Brandon M. Dawson
Matthew A. Franchek
Karolos Grigoriadis

Department of Mechanical Engineering,
University of Houston,
Houston, TX 77204

Robert W. McCabe
Mike Uhrich
Steve Smith

Ford Motor Company,
Dearborn, MI 48124

Automotive Three-Way Catalyst Diagnostics with Experimental Results

Presented is the model based diagnostics of a three-way catalyst (TWC). The proposed TWC model relates measurable engine inputs (engine air mass (AM) and catalyst temperature) to a metric that quantifies TWC oxygen storage capacity. The TWC model structure is based on the dynamics of the TWC and identified using orthogonal least squares (OLS). The model coefficients are estimated using an instrumental variables four step (IV4) approach. TWC diagnostics is realized by means of an information synthesis (IS) technique where changes in the adapted TWC model coefficients are utilized to estimate TWC health. The approach is experimentally validated on a federal test procedure (FTP) drive cycle for healthy (full useful life, FUL) and failed (threshold) TWCs. The results will show that a 100% accurate classification in TWC health estimation (FUL or threshold) is produced for the catalysts tested. [DOI: 10.1115/1.4004041]

Introduction

The Environmental Protection Agency (EPA) and the California Air Resource Board (CARB) began requiring onboard diagnostic (OBD) in the 1980s. Today, OBD-II requires every system that impacts tailpipe emissions be monitored. The main requirement from OBD-II is the detection of hydrocarbon (HC), carbon monoxide (CO), and oxides of nitrogen (NO_x) or evaporative emissions that exceed 1.5 times the federal test procedure (FTP) standards for that model year [1]. A catalyst monitor ensures that the TWC is functioning at an acceptable efficiency level keeping emissions within mandated requirements.

A comprehensive review of patented catalytic converter diagnostic methods, prior to 1998, is presented in Ref. [2]. This review details the three main modalities of TWC diagnosis: comparison of upstream and downstream exhaust gas oxygen (EGO) measurements, measurement of exhaust gas temperature to determine exothermic heat, and direct measurements of HC, CO, or NO_x. The comparison of EGO sensor readings is by far the most common method. The EGO signal comparisons is an indirect diagnosis method (meaning the exhaust gas composition is not directly analyzed, it is inferred) based on oxygen storage capacity (OSC) of the TWC. Temperature based diagnostics is also an indirect approach whereby measuring the heat release caused by the exothermic reactions taking place in the TWC indicates catalyst light off [2]. A properly functioning TWC generates heat during a purge event owing to the chemical reactions. This exothermic reaction leads to a difference between inlet and outlet exhaust gas temperatures. As the TWC ages, the main reaction zone moves from the front of the catalyst toward the back, representing a change in the temperature profile along the TWC length. Usually a temperature based method is rarely used since it requires long steady state driving conditions to accurately measure the temperature differences. Direct TWC diagnosis using HC, CO, or NO_x sensors seems like the obvious choice for catalyst diagnosis and emission monitoring. These sensors measure emission quantities directly and report a failure when the emission is 1.5 times its standard. However, because these sensors are generally bulky, temperature limited, and extremely cost prohibitive, they have not found their way to production.

On the road today, TWC diagnostic monitors utilize the EGO sensors, both upstream and downstream of the catalyst, to estimate TWC health. The TWC acts as a time varying emissions capacitor between the upstream and downstream EGO sensors. When the TWC is green (under 4000 miles) it is storing and releasing oxygen at a rate such that the downstream EGO senor measures a heavily attenuated response in comparison to the upstream EGO response. When the TWC is degraded, its ability to convert pollutants is dramatically reduced. In essence, a failed TWC behaves more like a straight pipe; "what goes in, comes out."

There are two basic classes of EGO based catalyst monitors: intrusive and passive. The intrusive monitors described in Ref. [3] use engine fueling to provide a persistently excited repeating cycle of lean and rich perturbations for a preset duration while the engine is running within a given speed/load window. The monitor then calculates the downstream signal length and divides it by a calibrated, worst case, low oxygen storage length. Failure is defined as when this ratio is near or greater than unity [3]. Passive monitors as described in Ref. [4] do not require the fuel perturbations to estimate TWC health. Instead, this monitor activates when speed, load, and AM reside within a prespecified window of operation. The monitor then calculates the signal ratio between the downstream signal arc length and the maximum allowable length. The catalyst is considered failed when this signal ratio is beyond a calibrated value [4]. There are pros and cons to both types of monitors. For example, when an intrusive monitor activates, closed-loop fueling is suspended, thereby leading to possible emission penalties. As the EPA implements more stringent emission requirements, intrusive monitors may become obsolete because of their potential to negatively impact emissions. Passive monitors, on the other hand, pose a challenge in that the persistency of excitation condition is contingent upon the drive cycle and the driver. Thus a passive monitor executes in batch mode once the engine has entered the window of operation.

