Dávid Hriadel, Bc.

Health Status Monitoring of Turbocharger for Passenger Vehicle Applications

Master's Thesis

Master's degree programme: Cybernetics and Robotics

Czech Technical University in Prague

Supervisor

Ing. Martin Gurtner

Department of Control Engineering

Prague, January 2019

MASTER'S THESIS ASSIGNMENT

I. Personal and study details

Student's name: **Hriadel Dávid** Personal ID number: **420200**

Faculty / Institute: **Faculty of Electrical Engineering**

Department / Institute: **Department of Control Engineering**

Study program: **Cybernetics and Robotics**

Branch of study: **Cybernetics and Robotics**

II. Master's thesis details

Master's thesis title in English:

Health Status Monitoring of Turbocharger for Passenger Vehicle Applications

Master's thesis title in Czech:

Monitorování stavu turbodmychadla v osobních vozech

Guidelines:

The goal of the thesis is to design a set of Health Indicators (HI) capable of monitoring selected turbocharger failure modes and to develop an algorithm estimating the current turbocharger health status based on the proposed HI. The solution should be deployable to production Electronic Control Unit (ECU) and it should use only commonly measured signals in passenger vehicles.
1. Study the modeling approaches for Prognostics and Health Management (PHM) used in the automotive domain; focus on air-path of Internal Combustion Engines (ICE).
2. Within a chosen ICE air-path model suitable for simulation, model and implement individual turbocharger failure modes.
3. Develop model-based and data-driven algorithms for turbocharger continuous degradation tracking, i.e. health status monitoring.
4. Validate the algorithms and discuss the benefits of the results with the stress on their applicability within the concept of predictive maintenance.

Bibliography / sources:

[1] Kumar, S., & Pecht, M. (2010). Modeling approaches for prognostics and health management of electronics. International Journal of Performability Engineering, 6(5), 467-476.
[2] Guzzella, L., & Onder, C. (2009). Introduction to modeling and control of internal combustion engine systems. Springer Science & Business Media.
[3] Adamkiewicz, A. (2012). An analysis of cause and effect relations in diagnostic relations of marine Diesel engine turbochargers. Zeszyty Naukowe/Akademia Morska w Szczecinie, 5-13.
[4] Prytz, R., Nowaczyk, S., Rögnvaldsson, T., & Byttner, S. (2015). Predicting the need for vehicle compressor repairs using maintenance records and logged vehicle data. Engineering applications of artificial intelligence, 41, 139-150.

Name and workplace of master's thesis supervisor:

Ing. Martin Gurtner, Department of Control Engineering, FEE

Name and workplace of second master's thesis supervisor or consultant:

Date of master's thesis assignment: **11.09.2018** Deadline for master's thesis submission: **08.01.2019**

Assignment valid until: **30.09.2019**

Ing. Martin Gurtner	prof. Ing. Michael Šebek, DrSc.	prof. Ing. Pavel Ripka, CSc.
Supervisor's signature	Head of department's signature	Dean's signature

III. Assignment receipt

The student acknowledges that the master's thesis is an individual work. The student must produce his thesis without the assistance of others, with the exception of provided consultations. Within the master's thesis, the author must state the names of consultants and include a list of references.

.	
Date of assignment receipt	Student's signature

Declaration

I declare that I worked out the presented thesis independently and I cited all references according to Methodical instruction n. 1/2009 about ethical principles for academic thesis writing.

Date	Signature

Acknowledgement

I would like to thank my thesis supervisor *Martin Gurtner* for his thoughtful guidance during the development of this thesis. He gave me valuable feedback always when I needed.

I would like to thank *Jaroslav* for the opportunity to write this thesis, creative consultations we had, and the data I was provided. Also, I would like to thank all of my *Garrett* colleagues and friends for the productive discussions about the topic of combustion engines.

At last, my gratitude belongs to *Simonka* and my family for their patience and support.

Abstract

This work focuses on the design and implementation of the health monitoring system of a turbocharger (primarily for passenger vehicles) with the emphasis on possible deployment of the proposed algorithms in production vehicles. The solution found in the thesis is ready to be embedded in the researched concept of Prognostics and Health Management (PHM). Possible turbocharger faults with their cause/effect relations and frequency of occurrence are investigated. Compressor efficiency drop (for instance, due to oil fouling) is identified as the only fault worth of further investigation. An approximate model of compressor efficiency drop due to oil fouling is proposed by regression on the real, clean and degraded compressor map data. Based on simulations of a diesel engine with this approximate model, three symptoms are found to arise when the compressor efficiency drop is present: increased command to open vanes in Variable Geometry Turbocharger (VGT), leading to the increased shaft rotational velocity, and increased compressor outlet temperature. For each of the symptoms, an indicator is designed, either by using a model, or data mining. At last, the indicators are employed in the engine simulation environment and their detection together with prognostics performance, involving prediction of Remaining Useful Life (RUL), is evaluated by simulating a series of emission driving cycles.

Keywords

turbocharger, diagnostics, prognostics, regression analysis, model-based design

Abstrakt

Táto práca sa zameriava návrhom a implementáciou systému na monitorovanie stavu turbodúchadla (predovšetkým pre osobné motorové vozidlá) s dôrazom na možné nasadenie navrhnutých algoritmov do vozidiel v produkcii. Riešenie nájdené v tejto práci je pripravené na začlenenie do konceptu Prognostiky. Potenciálne poruchy turbodúchadla so vzťahom príčina-porucha a ich frekvencia výskytu je preskúmaná v literatúre. Pokles efektivity kompresoru (napríklad z dôvodu znečistenia olejovými časticami) je porucha zvolená na analýzu v ďalších častiach diplomovej práce. Približný model poklesu efektivity kompresoru kvôli zaneseniu olejovými časticami je navrhnutý regresnou analýzou na reálnych, čistých a zanesených kompresorových máp. Na základe simulácií dieselového motora s týmto odhadnutým modelom sa zistilo, že tri príznaky vznikajú pri prítomnosti poklese efektivity kompresoru: zvýšený akčný zásah na otvorenie prívodných lopatiek v turbodúchadle s variabilnou geometriou VGT, čo vedie k zvýšeniu rýchlosti otáčania hriadeľa a k rastu teploty vzduchu na výstupe z kompresoru. Pre každý z týchto príznakov je navrhnutý indikátor buď pomocou modelu alebo hĺbkovej analýzy dát. Nakoniec sa indikátory implementujú v simulačnom prostredí motora a ich detekcia spolu s prognostikou, zahrňujúc predikciu zostávajúceho užitočného života RUL, je ohodnotená simuláciou série emisných cyklov.

Kľúčové slová

turbodmychadlo, diagnostika, prognostika, regresní analýza, návrh založený na modelu

List of Symbols

Internal Combustion Engine (ICE)
Charge Air Cooler (CAC)
Exhaust Gas Recirculation (EGR)
Mass Air Flow Sensor (MAF)
Manifold Absolute Pressure Sensor (MAP)
Electronic Control Unit (ECU)
Variable-Geometry Turbocharger (VGT)
Wastegate (WG)
Electronic Throttle (ET)
Center Housing Rotating Assembly (CHRA)
Mean-Value Models (MVM)
Variable Inlet Guide Vane (VIGV)
On-Board Diagnosis (OBD)
Fault Detection and Isolation (FDI)
Diagnostic Trouble Code (DTC)
Prognostics and Health Management (PHM)
Remaining Useful Life (RUL)
Crankcase Ventilation (CCV)
Foreign Object Damage (FOD)
Rotary Electronic Actuator (REA)
Low Cycle Fatigue (LCF)
High Cycle Fatigue (HCF)
Thermo-Mechanical Fatigue (TMF)
Cumulative Sum (CUSUM)
Exponentially Weighted Moving Average (EWMA)
Average Run Length (ARL)
Upper Control Limit (UCL)
Lower Control Limit (LCL)
World Harmonized Transient Cycle (WHTC)

Federal Test Procedure Transient Cycle (FTP)
End of Life (EOL)
Cumulative Distribution Function (CDF)
Probability Distribution Function (PDF)

ambient pressure, ambient temperature $p_{\text{amb}}, T_{\text{amb}}$
compressor inlet (air filter outlet) pressure, temperature p_{1c}, T_{1c}
compressor outlet (boost) pressure, temperature p_{2c}, T_{2c}
compressor pressure ratio $\Pi_c = \frac{p_{2c}}{p_{1c}}$
compressor, turbine isentropic efficiency η_c, η_t
air mass flow W_{HFM}, m_c
corrected air mass flow $\widetilde{m}_c$
inlet manifold pressure p_{inlet}
charge air cooler outlet temperature T_{CAC}
turbocharger shaft rotational velocity N_t
turbocharger shaft corrected rotational velocity $\widetilde{N}_t$
turbocharger shaft rotational inertia J_t
turbocharger map reference pressure, temperature $p_{\text{ref}}, T_{\text{ref}}$
turbine, compressor, friction, auxiliary source torque T_t, T_c, T_f, T_a
turbine, compressor power P_t, P_c
turbine inlet (exhaust manifold) pressure, temperature p_{1t}, T_{1t}
turbine outlet (after-treatment inlet) pressure, temperature p_{2t}, T_{2t}
exhaust mass flow m_t
turbine pressure ratio $\Pi_t = \frac{p_{1t}}{p_{2t}}$
VGT command input u_{VGT}
engine rotational velocity N_e
injected fuel quantity q_{inj}
compressor pressure ratio model $\Pi_c(\widetilde{N}_t, \widetilde{m}_c)$
compressor efficiency model $\eta_c(\widetilde{N}_t, \widetilde{m}_c)$
turbocharger (corrected) shaft rotational velocity model $\widetilde{N}_t(\Pi_c, \widetilde{m}_c)$
indicator, compressor efficiency approximated formula $\eta_c(\Pi_c, T_{2c}, T_{1c}), \eta_c$
indicator, turbocharger shaft velocity residual r_{N_t}
indicator, state of health regression model SOH
RUL probability distribution sampled from k-closest members RUL_k

Contents

Introduction

The automotive industry plays a key role in many economies around the world. Only in the Czech Republic, the automotive industry employs more than 150,000 people and accounts for more than 20% of both Czech manufacturing output and Czech exports[1]. With passenger car production at 128 vehicles per 1,000 persons, the Czech Republic has maintained its supreme position among the world, following only the country of Slovakia. However, the competition in the automotive industry and the legislative agencies constantly pushes the automotive to produce safer, cleaner, and more comfortable vehicles. But, at the same time, the car manufacturers pursue to offer a differentiated technology to maximize their profits.

One option for the vehicle operators to minimize the operational costs is to optimize the *maintenance* process. Since the beginning of this millennium, the concept of *predictive* maintenance has been replacing the condition-based and preventive maintenance. In some industries, such as the manufacturing plants, or the aerospace, predictive maintenance has proven to offer a more effective and cost optimizing maintenance process. However, there have been only a few reports of incorporation of the predictive maintenance in the automotive vehicle industry. A remarkable effort has been made by General Motors. In their report[2], they discuss the automotive-related prognostics and explain, for example, the need for advanced diagnosis in the more and more electrified vehicles.

In this work, we focus on the design and implementation of the health monitoring system of a turbocharger. The turbocharger is an Internal Combustion Engine's (ICE) component found in almost all passenger vehicles produced nowadays. A current diagnostic solution is implemented as part

[1] CzechInvest, 2018.
[2] Holland et al., 2010.

of On-Board Diagnosis (OBD) to detect faults leading to increased emissions through the observation of overboost and underboost conditions. Researched literature introduces model-based detection methods while relying on unmeasured variables and with insufficient failure mode analysis, whereas the solution presented in the thesis works with the latest knowledge of Prognostics and Health Management (PHM) concept, equivalent to the predictive maintenance. One of the thesis' goals is to demonstrate the readiness of using PHM for a turbocharger in current passenger vehicle configuration *without* physical intervention in the vehicle. Hence, the effort is made to prepare the software solution ready for embedding into the real, stringent ICE environment.

Outline of the Thesis

In the first chapter, we explain the role of the turbocharger in the ICE, list the related sensors available for monitoring, describe the air pressure boost feedback loop of the turbocharger and the Electronic Control Unit (ECU), pinpoint turbocharger's subcomponents, and define its performance characteristics.

In the second chapter, we investigate the general prognostics models as they are reported in the literature. Then, state of the art diagnostics of the turbocharger is introduced, where the diagnostic methods, used in the literature, rely on the model-based approach.

The third chapter offers extensive analysis of the turbocharger failure cause and effect relations. This is a crucial step in the prognostics concept. Based on the fault occurrence, its severity, and possibility to be observed, we grouped them into observable and unobservable faults; observable faults include actuator wear, compressor or turbine efficiency drop, and Low Cycle Fatigue (LCF) phenomenon, unobservable faults include shaft bearing wear or Foreign Object Damage (FOD). We conclude the third chapter with a recapitulation of most common faults, and if it is possible, suggest a method to observe the degradation. The rest of this thesis deals with a compressor efficiency drop fault modeling and monitoring. This fault has a direct and predictable impact on the performance of the engine as more power has

to be delivered by the turbine to the compressor to deliver same pressure boost.

In the fourth chapter, we perform a regression analysis on the real, healthy and faulty compressor map data to find an analytical expression of the compressor oil fouling, one of the causes of the compressor efficiency drop. Thanks to the efficiency drop prediction model we produce a degraded compressor map and verify, in the simulation, that three symptoms appear: increased command to open vanes in Variable Geometry Turbocharger (VGT), leading to the increased shaft rotational velocity, and increased compressor outlet temperature. We develop three health indicators, each covering one symptom mentioned before, by using model-based and data-driven methods.

In the last, fifth chapter, we discuss the on-board architecture of the health monitor and its actual implementation. Steps like normalization and filtering are required before any diagnostic decision is made. Afterward, health monitor's performance is evaluated by running a series of legally issued emission driving cycles World Harmonized Transient Cycle (WHTC), Federal Test Procedure Transient Cycle (FTP). The simulation is performed in the Simulink^{TM} environment, where the turbocharged diesel engine model plus its controller, identified on the real data, are available. Health monitor's diagnostic and prognostic capabilities are successfully validated. More precisely, the prediction of Remaining Useful Life (RUL) in the prognostics approach is made by searching for the vehicles with similar indicator history in the dataset of apriori simulated vehicles and processing their future behavior into the RUL probability distribution.

1 Turbocharger Description

In this chapter, we introduce a brief overview of a turbocharger's functionality. In the beginning, we describe the interaction of the turbocharger with other air-path loop components. Afterward, it is viewed as a standalone unit. We offer a more detailed description when we discuss individual failure modes.

A turbocharger is an engine's air-path component capable of producing additional pressure boost hence increases air-mass density swept to cylinders during an intake stroke. Engine's volumetric efficiency is enhanced[1], and more power can be supplied to a clutch from a flywheel while keeping the equivalent engine's displacement volume. This apparent advantage, however, imposes additional complexity to the whole engine system. And more complex systems require more advanced diagnostic methods to detect, isolate, and track the unpermitted changes in the system.

1.1 Turbocharger in the Air-Path Loop

A typical passenger vehicle turbo-charged diesel engine air-path is depicted in Figure 1.1. This case-study model has been used in extensive work on model-based diagnosis by Nyberg and Stutte (2004).

Ambient air is drawn into the engine and filtered through the air filter (not shown in the figure). Kinetic energy acquired from the compressor wheel rotational velocity is converted to static pressure in a diffuser, called *boost pressure*. Afterward, air is cooled by the Charge Air Cooler (CAC) to decrease its volumetric mass density and is passed to the intake manifold, where

[1] Opposed to naturally-aspirated engines.

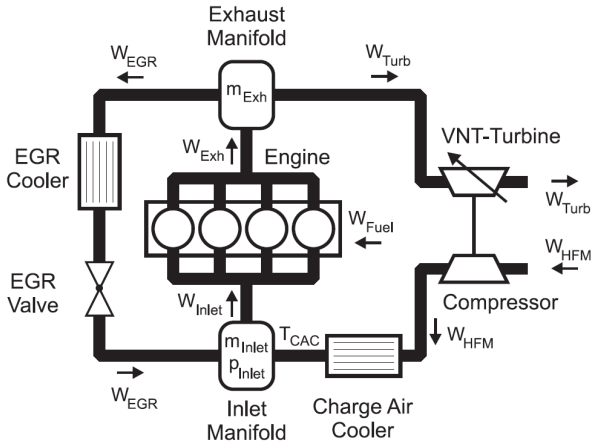


Figure 1.1: The Mercedes-Benz OM611 diesel engine layout. Figure adopted from Nyberg and Stutte (2004).

it is mixed with exhausts recycled via the Exhaust Gas Recirculation (EGR) path. Rest of the exhausts drives the turbine and leaves the engine system by passing through a series of after-treatment devices to reduce pollutant emissions (not shown in the figure).

In the scope of the thesis, we consider production engine is equipped with following (relevant) sensors:

1. Mass Air Flow Sensor (MAF). This sensor outputs raw mass flow rate of air entering the intake manifold. Recent applications show a trend where mass flow rate is estimated by means of *virtual sensor*[2]. One way or another, it is assumed the fresh air mass flow W_{HFM} is known.
2. Manifold Absolute Pressure Sensor (MAP). A pressure p_{inlet} measured at Inlet Manifold where mixed plenum is ready to be drawn into the cylinders.
3. Charge Air Temperature Sensor. This sensor measures CAC outlet temperature T_{CAC}, before entering the manifold.
4. Ambient Pressure Sensor. This sensor measures atmospheric pressure p_{amb}.

[2] Tabaček, 2016.

5. Ambient Temperature Sensor. This sensor measures air filter inlet temperature T_{amb}.

To complement, quantities generally *not* measured include:

1. Rotational velocity N_t of the shaft connecting compressor and turbine wheels.
2. Compressor discharge (CAC inlet) temperature T_{2c}.
3. Exhaust Manifold temperature T_{1t} and pressure p_{1t}.

Main control-loop involving turbocharger is a boost pressure p_{2c} control. Electronic Control Unit (ECU) demands p_{2c} set-point based on multiple inputs (for example, engine speed N_e, or driver's torque demand) and handles it to the underlying feedback control loop. Target boost pressure is achieved by manipulating an *actuator*[3] influencing the amount of exhaust gas energy transferred into the radial-flow turbine. However, another control loop which may involve slight manipulation with the turbocharger's actuator—within the air-path—is EGR mass flow control. Since the EGR control loop is introduced because of NO_x emission reduction, it is not considered relevant to the turbocharger analysis in the thesis.

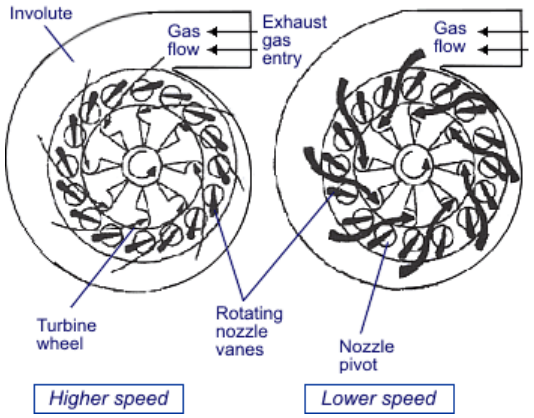
Two basic means of controlling the exhaust gas flow are recognized:

Variable-Geometry Turbocharger (VGT) A group of nozzle pivots changes angle-of-attack on commanded input. VGT optimizes performance over the different operating ranges, such as a large turbocharger performs like a small one at low speeds. Its schematic principle is illustrated in Figure 1.2a.

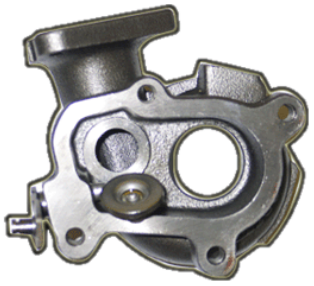
Wastegate (WG) A WG mechanism acts as a lever which vents exhaust gas, if needed. An example of WG built in the turbine housing is shown in Figure 1.2b.

