

The contribution of each input to the prediction accuracy can be indicated by feature importance. Feature importance is calculated by comparing the error increase of different cases that the corresponding feature sample values are randomly permuted among all the samples [32]. Feature importance of the eight inputs is shown in Fig. 17. It is observed that the exhaust gas mass flow is the most important feature, followed by one time step delayed mass flow rate. The reason why these

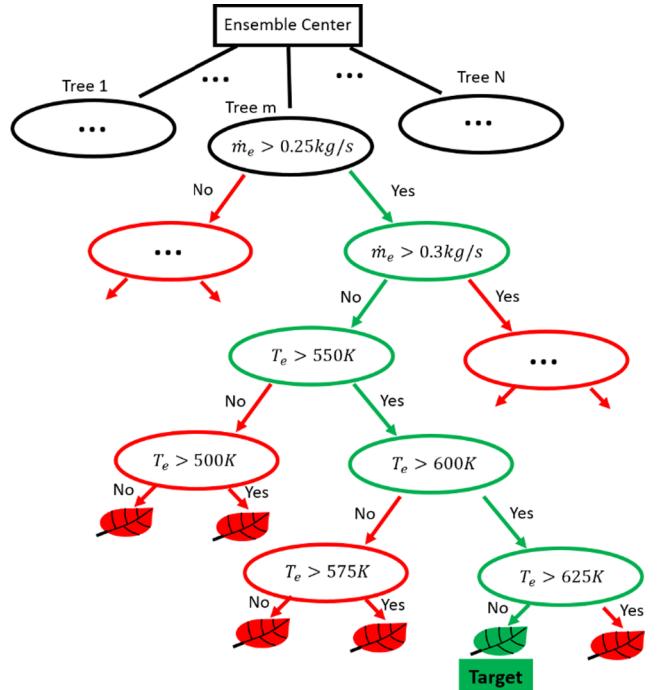


Fig. 14. Example of RF regression method. Ensemble center averages the outputs from all the decision trees. In each tree, the nodes direct an input sample to the target leaf, where the single working fluid pump speed is stored during the RF model training process.

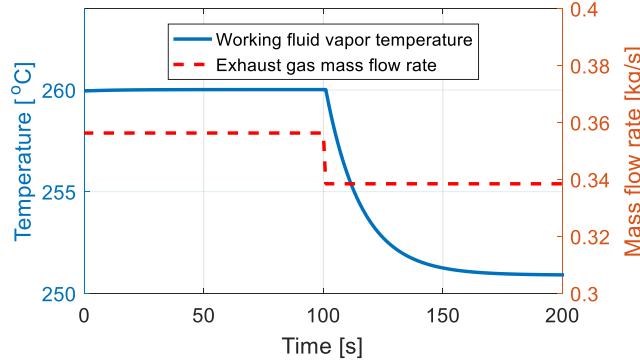


Fig. 15. Open loop simulation results on TP exhaust gas mass flow rate step change.

Table 2

RF model hyper-parameter lower and upper boundaries setup and results. Iterations are set to be 100.

| RF model hyper-parameters | Boundaries setup | | Optimal value | | Relative error | |
|-----------------------------------|------------------|-------|---------------|----------|----------------|----------|
| | Lower | Upper | 2 inputs | 8 inputs | 2 inputs | 8 inputs |
| Number of trees | 1 | 500 | 11 | | 1.08% | 0.85% |
| Maximum tree depth | 1 | 50 | 8 | | | |
| Minimum samples to be a leaf node | 2 | 50 | 1 | | | |

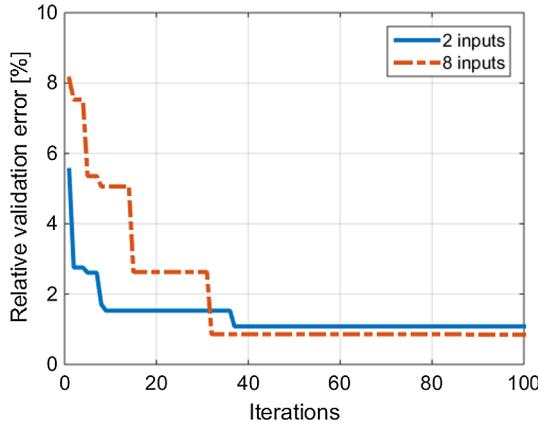


Fig. 16. RF model hyper-parameter optimization results utilizing random search method. y axis represents the lowest validation error percentage up to current iteration from the first iteration. The legend '2 inputs' represents the RF model with two inputs ($\dot{m}_e(k)$, $T_e(k)$), while '8 inputs' represents the RF model with eight inputs ($\dot{m}_e(k-3)$, $\dot{m}_e(k-2)$, $\dot{m}_e(k-1)$, $\dot{m}_e(k)$, $T_e(k-3)$, $T_e(k-2)$, $T_e(k-1)$, $T_e(k)$).

two exhaust gas mass flow rate inputs take up to more than 95% feature importance is because exhaust power is mainly affected by mass flow rate. This can be further explained using Eq. (19), which shows the exhaust gas power calculation. Exhaust gas heat capacity, $C_{p,e}$ and ambient temperature, T_{amb} is assumed to be constant in this study. According to Fig. 10, the exhaust gas mass flow rate is between 0.1 kg/s to 0.4 kg/s and the exhaust gas temperature is between 310 °C (583 K) to 390 °C (663 K). If the ambient temperature is fixed at 298 K, the temperature variation only leads to 28% change of waste power, whereas the mass flow rate change leads up to 400% change of waste power. It is also interesting to note that the one time step delayed mass flow rate passes the non-delayed temperature and ranks second. Most ORC-WHR optimization literature focused on steady state engine conditions,