There is a class of catalyst diagnostic methods that involve integrating subsystem models together to form an overall catalyst model. Brandt et al. proposed a model in Ref. [5] that combines three submodels: an oxygen storage model, a thermal model, and a static mapping model to predict exhaust emission levels thereby determining catalyst efficiency. The models use feedgas air/fuel ratio (AF), mass air flow (MAF), and feedgas temperature to estimate NO_x, HC, and CO conversion efficiencies. This work was expanded in Ref. [6] to incorporate a hypothesis test to estimate catalyst health. An assumption within this test is that the TWC is healthy until statistical evidence shows the TWC is malfunctioning. This method focuses on the effective catalyst volume or depleted OSC as the TWC failure mode for which the authors quantify by the empty or fill time of the OSC. The hypothesis test classifies the TWC as healthy until the TWC effective volume is

Contributed by the Dynamic Systems Division of ASME for publication in the JOURNAL OF DYNAMIC SYSTEMS, MEASUREMENT AND CONTROL. Manuscript received November 24, 2008; final manuscript received February 19, 2011; published online August 5, 2011. Assoc. Editor: Chris Gerdes.

greater than a prespecified quantity. A degraded TWC is classified as damaged when the effective volume is less than that required to meet emission regulations. One limitation of this methodology is that it requires longer than normal rich and/or lean excursions to determine the empty or fill characteristics required for diagnosis.

The TWC diagnostic knowledge base appearing in the open literature is significantly larger than presented. There are, however, commonalities among these many methods. Oxygen storage capacity estimation, in one form or another, appears in most of the proposed TWC diagnostics methods since OSC correlates to NO_x and HC efficiencies [7]. Moreover, OSC is generally estimated using models that compare upstream and downstream EGO responses. Another recurring theme in the TWC diagnostics literature is the use of engine operating information to make diagnostics more robust. Some of these parameters include engine speed, inlet AF ratio, catalyst temperature, exhaust gas temperature, and AM through the catalyst, all of which impact conversion efficiency of the TWC.

Three-Way Catalyst Primer

Three-way catalysts are engineered to direct reactants to specific products. The goal of the automotive TWC is to convert harmful engine feedgas emissions (NO_x, CO, and HCs) into benign gasses (N_2, O_2, CO_2, and H_2O). To promote oxidation and reduction reactions and enhance conversion efficiency, cerium (Ce) and various cerium oxides are added to the TWC washcoat. The cerium stores oxygen during lean engine conditions and releases the oxygen during rich engine conditions. The TWC OSC is a critical health characteristic influencing conversion performance. In particular, as the OSC decreases, the conversion capability of the TWC (the TWC health) decreases.

For a given health level, the TWC conversion efficiency is dominated by two primary parameters, exhaust gas space velocities and TWC temperature. As damage is introduced into the TWC, there is an irreversible decline in the TWC performance. The TWC damage considered in this work occurs through one of two possible paths: excessive TWC temperature leading to sintering or washcoat poisoning due to high sulfur fuels. Although these damage mechanisms are distinct, the resulting effects are similar. Namely, the number of active sites in the TWC is reduced, which manifest itself as a reduction in the OSC of the TWC.

Space Velocity and TWC Efficiency. The influence of space velocities on the chemical kinetics in a catalytic converter is described in Refs. [8] and [9]. These references assume a one-dimensional, isothermal, plug flow in the TWC during steady-state operation. Applying a material balance between the catalyst inlet (i) and outlet (o) yields

$$\ln\frac{C_{\beta_o}}{C_{\beta_i}} = -\frac{k'z}{v} = -k't \tag{1}$$

where v is the gas velocity (cm/s), C is the molar concentration of constituent [(g mol)/(cm^3)], z is the length of the catalytic reactor (cm), k' is the apparent rate constant, β is a particular reactant, and t is the residence time(s) of a reactant in the catalytic bed. In automotive TWCs, the residence time is a function of volumetric flow and catalyst volume. Space velocity (SV) through the catalyst is the volumetric gaseous flow rate at standard temperature and pressure (STP) divided by the volume of the catalyst plus void volume. Substituting the reciprocal of SV, which is the residence time, into Eq. (1), the rate expression becomes

$$\ln\frac{C_{\beta,o}}{C_{\beta,i}} = -\frac{k''}{\mathrm{SV}} \tag{2}$$

where k'' is the empirically determined ratio $\Delta\mathrm{SV}/C_\beta$. From Eq. (2), it can see that to increase conversion of reactant β (i.e., smaller concentration of β at outlet than at inlet) a decrease in space velocity must occur.

Temperature and TWC Efficiency. As the temperature of catalyst TWC increases from cold start conditions, the TWC oxygen storage capacity increases until the TWC reaches its peak efficiency temperature range, which is between 400 and 600°C. Under high load operation, the catalyst temperature can surpass 900°C at which point, the TWC oxygen storage capacity decreases due to the sintering of catalyst sites [10]. When this occurs, the catalytic sites begin to coalesce, thereby decreasing the surface to volume ratio which decreases the available catalytic sites. Therefore, active sites get buried within a growing crystal and with fewer sites available for a reaction, a decline in TWC performance will follow. Such TWC damage is permanent.