Previously, we focused on the means of the exhaust gas flow control with a little reference to an actual actuator mechanism; the mechanism, which both WG and VGT systems are physically connected to, can be implemented in various ways. According to Ford Motor Company (2016), these include:

[3] The simplest turbocharger may not contain an actuator. Called free-floating, this type does not permit limit on boost pressure nor turbospeed.



(a) The VGT principle schematics. Figure adopted from Jääskeläinen (2016).



(b) A WG mechanism example. A "flapper" valve (on the left) lifts up and leaves the gases through. Figure adopted from Honeywell International (2017a).

Figure 1.2: Two common ways of controlling the boost pressure at the exhaust side of the turbocharger.

Pneumatic A pneumatic mechanism consists of a diaphragm and a spring. High pressure applied to the diaphragm[4] pushes the spring upwards and manipulates WG/VGT through a crank arm and rod. The mechanism is placed at the compressor side to isolate it from turbine heat. ECU controls the amount of pressure at the diaphragm by opening solenoid valve.

Vacuum A vacuum mechanism works on the same principle as a pneumatic actuator except that the medium at diaphragm is vacuum pumped from the engine; it proposes an advantage of control in all engine's operating ranges.

Electric In an electric actuator, a mechanical system is replaced by a DC motor driven by H-bridge from the ECU. A *feedback* from actual rod position is available.

One can come up with various designs of the actuator mechanism depending on the expected performance, robustness and cost. Several of these will be discussed in the later analysis, though recent applications favor electric actuators.

[4]Source of the pressure is the boost built up at the compressor outlet.

In turbocharged *gasoline* engines, the working principle of air flow is very similar. Primary difference is the placement of Electronic Throttle (ET) right after the compressor outlet to sustain air/fuel ratio λ close to its stoichiometric level. This step introduces another sensor which measures directly p_{2c}, complementing already mentioned p_{inlet} sensor.

Even though a throttle can be seen in diesel air-path, its function is to develop a down-pressure in the manifold to allow EGR flow in case of low engine back-pressure.

Also, throttled air stream requires a pipe connecting compressor outlet and inlet through a *bypass valve*. The valve opens up when a boosted air is restricted to flow (throttle closes down) and *surge* phenomenon starts to develop.

Interested reader can find more information about modeling and control of ICE in the book by Guzzella and Onder (2009).

1.2 Turbocharger as a Component

Now when the role of the turbocharger within the air-path system is introduced, it is crucial to understand the turbocharger as a standalone component. A cut of the turbocharger is shown in Figure 1.3.

Turbocharger system can be decomposed into two parts:

Center Housing Rotating Assembly (CHRA) A CHRA supports the shaft connecting compressor and turbine wheel. It picks up the oil and lubricates the *journal bearings* for low-friction rotational movement. The oil, also, cools down the whole CHRA. Piston rings seal the assembly and prevent oil leaking into the wheel area.

Compressor (Turbine) Housing The housing directs the fresh air/exhaust within the engine air-path. Its ability to house the capacity of flow is expressed by A/R ratio, where A stands for area of the inlet tube

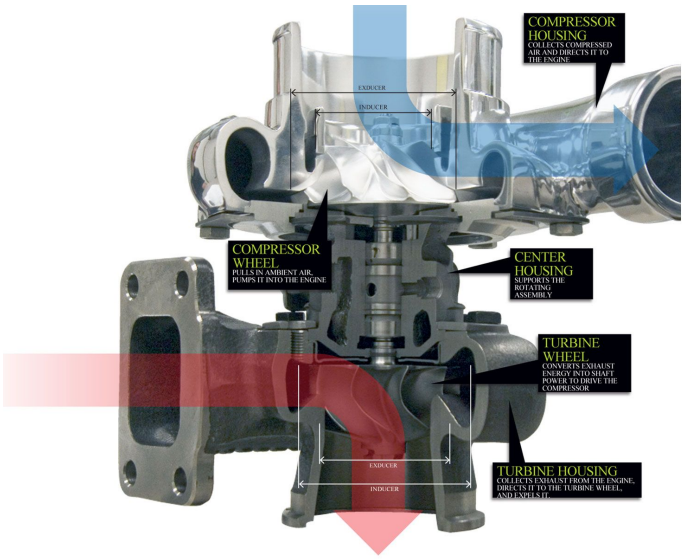


Figure 1.3: Garrett GTX3071R turbocharger cut and subcomponent description. An actuator is not displayed. Figure adopted from Munnell (2011).

surrounding the centroid and R is a distance from the turbocharger centerline to that centroid[5]. Also, it protects the other engine components from the wheel bursts.

The common shaft dynamics can be described by an equation of motion[6]

$$\frac{dN_t(t)}{dt} = \frac{1}{J_t}\Big(T_t(t) - T_c(t) - T_f(t) + T_{a(t)}\Big), \tag{1.1}$$

where J_t is a total shaft's rotational inertia, T_t is a torque generated by the turbine, T_c is a torque absorbed by compressor to compress the air, T_f covers a torque loss due to a shaft viscous friction, and T_a is a torque produced by an auxiliary source.

[5] Munnell, 2011.
[6] Guzzella and Onder, 2009.

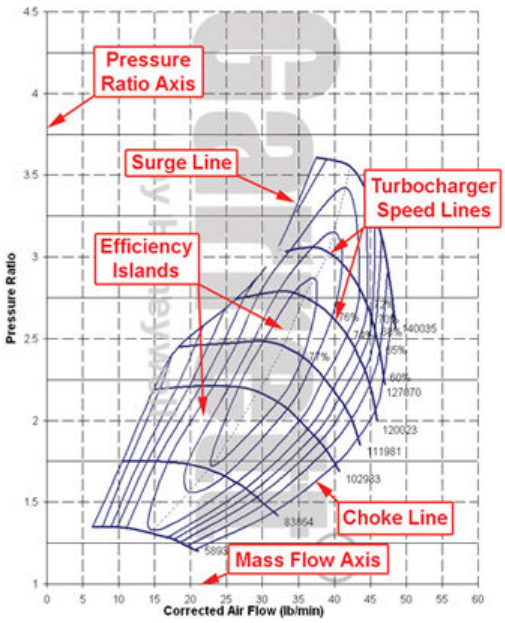


Figure 1.4: A compressor map graph. Figure adopted from Honeywell International (2017b).

1.2.1 Performance Characteristics

Every manufactured turbocharger is tested on a gas stand to obtain its performance characteristics. At reference temperature T_{ref} and pressure p_{ref}, a discrete map of compressor/turbine response is sampled. For a given turbocharger speed N_{t}, air mass flow entering the compressor is varied until boundary conditions of stable flow are achieved[7]. At each air mass flow value m_{c}, pressure ratio $\Pi_{\text{c}} = \frac{p_{2\text{c}}}{p_{1\text{c}}}$ is measured. Also, compressor outlet temperature measurement $T_{2\text{c}}$ helps to compute compressor isentropic efficiency η_{c}: a ratio of isentropic and actual work done by the system to compress the air. Turbocharger's speed is then increased until maximum mechanically-allowed rotational velocity is approached. A typical compressor map with indicated characteristics is illustrated in Figure 1.4.

Control design in ICE relies on Mean-Value Models (MVM), control-oriented

[7]Conditions when the mass flow is stalled (surge), or the volume flow at the compressor blade tips approaches sonic velocity (choke).

models accurate enough to capture transient behavior in temperatures and pressures, but, simultaneously, not too computationally complex. Since turbocharger manufacturers supply a test cell-measured map, a control design demands a control-oriented analytical model to be found. This implies compressor/turbine components to be *fitted* with functions describing their behavior continuously with respect to variables $\Pi_c, m_c, N_t, \eta_c, p_{1c}$, and T_{1c}[8]. Thus two functions $\Pi_c = f_{\Pi_c}(m_c, N_t, p_{1c}, T_{1c}), \eta_c = f_{\eta_c}(m_c, N_t, p_{1c}, T_{1c})$ are found by one of the methods available in the literature; for example, in Guzzella and Onder (2009).

1.2.2 Turbocharger Variants

Each car manufacturer expects different performance from the turbocharger system which may lead to various modifications while increasing the total cost. Supplementing the subcomponents aforementioned above, there are few improvements notable within the scope of the thesis.

A known issue with turbochargers is a so-called turbocharger lag. The larger turbocharger can provide higher boost pressure and increase the engine's rated power, but its inertia consumes more energy; thus the time from power demand to actual power delivery is increased. Their low-end performance is degraded as well, due to operation near the surge line in low engine speeds. Katrašnik et al. (2003) proposed adding an *electric motor* to the turbocharger's shaft to speed up the shaft by this motor when needed. Electric motor's developed torque corresponds to the term T_a in (1.1).

VGT utilizes functionality to guide the flow at the turbine wheel whereas Variable Inlet Guide Vane (VIGV) has the similar purpose at the compressor side—manages to lower A/R ratio when needed. It has been found out, by Uchida (2006), that VIGV and the casing treatment raises boost pressure at low engine speeds thus improving low-end torque. Optimal control can also be achieved in transient operations, and turbo-lag effect may be minimized compared to single vane turbocharger. The outcome of the system is similar to electric boosting and offers alternative to *two-stage* turbocharging, where

[8] Values p_{1c}, T_{1c} are used to correct air mass flow and turbocharger speed when there is difference from reference conditions p_{ref}, T_{ref}.

a small turbocharger optimizes performance in low air flows before large turbocharger starts providing high pressure ratio.

Honeywell International (2017a) proposes ball bearings as a replacement for journal bearings for high-speed applications. They claim resistance to shaft motion is decreased (term T_f in (1.1)) and, therefore, shaft can spin up faster. Additionally, less oil is required to lubricate the shaft, and assembly durability is increased.

Again, Honeywell International (2017a) states their Garrett turbochargers may be equipped with a water cooling system, supplementing the oil loop. Following hot-shutdown[9] a heat soak begins, eventually causing oil thermal instability, called *oil coking*, and degrading the engine oil with residues. Water-cooled center housing can acquire the heat energy from the system and prevent the oil from coking.

1.3 Turbocharger Description Summary

A turbocharger is the engine's air path component increasing the air pressure at the cylinder intake manifold, allowing the ECU to inject more fuel, hence produce more torque. The turbocharger is accompanied by several sensors required by ECU to compute optimal fuel injection properties. Therefore, the variables compressor pressure ratio Π_c, compressor inlet temperature T_{1c}, air mass flow m_c, and charge air temperature T_{CAC} are assumed to be known.

The ECU controls the level of boost pressure p_{2c} by manipulating the actuator. Either the VGT system or the WG mechanism can be deployed. The VGT adjusts vanes in the turbine to change the aerodynamics flow, the WG simply vents the exhaust gas if the boost pressure has to be decreased. Both mechanisms are at the turbine side, whereas the actuator is located at the cold compressor housing. The actuator is connected to the mechanism through a crank arm, bushings, and a rod. Pneumatic, vacuum, and electric means of actuation are possible.

[9] A state where the engine is shut down while the turbocharger is still very hot.

1 Turbocharger Description

The turbocharger itself is composed of a CHRA: the assembly of the shaft, nickel-based turbine wheel, and aluminum-based compressor wheel packed in the system of oil-lubricated bearings, plus a compressor and turbine housing. It is shipped together with a compressor (turbine) performance map. A compressor map links the m_c, N_t, Π_c, η_c measurements and is obtained at the laboratory gas test stand. For modeling and control purposes, this discrete map must be fitted with two functions $f_{\Pi_c}(m_c, N_t, p_{1c}, T_{1c})$, $f_{\eta_c}(m_c, N_t, p_{1c}, T_{1c})$.

2 Turbocharger State of the Art Diagnostics and Prognostics

In this chapter, we offer research of general prognostic models, related definitions, and a summary of current turbocharger diagnostic solutions. We have chosen two different authors dealing with the diagnostic and prognostic techniques and MATLAB's recent Predictive Maintenance Toolbox extensive documentation to support our solution. A current turbocharger implemented diagnostics is a part of OBD that detect faults leading to increased emissions through the observation of overboost and underboost conditions. Researched literature introduces model-based detection methods developed from the thermodynamic model of the turbocharger.

2.1 General Prognostic Models

In this section, we should clarify what is the role of this thesis within a Prognostics and Health Management (PHM) concept. This topic covers a broad spectrum of tools and methods for different engineering fields. Thus, it is essential to be clear what is the goal of the health monitoring system proposed in this thesis, how it can be related to current prognostic techniques performed, and what is *not* pursued in this work. We begin with the literature research of the general prognostic models, then we create a table that summarizes the researched models, and, in the end, we list a short example from the literature.

A work by Kumar and Pecht (2010) summarizes modeling approaches for PHM implementation. It is noted PHM is an approach that enables real-time health assessment of a system in its actual application conditions by sensing,

recording, and interpreting environmental, operational, and performance-related parameters that are indicative of a system's health. PHM can be further divided into three concepts:

Diagnostics The diagnostics estimates the system's health condition and provides FDI capability.

Prognostics The prognostics involves the assessment of system's health condition followed by modeling fault progression, health degradation, and Remaining Useful Life (RUL) estimation.

Health Management The health management provides the capability to make informed decisions about logistic actions based on diagnostics and prognostics information, available resources, and operational demand.

The authors grouped researched PHM models and algorithms into four categories, based on system's operation and data available:

Statistical Reliability-based Approach This approach is suitable for systems with insufficient sensor network, having short life, are non-critical, and involve low risk. Weibull distribution is the most appropriate statistical distribution which covers life cycle with respect to operational time.

Life Cycle Load-based Approach This approach consists of damage accumulation models depending on the usage profile. Includes fatigue life cycle computation as a function of operating conditions.

State Estimation-based Approach This approach takes advantage of the model already developed for control purposes. A single parameter, indicating system's health, can be estimated together with unmeasured states by means of an observer.

Feature Extraction-based Approach This approach derives a *feature* from the system data and assumes the feature varies when a fault is present.

A review of diagnostic and prognostic capabilities, by Vogl, Weiss, and Helu (2016), comments current PHM challenges; business level and human factor inputs are considered as well. Since PHM goal is to optimize the maintenance process, the approach must be carefully chosen because it is difficult to evaluate PHM cost savings.

Regarding measurement techniques, Vogl, Weiss, and Helu (2016) notes PHM designers must be able to use the available measurements placed for other functional purposes to assess health monitoring. This approach aligns with the hypotheses formulated in the beginning of the chapter, when list of available measurements was introduced. Identified PHM methods include

Experience-based Approach This approach uses human expertise for analysis and is the least complex.

Data-driven Techniques These techniques find non-linear relationships from the available data. Its advantage is that it might be applied to any system-subsystem-component level. Machine-learning and statistics-based approaches are counted within data-driven concept. Large amount of data is required to train the algorithms and prevent over- and under-fitting. Big data problem must be solved first.

Physics-based Approach This approach uses model-based solution and physics-of-failure-based failure propagation where accumulated damage is a function of the operating conditions.

The MathWorks, Inc. (2018) introduced in their 2018 releases a Predictive Maintenance Toolbox offering a set of models, algorithms, and guidelines to allow users to implement Predictive Maintenance[1] in an organized way. The core of a work-flow introduced by The MathWorks, Inc. (2018) is to assess the current system condition from the data by deriving a metrics called *condition indicator*. A condition indicator is a feature of system's data that distinguishes a healthy from a degraded system. Then a condition monitoring algorithm can perform fault detection and diagnosis based on current condition indicator value and known fault propagation in the system.

Prognostics algorithms process condition indicators to predict future behavior and assess most likely RUL. Depending on the source and type of the data available, Predictive Maintenance Toolbox offers three families of RUL models, namely:

Similarity Models These models make use of the available *run-to-failure* history of similar systems. Condition indicators show characteristic behavior as the similar systems degrade.

[1] Phrase Predictive Maintenance covers the same concept as PHM but is called differently.

	Identified Approach		
Available Data	Kumar and Pecht, 2010	Vogl, Weiss, and Helu, 2016	The MathWorks, Inc., 2018
Human Expertise	-	Experience-based	-
Recorded Lifespan	Statistical Reliability-based	Data-driven; statistics-based	Survival Models
Lifetime in Cycles	Life Cycle Load-based	Physics-based; physics-of-failure-based	Survival Models
Measurement Data	Feature Extraction-based	Data-driven; machine-learning	Similarity and Degradation Models
Mathematical Model	State Estimation-based	Physics-based; model-based	Degradation and Similarity Models

Table 2.1: An overview of researched PHM approaches and techniques. The first column keys out assumption on the type of available data from which we want to perform PHM concept. The body of the table lists concrete approaches identified in the literature.

Degradation Models These models fit condition indicator behavior to a linear or exponential function. The algorithm extrapolates the fitted function and searches for a most likely time when the indicator crosses failure-indicating threshold.

Survival Models These models perform statistical analysis to model time-to-event data. The models assume only lifetime data of similar systems (compared to run-to-failure indicator history in Similarity Models) are available. Working regimes and manufacturing batch may be included in the algorithm, too.

All of the three sources above made an effort to group their PHM knowledge into few models which can be applied in various systems. We have collected the methods and approaches proposed by the authors based on the assumptions they made on the data availability. We present the overview in Table 2.1.

We will develop health indicators from the mathematical model of the turbocharger. It may be the case that the directly measured signals will be insufficient for development of such an indicator. Then, we will need to use data-driven techniques to combine these measured signals with control inputs and, based on known turbocharger's dynamics, derive the indicator this way. No large datasets of lifespan records are available to us, therefore,

our methods should not rely on lifespan records. A possibility of a physics-of-failure-based approach will be analyzed when we identify turbocharger failure modes.

2.1.1 An Example of Data-Driven Prognostics Solution

A paper from Prytz et al. (2015) deals with a design of a predictive maintenance system that predicts the need for air compressor repair and thus minimizes commercial vehicle unexpected downtime and generates profit opposed to losses caused by unplanned stops. The authors present a supervised machine-learning approach that mines logged onboard data and original maintenance records from a fleet of commercial vehicles over a lifespan of three years. The solution falls to the second row in Table 2.1. Firstly, they discuss a problem of feature extraction from large and noisy data. Afterward, the authors show how feature selection affects accuracy and resulting profit. Prytz et al. (2015) conclude a motivation for predictive maintenance is cost saving which strongly depends on the real cost of planned and unplanned repairs, both challenging to estimate.

2.2 Turbocharger State of the Art Diagnostics

A faulty turbocharger will impact the exhaust emissions produced by the vehicle. This fact makes the turbocharger an On-Board Diagnosis (OBD) relevant component, therefore, its proper functionality has to be monitored.

2.2.1 Turbocharger Relevant On-Board Diagnosis

Legislative agencies all over the world demand vehicle manufacturers to implement Fault Detection and Isolation (FDI) system. Generally applied regulation for passenger cars is OBD in its OBD-II system version. A survey by Mohammadpour, Franchek, and Grigoriadis (2012) discusses common practice in designing OBD system. The goal of the system is to implement a FDI strategy for a variety of component, actuator, and sensor faults using

various data-driven and model-based methods. Also, the regulation states the system is calibrated to detect *single component* failure.

The paper above comments OBD-II is mainly deployed to ensure emission standards are followed in the vehicle operation. Even though the article deals mostly with fuel and after-treatment devices, it is noted sensor systems and air leaks must be monitored as well. Methods introduced by Nyberg and Stutte (2004) use a physics-based approach to model faulty air-path system components, such as MAP sensor multiplicative bias, EGR valve stuck closed, or manifold air-leakage.

