A Methodology for Fault Diagnosis of Diesel NO\textsubscript{x} Aftertreatment Systems

A. Pezzini, M. Canova, S. Onori, G. Rizzoni,∗ A. Soliman**

∗ The Ohio State University Center for Automotive Research, Columbus, OH 43212 USA
** The University of North Carolina at Charlotte, 9201 University City Blvd. DCH 103, Charlotte, NC 28223 USA

Abstract: Diesel engines are today considered leading candidates for the new generations of passenger vehicles due to their fuel efficiency and drivability. One of the key elements for the future acceptability is the compliance with emission standards (particularly on nitrogen oxides), which will require precise control of the aftertreatment system. Furthermore, in light of OBD-II regulations, considerable research must be devoted to the design of fault diagnosis algorithms. The definition of fault diagnosis strategies is a complex process that involves thorough studies of the system behavior in healthy and faulty conditions. Such studies can be done in multiple ways, including experimentation and mathematical modeling. In both cases, a thorough knowledge of the system components, sensors and actuators is required.

The proposed paper presents an approach to model-based fault diagnosis of Diesel NO\textsubscript{x} aftertreatment systems. The proposed methodology is based on a functional and structural analysis of the system, at the level of individual components and assemblies. This facilitates the mapping and characterization of system faults through FTA and FMEA methods, allowing for the design of control-oriented models to be used for fault detection and isolation. In this paper, the outlined approach is applied to a Lean NO\textsubscript{x} Trap system.

1. INTRODUCTION

Diesel engines are leading candidates for new generations of vehicles, due to their fuel efficiency and low-end torque characteristics enhancing drivability. One of the key elements for their market penetration is to control the emissions, particularly nitrogen oxides (NO\textsubscript{x}) and particulate matter (PM). The current US EPA Tier 2 emission regulations impose severe limits to NO\textsubscript{x} and PM emission, requiring a 65% NO\textsubscript{x} reductions for LD vehicles by 2009 and a 85% reduction for HD trucks by 2010 (Johnson [2008]).

The current state of the art shows a generalized use of Diesel Particulate Filters (DPF) to comply to PM emission standards. As for the NO\textsubscript{x} emissions, two systems are currently being studied and applied, namely Selective Catalytic Reduction (SCR) systems and Lean NO\textsubscript{x} Trap (LNT) catalysts. The current industry standards show that SCR catalysts are the technology of choice for HD application (Chi and DaCosta [2005], Willems et al. [2007], Devarakonda et al. [2008], Shost et al. [2008]). For light-duty applications, LNT are considered a viable alternative to SCR (Geckler et al. [2001], Nakagawa et al. [2004]). A basic cost analysis shows that both LNT and SCR could be utilized, with their economical feasibility depending mostly on the engine size.

The main challenges for the vehicle implementation of NO\textsubscript{x} aftertreatment lie in the design of robust control ensuring that the system operates with high conversion efficiency, regardless of variability in the operating conditions and with restrictions on the available sensors. In addition, dedicated algorithms are required to monitor the aftertreatment system for faults, in compliance with the recent OBD-II regulations (Baltus [2004]). Limited contributions have been presented to date on the subject (Siebenbrunner et al. [2008], Nebergall et al. [2005]), envisioning that consistent efforts will be devoted in the future to ensure compliance with the upcoming standards.

The paper illustrates an approach to the fault diagnosis of NO\textsubscript{x} aftertreatment systems, which from a control perspective are considered nonlinear systems with embedded feedback control. The presence of closed-loop control increases the complexity of the diagnostic problem, as its ability to compensate for errors may prevent a prompt detection of faults.

The methodology proposed is based on a preliminary functional and structural analysis, which facilitates Failure Mode and Effects Analysis (FMEA) and Fault Tree Analysis (FTA). This allows one to study the behavior of the system under healthy and faulty conditions, map the system faults and characterize their effects on the system performance. The analysis, strengthened by experimental evidences, leads to the formulation of mathematical models characterizing the system behavior, including fault modes. The math-based environment can then be used to design and test fault detection algorithms.

