Abstract
In this study, a methodology for optimal sizing of waste heat recovery (WHR) systems is presented. It deals with dynamic engine conditions. This study focuses on Euro-VI truck applications with a mechanically coupled Organic Rankine Cycle-based WHR system. An alternating optimization architecture is developed for optimal system sizing and control of the WHR system. The sizing problem is formulated as a fuel consumption and system cost optimization problem using a newly developed, scalable WHR system model. Constraints related to safe WHR operation and system mass are included in this methodology. The components scaled in this study are the expander and the EGR and exhaust gas evaporators. The WHR system size is optimized over a hot World Harmonized Transient Cycle (WHTC), which consists of urban, rural and highway driving conditions. The optimal component sizes are found to vary for these different driving conditions. By implementing a switching model predictive control (MPC) strategy on the optimally sized WHR system, its performance is validated. The net fuel consumption is found to be reduced by 1.1% as compared to the originally sized WHR system over the total WHTC.

Keywords: scalable models, component sizing, control, heavy-duty diesel engine

1. Introduction

Heavy-duty (HD) engines are the workhorse in the transport sector. Driven by societal concerns about global warming and energy security, this sector faces enormous challenges to dramatically reduce greenhouse gas emissions and fuel consumption over the upcoming decades. In the EU, CO₂ legislation for HD vehicles is in preparation. For 2050, a 60% CO₂ reduction sectorial target is set.
To meet these challenging targets for trucks, besides vehicle and logistic measures, increase of the powertrain efficiency is an important research area. In modern diesel engines, around 25% of the fuel energy is converted into heat and is wasted with the exhaust gases into the environment. Extracting this energy and converting it into useful propulsion energy will potentially lead to significant reductions in fuel consumption.

The Organic Rankine Cycle (ORC) seems a promising waste heat recovery (WHR) technology for heavy-duty applications [1, 2]. For future implementation, further optimization of the cost-benefit ratio is crucial. More precisely, optimal sizing of the WHR system is necessary to maximize the WHR power output and fuel economy of the vehicle. However, this is challenging, since there are many factors that affect the optimality of WHR system size, including: driving conditions, system constraints, and the control strategy. In [3], it is shown that dynamic operating conditions play an important role for optimization of WHR systems, especially in truck applications. A huge gap was observed between predictions based on steady-state and dynamic conditions. Similar results are found in [4]: performance evaluation in steady-state operating points derived by driving cycle reduction tends to overestimate the fuel gain induced by the WHR system. In addition, the coupling between system and control design has to be dealt with.

In the literature, publications dedicated to topology design and architecture, control and integration with the powertrain system can be found for ORC-based WHR systems in automotive applications. The studied physics-based models [5–9] are based on stock component models, which are already available commercially. The size of the components is chosen based on the packaging requirements and cost. In summary, models with scalable components and component sizing approaches are lacking.

In this study, a new methodology is presented for optimal component sizing of WHR systems in the presence of highly dynamic engine conditions. The main goal is to minimize overall powertrain fuel consumption, while meeting safety constraints. This study is an extension of the work done in [10], where models and control techniques are developed to enable waste heat recovery for a Euro VI heavy duty diesel engine. By following an alternating optimization approach, system and control design is separated. A general optimization framework is defined that deals with the impact of component size on overall fuel consumption, system costs and system mass. A new, scalable WHR system model is proposed to support this optimization methodology. It is noted that the optimization is performed using a stand-alone WHR system, since it is seen from our research that this does not affect the optimality of the results, compared to using the complete powertrain model.

This work is organized as follows. The studied engine with WHR system and the general WHR optimization problem are introduced in Sections 2 and 3, respectively. Section 4 presents the scalable WHR system model. In the proposed alternating optimization approach, optimal component sizing and control design are split. Section 5 introduces the sizing optimization problem, which is followed to determine optimal scaling factors for evaporators and expander using the developed, scalable WHR model. For a switching MPC controller, the optimally sized WHR system performance is validated over the hot start World Harmonized Transient Cycle (WHTC) in Section 6. Finally, the main conclusions are summarized in Section 7.
2. System description

Figure 1 shows the studied system, which is based on a 13 liter, 6 cylinder Euro-VI heavy-duty diesel engine. This engine is equipped with common rail fuel injection, a high-pressure exhaust gas recirculation (EGR) system, variable turbocharger geometry (VTG) and an aftertreatment system. This aftertreatment system consists of a diesel oxidation catalyst (DOC), a diesel particulate filter (DPF) and a selective catalytic reduction (SCR) system with ammonia oxidation catalyst (AMOX).

A waste heat recovery system is installed that recovers heat from the EGR line as well as the line downstream of the aftertreatment system using an EGR and Exhaust Gas (EXH) evaporator, respectively. The working principle of this WHR system is based on an Organic Rankine Cycle (ORC). The working fluid is ethanol. It is pumped from the open reservoir, which is at ambient pressure, through the evaporators by two electrically driven pumps. In the evaporators, heat is extracted from the exhaust gases and is used to vaporize the ethanol. This vapor expands in the two-piston expander and generates mechanical power. Note that the expander is mechanically coupled to the engine crankshaft. The expander is said to operate safely if vapor state is maintained before the expander, that is, the working fluid must be in superheated state. The presence of droplets can damage the expander. After expansion, the working fluid is cooled in the condenser. The resulting liquid working fluid flows back to the reservoir, where it is stored at atmospheric pressure. For WHR system control, both pumps are used. A throttle valve $u_t$ at the expander inlet is also available to accommodate gear shifting. When the driver’s requested power is less than the net power delivered by the WHR system, $P_{req} \leq P_{whr}$.

