A comprehensive review of organic rankine cycle waste heat recovery systems in heavy-duty diesel engine applications

Bin Xu\textsuperscript{a}, Dhruvang Rathod\textsuperscript{a,b,⁎}, Adamu Yebi\textsuperscript{c}, Zoran Filipi\textsuperscript{a}, Simona Onorib, Mark Hoffman\textsuperscript{a}

\textsuperscript{a} Clemson University, Department of Automotive Engineering, 4 Research Dr., Greenville, SC 29607, USA
\textsuperscript{b} Stanford University, Energy Resources Engineering Department, Stanford University, CA 94305, USA
\textsuperscript{c} Auburn University, Department of Mechanical Engineering, 1448 Wiggins Hall, 354 War Eagle Way, Auburn, AL 36849, USA

A R T I C L E   I N F O

Keywords:
- Waste heat recovery
- Organic Rankine cycle
- Heavy-duty diesel engine
- Review
- Simulation and experiments

A B S T R A C T

Effective recovery of heavy-duty vehicle waste heat is a key solution toward meeting the increasingly stringent fuel economy and CO\textsubscript{2} emission standards. Different from previous publications, this paper presents a preliminary introduction of organic Rankine cycle waste heat recovery (ORC-WHR) in heavy-duty diesel (HDD) vehicle applications in the past decade. It presents a wide range of topics in the HDD vehicle ORC-WHR system development, including system architecture evaluation, heat exchanger selection, expander selection, working fluid selection, power optimization, control strategy evaluation, simulation and experimental work overview, and limiting factors. In the system architecture selection, the tradeoff between fuel savings and system complexity dominates. In the heat exchanger design, besides the heat exchanger efficiency, transient evaporator response is critical factor for the system control and performance. The expander type and configuration is closely coupled to the expander power output type (i.e., electricity or mechanical power). WHR power production is most sensitive to working fluid mass flow rate, with less sensitivities to expander speed and condenser coolant mass flow rate. The integration of ORC-WHR control with engine control shows potential to improve the waste heat recovery system performance. The simulation studies predict higher power recovery levels than that in experimental work (0–60 kW vs. 0–14 kW), which could result from the large number of heat resources, optimistic expander and pump efficiencies and neglected heat losses in the simulations.

1. Introduction

Waste heat recovery (WHR) is a heavily researched topic in the Heavy-Duty Diesel (HDD) engine community. Driven by increasingly stringent fuel economy and emissions regulations, several works have been published in the field over the past twenty years [1–5]. In 2016, United States government agencies adopted regulations on HDD for 2018–2027, requiring a 25% decrease in fuel consumption and CO\textsubscript{2} emissions by model year (MY) 2027 compared with the MY 2018 baseline [6]. During normal HDD operation, more than 40% of fuel energy is wasted as heat [1,7], and hence WHR becomes attractive technology to save fuel and reduce CO\textsubscript{2} emissions. Among all WHR technologies, thermoelectric generators (TEG) [8–12], turbo-compounding [3,19–16] and organic Rankine cycles (ORC) [17–21] remain the most popular.

TEG utilize the temperature difference between the exhaust gas and engine coolant to generate electricity. TEG are compact and have a simple structure. However, current TEG materials predominantly possess low figure of merit values, limiting the efficiency of TEG [22]. A 1999 GMC Sierra pick-up truck was utilized to test TEG for WHR [10]. However, the thermal efficiency of the TEG measured during the test was less than 2% and the generated power was less than 0.3 kW. Similarly, Mack Trucks and Kenworth Trucks [23] tested a TEG device and the generated power was below 1 kW.

Turbo-compounding recovers part of the tailpipe (TP) exhaust gas waste heat energy downstream of the turbocharger. The devices can be either electrical or mechanical [16]. However, turbo-compounding only recovers waste heat from the TP exhaust gas [3,14], excluding waste

Abbreviation: AMOX, Ammonia Oxidation Catalyst; CP, Critical Point; DOC, Diesel Oxidation Catalyst; DPF, Diesel Particle Filter; EGR, Exhaust Gas Recirculation; FVM, Finite Volume Model; HD, Heavy-Duty; HDD, Heavy-Duty Diesel; HT, High Temperature; IPC, Integrated Powertrain Control; LT, Low Temperature; MBM, Moving Boundary Model; MY, Model Year; NMPC, Nonlinear Model Predictive Control; N\_HPP, High pressure pump speed, rpm; N\_Turb, Turbine expander speed, rpm; N\_CWP, Cooling water pump speed, rpm; ORC, Organic Rankine Cycle; SCR, Selective Catalytic Reduction; TEG, Thermoelectric Generator; TP, Tail Pipe; WF, Working Fluid; WHR, Waste Heat Recovery

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

https://doi.org/10.1016/j.rser.2019.03.012
Received 6 September 2018; Received in revised form 4 March 2019; Accepted 5 March 2019
Available online 08 March 2019
1364-0321/ © 2019 Elsevier Ltd. All rights reserved.
heat in the exhaust gas recirculation (EGR), charge air (CA), and engine coolant streams are not recovered. In addition, after the turbocharger expansion, the waste heat available thermal power is low in regards to further expansion.

ORC systems are similar to the Rankine cycle widely utilized in stationary power plants. ORC simply uses organic working fluid whereas the Rankine cycle utilizes water. Multiple heat exchangers can be utilized to recover heat from various HDD sources [24–27], maximizing the total waste heat utilization. While the power industry has a long history with the Rankine cycle, the transient nature of HDD heat sources, tight packaging constraints, system weight considerations, utilization of an organic working fluid, and the cost-conscious nature of the HDD industry create challenges for ORC-WHR mobile implementation.

In the literatures related to the reviews of ORC-WHR in mobile application, Chintala et al. overviewed the working fluid and heat exchanger selection, and back pressure limiting factor in compression ignition engine applications [28]. However, the factors considered in the paper were limited to working fluid and heat exchanger selection and not covered the architectures, power optimization, control, and limiting factors in the system development. Lion et al. analyzed the off-highway HDD vehicle ORC-WHR systems with the focus on variation of vehicle types and driving cycle profiles [29]. This paper presented excellent effort in different types of HDD vehicle and driving cycle profiles, but the challenges and possible solutions in the component level development were not described. Zhou et al. reviewed the ORC-WHR systems applied in on-road passenger cars with the focus on architectures, expander selection, working fluid selection and system integration challenges [30]. The passenger cars generally use spark ignition engine rather than compression ignition engine. The displacement of the passenger car engine is mostly in the range of 1–3 L, whereas the displacement of HDD engine is generally greater than 6 L. Besides the displacement difference, the weight and packaging requirement of ORC-WHR system in passenger cars have higher priority than that in HDD vehicle. Thus, the passenger car ORC-WHR system development is different than that in the HDD vehicles. Sprouse III et al. reviewed the ORC-WHR system in internal combustion engines based on the year of publication from 1970s to present [31]. This paper presented a good work in the history of ORC-WHR techniques in the internal combustion engine applications. However, the references were not gathered based on different components and topics. Therefore, it is hardly used as the preliminary guidance of in the component development, system optimization and control.

This work presents a preliminary introduction for the HDD vehicle ORC-WHR system development, including architectures, heat exchanger design, expander design, working fluid selection, power optimization, control strategies and limiting factors in the system development. In addition, simulation work and experimental work are listed in table and compared. This paper aims to cover a wide range of issues and important factors in the development of HDD ORC-WHR system, which can be used as the preliminary guidance for the people who are not familiar with the HDD ORC-WHR system and for the people who start developing the HDD ORC-WHR system.

This paper is organized as follows: Section 2 describes the ORC system architectures based on their respective evaporator, expander, condenser and working fluid pump configurations. Selection options for the heat exchanger, expander, and working fluid are presented in Sections 3–5, respectively. Section 6 overviews the challenges and possible solutions for HDD ORC-WHR system power optimization. In Section 7, ORC-WHR control strategies are discussed. Simulation and experimental works are reviewed in Section 8 and Section 9. The fuel economy and CO2 emissions are discussed and summarized in Section 10. The limiting factors in the system development are described in Section 11. Finally, the paper ends with summary and conclusions in Section 12 and Section 13, respectively.

2. ORC system architecture

System architecture is important for several reasons: complexity, efficiency, control, cost, and durability [32–35]. The architecture determines the layout of the entire HDD ORC-WHR system. The available room in the vehicle needs to be analyzed before the architecture design so that the architecture can fit into the HDD vehicle without interfering original components in the vehicle. Cummins integrated the ORC-WHR system under the hood [36] and AVL attached the system to the frame of the truck [37]. After deciding which heat sources to tap for WHR, the system configuration is better discussed on a component-by-component basis. This section is broken down into the following five subsections: (i) heat source available thermal power and quality analysis, (ii) evaporator configuration, (iii) expander configuration, (iv) condenser configuration, and (v) working fluid pump configuration.
2.1. Heat source available thermal power and quality analysis

Heat source available thermal power indicates the theoretical available energy using the ambient environment as the reference state. The available thermal power of a heat source can be calculated using Eq. (1) where \( m_{hs} \) and \( h_{hs} \) represent heat source mass flow rate and enthalpy, respectively. Subscript \( hs \) represents heat source and subscript 0 represents the reference condition. From Eq. (1), heat source available thermal power is proportional to heat source mass flow rate and enthalpy. At a given pressure, enthalpy increases with temperature. However, in the ORC-WHR system, more available energy does not mean more power output from the system.

Ideally, energy sources with higher available thermal power do translate to high power output from the ORC-WHR system. However, limitations to heat exchanger size can prevent this trend from coming to fruition. The heat transfer area between the heat source and working fluid is generally on the order of 1–10 m² in HDD ORC-WHR systems [38]. With this constraint, the heat source enthalpy at the heat exchanger outlet, \( h_{out} \), is often much larger than the reference enthalpy \( h_0 \) in Eq. (1). Thus, only a portion of the available thermal power is ultimately transferred to the heat exchanger wall, which is expressed in Eq. (2).

\[
E_{\text{avail}} = \dot{m}_{hs} (h_{hs} - h_0) \tag{1}
\]

Heat transfer area \( A \) depends on heat exchanger design and is not related to the heat source conditions. Heat transfer coefficient \( U \) depends critically on heat exchanger design, heat exchanger material, heat source fluid properties, and heat source flow conditions. The temperature difference, \( \Delta T \), depends on the heat source temperature and heat exchanger wall temperature. To compare the quality of different heat source streams, first assume \( A \) and \( U \) are constant. Then the heat transfer \( E_{\text{avail}} \) is proportional to \( \Delta T \). Thus, heat sources with high temperature (high quality) are highest priority for WHR.

Table 1 summarizes the energy analysis based on experiments conducted on a 2006MY 10.8 L Cummins ISX HDD [7] and a 2004MY 10.8 L Mack MP7-355E HDD [39]. Integrating the data from two HDDs, the portions of fuel energy wasted in the TP exhaust gas, engine coolant, EGR exhaust gas, and the CA are 19–41%, 10–20%, 6–15%, and 2–10%, respectively. In terms of energy quantity, the TP exhaust gas and engine coolant represent the most desirable streams for WHR. However, when available thermal power is considered, the top targets for WHR become the EGR exhaust gas and the TP exhaust gas due to their relatively high temperatures. For example, the average available thermal power of the various waste heat streams of a 2004MY 10.8 L Mack MP7-355E HDD operating over the ESC-13 mode as follows: TP exhaust gas (45.7 kW), EGR exhaust gas (20.7 kW), engine coolant (5.2 kW), and charge air (3.5 kW) [39].

To solidify this concept, consider the mass flow and temperature of the various heat sources from experiments with a 8.4L HDD engine [25]. The mass flow rate and temperature range are summarized in Table 2. Among the four heat sources, the TP exhaust gas and EGR exhaust gas have much higher temperatures than the CA and engine coolant. Therefore, the exhaust gas heat sources have good quality and will require the heat exchanger area per unit of energy recovered if all other factors are equal (i.e. equivalent heat transfer coefficients within the heat exchangers). Engine coolant has a relatively low temperature and therefore low quality, but contains a large quantity of waste energy and possesses a relatively high source fluid mass flow rate.

Table 1

<table>
<thead>
<tr>
<th>Energy flow path</th>
<th>% of total fuel energy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mechanical (crankshaft work)</td>
<td>25–42</td>
</tr>
<tr>
<td>TP exhaust gas</td>
<td>19–41</td>
</tr>
<tr>
<td>Engine coolant</td>
<td>10–20</td>
</tr>
<tr>
<td>EGR exhaust gas</td>
<td>6–15</td>
</tr>
<tr>
<td>Charge Air</td>
<td>2–10</td>
</tr>
</tbody>
</table>
Table 2
Heat source mass flow rate and temperature range for a HDD engine [25].