The paper is structured as follows. Section 2 presents a structural and functional analysis of NO\textsubscript{x} aftertreatment systems based either on LNT or SCR catalysts, in light of the current state of the art. Then, focusing only on LNT
systems, the formulation of FTA and FMEA is introduced to characterize the system faults. The final part of the paper describes the model-based fault diagnosis approach, focused on the characterization and detection of selected sensors faults and LNT parametric faults.

2. STRUCTURAL AND FUNCTIONAL ANALYSIS OF NO\textsubscript{x} AFTERTREATMENT SYSTEMS

In order to investigate the diagnostic problems associated to Diesel NO\textsubscript{x} aftertreatment, a survey of the state of the art is necessary. Once the typical architectures for LNT and SCR systems have been identified, a detailed analysis of the structure and functionality of the aftertreatment system is done. Since fault diagnosis operates on the entire system, the analysis must not be limited to the sole catalysts, but rather include sensors, actuators and control strategies.

Various urea-SCR system architectures are presented in (Chi and DaCosta [2005], Willems et al. [2007], Ghoest et al. [2008]), but the most noticeable one as a commercial-grade implementation is the system proposed by Bosch and supported by several automakers (Seher et al. [2003]). The structure of a SCR system is schematically represented in Figure 1.

From a functionality standpoint, the control strategies for urea-SCR focus on NO\textsubscript{x} reduction control and ammonia slip control (to prevent over-optimal urea injection). Such strategies typically rely on a catalyst-out feedback from a NO\textsubscript{x} sensor, which is also a requirement for OBD-II compliance.

The urea injection control is typically based on open-loop static maps accounting for engine-out NO\textsubscript{x} emissions, catalyst temperature and ammonia storage. The control is typically based on a simplified model of the ammonia storage in the catalyst. In order to accommodate for nonlinearities and modeling errors, or to calibrate the control parameters, adaptation strategies are typically adopted to perform parameters estimation.

For light-duty applications, LNT-based aftertreatment systems are a typical solution (Geckler et al. [2001], Nakagawa et al. [2004]). The most adopted layout, as represented in Figure 2, is composed by a Diesel Oxidation Catalyst (DOC), a DPF and a LNT catalyst. The use of in-cylinder regeneration methods is the general solution to provide rich exhaust conditions for trap regeneration, such as through late fuel injection or high EGR dilution.

The sensor set typically includes a NO\textsubscript{x} sensor at the catalyst outlet or catalyst mid-bed. Depending on the

![Fig. 2. Schematic of typical LNT aftertreatment system](image-url)

3. FTA AND FMEA OF NO\textsubscript{x} AFTERTREATMENT SYSTEMS

The structural and functional analysis detail the fundamental components of the aftertreatment system (catalytic converter, sensors and actuators). This initial step facilitates the identification of the faults that could affect each component. However, due to the high system complexity and the presence of feedback control, the task of describing the effects caused by each fault on the components and the system may become difficult. For this reason, Failure Mode and Effects Analysis (FMEA) and Fault Tree Analysis (FTA) can be adopted to streamline the characterization of faults and the creation of fault tables.

FMEA is an inductive method to analyze failure modes, using a bottom-up approach. The analysis starts from the basic system components for which accurate information about failure modes and their causes are available. By analyzing the functional relationships among the components, it is possible to identify the possibility of propagation of each type of failure and to predict its effects on the production performance of the entire system. In case of the LNT aftertreatment system (as shown in Figure 2), the detailed FMEA is reported in Table 1. The system is characterized in terms of its different components and for each of them the failure modes, failure causes, and failure effects are described. The scope of the FMEA is to prioritize the failures according to their consequences, how frequently they occur and how easily they can be detected.

Conversely, the FTA is a deductive top-down method that describes all possible combinations of events that produce a failure mode of the systems up to the failure modes of the individual components. A system model is used to identify the states where the system should be in at any point
in time. The causes of a system’s failure modes will be described in terms of the component states.