![Figure 1. Scheme of the complete powertrain with WHR system [11].](image-url)
this valve is closed to avoid unwanted torque responses. In this study, this valve is maintained at fully opened position, since we focus on realizing maximum power output. The system pressure is limited to 60 bar by a pressure relief valve. The EGR valve $u_{g1}$ is controlled by the engine control unit (ECU), whereas the exhaust gas bypass valve $u_{g2}$ is controlled, such that the condenser cooling capacity is not exceeded.

In previous work [10], an electrified WHR system is also studied. This WHR system is equipped with a battery for energy storage, and the expander is coupled to a generator instead of the engine crankshaft. However, to demonstrate the potential of the WHR component sizing methodology, the configuration shown in Figure 1 is chosen. This configuration is more attractive for short-term application due to its relatively low system costs and complexity.

3. Optimization problem

3.1. General problem definition

The high-level objective of this study is to minimize fuel consumption of the overall powertrain by optimal sizing and control of WHR system components over a transient drive cycle, while guaranteeing safe operation. In other words, optimal component scaling factors ($\lambda_i$) in combination with optimal speed settings for both pumps ($\omega_{p1}$ and $\omega_{p2}$) have to be determined:

$$\min_{\omega_{p1}, \omega_{p2}, \lambda_i} \int_0^{t_f} m_{fuel}(t)dt$$

where $t_f$ is the duration of drive cycle. The fuel mass flow is a function of engine torque $\tau_e$, engine speed $N_e$ and EGR valve and VTG positions:

$$m_{fuel} = f(\tau_e, N_e, \text{EGR\%}, \text{VTG\%})$$

The dynamic model of the engine with WHR system is shown in Figure 2. This scheme illustrates the components and their interaction. Ambient temperature and pressure ($T_{amb}$ and $p_{amb}$), the requested engine speed $N_d$ and the torque $\tau_d$ associated with the drive cycle are the external model inputs (in green). The variables to be optimized, that is, control inputs $\omega_{p1,2}$ and scaling parameters $\lambda_i$, are indicated in blue.

To meet the torque request $\tau_d$, the required engine torque is given by:

$$\tau_e = \tau_d - \tau_{whr}$$

with the torque $\tau_{whr}$ provided by the WHR system:

$$\tau_{whr} = \frac{P_{whr}}{\omega_e} = \frac{P_{exp} - P_{p1} - P_{p2}}{\omega_e}$$
As this study focuses on maximizing the WHR system performance, the fuel consumption in Eq. (2) can be reduced by lowering the engine torque $\tau_e$. This is done by maximizing the net WHR power output $P_{\text{whr}}$, Eq. (3)–(4). The external inputs to the WHR system are the EGR and exhaust gas flows from the engine and aftertreatment system, which are also a function of $\tau_e, N_e, \text{EGR}$ and VTG positions.

In conclusion, a combined design and control optimization problem is formulated:

\textbf{Problem 1}

\[ \text{minimize } \omega_p, \omega_p, \lambda \int_0^{t_f} P_{\text{whr}}(t) \, dt \]

subject to : \[ \omega_{p,1} \leq \omega_{p,2}(t) \leq \omega_{p,2}^{\text{max}} \], \[ \lambda_i^{\text{min}} \leq \lambda_i \leq \lambda_i^{\text{max}} \}; \; \{i = 1...n\} \]

\[ T_{\text{out}}(t) \geq 120^\circ \text{C} \]

\[ T_{f1,2}(t) \leq 270^\circ \text{C} \]

\[ \chi_f(t) \geq 1 \]  \hspace{1cm} (5)

with optimization variables:

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure2.png}
\caption{Scheme of the dynamic model for the studied engine-WHR system. WHR-related components are indicated by red blocks.}
\end{figure}
Pump speeds $\omega_{p1}(t)$ and $\omega_{p2}(t)$, which control the mass flow rate of the working fluid required to extract heat energy from both the evaporators;

Design variables $\lambda_i$: these time-independent scaling factors are applied to vary the size of different components of WHR system, where $n$ is the number of components to be scaled.

This optimization problem is subject to the following constraints:

- $\omega_{min}^p$ and $\omega_{max}^p$ are the minimum and maximum pump speeds to limit the mass flow rate of working fluid in the WHR system;
- Exhaust gas temperature $T_{\text{EGR}}^{\text{out}}(t)$ at the EGR evaporator outlet should be more than $120^\circ C$ to prevent condensation;
- Ethanol temperature should always be less than $270^\circ C$ to avoid degradation;
- $\chi_f(t)$ is the vapor fraction of the working fluid, which is given by:

$$\chi_f = \frac{h_f - h_l(p_f)}{h_v(p_f) - h_l(p_f)}$$

where $h_l(p_f)$ and $h_v(p_f)$ denote the specific saturated liquid and vapor enthalpy, respectively, as a function of system pressure $p_f$. To avoid damage by droplets, this fraction should be larger than 1 at the expander inlet.

Note that maximizing the WHR system power output by optimizing component sizes can lead to an increase in the needed cooling power in the condenser. However, in this work, we assume that this cooling capacity is always available (ideal condenser).