Estimating TWC State. Presented in this section is the introduction of a metric representing real-time estimate of a TWC state. The metric is an online adaptive TWC gain K_{TWC} defined as

$$\dot{V}_{\mathrm{HEGO}}(i) = K_{\mathrm{TWC}}\,\Delta\varphi(i) \tag{3}$$

where $V_{\mathrm{HEGO}}(i)$ denotes the sampled voltage from a narrowband HEGO sensor located in the middle of the TWC (also called the monitoring sensor) and $\Delta\phi(i)$ is deviations in the feedgas fuel-air ratio about unity.

This proposed metric is meaningful in that it essentially captures the state of the TWC OSC and lumps all other catalyst behavior such as space velocities and TWC temperature into one adaptive gain [11]. To illustrate, consider the K_{TWC} estimates using Eq. (3) for three different TWCs tested over the same engine cycle (Fig. 1). The green TWC gain is consistently low (near zero) over the engine cycle (Fig. 1). The interpretation of K_{TWC} in this case is that despite the incoming exhaust feedgas emissions, the resulting emissions reaching the monitoring sensor in the TWC are near zero. In fact, a low K_{TWC} gain indicates a high OSC for a TWC. As the aged and damage TWC were tested, a corresponding increase in the estimated K_{TWC} occurred, indicating a low OSC of the TWC. In response to an increase in K_{TWC}, the fueling controller modulates the air-fuel ratio, thus driving the chemical kinetics and thermal condition of the TWC in an effort to maintain regulated EPA emission requirements.

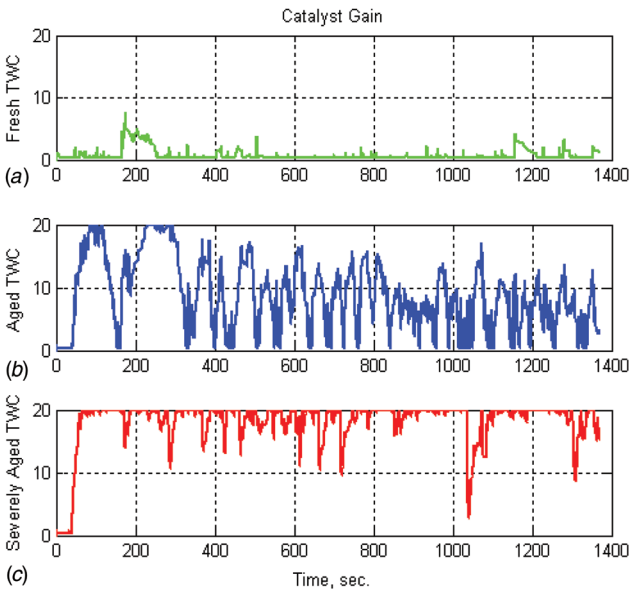

Fig. 1 K_{TWC} estimates for differently conditioned TWCs: (*a*) green TWC, (*b*) aged TWC, and (*c*) damaged TWC

The online estimation of K_{TWC} is based on a gradient descent approach. The sensor used in this estimation is a midbed (monitoring) narrowband EGO sensor (switching type) called the HEGO. To attenuate the switching frequency behavior of the HEGO and recover a DC trend, the voltage from the HEGO sensor, V_{HEGO}, is low pass filtered as

$$V^f_{HEGO}(k) = \frac{1}{1 - aq^{-1}} V_{HEGO}(k) \tag{4}$$

where $V_{HEGO}(k) = V_{HEGO}(t)|_{t=kT}$, T is the sampling interval in seconds, q^{-1} denotes a back-sample, and a is selected such that the filter break frequency corresponds to 1 rad/s. This value was preselected by the engine manufacturer. Similarly, the EEC fuel-air command must be filtered using the same low pass filtering giving

$$\Delta\varphi^f(k) = \frac{1}{1 - aq^{-1}} \Delta\varphi(k). \tag{5}$$

Using a first order backward difference equation to approximate the derivative in Eq. (3) and employing, the filter signals gives the following discrete equation:

$$V^f_{HEGO}(k) = T\,K_{TWC}(k)\Delta\varphi^f(k) + V^f_{HEGO}(k-1) \tag{6}$$

The discrete equation for the one step ahead *estimate* is

$$\hat{V}^f_{HEGO}(k) = T\,\hat{K}_{TWC}(k)\Delta\varphi^f(k) + V^f_{HEGO}(k-1) \tag{7}$$

Defining the error necessary for adapting the estimated TWC gain gives

$$e(k) = V^f_{HEGO}(k) - \hat{V}^f_{HEGO}(k) \tag{8}$$

Using the gradient descent method defined as

$$\Delta\hat{K}_{TWC}(k) = -\gamma\, e(k)\, \frac{\partial e(k)}{\partial \hat{K}_{TWC}} \tag{9}$$

the adapted TWC gain equation

$$\begin{aligned}\hat{K}_{TWC}(k+1) &= \hat{K}_{TWC}(k) + \gamma\, T\, \Delta\varphi^f(k) \\ &* \left(V^f_{HEGO}(k) - V^f_{HEGO}(k-1) - \hat{K}_{TWC}(k)\, T\, \Delta\varphi^f(k)\right)\end{aligned} \tag{10}$$

Thus, K_{TWC} in Eq. (3) is calculated as defined in Eq. (10).