The starting point for the FTA is based on the definition of an undesired effect as the root (‘top event’) of a logic tree. In case of a LNT-based aftertreatment system, a failure can be defined when an excessive slip of NOx or reductants is detected. Then, each situation that could cause system failure is added to the tree as a series of logic expressions. The FTA represents all possible combinations of events that produce a specific system failure, up to the failure modes of the individual components.

Fig. 3. FTA for LNT aftertreatment system

Figure 3 illustrates the FTA for the LNT aftertreatment system. It is worth observing that the catalyst is a critical component, being subject to faults (i.e., thermal deactivation and sulfur poisoning) that can immediately cause a system failure. This consideration motivates the model-based fault diagnosis approach described in the following sections, aimed at isolating faults on sensors (such as the catalyst-out temperature sensor, as in Figure 3), as well as the LNT parametric faults.

4. DIESEL ENGINE AFTERTREATMENT MODEL

For the fault diagnosis of Diesel aftertreatment systems, a model-based approach present significant advantages, for instance avoiding complex and costly experimental investigations on a catalyst test bench. Starting from a basic knowledge of the aftertreatment system, a control-oriented model can be designed and calibrated to characterize a selected set of faults and their most relevant effects on the system with very limited computation effort. Such model can be then utilized for fault detection, either as a “virtual catalyst” or as part of diagnostic algorithms.

In the proposed study, the available knowledge is a detailed model of a Diesel LNT aftertreatment system, designed for the characterization of emissions during transient operations and driving cycles (Canova et al. [2007]). The simulator is based on a detailed, physically-based LNT catalyst model, which accounts for the basic chemical reactions occurring at the catalyst surface, the storage and release dynamics of oxygen and NOx, and the temperature effects. The model also characterizes the storage and release dynamics of sulfur (SO2) adsorption and the most relevant effects of thermal deactivation on the gas species at the catalyst outlet. The model, validated from experimental data collected on a Diesel LNT system, applies the basic conservation principles with a phenomenological characterization of the reaction mechanisms during the storage and release phases. The resulting structure is formulated as a system of highly nonlinear, coupled differential-algebraic equations.

The LNT model is coupled with simple models for engine emissions, a Diesel oxidation catalyst, as well as oxygen, NOx and temperature sensors, providing a framework for the design of LNT regeneration control. A closed-loop controller schedules the LNT regeneration event, based on the feedback of the NOx concentration and AFR at the LNT outlet. The controller triggers the regeneration when the cumulative NOx storage efficiency of the catalyst crosses a minimum threshold. During the regeneration phase, the engine air/fuel ratio is modulated by a feedback PI controller to ensure a complete unloading of the trap and to prevent the reductant slip that could compromise CO and HC emissions (Canova et al. [2008]).

Finally, the aftertreatment model is coupled with a quasi-steady model of a vehicle longitudinal dynamics, including simple models for engine, transmission and powertrain
components. The simulator, calibrated on data from a midsize SUV powered with a 1.9l Diesel engine, allows one to estimate tailpipe emissions during driving cycles under healthy and faulty catalyst conditions (Canova et al. [2007]).

5. MODEL REDUCTION TO NONLINEAR STATE VARIABLE FORM

Despite the provided level of detail, the aftertreatment system simulator is rather complex for fault diagnosis purposes. Since a simpler structure would be desirable, model reduction can be operated to generate a control-oriented model from the more complex simulation model. The focus is on the ability to characterize the behavior of the catalyst in healthy and faulty conditions, subject to sulfur poisoning and thermal deactivation.

As Equations 4 and 6 show, the LNT temperature appears in a nonlinear fashion in the storage capacity, as well as the storage, release and conversion efficiencies. For this reason, a model for the temperature dynamics can be formulated, based on the energy conservation law:

\[
\frac{dT_{cat}}{dt} = \frac{1}{C_{max}} \exp \left[ - \left( \frac{T_{cat} - T_{in}}{T_{max}} \right)^2 \right]
\]

This parameter is identified on experimental data, as a function of the catalyst temperature (Choi et al. [2005]).