3.2. Optimization methodologies

The problem stated above is nonconvex and highly nonlinear where both control and design parameters are optimization variables. For combined plant and control design problems, three approaches can be distinguished [12]:

- **Alternating plant and control design**: the plant is optimized first, which is then followed by an optimal control design. Subsequently, this process is repeated until the coupled variables converge;
- **Nested optimization**: the control design is nested within the plant design, that is, for each evaluation of the plant, the controller design is optimized. Often, nested optimization architectures are also called bi-level, referring to the two design layers;
- **Simultaneous optimization**: optimization of plant and controller design is done simultaneously, that is, solving Eq. (5) all-in-one.

The WHR system shown in Figure 1 is controlled by a switching model predictive control (MPC) strategy. With an alternating optimization architecture, an MPC tuned for one system
size might not be functional for a different size, due to the changing heat exchanger system dynamics. When using a nested framework, for every evaluation of plant design, multiple MPCs must be obtained covering the WHR operating area. Moreover, high tuning effort is required to implement a switching MPC strategy to obtain good disturbance rejection.

Due to the high complexity of the optimization problem, the alternating optimization method is selected in this study. The main reasons are as follows: applicability to other WHR systems topologies and possibility to sequentially run the controller and plant optimization, which reduces the instantaneous computational burden.

**Remark 1.**
The sizing optimization requires significant controller tuning effort for different plant sizes. Therefore, a size independent feed forward controller is necessary to significantly reduce the tuning effort and thus the computational complexity of the optimization problem. Even though a feed forward controller does not give the full performance, it is still representative for solving sizing problems. Moreover, in Section 6, we will show that using such a controller produces results with acceptable validation properties.

### 3.3. Feedforward pump control

The low-level pump controllers have to guarantee that the working fluid at the expander inlet is at superheated state: \( \chi_f \geq 1 \). Both pumps control the mass flow of the working fluid through the evaporators and by that the heat transfer between the exhaust gas and the working fluid. The discussed MPC strategy needs relatively high tuning effort when the WHR system has to be simulated on a grid of design points.

Considering these issues, a feed forward (FF) pump controller is introduced, which is independent of the plant size. It is based on the measured EGR and exhaust gas heat flows from the engine, measured temperature of the working fluid at the evaporator inlet, and the working fluid system pressure. For stationary conditions, the amount of heat that needs to be transferred from the exhaust gas to the working fluid is determined from the energy balance:

\[
\dot{Q}_g - \dot{Q}_{g,\text{loss}} = \dot{Q}_f
\]  

where \( \dot{Q}_g \) is the heat flow from the exhaust gas, \( \dot{Q}_{g,\text{loss}} \) are the heat transfer losses and \( \dot{Q}_f \) is the heat flow toward the working fluid. From this equation, the required working fluid mass flow can be determined, using:

\[
m_f = \frac{\dot{Q}_g - \dot{Q}_{g,\text{loss}}}{h_{f,\text{in}}^{\text{ref}} - h_{f,\text{out}}^{\text{ref}}}
\]  

where \( h_{f,\text{in}} \) is the actual enthalpy of the working fluid at the evaporator inlet and \( h_{f,\text{in}}^{\text{ref}} \) is the estimated enthalpy of the working fluid corresponding to a post-evaporator temperature of 10°C above the saturation temperature. The feed forward pump controller realizes this required working fluid flow.
4. Scalable WHR system model

In this section, the Waste Heat Recovery model from [13] is made scalable for component size. For each component, the physical parameters that have the biggest impact on the WHR power output are identified to scale the overall size of these components. The WHR system model is described using a component-based approach. The pumps and expander are map-based components. The remaining components, that is, evaporators, condenser, valves, and pressure volumes, are based on conservation of mass and energy principles.

The following assumptions are made in the model:

- Transport delays and pressure drops along the pipes are neglected;
- Change in exhaust gas density as a function of temperature and pressure is neglected;
- Pressure dynamics in the heat exchangers are not considered because of small time scales compared to temperature phenomena;
- Temperature along the transverse direction is considered to be uniform for both exhaust gas and working fluid;
- Condenser model is ideal, such that the reservoir provides the working fluid at an ambient pressure of 1 bar and temperature of 65°C. Hence, condenser sizing is not considered in this study.

4.1. Pumps

There are two identical pumps in the WHR system to pump the working fluid from the reservoir to the EGR and exhaust gas evaporators. Pumping power $P_{p1,2}$ is directly proportional to the displacement volume and rotational speed of the pump. Thus, it can be inferred that a smaller pump can rotate at higher rotational speeds to meet the demands of mass flow rates of working fluid, while maintaining the same pressure difference without necessarily affecting the power output. Therefore, any variation in their displacement volume would not affect the required working fluid mass flow rate for the same operation cycle. Hence, sizing of the pumps is not considered here.

4.2. Expander

The expansion process in the two-piston expander is illustrated in Figure 3. This cycle consists of two isobaric strokes (1→2 and 4→5), two isentropic stokes (2→3 and 5→6) and two isenthalpic mass transfers at the end of the strokes (3→4 and 6→1).