Section Summary. Reviewed in this section is the operation of a TWC, an adaptive gain of the TWC quantifying its state and those engine parameters (SV and TWC temperature) that influence the TWC conversation efficiency. In the "Proposed TWC Diagnostic Approach" section, an online adaptive TWC model predicting K_{TWC} will be developed based on SV, TWC temperature, and past calculations of K_{TWC} via Eq. (3). It is expected that K_{TWC} will be proportional to SV since higher level SVs will lead to more emissions reaching the midbrick of the TWC thus giving a higher K_{TWC} gain. For the TWC temperature, the conversion efficiency of the TWC increases as the catalyst reaches its light-off. Thus, as TWC temperature increases, K_{TWC} is expected to decrease indicating improved TWC performance. Therefore, the inverse of TWC temperature is expected to appear in the adaptive diagnostics model involving K_{TWC}.

Proposed TWC Diagnostic Approach

This work focuses on the development of an adaptive TWC health model using healthy TWC data from the FTP cycle. The health monitoring model uses SV, the inverse of TWC temperature and past samples of K_{TWC} to estimate the current K_{TWC} value calculated using Eq. (10). Based on the changes in the coefficients of this model, a diagnostic algorithm will estimate TWC health. To develop the TWC health model, orthogonal least squares (OLS) is used to determine the model regressors. These regressors are then used in the instrumental variables four step (IV4) method to identify the diagnostics model coefficients from which an information synthesis (IS) technique will be used to estimate TWC health. The primary advantage of using adaptive models is that both slowly time varying dynamics and sensor noise attenuation are directly addressed. The latter feature is addressed by consistency in the model parameter estimated via system identification methods. This improves the robustness of the diagnostics method. Presented herein are the three primary steps in developing the TWC adaptive model diagnostics method. In particular, orthogonal least squares will be used to identify the adaptive model structure. Instrumental variables four step approach will be used to obtain consistent estimates of the adaptive model parameters. Finally, information synthesis is employed to realize the diagnostics approach.

Orthogonal Least Squares. Least squares (LS) techniques have been employed in the past to identify model parameters of real world systems. The main problem with LS is that it requires a priori knowledge of the system: model structure and terms to be included in the model must be known or hypothesized. Once these unknowns have been determined only the model parameters are unknown and LS can then solve the problem. However, the model structure of real systems is rarely known ahead of time and therefore methods of model structure determination must be included as an important part of the identification process.

An orthogonal algorithm efficiently and effectively identifies model structure. This algorithm was originally developed in Ref. [12] and made more reliable in Ref. [13]. Inputs to the OLS algorithm are comprised of the anticipated model regressor variables (v), the maximum number of terms desired in the model (*terms*), and the maximum integer power to which the regressor term is raised (p). Based on these parameters, OLS systematically searches for the regressor combinations and permutations that result in the lowest error between actual output and model output. The most significant regressor is added to the model, removed from the orthogonal equation, and the method searches for the second most significant regressor. This process repeats until the specified number of terms is reached or until the error between the actual output and model output is within a desired tolerance.

Online TWC Health Monitoring Model Adaptation. The coefficients of the model are identified online using the IV4 approach for parameter estimation outlined in Ref. [14]. The first step in the IV4 process uses a standard LS estimation technique to calculate an initial estimate of the model parameters. The second step performs an IV parameter estimate using the model output from step one as the input to step two. The instrumental variables are correlated to the system input and uncorrelated to system noise. Next a noise filter is identified based on the error of the second model. A final estimate is recovered using the filtered instrumental variables and filtered data. This four step process has a proven reliability in obtaining consistent parameter estimates.

Information Synthesis. IS is a fault detection, isolation, and estimation methodology presented in Ref. [15]. Fault detection, isolation, and estimation are realized by adapting the model coefficients online, which in turn addresses issues such as sensor noise. A coefficient error vector is produced by comparing healthy model parameters (H) to the adapted model parameters (F) that may or may not contain a fault Ref. [15]

$$\mathrm{E} = \mathrm{S}(\mathrm{H} - \mathrm{F}) \tag{11}$$

where

$$S = diag\left(\frac{1}{h_1} \quad \frac{1}{h_2} \quad \ldots \quad \frac{1}{h_i} \quad \ldots \quad \frac{1}{h_X}\right) \tag{12}$$

$$H = [h_1 \quad h_2 \quad \ldots \; h_i \ldots \; h_X]^T \tag{13}$$

and

$$F = [a_1 \quad a_2 \quad \ldots \; a_i \ldots \; a_X]^T. \tag{14}$$

Fault Size Estimation. Detecting the presence of a fault is achieved by evaluating the magnitude M of E and projecting it onto directional vectors (fault isolation). Ensuring that the error is not due to normal system variability or sensor noise, a statistically determined length ε is used as

$$M = \begin{cases} \|E\|_2 - \varepsilon & \text{provided } \|E\|_2 > \varepsilon \\ 0 & \text{else} \end{cases} \tag{15}$$

where the 2-norm, or length, of the error vector is

$$\|E\|_2 = \sqrt{\sum_i |E(i)|^2} \quad i \in I_{+/(X+1)} \tag{16}$$

and ε is the statistically determined threshold. For this work, ε is defined by

$$\varepsilon = 3\sigma_T \tag{17}$$

where σ_T is the worst case standard deviation of the $\|E\|_2$ for a threshold catalyst. This value of 3 σ_T represents a 99.73% confidence interval assuming a normally distributed $\|E\|_2$.