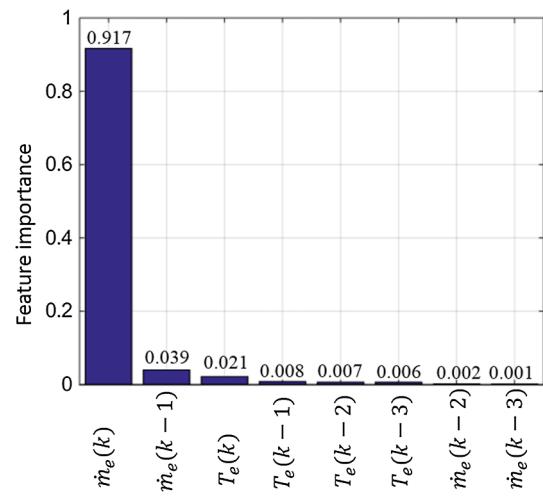


Fig. 17. RF model feature importance ranking based on training data.

where thermal delays were not considered. Thus, the steady state optimization results might not hold true in the transient conditions. The optimization conducted in transient conditions are preferred.

$$\frac{dQ_{waste}}{dt} = \dot{m}_e C_{p,e} (T_e - T_{amb}) \quad (19)$$

Since the presence of liquid droplets in working fluid is detrimental to the turbine operation, a rule-based (RB) method is added on top of the RF model to prevent the appearance of liquid working fluid and the combination of RF model plus RB method is named RF-RB model in this study. The main principle of the RB method is to reduce the working fluid pump speed when vapor temperature is close to saturation temperature (i.e. low superheat temperature). The RB method expression is shown as follows:

$$N_{RF_RB} = \begin{cases} N_{RF}, T_{vap} - T_{sat} \geq T_{sh,1} \\ N_{RF} - \Delta N_1, T_{sh,1} > T_{vap} - T_{sat} \geq T_{sh,2} \\ \vdots \\ N_{RF} - \Delta N_{i-1}, T_{sh,a-1} > T_{vap} - T_{sat} \geq T_{sh,a} \\ N_{RF} - \Delta N_a, T_{vap} - T_{sat} < T_{sh,a} \end{cases} \quad (20)$$

where N_{RF_RB} is the working fluid pump speed calculated by RF model plus RB method, N_{RF} is the pump speed purely calculated by RF model, T_{vap} and T_{sat} are working fluid vapor and saturation temperature, respectively, a is the number of superheat temperature layers to be split, $T_{sh,i}$ is the i^{th} superheat threshold, $i = (1, 2, \dots, a)$, ΔN_j is the j^{th} pump speed reduction, $j = (1, 2, \dots, a)$. In Eq. (20), a is the first parameter to be determined or optimized. After a is determined, there are $2a$ parameters to be optimized, $(T_{sh,1}, T_{sh,2}, \dots, T_{sh,a}, \Delta N_1, \Delta N_2, \dots, \Delta N_a)$. a is selected to be 2 for the saving of computation cost in the parameter optimization. To optimize the $2a$ parameters, the particle swarm optimization method is chosen for its global optimal characteristic and its high efficiency in multi-variable optimization [35]. Details of optimization process are not given and the optimization results are given in Table 3.

Table 3

Particle swarm optimization results on RB parameter in Eq. (20).

| $T_{sh,1}$ | $T_{sh,2}$ | ΔN_1 | ΔN_2 |
|------------|------------|--------------|--------------|
| 4 | 1 | 93 | 61 |

6. Extracted rule validation

In the validation process, RF-RB model is validated over a physics based high-fidelity plant model. The details of the plant modeling, identification/validation can be found in [20].

6.1. Validation setup

The training flowchart is shown in Fig. 18, where the DP generates optimal input trajectory (working fluid pump speed) at given transient engine exhaust conditions. Then the exhaust conditions and optimal input trajectory are utilized to training the RF model. After the RF model is trained in Python, the model is saved and can be directly loaded and generates predictions at given inputs. The ORC-WHR plant model is built in Matlab. A GT-POWER-Python-Matlab co-simulation is conducted to validate the performance of the RF-RB model. The procedure is explained in Fig. 19. At each time step, engine speed and torque is interpolated from the transient profiles and serve as the inputs to the GT-POWER engine model, which predicts the exhaust gas mass flow rate, temperature and their respective time change rate. Python predicts working fluid pump speed given the four exhaust gas inputs and RB method corrects the pump speed. Finally, pump speed is executed in the ORC-WHR plant model. Then, time is increased by one time step and repeat the processes until the end of the engine operating profiles.

For comparison, RF model only and RF-RB model are compared with offline DP results. The validation compares working fluid vapor temperature, working fluid pump speed, net power production and accumulative energy among the RF model only, RF-RB model and DP.

Constant Speed Variable Load (CSVL) is considered as a representative driving cycle for long haul trucks in validation process. Engine model is simulated in GT-Power environment over CSVL driving cycle to obtain exhaust conditions for RF model validation. The corresponding exhaust gas conditions are shown in Fig. 20.