The basic model equations are based on the conservation principles. In particular, the NO\textsubscript{x} storage dynamics model is based on the continuity equation, applied first to the mass of NO\textsubscript{x} stored on the catalyst in solid phase, then to the gas phase. This allows one to write the state and output equations:

\[
\frac{d\xi_{NOx}}{dt} = \frac{1}{C_{NOx}} (\dot{m}_{NOx,stor} - \dot{m}_{NOx,rel})
\]

\[
\dot{m}_{NOx,out} = \dot{m}_{NOx,in} - \dot{m}_{NOx,rel} + \dot{m}_{NOx,slip}
\]

where \( \xi_{NOx} \) is the normalized catalyst fill ratio and \( C_{NOx} \) the NO\textsubscript{x} storage capacity:

\[
C_{NOx} = C_{max} \exp \left[ - \left( \frac{T_{cat} - T_{in}}{T_{max}} \right)^2 \right]
\]

The terms on right-hand side of Equations (3) are the mass flow rates of the NO\textsubscript{x} stored on and released from the trap, and of the NO\textsubscript{x} slipping during regeneration. Analytical formulations can be found from physical and empirical observations (Nieuwstadt and Yanakiev [2004], Brandt et al. [2000]):

\[
\begin{align*}
\dot{m}_{NOx,stor} &= k_{stor} \cdot \eta_{stor} \cdot \dot{m}_{NOx,in} \\
\dot{m}_{NOx,rel} &= k_{rel} \cdot \eta_{rel} \cdot x_{CO-HC} \cdot \dot{m}_{exh,in} \\
\dot{m}_{NOx,slip} &= [1 - \eta_{conv}] \dot{m}_{NOx,rel}
\end{align*}
\]

The NO\textsubscript{x} storage, release and conversion efficiency are complex expressions depending on the catalyst fill ratio, temperature and redundant concentration (during the regeneration phase):

\[
\begin{align*}
\eta_{stor} &= \frac{1 - \exp(a_1 \xi_{NOx})}{\exp(a_1) - 1} \\
\eta_{rel} &= \frac{\exp(-a_2 \xi_{NOx}) - 1}{\exp(-a_2) - 1} \\
\eta_{conv} &= \frac{\exp(-a_3 \xi_{NOx}) - 1}{1 - \exp(-a_4)}
\end{align*}
\]

where \( a_i = a_{i1}T_{cat} + a_{i2} \). The parameters \( a_{ij} \) are identified using the parent model as a “virtual catalyst”, simulating storage and regeneration cycles at several engine operating conditions.

With similar methods, the parameters: \( k_{stor}, k_{stor}, x_{CO-HC} \) can also be identified. In particular, the net concentration of reductants \( x_{CO-HC} \) during the trap regeneration phase, was identified based on a normalized air/fuel ratio:

\[
x_{CO-HC} = k_0 \cdot \exp(-k_1 \lambda_{in})
\]

As Equations 4 and 6 show, the LNT temperature \( T_{cat} \) appears in a nonlinear fashion in the storage capacity, as well as the storage, release and conversion efficiencies. For this reason, a model for the temperature dynamics can be formulated, based on the energy conservation law:
\[
\frac{dT_{\text{cat}}}{dt} = \frac{1}{C_{\text{cat}}} \left[ \dot{m}_{\text{exh,in}, cp} (T_{\text{in}} - T_{\text{cat}}) - Q_{ht} + Q_{\text{reac}} \right]
\] (8)

where the thermal losses are due to convective heat transfer (Canova [2006]) and a simple correlation is used to estimate the energy released by the conversion reactions during regeneration (Canova et al. [2007]):

\[
Q_{\text{reac}} = m_{\text{NOx,rel}} h_{\text{NOx}}
\] (9)

The term \( h_{\text{NOx}} \) is the enthalpy associated to the chemical reactions that lead to the release of the stored \( \text{NO}_x \) and the following combination with the available reductants in the exhaust. This parameter is also identified on the parent model.

The validation of the model was done by comparing the values of the states and outputs to the corresponding variables in the parent model. As an example, Figure 6 shows the validation results during a portion of a test comprising repeated FTP driving cycles. As can be observed, very good agreement is achieved on the catalyst-out \( \text{NO}_x \) (here represented as mass flow rate), which is a critical output to be used for fault diagnosis.