To identify the coefficients of the healthy model, a series of experiments is performed producing a family of healthy model coefficients $H_{(z,X)}$, where z is the number of experiments performed and X is the number of model coefficients. The mean of each element in $H_{(z,X)}$ gives the healthy model coefficients of Eq. (13).

Fault Isolation. The magnitude of the error vector alone does not contain enough information to guarantee a fault. Fault isolation will be realized through the projection of E onto predefined vectors that facilitate isolation as described in Ref. [15]. The predefined vectors, Dir_j, are defined to isolate a fault to a specific path that exists between the model input(s) and output. The general unity fault direction vector for the jth path is defined by

$$Dir_j = \frac{1}{\|Dir_j\|_2}[d_1 \quad d_2 \quad \cdots \quad d_i \quad \cdots \quad d_X] \tag{18}$$

where d_i's are chosen based on the direction of the fault. If the fault direction is unknown or if specific fault path isolation is desired, a method described in Ref. [15] chooses d_i to be either 0 or 1 based on the influence of the jth input on the model regressors. Fault isolation can be isolated to a specific path through the projection of E on Dir_j

$$\Psi(j) = Dir_j \bullet E \tag{19}$$

where $\bullet$ denotes the usual dot-product. The location detected fault will be isolated as the largest projection, defined as

$$\Psi(l) = \max[\Psi(j)] \tag{20}$$

A schematic representation of the IS fault space is shown in Fig. 2 to illustrate the ability of IS to detect and isolate a fault to a specific path. This generalized representation shows three distinct fault spaces, which could correspond to three different possible faults of the system. Fault isolation gives realization to the identification of these different faults. The radius of each fault space is a result of system variability possibly due to, among other things, sensor noise and accuracy.

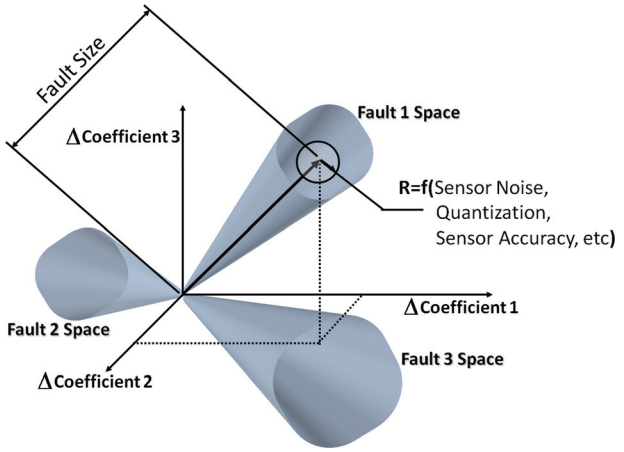

Fig. 2 Example of information synthesis fault space projections

Experimental Results

The full useful life (FUL) TWC used in this work was created on an engine dynamometer test where the fuel is doped with phosphorus (P_2O_5) to simulate 150,000 mile aging. The threshold catalyst was created by first oven aging the brick for 8 h at 1250°C. The brick was then submerged in water to determine the H_2O absorbed per volume of the TWC. Once dried, the brick was submerged into a dilute H_3PO_4 solution to achieve 5% phosphorus, dried again, and calcined.

The FTP data for the FUL TWC was used for the identification of the model regressors. A total of ten FTP cycle data sets were analyzed; five equipped with FUL TWCs and five threshold TWCs. The possible model regressors for OLS to identify were chosen based on the physics of the TWC operation and include engine air mass (*AM*) (correlated to SV), the inverse of catalyst temperature (*invTEMP*) since temperature and conversion efficiency are inversely related, and four back samples of K_{TWC}: *AM(k), invTemp(k),* K_{TWC} *(k-1),* K_{TWC} *(k-2),* K_{TWC} *(k-3),* and K_{TWC} *(k-4)*. With these possible regressors, a maximum power of $p = 3$ is selected thereby giving the highest power regressor term as

$$AM^3(k) * invTemp^3(k) * K^3_{TWC}(k-1) * K^3_{TWC}(k-2) * K^3_{TWC}(k-3) * K^3_{TWC}(k-4)$$