6.2. Validation results

The comparison between RF model only and offline DP are shown in Fig. 21. There is significant discrepancy in net power and accumulative energy. The RF model produces 33.5% less energy than the offline DP over the 1200 s driving cycle. This is due to multiple turbine shutdowns during the driving cycle when the vapor temperature touches saturation temperature. During these shutdowns, the turbine power is zero as noted in Fig. 21(c). In Fig. 21(b), the offline DP vapor temperature is only slightly above the saturation temperature with an average temperature difference of 6.1 °C only, which converts to 1.4% using offline DP as the reference and Kelvin as the temperature unit. When the RF model learns from the DP results, the RF model hardly keeps the error

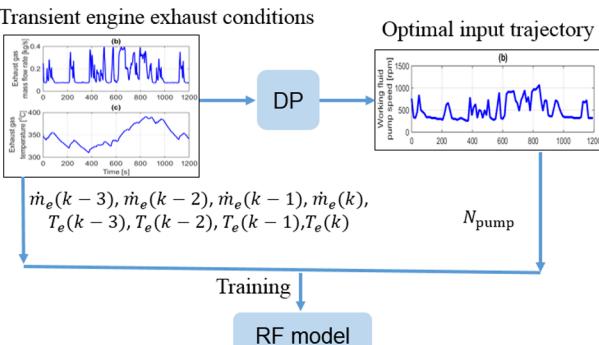


Fig. 18. RF model training process: input- exhaust mass flow conditions, output – ORC pump speed.

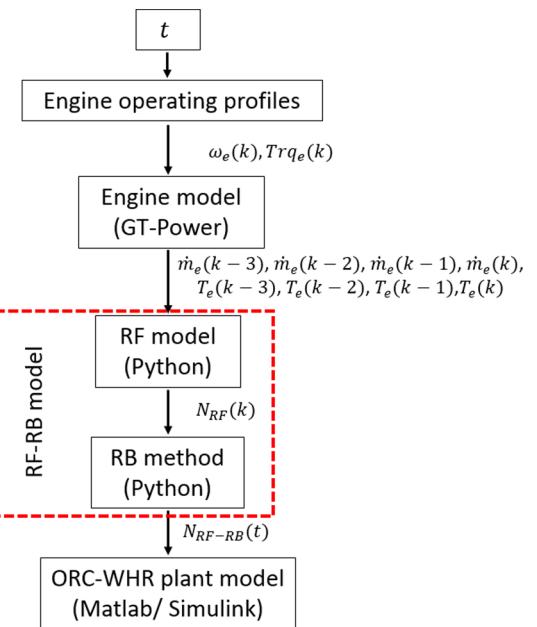


Fig. 19. RF-RB model validation process: engine operating condition – engine model – RF model – RB method – ORCWHR plant model.

below 1.4%. In Fig. 21(a), the working fluid pump speed error between the RF model and offline DP is 1.2%. When the RF model pump speed prediction has substantial error, the working fluid vapor temperature is then affected and shows error greater than 6.1 °C, which results in the saturation of the working fluid and turbine shutdowns. Thus, the shutdown is caused by the small variation of working fluid temperature,

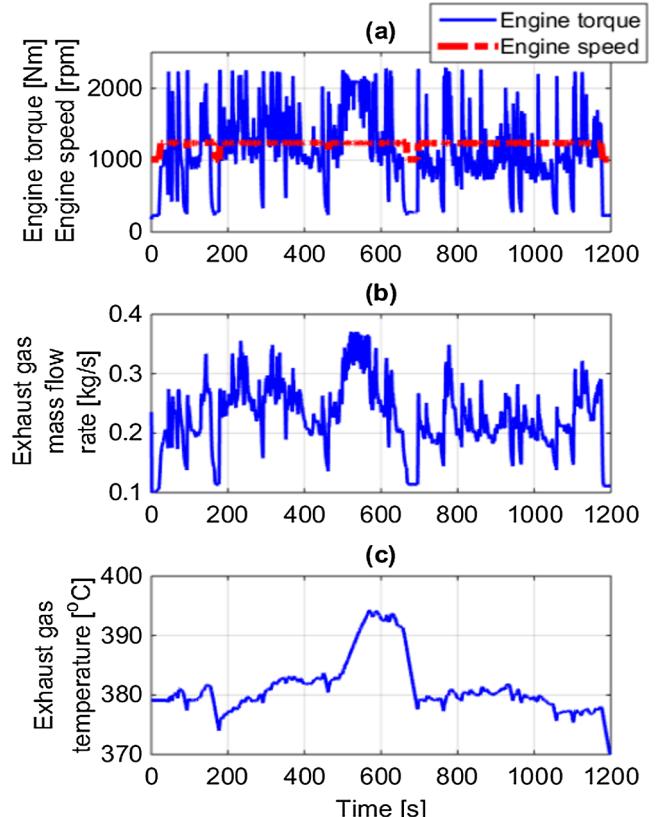


Fig. 20. Exhaust gas conditions from CSVL heavy-duty diesel engine driving cycle: (a) engine speed and torque, (b) exhaust gas mass flow rate, and (c) exhaust gas temperature.