The expander power is ideally calculated by multiplying net work done in the cycle with the expander speed, given by

$$P_{\text{exp, ideal}} = \frac{W_{\text{net, ideal}} \cdot N_{\text{exp}}}{60}$$

where
\[ W_{net, ideal} = W_{12} + W_{23} + W_{45} + W_{56} \quad (10) \]

Using the ideal gas law, work delivered for different sub processes in the cycle are given by [14]:

\[ W_{12} = p \cdot (V_2 - V_{\text{head}}) \]

\[ W_{23} = \frac{p V_2^\kappa}{(1 - \kappa)} \cdot \left( \frac{1}{(V_{\text{head}} + V_d)^{\kappa-1}} - \frac{1}{V_2^{\kappa-1}} \right) \]

\[ W_{45} = p_{\text{atm}} \cdot (V_5 - V_{\text{head}} - V_d) \]

\[ W_{56} = \frac{p_{\text{atm}} V_5^\kappa}{(1 - \kappa)} \cdot \left( \frac{1}{(V_{\text{head}})^{\kappa-1}} - \frac{1}{V_5^{\kappa-1}} \right) \quad (11) \]

where \( \kappa \) is the adiabatic index, and \( p_{\text{atm}} \) is the atmospheric pressure.

The physical parameter that is affecting the power output of the expander is its volume, as indicated by Eq. (11). Thus, applying the same scaling factor to \( V_{\text{head}} \) and \( V_d \), will change the overall dimensions of the expander. By applying the scaling factor to these volumes, it is assumed that the bore-to-stroke ratio of the cylinder is not changed. Hence, new volumes are defined:

\[ V_{\text{head}}^* = \lambda_{\exp} \cdot V_{\text{head}} \]

\[ V_d^* = \lambda_{\exp} \cdot V_d \quad (12) \]

To keep the valve timings of the expander same as the original system, the same scaling factor, \( \lambda_{\exp} \), is applied to \( V_2 \) and \( V_5 \).

\[ \text{Figure 3. p-V diagram of the expansion process (} V_2 = V_{\text{IC}}: \text{intake valve closing volume;} \ V_5 = V_{\text{EC}}: \text{exhaust valve closing volume;} \ V_{\text{head}}: \text{clearance volume;} \ V_d: \text{displacement volume).} \]
\[ V_2^* = \lambda_{\text{exp}} \cdot V_2 \]
\[ V_5^* = \lambda_{\text{exp}} \cdot V_5 \]

(13)

The ideal physics-based model of the expander deviates from the measurements because the model simulates an ideal cycle, and a number of adverse effects are not taken into account, for example, drag in the outlet, formation of droplets, Van der Waals interactions and volumetric efficiency. Hence, a steady state physics-based model of the expander was estimated in [14] based on the measurement data. This data were obtained during steady-state dynamometer testing for different values of expander speed \( N_{\text{exp}} \) and system pressure \( p_f \) [10]. Results for different expander sizes are provided by the expander manufacturer. It suggests that the nominal power output increases linearly with increase in displacement volume. This is due to the modular design of the expanders. Therefore, for this study, the losses \( P_{\text{loss}} \) are considered equal for different expander sizes \( \lambda_{\text{exp}} \), so Figure 4 is used:

\[ P_{\text{whr}} = P_{\text{whr, ideal}} \cdot \left[ 1 - P_{\text{loss}} \left( N_{\text{exp}}, p_f \right) \right] \]

(14)

4.3. Evaporators

The evaporator model [15] is based on the conservation principles of mass and energy. To scale the evaporators size, the scaling factor needs to be applied on the volume of the evaporator. And the volume of evaporator can be varied by changing either length or width or height of the evaporator. The general structure of the studied evaporator is shown in Figure 5.

Figure 4. Expander power losses.
Scaling factors $\lambda_l$, $\lambda_w$ and $\lambda_h$ are applied to length $l$, width $w$ and height $h$ of the evaporators:

$$l_{\text{new}} = \lambda_l \cdot l,$$
$$w_{\text{new}} = \lambda_w \cdot w,$$
$$h_{\text{new}} = \lambda_h \cdot h$$

Consequently, the following model parameters are affected. The number of plates inside the evaporator will vary, when its height is changed,

$$n_{p,\text{new}} = \lambda_h \cdot n_p$$

Note that number of plates should be a discrete number, that is, $n_{p,\text{new}} \in \mathbb{N}$. But in this study, it is varied continuously with the scaling factors, as it does not affect the end results.

In addition, the surface area available for the working fluid to extract the heat energy from exhaust gases through the wall or plates is affected:

$$S_f = 2 \cdot n_{p,\text{new}} \cdot w_{\text{new}} \cdot l_{\text{new}}$$

This also impacts the surface area available for the exhaust gas to transmit its heat to the working fluid through wall or plates, which is given by:

$$S_g = S_f + S_{g,\text{fins}}$$

where $S_{g,\text{fins}}$ is the surface area of fins at exhaust gas side. The surface area of the wall is linear dependent on evaporator length and width:

$$S_w = l_{\text{new}} \cdot w_{\text{new}}$$

Finally, the flow cross-sectional area, $A_i$, is given by,

$$A_{i,\text{new}} = w_{\text{new}} \cdot h_{\text{new}} \quad \{i = \text{fluid, gas}\}$$

Accordingly, the Reynolds number for the working fluid and gas side is affected:

\[\text{Figure 5. Exhaust gas recirculation heat exchanger modular design.}\]
\[ Re_i = \frac{\dot{m}_i \cdot d_{hi}}{A_{i,\text{new}} \cdot \eta_i} \]  

(21)

where \( \eta_i \) is the viscosity, \( d_{hi} \) is the hydraulic diameter, that is, outer gap height of one plate and \( \dot{m}_i \) is the mass flow rate of the working fluid and gas. The heat transfer coefficient \( \alpha_i \) for the working fluid and exhaust gas also depends on \( A_{i,\text{new}} \):

\[ \alpha_i = \frac{Nu_i \cdot \Lambda_i}{dh_i} = \frac{Nu_i \cdot \dot{m}_i \Lambda_i}{Re_i \cdot A_{i,\text{new}} \cdot \eta_i} \]

(22)

where \( \Lambda_i \) is the thermal conductivity of the working fluid and exhaust gas.