6. FRAMEWORK FOR MODEL-BASED FAULT DIAGNOSIS

Although unable to provide information on faults affecting certain components, the reduced control-oriented model allows one to focus on the detection and isolation of the most critical LNT catalyst faults, particularly on the outputs and parameters. Therefore, model-based strategies may be designed to isolate faults on sensors (such as the catalyst-out temperature sensor), as well as sulfur poisoning and thermal deactivation.

The fault diagnosis scheme is based on a parity equation approach, comparing the behavior of the system with the behavior of the aftertreatment model to produce residuals containing fault information (Chen and Patton [1999], Chiang et al. [2001], Kimmich et al. [2005]). The scheme is based on the following assumptions:

- residuals are considered only during storage phase;
- the catalyst-out \( \text{NO}_x \) mass flow rate and temperature can be measured;
- inputs are assumed measurable (or estimated) and fault-free.

Table 2. FDI scheme: error signature

<table>
<thead>
<tr>
<th>Fault</th>
<th>( R_1 )</th>
<th>( R_2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Temp. Sensor</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Thermal damage</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>( \text{SO}_2 ) poisoning</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

Figure 7 shows a basic structure for fault diagnosis based on parity equation method. Thresholds for residuals have to be set to accommodate for noise and model uncertainties. The threshold levels are derived statistically to minimize false alarms by calculating the normal deviations of the estimates from the measured variables during several tests under no fault conditions (Canova et al. [2008]). Under the assumption of single fault occurrence, the fault isolation logic shown in Table 2 is derived.

Fig. 8 shows the residuals for temperature sensor fault diagnosis.

To test the effectiveness of the FDI scheme, a number of simulations were run inducing faults in the system. The results presented are related to tests for a catalyst-out temperature sensor fault and a thermal deactivation fault. The tests were performed in transient conditions, operating during a series of FTP driving cycles.

A fault on the temperature sensor was simulated by introducing a +20°C offset after about 4000s of simulated time. Figure 8 shows that the FDI scheme correctly detects the fault, as the residuals R2 (based on the outlet...
A thermal deactivation fault was introduced in the system after 2000s of simulation, by increasing the reaction heat $Q_{\text{react}}$ (in Equation 8), during one regeneration. As shown in Figure 9, the FDI scheme correctly detects the thermal damage fault, as the residuals R1 and R2 exceed the pre-set thresholds.

7. CONCLUSIONS

The paper illustrates a step-by-step methodology for fault diagnosis of complete Diesel NO\textsubscript{x} aftertreatment systems, which is applicable to LNT or SCR catalytic converters with relative sensors, actuators and feedback regeneration control.

The proposed approach is based on a preliminary analysis, which identifies the main components of the aftertreatment system and their functionalities. This facilitates the application of FTA and FMEA techniques, which identify the faults on each individual components and their effects on the performance of the entire system. This enables for the design of control-oriented models to be applied to the definition and testing of fault diagnosis algorithms.

The paper illustrates the methodology through a case study of a Diesel LNT aftertreatment system. Starting from a physically-based detailed model previously developed, reduction methods are applied to formulate a control-oriented model of the LNT catalyst capturing (in a phenomenological fashion) the parametric faults, such as thermal damage and sulfur poisoning. Using this model, a fault diagnosis scheme is defined based on the parity equations approach. The scheme, tested in simulation, is able to detect and isolate temperature sensor faults from the more critical LNT parametric faults.

The on-going research is devoted to extend the developed fault diagnosis scheme to encompass the complete set of system faults, based on the estimation of system states and parameters such as the NO\textsubscript{x} storage capacity. The methodology illustrated in the paper will also be applied to Urea-SCR systems, and validated on an experimental setup.

REFERENCES


M. Canova, P. Pis, and A. Soliman. Model-Based Fault Diagnosis of a NO\textsubscript{x} Aftertreatment System. *17th IFAC World Congress*, 2008.