The maximum number of terms is selected as 10. The resulting OLS model is

$$\hat{K}_{TWC}(k) = \phi^T(k)\Theta \tag{21}$$

where

$$\phi(k) = [K_{TWC}(k-1) \quad K_{TWC}(k-2) \quad K_{TWC}(k-3) \ldots \quad K_{TWC}(k-4) \quad AM(k) * invTEMP(k)]^T \tag{22}$$

and

$$\Theta = [a_1 \quad a_2 \quad a_3 \quad a_4 \quad a_5]^T \tag{23}$$

Based on the dataset, the identified regressors best accounting for the output variance in the current sample of K_{TWC} is a linear model

Table 1 Sum of the square error for various number of terms and powers

	Maximum power (p)		
Terms	1	2	3
1	5.3690	5.3690	5.3690
2	0.1688	0.1688	0.1688
3	0.0240	0.0240	0.0240
4	0.0124	0.0124	0.0124
5	0.0108	0.0108	0.0108
6	0.0106	0.0106	0.0106
7	0.0106	0.0106	0.0106
8	0.0105	0.0105	0.0105
9	0.0103	0.0103	0.0103
10	0.0103	0.0103	0.0103

involving all four past samples and one input. Note that each K_{TWC} is calculated via Eq. (10). The coefficients a_i are estimated online using IV4. Notice that OLS analysis showed that powers greater than unity and more than five terms had little effect on reducing the error as seen in Table 1. Models were also developed using additional past samples of K_{TWC}, but it was found that more than four past samples did not reduce the error enough to justify their inclusion in the model. The measure of model accuracy is the sum of the square error (sqerr), which should be minimized

$$\mathrm{sqerr} = \sum_{i=1}^{N} \left(K_{\mathrm{TWC}}(k) - \hat{K}_{\mathrm{TWC}}(k)\right)^2 \tag{24}$$

where N is the number of data points.

Results. Prior to running the IV4 algorithm, the FTP data are segmented into 300 s windows, and each regressor is normalized by its mean value. Beginning just after the TWC reaches light-off temperature, an IV4 coefficient calculation is performed for each window segment, the results stored, and the window shifts over 60 s to perform another IV4 calculation. This procedure results in 100 coefficient estimates per FTP cycle, which will be used to create the healthy case coefficient set for the diagnostic method presented in the following section.

To depict the accuracy of the IV4 results one data segment is shown in Fig. 3. Compared in this plot is the actual output with the identified model output for a data set that was not used for model identification. The model represents a 98.79% match to the actual output, which yields a total sqerr of 1.589.

When IV4 is performed over the entire FTP cycle for the 300 s windows, this method has shown greater repeatability and greater model accuracy. The average FUL TWC coefficients, H, are

$$\mathrm{H} = [7.22 \quad -6.74 \quad 2.86 \quad -0.47 \quad 0.01]^T \tag{25}$$

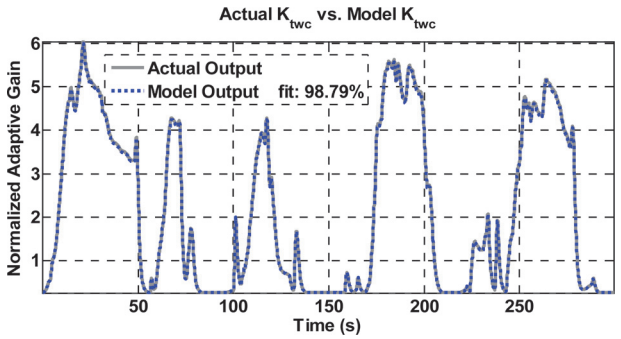

Fig. 3 Actual K_{TWC} vs. model K_{TWC} for one 300 s window

With the FUL set of coefficients determined, model parameter identification is applied to other FTP cycle data for FUL and threshold TWCs.

IS Results. To determine σ_T of Eq. (17), FUL catalyst data were analyzed and the maximum standard deviation of $\|\mathrm{E}\|_2$ was found to be 1.72 and therefore ϵ is rounded to be 5. Thus, M is defined as

$$M = \begin{cases} \|\mathrm{E}\|_2 - 5 & \text{provided } \|\mathrm{E}\|_2 > 5 \\ 0 & \text{else} \end{cases} \tag{26}$$

To validate the diagnostic, the method was applied to several different FTP data sets for FUL and threshold catalysts. For one particular segment of FTP data equipped with a FUL catalyst, the coefficient set identified by IV4 is

$$F_{\mathrm{FUL}} = [6.69 \quad -5.06 \quad 1.40 \quad -0.12 \quad 0.02]^T \tag{27}$$

and the corresponding error vector is

$$E_{\mathrm{FUL}} = [0.07 \quad 0.25 \quad 0.51 \quad 0.74 \quad -1.15]^T \tag{28}$$

which has a length, $\|\mathrm{E}\|_2$, of 1.48 and when this is applied to the fault detection criterion of Eq. (26), M is 0. Therefore, no fault is correctly detected.