For varying evaporator length, there will be no change in the flow cross sectional area as well as in the exhaust gas side cross sectional area. As a result, there will be no effect on Reynolds number, \( Re \), and the Nusselt number, \( Nu \), which is directly proportional to \( Re \). The surface areas available for exhaust gas and working fluid increase with increasing \( l \) and vice versa (see Eqs. (17)–(19)). These areas directly affect the working fluid temperature at the evaporator’s outlet.

The cross-sectional areas \( A_{i,\text{new}} \) varies with width as well as with height and is inversely proportional to the heat transfer coefficients \( \alpha_{\text{fluid}} \) and \( \alpha_{\text{gas}} \). The surface areas \( S_f \) and \( S_g \) increase with \( w \) and \( h \) and, hence, the working fluid temperature at the evaporator’s outlet will behave in the same direction through equations for conservation of energy. However, \( S_w \) will stay the same with change in height, because the number of plates will change with this dimension.

With the introduced scaling factors, the evaporator size can be changed by varying its length, width or height depending on the requirements from the system, input heat flows, and type of working fluid. These three parameters have different impact on the evaporator’s performance. Therefore, a sensitivity analysis has to be done to select the parameter that has the most positive impact on WHR system power output.

### 5. WHR system size optimization

This section presents a methodology to optimize component size for WHR systems under transient driving conditions. Figure 6 gives an outline of the approach that is followed in this study. The scalable WHR system model developed in the previous section is crucial input for this approach. From a sensitivity analysis for the exhaust evaporator, scaling of the evaporator length is identified as the most promising route to maximize WHR power output. Details can be found in [16]. As a result, evaporator width and height will be set to their original system values in the sequel of this study. In summary, the following parameters are considered for optimal component sizing in order to maximize WHR power output:

- Expander scaling \( \lambda_{\text{exp}} \);
- Exhaust gas evaporator length scaling \( \lambda_{\text{exh}} \);
- EGR evaporator length scaling \( \lambda_{\text{EGR}} \).
The component sizes are optimized based on fuel consumption and investment cost criteria. In the next section, the alternating optimization method is described in detail. The optimal sizes $\lambda_{\text{opt}}$ that are finally determined are input for the overall performance analysis with switching Model Predictive Control (MPC) strategy in Section 6.

5.1. Alternating system optimization

Optimizing the controller and plant iteratively to converge the coupled variables can be computationally too expensive. Therefore, an alternating architecture is followed for only one complete loop, see Figure 7. The controller designed for a specific WHR system will give different performance for a resized WHR system. Consequently, the feed forward controller from Section 3.3 is applied, which calculates the pumps speeds, such that $\chi_f \geq 1$ at the outlet of both the evaporators for given engine exhaust heat flows. Although with lower performance

Figure 6. Approach for optimal sizing and control of WHR systems.

Figure 7. Optimization architecture.
compared to PI control or MPC, it is seen to maintain the same trend for fuel consumption with different components sizes, without affecting the optimality of WHR system components sizes.

The standalone WHR system with feed forward controller is simulated on a 3D design grid of different sizes of EGR evaporator, exhaust evaporator and expander for a hot-start WHTC. WHR system performance is strongly affected by operating conditions. Therefore, besides overall (complete cycle) performance, also optimization is performed on urban, rural and highway driving parts, which are illustrated in Figure 8. The design grid is chosen such that: (1) it captures the main trend in outputs due to component sizing and (2) costs and total system mass remain acceptable. For the scaling factors, the following grid is chosen: $\lambda_{\text{EGR}} = [0.4 : 0.1 : 1.5]$ and $\lambda_{\text{exp}} = [0.4 : 0.1 : 2.5]$.

5.1.1. Objective functions

Using exhaustive search, also referred to as brute force search, the WHR system performance is determined for each point on the grid. To obtain the optimal size of the plant, the following objective functions are defined:

- **Fuel consumption** (FC), which is determined from Eqs. (2)–(3). For the specified design space, WHTC results are summarized in Figure 9. Note that the net fuel consumption is normalized by using engine-only (without WHR system) results. Minimal fuel consumption is found at maximum evaporator size, although the reduction in fuel consumption decreases with increasing size. For highway conditions, where engine exhaust heat flows are relatively high, fuel economy increases for increasing expander sizes. The exhaust evaporator’s size has a significant impact on fuel consumption, especially in the urban

![Figure 8](image-url) Engine torque, engine speed and heat flows for hot start WHTC.
region. Due to low exhaust gas heat flows in this region, the time for the working fluid to extract heat increases with increase in length (or surface area) of evaporator.

