Incorporating a Model of Vehicle Dynamics in a Diagnostic System for Steer-by-Wire Vehicles

Christopher D. Gadda **Stanford University**
Paul Yih **Stanford University**
J. Christian Gerdes **Stanford University**

Dept. of Mechanical Engineering
Stanford, California 94305-4021, USA
Phone: 650-906-6417
Fax: 650-723-3521
Email: cgadda@stanford.edu

This paper examines the benefits of incorporating vehicle dynamics modeling into the design of a diagnostic system for a steer-by-wire vehicle. The use of a model of vehicle dynamics improves the speed and accuracy of the diagnoses, by eliminating a significant source of disturbance input to the steer-by-wire system model. A method for reducing the effects of modeling uncertainty on diagnostic system performance based on spectral fault characteristics is also presented. The techniques discussed are demonstrated on an experimental steer-by-wire vehicle.

Topics / Vehicle Diagnostics, Active Safety

1. INTRODUCTION

The numerous benefits of steer-by-wire technology have yet to be realized in a production vehicle. The potentially catastrophic nature of a steering system failure mandates that any practical steer-by-wire system be extremely reliable. One approach for achieving the necessary level of reliability incorporates a fast and accurate diagnostic system that can isolate a fault and respond by switching to a redundant component or a modified control law that can accommodate the fault.

The area of drive-by-wire research draws from the years of work done on fly-by-wire technologies, diagnostic systems included, and many of the results apply directly. A probabilistic analysis of the failure rates of steer-by-wire systems using various forms of redundancy coupled with diagnostic techniques described in [1], shows that steer-by-wire systems can be designed to have an overall reliability rate of 10^{-9} failures/hour, the same as imposed on the aviation industry. But diagnostic systems for aircraft have certain design freedoms that are not available to those for ground vehicles. The expense of triply redundant sensors, actuators, and controllers, all common practices in fly-by-wire designs, are prohibitive in production automobiles. Furthermore, aircraft are typically tens of seconds or more from any possible source of collision, so diagnostic systems for aircraft have ample time to correctly identify the source of a fault and choose appropriate action. This is clearly not the case in an automobile where the decision must be made much more quickly to prevent a collision.

Previous work in this area includes a baysian network approach to diagnostic systems [7], [8]. In [5], a technique for diagnosing actuator faults in the presence of unknown disturbances and modelling errors is developed. Isermann presents a discussion of many of these diagnostic techniques in [2], [3], and their use in fault-tolerant drive-by-wire systems in [4].

In any practical diagnostic system, noise and modeling uncertainty ultimately limit the accuracy and speed of fault detection and isolation. In the case of a steer-by-wire system, the lateral forces on the road wheels create disturbance torques on the steering motor. If the diagnostic system is designed without knowledge of the vehicle dynamics, these disturbance torques can only be regarded as a noise source. This reduces the accuracy of the diagnoses and/or increases the fault detection time. By accurately modeling the vehicle dynamics and estimating the state of the vehicle, these disturbance torques can be included in the model and need not adversely affect performance.

We use simple linear models of the steering system and vehicle dynamics in the design of an example diagnostic system for a steer-by-wire vehicle. The imperfections in the models used result in a diagnostic system that produces residuals that respond to the driver command signal, even when there is no fault present. This effect is addressed through the use of residual filters designed to exploit the differences in spectral content between the resid-

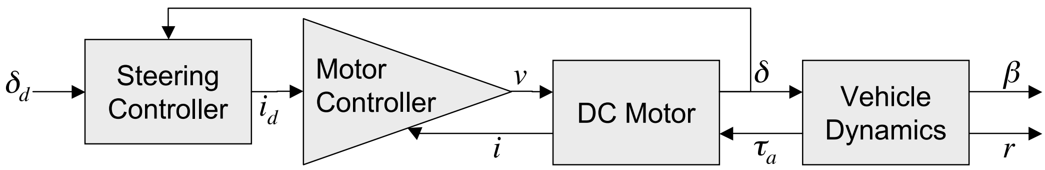

Fig. 1: Steer-by-wire system block diagram

uals produced by an actual fault and those which result from modeling inaccuracies. Experimental results collected on a steer-by-wire vehicle demonstrate the effectiveness of the techniques presented.