<table>
<thead>
<tr>
<th>Heat sources</th>
<th>Mass flow rate (kg/s)</th>
<th>Temperature (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TP exhaust gas</td>
<td>0.099–0.380</td>
<td>206–471</td>
</tr>
<tr>
<td>EGR exhaust gas</td>
<td>0.000–0.060</td>
<td>282–594</td>
</tr>
<tr>
<td>CA</td>
<td>0.140–0.390</td>
<td>57–153</td>
</tr>
<tr>
<td>Engine coolant</td>
<td>2.400–3.800</td>
<td>74–84</td>
</tr>
</tbody>
</table>

2.2. Evaporator configuration

The heat exchanger configuration is related to the waste heat sources utilized, their accessibility, energy quality, and packaging constraints. In HDD, potential heat sources include TP exhaust gas, EGR exhaust gas, charge air, and engine coolant [39,45,47]. The number of heat sources selected and the medium supplying the waste heat affect the evaporator configuration. As more heat sources are considered, more room is needed for the ORC-WHR system installation in the vehicle. In addition, the order of the heat sources that working fluid flows through requires analysis, which should consider multiple factors of the heat sources such as temperature, mass flow, location, temperature restriction at the heat exchanger exit. In [48], a very complex evaporator configuration is utilized, where recuperator and charge air cooler are connected in parallel as preheater. Exhaust gas located downstream of recuperator as boiler. The working fluid coming out of charger air cooler and exhaust boiler mixed up, entered the EGR boiler and then EGR superheater. The EGR heat exchanger is composed of boiler and superheater. Similar two-stage heat exchanger can be found in [7]. If more than one heat source is utilized, parallel or series heat exchanger layout also needs analysis.

2.2.1. TP and EGR as heat sources

As the highest quality heat sources, many systems attempt to recover waste heat from both the TP and EGR gases. These two heat sources can be connected in series or in parallel. The advantages of a series configuration include the ability to utilize a single working fluid mass flow control, and a better guarantee of superheated working fluid prior to the expansion device. Unfortunately, series configurations also induce larger pressure drops in the working fluid side, and produce slightly lower ORC efficiencies according to [35]. The working fluid side pressure drop increases proportionally with mass flow rate [49]. For the series configuration, all the working fluid exiting the high pressure pump passes through two evaporators. For the parallel configuration, the working fluid is split into the two evaporators, thus the working fluid mass flow rate for each of single evaporator is less than that of series configuration. For two identical evaporators and working fluid mass flow rate, the pressure drop of parallel configuration is ¼ of that in series configuration. For the exhaust gas side, instead of linear relation, the pressure drop showed second order relation with the exhaust gas mass flow rate as measured in [50,51]. Teng et al. developed and tested a HDD ORC-WHR system with TP and EGR evaporators configured in series [7]. The experimental results are shown in Table 3. At different engine speed and torque values, the indicated WHR turbine power is up to 9.1 kW. Cummins, in 2013, presented a HDD ORC-WHR system with series connected TP and EGR evaporators [52]. The series configuration increased absolute brake thermal efficiency by 2.8% (or 6% relative improvement considering 47% as the baseline brake thermal efficiency).

In contrast, a parallel evaporator configuration added complexity to the working fluid mass flow control, but provided benefits in ORC efficiency [35]. Grelet et al. presented a simulation comparison between series and parallel configured TP and EGR evaporators [53]. The results showed that parallel configuration outperforms the series configuration for the investigated working fluids, with the parallel configuration achieving a 3–5% fuel economy improvement and the series configuration achieving a 2–4% fuel economy improvement. Bosch also reported HDD ORC-WHR system experimental results with TP and EGR evaporators configured in parallel [54]. The WHR turbine power values are in the range of 2.1 kW, 5.3 kW and 9 kW at 25%, 50% and 75% load with B engine speed points.

2.2.2. Multi-loop ORC systems

Even though TP and EGR exhaust gas are the two most popular heat sources for WHR, utilization of three or more heat sources can provide incremental benefits in HDD ORC-WHR systems. Multi-loop ORC systems attempt to maximize the cycle efficiency by tailoring each system loop to the respective heat source. Some works have utilized a different working fluid in each loop, with each fluid selected based on the temperature level of the heat source utilized. While multi-loop ORC systems add to system cost and complexity, they can utilize a higher fraction of the waste heat for power generation.

Shu et al. considered TP exhaust gas, EGR exhaust gas, CA and cooling jacket water as heat sources [25]. The authors claimed that the diversity and multi-grade characteristics of the multiple heat sources make it difficult to recover the waste heat efficiently. Thus, a dual-loop ORC system was chosen. First, the high temperature (HT) loop recovered the high quality heat from TP and EGR exhaust gas. Afterwards, the TP and EGR exhaust gas flowed through the low temperature (LT) loop to evaporate the LT loop working fluid. Meanwhile, the CA and jacket water preheated the LT loop working fluid. A recuperator was utilized in the HT loop. The engine in the study was 8.4 L HDD with rated power of 243 kW. The heat source mass flows and temperatures were experimentally collected and utilized as inputs to a dual-loop ORC simulation. The TP evaporator was located at the turbocharger exit rather than downstream of the aftertreatment system, which is not common because the aftertreatment system generally operates at high TP exhaust gas temperatures [55,56]. The working fluid in each loop was selected separately and the net power maps for all the working fluid candidates were given. The results showed that the maximum net power output from the dual loop ORC was in the range of 27.9–33.9 kW depending on the working fluid selection and evaporation pressures of the respective loops. In addition, the HT loop generated less power than the LT loop.

In contrast with Shu’s study, Yang et al. presented a two-loop (HT and LT) ORC-WHR system with a single working fluid – R245fa [24]. R245fa was chosen for its merits regarding environmental impact, safety, system efficiency, and stability. The 11.6 L HDD engine under investigation was rated at 247 kW. The HT ORC loop utilized the TP exhaust gas as a heat source and a recuperator. The LT ORC loop was utilized as the condenser fluid for the HT loop to preheat the LT working fluid. Additional preheat for the LT loop was provided by the CA while engine coolant was utilized to evaporate the working fluid. In addition, the HDD aftertreatment system was not considered in the architecture. Mass flow and temperature of the respective heat sources were collected, and an energy balance analysis was conducted (detailed models were not provided). Available thermal power destruction rate map of each component was given as function of engine speed and torque for both HT and LT loops. The results were similar to Shu’s study, namely, the HT loop produced less power than LT loop. This was a result of utilizing multiple LT heat sources, which contained a high

Table 3
Experiment summary for the TP and EGR exhaust gas evaporators ORC-WHR system [7].

<table>
<thead>
<tr>
<th>Engine speed (rpm)</th>
<th>Engine torque (Nm)</th>
<th>Indicated turbine power (kW)</th>
<th>Fuel consumption benefits (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1615</td>
<td>925</td>
<td>3.1</td>
<td>2.1</td>
</tr>
<tr>
<td>1615</td>
<td>1385</td>
<td>3.7</td>
<td>1.6</td>
</tr>
<tr>
<td>1615</td>
<td>1825</td>
<td>7.8</td>
<td>2.6</td>
</tr>
<tr>
<td>1900</td>
<td>1555</td>
<td>9.1</td>
<td>3</td>
</tr>
</tbody>
</table>
total energy quantity. In contrast, the HT loop only utilized the high
available thermal power TP exhaust gas as heat source.

The interesting finding from above two references is that the LT
loop produces more power than the HT loop. However, there are a few
experimental studies that support this finding. In [25], the author as-
sumed the pinch point between the gas-state heat sources and the
working fluid was 30°C and the pinch point between the liquid-state
heat/sink sources and working fluids was 5K. These two values might
be too small, considering the size constraints placed on vehicle heat
exchangers. In [24], the LT loop condensation temperature was around
20 °C, which is difficult to achieve in the vehicle operation when the
ambient temperature is above 20 °C. However, these multi-loop ORC
systems prove the potential of considering multiple heat sources.

2.3. Expander configuration

The HDD ORC-WHR configuration and operability regime is related
to the expander type and the power output type. The two most common
expander types in HDD ORC-WHR applications are reciprocating piston
expanders and radial-inflow turbine expanders. The WHR power output
type can be electrical or mechanical. For electrical power generation,
an electrical generator is connected to the expander shaft to generate
electricity. Due to the energy is in electricity type, the energy can be
transmitted via the wires to the electrical auxiliaries or electrical
powertrain. The location of the expander in the vehicle is not limited by
the location of energy consumption components. However, for the
mechanical power output, a gearbox connects the expander shaft and
the engine crankshaft. In order to reduce the cost of the transmission
system, the mechanical expander is generally installed near the engine
crankshaft, which has more restriction than the electrical expander in
terms of vehicle integration. The distance between the expander and
the evaporator should be minimized to reduce the heat loss of the
working fluid vapor through the connecting pipes. The following sub-
sections first introduce the piston expander and turbine expander in-
dependently and then compare them in the last subsection.

2.3.1. Volumetric expanders

The volumetric expanders generally have fixed expansion ratios and
low rotational speed. Some common volumetric expanders in the HDD
field include reciprocating piston expander, scroll expander, screw ex-
ander, rotary vane expander and swash plate expander. Volumetric
expanders generally operate in a speed range of 500–6000 rpm, which
is compatible with the typical HDD engine crankshaft speed range of
500–2500 rpm. In general, volumetric expanders are mechanically
connected to the HDD crankshaft through a gearbox. AVL reported
results from a HDD ORC-WHR system utilizing a piston expander
mechanically coupled via a gearbox [57]. The ORC-WHR architecture is
shown in Fig. 1. The advantage of this configuration is the high power
transmission efficiency resulting from the low energy loss of the me-
chanical coupling. However, utilization of a gearbox adds system
complexity, cost, and puts spatial constraints on the physical expansion
machine location, i.e. the expander should be in close proximity to the
HDD crankshaft or power take off. Moreover, after coupling with the
crankshaft, the expander speed is directly correlated to the crankshaft
speed, making it difficult for the expander to operate in its maximum
efficiency range during transient engine operating conditions.

2.3.2. Turbine expanders

Radial-inflow turbine expanders utilized for ORC typically operate
in the speed range of 10,000–120,000 rpm, which is much higher than
the HDD engine crankshaft speeds. Thus, mechanical coupling of tur-
bine expanders requires a large gear ratio reduction. Both mechanical
power output and electrical power have been utilized in studies em-
ploying radial-inflow turbine expanders in HDD applications. Cummins

![Fig. 1. Schematic diagram of a HDD ORC-WHR system from AVL and Iveco [57]. (1): expander with bypass, (2): condenser, (3) reservoir and pump, and (4) exhaust evaporator with bypass.]

![Fig. 2. Schematic Diagram of a HDD ORC-WHR system from Cummins [52].]
mechanically coupled an ORC-WHR radial-inflow turbine expander to a HDD via a gearbox between the turbine shaft and engine crank, as shown in Fig. 2 [52]. Similar to the piston expander mechanical power output configuration, a mechanically coupled turbine expander yields low energy loss in the power transmission path. However, a mechanically coupled turbine expander configuration utilizes a complex gearbox and can experience non-optimal efficiencies due to limited speed flexibility. In addition, the gearbox has high gear ratio reduction due to the large speed difference between crankshaft and turbine expander. In [43], a mechanically coupled turbine and a fully electrified turbine were compared in a experimentally validated ORC-WHR system model. The constant gear ratio mechanically-coupled turbine exhibited turbine efficiencies very near those of an electrically coupled turbine when the engine operated over the steady state AVL 8 model cycle and a constant speed variable load cycle, eliminating the concern of non-optimal efficiencies by the mechanically coupled turbine expander.

2.3.3. Comparing reciprocating piston and turbine expanders

A reciprocating piston expander and a radial-inflow turbine expander were compared in [58], where the piston expander outputted mechanical power and the turbine expander outputted electrical power. The mechanical piston expander configuration requires a gearbox, increasing the mechanical complexity, although avoiding power electronics. The expander speed adjustment is limited by the gearbox and not as flexible as electrical turbine expander. In contrast, the turbine speed was optimized for maximum efficiency in real-time. However, an energy storage system is required if the electricity is not directly utilized real-time or if the electric power would be better utilized at a different operational point. Moreover, when a battery is added, a battery management system is needed, which adds complexity in terms of software. Energy loss for electrical ORC power output can be larger than mechanically coupled devices if multiple conversions of ORC power are required (e.g. transforming from turbine shaft mechanical power – generator electrical power – battery storage – electrical power).

2.4. Condenser configuration

The ORC heat rejection medium affects the final condenser configuration. The condenser can reject working fluid heat to the engine coolant or to ambient AIR either directly or through a secondary cooling loop. Many studies [35,57,59], cool down the condenser with engine coolant as shown in Fig. 3(a). This configuration minimizes system complexity, and provides flexibility regarding the physical condenser location in the vehicle integration. However, condenser heat rejection is limited by the engine coolant temperature. Others have cooled the condenser via a secondary cooling circuit [4,35,52,59–61]. The sub-radiator of the secondary cooling circuit is then located near the engine radiator as shown in Fig. 3(b), inducing spatial limitations and increasing system complexity. Even though this configuration adds complexity, it decouples the working fluid condenser outlet temperature from engine coolant temperature. Volvo utilized a water-glycol mixture as the heat rejection fluid in a secondary cooling circuit, enabling heat rejection fluid temperatures below 65ºC [4]. The ORC condenser can also be configured to reject heat to ambient AIR [62,63]. However, this increases uncertainty to the heat rejection capability of condenser and directly couples ORC operability to ambient conditions.