OBD documents created by Ford Motor Company (2017) for both diesel and gasoline engines illustrate implementation of a turbocharger monitor. Actually, it is a boost pressure control system which is monitored and where possible turbocharger-related faults manifest themselves. Diesel engine, containing oil-pressure actuated VGT without position sensor, performs three monitors on the ECU:

Intrusive The intrusive response monitor executes a sweep (when idling) of VGT from 85% to 25% with EGR valve shut and expects a separation of $>$ 2 kPa in MAP readings after 4 seconds. If the separation is not achieved, a Diagnostic Trouble Code (DTC) "Turbocharger Boost Control Performance" is set.

Overboost The overboost monitor continuously checks if actual boost pressure and desired boost pressure are not separated by a predefined threshold for a predefined duration. If so, a DTC "Turbocharger Overboost Condition" is set. Cause of the failure may be VGT stuck closed.

Underboost The underboost monitor works similarly as the overboost monitor except that desired boost pressure is now not achieved within a predefined duration. In addition, monitor checks if control effort in different operating ranges is not above calibrated threshold. DTC "Turbocharger Underboost Condition" or "Turbocharger Boost Pressure Low" is set. This failure may be due to air-path or exhaust leaking, slow response, or VGT stuck open.

Ford gasoline engines, on the other hand, limit boost pressure with WG actuator only. Multiple options of actual actuator implementations have been discussed in section 1.1; each of them has its own monitor. Pneumatic and vacuum systems are monitored only from high-level performance

(underboost/overboost); if an actuator pressure sensor is available, low-level control loop performance is monitored as well. Solenoid regulating amount of pressure at diaphragm is checked for short/open circuits. Electric actuators are supplied with position feedback, thus monitoring difference of desired and actual position is possible. Excessive mechanical wear is detected when sensor readings are out of operating range.

Ford Motor Company (2017) lists the sensors that must be fault-free to perform turbocharger monitor in OBD. If the conditions are not met, the monitor can not be performed. We adopt the same assumption; the whole air-path is working in *fault-free* operation[2], leaving the health monitoring focused on the turbocharger component alone, in the ECU software layer above the OBD system.

2.2.2 Model-based Turbocharger Fault Detection

A PhD thesis by Ceccarelli (2012) describes a design of adaptive *observer* for fault-detection of intake leakage and fault-diagnostics for turbine isentropic efficiency η_t loss. The cause of η_t drop is not discussed in the report. Ceccarelli (2012) states on-line state estimation algorithms are popular in domain of diagnostics, but realizes that working with accurate physical turbocharger model is not feasible on car's ECU. Even the already simplified control-oriented MVM models must be approximated further to count with few production engine sensors. A model used in the work is created by modifying the equation (1.1), acquiring state-space equation

$$\frac{\mathrm{d}N_t^2}{\mathrm{d}t} = \frac{2}{J_t}(P_t - P_c), \tag{2.1}$$

$$y = N_t^2, \tag{2.2}$$

[2]No leaks has occurred and sensors are fault-free.

where P_t, P_c are turbine and compressor powers, related to torques (defined in (1.1)) by relation $P_{t(c)} = T_{t(c)} N_t$. These powers are computed as

$$P_t = m_t c_p \eta_t T_{1t} \left[1 - \left(\frac{1}{\Pi_t} \right)^{\frac{\gamma-1}{\gamma}} \right], \tag{2.3}$$

$$P_c = m_c c_p \frac{1}{\eta_c} T_{1c} \left[\Pi_c^{\frac{\gamma-1}{\gamma}} - 1 \right], \tag{2.4}$$

where c_p is a specific heat at constant pressure, γ is a ratio of specific heats, $\Pi_t = \frac{p_{1t}}{p_{2t}}$ is a pressure ratio across the turbine, m_t is exhaust mass flow, and $m_c = W_{HFM}$ is air mass flow.

Ceccarelli (2012) proposed following observer, where $x = N_t^2$ is the state from state-space equation (2.1),

$$\dot{\hat{x}} = \frac{2}{J_t} \left(m_t c_p \hat{\eta}_t T_{1t} \left[1 - \left(\frac{1}{\Pi_t} \right)^{\frac{\gamma-1}{\gamma}} \right] - P_c \right) + K(y - \hat{x}), \tag{2.5}$$

$$\dot{\hat{\eta}}_t = \frac{2}{J_t} m_t c_p T_{1t} \left[1 - \left(\frac{1}{\Pi_t} \right)^{\frac{\gamma-1}{\gamma}} \right] K_{\eta_t} (y - \hat{x}), \tag{2.6}$$

K, K_{η_t} are tuning scalar constants. In order to compute turbine efficiency estimate $\hat{\eta}_t$, assumptions are introduced

1. $T_{amb} = T_{1c} = T_{2t}$; compressor inlet/turbine outlet temperature is equal to ambient temperature,
2. $p_{amb} = p_{1c} = p_{2t}$; compressor inlet/turbine outlet pressure is equal to ambient pressure, thus no pressure drop at filter/after-treatment system is presumed,
3. $y = N_t^2 = a\Pi_c + b$; a shaft speed is not measured but is estimated from pressure ratio Π_c measurement and calibrated constants a, b,
4. compressor efficiency η_c is computed from a fitted function, such as $f_{\eta_c}(m_c, N_t, p_{1c}, T_{1c})$ described in section 1.2.1,
5. turbine mass flow is calculated as $m_t = m_c + m_f$, where m_f is fuel flow supposed to be known, and
6. T_{1t}, p_{1t} are exhaust manifold's temperature and pressure, respectively. These two variables are assumed to be *measured* and violate our

hypothesis introduced in section 1.1, where we state exhaust pressure/temperature reading is not available in production passenger vehicle.

The work by Ceccarelli (2012) then deals with a design of an adaptive threshold—threshold dynamically adjusted in different operating points to prevent false alarm triggering and to count with sensor and modeling uncertainties. On-line diagnosis algorithm computes *residual* $r_{\eta_t} = \hat{\eta}_t - f_{\eta_t}$[3] and compares it with the threshold.

A recent book on model-based ICE condition monitoring by Isermann (2016) deals with the diesel engine turbocharger diagnosis in a similar manner as Ceccarelli (2012). Model used in the work is the same as in (2.1) but with bearings friction losses covered by a power loss $P_f = c_f(T_{oil})N_t^2$, where $c_f(T_{oil})$ is a viscous friction coefficient dependent on lubricating oil temperature. Alternative way to express non-ideal rotational movement of the shaft is to introduce mechanical efficiency η_m and a new turbine power $P_{tm} = \eta_m P_t$.

Isermann (2016) proposes a model-based symptom generation concept—residuals are computed as a difference of variable computed from measurements, for example, power P_i, turbocharger speed N_t, or charging pressure p_{2c}, and the same variable predicted from a calibrated model at actual operating condition[4]. For VGT turbocharger it is claimed a fault "VGT blocked middle position" is detected and isolated from power residuals r_{P_i}, and "compressor blades damaged" is observable only in charging pressure residual $r_{p_{2c}}$. For WG turbocharger diagnosis is applied only in closed position where models are simpler. Through the charging pressure residual $r_{p_{2c}}$ faults "decreased turbocharger efficiency" and "leaky wastegate" can be detected.

Both of the works from authors Ceccarelli (2012) and Isermann (2016) focus only on the fault detection, none of them extended their work to the prognostics concept. Ceccarelli (2012) designed an observer only for a single fault, a turbine efficiency drop, whereas Isermann (2016) works

[3]Residual is a difference between observed $\hat{\eta}_t$ and modeled efficiency which is computed from fitted efficiency map $f_{\eta_t}(m_t, N_t, p_{1t}, T_{1t}, u_{VGT})$.

[4]More precisely, charging pressure is computed from *steady-state* power balance equation (2.1). Author expects measurement p_{1t}, T_{1t} are available.

with the parity equation concept to diagnose several turbocharger faults. Their models rely on the turbocharger power balance model defined in (2.1). This model takes into the account the turbine, mechanical, and compressor powers. However, in the production engine, turbine side measurements are unaccessible. Both authors rely on them but do not address their estimation nor the error analysis. From our point of view, the accuracy required to perform the prognostics in the turbocharger can not be attained with the model counting on the turbine pressure/temperature measurements T_{1t}, p_{1t} and the turbocharger efficiencies η_t, η_c, and η_m. This approach certainly lacks redundancy, robustness, and fault isolation properties.

3 Turbocharger Fault Cause and Effect Analysis

In this chapter we present an analysis of relations between turbocharger-relevant fault and its produced effect. It is crucial to understand individual fault behavior in PHM concept to develop tailored features that deviate as soon as an unpermitted change occurs. Primarily we focus on a set of observable faults, i.e. faults producing a *symptom*—visible change in the nominal behaviour of the system. Although we expect some of the faults–primarily those related to material fatigue–to be unobservable from their nature[1], we will discuss them as well.

3.1 Failure Modes

Before grouping the faults caused by a similar source or having a similar effect on the system, we analyze most common failure modes in the field, reported by largest turbocharger distributors.

Owen (2012) lists following relations: the correct pressurized *oil supply* in the bearings must be ensured otherwise quick bearings failure will follow; *oil contamination* induces seals and bearing wear; *compressor wheel damage* results from injection of a foreign object, or from fine particle impact in the form of pitted and worn blades, or due to the bearing failure and direct contact with the housing. *Performance loss* is produced by carbon and oil contamination on the volute surfaces of a turbine and a compressor leading to the wheel efficiency drop; *turbine wheel failure* results from overheating, over-speeding, or suction of broken engine parts; *boost pressure* control malfunction due

[1] A crack developing at micro-structural level.

to a failed actuator; *turbocharger overheating* caused primarily by incorrect fueling.

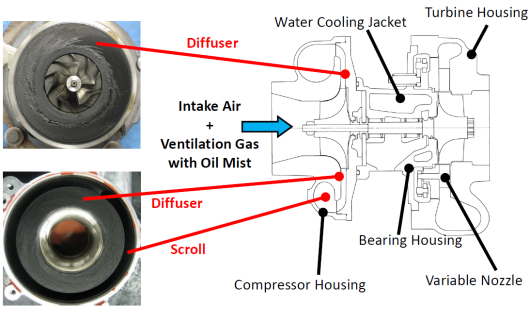
The distributor BTN Turbo (2012) notes turbochargers are very reliable and that less than 1% of turbos fail due to a manufacturing fault with the turbo itself. 95% of turbocharger failures are blamed to be caused by *oil starvation*, *oil contamination*, or *foreign object damage*, aligning with the report by Owen (2012). Within the remaining failures, BTN Turbo (2012) states severe *oil leaks* past the piston rings can occur when the low pressure at the compressor side remains for a long time; *over-speeding* as a consequence of the other root cause; or *electronic actuator* malfunction being a common turbocharger field issue.

3.1.1 Turbocharger Performance Degradation

We define that turbocharger operates with a degraded performance when the work that has to be done by the system to perform a specific task is *increased* compared to the work done to perform the same task at the beginning of turbocharger's operational life. The degradation is therefore linked with an *efficiency* parameter. To recall, we defined compressor isentropic efficiency η_c as a ratio of isentropic and actual work done by the system to compress the air. Thus for the degraded turbocharger, more work has to be delivered to maintain the equal level of boost pressure, resulting in higher shaft velocity N_t, smaller η_c parameter, and higher compressor outlet temperature T_{2c}. For the turbine, the working principle is similar but with "opposite direction"; more work has to be extracted from the exhaust flow to spin the shaft to equal velocity.

We can reformulate the problem to a compressor and a turbine wheel efficiency drops η_c and η_t, respectively. We treat both issues independently: an η_c loss is decoupled from η_t loss. A mechanical efficiency η_m related to the performance of the shaft-bearing coupling will be discussed in the bearing issues analysis.

Multiple causes lead to an effect of compressor/turbine efficiency loss. They may originate from



(a) The oil residues primarily lays in the diffuser, reducing cross-sectional area. The source of the oil particles is a Crankcase Ventilation (CCV) system. Blow-by gases from crankcase, including oil droplets, are brought back to the compressor inlet. Figure adopted from Hirano and Decker-Brentano (2017).

(b) A wheel wear caused by small object damage. If the blades are uniformly damaged and the wheel remains balanced, the turbocharger may continue in its operation. The material gets weakened after the impact and fatigue life decreases. Figure adopted from Polichronis et al. (2013).

Figure 3.1: Illustration of various sources of decreased compressor/turbine wheel efficiency.

1. deposit of oil/carbon particles within the flow-carrying ducts (i.e. diffuser, Figure 3.1a),
2. reduced surface smoothness due to the fine particle impact, and
3. minor blade damage due to the poor air filtration (compressor) or worn engine debris release (turbine). Effect displayed in Figure 3.1b.

All of the sources contribute to the additional *restriction* in the flow, eventually causing irreversible loss of the energy and consequent wheel efficiency drop. Adamkiewicz (2012) analyses diagnostic relations of marine Diesel engine turbochargers and states compressor's diffuser cross-sectional area can be decreased by $10 - 20\%$ because of oil deposits. Although larger engines with turbocharger designed for heavy-duty applications show a more severe impact of the efficiency drop, it does not exclude passenger vehicles from victims of performance degradation.

OBD detects degraded performance in high-load operations when the actuator is no longer able to compensate power loss and desired boost pressure is not achieved. A DTC linked to Underboost is triggered.

3.1.2 Turbocharger Bearing Wear

A recent passenger vehicle's turbocharger bearing system is equipped with a thrust bearing supporting the shaft in the axial direction and a couple of semi-floating journal bearings supporting the shaft in the radial direction. Zhang et al. (2010) comment that the thrust bearing compensates axial thrust load caused by pressure differential at compressor and turbine rotors. A major turbocharger failure occurs when the thrust bearing is no longer able to withstand the axial load and allows the axial displacement to exceed the limit, resulting in the turbine's psychical contact with the housing. Such a failure is shown in Figure 3.2a.

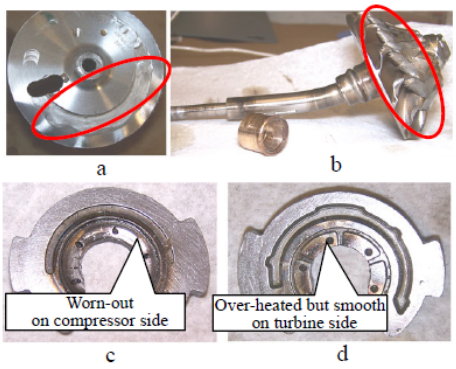
Semi-floating journal bearings, one at each side of the shaft, allow the lubricant to flow between the bearing housing and the shaft to simultaneously permit friction-less motion, cool the whole CHRA, and act as a damper. Polichronis et al. (2013) blame 75% of all turbocharger failures originating from lubricating issues, which seems as biased statistics to us. Worn and damaged bearings may arise from

1. insufficient lubrication flow,
2. lubrication feeding lag during heavy transient operations,
3. oil contamination by foreign particles. These particles are either collected in the oil loop (having degraded oil filter) or developed from a coking process under excessive oil temperature. Additionally, charred oil residues block the oil stream and contribute to causes 1. and 2.

In Figure 3.2b an over-heated shaft and worn bearings, due to the lack of lubrication and consequent metal-to-metal contact, are illustrated.

We can conclude oil quality/quantity and bearing damage relation is covered by complex behavior. We have not found any proof that worn and damaged bearing leads to a decreased mechanical efficiency η_{m}. Instead, the rotor starts vibrating because of increased shaft-bearing clearance, thus leading to the reduced radial wheel-housing clearance and eventually causing turbocharger failure.

Many researchers pursue to develop a method to estimate the shaft motion.



(a) Worn-out thrust bearing, leading to a major turbocharger failure. Figure adopted from Zhang et al. (2010).

(b) Shaft and bearing wear caused by lack of lubrication supply. Blue discoloring results from over-heating of the shaft due to metal-to-metal contact. Figure adopted from Polichronis et al. (2013).

Figure 3.2: Illustration of various turbocharger bearing system cause-effect relations.

A promising non-invasive[2] optical technique, introduced by Pastor et al. (2012), uses a camera and two bolts to detect a shaft whirl when the lubrication is cut. An increase of maximum shaft's eccentricity is recorded when the turbocharger is close to its end-of-life, as illustrated in Figure 3.3.

3.1.3 Foreign Object Damage

A solid object entering the compressor/turbine inlet has a fatal consequence for the turbocharger. If the energy of the object is sufficiently low, the wheel blades may get bent, worn, and structurally weakened—but the turbocharger remains functional with the effect we have discussed in 3.1.1. When the object is sucked at high speeds, it causes a blade to get broken and induces collateral damage failing all blades. An OBD must quickly react to prevent further engine damage.

Failure due to a Foreign Object Damage (FOD) results from improper engine maintenance. The FOD is an instantaneous process—no transient effect occurs between fault origin and turbocharger failure. Also, it depends only

[2]Opposed to accelerometer-based or infrared-camera techniques where significant modification of the system must be done.

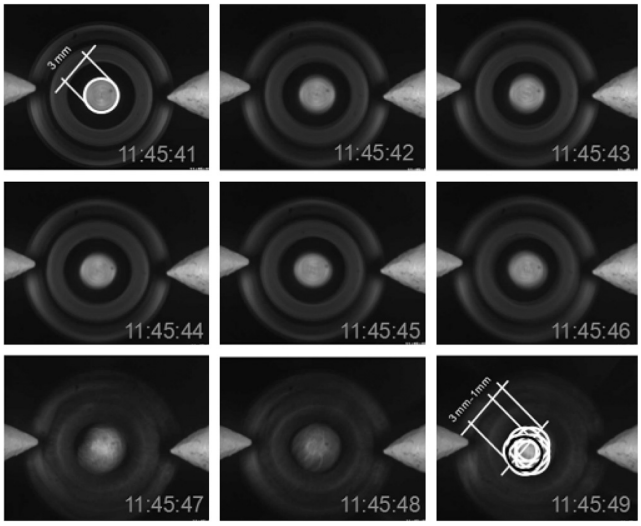
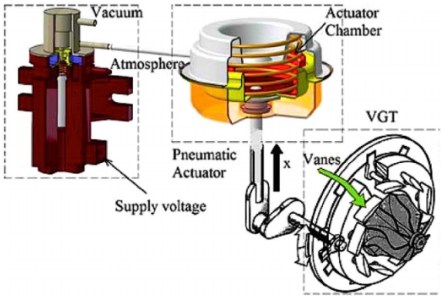


Figure 3.3: Shaft's whirl motion recorded when the bearing lubrication was cut. Figure adopted from Pastor et al. (2012).

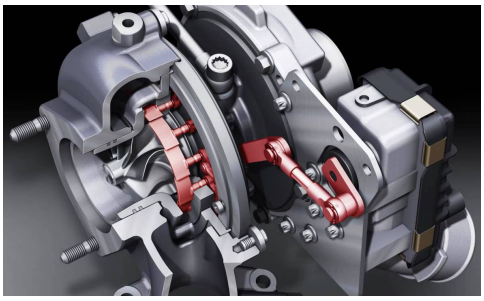
on the external conditions. Therefore we rule out FOD from our monitoring analysis.

3.1.4 Actuation Issues

In the section 2.2.1 we have briefly discussed implemented solutions for detecting actuator malfunction. A chosen fault diagnosis approach strongly depends on the actuation mechanism applied, and it would be exhausting for us to analyze diagnostic relations for all the possible components and subcomponents. A detailed overview of existing VGT systems can be found in work by Feneley, Pesiridis, and Andwari (2017). The most common failure mode for VGT mechanisms is *vane sticking* to the backplate—the vanes get stuck and do not react to a commanded input. A Rotary Electronic Actuator (REA) manufacturer, Melett Ltd. (2017), states carbon and sludge built up at the vanes will cause intermittent sticking which DC motor tries to compensate. The motor then overheats and fail other components such as gearbox, or electronics board. Feneley, Pesiridis, and Andwari (2017) mention metal-to-metal contact, causing sticking, is induced at elevated



(a) Principle of vacuum-actuated VGT mechanism. Figure adopted from Feneley, Pesiridis, and Andwari (2017).



(b) Mechanical linkage between the electric actuator (box with the black cover) and the vanes, highlighted in red color. Figure adopted from Arbore (2015).