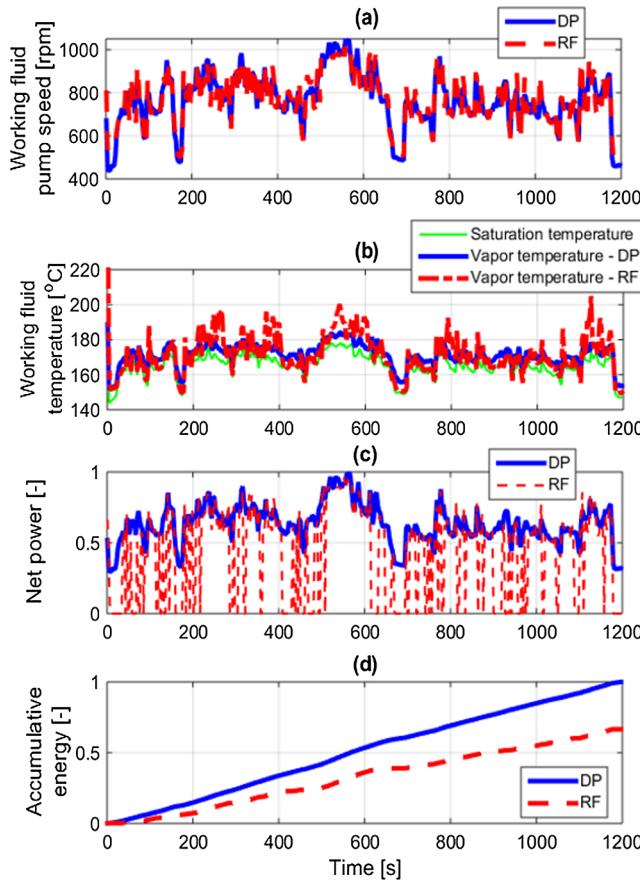


Fig. 21. Validation of RF model with eight inputs at CSVL driving cycle and DP results are the references: (a) working fluid pump speed comparison, (b) working fluid vapor temperature comparison between RF model prediction and DP results, (c) net power comparison, and (d) accumulative energy comparison.

which traces back to the RF model pump speed prediction error. However, the number of shutdowns can be reduced by adding RB method to the RF model.

The comparison between RF-RB method and offline DP are shown in Fig. 22. In Fig. 22(d), as the key performance parameter, the accumulative energy from RF model plus RB method is only 2.8% less than the offline DP optimization results. With the added RB method, the net power only cuts off once at the beginning of the driving cycle as shown in Fig. 22(c). For the working fluid pump speed prediction, due to the RB method reduces working fluid pump speed, the error increases to 1.7% compared with 1.2% in the RF model only simulation. This pump speed error increase is beneficial to the working fluid vapor duration in the operation. The pump speed difference between the RF only method and RF-RB method is shown in Fig. 23. Multiple pump speed reductions occur through the entire driving cycle. These speed reduction by RF-RB deviates the HPP speed from the prediction, thus slightly increases the speed error. In Fig. 22(b), with the addition of RB method, the vapor temperature stayed above the saturation temperature. This result reveals that the RF-RB model has satisfactory performance in real-time realizing the DP power optimization rules.

At each time step, RF-RB model execution time is 14.9 ms and working fluid pump speed update time is 200 ms. As long as the RF-RB model execution time is shorter than the working fluid pump speed update time, the real-time execution is feasible. During the RF model parameter optimization, three parameters are analyzed which include the number of estimators, the minimum samples split, and the minimum weight fraction leaf. In the parameter optimization, the model computation time is found to be only sensitive to the number of estimators. The computation time and percentage error versus the

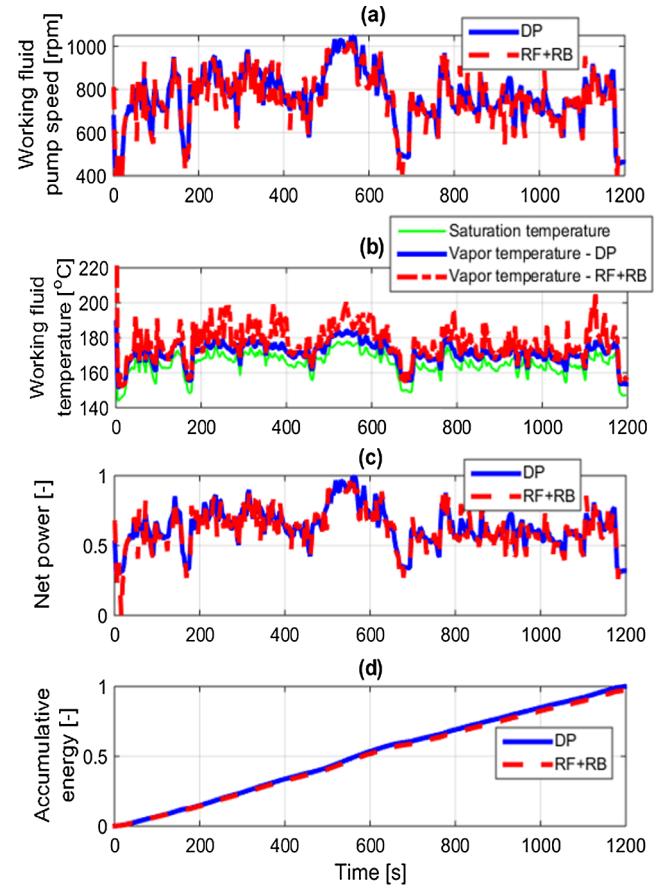


Fig. 22. Validation of RF model with eight inputs plus RB method at CSVL driving cycle and DP results are the references: (a) working fluid pump speed comparison, (b) working fluid vapor temperature comparison between RF model prediction and DP results, (c) net power comparison, and (d) accumulative energy comparison.