2.5. Working fluid pump configuration

The working fluid pump can be coupled to the engine crankshaft, the expansion device or driven by a stand-alone electric motor in HDD ORC-WHR systems. Willems et al. presented an ORC-WHR system configuration which coupled both the working fluid pump and the expansion device to the engine crankshaft, as shown in Fig. 4 [64]. This configuration saved the electrical motor and thus reduced cost. However, this configuration restricted working fluid pump operational flexibility and required other actuators to adequately control the working fluid mass flow rate. In Fig. 4, valves after the pump were utilized to bypass any working fluid supplied by the pump in excess of the evaporator demand to obtain proper working fluid mass flow control. Additionally, bypassing the surplus working fluid mass flow resulted in unnecessary energy expenditure via the pump.

In [65], the working fluid pump was coupled with a turbine expander and electric generator. Hence, the pump and turbine were both decoupled from the engine crankshaft and flexibility of their respective speed controls enhanced both energy harvesting efficiency and minimized parasitic pump losses. In addition, like the above configuration, it saved the electric motor and reduced the cost of the system.

Electrically driven working fluid pumps were utilized in [52,57,58]. Independent pump speed control maximizes working fluid mass flow rate flexibility, which improves ORC-WHR transient response. In addition, without direct or gearbox coupling to another device, the pump location is not limited, improving system packaging and contributing to the system integration.

Fig. 3. Schematic diagram of different cooling methods in HDD ORC-WHR systems: (a) condenser rejecting to engine coolant, and (b) condenser rejecting to a secondary cooling circuit [35].
3. Heat exchanger designs

HDD ORC-WHR systems utilize three categories of heat exchangers corresponding to their respective system designs: (i) evaporators, (ii) condensers, and (iii) recuperators. Evaporators and condensers exist in all ORC-WHR systems, while recuperators are utilized less frequently. Generally, an evaporator is installed downstream of the working fluid pump and upstream of the expansion device to transfer heat to the working fluid. Inside the evaporator, working fluid experiences a phase change from liquid to vapor. A condenser is installed downstream of the expansion device and upstream of the reservoir and working fluid pump. It transfers unutilized heat from working fluid to the cooling medium of choice, i.e. engine coolant or the ambient. Inside the condenser, working fluid changes phase from vapor to liquid. A recuperator transfers heat from the working fluid to the cooling medium of choice, i.e. engine coolant or the ambient. Inside the condenser, working fluid changes phase from vapor to liquid. A recuperator transfers heat from the working fluid at expander exit to the working fluid prior to the evaporator inlet, making use of otherwise unutilized post expansion heat to increase ORC efficiency and reduce condenser load.

Two types of heat exchangers are common in HDD ORC-WHR systems: shell-and-tube type [66–68], as shown in Fig. 5, and plate-and-fin type [69–72], as shown in Fig. 6. Shell-and-tube type heat exchangers are generally utilized in large-scale applications, whereas plate-and-fin type heat exchangers are used in small-scale applications due to their compactness [73]. Mavridou et al. calculated the performance of different types of exhaust gas heat exchangers given a 1300 rpm heavy duty diesel engine operating at 1300 rpm. At this operating condition, the exhaust gas temperature was 503ºC, the exhaust gas mass flow rate was 0.49 kg/s, and the heat transfer at the exhaust gas heat exchanger was 136.44 kW [74]. In this comparison, the standard plate-and-fin type heat exchanger had 66.5% less volume, 20.6% more mass and 97.5% less pressure drop than the shell-and-tube type heat exchanger. Some of the drawbacks in the standard geometry heat exchangers can be reduced by redesign. For instance, the plate-and-fin type heat exchanger mass can be reduced by replacing the fins with metal foam.

3.1. Acid dew point

In the heat exchanger design, the acid dew point [75] should be avoided by ensuring the evaporator outlet EGR exhaust gas temperature is always greater than 120–180 ºC depending on the EGR exhaust gas pressure and sulfur content. In addition, the EGR exhaust gas temperature should be kept as low as possible to ensure the high engine volumetric efficiency and enhanced NOx reduction.

3.2. Working fluid

Working fluid temperatures should not exceed working fluid decomposition temperature (e.g. 300–350 ºC for ethanol) [76]. In addition, the working fluid should reach certain superheat temperature to maintain the high expansion efficiency. At the design point, the heat exchanger heat transfer area should be able to deliver the working fluid superheated for a given working fluid mass flow rate. Working fluid pressure level places an additional constraint on the structure and material of the heat exchangers [77]. Each working fluid has its own ideal evaporating pressure with respect to maximum ORC-WHR system efficiency. Wei et al. reported that working fluid pressure range affected the ORC cycle efficiency in the heat exchanger design [78]. Higher pressure was beneficial to the ORC cycle efficiency. Thus, the heat exchanger structure and material should be reliable at high pressure working fluid operating conditions. As the working fluid
dependent evaporating pressure design point increases, the cost of the heat exchanger increases accordingly. If the cost of heat exchanger for optimal working fluid pressure is too high, the working fluid pressure should be reduced such that the cost drops to an acceptable level.

There are many types of working fluid [79–82] studied in the HDD ORC-WHR field, such as R245fa, ethanol, water, R134a, and toluene. Besides the aforementioned evaporating pressure disparities, these working fluids differ in chemical and thermodynamic properties, thus requiring disparate mass flow rates for the same heat source. For refrigerants, the required mass flow rate is high relative to ethanol or water. Hence, the working fluid pressure drop in the evaporator is high. Heat exchanger design should consider this effect. In addition, the differing corrosive natures of various working fluids affect the heat exchanger material selection and the evaporator weight. Heat exchanger weight is much more important in mobile applications than the stationary applications as it influences total hauling capacity and fuel economy. The weight of an entire ORC-WHR system is generally in the range of 50–200 kg depending on the complexity of the system. The entire ORC-WHR system from Cummins SuperTruck II was reported to be 300 lb (136 kg) as shown in Fig. 7 [5,36].

3.3. Fouling effects

Condensation of partially burned HC and soot particulates within the EGR evaporator should be considered during the full useful life of any ORC WHR system [83–86]. Hoard et al. reports that the accumulation of EGR cooler deposits reduces the device’s heat transfer performance by 20–30% for diesel engines [83]. A similar reduction in heat exchanger effectiveness within an EGR evaporator would reduce total recovered energy by a similar amount. Overall, the fouling effect is related to the chosen EGR evaporator geometry and the chosen working fluid. Utilization of a low temperature working fluid increases the deposition accumulation. Hence, the EGR evaporator design should consider fouling effects and try to minimize the fouling.

3.4. Thermal inertia

The heat exchanger thermal inertia affects the transient ORC-WHR system performance. Evaporators of greater thermal inertia require a longer time to reach designed operating temperature, but also provide a buffer against transient decreases in available waste heat. Fig. 8 shows the impact of evaporator thermal inertia on working fluid temperature response to a working fluid mass flow step change given a 0.3 kg/s mass flow rate of 320°C exhaust gas using the TP evaporator model developed in [87]. The shell-and-tube heat exchanger is modeled with the moving boundary method and the tube wall mass is manipulated to create varying thermal inertia values. Note that utilizing constant geometric variables and altering the thermo-physical properties of the wall material produces the same trends. In Fig. 8, as the tube wall mass increases, the working fluid outlet temperature exhibits slower response to the step change occurring at 100 s. When the normalized mass is 0.05, less than 200 s is required to complete the transient response. However, when the normalized mass is 1.0, the response time increases to 700 s. Thus, increasing the evaporator’s thermal inertia dampens the working fluid temperature dynamics.

The heat exchanger response time is critical for ORC-WHR control over HDD driving cycles [70,88]. Different heat exchanger response times require different control response speeds to maintain the control target (e.g. constant working fluid temperature, constant superheat, or an optimized working fluid temperature trajectory). Thus, the thermal inertia is important for the transient performance in heat exchanger design. Increasing the evaporator’s thermal inertia dampens the working fluid temperature dynamics. Pandiyarajan et al. [89] overviews an exhaust gas waste energy storage system that intentionally adds thermal inertia to the exhaust system, serving as a buffer of the transient waste energy for ORC-WHR control.

3.5. Heat exchanger sizing

The size of the heat exchanger is important in terms of vehicle packaging constraints [38] and thermal inertia. Evaporators should be small enough to integrate with the existing vehicle structure and avoid interference. However, heat exchanger size is directly related to the total heat transfer area and total thermal mass. For evaporators, if the heat transfer area is too small, the ORC-WHR efficiency will suffer [90].

Meanwhile, the condenser heat transfer area design should be designed based on condenser load and ambient conditions. Extreme operating conditions (high condenser load and high ambient temperature) need to be considered, which has seldom been reported in current literature. Condensers can reject heat to the engine coolant (90–120 ºC), a secondary cooling circuit, or ambient Air (10–50 ºC). Different heat rejection strategies require different condenser structures and heat transfer area. The working fluid temperature at the condenser outlet is limited by the temperature of cooling medium. For passive condensation pressure control, the working fluid temperature at the condenser outlet directly determines the condensation pressure, (i.e. lower working fluid outlet temperature begets a lower condensation pressure). The importance of condensation pressure is described in Section 4.2.

4. Expander designs

The expansion device extracts energy from the HDD ORC-WHR working fluid and produces either electricity or mechanical power. Expander selection and design is closely related to the type of working fluid, output power type, expansion ratio and output power level. In Fig. 7. Installation of the WHR system in a Class 8 tractor by Cummins [36].
HDD ORC-WHR systems, some common volumetric expander types are: reciprocating piston, scroll, screw, rotary vane and swashplate in design. Reciprocating piston expanders (Fig. 9) and radial-inflow turbine expanders (Fig. 10) are shown in this paper. Section 2.2 discussed the influence of volumetric and turbine expander power output type on the ORC-WHR system configuration. The current section overviews more details about these two expander types in HDD ORC-WHR applications.

Among volumetric expanders, piston expanders have the highest built in expansion ratios (6:1 – 14:1) [91]. Meanwhile, scroll expanders, although of complex geometry, do not require extra lubrication as the working fluid acts as the lubricant [92]. In addition, scroll expanders have less friction than piston expanders and require no intake and exhaust valves [93]. Screw expanders have a simple configuration and produce high efficiency under the off-design operating points with good durability [91]. Unfortunately, scroll expanders require precise machining during manufacturing to reduce leakage. Rotary vane expanders have a simple structure, are low cost and possess a flat efficiency curve for a large range of expander speeds [94]. Unfortunately, rotary vane expanders are of low capacity [80]. Swash plate expanders are compact, with low leakage and low friction [95]. However, swash plate expander efficiency is limited by their relatively low expansion ratio. Finally, unlike volumetric expanders, turbine expanders have high expansion ratios which are flexible and operate at much higher rotational speeds.

In expander design or selection, expander performance can be expressed as a function of specific speed $N_s$ and specific diameter $D_s$ [96,97]. The specific speed and specific diameter are defined in Eqs. (5) and (6) where $N_s$ is expander rotational speed in $\text{rpm}$, $v_{\text{flow}}$ is fluid volume flow in $\text{ft}^3/\text{s}$, $\Delta h_s$ is isentropic enthalpy drop in feet, and $D$ is rotor or piston diameter in feet. Latz et al. compared volumetric expanders and turbine expanders based on different HDD heat source configurations and operating conditions [20]. In [20], the authors selected $N < 4000 \text{ rpm}$ as the speed range of volumetric expanders based on [98,99] and $5000 \text{ rpm} < N < 120,000 \text{ rpm}$ as the speed range of turbine expander based on [98,100,101]. The resulting specific speed of the volumetric expander was less than the turbine expander. The authors concluded that the volumetric expander showed higher efficiency than the turbine expander. (The expander speed-efficiency map can be found in [97].) To improve the turbine expander efficiency, either the specific speed $N_s$ should increase based on [97]. Thus, either fluid volume flow $v_{\text{flow}}$ needs to be increased or isentropic enthalpy drop $\Delta h_s$ should be reduced according to Eq. (5).

\[
N_s = N \left( \frac{v_{\text{flow}}}{\Delta h_s} \right)^{0.5}
\]  

(5)

4.1. Vapor quality restrictions

In addition to the speed disparity between two categories of expander, they also tolerate different ranges of working fluid vapor quality. Volumetric expanders are not sensitive to the working fluid vapor quality and can operate with mixed liquid and vapor [30]. However, turbine expander blades are easily damaged by liquid working fluid at high rotational speeds. Thus, operation with saturated working fluid can damage the turbine expander, reducing its useful lifespan and efficiency during operation post damage. The vapor quality requirement disparity between turbine and volumetric expanders is more important during engine transients than steady state conditions. Thus, to achieve good performance during highly transient engine conditions, HDD ORC-WHR systems utilizing turbine expanders require more vapor quality control design effort than the systems with volumetric expanders. Generally, if the vapor quality drops below 1.0, the working fluid is bypassed from the turbine expander and turbine expander shuts down [105]. Xie et al. analyzed the turbine operation procedures including startup, expander turning, power generation, and turbine protection using dynamic heat exchanger models [106]. Park et al. described the HDD ORC-WHR system startup and shutdown including turbine operation for experimental purposes [7].