Figure 3.4: A comparison of VGT actuators, focused on the linkage connecting the actuator and the vanes.

temperatures. Figures 3.4a and 3.4b illustrate pneumatic and electric VGT actuator implementations.

On the other hand, WG flapper does not suffer from sticking but has to cope with the high pressure exerted by exhaust gas in a closed position. The actuator must, therefore, withstand heavy cyclic loading. ANSYS, INC. (2016) developed a computer-aided design software which is used by Borg-Warner engineers to predict wear life of the mechanism. Force acting on the flapper is simulated by 12.8 kg mass. A shaft connecting the flapper and the actuator, a crank arm, and bushings are the most critical elements that wear out over the lifetime. The software can predict accumulated wear as a function of operating cycles resembled by artificial crankshaft rotation in the test rig. Thus the wear can be extrapolated in the future and component lifetime estimation in cycles is acquired. The wear accumulated over two open-and-close cycles is shown in Figure 3.5.

We conclude an actuation mechanism is susceptible to faults of various sources depending on the actual actuator design. However, every design contains a mechanical linkage connecting the actuator at cold compressor side and the flapper or vane assembly at hot turbine side. This linkage consists of a shaft, crank arms, and bushings that wear out over the cyclic life of the actuator in a predictable way. From the actuator's point of view, the wear is observable as a *hysteresis*. An existing OBD implementation, by

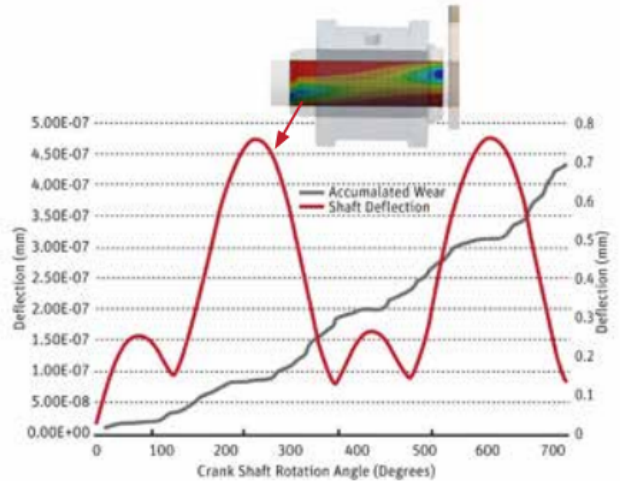
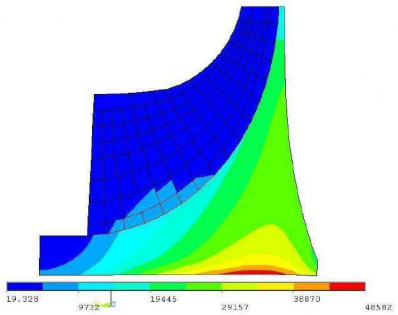


Figure 3.5: Overall accumulated linkage mechanical wear (gray line) and stress acting on the shaft (produced shaft deflection as a red line) with respect to two open-and-close cycles. Figure adopted from ANSYS, INC. (2016).

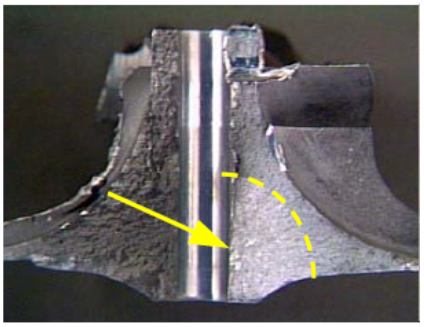
Ford Motor Company (2016), detects worn mechanical linkage through observation of position sensor feedback in electric actuator. When the rod travels out of operational range, the ECU triggers corresponding DTC.

3.1.5 Material Fatigue

Floren (2011) discusses turbocharger housing must be designed to protect the engine when the wheel bursts. The wheel's rotational kinetic energy has to be absorbed by the housing thus the debris remains contained within the turbocharger without harming other components. Floren (2011) states either blade burst or hub burst occur. A single blade may detach from the hub when the wheel's rotational velocity N_t exceeds the maximum allowed rotational velocity $N_{t_{max}}$ where the material's strength at blade root section can no longer resist exerted centrifugal force. The cause of a wheel hub burst is material fatigue; more precisely, a Low Cycle Fatigue (LCF) phenomenon. Both failure modes affect mostly aluminum-based compressor wheels because of their lower material strength limits, compared to nickel-based Inconel turbine wheels. Highest stress region of the compressor wheel



(a) Stress concetration in a wheel spinning at full speed. Peak is present at bore section of wheel. Figure adopted from Floren (2011).



(b) A wheel hub failure due to LCF phenomenon. Half-moon shaped fatigue crack is highlighted. Figure adopted from Engels (2002).

Figure 3.6: LCF cause-effect relation.

is at bore section of the wheel, as illustrated in 3.6a. Engels (2002) further explains it is *low-frequency cycling* of the wheel that causes LCF failure. An accumulation of cyclic deformation leads to micro-structural fatigue cracks development which most likely starts at the bore section. Then, these cracks propagate into macro-structural cracks up to the point where the material is not able to withstand the centrifugal force causing the wheel to fail[3]. Figure 3.6b shows a cross-sectional view of a failed wheel due to LCF.

The number of cycles compressor wheel undergoes until it fails depend heavily on the cycle amplitudes and the wheel's temperature. Expected lifetime, for a predicted mean vehicle's duty cycle, is given in kilometers or hours as a probabilistic distribution. A medium-duty truck, or a bus, driving in the city has several-times lower expected lifetime than a highway or country-road driving vehicle. It is a "full load-brake" driving cycle killing the wheel fastest. Engels (2002) proposes titanium wheel with lifetime approaching infinity. To illustrate an example of typical aluminum wheel's life expectancy: a probability of failure equal to 2% for a city bus is 98 000 km and for a country road vehicle is 1 012 000 km.

Another known turbocharger-related fatigue phenomenon is High Cycle Fatigue (HCF). HCF is induced by high-frequency loading of the blades and

[3]The operating point when the wheel bursts is within designed limits. It is the cycling that weakened the material, opposed to a blade burst.

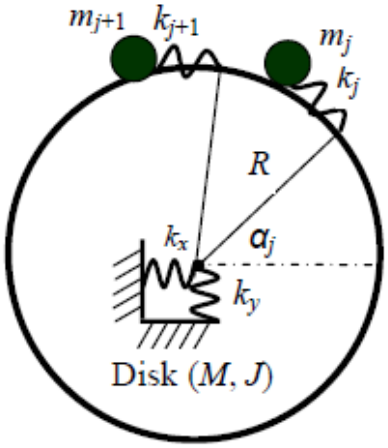
becomes significant when the loading frequency matches wheel's natural frequency. Because of the flow dynamics in the turbocharger, only the turbine wheel is affected (see Figure 1.3 where flow direction is illustrated). During the turbocharger design and engine-matching process designers ensure that over the whole operating range of the engine these critical vibration modes are not achieved.

Sheng, Clay, and Allport (2006) discuss a turbine wheel mistuning problem. Even though the wheel is designed correctly, a manufacturing process[4] may cause a shift in wheel's natural frequencies, and then HCF becomes a serious issue. Sheng, Clay, and Allport (2006) introduce a method that identifies mistuning from the wheel's response. They use a low-order lumped parameter model of the wheel where the blades are modeled as masses m_j connected to the hub through springs with stiffness k_j. The model is illustrated in Figure 3.7a.

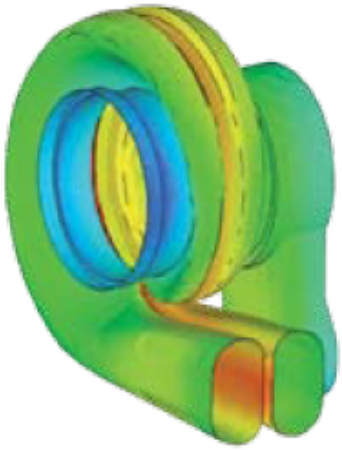
The last type of fatigue is called Thermo-Mechanical Fatigue (TMF). Eastwood and Allport (2007) comment TMF is caused by exhaust temperature cycling during engine's operation. Turbine's housing experiences high thermal gradients in high-load operations cycled with low exhaust temperatures while idling. Each cycle contributes to thermal stress, and when enough damage is accumulated, the housing starts cracking. A temperature distribution, predicted by finite element analysis software, is illustrated in Figure 3.7b. The area highlighted in red color is where cracks due to TMF will start.

We believe a considerable effort is made to develop a turbocharger with expected fatigue life demanded by a vehicle manufacturer. Specifically, HCF should not occur because the turbine wheel's natural frequency does not get excited. TMF has very long expected lifetime, and if a crack in the housing happens to develop, it does not lead to a failure. Compressor wheel's LCF is a fatigue mode most likely to happen, but with a lifetime long enough the wheel outlives most of the other engine components (at least in a typical passenger vehicle).

[4] Generally, turbine wheels are casted.



(a) Turbine wheel low-order model (opposed to high-order finite element analysis model) with blades modeled as masses m_j connected to the hub through springs with stiffness k_j. Figure adopted from Sheng, Clay, and Allport (2006).



(b) Temperature distribution in turbine volute, predicted by finite element analysis software (red is the hottest). Figure adopted from Eastwood and Allport (2007).

Figure 3.7: An illustration of housing's temperature model, used in TMF analysis, and turbine wheel's lumped parameter model, used in HCF mistuning identification.

3.1.6 Other Turbocharger Issues

In sections from 3.1.1 to 3.1.5, we have systematically grouped failure modes originated from a similar source or having a similar effect on the turbocharger. These failure modes were reported to have the highest rate of occurrence in the field. We close the section by complementing the list of failure modes to cover the whole turbocharger system. Other turbocharger issues include manufacturing issues, mismatched operating range, and improper turbocharger operation and maintenance.

We have already discussed an impact of manufacturing issue when dealing with HCF. HCF is not a field issue unless a turbine wheel is cast outside of the process limits. The same applies for each turbocharger's subcomponent. High manufacturing process quality must be followed to fulfill the predicted turbocharger lifetime.

A turbocharger is designed to efficiently work in a variety of operating conditions characterized by compressor/turbine maps (see compressor

map example in Figure 1.4). Ideally, the operating point should be close to the peak efficiency. Map boundaries represent physical limits of a turbocharger's operation. A mistuned controller or incorrect operation that makes turbocharger surge, choke, or overspeed, will significantly reduce expected lifetime. Also, for example, exposing the turbine to temperatures not accounted with in the TMF analysis will impact the thermal durability of the turbine's housing.

Since a turbocharger is a subcomponent in the ICE system, it is affected by other components' behavior as well. Failure modes performance degradation, bearing wear, and FOD originate only because the vehicle is not maintained the way the manufacturer recommends.

We conclude the section with a statement: If a turbocharger is manufactured according to the demanded standards, properly matched to the engine and its controller, and the engine and other relevant subcomponents are maintained regularly, one can expect the turbocharger to work as long as LCF permits.

3.2 Fault Observability Summary

Turbocharger performance degradation results from a compressor/turbine efficiency drop caused by either minor FOD or oil particle deposit at the diffusor area. The FOD occurs randomly and has an unpredictable effect since the size and shape of an object entering the compressor is unknown. On the other hand, oil fouling of the compressor impacts turbocharger's performance in a systematic way as the oil droplets accumulate a deposit of a certain thickness at the diffusor area. Since turbocharger's performance is linked with overall engine's performance, the effect of efficiency drop produces a symptom in the engine's response that our health monitor must acquire and track since then. We will design the health monitor system for the compressor efficiency drop in the following chapters.

Bearing wear is a frequent and serious issue that leads to a failure of the turbocharger without a warning. The shaft's axial and radial displacement,

due to the bearing wear, is not obervable in production vehicles. An accumulated wear is linked to an oil contamination and delay of sufficient lubrication especially in transient and high-load operations. An oil quality sensor or oil pressure feedback (before the turbocharger) is missing which makes it infeasible to develop models predicting bearing wear from oil properties. These are the reasons we can not implement bearing wear monitor in production vehicles at the moment. This issue will no longer be discussed in the thesis.

Turbocharger's life expectancy distribution due to LCF is computed in the design stage to match vehicle manufacturer's requirements (for predicted mean vehicle's duty cycle). Engels (2002) explains the steps to calculate lifetime are:

1. from the test cell, acquire a Woehler-Curve (S-N-Curve) for each different type of compressor wheel,
2. from the vehicle's manufacturer, collect the turbo speed measurement from the expected mean vehicle's duty cycle,
3. use rain-flow method for statistical counting of the load cycles, and
4. use Miner's rule for linear damage accumulation.

LCF failure mode can be modeled by Life Cycle Load-based or Physics-of-failure-based approach, described in Table 2.1, in PHM approach. We could collect the actual load cycles on-board and perform steps 3 and 4 to update the damage accumulated and correct for real driving cycle difference from the expected one. However, LCF has no direct impact on the engine's performance. Thus, we leave it for future work.

Actuator's mechanical linkage is subject to wear as well. One option to predicted the linkage wear is to use the load cycle model (an example of cycle-to-wear model is illustrated in Figure 3.5). Then the PHM approach is similar to LCF life expectancy model. If the turbocharger contains an electric actuator (REA), a position measurement from the mechanism is received, and accumulated wear can be predicted directly from this feedback measurement. For example, having an electric WG actuator in the turbocharger, one can sample the rod position when WG is closed (flapper sits down) and fit it with a function of time to extrapolate the behavior. We expect that the measurement when flapper is closed will vary as the wear increases. A real actuator's mechanical wear is essential to predict because the ECU

must compensate the hysteresis to maintain the same performance over the whole life of the engine. To conclude, a fatigue model and sensors available depend strongly on the actuator configuration. We do not possess any concrete actuator data; hence the actuator monitor wear is left for future work.

4 Monitoring of Compressor Efficiency Drop

From now on, we will deal only with the compressor performance degradation failure mode. That is, we propose a model of the fault and then introduce an architecture of our health monitor that detects and tracks the level of performance drop. Performance (efficiency) drop due to oil fouling issue is an excellent candidate to model because the effect (efficiency drop) is predictable and once the fouling process has started, the deposit accumulates over a time-span long enough to perform the prognostics.

4.1 Modeling of Compressor Efficiency Drop due to Oil Fouling

We were provided a set of four compressor map pairs; each pair consists of a fouled and a clean compressor map. These compressors come directly from the field operation; the level of fouling was high enough that the lack of performance was registered either by a driver or by an ECU. After the gas stand measurement (to obtain fouled map characteristics), the oil deposits had been washed away, and a clean compressor's map was collected. Therefore we can now evaluate diagnostic relation of a fouled-clean compressor. It was also demonstrated that a cleaned compressor has identical performance characteristics as a new one of the same type.

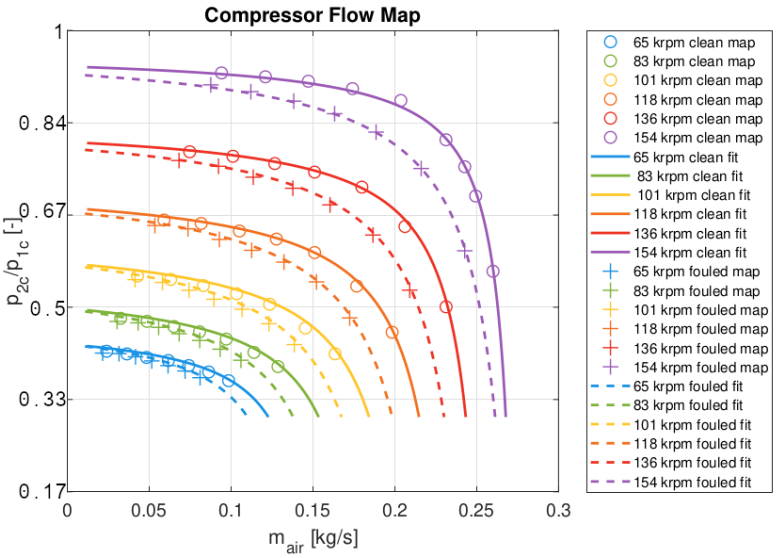


Figure 4.1: A normalized pressure ratio as a function of turbocharger rotational velocity and mass air flow for an example turbocharger's fouled and clean map, plus a model identified by non-linear regression on the sampled compressor data. The normalized pressure ratio 1 means maximum pressure ratio for this particular map.

4.1.1 Compressor Maps and Models

These provided four pairs of compressor maps include medium-size wheel diameters from 59 mm to 76 mm, used in both diesel and gasoline engines. A measured pressure ratio Π_c for a selected fouled-clean map pair is scattered in Figure 4.1.

Figure 4.2 illustrates acquired and normalized compressor efficiency η_c from the same compressor map. We can see the measurement point is described by tuple (N_c, m_c) and that N_c remains same for fouled-clean pair but m_c differs. The whole map of the fouled compressor is shifted towards the smaller values of mass air flow m_c because of change in flow aerodynamics due to oil deposit. Since we want to evaluate the difference of pressure/efficiency, an operating point tuple (N_c, m_c) for both maps must be the same. This leads us to *fitting* of a clean compressor map with continuous functions f_{Π_c}, f_{η_c}, and then sampling it in fouled map tuples (N_c, m_c).

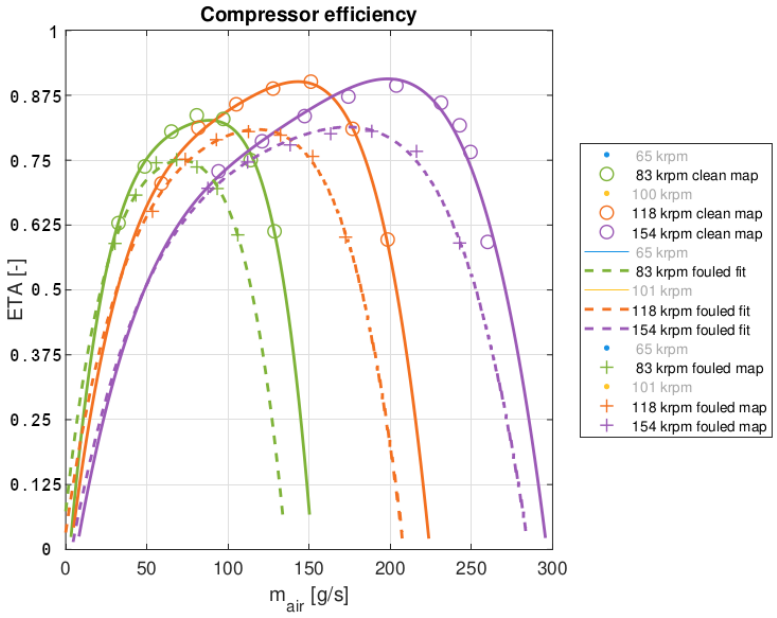


Figure 4.2: A normalized isentropic efficiency as a function of turbocharger rotational velocity and mass air flow for an example turbocharger's fouled and clean map, plus a model identified by non-linear regression on the sampled compressor data. Some turbocharger speed curves were omitted for better readability, but the quality of the fit is comparable. The normalized efficiency 1 means maximum efficiency for this particular map.