2. DIAGNOSTIC SYSTEM

2.1 Fault Types

While there exist many ways to classify different failure modes for a steer-by-wire system, one particularly useful classification approach is the simple dichotomy of time scale: gradual faults and sudden faults. Gradual faults occur as components wear out or overheat, but do not require immediate detection, since by definition they are not changing very quickly. Fault detection techniques which are suitable for detecting gradual faults can operate over a longer time scale. A sudden fault, however, such as a wiring harness failure may require immediate corrective action to maintain vehicle stability and driver control. Thus, detection time is of critical importance for sudden faults.

Techniques which focus on parameter estimation or adaptive filtering such as recursive least-squares, extended Kalman filtering, or instrumental variables are well suited for the detection of a gradual fault, but may result in detection times which are unacceptably long in response to a sudden fault. The focus of this paper is on detection and isolation techniques for sudden faults, which when used in parallel with parameter estimation techniques will provide a system capable of diagnosing faults that develop over any time scale.

2.2 System Description

We develop an example diagnostic system for a 1997 Corvette which has been adapted to steer-by-wire. A block diagram of the system is shown in figure 1. The steer-by-wire system uses a DC motor with a gearbox to drive the pinion of the original power steering system, so the following linear model of the motor dynamics is used:

$$\dot{x_m} = A_m x_m + B_m \begin{bmatrix} i & \tau_a \end{bmatrix}^T \tag{1}$$

where

$$\begin{aligned} x_m &= \begin{bmatrix} \delta & \dot{\delta} \end{bmatrix}^T \\ A_m &= \begin{bmatrix} 0 & 1 \\ 0 & -\frac{b_m}{J_m} \end{bmatrix} \\ B_m &= \begin{bmatrix} 0 & 0 \\ \frac{r_s r_p r_g \eta k_m}{J_m} & -\frac{1}{J_m} \end{bmatrix} \end{aligned}$$

where δ is the steering angle, J_m is the effective moment of inertia of the steering system, b_m is the effective damping of the steering system, r_s is the steering ratio, r_p is the torque amplification factor of the power steering system, r_g is the gear ratio of the gearbox connecting the DC motor to the pinion, η is the combined efficiency of the motor and gearbox, and k_m is the motor constant relating torque to current. The inputs to this model are the current to the motor, i, and the aligning torque, τ_a.

The aligning torque is a function of the vehicle state and represents a significant disturbance to the steer-by-wire system. In the absence of a suitable model of vehicle dynamics, the aligning torque could be modeled as a stochastic process or a norm-bounded unknown signal, but due to the magnitude of the aligning torque, the resulting diagnostic system would be very insensitive to faults. For this reason, we incorporate a model of vehicle dynamics into the design of our diagnostic system.

The planar dynamics of the vehicle are modeled using the bicycle model, where left and right tire forces are considered in aggregate. Small angle approximations are used and lateral tire force is assumed to be proportional to the tire slip angle, so that a linear model of the planar vehicle dynamics is developed, given by the following:

$$\dot{x_v} = A_v x_v + B_v \delta \tag{2}$$

where

$$\begin{aligned} x_v &= \begin{bmatrix} \beta & r \end{bmatrix}^T \\ A_v &= \begin{bmatrix} -\frac{C_0}{mV} & -1 + \frac{C_1}{mV^2} \\ \frac{C_1}{I_z} & -\frac{C_2}{I_z V} \end{bmatrix} \\ B_v &= \begin{bmatrix} \frac{C_f}{mV} \\ \frac{C_f a}{I_z} \end{bmatrix} \end{aligned}$$

and

$$\begin{aligned} C_0 &= C_{\alpha f} + C_{\alpha r} \\ C_1 &= C_{\alpha r} b - C_{\alpha f} a \\ C_2 &= C_{\alpha f} a^2 + C_{\alpha r} b^2 \end{aligned}$$

where β is the sideslip angle, r is the yaw rate, I_z is the polar moment of inertia of the vehicle, $C_{\alpha f}$ and $C_{\alpha r}$ are the front and rear cornering stiffnesses, respectively, a and b are the distances from the center