4.2. Expansion ratio

Operating expansion ratios differ between volumetric expanders and turbine expanders. Among the volumetric expanders: piston expanders have built-in expansion ratios (generally between 6 and 14 [33]), screw expanders expansion ratios are less than 5, scroll expanders have 1.5–3.5 built-in expansion ratios [91], and swash plate expansion ratios ranges from 5 to 15 [95]. Meanwhile, turbine expanders possess variable expansion ratio between 5 and 30 [107,108] with downstream pressures in the range of 1–3 bar [109].

Expander efficiency depends upon downstream pressure regardless of the device chosen. For the piston expanders, pressure downstream of the expander should be minimized to avoid large pumping losses. In the case of turbine expanders, downstream pressure critically influences the overall expansion ratio as small changes to the downstream pressure create large expansion ratio variations. For example, if turbine upstream and downstream pressure are 13 bar and 1.3 bar, respectively, then the expansion ratio is 10. If downstream pressure increases by 0.1 bar, the expansion ratio decreases by 7% to 9.3. By contrast, an identical 0.1 bar decrease in the upstream pressure only leads to a 1% decrease of expansion ratio. In addition, given a fixed upstream pressure, minimizing the downstream pressure maximizes the expansion ratio and generally produces higher turbine efficiency. While turbine expanders enjoy an expansion ratio advantage in HDD ORC-WHR applications, the pressure downstream of the expander should be minimized to increase the efficiency for both piston and turbine expanders.

\[ D_s = \frac{D \Delta h_s^{0.25}}{v_{\text{flow}}} \]  

(6)
For the scroll expanders, the highest efficiency does not occur at the minimum expander downstream pressure, but at the pressure corresponding to the nominal expansion ratio. Otherwise, the under-expansion or over-expansion would affect the scroll expander efficiency.

4.3. Power output range

The expander output power range for HDD ORC-WHR systems varies between experiments and theoretical analyses. Experimental studies have produced between 1 kW and 15 kW [7,54], whereas theoretical analyses rarely consider heat losses and studies have produced between 1 kW and 15 kW [7,54], whereas the-...
hydrocarbons by blending with CO₂. Garg et al. studied the CO₂ blends of isopentane and propane as working fluids [122]. It was found that adding CO₂ to isopentane causes large temperature glide and shifts pinch temperature towards warmer end of the regenerator causing a large loss in irreversibility. While adding CO₂ to propane causes the pinch temperature towards colder end of the regenerator and reduces irreversibility resulting in improved cycle efficiency. For a given source temperature of 573 K, propane/CO₂ mixture produced 17% cycle efficiency in comparison to isopentane/CO₂ mixture which generated an efficiency of about 12%. In general, for working fluid with CO₂ blend, hydrocarbon with less carbon atoms is preferred to reduce irreversibility. In a similar study conducted by Fenga et al. for six binary mixtures: CO₂ + propane/n-butane/iso-butane/n-pentane/iso-pentane/neopentane in a Trans-critical ORC system also revealed that CO₂ blend, hydrocarbon with less carbon atoms is preferred to reduce irreversibility. In a different study, Lecompte et al. thermodynamically investigated 67 types of working fluids [126]. Four cases were compared as shown in Table 4. Two heat source temperatures were considered: (i) charge air, (ii) engine coolant, (iii) EGR exhaust gas, and (iv) TP exhaust gas. Among the four working fluids, ethanol produced the highest cycle efficiency and net power. In a simulation study, Rachel et al. optimized the most energy efficient pure fluid to be used for different turbine inlet temperatures by comparing the first law efficiency (η₁) shown in Fig. 15 [120].

The general trend was that the higher molecular weight compounds produced a lower efficiency. However, higher molecular weight compounds could operate at higher temperatures than that of lower molecular weight. This might be due to the higher critical temperature and higher boiling point. Critical pressure decreased as the molecular weight increased, subsequently reducing the entropy difference between the two pressures. Therefore, the efficiency of the cycle was reduced.

In another study, Lecompte et al. thermodynamically investigated 67 types of working fluids [126]. Four cases were compared as shown in Table 4. Two heat source temperatures were considered: (i) 350°C representing typical long haul truck HDD TP exhaust gas temperature downstream of the aftertreatment system, and (ii) 500°C
Fig. 14. Transcritical ORC thermal efficiencies (a) and relative efficiencies (b) for various mixtures for a turbine inlet temperature of 453.15 K and a turbine inlet pressure of 1.3 psia [123].

Fig. 15. First law efficiency ($\eta_f$) comparison for various pure fluids for varying turbine inlet temperature ($T_i$) [120].

Table 4

<table>
<thead>
<tr>
<th>Case</th>
<th>Waste heat temperature (°C)</th>
<th>Cooling water temperature (°C)</th>
<th>Constraints</th>
<th>Expander model</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>350, 500</td>
<td>25, 50, 75, 100</td>
<td>$p_{\text{max}} &lt; 32\text{bar}$ Fixed efficiency</td>
<td>Volumetric</td>
</tr>
<tr>
<td>II</td>
<td>350, 500</td>
<td>25, 50, 75, 100</td>
<td>$p_{\text{max}} &lt; 30\text{bar}$ Fixed efficiency</td>
<td>Volumetric</td>
</tr>
<tr>
<td>III</td>
<td>350, 500</td>
<td>25, 50, 75, 100</td>
<td>$p_{\text{max}} &lt; 32\text{bar}$ VR = 5 Fixed efficiency</td>
<td>Volumetric</td>
</tr>
<tr>
<td>IV</td>
<td>350, 500</td>
<td>25, 50, 75, 100</td>
<td>$p_{\text{max}} &lt; 50\text{bar}$ VR = 5 Volumetric</td>
<td>Volumetric</td>
</tr>
</tbody>
</table>

representing EGR temperature upstream of the EGR intercooler. In addition, two types of expanders were compared: (i) a turbine expander model with fixed isentropic efficiency, and (ii) a double screw expander model with fixed volumetric ratio. The results showed that increasing condenser cooling water temperature reduced the second law efficiency in the case of a fixed isentropic efficiency expander. However, the varying condenser cooling water temperature had a negligible influence on the second law efficiency when utilizing a fixed volumetric expander efficiency. For all four cases, increasing maximum evaporation pressure increased the second law efficiency slightly. The fixed volumetric ratio (cases III and IV) showed drastically less second law efficiency compared with fixed isentropic efficiency due to a low built-in volume ratio (cases I and II). The low built-in volume ratio was constrained by the type of expander – double screw expander.

The final step in working fluid selection is experimental investigation for selected working fluids from the first two steps. Shu et al. compared R123 and R245fa in HDD ORC-WHR experiments [127]. For different working fluids, the heat exchangers did not change. Thermodynamic properties, power production, and expansion ratio were compared. The authors concluded that R123 was better suited for long-haul HD trucks, while R245fa was more suitable for city buses. Utilizing R123 as the working fluid achieved 2.5% fuel consumption reduction in the experiments.

In some other studies, researchers studied the comparison between hydrocarbons and refrigerants. Aljundi in his study on dry hydrocarbons for ORC found that thermodynamically, hydrocarbons are superior to some refrigerants [128]. It was also found that the thermal efficiencies for each working fluid correlate with its critical temperature. The best working fluid in the studied system was n-hexane. Dai et al. experimentally evaluated the thermal stability of hydrocarbon working fluid for supercritical ORC systems and suggests screening methodology of hydrocarbons based on thermal stability [129]. The results are summarized in Table 5. This study showed that the thermal stability is a function of molecular structure of a given hydrocarbon compound. Hydrocarbons with shorter chain possess higher thermal stability. The isoalkanes and cycloalkanes do not have good thermal stability, and were not recommended for supercritical ORCs.

More recent studies [130,131], have focused their efforts on a new generation of working fluids with almost no Ozone Depletion Potential (ODP) and very small Global Warming Potential (GWP). Sebastian et al. [131] and Yang et al. [130] both evaluated R1233zd-E as a drop-in replacement fluid for R245fa mainly due to similarities in their thermophysical properties. The comparison is shown in Fig. 16.

Sebastian et al. experimentally showed that using R1233zd-e increases thermal efficiency by 6.92% in comparison with R245fa [131]. In their study, the maximum gross thermal efficiency with R245fa fluid was 4.77% at a mass flow rate of 29.2 g/s whereas for R1233zd-E the recorded thermal efficiency was 5.1% at mass flow rate of 25 g/s. The difference in the mass flow rate at which the maximum efficiency

Table 5

<table>
<thead>
<tr>
<th>Temperature/°C</th>
<th>240</th>
<th>260</th>
<th>280</th>
<th>300</th>
<th>320</th>
</tr>
</thead>
<tbody>
<tr>
<td>n-Hexane</td>
<td>None</td>
<td>None</td>
<td>CH4</td>
<td>CH4 &amp; H2</td>
<td>None</td>
</tr>
<tr>
<td>n-Pentane</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>CH4</td>
<td>CH4 &amp; H2</td>
</tr>
<tr>
<td>Isopentane</td>
<td>None</td>
<td>None</td>
<td>CH4</td>
<td>CH4 &amp; H2</td>
<td>None</td>
</tr>
<tr>
<td>Cyclopentane</td>
<td>None</td>
<td>None</td>
<td>CH4</td>
<td>CH4 &amp; H2</td>
<td>None</td>
</tr>
<tr>
<td>n-Butane</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>CH4 &amp; H2</td>
</tr>
<tr>
<td>Isobutane</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>CH4 &amp; H2</td>
</tr>
</tbody>
</table>
| “None”: no product; “CH4”: organic products were detected; “CH4&H2”: hydrogen was detected.

156
occurs is due to the utilization of the same system components, affecting the filling factor in the scroll type expander. For maximum power output from turbine, it was recommended to check for material compatibility and to take into account the volume flow dependent filling factor for accurate prediction of expander performance. Similar observations were made by Yang et al. in their experimental evaluation of R1233zd-E fluid [130]. It was shown that utilization of R1233zd-E leads to approximately 3.8% better thermal efficiency in comparison to R245fa. In this study, R1233zd-E generated 4.5% more electrical power compared to R245fa due to the higher working fluid mass flow rate at the same expander rotational speed. Yang also noted that expander isentropic effectiveness could be improved by increasing the expander volume ratio as required by R1233zd-E.

The general methodology for selecting the best working fluid remains the same across published literature. Most of the comparisons in the literature are conducted under predefined temperature conditions. Proper bounding of the study influences the simulation results. Claims for the best working fluid and the cycle with highest efficiency may not hold true under operating conditions outside those considered for each study. Important factors to be carefully considered for working fluid selection include: (i) the maximum pressure constraint, (ii) expansion ratio of the expander, (iii) evaporator efficiency and sizing, and (vi) condenser outlet temperature and pressure range.

6. ORC power optimization

During the ORC-WHR architecture design, heat source selection, fluid selection, and component selection, operational optimization is important for maximum power generation. There are several parameters to be optimized: (i) evaporation pressure, (ii) superheat temperature (i.e. vapor quality), (iii) expander speed (i.e. generator speed for electrical power output or transmission ratio for mechanical power output), (vi) condensation pressure, and (v) condenser subcooling temperature.

6.1. Evaporation pressure effects

At a given engine exhaust waste heat power and working fluid mass flow rate, evaporation pressure affects the evaporation latent heat, superheat temperature, expander inlet temperature, and expander inlet pressure, which affect the expander efficiency and influence the expander power output. Fig. 17 shows a typical architecture for a parallel-evaporator HDD ORC-WHR system. In this architecture, evaporation pressure is related to the type of working fluid, working fluid pump speed, expander speed (except turbine expanders under choked flow conditions), expander bypass valve opening, and expander inlet valve opening. Lang et al. optimized the HDD ORC-WHR net power by changing the evaporation pressure [132]. In their ORC system, the TP and EGR evaporators were connected in parallel, a turbine was utilized as the expansion device, turbine inlet superheat temperature was maintained at 5K (above saturation), a regenerator was utilized, and three working fluids were compared: (i) water, (ii) cyclic siloxane D4 and (iii) cyclic siloxane D5. For water and D5, as the evaporation pressure increased, the net power first increased and then decreased. For D4, net power kept increasing with evaporation pressure and the slope gradually decreased. Among the three working fluids, water produced the maximum net power, harvesting 9.8 kW out of 56.7 kW transferred to the working fluid in the evaporator a an evaporation pressure of 29 bar.

The effects of evaporation pressure vary with the selected working fluid. Wei et al. analyzed the relationship between working fluid evaporation pressure, mass flow rate and heat source temperatures based on the constraints of different working fluids [133]. The results showed that the system utilizing R123 had higher thermal efficiency than the R245, and the ORC system mass flow rate varied with evaporation pressure. Given a fixed evaporation pressure, the thermal efficiency did not change much at different exhaust gas temperatures. Espinosa et al. analyzed the evaporation pressure influence on net power production and the results are shown in Fig. 18 [134]. Their results indicated that net power increases with evaporation pressure for all three working fluids (ethanol, water and R245fa). The ORC system utilizing ethanol as its working fluid provided the maximum net power.