A systematic approach to modeling of the compressor is discussed in master's thesis by Tabaček (2016). Inspired by this thesis, we define an empirically derived pressure ratio model $\Pi_c(\widetilde{N}_t, \widetilde{m}_c)$ and efficiency model $\eta_c(\widetilde{N}_t, \widetilde{m}_c)$ as

$$\Pi_c(\widetilde{N}_t, \widetilde{m}_c) = \frac{A(\widetilde{N}_t, \widetilde{m}_c)}{B(\widetilde{N}_t, \widetilde{m}_c)}, \tag{4.1}$$

$$\begin{aligned} \eta_c(\widetilde{N}_t, \widetilde{m}_c) &= c_0 + c_1\widetilde{N}_t^{-1}\widetilde{m}_c + c_2\widetilde{N}_t^{-2}\widetilde{m}_c^2 + c_3\widetilde{N}_t^{-3}\widetilde{m}_c^3 + c_4\widetilde{N}_t^{-4}\widetilde{m}_c^4 \\ &+ c_5\widetilde{N}_t + c_6\widetilde{m}_c + c_7\widetilde{N}_t^{-1}\widetilde{m}_c^2 + c_8\widetilde{N}_t^{-2}\widetilde{m}_c^3 + c_9\widetilde{N}_t^2 \\ &+ c_{10}\widetilde{N}_t\widetilde{m}_c + c_{11}\widetilde{m}_c^2 + c_{12}\widetilde{N}_t^3 + c_{13}\widetilde{N}_t^2\widetilde{m}_c + c_{14}\widetilde{N}_t^4, \end{aligned} \tag{4.2}$$

where

$$\begin{aligned} A(\widetilde{N}_t, \widetilde{m}_c) &= a_0\widetilde{N}_t + a_1\widetilde{m}_c + a_2\widetilde{N}_t^2 + a_3\widetilde{N}_t\widetilde{m}_c + a_4\widetilde{N}_t^3 \\ &+ a_5\widetilde{N}_t^2\widetilde{m}_c + a_6\widetilde{N}_t^4 + a_7\widetilde{N}_t^3\widetilde{m}_c + a_8\widetilde{N}_t^5 + a_9\widetilde{N}_t^4\widetilde{m}_c, \end{aligned} \tag{4.3}$$

$$\begin{aligned} B(\widetilde{N}_t, \widetilde{m}_c) &= b_0\widetilde{N}_t + b_1\widetilde{m}_c + b_2\widetilde{N}_t^2 + b_3\widetilde{N}_t\widetilde{m}_c + b_4\widetilde{N}_t^3 \\ &+ b_5\widetilde{N}_t^2\widetilde{m}_c + b_6\widetilde{N}_t^4 + b_7\widetilde{N}_t^3\widetilde{m}_c + b_8\widetilde{N}_t^5 + b_9\widetilde{N}_t^4\widetilde{m}_c, \end{aligned} \tag{4.4}$$

a_i, b_j, c_k are coefficients found by solving *non-linear regression* problem on sampled compressor map data, and $\widetilde{N}_t, \widetilde{m}_c$ are corrected rotational velocity and corrected air mass flow for compressor inlet conditions p_{1c}, T_{1c} different from compressor map reference conditions p_{ref}, T_{ref}, defined as

$$\widetilde{N}_t = \sqrt{\frac{T_{ref}}{T_{1c}}} N_t, \tag{4.5}$$

$$\widetilde{m}_c = \frac{p_{1c}}{p_{ref}} \sqrt{\frac{T_{1c}}{T_{ref}}} m_c. \tag{4.6}$$

We verified that polynomial $B(\widetilde{N}_t, \widetilde{m}_c) \neq 0$ for each tuple $(\widetilde{N}_t, \widetilde{m}_c)$ in eligible operating condition range. Modeled $\Pi_c(\widetilde{N}_t, \widetilde{m}_c)$ for clean and fouled compressor maps is included in Figure 4.1, $\eta_c(\widetilde{N}_t, \widetilde{m}_c)$ is included in Figure 4.2.

4.1.2 Efficiency Drop Prediction

At the moment, we are not able to express the fouling effect using an analytical expression. The only asset we have is the fouled-clean measured compressor map. Thus, we need to extract the useful information only from the map data available. Sinha (2013) explains this process is called *Data Mining* and that a concept of regression is used for this purpose. Regression analysis is composed of four different stages: identification of dependent and independent variables; identification of the nature of the relationship among the variables; computation of regression equation; and error analysis.

We model pressure drop as Second-order Multivariate Polynomial

$$\begin{aligned}\Delta\Pi_{c_i}(\Pi_c, m_c, N_t) &= \beta_0 + \beta_1\Pi_{c_i} + \beta_2 N_{t_i} + \beta_3 m_{c_i} + \beta_{11}\Pi_{c_i}^2 \\ &+ \beta_{12}\Pi_{c_i}N_{t_i} + \beta_{13}\Pi_{c_i}m_{c_i} + \beta_{22}N_{t_i}^2 + \beta_{23}N_{t_i}m_{c_i} \\ &+ \beta_{33}m_{c_i}^2,\end{aligned} \tag{4.7}$$

where $\Delta\Pi_{c_i}$ is the ith observed pressure drop computed as

$$\Delta\Pi_{c_i} = \Pi_{c_{clean}}(N_{t_{fouled}}, m_{c_{fouled}}) - \Pi_{c_{fouled}}, \tag{4.8}$$

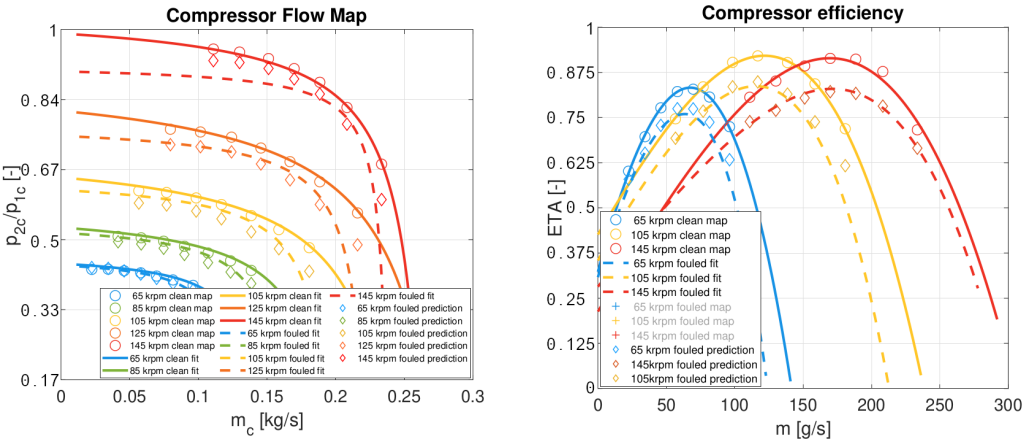
$\Pi_{c_{clean}}(N_{t_{fouled}}, m_{c_{fouled}})$ is a model identified on the clean compressor map sampled in fouled map point tuple $(N_{t_{fouled}}, m_{c_{fouled}})$, $\Pi_{c_{fouled}}$ is a measured pressure ratio taken from the fouled map, β_j are regression parameters, and variables with subscript i in (4.7) represent the map operating point at which the pressure ratio drop is investigated. We have chosen a parabolic relationship because of the "eyebrow" shape visible in Figure 4.1 and 4.2.

In Figure 4.3a we depict normalized estimated fouled Π_c compared with fitted $\Pi_c(N_t, m_c)$ on the fouled map data. We apply same steps to produce the fouled efficiency prediction algorithm; therefore, we do not repeat the procedure. An illustration of normalized estimated fouled η_c compared with fitted $\eta_c(N_t, m_c)$, on the fouled map, is depicted in Figure 4.3b.

From the Figures 4.3a and 4.3b we can infer that the accuracy of the prediction model is reasonable. To improve the algorithm robustness, a more extensive training set is required. The largest error is produced at the clean map choke line, which is decreased in the fouled compressor. Besides, our goal is to develop a compressor map model that resembles a physics-based fault for any given clean compressor map and allows us to create and verify performance drop-related health indicators, since emulating the fault at the test-stand is time and cost consuming.

4.1.3 Efficiency Drop Prediction Algorithm Summary

We present the summary of our degraded compressor map prediction algorithm. The algorithm is trained as follows:



(a) A normalized pressure ratio as a function of turbocharger rotational velocity and mass air flow for an example turbocharger's fouled and clean map. Diamond marker points are estimated $\Pi_{c_{fouled}}$ points from the given clean compressor map; dashed line is the modeled Π_c identified on acquired fouled compressor map.

(b) A normalized isentropic efficiency as a function of turbocharger rotational velocity and mass air flow for an example turbocharger's fouled and clean map. Diamond marker points are estimated $\eta_{c_{fouled}}$ points from the given clean compressor map; dashed line is the modeled η_c identified on acquired fouled compressor map.

Figure 4.3: A comparison of fouled Π_c, η_c—predicted and fitted, on the fouled map.

1. receive compressor map measurements of both, clean and fouled turbochargers. The fouled compressors come from the *real* operation,
2. find the pressure ratio $\Pi_c(\widetilde{N}_t, \widetilde{m}_c)$ and efficiency $\eta_c(\widetilde{N}_t, \widetilde{m}_c)$ models of the clean compressor map, as defined in (4.1),
3. resample clean compressor map's pressure ratio and efficiency to the fouled map operating point tuple (N_c, m_c) and compute the observed pressure ratio/efficiency drop, as defined in (4.8),
4. solve second-order multivariate polynomial regression to get coefficients β_j of second-order multivariate polynomial that predicts pressure ratio and efficiency drops $\Delta\Pi_c, \Delta\eta_c$.

The algorithm is then applied as follows:

1. get a discrete, clean compressor map with unknown efficiency drop pattern,
2. apply the pressure ratio drop $\Delta\Pi_c$ and efficiency drop $\Delta\eta_c$ to the given map at its original operating point tuples (N_c, m_c).

4.2 Engine Simulation Environment

To design and validate the algorithms we have a calibrated engine model, coupled with a controller, available in the Simulink environment. The architecture of the Simulink model is shown in Figure 4.4. Block Driving Cycle specifies the simulation operating conditions: engine speed N_e and fuel injection quantity per stroke $\dot{q}_{inj}$. Block Diesel Engine is a turbocharged diesel control-oriented engine model, having 2.0 l engine displacement, four cylinders, and EGR and VGT actuators[1]. The parameters of the model are calibrated on the test cell-measured data. Block Engine Control is the controller ensuring that the engine stays in the operating conditions specified by the driving cycle, by commanding u_{EGR} and u_{VGT}, while optimizes fuel consumption, emissions, and driving comfort. The model and the controller are calibrated and designed in software package OnRAMP^{TM}[2]. Our developed block Turbocharger Health Monitor acts as a turbocharger health observer; it collects the actuator commands and sensor measurements and then computes health indicators.

4.2.1 Health Degradation

A fouling fault can be introduced to perturb the nominal turbocharger model[3]. The fouling prediction model from (4.7) is applied to the original compressor map. We produce ten levels of fouling by scaling of (4.7), that is

$$\Delta\Pi_{c_\alpha} = \alpha\Delta\Pi_c, \tag{4.9}$$

where $\alpha = 0.1, 0.2, ..., 1$.

[1] Such as the OM611 model illustrated in Figure 1.1

[2] GARRETT MOTION INC., 2018.

[3] The original turbocharger can provide maximum pressure ratio equal to 3 at 193 krpm. The diameter of the wheel is 52 mm.

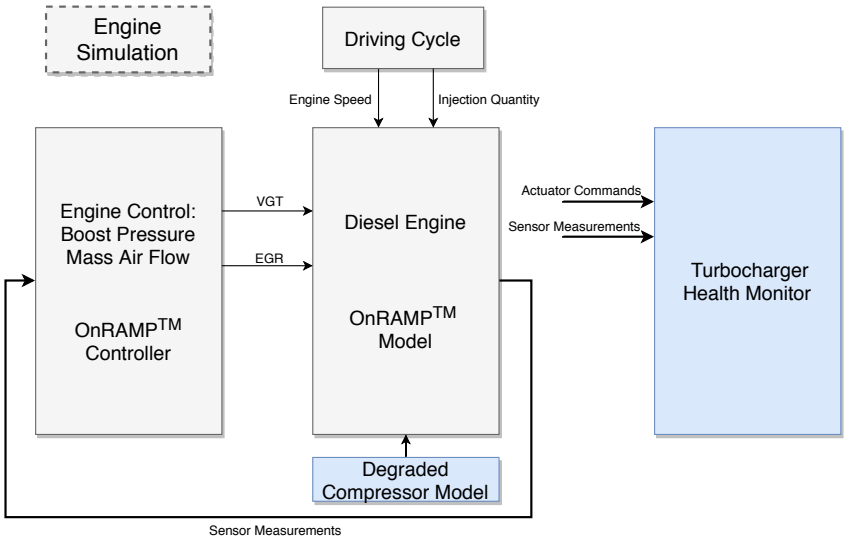


Figure 4.4: The architecture of our engine simulation environment. We specify the driving cycle which controller ensures to follow. Our health monitor observes the turbocharger by collecting actuator commands and sensor measurements.

4.2.2 Calibration Sequence

For further indicator calibration we generate engine's response by adjusting calibrating driving cycle such that the whole operating range of the engine is covered. Specifically, for each engine speed step

$$N_e \in [1000, 1250, 1500, ..., 2500, 2750, 3000] \text{ rpm} \tag{4.10}$$

we set fuel injection quantity to

$$\dot{q}_{inj} \in [10, 12.5, 15, ..., 55, 57.5, 60] \text{ mg per stroke.} \tag{4.11}$$

4.3 Health Indicator Design

We swap the original turbocharger with the degraded one (based on modified compressor map from (4.9)), thus we obtain system's response to the *faulty condition*. The controller compensates the energy loss but generates following symptoms:

1. to follow equal boost pressure setpoint, turbocharger's shaft velocity N_t increases,
2. command to open VGT u_{VGT} increases, and
3. temperature after the compressor T_{2c} increases.

We proceed to the design of health indicators for efficiency drop fault. We use system and fault dynamics knowledge while developing the indicators. Each of the symptoms, listed above, guides us to the health indicator implementation. First, increased turbocharger speed N_t can be tracked when *redundancy* information is available in the system. That is, we develop analytical expression to estimate N_t, and we compare it with the measured N_t, loosely following residual concept by Isermann (2016). Even though we have already stated that N_t measurement is not currently available in production passenger vehicles, we consider this indicator as a baseline for other indicators. Furthermore, N_t is measured in soon to be appearing vehicles with an electric motor assisted turbochargers, or in commercial vehicles. Second, in VGT, the controller produces increased u_{VGT} command to achieve desired setpoint (in other words, controller adjusts u_{VGT} to increase N_t by extracting more exhaust energy). We can use data mining method to train the state of health model that captures this relationship. Last, due to reduced efficiency, wasted energy heats the air mass, thus the compressor outlet temperature T_{2c} gets increased. We investigate a possibility to observe this symptom through approximated compressor efficiency formula.

4.3.1 Turbocharger Shaft Velocity Residual

Similar to pressure ratio and efficiency models, introduced in 4.1.1, we define a corrected turbocharger shaft velocity model $\widetilde{N}_t(\Pi_c, \widetilde{m}_c)$ (taken from Tabaček (2016))

$$\begin{aligned} \widetilde{N}_t(\Pi_c, \widetilde{m}_c) &= d_0 + d_1\widetilde{m}_c + d_2\Pi_c + d_3\widetilde{m}_c^2 + d_4\widetilde{m}_c\Pi_c + d_5\Pi_c^2 \\ &+ d_6\widetilde{m}_c^3 + d_7\widetilde{m}_c^2\Pi_c + d_8\widetilde{m}_c\Pi_c^2 + d_9\Pi_c^3, \end{aligned} \tag{4.12}$$

where Π_c is a measured compressor pressure ratio, $\widetilde{m}_c$ is a corrected measured mass air flow, and coefficients d_i are found by non-linear regression on compressor map data.

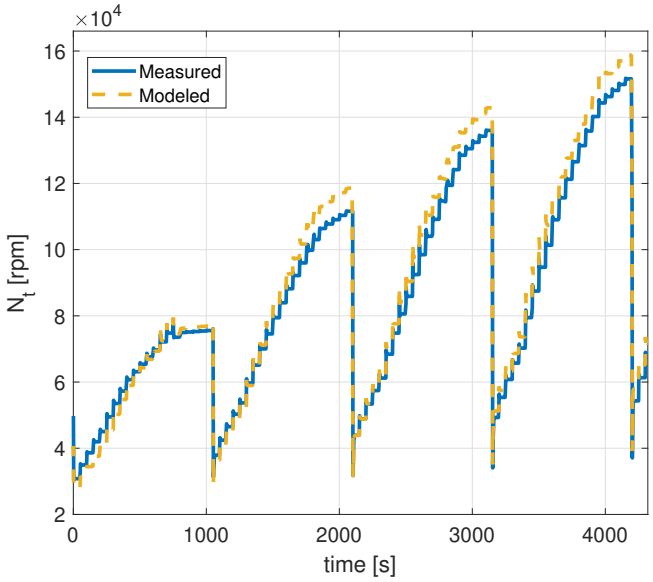


Figure 4.5: Measured and modeled turbocharger speed response to the calibration sequence, defined in 4.2.2, with step length $T = 50$ s.

In Figure 4.5 we display measured and modeled turbocharger speed response to the calibration sequence, defined in 4.2.2, with step length $T = 50$ s. The estimation of N_t is biased; the bias however varies with the operating condition. Speed transients are captured accurately by the model.

We construct the turbocharger shaft velocity residual as a difference of the measured and predicted signal

$$r_{N_t} = N_{t_{meas}} - \widetilde{N}_t(\Pi_c, \widetilde{m}_c). \tag{4.13}$$

4.3.2 State of Health Regression

Now we assume turbocharger speed is not measured. We need to replace information coming from this measurement with other source, such as actuator command u_{VGT}. Again, we use data mining technique to link information-carrying, measured variables Π_c, $\widetilde{m}_c$, and u_{VGT}. In other words, we use supervised learning to train the function that maps input variables

$\Pi_c, \widetilde{m}_c$, and u_{VGT} to output variable indicating the state of health of the compressor.

We define a variable, health indicator, SOH_k that represents health of the system with respect to efficiency drop, i.e. $SOH_1 = 1$ when original turbocharger model is present, $SOH_{0.5} = 0.5$ when 50% efficiency drop is present, and $SOH_0 = 0$ when maximum of effiency drop is applied. We generate calibration sequence (4.2.2) response to $0\%, 50\%, 100\%$ fouled compressor model and record variables $\Pi_{c_k}, \widetilde{m}_{c_k}, u_{\text{VGT}_k}$. Then we build a set of equations

$$\begin{bmatrix} SOH_1 \\ SOH_{0.5} \\ SOH_0 \end{bmatrix} = \begin{bmatrix} P_{SOH}(\Pi_{c_1}, \widetilde{m}_{c_1}, u_{\text{VGT}_1}) \\ P_{SOH}(\Pi_{c_{0.5}}, \widetilde{m}_{c_{0.5}}, u_{\text{VGT}_{0.5}}) \\ P_{SOH}(\Pi_{c_0}, \widetilde{m}_{c_0}, u_{\text{VGT}_0}) \end{bmatrix} \tag{4.14}$$

and solve fourth-order multivariate polynomial regression problem to get coefficients of fourth-order multivariate polynomial $P_{SOH}(\Pi_c, \widetilde{m}_c, u_{\text{VGT}})$. The state of health indicator is then computed, from available on-board measurements and VGT command, as

$$SOH = P_{SOH}(\Pi_c, \widetilde{m}_c, u_{\text{VGT}}). \tag{4.15}$$

4.3.3 Compressor Efficiency Computation

A difference between the healthy and degraded compressor is the increased compressor outlet temperature T_{2c} when a compressor efficiency drop is present. A relationship that links the temperature and the health of the turbocharger is the efficiency formula itself. We can ask ourselves why do we not compute efficiency directly when we want to observe the efficiency drop. The International Council on Combustion Engines (2007) answers the real turbocharger efficiency can not be reconstructed on the engine. It requires controlled laboratory conditions where straight pipes are fitted, a high-quality set of sensors is equipped, and steady air flow is produced. None of these occurs in the real turbocharged running engine. However, our goal is to observe a difference from the nominal condition, and an approximated efficiency formula can be handy for this task; next, we will introduce one.