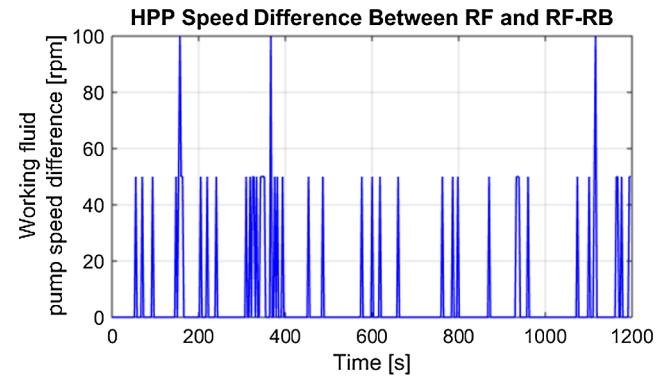


Fig. 23. HPP speed difference between RF simulation and RF + RB simulation.

number of estimators are shown in Fig. 24. The CPU computation time linearly increases with the number of estimators. The validation error decreases sharply when number of estimators increase from 5 to 30. Further increase of estimators does not have significant benefits to the validation error. In this study, 30 is selected as the final estimator number for its low error and low computation cost. Note that the simulation is implemented on a high-end Dell Workstation 3530 (Intel(R) Xeon CPU@2.80GHZ, 8 GB RAM, 1 TB Hard Disk). The RF algorithm hardware implementation was well researched in literature [36]. Field Programmable Gate Arrays (FPGA) is reported to be a very competitive hardware in the machine learning application. In [37], an FPGA only

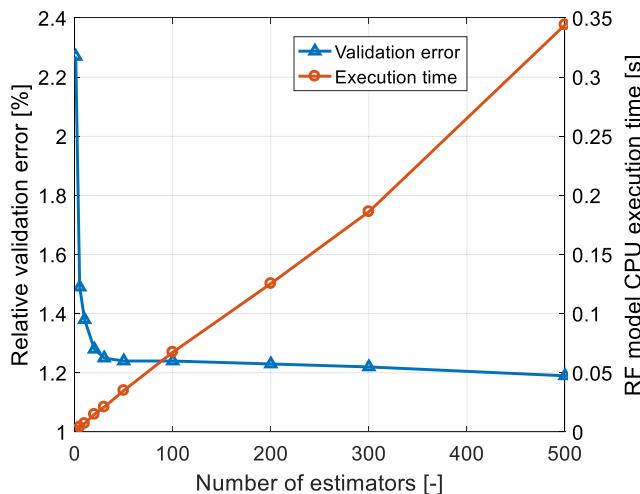


Fig. 24. Relative validation error of the optimal pump speed and CPU computation time at different number of estimators in RF model.

took 1.48×10^{-6} s to finish a single if-tree computation. In [38], the author achieved 0.03 s training time using a 1000-estimator RF model in FPGA hardware. In [36], the author reported the computation time by RF model in CPU, GPU and FPGA. The slowest hardware computing RF model is CPU and it can finish one prediction within 0.005 s at most in all seven different application datasets. GPU and FPGA take 0.003 s and 0.0003 s to finish one prediction, respectively. Besides the existing hardware, the hardware development in machine learning field has been soaring since early 2010 s. Deep Neural Networks training and prediction require much more computation than RF due to the high complexity of model architecture. Many Deep Neural Networks hardware implementation work have been published so far [39]. Therefore, the RF real-time hardware application is achievable.

6.3. Discussion

The machine learning algorithms not only extracts rules from DP results, but also rank the feature importance for all the input parameters. This feature importance ranking exhibits the relation strength between the inputs and the optimal pump speed. The exhaust gas mass flow rate shows the strongest relation with the pump speed. More importantly, the one time step delayed exhaust gas mass flow rate is more important than the exhaust gas temperature in the pump speed prediction, which has never been found in literature. This discovery proves the difference between the steady state optimization and transient optimization. For steady state optimization, the exhaust gas mass flow and temperature are kept constant, which means there is no time delayed parameter. However, the machine learning algorithms reveal that the time-delayed mass flow rate plays a more important role in optimal pump speed prediction than the temperature. In addition, the transient parameters like time-delayed exhaust gas mass flow rate and temperature can be considered in the rule extraction of the transient optimization. According to the results presented in Table 2, optimal pump speed error decreases from 1.05% to 0.85% by considering the time-delayed parameters. Therefore, in order to fully explore the power optimization, transient optimization should be considered.

The validation of pure RF model shows extremely large accumulative energy error, which reveals challenge in real-time implementation of ORC-WHR system optimization. The challenge is that the optimal vapor temperature is too close to the saturation temperature and the possibility of working fluid temperature saturation is high in this optimal temperature zone. The combination of RF model and RB method resolves this challenge by reducing the pump speed whenever the vapor

temperature is very close to the saturation temperature. Even though the addition of RB method increases the pump speed prediction error, it substantially reduces the accumulative energy error from 33.5% to 2.8%. In other words, the RF-RB method recovers 97.2% of energy recovered by offline DP method. Without the help of RB method, the RF model alone can only recover 66.5% of the energy by offline DP. This reveals that RF-RB method is suitable for real-time implementation.