![Fig. 16. (a)Vapor pressure curves for R235fa and R1233zd-E (b) T-s diagram with isobaric lines and lines with constant vapor content for R245fa and R1233zd-E [131].](image1)

![Fig. 17. Schematic of ORC-WHR system.](image2)
6.2. Superheat temperature impacts

Superheat temperature affects the expander inlet temperature, quality at the end of expansion and the expander power output. Superheat temperature is closely related to the working fluid mass flow rate, which is generally controlled through working fluid pump speed for positive displacement pumps. For the centrifugal pumps, besides the pump speed, the pressure differential across the pump also affects the working fluid mass flow rate. Because both the superheat temperature and the evaporation pressure are related to the working fluid pump speed, these two parameters are coupled. In real-world operation, the turbine bypass valve in Fig. 17 is fully closed during power production, ensuring all the working fluid flows through the expander. In addition, the turbine valve is generally fully open to minimize the pressure drop across the valve. Therefore, the evaporation pressure is only affected by the working fluid pump speed and the turbine speed. At a given turbine speed, higher working fluid pump speed produces more working fluid mass flow resulting in less superheat. The superheat temperature should be positive for the turbine expander protection. Meanwhile, for piston expanders, vapor qualities less than 1.0 are acceptable.

Several researchers have sought to optimize the ORC net power recovery via working fluid superheat optimization. Working fluid selection and expander design critically influence the level of superheat necessary for maximum ORC power generation. Variances in the latent heat of the working fluid alter the optimized level of superheat for maximum power generation of any ORC-WHR system. Quoilin et al. optimized the evaporating temperature and working fluid mass flow rate at various steady state engine conditions and developed two correlations to capture the relationships among different parameters respective to the maximum waste heat energy, as shown in Eqs. (7) and (8) [135]. These methods are evaluated over 1500s of transient heat source conditions. PID controllers are utilized to control the expander and pump speeds. Both methods recovered 6–7% available heat source energy into ORC-WHR net power with a piston expansion device.

\[ T_{\text{evap}} = a_0 + a_1 m_{\text{wf}} + a_2 T_{\text{cond}} + a_3 \log(T_{\text{hs}}) \]  

(7)

\[ m_{\text{wf}} = a_4 + a_5 T_{\text{hs}} + a_6 N_{\text{expand}} + a_7 T_{\text{cond}} \]  

(8)

In Fig. 19, Feru et al. showed that maximum net power appears when working fluid superheat is zero (vapor quality \( \chi = 1 \)) at evaporator exit in a HDD ORC-WHR system utilizing ethanol as working fluid and piston expander [88].

Using ethanol as the working fluid and a turbine expander, Xu et al. presented three working fluid temperature trajectory strategies for a HDD ORC-WHR system: (i) constant turbine inlet temperature, (ii) constant superheat temperature, (iii) rule-based superheat temperature based on exhaust waste power level [105]. The three trajectories were tracked by PID controller. The engine operated over a 1200s transient drive cycle representing typical highway driving conditions for long haul truck. As shown in Fig. 20, the rule-based working fluid superheat strategy outperformed both the constant turbine inlet temperature strategy and constant superheat temperature strategy by 2% and 1%, respectively. In addition, accumulated power was closely related to the turbine operation duration, which in turn was governed by vapor quality upstream of the turbine expander. When the working fluid fell into saturation at the evaporator outlet, the turbine was bypassed for protection. During the highly transient driving cycle, if the superheat temperature reference was not large enough, vapor quality dropped below 1.0 and no power was harvested. This behavior dominated the accumulated energy trend for low superheat references, as shown in Fig. 21. For the low latent heat working fluid utilized in [95] Xu et al. noted that, once the evaporator outlet working fluid superheat setpoint was elevated enough to maintain turbine operation for the entire transient cycle, further increases to superheat level actually reduced the total cumulative energy due to the tradeoff with working fluid mass flow rate.

A controller with more accurate reference tracking than the PIDs utilized in [95], such as nonlinear model predictive control (MPC) [87], increases the turbine operation duration during transient drive cycles, increasing cumulative waste heat power. Yebi et al. reports that MPC increases the turbine operation time and the power generation by 12% and 9%, respectively, compared with PID control [87]. Feru et al. notes a 10–15% thermal energy gain utilizing MPC and nonlinear MPC compared with PID for HDD ORC-WHR working fluid vapor temperature control [42].

6.3. Effect of expander speed on power generation

After selecting the desired ORC expansion device, the speed of the expander is optimized based on the device efficiency at the expected engine operating conditions [43]. If the expander is mechanically coupled to the engine crankshaft, like the configurations in Figs. 1, 2, and 4, the transmission ratio should be designed based on the optimized expander speed range and main engine operating speed range. If a single transmission ratio is not able to ensure the expander operates at high efficiency, multiple gear transmissions may be required, increasing the cost and complexity of the transmission integration. An expansion machine study showed that a single transmission ratio can collect 90–97% of the possible electrically generated power at large range of engine steady state operating conditions [43]. If the expander is directly connected to an electric generator, like the configuration in Fig. 17, the...
generator speed can be controlled real-time to optimize the harvested energy. The volumetric ratio of the expansion has significant impact on the net power production. Espinosa et al. studied the net power production at different volumetric ratios [134]. The results showed a net power increase with increasing volumetric ratio when using ethanol, water and R245fa as working fluids.

6.4. Influence of condensation pressure on power generation

The condensation pressure is critical for the expander power generation as it directly affects pressure downstream of the expander. Section 4.2 discusses the role of downstream pressure on pumping loss for a piston expander and expansion ratio for the turbine expander. Overall, the condensation pressure should be minimized for ORC-WHR power maximization.

Yang et al. studied the condensation pressure impact on net power output from a HDD ORC-WHR system with R123 as working fluid and a turbine expansion device [61]. The impact of condensation pressure on net power varied with engine load, where the net power considers the turbine expander power generation, working fluid pump power consumption, condenser pump power consumption and condenser cooling fan power consumption. From 25% to 50% engine load, the net power output decreased with increasing condensation pressure. During low engine load conditions, both cooling fan and condenser cooling pump were off. The increasing condensation pressure reduced the turbine expansion ratio and turbine efficiency. Thus, the net power decreased with increasing condensation pressure. However, when the engine load was in the range of 50–100%, the net power first increased and then decreased with increasing condensation pressure. At the high engine load conditions, both the condenser cooling fan and cooling pump were on. The cooling fan and pump power consumption were related to the temperature difference between the working fluid and coolant. On the working fluid side, as the condensation pressure increased, the condensation temperature increased as well. On the coolant side, the inlet coolant temperature was fixed. Thus, the increasing condensation pressure created a larger temperature difference between the working fluid and the coolant. The large temperature difference was beneficial to the heat transfer and reduced the cooling power consumption. Therefore, the net power increased. As the cooling power consumption reduced to certain level, the room for further power reduction was minimal. Thus, the net power decreased due to the lower turbine efficiency.

6.5. Impacts of condenser subcooling temperature

The condenser subcooling temperature directly affects the evaporator inlet temperature. Utilizing less subcooling (higher condenser outlet temperatures) leads to higher evaporator inlet temperatures [43]. At given values of post evaporator working fluid superheat and exhaust gas waste heat power, the working fluid mass flow rate will increase with elevation of working fluid temperature upstream of the evaporator, increasing the expander power output. For a given working fluid mass flow rate and exhaust gas waste power, reduction in subcooling and the subsequently higher evaporator inlet temperatures will increase the level of post evaporator superheat, increasing the expander power output.

Subcooling temperature also influences condensation pressure. As the level of subcooling decreases (higher condenser outlet temperatures), the condensation pressure increases. The principle is similar to the evaporation pressure relation with superheat temperature at constant working fluid mass flow rate and varying exhaust gas waste power. When the exhaust gas waste power increases, both the superheat temperature and evaporation pressure increase and vice versa. The resulting variation of condensation pressure affects the expander efficiency.

Condensers can either utilize engine coolant or a secondary refrigerant circuit to reject heat. The condenser subcooling temperature is controlled by the coolant mass flow rate and temperature entering the condenser. If the condenser is directly cooled by ambient AIR, the subcooling temperature is affected by vehicle speed, condenser location, and ambient conditions. Regardless of which condenser heat rejection method is utilized, subcooling temperature is eventually coupled to the environment temperature.

6.6. Combining different factors

Complex ORC-WHR architectures contain several actuators and often actuator coupling for proper functionality. Thus, the independent impacts of each actuator and combinations of all the actuators should be investigated for ultimate ORC-WHR power optimization. Xu et al. investigated the sensitivity expander power generation to four actuators in a parallel evaporator HDD ORC-WHR system given two steady state engine conditions [43]. The system architecture is shown in Fig. 17 and the system model is a dynamic model, which can be found in [136]. The four actuators included: (i) the high pressure pump speed determining overall working fluid mass flow, (ii) the distribution valves splitting the working fluid mass flow to the parallel evaporators, (iii) the turbine expander speed, and (iv) the cooling water pump speed. The upper and lower boundaries for the high pressure pump speed and distribution valve splitting ratio were determined by the working fluid saturation and decomposition temperature limits. Turbine expander speed was constrained by the gearbox reduction ratio and engine speed. The cooling water the pump speed lower and upper boundaries were determined by the working fluid evaporation temperature at the high pressure pump inlet and characteristics of condenser coolant pump. The results of the optimization are shown in Fig. 22. When operating within the aforementioned constraints, the working fluid pump speed.
influenced turbine power generation more than any other actuator. Turbine speed and cooling water pump speed also produced a noticeable impact on the turbine power generation. However, the working fluid split ratio across the TP and EGR evaporators had negligible effect on the power generation provided the working fluid was maintained in a vapor state prior to the turbine expander inlet.

Peralez et al. applied a dynamic programming optimization (DP) algorithm to a single evaporator HDD ORC-WHR system, which generated global maximum power trajectories given transient engine conditions [62]. The control actuators considered were the exhaust gas vaporator bypass valve and the mass flow rate of condenser cooling air. Optimal trajectories were obtained for expander inlet temperature, condenser temperature, evaporator pressure, and condensation pressure over the transient conditions. The DP strategy improved energy recovery by 7% compared with a baseline control over the 800 s transient engine conditions.

7. ORC control strategies

The HDD ORC-WHR system often undergoes rapid transitions from the nominal operating point to different off-design conditions, due to the transient heat source. Optimal system operation is only possible in a narrow range of working fluid evaporating pressure and temperature. The maximum applicable operating temperature is constrained by working fluid degradation while the lower temperature constraint is determined by condensation of working fluid in the turbine expander. By comparison, piston expanders have more relaxed constraints on the lower working fluid temperature. To address the control challenges of maintaining optimal/efficient operation within acceptable safety margins, some control design schemes have been proposed in the literature. Many existing control schemes focus on very common control approaches, namely, feedback plus feedforward [137], PI-based decentralized control [135], and gain-scheduled PI type control [138]. These studies agree that the best ORC-WHR performance is obtained with regulation of evaporating pressure and superheating temperature. In these control designs, maintaining a minimum superheating temperature at the evaporator outlet is sought to guarantee both safe operation and maximum energy recovery. However, such control schemes may not offer satisfactory results for highly transient HDD heat source profiles for two main reasons: (i) the reference trajectories for evaporation pressure and temperature are generated through steady state optimization, and (ii) the disturbance rejection capability of traditional feedback control methodologies is very limited.

In order to address the limitations of traditional control strategies, recent studies proposed the use of advanced control strategies including: nonlinear/linear model predictive control [88,139,140], supervisory predictive control [141], and extended prediction self-adaptive control [142]. In most of these works, simulation studies confirmed the superior control performance of advanced control over traditional PID control. Most of these studies were focused on precise control of either a minimum superheating temperature or a vapor quality close to unity. These ORC control variables were largely assumed to be both safety and performance indicators. Although this assumption is true for ORC systems utilizing low latent heat working fluids, there is no systematic analysis for a large group of working fluid choices with the exception of steady state optimization work for a very limited working fluid selection [143]. In the authors’ recent steady state ORC optimization with ethanol as working fluid [43], maximum power production is not correlated with minimization of superheating temperature. Instead, a dome-shaped power curve was discovered which indicates a decrement of power only after a minimum level of superheating temperature.

Two-level control strategies are proposed in literature to address the aforementioned limitations of the above cited advanced control schemes. These two-level strategies consider optimal evaporation pressure or temperature trajectory generation in the high level and reference tracking output feedback control or model predictive control in the lower level [62,144]. Different methods are considered for reference generation including offline and online methods for formulating optimization problems to maximize net power within defined safety constraints. Experimentally determined empirical relations of optimal evaporating temperature and maximum power were considered as one of the offline methods [145] while dynamic programming was utilized in [62] for a systematic and global optimization result. In [62], dynamic real-time optimization generated an optimal trajectory with a slow update rate, and produced a significant net power gain in the presence of a process disturbance when compared to the operational trajectory generated offline. However, the model used for real-time optimization in [62] was a simplified, 2-state (0D heat exchanger model), which might compromise the optimal control performance for operating conditions away from the calibration points and for highly transient heat sources. In [144], a perturbation-based extremum-seeking algorithm was considered to generate an online optimal temperature trajectory based on a lumped model developed by system identification method.