- **Specific investment cost** (SIC, in €/kJ): in this study, we focus on installation (Cost\textsubscript{labor}) and material and production cost of the components (Cost\textsubscript{comp}) corresponding a specific WHR energy output:

\[
SIC = \frac{\text{Cost}_{\text{labor}} + \text{Cost}_{\text{comp}}}{\int_0^{t_v} P_{\text{whr}}(t) dt}
\]  

where \(P_{\text{whr}}(t)\) is the instantaneous net WHR power output and \(t_v\) is cumulative time in vapor. For the evaporators and expander, cost correlations are taken from [17]:

\[
\begin{align*}
\text{Cost}_{\text{evap}} &= 190 + 310 \cdot A_{\text{evap}} \cdot \lambda_{\text{evap}} \\
\text{Cost}_{\text{exp}} &= 1.5 \cdot (225 + 170 \cdot V_{\text{exp}} \cdot \lambda_{\text{exp}})
\end{align*}
\]  

These equations clearly show that the component costs are proportional to the scaling factors to be applied. The cost of other components, such as pumps, piping, condenser and valves, are not included in the SIC. Their sizes are assumed to be fixed in this study, which will not affect the objective function. For SIC, similar graphs are generated as for FC. From these results, it is concluded that \(\lambda_{\text{EGR}}\) has negligible effect on SIC, whereas \(\lambda_{\text{exp}}\) has the biggest impact, because of its dominant share in the total system cost. Details can be found in [16].

\(\text{Figure 9.}\) Normalized fuel consumption for different sizes of WHR system components on different driving conditions from hot-start WHTC.
5.1.2. Sizing optimization problem

Having defined the objective functions for FC and SIC, the optimization problem boils down to:

\[
\min_{\lambda_{\text{exh}}, \lambda_{\text{EGR}}, \lambda_{\text{exp}}} J
\]

subject to,

\[
m_{\text{eq}}(\lambda_{\text{EGR}}) + m_{\text{exh}}(\lambda_{\text{exh}}) + m_{\text{exp}}(\lambda_{\text{exp}}) \leq M_{\text{tot}}^{\text{max}}
\]

where the multi-objective function \( J \) is given by:

\[
J = \begin{cases} 
\int_{0}^{t_f} \dot{m}_\text{fuel} dt \\
\text{SIC} 
\end{cases}
\]

Note that mass of the WHR system is directly proportional to the components sizes and directly affects the overall load carrying capacity of a truck. Therefore, a limit, \( M_{\text{tot}}^{\text{max}} \), is defined on the component mass associated with sizing.

5.1.3. Sizing optimization for different system mass

As a next step to solve Eqs. (25)–(27), optimal \( \lambda \) setting is determined for different mass constraints \( M_{\text{tot}}^{\text{max}} \). This is done using the lambda sweep plots for FC as well as SIC, similar to Figure 9. For each objective function, the best \( \lambda_{\text{EGR}}, \lambda_{\text{exh}} \) and \( \lambda_{\text{exp}} \) combination is determined, which gives lowest FC or SIC while meeting a varying \( M_{\text{tot}}^{\text{max}} \). Figure 10 shows an example of results for using the FC objective function. The resulting scaling parameters are given for four different operating conditions associated with the hot WHTC. For the overall WHTC, the corresponding fuel consumption is shown in the lower plot. Similar plots are made for SIC.

5.1.4. Best WHR sizing per objective function

Final step in the optimization approach is to select \( M_{\text{tot}}^{\text{max}} \). For the purpose of benchmarking, the optimal component sizes associates with the two different optimization criteria are compared for a mass constraint equal to the original mass of the system: \( M_{\text{tot}}^{\text{max}} = 210 \) kg, which is indicated in the plots of Figure 10 by the blue vertical lines. The results of this final step are summarized in Figure 11.

The results clearly indicate that different operating conditions and different optimization criteria lead to different component sizing. However, optimal exhaust evaporator size is smaller (i.e., \( \lambda_{\text{exh}} < 1 \)) than its original size for all the driving conditions and both FC and SIC optimization. The exhaust evaporator has the biggest mass of the three scaled components. Although longer exhaust evaporator have better performance, the increment reduces with
increasing $\lambda_{\text{exh}}$. Hence, $M_{\text{max}}^{\text{tot}}$ plays a bigger role in sizing because increasing the size of other components is much more beneficial in terms of energy recovery.

Apart from highway driving conditions, the optimal scaling of EGR evaporator is found to be bigger than the original one. During highway driving, the amount of heat that needs to be extracted from the exhaust gases is high, which leads to higher ethanol flows. The results

![Figure 10. Optimal component sizes to realize minimal fuel consumption for different mass constraints.](image)

![Figure 11. Optimal components sizes for $M_{\text{max}}^{\text{tot}} = 210\text{kg}$ and for different objective functions.](image)
indicate that the original evaporator is over dimensioned for this condition. For urban and rural conditions, where exhaust heat flows are low, the ethanol flow needs to be low to extract the maximum amount of heat. However, when mass flows reach the lower boundary condition and no vapor is generated, increased evaporator length would provide more surface area and hence more time for the working fluid to extract heat. This effect is confirmed for the EGR evaporator with bigger optimal sizes for urban and rural regions. As these regions play an important role in the overall cycle result, it is expected that similar $\lambda_{EGR}$ is found for the overall cycle.