In this simulation study, it is found that the RF model only requires 14.9 ms to make one working fluid pump speed prediction, while the working fluid pump speed updates every 200 ms. Therefore, the RF model can run in real-time without computation time issue. During the RF model parameter optimization, it is found that the computation time mostly depends on the number of estimators as shown in Fig. 24. The benefit of computation time saving is that the relative validation error of the optimal pump speed drops sharply when number of estimators is less than 30. As the number of estimators keeps increasing above 30, the validation error decreases slowly. Considering the linear increase of computation time as the number of estimators increases, a small estimator number can be selected for its low computation cost and error.

7. Conclusion

This study utilizes machine learning algorithms to extract rules from offline Dynamic Programming results for possible real-time realization of the Dynamic Programming algorithm. This study thus bridges the gap between offline Dynamic Programming analysis and its real-time implementation. After extracting rules from offline Dynamic Programming results on a Federal Test Procedure driving cycle, the Random Forest model is then evaluated on a Constant Speed Variable Load driving cycle. The Random Forest model recovers 66.5% of energy recovered by the offline Dynamic Programming. The reason of the large error is that the optimal vapor temperature is too close to saturation temperature. The turbine shuts down when working fluid temperature saturates.

One of the key discovery of this study is that during the transient engine conditions, a one time-step delayed exhaust gas mass flow rate is more important than the exhaust gas temperature for power optimization in an organic Rankine Cycle waste heat recovery system. Further, the addition of time delayed parameters reduces optimal pump speed prediction error from 1.05% to 0.85%. This discovery reveals that steady state optimization results are not necessarily suitable for transient cases in organic Rankine Cycle waste heat recovery system. It is also shown that to improve power production in transient driving conditions, transient parameters like time delayed exhaust gas mass flow rate and temperature should be incorporated. The other important observation is that the energy recovery of Random Forest model significantly increases to 97.2% by adding rule-based method to the Random Forest model, which reduces the working fluid pump speed and increase the vapor temperature. Thus, a combination of Random Forest model and rule-based method could be a viable surrogate for implementing Dynamic Programming algorithm like performance in real-time.

To further improve the proposed methodology, a wider range of engine operating conditions and experimental exploration can be considered to exploit the Random Forest algorithm performance.

CRediT authorship contribution statement

Bin Xu: Conceptualization, Methodology, Software, Data curation, Investigation, Writing - original draft. **Dhruvang Rathod:** Writing - review & editing. **Adamu Yebi:** Writing - review & editing. **Zoran Filipi:** Supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial

interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A

Linear Regression: It is a linear approach to model the relationship between one dependent variable and one or more independent variables. The relationships are often fitted using the least squares approach.

Ridge Regression: It is a linear approach similar to linear regression, except the way of training the model. In general, the linear regression does not penalize the coefficients of the inputs. However, ridge regression penalizes the coefficients to reduce the flexibility of coefficients. The reduced flexibility means better generalization and less overfitting. Ridge regression uses L2 norm in the coefficients penalty.

Kernel Ridge Regression: Similar to ridge regression except the training methods. Kernel ridge regression utilizes Lagrange multiplier in the training loss function. In addition, the coefficients of the linear terms are not explicitly calculated and are substituted by Lagrange multiplier and training data.

K-Nearest Neighbor Regression: The prediction is the weighted average of the k nearest neighbor points from the training dataset.

Gaussian Process Regression: It provides a method for modeling probability distributions over functions.

Decision Tree Regression: It is a decision support tool that uses a tree-like model of decisions and their possible consequences. It is one way to display an algorithm that only contains conditional control statements.

Random Forest Regression: It is an ensemble learning method for regression that operates by constructing a multitude of decision trees at training time and outputting the mean prediction of the individual trees. By outputting the mean prediction, it reduces the overfitting probability of the decision trees.

Bagged Trees Regression: It is a machine learning ensemble meta-algorithm designed to improve the stability and accuracy of machine learning algorithms used in regression. It also reduces variance and overfitting.

Adaboost Regression: It combines multiple weak learners to improve the performance. It gradually adds weak learning algorithms by minimizing the large errors from the existing learners. The individual learners can be weak, but as long as the performance of each one is slightly better than random guessing, the final model can be proven to converge to a strong learner.

Xgboost Regression: Similar to Adaboost regression, Xgboost considers multiple weak learners. Beyond that, regularization is also included to reduce the overfitting. In addition, it includes a sparsity-aware split finding algorithm to handle different types of sparsity patterns in the dataset. Its computation cost is minimized by parallel learning, cache awareness and out-of-core computing features.

Artificial Neural Network Regression: It is based on a collection of connected units (artificial neurons), which is inspired by the brain biological structure. It has different layers of neurons and neurons are inter-connected layer by layer. The information is transported through the neurons by different weights, bias and activation functions.

During the model selection, the space of parameters and common variations for each learning algorithm are explored as thoroughly as possible. The parameters utilized for each model are summarized as follows:

LR: Fit intercept varies ('True', 'False') and normalization varies ('True', 'False').