In addition to the well-researched ORC control problem of efficient and safe operation, there are additional challenges depending on the ORC system design and configuration. If both EGR and TP exhaust gas are considered as heat sources in a HDD ORC-WHR system, parallel evaporators are often utilized and are reported to be more efficient than series evaporator configurations [35]. For a parallel evaporator ORC system design, a coordinated control effort is needed to split working fluid mass flow through each evaporator in proportion to the transient heat source power supplied to each evaporator. For parallel evaporator configurations utilizing a single pump, coordinated actuation of working fluid mass flow through both evaporators for mixed evaporator outlet working fluid temperature control is not trivial. In addition to the need for coordinated actuation, the coupled nonlinear system dynamics and the distinctly different time constants of the two parallel evaporators interacting with different heat sources makes control of the mixed working fluid temperature at the expander inlet a difficult problem. In this regard, only a single study has been published [146], which considered explicit penalization of the temperature difference between evaporator outlets in nonlinear model predictive control formulation.

There are several interesting ORC-WHR control research studies from Willems and Perus’s group integrating the ORC-WHR control with the engine control [147]. Even though this concept increases the complexity of the WHR control, incorporating the engine control facilitates changes to the TP exhaust gas and EGR exhaust gas mass flow. These strategic operational alterations could benefit the WHR energy recovery if the heat source mass flow increases or temperature increases without compromising the engine performance. Willems et al. developed an Integrated Powertrain Control (IPC) in [147], where the WHR system model only considered the energy balance in TP and EGR evaporators. The results reported 2.8% CO₂ emission reduction by adding the ORC-WHR system. Later, a dynamic ORC-WHR system model was added and the piston expander was mechanically coupled to the engine crankshaft [148], resulting in a 2.6% CO₂ reduction while
simultaneously meeting NOx limits. In [149], the piston expander was decoupled from the engine crankshaft and an electric generator was connected to the expander. The results showed the IPC achieves 3.5% CO₂ reduction, 2.5% AdBlue dosage reduction, and 19% DPF particulate matter reduction compared with baseline powertrain without WHR system. The diagram of the IPC design from [149] is presented in Fig. 23. The goal of the optimization was to minimize the weighted sum between the fuel mass flow \( m_f \), the AdBlue mass flow \( m_A \), and an equivalent cost associated with the fuel consumption for active DPF regeneration \( m_{DPF} \).

8. ORC simulation studies

The HDD ORC-WHR simulation literatures are listed in Table 6. ORC system simulation can be categorized by the heat exchanger modeling methodology utilized, namely static versus dynamic models. Generally, the ORC system development procedure can be illustrated by Fig. 24. Static models are best implemented to analyze the energy flow and cycle efficiency at the concept development phase. With the help of static models, heat source selection, working fluid screening, expansion device selection, and cycle efficiency analyses can be conducted. Dynamic models are employed in the component development, control development, and power optimization phases due to their enhanced transient capabilities and accuracy.

8.1. Simulation with static heat exchanger models

Static models do not consider system dynamics (i.e. no ordinary differential equation in the model). Thus, static models are only utilized in steady state engine analyses. In general, static models only consider the energy balance between the heat sources and working fluid, which makes the models easy to implement and reduces simulation time. Due to this advantage, static models are generally utilized in the concept phase of the ORC-WHR system development as shown in Fig. 24. In this phase, the static models provide rough calculations of ORC system efficiency and net power in different scenarios, such as: varying the number of heat sources [1,21], comparing working fluids [35], and evaluating different heat exchangers [78]. Teng et al. [1] and Arunachalam et al. [21] utilized static models to evaluate the potential of different heat sources including TP, EGR, and CA gases. Teng et al. found that utilizing different working fluid pressures in different heat exchangers could improve the ORC-WHR system performance with multiple heat sources. In addition, they determined that it was not practical to consider engine coolant as a heat source due to packaging issues introduced by the engine coolant heat exchanger and corresponding condenser. Meanwhile, Arunachalam et al. concluded that it was economical to consider only the EGR as an ORC heat source due to the high cycle efficiency. By considering TP exhaust gas and CA as additional heat sources, the cycle efficiency dropped from 22.2% to 15.6% without considering increases in system cost, weight and complexity caused by the utilization of additional heat sources. Grelet et al. utilized a static model to quickly analyze 13 working fluids at different steady state engine conditions and ORC condensing temperatures [35]. Ethanol and acetone were found to be the best working fluids based on the weighted net power production over 13 operating conditions. Finally, Wei et al. utilized a static model to evaluate heat exchanger design parameters before the design finalization [78]. The optimal design of heat exchanger was found to achieve 10–15% cycle efficiency, where the cycle efficiency is defined in Eq. (9).

\[
\eta_{ORC} = \frac{P_{\text{expander}} - P_{\text{pump}} - P_{\text{cooling}}}{P_{\text{evaporator}}} \times 100\%
\]  

(9)

where \( P_{\text{expander}} \) is the expander power generation, \( P_{\text{pump}} \) is working fluid pump power consumption, \( P_{\text{cooling}} \) is condenser cooling system power consumption, \( P_{\text{evaporator}} \) is the power absorption by preheater, evaporator, and superheater.

8.2. Simulation with dynamic heat exchanger models

The main advantage of dynamic heat exchanger models is their capability of simulating over transient engine operating conditions. One popular dynamic heat exchanger model is the finite volume model (FVM). The advantages of a FVM are its high accuracy and stability. However, FVMs obtain those advantages through a compromise with computational cost. Grelet et al. built and validated two FVMs for their TP and EGR evaporators, respectively [35]. These FVMs were utilized to evaluate condenser cooling configurations, parallel and series evaporator configurations, and net ORC-WHR power production during both steady state and transient engine operation. The results showed that, utilizing a condenser coolant-loop that was independent from the engine coolant circuit increased the ORC-WHR net power production relative to a system where the condenser was directly rejecting heat to the engine coolant. In addition, the parallel-evaporator configuration produced more net power than the serial-evaporator configuration. Compared with steady state engine conditions, transient driving cycles produced less net power, which was also reported by Xie et al. in [106]. Finally, Xu et al. developed FVM models for TP and EGR evaporators, which were validated via ORC-WHR system experiments. In these FVMs, a novel pressure drop model was included and validated [136]. In addition, an engine model was built and co-simulated with the ORC system model over a Constant Speed Variable Load HDD engine driving cycle to simulate long haul operation.

Another popular dynamic heat exchanger model, the moving boundary model (MBM) is computationally efficient, but compromises its accuracy and stability. Due to its low computation cost, MBMs are widely utilized in ORC-WHR model-based control. Yebi et al. implemented a MBM within a nonlinear model predictive control (NMPC) design and revealed the advantages of NMPC over PID in ORC-WHR system power generation and cycle safe operation [87]. Peralez et al. utilized a MBM in a feedforward plus feedback control, which outperformed a stand-alone PID control in the ORC-WHR vapor temperature control [138].

Based on the literature listed in the simulation table, the fuel economy improvement brought by the ORC-WHR system in the HDD

---

Fig. 23. Schematic representation of the integrated powertrain control with an electrified HDD ORC-WHR system [149]. DOC is the diesel oxidation catalyst, DPF is the diesel particle filter, SCR is the selective catalytic reduction system, and AMOX is ammonia oxidation catalyst. \( m_B \) is fuel mass flow, \( u_{TPG} \) is variable turbine geometry rake position, \( u_{EGR} \) is EGR valve position, \( m_A \) is urea dose mass, \( P_E \) is engine alternator power supply to the battery, \( P_B \) is power transmission between WHR generator/motor and battery, \( u_{EGP} \) is exhaust gas bypass valve position, \( u_{PTG} \), \( u_{ETG} \), \( u_{EGP} \) are the speed of EGR evaporator working fluid pump, TP evaporator working fluid pump and expander, respectively, \( P_{WHR} \) is the power output from WHR system.


<table>
<thead>
<tr>
<th>Year</th>
<th>Reference</th>
<th>HDD displacement/rated power/model</th>
<th>Heat source</th>
<th>E(^a) type</th>
<th>E(^b) type/PO</th>
<th>Working fluid</th>
<th>EC(^e) (kW)</th>
<th>ORCE(^f)/FEI(^g) (%)</th>
<th>HM(^h)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2006–2007</td>
<td>Teng et al. [1,150]</td>
<td>275 kW/Cummins ISX</td>
<td>TP, EGR, CA</td>
<td>–</td>
<td>–/–</td>
<td>Water, R134a, R245fa</td>
<td>SS –50–55</td>
<td>29.5/–</td>
<td>S</td>
</tr>
<tr>
<td>2009</td>
<td>Teng et al. [86]</td>
<td>–/-Cummins ISX</td>
<td>EGR, CA</td>
<td>–</td>
<td>–/E</td>
<td>R245fa, ethanol</td>
<td>SS –/–</td>
<td>15.8–25.5/–</td>
<td>S</td>
</tr>
<tr>
<td>2011</td>
<td>Wei et al. [78]</td>
<td>8.6 L/258 kW/</td>
<td>TP</td>
<td>–</td>
<td>–/–/S(^k)/M(^l)</td>
<td>R123</td>
<td>SS 2–20/–</td>
<td>12–16/–</td>
<td>S</td>
</tr>
<tr>
<td>2012</td>
<td>Aramachalam et al. [21]</td>
<td>13 L/367 kW/Volvo D–13</td>
<td>TP, EGR, CA</td>
<td>–</td>
<td>–/T/E</td>
<td>Mixture (80% water and 20% Methanol)</td>
<td>SS 2–22/–</td>
<td>9–23/–</td>
<td>S</td>
</tr>
<tr>
<td>2012</td>
<td>Dole et al. [26,34]</td>
<td>12 L/-/–</td>
<td>TP, EGR, CA</td>
<td>–</td>
<td>–</td>
<td>R245fa, water</td>
<td>SS 46–59/–</td>
<td>5.7–26.3/–</td>
<td>S</td>
</tr>
<tr>
<td>2013</td>
<td>Lang et al. [132]</td>
<td>–/-</td>
<td>TP, EGR</td>
<td>Plate T/E</td>
<td>Siloxane (D4, D5), water</td>
<td>SS 6–11</td>
<td>5–19/–</td>
<td>S</td>
<td></td>
</tr>
<tr>
<td>2013</td>
<td>Xie et al. [106]</td>
<td>11.6 L/353 kW/Weichai WP12-480</td>
<td>TP</td>
<td>Tube T/E</td>
<td>R123, R245fa, water</td>
<td>TS 0–25</td>
<td>3–4/–</td>
<td>D(^o)</td>
<td></td>
</tr>
<tr>
<td>2014</td>
<td>Grelet et al. [53]</td>
<td>–/-</td>
<td>TP, EGR</td>
<td>Tube T/E</td>
<td>Ethanol, alicyclic, cyclopentane, water</td>
<td>SS –/–</td>
<td>2–5/–</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>2014</td>
<td>Luong et al. [151–153]</td>
<td>13 L/-/–</td>
<td>EGR</td>
<td>Tube -/M</td>
<td>Mixture (52% ethanol and 48% water)</td>
<td>TS 0–5</td>
<td>–/–</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>2015</td>
<td>Amicable et al. [109]</td>
<td>10.8 L/2006MY Cummins ISX</td>
<td>EGR</td>
<td>Plate T/E</td>
<td>Ethanol, R245fa, pentane</td>
<td>SS 6–9</td>
<td>–/–</td>
<td>S</td>
<td></td>
</tr>
<tr>
<td>2015–2016</td>
<td>Grelet et al. [35,56,154]</td>
<td>11 L/320 kW/-</td>
<td>TP, EGR</td>
<td>Tube T/E</td>
<td>Mixture (20% ethanol and 80% water)</td>
<td>TS –/–</td>
<td>–/–</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>2015</td>
<td>Ayagi et al. [45]</td>
<td>10.5 L/-/–</td>
<td>TP, EGR, CA</td>
<td>–</td>
<td>–/–/E</td>
<td>Water</td>
<td>TS 0–21</td>
<td>&lt;27–2.9</td>
<td>D</td>
</tr>
<tr>
<td>2016</td>
<td>Feru et al. [149]</td>
<td>–/-six cylinder Euro-VI</td>
<td>TP, EGR</td>
<td>Plate-fin P(^n)/E</td>
<td>Ethanol</td>
<td>TS 0–15</td>
<td>&lt;35</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>2016</td>
<td>Robertson et al. [115]</td>
<td>11.7 L/Scania DCG2 Euro III</td>
<td>TP</td>
<td>Tube-shell T/E</td>
<td>R245fa</td>
<td>TS 0–5</td>
<td>6–9/–</td>
<td>S</td>
<td></td>
</tr>
<tr>
<td>2016</td>
<td>Stanzel et al. [59]</td>
<td>12.8 L/315 kW/Mercedes-Benz Actros OM 471</td>
<td>TP, EGR</td>
<td>Plate T/E</td>
<td>Ethanol</td>
<td>TS 2–9</td>
<td>&lt;2–3</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>2016</td>
<td>Xu et al. [43]</td>
<td>13 L/-/–</td>
<td>TP, EGR</td>
<td>Tube-shell T/E</td>
<td>R245fa</td>
<td>TS 2–9</td>
<td>&lt;2–3</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>2017</td>
<td>Andwari et al. [155]</td>
<td>10.3 L/316 kW/IVECO Cursor 10</td>
<td>TP</td>
<td>Tube-shell T/E</td>
<td>R245fa</td>
<td>TS 2–16</td>
<td>&lt;2–8</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>2017</td>
<td>Chen et al. [111]</td>
<td>11.1 L/298 kW/-</td>
<td>TP, CA</td>
<td>Plate T/M</td>
<td>R245fa</td>
<td>TS 6–29</td>
<td>&lt;5–9</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>2017</td>
<td>Lion et al. [110]</td>
<td>–/-300 kW/-</td>
<td>TP, EGR</td>
<td>–/–/–/–/–/–</td>
<td>Ethanol, water, toluene, ethanol</td>
<td>SS 6–36</td>
<td>&lt;11</td>
<td>S</td>
<td></td>
</tr>
<tr>
<td>2017</td>
<td>Rijpkema et al. [125]</td>
<td>12.8 L/-/–</td>
<td>TP, EGR, CA, engine coolant</td>
<td>Tube T/E</td>
<td>Cyclopentane, ethanol, R245fa, water</td>
<td>SS 2–8</td>
<td>–/–</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>2017</td>
<td>Trabucchi et al. [156]</td>
<td>–/-</td>
<td>TP, EGR</td>
<td>Plate T/M</td>
<td>Hexamethyldisiloxane</td>
<td>TS 5</td>
<td>–/–</td>
<td>D</td>
<td></td>
</tr>
</tbody>
</table>