The optimal expander scaling is found to be bigger than the original one for all driving conditions. This especially holds for FC optimization. For highway conditions, $\lambda_{exp}$ is twice the original one due, in order to exploit the high heat flows. The expander has the biggest impact on WHR system performance. With mass comparable to other components, this leaves room for increasing $\lambda_{exp}$.

Finally, these results also indicate that optimal sizing of a single component strongly depends on the performance of all components. Interaction between evaporators and expander as well as the total mass constraint play an important role.

5.2. Selected optimal scaling of WHR components

In order to make a final decision on optimal component scaling, a trade off has to be made between the optimization criteria. Therefore, the impact on FC and SIC are analyzed for both criteria. Focus is on the overall cycle result, since this is assumed to be representative for real-world performance. As expected, fuel consumption can be reduced by 0.85% compared to the originally sized system (and 2.78% compared to engine-only mode) in case of FC minimization. In case of SIC minimization, compared to the original system, there is no FC reduction, but system costs are reduced by 25%. The SIC of the FC optimal system is 60 €/kJ higher than that for the SIC minimal case.

Comparison of both cases learns that the additional system costs associated with the FC minimum case, requires an additional 1 month truck operation for return on investment. Therefore, the final optimal components scaling for the WHR system are based on the values for FC minimization:

$$\lambda_{exh, opt} = 0.63 \quad \lambda_{EGR, opt} = 1.47 \quad \lambda_{exp, opt} = 1.50$$

These values will be used in the sequel of this study.

6. Simulation results

In this section, the switching MPC strategy from [18] is applied to the standalone WHR system with optimal component sizes. A comparison is made with the original WHR system. The MPC strategy is first evaluated on a simple stepwise cycle data from a real Euro VI heavy-duty diesel engine and then on a hot-start World Harmonized Transient Cycle (WHTC) [16].
Considering the real time system dynamics of WHR system, the controller sampling time is chosen to be $T_s = 0.4$ seconds. The prediction and control horizon are $N_y = 50$ and $N_u = 4$ time steps, respectively, for all the three linear models.

### 6.1. Controller performance validation

The proposed switching MPC strategy is validated on the highly dynamic, hot start WHTC. Disturbances from the engine, that is, engine speed and EGR and exhaust gas heat flows, are inputs for the simulation. The objective of the controller is to maintain vapor state. However, due to the highly dynamic disturbances and limitations of the control input, it is challenging to achieve this target for all the time. To avoid damage, the expander is bypassed using the bypass valve, such that net power output: $P_{whr} = 0$.

Figure 12 shows the vapor fraction after EGR and EXH evaporators, and the mixing junction for the original and optimized system. Vapor fraction is not meeting the reference (indicated by dashed line) between 200 and 400 s, due to the low heat flows in the urban region. However, the controller shows improved overall performance in terms of disturbance rejection, where the controller specifications, $\chi_{f,\text{mix}} \geq 1$, are met with short periods of time reaching at 0.9 (around 800 and 1200 s).

![Figure 12. Comparison of vapor fraction after the EGR and EXH evaporators, mixing junction for MPC strategy between original and optimally sized system (with vapor fraction according to Eq. (6)).](image-url)
The main objective of MPC tuning is to keep the vapor fraction of working fluid after both the evaporators’ outlets close to reference data with good disturbance rejection properties. Due to different system dynamics, the values of the weighting matrices $W_{\Delta u}$ and $W_{y}$ vary from the original to the optimally sized system. Hence, the performance of the two controllers, that is, MPC for original system and optimally sized system, is quantified in terms of net thermal energy recovered and total time in vapor state, $t_v$ for different parts of the WHTC. Figure 13 illustrates that the optimally sized system outperforms the original system in terms of recovered thermal energy for all the driving conditions. Note that the recovered energy is almost doubled when complete cycle is considered. This is due to the increased expander size leading to more power output. In terms of time in vapor, both systems behave similarly, with slightly increased $t_v$ for the optimal system over the full WHTC.

6.2 Powertrain performance validation

The net fuel consumption results for the studied cases are compared with the engine only mode in Figure 14. The original sized WHR system gives a 1.94% reduction in fuel consumption using the feed forward controller ($\lambda_i = 1(FF)$). An additional 0.8% reduction is found in

Figure 13. Performance indices in terms of recovered thermal energy (left) and time in vapor (right) compared with original system.

Figure 14. Fuel consumption for different driving conditions from hot-start WHTC.
case a switching MPC strategy \((\lambda_i = 1(MPC))\) is applied. The optimally sized WHR system with feed forward control strategy \((\lambda_i = \text{optimal(FF)})\) reduces fuel consumption by 2.78%. Using a switching MPC strategy \((\lambda_i = \text{optimal(MPC)})\) gives a fuel consumption reduction of 3.82% as compared to the engine only mode. In summary, by optimizing the size of WHR system components, an additional 1.08% reduction in fuel consumption can be achieved compared to the original WHR system using the methodology given in this study.