RR: Regularization strength varies (10^{-4} , 10^{-3} , 10^{-2} , 10^{-1} , 0). Fit intercept varies ('True', 'False') and normalization varies ('True', 'False').

KRR: variance reduction alpha varies (10^{-4} , 10^{-3} , 10^{-2} , 10^{-1}). Internal kernel mapping varies ('linear', 'polynomial') and degree of

polynomial kernel varies (1, 2, 3, 4, 5). Zero coefficient for polynomial kernel varies (10^{-4} , 10^{-3} , 10^{-2} , 10^{-1} , 0, 1).

KNN: The K value varies (1, 2, 4, 8, 12, 16, 20) and leaf size varies (5, 10, 20, 30, 50, 100). Weight varies ('uniform', 'distance').

GPR: Diagonal value of the kernel matrix during fitting varies (10^{-10} , 10^{-8} , 10^{-4} , 10^{-2} , 10^0).

DT: The maximum depth of the tree varies (1, 5, 10, 20, 30, 40, 50, 100, 200, 500). Minimum samples split varies (1, 5, 10, 20, 30, 40, 50, 100)

RF: The number of estimators varies (2, 5, 10, 20, 30, 50). Minimum samples split varies (1, 2, 4, 6, 8, 10, 15, 20). Minimum weight fraction leaf varies (10^{-4} , 10^{-3} , 10^{-2} , 10^{-1} , 0).

BT: Number of estimators varies (2, 5, 10, 20, 30, 50) and maximum sample ratios varies (0.1, 0.3, 0.5, 0.7, 0.9, 1.0).

ADB: Number of estimators varies (2, 5, 10, 20, 30, 50) and learning rate varies (0.1, 0.3, 0.5, 0.7, 0.9, 1.0). Loss function for weights update varies ('linear', 'square', 'exponential').

XGB: Number of estimators varies (10, 50, 100, 200, 500, 1000, 2000, 5000, 10000). Minimum child weight varies (1, 3, 5, 10, 20, 30, 50, 100). Subsample ratio of columns in constructing each tree varies (0.1, 0.3, 0.5, 0.7, 0.9, 1.0). Subsample ratio of the training instances varies (0.1, 0.3, 0.5, 0.7, 0.9, 1.0).

ANN: The ANN is trained with gradient descent backpropagation. The number of neurons in each layer varies (10, 50, 100, 200, 500) and the number of hidden layer varies (1, 2, 4, 8, 16).

References

- [1] Xu B, Rathod D, Yebi A, Filipi Z, Onori S, Hoffman M. A comprehensive review of organic rankine cycle waste heat recovery systems in heavy-duty diesel engine applications. *Renew Sustain Energy Rev* 2019;107:145–70.
- [2] Wang XZ, Li B, Yan YY, Liu S, Li J. A study on heat transfer enhancement in the radial direction of gas flow for thermoelectric power generation. *Appl Therm Eng* 2016;102:176–83.
- [3] Kant Manuel, Romagnoli Alessandro, Mamat Aman MI, Martinez-Botas Ricardo F. Heavy-duty engine electric turbocompounding. *Proc Inst Mech Eng, Part D: J Automobile Eng* 2015;229(4):457–72. <https://doi.org/10.1177/0954407014547237>.
- [4] Xu B, Yebi A, Onori S, Filipi Z, Liu X, Shatty J, et al. Power maximization of a heavy duty diesel organic Rankine cycle waste heat recovery system utilizing mechanically coupled and fully electrified turbine expanders. *ASME 2016 internal combustion engine division fall technical conference*. 2016. V001T05A005-V001T05A005.
- [5] Xu B, Rathod D, Yebi A, Filipi Z. A comparative analysis of real-time power optimization for organic Rankine cycle waste heat recovery systems. *Appl Therm Eng* 2020;164: 114442.
- [6] Quoilin S, Aumann R, Grill A, Schuster A, Lemort V, Spleithoff H. Dynamic modeling and optimal control strategy of waste heat recovery Organic Rankine Cycles. *Appl Energy* 2011;88:2183–90.
- [7] Nazari N, Heidarnajad P, Porkhial S. Multi-objective optimization of a combined steam-organic Rankine cycle based on exergy and exergo-economic analysis for waste heat recovery application. *Energy Convers Manage* 2016;127:366–79.
- [8] Dai YP, Wang JF, Gao L. Parametric optimization and comparative study of organic Rankine cycle (ORC) for low grade waste heat recovery. *Energy Convers Manage* 2009;50:576–82.
- [9] Sadeghi M, Nemati A, Ghavimi A, Yari M. Thermodynamic analysis and multi-objective optimization of various ORC (organic Rankine cycle) configurations using zeotropic mixtures. *Energy* 2016;109:791–802.
- [10] Yang F, Zhang H, Bei C, Song S, Wang E. Parametric optimization and performance analysis of ORC (organic Rankine cycle) for diesel engine waste heat recovery with a fin-and-tube evaporator. *Energy* 2015;91:128–41.
- [11] Xu B, Yebi A, Onori S, Filipi Z, Liu XB, Shatty J, et al. Transient power optimization of an organic Rankine cycle waste heat recovery system for heavy-duty diesel engine applications. *Sae Int J Altern Powertrains* 2017;6:25–33.
- [12] Yebi A, Xu B, Liu X, Shatty J, Anschel P, Filipi Z, et al. Estimation and predictive control of a parallel evaporator diesel engine waste heat recovery system. *IEEE Trans Control Syst Technol* 2017;1–14.
- [13] Hernandez A, Desideri A, Ionescu C, Quoilin S, Lemort V, De Keyser R. Experimental study of predictive control strategies for optimal operation of organic rankine cycle systems. *2015 European control conference (ECC)*. 2015.
- [14] Feru E, Willems F, de Jager B, Steinbuch M. Model predictive control of a waste heat recovery system for automotive diesel engines. *Proceedings of the 18th international conference on system theory*. 2014. p. 658–63.
- [15] Esposito MC, Pompini N, Gambarotta A, Chandrasekaran V, Zhou J, Canova M. Nonlinear model predictive control of an organic rankine cycle for exhaust waste heat recovery in automotive engines. *IFAC-PapersOnLine* 2015;48:411–8.
- [16] Merino EG, Schloder JP, Kirches C. A nonlinear model-predictive control scheme for