When applied to FTP data equipped with a failed (threshold) catalyst, the coefficient set is

$$F_{\mathrm{Failed}} = [31.4 \quad -15.3 \quad 10.0 \quad -6.36 \quad 0.002]^T \tag{29}$$

with the error vector being

$$E_{\mathrm{Failed}} = [-3.35 \quad -1.28 \quad -2.51 \quad -12.5 \quad 0.74]^T \tag{30}$$

which has a length of 13.3 thus giving $M = 8.27$. Hence a fault is *possible*.

Fault isolation is realized by defining Eq. (18) in the direction of the fault relative to the healthy case. To realize this vector, the *center of gravity* for the healthy case error coefficients, denoted as $(\otimes_{E,H})$, are subtracted from the center of gravity for the threshold case error coefficients $(\otimes_{E,T})$ giving realization to the error direction

$$\otimes_{E,H} = [-0.23 \quad -0.25 \quad -0.30 \quad -0.39 \quad -0.12]^T \tag{31}$$

and

$$\otimes_{H,T} = [-4.40 \quad -1.95 \quad -2.63 \quad -12.3 \quad 0.74]^T \tag{32}$$

leading to a unit fault vector of

$$\mathrm{Dir} = [-0.32 \quad -0.13 \quad -0.18 \quad -0.92 \quad 0.07]^T \tag{33}$$

The direction of the fault will point to the coefficient or a combination of coefficients most responsible for fault. For a failed TWC, the projection of the E vectors for the threshold case onto Dir is upper and lower bounded as

$$1 \leq \Psi(j) \leq 25 \tag{34}$$

with a mean projection of 15, else TWC is healthy regardless of the error vector length, $\|\mathrm{E}\|_2$.

To validate the method, the entire FTP cycle was analyzed using each FTP cycle test. Of the over 500 window segments that were analyzed for each case, every threshold catalyst was properly diagnosed as failed. In contrast, four false classifications were

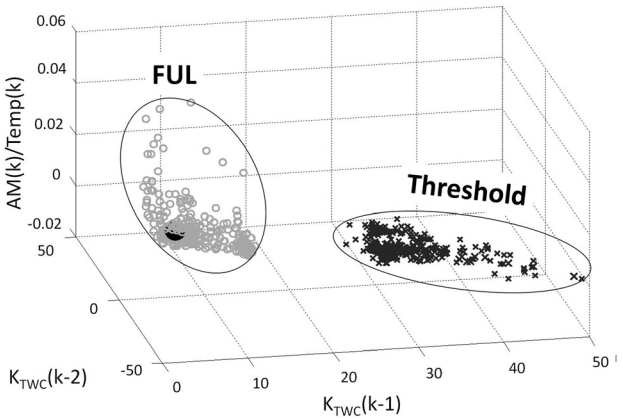

Fig. 4 Model coefficients in the coefficient space for healthy, FUL, and threshold TWCs

found for the FUL case when $\|E\|_2$ alone is used for diagnostics. When the fault is isolated using Eq. (19) and applied to the criterion in Eq. (34), all false classifications are rejected and properly diagnosed, *leading to a* 100% *accurate classification rate*.

Reduced Order TWC Health Monitoring Model. For diagnostic purposes, the entire regressor and coefficient set, determined by OLS, was reduced. It was found that using three coefficients of the original model can provide as accurate classification of the healthy and threshold cases. The coefficients corresponding to the K_{TWC} *(k-1)*, K_{TWC}*(k-2)*, and *AM*invTemp(k)* regressors were chosen for diagnostics. The inclusion of the *AM*invTemp(k)* coefficient is to accommodate for possible excursions in engine operation. Shown in Fig. 4 is the projection of the model coefficients in a fault space, showing clear separation between FUL and threshold catalysts. The normalized error vector coefficients, depicted in Fig. 5, demonstrate the significant separation between healthy and threshold cases as well. Also from this figure, a distinct direction of the fault, drawn between the center of gravities of the healthy and threshold regions, is evident.

For this reduced diagnostic model the value of ε of Eq. (17) is 3 and a fault is now considered present when

$$\tilde{M} = \|E\|_2 - 3 > 0 \tag{35}$$

and when the projection of E onto Dir_2, denoted as Ψ_3, is bounded as

$$2 \leq \Psi_3(j) \leq 10 \tag{36}$$

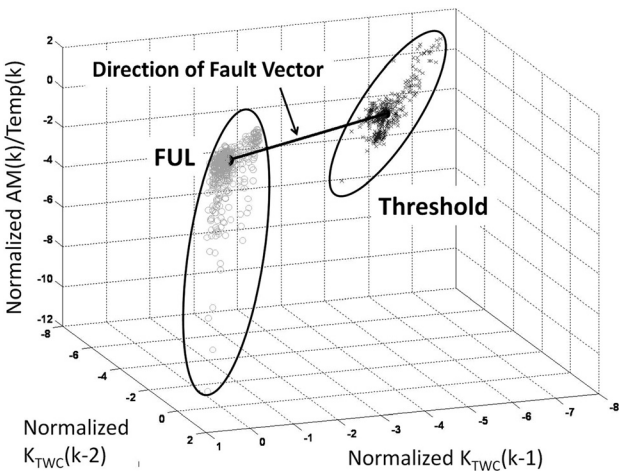

Fig. 5 Reduced model normalized error vector coefficients shown with direction of fault

where

$$\tilde{\text{Dir}} = [\,-0.91 \quad -0.37 \quad 0.20\,]^T \tag{37}$$

When this reduced order health monitoring model was applied to the same procedure outlined in the previous section, similar results to the full model set were found. Every window segment for both healthy and threshold TWCs was properly classified as such when coefficient estimates where applied to the criterion of Eqs. (35) and (36).