The symptom of reduced efficiency is increased compressor outlet temperature T_{2c}, thus we must begin with estimation of this temperature from the available mass air flow measurement m_c and intercooler outlet temperature T_{CAC}. We use the heat exchanger model, introduced by Tabaček (2016)

$$T_{CAC} - T_{2c} = \frac{e_0 + e_1 T_{2c}}{e_2 + e_3 m_c}, \tag{4.16}$$

where coefficients e_i are identified by regression on the generated data. T_{2c} prediction is accurate when the mass air flow m_c through the intercooler CAC is high enough, intercooler's efficiency decreases, and allows the flow to "carry" the information about increased temperature before the intercooler.

Afterward, we compute approximated isentropic efficiency, according to Isermann (2016), as

$$\eta_c(\Pi_c, T_{2c}, T_{1c}) = \frac{\Pi_c^{\frac{\kappa-1}{\kappa}} - 1}{\frac{T_{2c}}{T_{1c}} - 1}, \tag{4.17}$$

where $\kappa \approx 1.4$ is the air specific heat ratio.

We expect that equation (4.17) may serve as a health indicator; during the high-load operations, the same pressure ratio is accompanied by higher outlet temperature which can be observed through the temperature after the intercooler measurement. We note that the flow thermodynamic condition must be in steady-state.

To conclude, we do not expect that the computed $\eta_c(\Pi_c, T_{2c}, T_{1c})$ will match the true compressor efficiency η_c but for the same engine operating conditions, computed $\eta_c(\Pi_c, T_{2c}, T_{1c})$ will deviate when the efficiency drop is present.

Sensitivity Analysis

In Figure 4.6, we illustrate computed $\eta_c(\Pi_c, T_{2c}, T_{1c})$, perturbed in T_{2c}, in steady-state. If we want to observe a difference of 5% in $\eta_c(\Pi_c, T_{2c}, T_{1c})$, we must observe the difference of 5 K in T_{2c}, too.

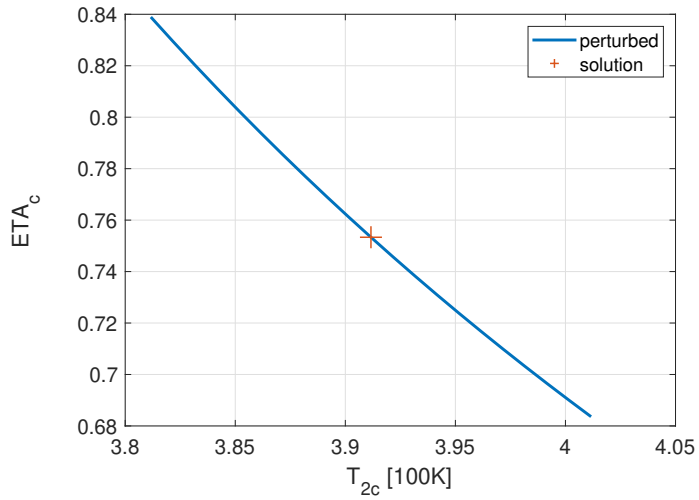


Figure 4.6: Computed compressor efficiency, from (4.17), T_{2c} perturbed about the real value by ± 10 K. Steady-state flow conditions, $\Pi_c = 2.2$.

5 Health Monitor Implementation and Evaluation

In the preceding chapters, we converged from turbocharger failure mode analysis to the design of health indicators based on the physics of failure. In this chapter, we will discuss the implementation of the indicators into the engine environment. Health indicators must be processed and filtered before we make any decision. To verify indicators' performance, we execute an emission driving cycle simulation with fault injection. We offer a short PHM analysis, as well.

5.1 Health Indicator Processing

The indicators, described in sections 4.3.1, 4.3.2, and 4.3.3, are functions of measurements and commands with one output, the indicator. The functions are evaluated at ECU clock ticks when enough processing power is available. Collected variables from the ECU are

$$\mathbf{z} = \begin{bmatrix} N_t, & m_c, & \Pi_c, & T_{CAC}, & u_{VGT} \end{bmatrix} \tag{5.1}$$

and the produced raw indicators are[1]

$$\mathbf{y} = \begin{bmatrix} r_{N_t}, & SOH, & \eta_c \end{bmatrix}. \tag{5.2}$$

The architecture of the health monitor is illustrated in Figure 5.1. After obtaining vector $\mathbf{y}$ we check if the value is feasible, resample it to common

[1]For the sake of brevity, we drop using $\eta_c(\Pi_c, T_{2c}, T_{1c})$ for computed approximated compressor efficiency and use only shortened η_c in the rest of this thesis.

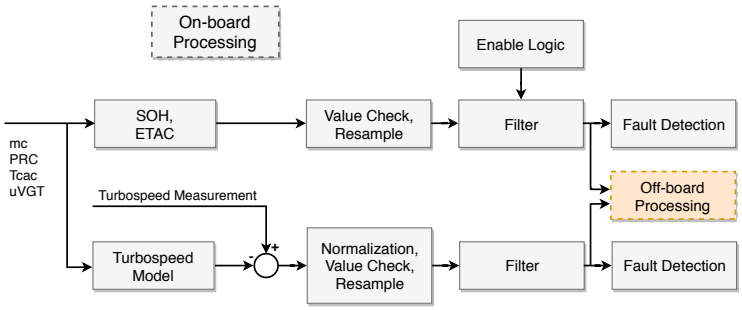


Figure 5.1: A structure of the on-board part of the turbocharger health monitor.

indicator sampling time T_s, and apply low-pass filtering. Normalization and Enable Logic actions make sure the health monitoring is not affected by operating condition variation. Decision logic is applied only after the health indicator is processed.

From now on, we consider the filtered value to be the health indicator as only after processing it becomes a standalone value that holds information about the system's health. We mark filtered values with the bar,

$$\bar{\mathbf{y}} = \left[\bar{r}_{N_t},\ \overline{SOH},\ \bar{\eta}_c\right]. \tag{5.3}$$

5.1.1 Value Check & Resample

Purpose of value check is to validate that incoming data y are within predefined limits. It also serves as a protection to the external faults.

Raw indicator data are resampled to common sampling time $T_s = 1$ s. This ensures that

1. data incoming to the filter input are uniformly spaced, independent of ECU sampling time. Filter calibration and post-processing rely on the equal sampling time in the data;
2. indicator data size is compressed. Dynamics of the health degradation is very slow compared to engine dynamics, thus we can decrease memory requirements of the health monitor by increasing sampling time.

5.1.2 Normalization & Enable Logic

We normalize generated turbocharger shaft velocity residual, as

$$r_{N_{t,n}} = \frac{r_{N_t} - r_{N_{t,0}}(N_e, \dot{q}_{inj})}{s}, \tag{5.4}$$

where $r_{N_{t,0}}(N_e, \dot{q}_{inj})$ is the nominal, operating condition-dependent residual bias obtained from the calibration (Figure 4.5), and s is the residual sample standard deviation.

For indicators η_c and SOH we propose a concept of enabling; the indicators are computed only when conditions suitable for indicator computation are met. We have already mentioned before that efficiency η_c calculation requires steady-state flow. This is rather stringent constraint as the real steady-state can not be achieved due to the flow pulsation. In addition, engine is equipped with high-inertia temperature sensors, having time constant of several seconds. We must search for conditions that do not limit indicator processing too much but also that minimize false alarm rate. However, the "steady-state" is not the only constraint; a minimum amount of power at the compressor must be present, too. That is, when the sensitivity from fault-affected variable to the indicator is high, such as the η_c sensitivity on T_{2c} illustrated in Figure 4.6.

A simple enable logic, using the compressor power P_{c_k} evaluated at sample k, can be implemented as

$$\text{enable}_k = P_{c_k} > P_{c,\min} \cap |P_{c_k} - P_{c_{k-1}}| < P_{c,\text{diffmax}}. \tag{5.5}$$

5.1.3 Filtering & Fault Detection

A raw indicator y_k, observed at sample k, can be viewed as the individual observation from a random process. There are few control schemes available that can detect a change in the process, for example, Shewhart, Cumulative Sum (CUSUM), or Exponentially Weighted Moving Average (EWMA). Lucas and Saccucci (1990) describes EWMA properties and shows EWMA to be useful for detecting small shifts in the mean of a process, it is easy to interpret

and implement, and its control limits and Average Run Length (ARL) can be analytically computed.

The EWMA is based on the statistic

$$\overline{y}_k = \lambda y_k + (1-\lambda)\overline{y}_{k-1}, \qquad 0 < \lambda \leq 1, \tag{5.6}$$

together with Upper Control Limit (UCL) and Lower Control Limit (LCL). The process is considered out of control whenever $\overline{y}_k$ falls outside the range of UCL and LCL. When the change in the process occurs, the average number of samples that filter needs to trigger out-of-control, is the ARL value.

Lucas and Saccucci (1990) further explains very small values of λ allows EWMA to detect small shifts in the process at the expense of long ARL. For our application, this property is desired as our goal is to observe the performance degradation before it reaches the maximum permitted value and we do not mind longer ARL because the degradation dynamics is very slow as well. The filtered value $\overline{y}_k$ is tracked even after it crosses the threshold. The level of fault severity is of our interest, but the information when the degradation has begun is essential in PHM concept.

Filter Calibration

The filter calibration involves determination of limits {LCL,UCL}, weighting coefficient λ, and *Target* values of $\overline{\mathbf{y}}$. Target value is the value that filter is expected to hold in fault-free case. In addition, the initial value $\overline{\mathbf{y}}_{\mathbf{0}}$ is commonly set as the Target value. We set Target values as $\overline{r}_{N_{t_0}} = 0, \quad \overline{SOH}_0 = 1, \quad \overline{\eta}_{c_0} = 0.74$.

The weighting coefficient λ determines weight of the current observation sample y_k in the filtered value computation. We found an optimal value of $\lambda = 0.01$ to match our expected filter performance. Such a small λ allows small shifts in the mean of the process to impact the filter trend.

To compute EWMA limits, Lucas and Saccucci (1990) propose: when y_k are independent and identically distributed with common variance σ_y^2, the variance of the control statistic converges to $\sigma_{\overline{y}}^2 = \frac{\lambda}{2-\lambda}\sigma_y^2$. The control limits are then based on the asymptotic standard deviation of the control statistic

as LCL = Target$-L\sigma_{\overline{y}}$ and UCL = Target$+L\sigma_{\overline{y}}$. L can be found in the tables provided by Lucas and Saccucci (1990). However, the assumption that y_k are independent and identically distributed does not hold in our case. The dependency of the indicator on current operating condition is not completely eliminated by normalization and enabling (5.1.2). Therefore, we could not compute robust limits {LCL,UCL}. Instead, we find the limits empirically so that they are not crossed by the indicators $\overline{y}$ for broad range of driving cycles.

5.1.4 Health Indicator Algorithm Summary

In this subsection we offer a short summary of the design and computation of the health indicators.

Turbocharger Shaft Velocity Residual

Main source of the information about the fault is coming from increased turbocharger shaft velocity N_t. The algorithm is trained as follows:

1. find an analytical model of the turbocharger shaft velocity $\widetilde{N}_\mathrm{t}(\Pi_\mathrm{c}, \widetilde{m}_\mathrm{c})$, defined in (4.12),
2. obtain nominal, operating condition-dependent residual bias $r_{N_{\mathrm{t},0}}(N_\mathrm{e}, \dot{q}_\mathrm{inj})$, from the calibration (Figure 4.5),
3. obtain sample standard deviation s of the residual cleared from the bias $r_{N_\mathrm{t}} - r_{N_{\mathrm{t},0}}$, from the response to the test cycle.

The algorithm is computed on-board as follows:

1. collect the turbocharger shaft velocity measurement $N_{\mathrm{t}_\mathrm{meas}}$,
2. estimate turbocharger shaft velocity from the pressure, temperature, and mass air flow measurements as $\widetilde{N}_\mathrm{t}(\Pi_\mathrm{c}, \widetilde{m}_\mathrm{c})$,
3. compute the raw residual $r_{N_\mathrm{t}} = N_{\mathrm{t}_\mathrm{meas}} - \widetilde{N}_\mathrm{t}(\Pi_\mathrm{c}, \widetilde{m}_\mathrm{c})$,
4. apply normalization, defined in (5.4),
5. check the value of normalized residual $r_{N_{\mathrm{t,n}}}$, resample to common sampling time T_s, and pass to EWMA filter.

State of Health Regression

Main source of the information about the fault is coming from increased VGT actuator command u_{VGT}. The algorithm is trained as follows:

1. create a set of degraded compressor maps, by applying fouling level $\alpha = 0\%, 50\%, 100\%$ in (4.9),
2. record the response of variables $\Pi_{c_\alpha}, \widetilde{m}_{c_\alpha}, u_{\text{VGT}_\alpha}$ to the calibration sequence (4.2.2),
3. solve fourth-order multivariate polynomial regression to get coefficients of fourth-order multivariate polynomial $P_{SOH}(\Pi_c, \widetilde{m}_c, u_{\text{VGT}})$, defined in (4.14).

The algorithm is computed on-board as follows:

1. collect the measurements $\Pi_c, \widetilde{m}_c$ and commanded input u_{VGT},
2. compute the raw state of health indicator $SOH = P_{SOH}(\Pi_c, \widetilde{m}_c, u_{\text{VGT}})$,
3. check the value of the indicator, resample to common sampling time T_s, and when enabling conditions are met (defined in (5.5)), pass the indicator value to EWMA filter.

Compressor Efficiency Computation

Main source of the information about the fault is coming from increased compressor outlet temperature T_{2c}. The algorithm is trained as follows:

1. train the CAC model, defined in (4.16), on the calibration data,
2. verify, from the sensitivity analysis in Figure 4.6, what is the desired accuracy of T_{2c} estimation and if it is feasible to achieve.

The algorithm is computed on-board as follows:

1. collect the measurements $\Pi_c, T_{1c}, m_c, T_{\text{CAC}}$,
2. estimate compressor outlet temperature T_{2c} from the charge air cooler model, defined in (4.16),
3. compute the raw approximated compressor efficiency $\eta_c(\Pi_c, T_{2c}, T_{1c})$, defined in (4.17),

4. check the value of the indicator, resample to common sampling time T_s, and when enabling conditions are met (defined in (5.5)), pass the indicator value to EWMA filter.

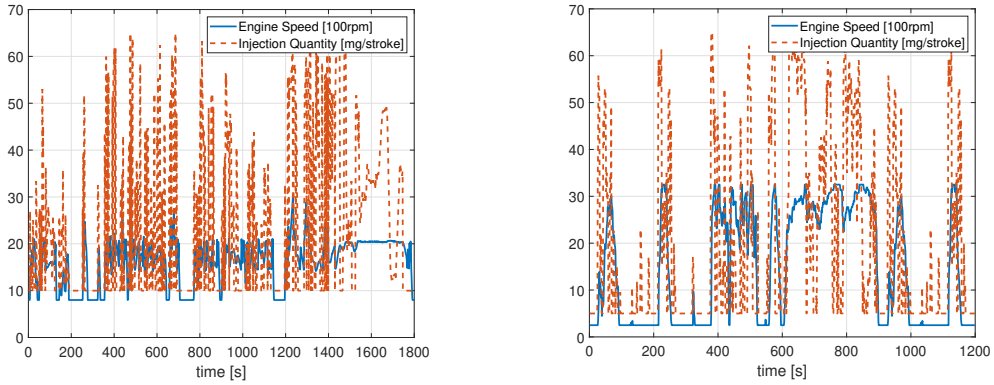
5.2 Monitor Performance Evaluation

In this section, we will validate our health indicators in the simulation. Simulated turbocharged engine model, with the boost pressure controller, is identified on the real, measured data. We emulate the fault by replacing the original compressor model with the faulty one. This faulty compressor model is produced by applying the efficiency drop algorithm, described in 4.1.2, on the original compressor model. Then we evaluate the performance of the health monitor by running a series of emission test cycles on the engine. We note that these driving cycles are different from the training driving cycle which we created to calibrate the health indicators. The simulations are performed in the Simulink^{TM} environment. The block structure of the engine model is displayed in appendix Figure 5.7, the block structure of our health monitor is displayed in appendix Figure 5.8, and the overview of the Simulink^{TM} model is displayed in appendix Figure 5.9.

5.2.1 Test Cycle Description

Emission test cycles we have selected for testing are World Harmonized Transient Cycle (WHTC)[2] and Federal Test Procedure (FTP) Transient Cycle[3]. WHTC is issued worldwide for heavy-duty engines, FTP is issued by the United States for heavy-duty vehicles—the test resembles driving in New York and Los Angeles. Both tests are characterized by multiple transients to high-load regions, which is rather the case for heavy-duty vehicles, but they can serve as a baseline for worst case passenger vehicle driving as well. Time series of variables $\{N_e, \dot{q}_{inj}\}$ defining WHTC is shown in Figure 5.2a, FTP is shown in Figure 5.2b.

[2] DieselNet, 2007.
[3] DieselNet, 1999.



(a) WHTC emission test cycle.

(b) FTP emission test cycle.

Figure 5.2: Emission test cycles, issued by legislative agencies, used for engine driving cycle simulation and health monitor validation.

5.2.2 Basic Test

First test serves to verify crucial indicators' property, if the indicator statistic deviates when the fault is present. We perform standard WHTC and FTP tests with the maximum efficiency drop fault injected in the half-time of the cycle. The response of the indicators $\overline{\mathbf{y}}$ is displayed in Figure 5.3. Turbocharger speed residual indicator $\overline{r}_{N_t}$, which uses normalization so the indicator is enabled for the whole duration, crosses the threshold UCL in < 20 s in both tests. On the other hand, indicators $\overline{SOH}$ and $\overline{\eta}_c$ uses operating condition-dependent enabling which leads to higher run length until the threshold LCL is crossed and degradation is detected. In the beginning, we can see a delay between cycle start and first filter enable. When the fault is injected, the run length is in order of 100 s for these indicators. However, we notice that from time 1500 s in WHTC run, where favored operating conditions are met, a significant indicator trend is visible. Apparently, this basic test suggests that the proposed indicators are applicable for turbocharger fault detection.

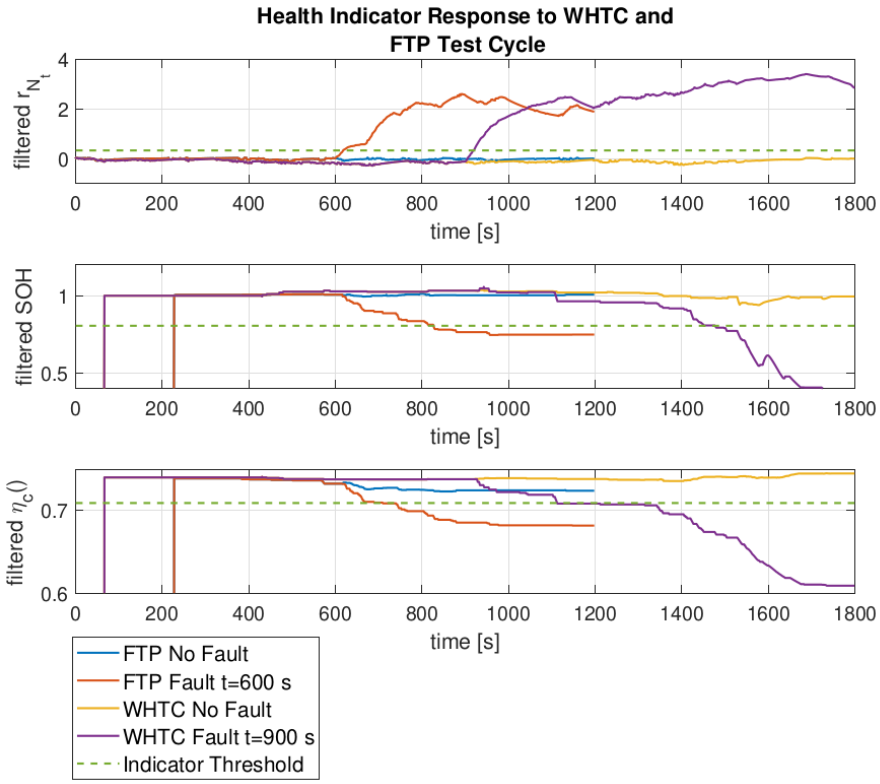


Figure 5.3: WHTC and FTP test cycle run. The indicators follow the expected statistic when no fault is present. At half-time of the cycle a maximum efficiency drop fault is applied. The indicators deviates from its expected value when there is the fault present in the turbocharger.