- a heavy duty truck's waste heat recovery system featuring moving horizon estimation. 2018 annual American control conference (ACC). 2018. p. 6329–34.
- [17] Peralez J, Tona P, Sciarretta A, Dufour P, Nadri M. Optimal control of a vehicular organic rankine cycle via dynamic programming with adaptive discretization grid. IFAC Proc Volumes 2014;47:5671–8.
- [18] Bellman R. Dynamic programming. Science 1966;153:34–7.
- [19] Scornet E, Biau G, Vert J-P. Consistency of random forests. Ann Statistics 2015;43:1716–41.
- [20] Xu B, Rathod D, Kulkarni S, Yebi A, Filipi Z, Onori S, et al. Transient dynamic modeling and validation of an organic Rankine cycle waste heat recovery system for heavy duty diesel engine applications. Appl Energy 2017;205:260–79.
- [21] Rathod D, Belwariar U, Xu B, Hoffman M. An enhanced evaporator model for working fluid phase length prediction, validated with experimental thermal imaging data. Int J Heat Mass Transf 2019;132:194–208.
- [22] Desideri A, Dechesne B, Wronski J, van den Broek M, Gusev S, Lemort V, et al. Comparison of moving boundary and finite-volume heat exchanger models in the modelica language. Energies 2016;9.
- [23] Rathod D, Xu B, Yebi A, Vahidi A, Filipi Z, Hoffman M. A look-ahead model predictive control strategy for an organic rankine cycle-waste heat recovery system in a heavy duty diesel engine application. SAE Technical Paper 0148-7191; 2019.
- [24] Horst TA, Tegethoff W, Elts P, Koehler J. Prediction of dynamic Rankine Cycle waste heat recovery performance and fuel saving potential in passenger car applications considering interactions with vehicles' energy management. Energy Convers Manage 2014;78:438–51.
- [25] Xu B, Yebi A, Hoffman M, Onori S. A rigorous model order reduction framework for waste heat recovery systems based on proper orthogonal decomposition and Galerkin projection. IEEE Trans Control Syst Technol 2018.
- [26] Li Y, Zou C, Berencabar M, Nanini-Maury E, Chan JC-W, van den Bossche P, et al. Random forest regression for online capacity estimation of lithium-ion batteries. Appl Energy 2018;232:197–210.
- [27] Ahmad MW, Reynolds J, Rezgui Y. Predictive modelling for solar thermal energy systems: A comparison of support vector regression, random forest, extra trees and regression trees. J Cleaner Prod 2018;203:810–21.
- [28] Wang La, Zhou X, Zhu X, Dong Z, Guo W. Estimation of biomass in wheat using random forest regression algorithm and remote sensing data. The Crop J 2016;4:212–9.
- [29] Weisberg S. Applied linear regression 528. John Wiley & Sons; 2005.
- [30] Kohavi R. A study of cross-validation and bootstrap for accuracy estimation and model selection. In: Ijcai, 1995. p. 1137–45.
- [31] Deng J, Dong W, Socher R, Li L-J, Li K, Fei-Fei L. Imagenet: A large-scale hierarchical image database. 2009 IEEE conference on computer vision and pattern recognition. 2009. p. 248–55.
- [32] Breiman L. Random forests. Mach Learning 2001;45:5–32.
- [33] Buitinck L, Louppe G, Blondel M, Pedregosa F, Mueller A, Grisel O, et al., API design for machine learning software: experiences from the scikit-learn project, arXiv preprint arXiv:1309.0238; 2013.
- [34] Bergstra J, Bengio Y. Random search for hyper-parameter optimization. J Mach Learning Res 2012;13:281–305.
- [35] Kennedy J. Particle swarm optimization. Encyclopedia of machine learning. Springer; 2011. p. 760–6.
- [36] Nakahara H, Jinguiji A, Sato S, Sasao T. A random forest using a multi-valued decision diagram on an FPGA. 2017 IEEE 47th international symposium on multiple-valued logic (ISMVL). 2017. p. 266–71.
- [37] Buschjäger S, Morik K. Decision tree and random forest implementations for fast filtering of sensor data. IEEE Trans Circuits Syst I Regul Pap 2017;65:209–22.
- [38] Cheng C. Random forest training on reconfigurable hardware; 2015.
- [39] Gysel P. Ristretto: Hardware-oriented approximation of convolutional neural networks. arXiv preprint arXiv:1605.06402; 2016.