7. Conclusions

A methodology for optimal components sizing is presented for waste heat recovery systems operated during dynamic engine conditions. Optimality was defined in terms of minimizing the fuel consumption of the overall powertrain system. The main challenge in developing this methodology is the coupling between system design and control parameters. Focus is on Euro-VI heavy-duty engines with a mechanically coupled WHR system. Based on this work, the following conclusions are drawn:

- An existing WHR system model [13] is extended with a detailed expander model and is made scalable for component size. Expander volume as well as evaporator length, width and height can be varied;
- Sensitivity analysis shows that length is the most promising route to optimize power output for evaporators;
- An alternating optimization architecture is presented, which uses the standalone, scalable WHR system model. This methodology combines an constrained optimization problem based on fuel consumption, system costs and system mass considerations with a feed forward pump controller in order to isolate system design from control design;
- This methodology is successfully followed for optimal design and control of WHR system for transient driving conditions while satisfying safe operation. The components scaled in this study are EGR and exhaust gas evaporator, and expander. Different optimal component sizes are found for city, urban, rural and overall hot-start WHTC driving conditions;
- By implementing a switching model predictive control (MPC) strategy on the optimally sized WHR system, time in vapor state is identical, while the net fuel consumption, as compared to the originally sized WHR system, is reduced by:
  - Overall hot-start WHTC: 1.08%.
  - Urban: 0.30%.
  - Rural: 1.46%.
  - Highway: 1.61%.
Author details

Emanuel Feru\textsuperscript{1,2}, Srajan Goyal\textsuperscript{2,3} and Frank Willems\textsuperscript{1,2*}

*Address all correspondence to: frank.willems@tno.nl

1 TNO Automotive, Helmond, The Netherlands
2 Eindhoven University of Technology, Eindhoven, The Netherlands
3 FlandersMake, Lommel, Belgium

References

[1] Park T, Teng H, Hunter GL, van der Velde B, Klaver J. A rankine cycle system for recovering waste heat from hd diesel engines-experimental results, Technical report. SAE Technical Paper. 2011-01-1337. 2011

[2] Bredel DE, Nickl IJ, Bartosch D-IS. Waste heat recovery in drive systems of today and tomorrow. MTZ Worldwide eMagazine. 2011;72(4):52-56

[3] Grelet V. Rankine cycle based waste heat recovery system applied to heavy duty vehicles: Topological optimization and model based control. Ph.D. thesis, Université de Liège, Liège, Belgique; 2016

[4] Goyal S. Optimal sizing of waste heat recovery system for HD truck: Steady state analysis, Technical report. Eindhoven University of Technology; 2016

[5] Horst TA, Rottengruber H-S, Seifert M, Ringler J. Dynamic heat exchanger model for performance prediction and control system design of automotive waste heat recovery systems. Applied Energy. 2013;105:293-303

[6] Horst TA, Tegethoff W, Eilts P, Koehler J. Prediction of dynamic rankine cycle waste heat recovery performance and fuel saving potential in passenger car applications considering interactions with vehicles energy management. Energy Conversion and Management. 2014;78:438-451

[7] Lecompte S, Huisseune H, van den Broek M, De Schampaheleire S, De Paepe M. Part load based thermo-economic optimization of the organic rankine cycle (orc) applied to a combined heat and power (chp) system. Applied Energy. 2013;111:871-881

[8] Seher D, Lengenfelder T, Gerhardt J, Eisenmenger N, Hackner M, Krinn I. Waste heat recovery for commercial vehicles with a rankine process. In: 21st Aachen Colloquium on Automobile and Engine Technology, Aachen, Germany. Oct, 2012. pp. 7-9

[9] Quoilin S, Aumann R, Grill A, Schuster A, Lemort V, Splethoff H. Dynamic modeling and optimal control strategy of waste heat recovery organic rankine cycles. Applied Energy. 2011;88(6):2183-2190
[10] Feru EE. Auto-calibration for efficient diesel engines with a waste heat recovery system. Ph.D. thesis, Eindhoven University of Technology; 2015

[11] Feru E, Willems F, de Jager B, Steinbuch M. Modeling and control of a parallel waste heat recovery system for Euro-VI heavy-duty diesel engines. Energies. 2014;7(10):6571-6592

[12] Silvas E, Hofman T, Murgovski N, Etman LFP., Steinbuch M. Review of optimization strategies for system-level design in hybrid electric vehicles. In: IEEE Transactions on Vehicular Technology. Jan 2017;66(1):57-70. DOI: 10.1109/TVT.2016.2547897

[13] Feru E, Kupper F, Rojer C, Seykens X, Scappin F, Willems F, Smits J, De Jager B, Steinbuch M. Experimental validation of a dynamic waste heat recovery system model for control purposes, Technical report. SAE Technical Paper; 2013-01-1647. 2013

[14] Oom M. Identification and validation of an expander model for automotive waste heat recovery systems, Technical report. Eindhoven University of Technology; 2014

[15] Feru E, de Jager B, Willems F, Steinbuch M. Two-phase plate-fin heat exchanger modeling for waste heat recovery systems in diesel engines. Applied Energy. 2014;133:183-196

[16] Goyal S. Optimal sizing of waste heat recovery system for dynamic engine conditions, Master thesis CST 2016.137. Eindhoven University of Technology; 2016

[17] Quoilin S, Declaye S, Tchanche BF, Lemort V. Thermo-economic optimization of waste heat recovery organic rankine cycles. Applied Thermal Engineering. 2011;31(14):2885-2893

[18] Feru E, Willems F, de Jager B, Steinbuch M. Model predictive control of a waste heat recovery system for automotive diesel engines. In: System Theory, Control and Computing (ICSTCC), 2014 18th International Conference. IEEE; 2014. pp. 658-663