Conclusion

Developed is a model based diagnostic method for automobile three-way catalysts health. The method is founded on automated processes for model structure identification; model parameters; and fault detection, isolation, and estimation (TWC classification). The results and analysis presented have shown this goal to be accomplished. The theory behind catalyst deactivation, oxygen storage phenomenon, catalyst dynamics, and modeling techniques have been described in detail and experimentally validated. The method was proven using actual vehicle data, equipped with different aged catalysts, generated on the federal test procedure chassis dynamometer drive cycle.

The orthogonal least squares method proved to be reliable in terms of choosing model regressors and the four-step instrumental variables method proved effective for estimating parametric model coefficients. The accuracy of these techniques was proven, even in case of noise corrupted data, using benchmark TWCs. The information synthesis technique proved to be a simple method to mine the model coefficients for diagnostic information.

One main advantage of this diagnostic method is that it does not require specific operating windows or intrusive fueling control to effectively diagnosis the health of the TWC; features that are extremely advantageous in on-board diagnosis. The recursive nature of the IV4 and IS processes lends themselves to online identification and diagnosis.

Acknowledgment

This research was performed as part of a University of Houston and Ford partnership where mechanical engineers in graduate school take chemical engineering courses. This work was also partially supported by the National Science Foundation Grant Number 0727999. *Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.*

References

[1] EPA Report, "Control of Air Pollution From Motor Vehicles and New Motor Vehicle Engines; Modification of Federal On-board Diagnostic Regulations for Light-Duty Vehicles and Light-Duty Trucks; Extension of Acceptance of California OBD II Requirements," Vol. **63**, No. 245, December 22, 1998. Report Number is 63 FR 70681.

[2] Sideris, M., *Methods for Monitoring and Diagnosing the Efficiency of Catalytic Converters* (Elsevier, New York, 1998).

[3] Makki, I., Kerns, J., Uhrich, M., and Kluzner, M., 2007, "Control and Diagnostic Approach for Emission Control," U.S. Patent No. 2007/0256406 A1.

[4] Jerger, R., 1999, "Catalyst Monitor Using Arc Length Ratio of Pre- and Post-Catalyst Sensors Signals," U.S. Patent No. 5,899,062.

[5] Brandt, E., Wang, Y., and Grizzle, J., 2000, "Dynamic Modeling of a Three-Way Catalyst for SI Engine Exhaust Emission Control," IEEE Trans. Control Systems Technol., **8**(5) 767–776.

[6] Brandt, E., Wang, Y., and Grizzle, J., 2001, "Three-Way Catalyst Diagnostics for Advanced Emissions Control Systems," *Proceedings of the American Control Conference*, Vol. **5**, pp. 3305–3311.

[7] Hepburn, J., and Gandhi, H., 1992, "The Relationship Between Catalyst Hydrocarbon Conversion Efficiency and Oxygen Storage," SAE Paper 920831.

[8] Heck, R., and Farrauto, R., 2002, *Catalytic Air Pollution Control*, 2nd ed., Wiley, New York.

[9] Dawson, B., Franchek, M., and Grigoriadis, K., 2007, "Data Driven Simplified Three-Way Catalyst Health Diagnostic models: Experimental Results," *Proceedings of the Dynamic Systems and Control Conference*, ASME IMECE2007- 42281.
[10] Lassi, U., 2003, "Deactivation Correlations of Pd/Rh Three-Way Catalysts Designed for EURO IV Emission Limits," Dissertation, Oulu University Press.
[11] Makki, I. H., Kerns, J., Olson, B., Reed, D., and Smith, S., 2003, "UEGO Based A/F Controls for SULEV/PZEV Tailpipe Emissions Applications," Ford Research and Advanced Engineering Technical Rep., SRR–2003–0072.
[12] Korenberg, M., 1985, "Orthogonal Identification of Nonlinear Difference Equation Models," *Midwest Symposium on Circuits and Systems*. pp. 90–95.
[13] Billings, S., Korenberg, M., and Chen, S., 1988, "Identification of Non-Linear Output-Affine Systems Using an Orthogonal Least-Squares Algorithm," Int. J. Syst. Sci., **19**(8). pp. 1559–1568.
[14] Ljung, L., 1999, *System Identification: Theory for User*, 2nd ed., Prentice-Hall, Englewood Cliffs, NJ.
[15] Franchek, M., Buehler, P., and Makki, I., 2007, "Intake Air Path Diagnostics for Internal Combustion Engines," J. Dyn. Syst., Measure., Control, **129**(1), pp. 32–40.