5.2.3 Continuous Degradation

Second test involves multiple cycle duration and continuous application of the efficiency drop. WHTC and FTP cycles are repeated five times, but every cycle is slightly different. Operating condition $\{N_e, \dot{q}_{inj}\}$ is perturbed around the nominal condition with zero-mean normal distributed noise.

The test starts with the clean compressor map. After completing first cycle, the degradation starts at a random time during the duration of second cycle. Then, the duration of individual levels of efficiency drop varies randomly, until the maximum of efficiency drop is applied and kept until the end of the

simulation. Response of the health indicators to this scenario is displayed in Figure 5.4. We make two main observations:

1. the health indicators are sensitive enough so even the smallest change in the compressor model (efficiency drop) affects the filter statistic, and
2. residual indicator $\bar{r}_{N_t}$ firstly decreases and increases only afterwards (as in the first test, Figure 5.3). The reason is that real turbocharger speed $N_{t_{meas}}$ increases, due to efficiency drop, reaching the modeled value, and then rises further (residual, before calibration, is shown in Figure 4.5). Therefore, the filtered residual reaches negative values in the beginning, so we implement bottom control limit—LCL.

5.2.4 Approach to Prognostics

In the last test, we assess the current unknown health of the vehicle by predicting its Remaining Useful Life (RUL). That is, we are interested in estimate how long our vehicle will operate until the maximum fault is present. We look back at the prognostics models overview, in Table 2.1, where we see few possibilities how to perform RUL estimation.

One option is to fit the health indicator with a model that describes the evolution of the faulty indicator and then extrapolate the model to see when it reaches the threshold defining End of Life (EOL). In our turbocharger application, the health indicator value is a function of the fault and the operating condition. This would require a set of models and thresholds, each suitable for different driving cycle.

Second option to predict RUL is to combine the statistics, current indicator value, and the information about similar vehicles—the fleet. This leads us to a prognostics solution, inspired by *Similarity* model introduced by The MathWorks, Inc. (2018), where RUL estimation is based on known behavior of similar machines. Specifically, a dataset of vehicles with indicator *run-to-failure* history is available. That is, the history of the indicators from the start of degradation to the vehicle EOL. When we query a RUL for the vehicle with unknown health status, we search for k vehicles in the dataset that have

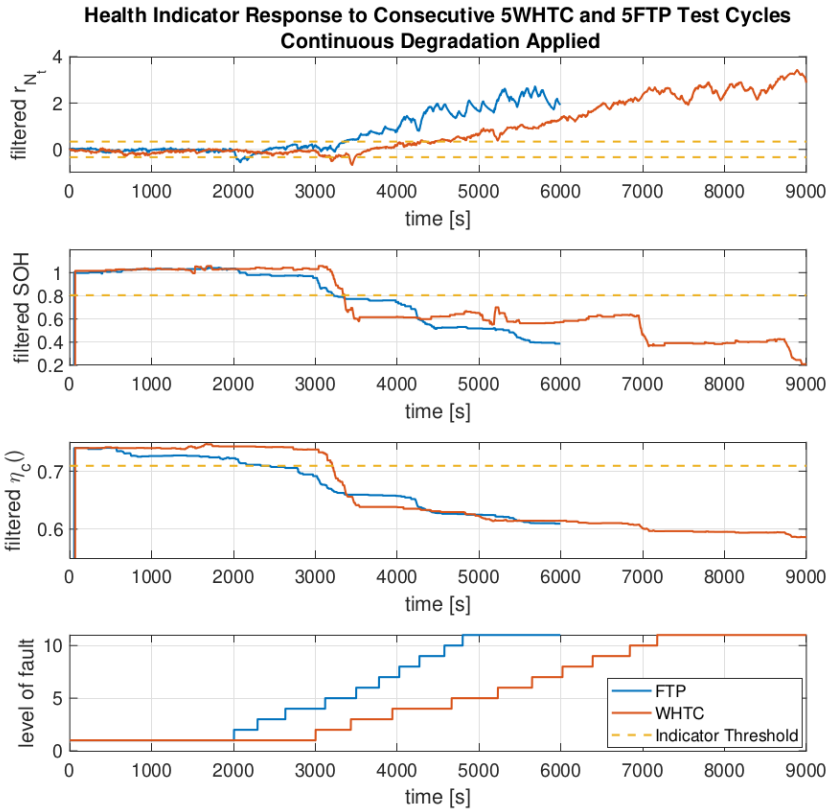


Figure 5.4: Five times consecutive WHTC and FTP test cycle run with additive random operating condition variation. Level of fault indicates current applied efficiency drop fault. Value 1 means no efficiency drop is present, 2 stands for 10% of maximum drop ($\alpha = 0.1$ in (4.9)), and so on.

similar indicator pattern and estimate the RUL from these vehicles' known future degradation.

For any prognostics model chosen, its output, the RUL, is a *random variable*. Thus, it is described by a Probability Distribution Function (PDF). In the data-driven, statistics-based prognostics solution which we discussed in subsection 2.1.1 the authors assessed the likelihood of turbocharger failure occuring before the next service visit. We adopt this idea but scale it to our test cycle scenario. Besides an estimate of RUL, it is the probability of failure within a specified time horizont that the fleet operator or vehicle owner are interested in.

We run the same test scenario as in the continuous degradation test, but after each WHTC cycle, we evaluate the probability that the maximum efficiency drop fault occurs in the next cycle. We reformulate the problem to computation of probability

$$P(RUL < T_{\text{WHTC}}), \tag{5.7}$$

where RUL is the expected Remaining Useful Life after the WHTC cycle and $T_{\text{WHTC}} = 1800$ s is the cycle duration. The steps of our prognostics algorithm are:

1. check that the indicators crossed the thresholds and the degradation process has begun,
2. transmit the time series of the indicator to Off-Board Processing part (see architecture of the indicators in Figure 5.1). It can be a remote server that collects data from the fleet of the vehicles, has high computational power, large storage capacity, and holds run-to-failure indicator history of a set of vehicles—these are the properties not found on car's ECU,
3. fit the time series of the indicator with second order polynomial to obtain $\hat{y}$,
4. compute Euclidean distances $d_i = ||\hat{y} - \hat{y}_i||_2$, where $\hat{y}_i$ is the second order polynomial fit of the vehicle from set i with *known* indicator history and EOL,
5. choose k-closest vehicles by selecting k smallest distances from d_i,
6. compute individual RUL_j for these k-closest vehicles as $RUL_j = EOL_j - T_{\text{eval}}$, where T_{eval} is the time of the evaluation and EOL_j is the jth vehicle's associated EOL. RUL_j are samples from a probability distribution RUL_k,
7. estimate the RUL for the vehicle with unknown health status as median of the RUL samples, i.e. $\widehat{RUL} = \text{median}(RUL_k)$,
8. compute probability of EOL event in the next cycle as defined in (5.7). This step requires finding the PDF of RUL_k distribution by fitting a normal distribution on RUL_j data by estimating distribution's mean μ and variance σ^2 with minimum variance unbiased estimator. At last, the probability of EOL in the next cycle is computed from Cumulative Distribution Function (CDF) of distribution RUL_k as

$$\Phi_{\mu,\sigma^2}(T_{\text{WHTC}}) = P(RUL < T_{\text{WHTC}}). \tag{5.8}$$

We apply proposed prognostics algorithm in the five times WHTC test, specified in 5.2.3. Only WHTC driving cycle and residual indicator $\bar{r}_{N_t}$ are processed, but we have verified that the algorithm has same performance for each combination of the cycle and the indicator. A set of $n = 200$ vehicles with run-to-failure indicator history is loaded. Algorithm searches for $k = 25$ closest neighbors.

After completing first WHTC cycle, there is no efficiency drop; thus, the threshold is not crossed and degradation is not detected. After completing second WHTC cycle, the fault detection part detects crossed LCL at time $T_{\text{fault}} = 3422$ s. We find $\hat{y}$, a second order polynomial fitted to the data $\bar{r}_{N_t}$ from time T_{fault} to time of evaluation $T_{\text{eval}} = 3600$ s. Then, we search for k-closest vehicles from the n vehicles with known degradation and computes the distribution RUL_k. Evolution of the indicator $\bar{r}_{N_t}$ and the history of selected k-closest vehicles is shown in Figure 5.5a (blue are 25-closest vehicles, cyan are 26-to-50-closest, out of total $n = 200$ vehicles.). A histogram RUL_k samples, normalized to PDF, and the Gaussian PDF fitted on the RUL_k data (estimated μ, σ^2) is shown in Figure 5.5b. Predicted RUL, the median of the distribution is $\widehat{RUL} = 3972$ s, true RUL is $RUL_{\text{true}} = 4399$ s. Both Figures 5.5a and 5.5b illustrate $\widehat{RUL}$ in yellow color and RUL_{true} in black. Finally, the probability of EOL event in the next cycle is $\Phi_{\mu,\sigma^2}(T_{\text{WHTC}}) = P(RUL < T_{\text{WHTC}}) = 0\%$.

After the third cycle, at time $T_{\text{eval}} = 5400$ s, we repeat the same procedure. The relative error of RUL prediction remains $\frac{RUL_{\text{true}} - \widehat{RUL}}{RUL_{\text{true}}} = 10\%$ as in the previous case, but in Figure 5.5c we can see that when the longer history of the indicator is available, vehicles having more similar indicator pattern are selected. We believe this is a crucial property of any prognostics algorithm. Figure 5.5d illustrates PDF of RUL_k; the probability of EOL event in the fourth cycle is $\Phi_{\mu,\sigma^2}(T_{\text{WHTC}}) = P(RUL < T_{\text{WHTC}}) = 19\%$. This is key *information* for the Health Management system within PHM concept to make informed, operation and cost-optimizing decision.

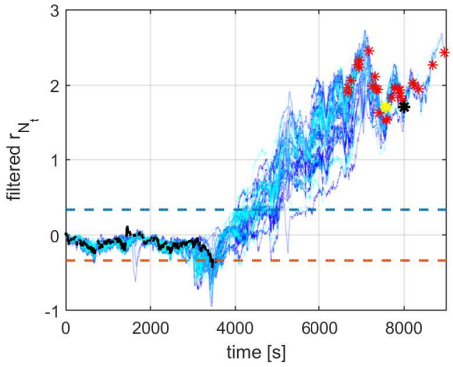
After completing fourth cycle, at time $T_{\text{eval}} = 7200$ s, the efficiency drop fault is close to its maximum. In Figure 5.5e, it is visible from the absolute value of the residual indicator $\bar{r}_{N_t}$ that it is approaching a cluster of EOL values. If we have a look at the PDF in Figure 5.5f and compute the probability of EOL

in the fifth cycle, we confirm that with 100% probability the EOL event, with respect to efficiency drop, will occur in the next $T_{\text{WHTC}} = 1800$ s.

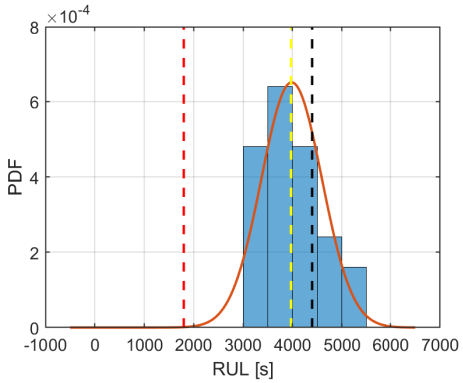
5.2.5 Monitor Performance Evaluation Summary

Our goal was to detect and track the compressor efficiency drop fault. This fault causes the compressor to deliver less power compared to the nominal condition. Driving cycles were chosen to represent official issued emission transient cycles. In the basic test, we successfully verified that the health monitor serves as a diagnostic tool—when the compressor model affected by the efficiency drop fault was injected, all the three indicators deviated from their nominal, bounded values (Figure 5.3). Then, in the continuous degradation test, we simulated compressor degradation by scaling the applied efficiency drop (Figure 5.4). Our observation confirmed that the indicators responded in the early stages of the fault development, long before the maximum efficiency drop (with the magnitude of approximately 8%) was present. Of course, we do not expect detection of the efficiency drop 1% in the real engine application, but we are confident that our monitor can capture the deviation earlier than the current solutions. This early detection is crucial in the concept of prognostics; observing the degradation as soon as possible leads to computation of RUL probability (Figure 5.5) which helps the health management to optimize the repair or exchange process and avoid unexpected downtimes.

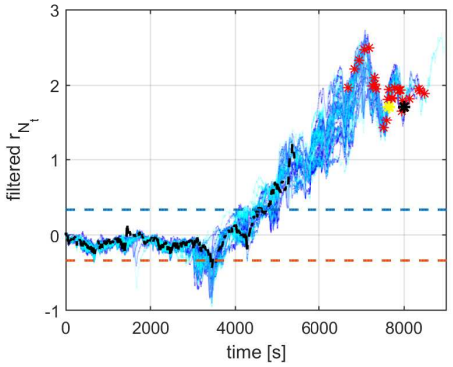
The architecture of the health monitor we developed consists of a set of three indicators. Its advantage is the ability to isolate the fault and protect the monitor against other faults. The indicators η_c and r_{N_t} rely on the measurements from the compressor side of the air path, but the *SOH* indicator takes into account also the turbine's VGT command input. Therefore, a turbine efficiency drop fault, or an uncompensated VGT actuator hysteresis will impact *only* the *SOH* indicator. For the fault detection, a concept of voting can be established: at least two out of three indicators must agree on the diagnosis (healthy or faulty compressor). When one of them does not agree, an implication can be made: for example, there is a high probability of a turbine/actuator issue when *SOH* drifts away, CAC wear when η_c deviates, or the turbocharger shaft speed sensor bias when r_{N_t} diverges.



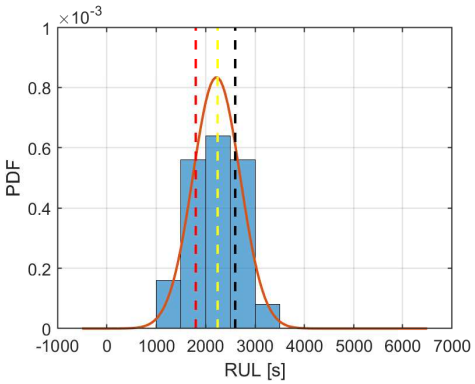
(a) History of $\bar{r}_{N_t}$ at $T_{\text{eval}} = 3600$ s. Performance degradation is detected when the indicator crosses the LCL. We can see a higher variation of selected vehicles' indicator history as only a short indicator sample is available for comparing.



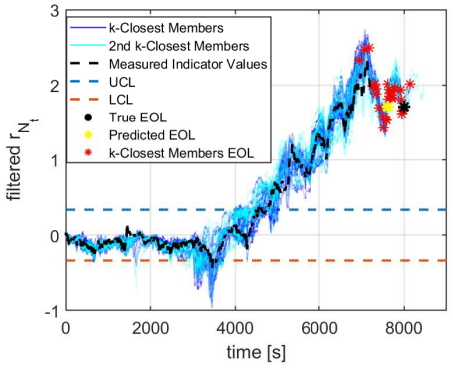
(b) PDF of RUL_k distribution sampled from available k-closest vehicles' history, at $T_{\text{eval}} = 3600$ s. $P(RUL < T_{\text{WHTC}}) = 0\%$.



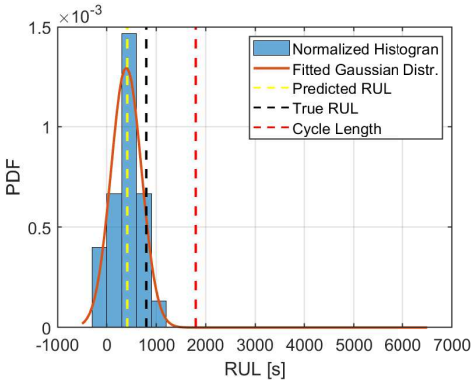
(c) History of $\bar{r}_{N_t}$ at $T_{\text{eval}} = 5400$ s. We are now able to select more similar vehicles and estimate more robust RUL when we have longer indicator history at hand.



(d) PDF of RUL_k distribution sampled from available k-closest vehicles' history, at $T_{\text{eval}} = 5400$ s. $P(RUL < T_{\text{WHTC}}) = 19\%$.



(e) History of $\bar{r}_{N_t}$ at $T_{\text{eval}} = 7200$ s. The vehicle is close to its EOL.



(f) PDF of RUL_k distribution sampled from available k-closest vehicles' history, at $T_{\text{eval}} = 7200$ s. $P(RUL < T_{\text{WHTC}}) = 100\%$. Now, the vehicle should not proceed in its operation.

Figure 5.5: Overview of our prognostics algorithm test. The idea is to assess the current unknown health of the vehicle after each of the five WHTC cycles. That is, we find an estimate of RUL and evaluate the probability that the vehicle finishes the next cycle without having maximum efficiency drop fault. The algorithm makes use of the similar vehicles' known degradation history available at off-board processing part.

Conclusion

In this thesis, we designed and implemented the health monitoring system of a typical turbocharger found in passenger vehicles with the stress on applicability in the production unit. That is, the proposed algorithms are feasible under the stringent conditions imposed by common ECUs and respect the configuration of the engine and its sensor network. This is where the main contribution of this thesis lies as other published diagnostic and prognostic methods rely on the commonly unavailable measurements or data.

We explored the prognostic models and methods reported in the literature and summarized them; none of the methods complies with available processing power and measured signals in production ECUs.

To design a turbocharger health monitoring system, we carefully identified and grouped all the possible failure modes based on the fault occurrence rate, fault cause/effect relation, and fault observability. We classified the faults into five groups: a turbocharger performance degradation, turbocharger bearing wear, FOD, actuation issues, and material fatigue. For each of the fault group we examined the possibility to observe the fault in the production vehicle and proposed a prognostic approach to monitoring the degradation, if possible. We chose the compressor efficiency drop due to oil fouling as the only fault worth of further analysis as this seems to be the only fault with an observable and predictable impact on engine's performance. We collected four clean-fouled compressor map pairs directly from the field operation and applied polynomial regression to these maps to obtain an analytical model describing the loss of turbocharge's performance based on the level of the oil fouling.

We proposed a health monitor consisting of a set of three indicators: turbocharger shaft velocity residual, state of health regression, and approxi-

mated compressor efficiency formula. These indicators were designed to capture the physics of the efficiency drop fault by either defining a model or training the indicator. The proposed health monitor implementation and performance evaluation were carried out in the Simulink^{TM} environment on the diesel engine model coupled with the controller, identified on the real data and created in the OnRAMP^{TM} software package. Running a series of official WHTC and FTP validated key prognostic properties of the proposed health monitor: it serves for onboard fault detection and isolation, provides a robustness thanks to the indicator triplet, and is ready to be embedded in the PHM framework. Also, we proposed and validated a similarity prognostic model estimating the Remaining Useful Life (RUL) based on the so-called similarity model approach. This model relies on a dataset of lifetime records of vehicles of the same type and estimates the RUL based on the most similar vehicles (with the most similar indicator history) to the vehicle at hand.

In Figure 5.6, a flow of the development of a prognostic application, as pictured by The MathWorks, Inc. (2018), is displayed. The output of this thesis is closing the "Develop Detection or Prediction Model" phase. The future steps are to deploy & integrate the algorithms into the engine's ECU. Afterward, we use the control theory's feedback principle to improve the proposed health monitor: a discovered health monitor's error accumulated from the modeling uncertainties will be corrected by recalibrating the algorithms on the new sensor data coming when the algorithms get integrated, as the Figure 5.6 illustrates.