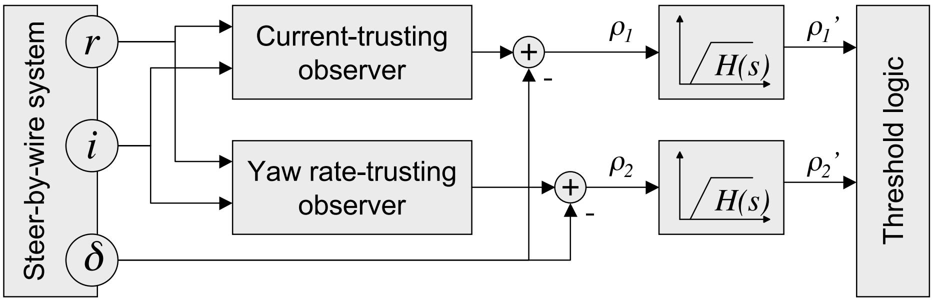

Fig. 3: Diagnostic system block diagram

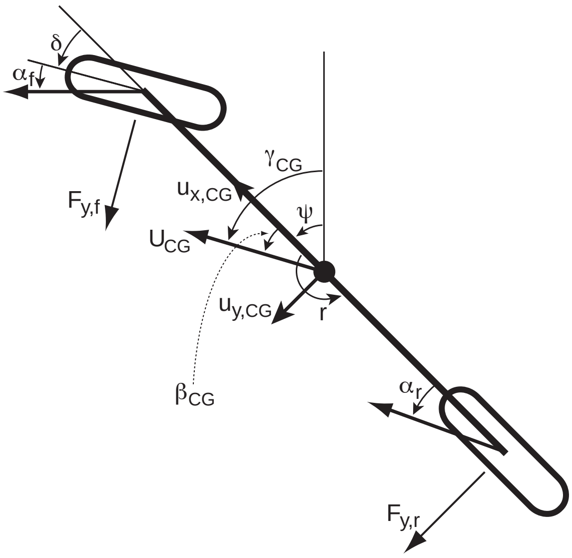

Fig. 2: Bicycle model

of gravity to the front and rear axles, respectively, m is the mass of the vehicle, and V is the vehicle velocity.

The aligning torque is related to the vehicle state by the following equation:

$$\tau_a = -C_{\alpha f}(t_p + t_m)(\beta + \frac{a}{V}r - \delta) \qquad (3)$$

where t_p and t_m are the pneumatic and mechanical trail of the tire, respectively. In order to arrive at a linear model, t_m and t_p are assumed to be constant and known, as are $C_{\alpha f}$ and $C_{\alpha r}$. Combining (1), (2), and (3) yields the following state-space model of the vehicle and steering system dynamics:

$$\dot{x} = Ax + Bi \qquad (4)$$

where

$$x = \left[\begin{array}{cccc} \beta & r & \delta & \dot{\delta} \end{array} \right]^T$$

$$A = \left[\begin{array}{cccc} -\frac{C_0}{mV} & -1 + \frac{C_1}{mV^2} & \frac{C_{\alpha f}}{mV} & 0 \\ \frac{C_1}{I_z} & -\frac{C_2}{I_z V} & \frac{aC_{\alpha f}}{I_z} & 0 \\ 0 & 0 & 0 & 1 \\ \frac{C_3}{J_w} & \frac{aC_3}{J_w V} & -\frac{C_3}{J_w} & -\frac{b_w}{J_w} \end{array} \right]$$

$$B = \left[\begin{array}{c} 0 \\ 0 \\ 0 \\ \frac{r_s r_p k_M r_g \eta}{J_w} \end{array} \right]$$

and

$$C_3 = (t_p + t_m)C_{\alpha f}$$

A block diagram presented in figure 3 shows the basic structure of the diagnostic system. Measurements of yaw rate, motor current, and steering angle are provided to a pair of state observers, each of which computes separate estimates of the state variables. From each of these state estimates, the steer angle estimates are compared against the measured steer angle to produce two residuals. These residuals are then filtered (as described in section 2.4) and then compared against thresholds to detect and isolate a fault.

Each of the observers have the following structure:

$$\dot{\hat{x}} = A\hat{x} + Bi + L(\hat{r} - r) \qquad (5)$$

The gain vector L is chosen such that the observer dynamics will be stable and "trust" either the measured current or the measured yaw rate, while "distrusting" the other. This is accomplished by solving the steady-state continuous-time Kalman filter design problem with process and sensor noise weighting factors adjusted appropriately. To create the observer which depends heavily on the measured current, the process noise variance is set to be 1/10,000 the value of the sensor noise variance. To create the observer which depends heavily on the measured yaw rate, the process noise variance is set to be 10,000 times the value of the sensor noise variance. This technique obviates the difficult process of choosing pole locations for a pole-placement approach or manually adjusting gains. Furthermore, it always results in a stable observer. It should be noted that the process and sensor noise variance values are not chosen based on actual noise models for either sensor, and the resulting observer is not an optimal state estimator. The noise variances are simply a convenient way of tuning the resulting observer's sensitivity to a particular input.