\(\text{a} \) Evaporator.  
\(\text{b} \) Expander.  
\(\text{c} \) Expander power output.  
\(\text{d} \) Engine conditions.  
\(\text{e} \) Expander power.  
\(\text{f} \) ORC efficiency.  
\(\text{g} \) Fuel economy improvement.  
\(\text{h} \) Heat exchanger model types.  
\(\text{i} \) Steady state engine conditions.  
\(\text{j} \) Static heat exchanger model.  
\(\text{k} \) Turbine expander.  
\(\text{l} \) Electrical power output.  
\(\text{m} \) Scroll expander.  
\(\text{n} \) Mechanical power output.  
\(\text{o} \) Transient engine conditions.  
\(\text{p} \) Dynamic heat exchanger model.  
\(\text{q} \) Piston expander.
vehicle ranges from 2% to 11%. The majority of simulation work showing fuel economy benefits consider TP and EGR gases as their heat resources.

9. Experimental ORC investigations

This section discusses the HDD ORC-WHR experimental studies from the past decade. Due to the cost and complexity of HDD ORC-WHR systems, experimental results are few and far between in the literature. A summary of HDD ORC-WHR studies is provided in Table 7.

In 2009, Cummins presented 5–10% HDD fuel economy improvement by implementing ORC-WHR technology with a 2007MY Cummins ISX HDD engine, as shown in Fig. 25 [157]. The system architecture was compact and the ORC components, utilizing both EGR and TP exhaust gases as heat source, were closely attached to the engine. R245fa was selected as the working fluid and waste heat power was generated with a turbine expander coupled to an electric generator. Test results from single heat source (EGR exhaust gas only) and two heat sources (EGR plus TP exhaust gas) were compared, revealing that the EGR heat source contributed to the fuel economy gains more than the TP exhaust gas. Fuel economy improvement peaked at the B25 engine condition for this system.

In 2013, Cummins presented the results from the US SuperTruck-I program [5]. As shown in Fig. 26, the ORC-WHR system contributed a 3.5% absolute brake thermal efficiency improvement. In addition to the efforts of Cummins, other HDD manufacturers (Daimler, Volvo, and Navistar) also demonstrated 1–2% absolute brake thermal efficiency improvements utilizing WTHR technology in the SuperTruck-I program [5]. Fig. 7 shows the vehicle installation of the Cummins SuperTruck-I ORC-WHR system.

AVL presented HDD ORC-WHR experimental results in 2011, achieving a 3–5% fuel economy improvement [7,101]. They utilized a 10.8 L 2006MY Cummins ISX HDD engine, ethanol as the working fluid, and a turbine expansion device in the architecture shown in Fig. 27. TP and EGR evaporators were connected in series and an air brake compressor was used as the turbine expander load. Startup and shutdown procedures were presented in detail. The experiments were conducted over four steady state engine operating conditions and produced 3–12 kW. In addition to the power generation results, system payback time was analyzed based on the fuel economy improvement. Assuming a system cost of $8500 and a diesel price of $3.0 per gallon, the 3.5% fuel consumption reduction resulted in 700 gallons of annual fuel savings, which converted to a four-year payback time. If the fuel reduction was 6.0%, the annual fuel savings would be 1200 gallons and the payback time decreased to 2.4 years.

Bosch presented HDD ORC-WHR experimental results in 2012 which recovered 2–9 kW of power with a turbine expander and water as the working fluid [54]. Their system utilized parallel TP and EGR exhaust gas evaporators as shown in Fig. 28. The piston expander referenced in Fig. 28 was only simulated. The experiments were conducted at three steady state engine conditions.

Bettoja et al. presented experimental data for two HDD ORC-WHR systems from Volvo and Centro Ricerche Fiat [4]. In the Volvo HDD ORC-WHR system, a 12.7L Euro VI US10 was outfitted with TP and

![Fig. 24. ORC-WHR system development procedures.](image)

### Table 7

<table>
<thead>
<tr>
<th>Year</th>
<th>Reference</th>
<th>HDD displacement/rated power/model</th>
<th>Heat source</th>
<th>E\text{\textsuperscript{a}} type</th>
<th>EC\text{\textsuperscript{b}}</th>
<th>E\text{\textsuperscript{b}} type</th>
<th>Working fluid</th>
<th>ORC\text{\textsuperscript{c}}</th>
<th>EFF\text{\textsuperscript{d}} (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2009</td>
<td>Nelson C.R. [48]</td>
<td>-/-/Cummins ISX</td>
<td>TP, EGR, CA</td>
<td>-</td>
<td>T/E</td>
<td>R245fa</td>
<td>SS</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2010</td>
<td>Nelson C.R. [48]</td>
<td>-/-/Cummins ISX</td>
<td>TP, EGR, CA</td>
<td>-</td>
<td>T/E</td>
<td>R245fa</td>
<td>SS</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2011</td>
<td>Pang et al. [73]</td>
<td>12L/13B Toyota</td>
<td>TP, EGR</td>
<td>-</td>
<td>T/E</td>
<td>Water</td>
<td>TS</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2012</td>
<td>Saber et al. [80]</td>
<td>12L/13B Toyota</td>
<td>TP, EGR</td>
<td>-</td>
<td>T/E</td>
<td>Water</td>
<td>TS</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2013</td>
<td>Ruscet et al. [86]</td>
<td>15L/13B Cummins ISX</td>
<td>TP, EGR</td>
<td>-</td>
<td>T/E</td>
<td>Water</td>
<td>TS</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2014</td>
<td>Bari et al. [90]</td>
<td>3.4L/13B Toyota</td>
<td>TP, EGR</td>
<td>-</td>
<td>T/E</td>
<td>Water</td>
<td>TS</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2015</td>
<td>Stanton D.W. [5,36]</td>
<td>15L/-/Cummins ISX</td>
<td>TP, EGR</td>
<td>-</td>
<td>T/E</td>
<td>R245fa</td>
<td>SS</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2016</td>
<td>Feru et al. [42]</td>
<td>-/-/six cylinder Euro-VI TP, EGR</td>
<td>Plate, Shell</td>
<td>PP</td>
<td>P/M</td>
<td>Ethanol</td>
<td>TS</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2017</td>
<td>Furukawa et al. [60]</td>
<td>9L/-/TP, EGR, engine coolant</td>
<td>TP, EGR</td>
<td>-</td>
<td>T/E</td>
<td>Hydrofluro ether</td>
<td>SS</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2018</td>
<td>Bettoja et al. [4]</td>
<td>12.7L/317 kW/Euro VI US10</td>
<td>Plate, Shell and plate</td>
<td>PP</td>
<td>P/M</td>
<td>Mixture (Ethanol and water)</td>
<td>SS</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2019</td>
<td>Can et al. [61]</td>
<td>11.6L/353 kW/Weichai WP12.480</td>
<td>TP, EGR</td>
<td>-</td>
<td>T/E</td>
<td>Water</td>
<td>TS</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2020</td>
<td>Xu et al. [136]</td>
<td>13L/13B Toyota</td>
<td>TP, EGR</td>
<td>-</td>
<td>T/E</td>
<td>Water</td>
<td>TS</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2021</td>
<td>Alshammari et al. [160]</td>
<td>7.3L/284 kW/SIgula</td>
<td>TP, EGR</td>
<td>-</td>
<td>T/E</td>
<td>Ethanol</td>
<td>SS</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2022</td>
<td>Li et al. [161]</td>
<td>8.4L/243 kW/-</td>
<td>TP, EGR</td>
<td>-</td>
<td>T/E</td>
<td>CO\text{\textsubscript{2}}</td>
<td>SS</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
EGR exhaust gas evaporators in series. A water and ethanol mixture was chosen as working fluid. The experimental results showed 9–11% ORC efficiency over large range of steady state engine conditions, producing 1–8 kW via a turbine expander. In the Centro Ricerche Fiat HDD ORC-WHR system, an 11.1 L Euro VI IVECO Cursor 11 engine was outfitted with only a TP exhaust gas evaporator for the purpose of system simplicity. The experimental results showed 1–3 kW turbine power generation, which corresponds to 2–4% ORC cycle efficiency.

Xu et al. validated a HDD ORC-WHR system dynamic model using experimental data from a 13L HDD engine and ORC-WHR system utilizing parallel EGR and TP exhaust gas evaporators [136]. A turbine expander was integrated with a generator for electrical power generation using ethanol as the working fluid.

Considering all the experimental work stating fuel economy benefits numbers, the experimental fuel economy improvements made by ORC-WHR system in HDD vehicle are 2–10%. Unlike simulation studies, nearly half of experiments showing fuel economy benefits were conducted with TP heat source only. On the other end of the complexity spectrum, two experiments utilized more than two heat sources: TP, EGR and CA in reference [48], TP, EGR and engine coolant in reference [60]. In these experiments, cost was not considered as the constraint and the goals were to maximize the fuel economy benefits.

10. Fuel economy and CO₂ emission impacts

The ultimate goal of HDD ORC-WHR for the vehicle manufacturers and customers is the fuel economy improvement. For the society, the goal is the reduction of greenhouse gases and pollutants emission like CO₂, CO, and NOx. The fuel economy improvement comes from the engine fuel saving, which is compensated by the power production of the ORC-WHR system. To quantify the fuel economy and CO₂ emission benefits, this section organized the literatures which addressed the benefits of these two factors. The overall summary is shown in Table 8. The table is grouped by maximum fuel economy improvement and maximum CO₂ reduction. Each group is further separated by percentage changes. In each column, the references are differentiated by simulation, experiments in test rig and experiments on road. It shows that very few experiments are conducted on the vehicle and most of experiments are in test rig with engine and ORC system. Most of references are based on simulation. The maximum fuel economy improvement are all above 2% for simulations and experiments, which shows the minimum boundary for this ORC-WHR technology in the HDD application and it is a substantial improvement. The fuel economy improvement distributes uniformly among the three regions of percentage. The difference of fuel economy improvement could results from multiple factors such as heat sources, working fluid, expander, and engine operating conditions. Most of literatures mentioned CO₂ reduction in the motivation of the HDD ORC-WHR development, but only a few presented the CO₂ reduction in their results [45,55,57,59]. Due to the less fuel consumption, less greenhouse gases and pollutants are produced to the environment as referred in [157].

11. Limiting factors

Even though ORC-WHR technology is promising in terms of improving fuel economy, some limiting factors need to be carefully considered. The noticeable limiting factors include: (i) system complexity and interaction with an existing powertrain [172–174], (ii) safety [175,176], (iii) temperature control during driving cycles [35,177,178], (iv) system durability [179,180], (v) cost [73,77,181–183] and (v) system weight [172,173,184]. Unlike the TEG
or turbo-compounding, ORC-WHR adds multiple components to the existing powertrain. As shown in Fig. 7, the ORC components are tightly packaged around the engine. The added components will change under hood aerodynamics, which affects the heat release to the environment and the aerodynamic resistance [63,172]. A backpressure increase results from the ORC-WHR evaporator pressure drop and leads to the increase of engine pumping loss. To minimize the backpressure increase, the evaporator designs need to reduce the flow resistance on the heat source side, which may reduce the evaporator’s heat transfer capability [185]. In the engine side, variable geometry turbine seems to be a good solution to withstand increased backpressure [174]. If the condenser coolant is independent of the existing engine coolant system, there is no interaction between the two coolant systems. However, in order to reduce the complexity and avoid independent condenser coolant radiator, the engine coolant may be utilized as the condenser heat rejection medium as shown in Fig. 3(a). This configuration changes the existing engine coolant operating conditions and may require recalibration of the engine cooling system.