Appendix

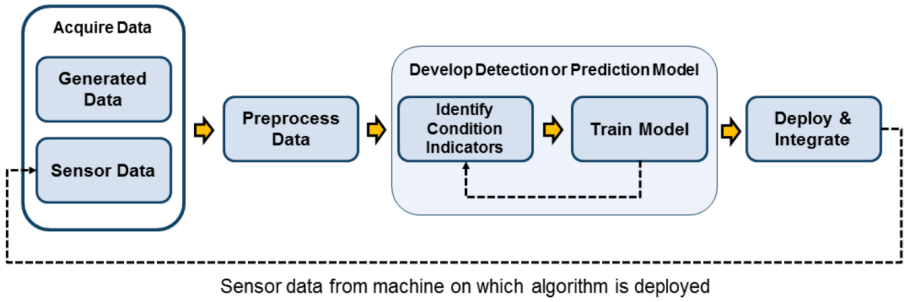


Figure 5.6: The prognostics feedback. Picture taken from The MathWorks, Inc. (2018).

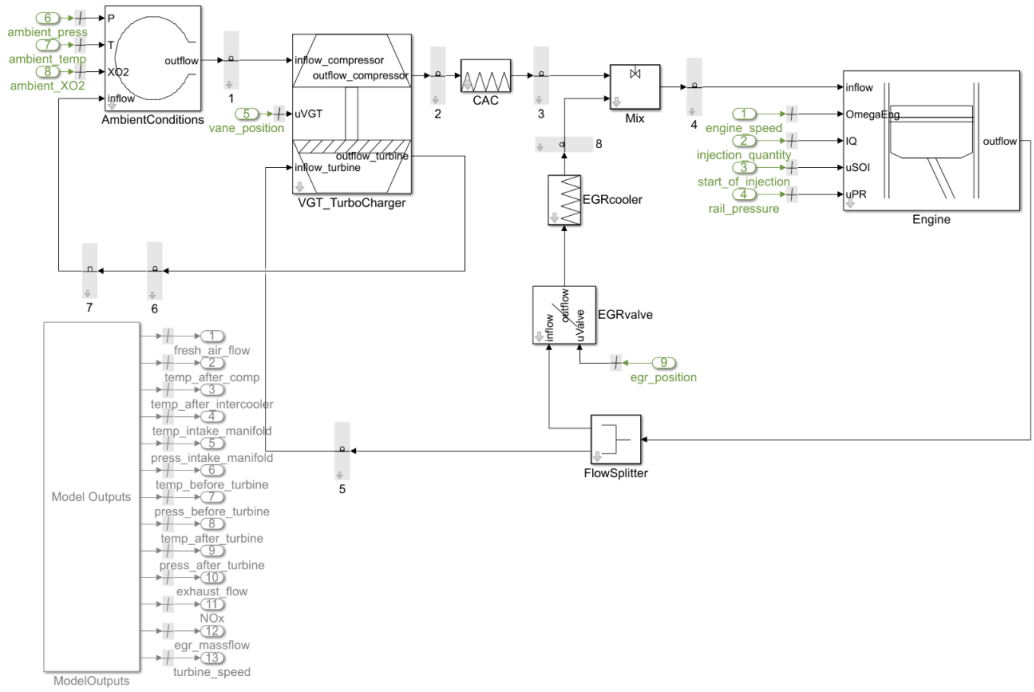


Figure 5.7: A Simulink™ environment where the OnRAMP™ Model is implemented.

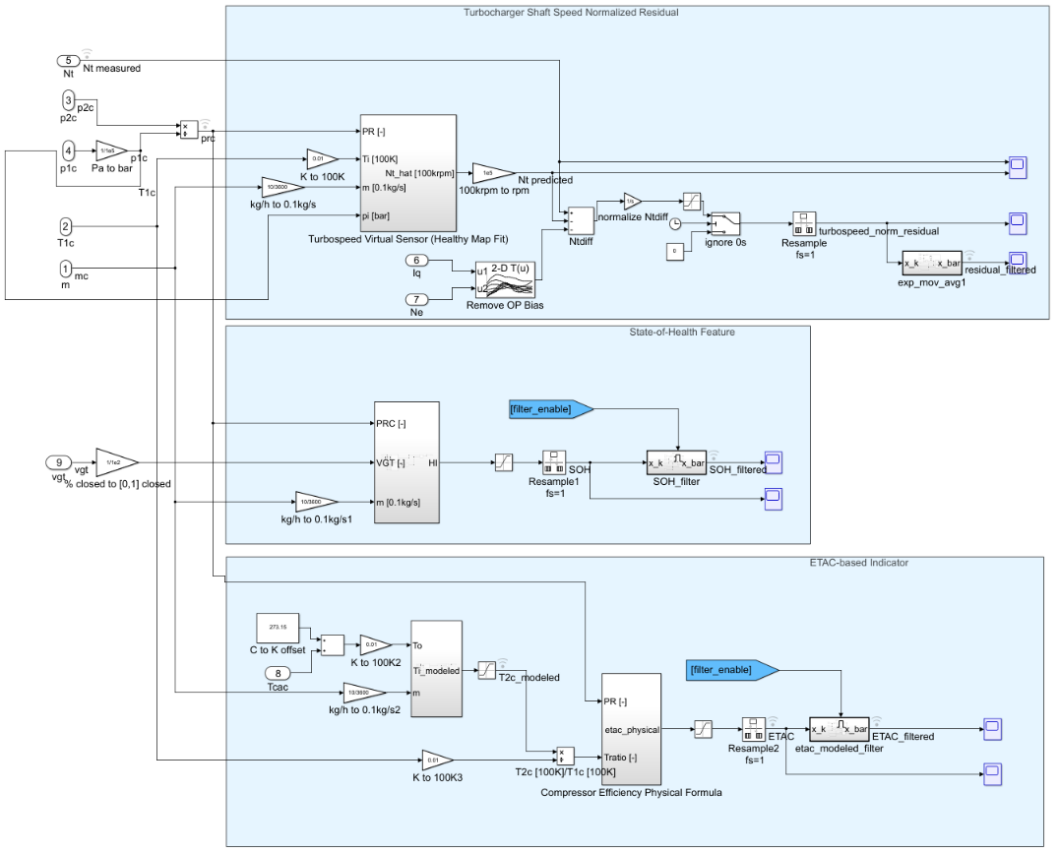


Figure 5.8: A Simulink^{TM} environment where our health monitor is implemented.

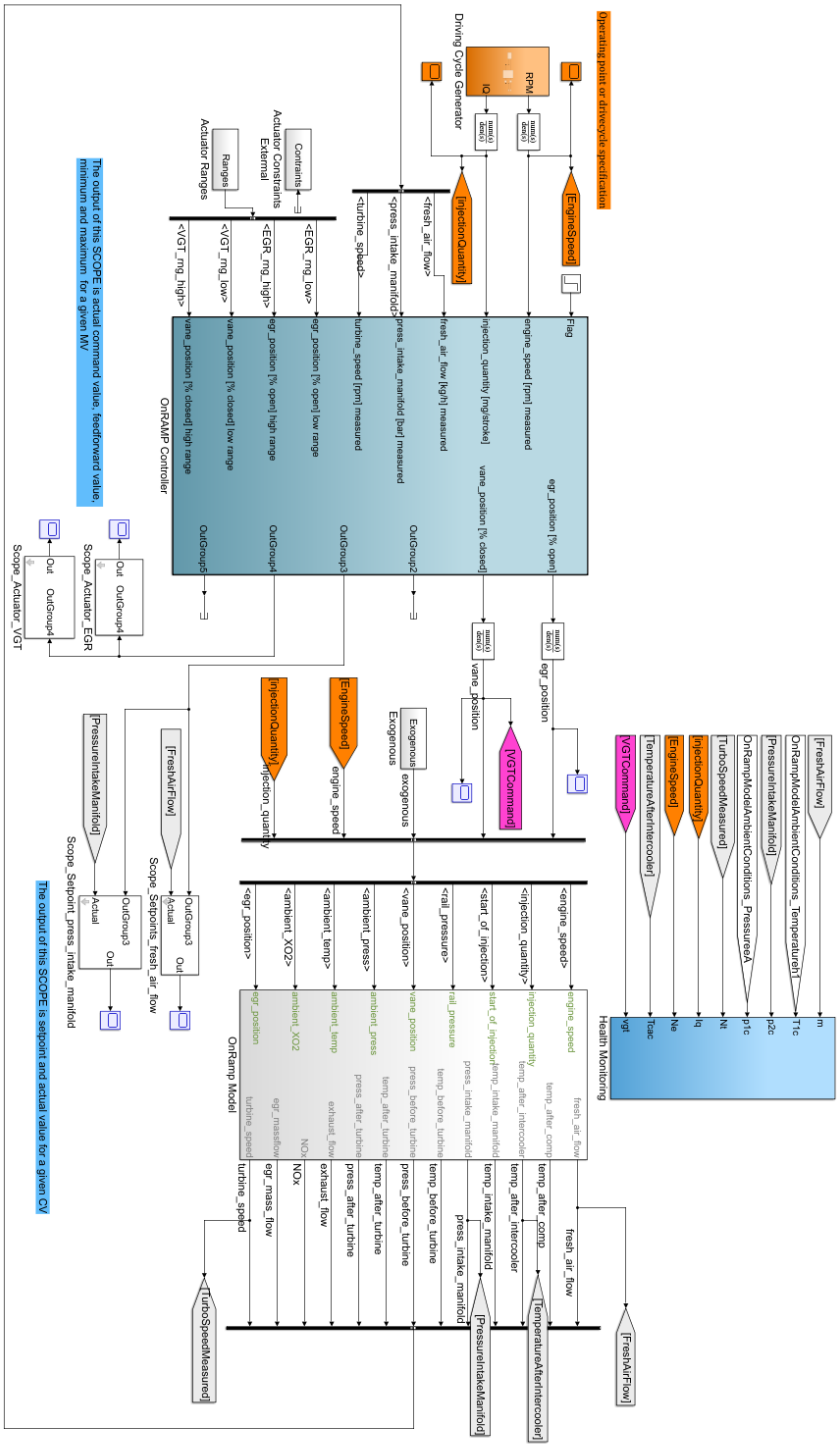


Figure 5.9: A high-level overview of the Simulink™ engine model, its controller, and our implemented health monitor.

Bibliography

Adamkiewicz, Andrzej (2012). "An analysis of cause and effect relations in diagnostic relations of marine Diesel engine turbochargers." In: *Zeszyty Naukowe/Akademia Morska w Szczecinie*, pp. 5–13 (cit. on p. 27).

ANSYS, INC. (2016). *Wearing it WELL*. URL: https://www.ansys.com/-/media/ansys/corporate/resourcelibrary/article/wearing-it-well-aa-v10-i3.pdf (cit. on pp. 31, 32).

Arbore, Luigi (2015). *Turbocompressori: gli attuatori pneumatici ed elettronici della geometria variabile (VNT)*. URL: http://www.notiziariomotoristico.com/approfondimenti/6223/turbocompressori-gli-attuatori-pneumatici-ed-elettronici-della-geometria-variabile-vnt (cit. on p. 31).

BTN Turbo (2012). *Fault Fact Sheets*. URL: http://www.btnturbo.com/spindoctor/faultFactSheets.aspx?ck=1 (cit. on p. 26).

Ceccarelli, Riccardo (Nov. 2012). *Model-based fault detection in Diesel engines air-path*. URL: https://tel.archives-ouvertes.fr/tel-00757525 (cit. on pp. 21–23).

CzechInvest (2018). *Land of Automotive*. URL: https://www.czechinvest.org/en/Key-sectors/Automotive (cit. on p. 1).

DieselNet (1999). *Heavy-Duty FTP Transient Cycle*. URL: https://www.dieselnet.com/standards/cycles/ftp_trans.php (cit. on p. 59).

DieselNet (2007). *World Harmonized Transient Cycle (WHTC)*. URL: https://www.dieselnet.com/standards/cycles/whtc.php (cit. on p. 59).

Eastwood, Mike and John Allport (2007). *Understanding Thermo-Mechanical Fatigue*. URL: http://caboturbo.nl/wp/wp-content/uploads/2012/03/HTi_08.pdf (cit. on pp. 34, 35).

Engels, Dr. Bertold (2002). *Lifetime prediction for Turbocharger Compressor Wheels - Why Use Titanium-?* URL: https://www.turbos.borgwarner.com/tools/download.aspx?t=document&r=107&d=109 (cit. on pp. 33, 37).

Feneley, Adam J, Apostolos Pesiridis, and Amin Mahmoudzadeh Andwari (2017). "Variable geometry turbocharger technologies for exhaust energy recovery and boosting-A Review." In: *Renewable and sustainable energy reviews* 71, pp. 959–975 (cit. on pp. 30, 31).

Floren, Chris (2011). *Burst & Containment: Ensuring Turbocharger Safety*. URL: https://www.turbobygarrett.com/turbobygarrett/sites/default/files/Garrett_White_Paper_02_Burst__Containment.pdf (cit. on pp. 32, 33).

Ford Motor Company (Dec. 2016). *2017 MY OBD System Operation. Summary for Gasoline Engines*. URL: http://www.fordservicecontent.com/ford_content/catalog/motorcraft/OBDSM1704.pdf (cit. on pp. 7, 32).

Ford Motor Company (July 2017). *2017 MY OBD System Operation. Summary for Diesel Engines*. URL: http://www.fordservicecontent.com/ford_content/catalog/motorcraft/DOBDSM1701.pdf (cit. on pp. 20, 21).

GARRETT MOTION INC. (2018). *OnRAMP Design Suite: Virtual Sensing, Control Design & Simulation*. URL: https://www.garrettmotion.com/connected-vehicle/oem-diagnostic-solutions/ (cit. on p. 45).

Guzzella, Lino and Christopher Onder (2009). *Introduction to modeling and control of internal combustion engine systems*. Springer Science & Business Media (cit. on pp. 9, 10, 12).

Hirano, Satoshi and Paul Decker-Brentano (2017). *Engine Oil Requirements from an OEM Point of View - Prevention of Diesel Turbocharger Compressor Deposit*. URL: http://docplayer.net/storage/65/53569211/1541075701/iVp5wipbmIP4ZZh6P-na7g/53569211.pdf (cit. on p. 27).

Holland, Steven W et al. (2010). *PHM for Automotive Manufacturing % Vehicle Applications*. URL: https://www.phmsociety.org/sites/phmsociety.org/files/Holland_GM_PHM2010_Fielded_Sys.pdf (cit. on p. 1).

Honeywell International (2017a). *Turbo Tech 101 (Basics)*. URL: https://www.garrettmotion.com/turbocharger-technology/how-a-turbo-works/basic/ (cit. on pp. 8, 13).

Honeywell International (2017b). *Turbo Tech 103 (Expert)*. URL: https://www.garrettmotion.com/turbocharger-technology/how-a-turbo-works/expert/ (cit. on p. 11).

Isermann, Rolf (2016). *Combustion engine diagnosis*. Springer (cit. on pp. 23, 47, 50).

Jääskeläinen, Hannu (Aug. 2016). *Variable Geometry Turbochargers*. URL: https://www.dieselnet.com/tech/air_turbo_vgt.php (cit. on p. 8).

Katrašnik, Tomai et al. (2003). "Improvement of the dynamic characteristic of an automotive engine by a turbocharger assisted by an electric motor." In: *Journal of engineering for gas turbines and power* 125.2, pp. 590–595 (cit. on p. 12).

Kumar, Sachin and Michael Pecht (2010). "Modeling approaches for prognostics and health management of electronics." In: *International Journal of Performability Engineering* 6.5, pp. 467–476 (cit. on pp. 15, 18).

Lucas, James M and Michael S Saccucci (1990). "Exponentially weighted moving average control schemes: properties and enhancements." In: *Technometrics* 32.1, pp. 1–12 (cit. on pp. 55–57).

Melett Ltd. (2017). *Common turbo failure – REA/SREA*. URL: https://www.melett.com/wp-content/uploads/2014/07/REASREA_EN.pdf (cit. on p. 30).

Mohammadpour, Javad, Matthew Franchek, and Karolos Grigoriadis (2012). "A survey on diagnostic methods for automotive engines." In: *International Journal of Engine Research* 13.1, pp. 41–64 (cit. on p. 19).

Munnell, Luke (Feb. 2011). *Turbo Tech 101 - Tech-Knowledge: The first step in understanding boost.* URL: http://www.superstreetonline.com/how-to/engine/impp-1103-turbo-tech-101/ (cit. on p. 10).

Nyberg, Mattias and Thomas Stutte (2004). "Model based diagnosis of the air path of an automotive diesel engine." In: *Control Engineering Practice* 12.5, pp. 513–525 (cit. on pp. 5, 6, 20).

Owen, Brian (2012). *WHY HAS MY TURBO FAILED?* URL: https://www.owendevelopments.co.uk/uploads/PDF%20files/Why-has-my-turbo-failed.pdf (cit. on pp. 25, 26).

Pastor, José V et al. (2012). "Study of turbocharger shaft motion by means of non-invasive optical techniques: Application to the behaviour analysis in turbocharger lubrication failures." In: *Mechanical Systems and Signal Processing* 32, pp. 292–305 (cit. on pp. 29, 30).

Polichronis, Dellis et al. (2013). "Turbocharger Lubrication—Lubricant Behavior and Factors That Cause Turbocharger Failure." In: *Int. J. Autom. Eng. Technol* 2.1, pp. 40–54 (cit. on pp. 27–29).

Prytz, Rune et al. (2015). "Predicting the need for vehicle compressor repairs using maintenance records and logged vehicle data." In: *Engineering applications of artificial intelligence* 41, pp. 139–150 (cit. on p. 19).

Sheng, X, DC Clay, and John Allport (2006). "Dynamics of mistuned radial turbine wheels." In: Elsevier (cit. on pp. 34, 35).

Sinha, Priyanka (2013). "Multivariate polynomial regression in data mining: methodology, problems and solutions." In: *International Journal of Scientific and Engineering Research* 4.12, pp. 962–965 (cit. on p. 42).

Tabaček, Jaroslav (May 2016). *Exhaust Gas Recirculation and Air Flow virtual sensors for internal combustion turbocharged engines*. URL: https://support.dce.felk.cvut.cz/mediawiki/images/d/dd/Dp_2016_tabacek_jaroslav.pdf (cit. on pp. 6, 41, 47, 50).

The International Council on Combustion Engines (2007). *TURBOCHARGING EFFICIENCIES - DEFINITIONS AND GUIDELINES FOR MEASUREMENT AND CALCULATION*. URL: https://www.cimac.com/cms/upload/Publication_Press/Recommendations/Recommendation_27_rev_081007.pdf (cit. on p. 49).

The MathWorks, Inc. (2018). *Designing Algorithms for Condition Monitoring and Predictive Maintenance*. URL: https://www.mathworks.com/help/pdf_doc/predmaint/predmaint_gs.pdf (cit. on pp. 17, 18, 62, 70, 71).

Uchida, Hiroshi (2006). "Transient performance prediction for turbocharging systems incorporating variable-geometry turbochargers." In: *R&D Rev Toyota CRDL* 41.3, pp. 22–8 (cit. on p. 12).

Vogl, Gregory W, Brian A Weiss, and Moneer Helu (2016). "A review of diagnostic and prognostic capabilities and best practices for manufacturing." In: *Journal of Intelligent Manufacturing*, pp. 1–17 (cit. on pp. 16–18).

Zhang, Jizhong et al. (2010). "Fault diagnosis and failure prediction by thrust load analysis for a turbocharger thrust bearing." In: *ASME Turbo Expo 2010: Power for Land, Sea, and Air*. American Society of Mechanical Engineers, pp. 491–498 (cit. on pp. 28, 29).