Safety is another concern for onboard ORC-WHR system. The main safety concern is the working fluid flammability and toxicity [81,186]. For example, ethanol is flammable when exposed to oxygen at elevated temperatures. When ethanol is considered as working fluid, the system sealing is extremely important.

Many ORC-WHR studies control showed satisfactory working fluid superheat control performance [87,88]. However, the engine operating conditions were either step changes or mild transients, not driving cycles. For the HDD engine driving cycles, the heat source power profile is much more dynamic and challenges the working fluid superheat controller [178]. Many ORC-WHR HDD vehicle application conclusions were drawn from steady state analysis. Considering the disparate power recovery between the steady state and transient shown in [35], the power recovered from steady states may be discounted by 50–60% during transient operation.

Since vehicle application of ORC-WHR system is in the initial stages, system durability is not well researched. ORC systems installed on vehicles are exposed to vibration from engine, transmission, and suspension system during their lifetime [179]. As the systems age, sealing of the working fluid high pressure section, corrosion between working fluid and components material [180], lubrication of the expander, fouling of the evaporators, leakage of the working fluid need to be considered.

The addition of the weight is another concern in the component selection and design. The main weight contributing components are heat exchangers [172], expander [180] and working fluid. The negative fuel economy effect of ORC system weight increase is not well researched at this time. Current experimental campaigns were conducted on a test rig [187] or in a prototype vehicle [57] without separating the effect of ORC-WHR system weight. Component weight and fuel economy impact should be considered during the ORC system development phase.

12. Summary

12.1. HDD ORC-WHR system architecture

There are four main components dictating HDD ORC-WHR system architectures, namely: the evaporator(s), expander, working fluid pump and condenser. The number of heat sources delivering heat to the working fluid through the evaporators needs to be carefully considered. Most HDD ORC-WHR studies consider only TP and EGR exhaust gas as heat sources due to the high temperature and mass flow of the TP exhaust gas and high temperature of the EGR gas.

In order to increase the ORC-WHR fuel economy and emissions benefits, extra heat sources can be considered, such as charge air and engine coolant. Generally, charge air and engine coolant are utilized to preheat the working fluid upstream the TP and EGR exhaust gas evaporators due to the relatively low available thermal power of the heat sources. However, the ORC-WHR system complexity and cost increases directly with the number of heat sources implemented.

The chosen expansion machine and expander power output type also affects the system architecture. Volumetric expanders are generally mechanically connected to the crankshaft through a gearbox, while turbine expanders are generally connected to electric generators to avoid a complex gearbox. Expander selection eventually influences system operational constraints as turbine expanders can only accept working fluid in the vapor phase.

Condenser configuration changes with the chosen heat rejection medium. If the condenser rejects heat to the engine coolant, the ORC configuration is relatively simple, but the condensation temperature is constrained by the coolant temperature, negatively affecting the cycle efficiency. Enhanced condenser heat rejection can be achieved via incorporation of a secondary cooling circuit for the condenser or by direct air cooling of the condenser. However, these two cooling methods add complexity and cost to the existing system.

The working fluid pump can be electrical or mechanical. Electrical pumps incur a higher cost and are less reliable than mechanical pumps. However, electric pumps enable simpler configurations and better working fluid mass flow control than mechanical pumps.

12.2. Heat exchangers

Heat exchangers comprise ~20% of total HDD ORC-WHR system cost [4] and significantly affect the ORC efficiency and power output. Both the working fluid and heat source pressure drops are important for heat exchanger design, especially the TP exhaust gas pressure drop because it influences engine operation. For the EGR exhaust gas evaporator, the acid dew point in the EGR exhaust gas should be avoided and carbonaceous fouling effects need to be considered.

The heat exchanger dynamics are important in the real-world HDD operating conditions [105]. The thermal inertia of the heat exchanger attenuates the transient heat source dynamics, which can be beneficial for ORC-WHR control. However, heat exchangers with high thermal inertia limit ORC operation by extending system warm-up time. The thermal inertia of the heat exchangers should consider the proposed vapor temperature controller for optimal ORC operation. Finally, the size of the heat exchanger should package within the existing vehicle architecture.

12.3. Expansion devices

The expander and its corresponding drivetrain generally account for 26–28% of total HDD ORC-WHR system cost [4]. Reciprocating piston
and radial-inflow turbine expanders are the most researched HDD ORC-WHR expansion devices. Volumetric and turbine expanders operate distinctly disparate speed regimes, < 6000 rpm and 10,000–120,000 rpm, respectively, providing different levels of synergy with mechanical coupling to the engine crankshaft. Namely, low speed volumetric expanders are more compatible with engine crankshaft coupling whereas high speed turbine expanders would require a reduction gearbox. Mechanical coupling with the engine crankshaft increases the complexity of system integration and restricts the design through packaging constraints whereas electrical power output requires energy management for efficient electricity utilization.

While volumetric expanders can operate with mixed phase working fluid, their built-in expansion ratio is generally less than turbine expanders, limiting their efficiency. However, the stringent working fluid vapor quality constraints of turbine expanders places additional burden on system control strategies. In addition, due to their high-speed characteristics, turbine bearing cooling should be carefully considered.

### 12.4. Working fluid selection

Working fluid selection is a daunting task in HDD ORC-WHR system development because there are hundreds of available working fluids and there are many criteria affecting the working fluid selection. Key factors in working fluid selection for HDD ORC-WHR systems include: (i) temperature compatibility between the working fluid and the heat sources (TP exhaust gas, EGR exhaust gas, CA and engine coolant), (ii) temperature compatibility with the condenser heat rejection medium (engine coolant, refrigerant in secondary cooling circuit, or ambient AIR), (iii) condensation pressure (1–3 bar), (iv) cycle efficiency, (v) freezing point at atmospheric pressure, (vi) safety, and (vii) cost. The most researched working fluids satisfying all those criteria for HDD include R245fa, ethanol, R123 and water-ethanol mixtures. Water is generally utilized as reference working fluid for its low cost, high safety, and high stability.

#### 12.5. Power optimization

Power optimization is the key to understanding ORC-WHR system potential and is important to the component selection process and ORC system design. HDD ORC-WHR system net power increases with increasing evaporation pressure when utilizing ethanol, R245fa or water. For these working fluids, maximum net power is generated with low levels of superheat upstream of the expander. During transient engine conditions, controllers cannot perfectly track a minimally superheated temperature reference. Thus, for ORC systems utilizing a turbine expander, the required level of superheat should be increased to maximize the net power by avoiding working fluid saturation and subsequent turbine shutdown. Advanced control strategies like MPC increase the net power production by improving the superheating temperature reference tracking performance.

Expander speed requires optimization to maximize the net power production. Mechanically coupled expander configurations require optimization of the gear ratio between the expander and crankshaft based on the expander efficiency map. Electrical turbine generators allow real-time generator speed optimization based on the expander efficiency map by removing the expander-crank shaft coupling.

Condensation pressure should be as low as possible to increase the efficiency, and net power production. Working fluid subcooling should be minimized to facilitate either an increase the working fluid mass flow rate or an increase the turbine inlet temperature. However, the effect of subcooling on the condensation pressure should also be considered. For instance, if condensation pressure is not controlled by any specific actuator, it would increase as subcooling decreases and vice versa. In this case, there is an optimal combination of subcooling temperature and condensation pressure, which maximizes the expander power output.

The working fluid pump speed governing the working fluid mass flow rate exhibits the largest impact on the net power production. Expander speed and condenser coolant pump speed produce smaller, but noticeable, impacts on the net power production. Overall, integrated optimization of the ORC system, engine, and aftertreatment system could bolster ORC power production and improve the potential of ORC-WHR systems in HDD applications.

#### 12.6. Simulation and experimental analyses

Table 6 lists the simulation (Table 6a) and experimental studies (Table 6b) of HDD ORC-WHR systems in the past decade. Experimental work is still lacking relative to the quantity of simulation studies. Experiments are predominantly conducted by companies rather than research institutes or universities, which could be caused by the high cost of HDD ORC-WHR facilities and operation. Power harvested from expanders during experiments ranges from 0 kW to 14 kW, whereas the simulations show the potential for 0–60 kW power production. This disparity in power output may stem from assumptions made in the simulation studies, such as neglected heat loss, overestimated expansion device efficiencies, ideal condensation pressure and condenser coolant temperature. The predominant experimental working fluids are ethanol and R245fa with a limited number of experimental studies utilizing water, water-ethanol mixtures, or R123. In addition, the most popular experimentally utilized heat sources are TP and EGR gases, whereas simulations often consider CA and engine coolant, creating another disparity in expected ORC-WHR power levels. In terms of fuel economy improvement, both simulation and experimental results range from 2% to around 10%. The fuel economy improvements are the combined contribution from design and selection of components, architectures, and working fluid. Even though few studies discussed the CO₂ emission reduction, fuel economy benefits are correlated with CO₂ emission via the percentage of carbon in diesel fuel. As long as fuel consumption is reduced, the CO₂ emissions are proportionally reduced.

#### 12.7. Limiting factors

Among the limiting factors in the HDD ORC-WHR system development, engine backpressure and weight increase attract most of attention. The engine backpressure issue can be addressed in both ORC side and engine side [174]. In ORC side, the heat exchanger geometry design, working fluid selection can be combined to minimize the backpressure and maximize energy recovery. In the engine side, different turbocharging strategies like fixed turbine, waste gate, and variable geometry turbine can be considered to evaluate the backpressure by ORC-WHR system. Even though the weight increase has been mentioned multiple times in different literatures, the weight impact on fuel economy in the HDD vehicle ORC-WHR application is merely analyzed. Besides the backpressure and weight increase, working fluid safety, reliability, transient condition efficiency, and cost are also important limiting factors.

#### 13. Conclusion

This paper reviews for a wide range of important factors in the HDD vehicle ORC-WHR system development. In the selection of the system architecture, a number of factors are important including heat sources, expansion device selection, expander power output type, condenser cooling methods, and working fluid pump actuation methods. The tradeoff between the power recovery and complexity must be determined during HDD ORC-WHR design. In the heat exchanger design, thermal efficiency is not the only design factor. Other parameters like exhaust gas pressure drop, thermal inertia, and sizing constraints are also important. The thermal inertia influences the ORC-WHR system control. The combination of heat source mass flow rates and temperatures determines the waste power dynamics, which the working fluid...
temperature control should consider. In the expander selection, speed range, efficiency, cooling requirements are important. In addition, whether the expander power is directly coupled to the crankshaft or converted to electricity is a critical decision impacting system complexity, cost, and efficiency. Even though a large number of working fluids are researched and discussed in simulation, less than half dozen working fluids are frequently utilized in HDD ORC-WHR experiments, including water, R245fa, ethanol, water-ethanol mixtures, R123 and R123zd-E. When the entire ORC system is considered during power optimization, evaporation pressure, superheat temperature, expander speed, condensation pressure, and condenser subcooling temperature all play critical roles. Advanced ORC control benefits power production and safety operation in transient driving cycles. There is still a large gap between predicted ORC power production in simulation and experimental studies. This disparity requires further investigation to identify avenues to improve the experimental power output based on the findings and suggestions from simulation.

References

[5] Delgado O, Lutsey N. The US SuperTruck program: Expanding the development of advanced heavy-duty vehicle efficiency technologies. The International Council on Clean Transportation, June C-70C1C7C3C7C5C7C6C7C8C7C9; 2014.


Peralez J, Nadri M, Dufour P, Tona P, Sciarretta A. Control of Automot Turbine
Gibble J. Very High Fuel Economy, Heavy Duty, Truck Engine Utilizing Biofuels and
Hybrid Vehicl Technologies, Mack Trucks, Incorporated; 2011.
Lecompte S, Ameel B, Zvivani D, van den Broek M, De Paepe M. Energy analysis of
zoogetic mixtures as working fluids in Organic Rankine Cycles. Energy Convers
Robertson MC, Costall AW, Newton PJ, Martinez-Botas RF. Radial Turboexpanded
Feng L, Zheng D, Chen J, Dai X, Shi L. Exploration and Analysis of CO2+
Shu GQ, Zhao MR, Tian H, Huo YZ, Zhu WJ. Experimental comparison of R123
Willems F, Kupper F, Cloudt R. Integrated Powertrain Control for optimal CO2
Willems F, Kupper F, Cloudt R. Integrated Powertrain Control for optimal CO2
Willems F, Kupper F, Cloudt R. Integrated Powertrain Control for optimal CO2
Willems F, Kupper F, Cloudt R. Integrated Powertrain Control for optimal CO2