By using two separate observers, one which depends primarily on the yaw rate measurement and one that depends primarily on the motor current measurement, a single fault can be both detected and, to an extent, isolated. When the steer-by-wire system is working correctly, the observers will converge to accurate state estimates and both residuals will be small. If a fault occurs which affects the yaw rate sensor the residual which relies on yaw rate, ρ_2, will have much larger response than the residual which relies on the motor current, ρ_1. When a fault occurs which affects the motor current sensor or some aspect of the motor model, such as a change in friction or motor constant, ρ_1 will have a much larger response than ρ_2. If a fault occurs which affects the steering angle sensor, both ρ_1 and ρ_2 will have a significant response.

With this simple example diagnostic system there are only two residuals, hence only four possible diagnoses using this technique. In a practical application of this approach, additional residuals would be formed from other measurements, increasing the variety of faults that can be detected and the degree to which a fault condition can be isolated. The technique presented here focuses on the diagnosis of the motor, as an accurate diagnosis of the motor facilitates the diagnosis of other components of the steer-by-wire system.

2.3 Modeling Uncertainty

Including a model of the vehicle dynamics in the models used by the diagnostic system significantly reduces the uncertainty in the aligning torque τ_a, but also creates a dependence on the accuracy of the vehicle dynamics model. Several factors limit the accuracy of this model. The dynamics are linearized and thus cannot accurately represent coulomb friction in the steering system, nor tire saturation near the limits of handling. Also, the model depends on a number of parameters such as vehicle mass, vehicle center of gravity, or tire cornering stiffness, which may not be known very precisely, or may change over time.

System identification techniques, such as recursive least-squares or instrumental variables can be used to estimate the various uncertain parameters in the model. These parameter estimates may already be available as the result of a portion of the diagnostic system looking for gradual faults. This approach will reduce modeling errors, but even with perfect parameter estimates, nonlinear and other unmodeled dynamics still limit how accurately the models will be able to predict the behavior of the actual system. This, in combination with sensor noise, will limit how small the residuals will be for an unfaulted system, which in turn limits the sensitivity of the diagnostic system to an actual fault.

2.4 Residual Filtering

Filtering the residuals before comparing them to a set of fixed thresholds is another approach for handling modeling errors. If the residuals for the un-faulted system are too large, linear filters can be used to remove energy in select frequency ranges to reduce the overall amplitude of the residual. When the frequency content of an un-faulted residual differs significantly from that of the residual when a fault is occurring, residual filtering can improve detection performance significantly.

The input to the actual system is primarily the driver command, which generally does not have significant frequency content above ~ 2 Hz. The actual system and the observers are driven by this same signal, so the component of the residuals formed due to model inaccuracies is expected to also have little frequency content above ~ 2 Hz. A suddenly occurring fault, however, produces a broad range of frequency content in the residuals. This spectral difference between faulted and unfaulted residuals makes it possible to design linear filters which reject the frequency content of a nominal residual while passing the frequencies associated with a suddenly occurring fault, thereby reducing the effects of model inaccuracy and increasing the sensitivity of the fault detection. While by no means optimal, a simple second-order high-pass filter with a cutoff frequency of 10 rad/s, given by the following transfer function, significantly improves the performance of the diagnostic system:

$$H(s) = \frac{s^2}{s^2 + 20s + 100} \tag{6}$$

This filter technique has some disadvantages as well. The use of a filter which eliminates the DC component of a signal means that all residuals will return to zero a short time after a fault occurs, even if the fault is still present. Also, the use of high-pass filtering renders the diagnostic system insensitive to faults which develop on a time-scale below the cutoff frequency of the filters. As mentioned previously, the technique presented here is only one part of complete system and must be used in combination with other techniques to detect more gradually occurring faults.

3. SIMULATION

This technique is demonstrated in simulation for a vehicle traveling at a constant speed of 6.5 m/s for two different types of fault. Figure 4 shows the residuals for a simulated bias shift of 1° on the steering angle sensor. The horizontal dotted lines represent the detection threshold level. The simulated fault occurs at 5 seconds into the simulation, and as expected, both residuals respond immediately. Figure 5 shows the residuals for a simulated bias shift of 2.8°/s on the yaw rate sensor. Again, the simulated

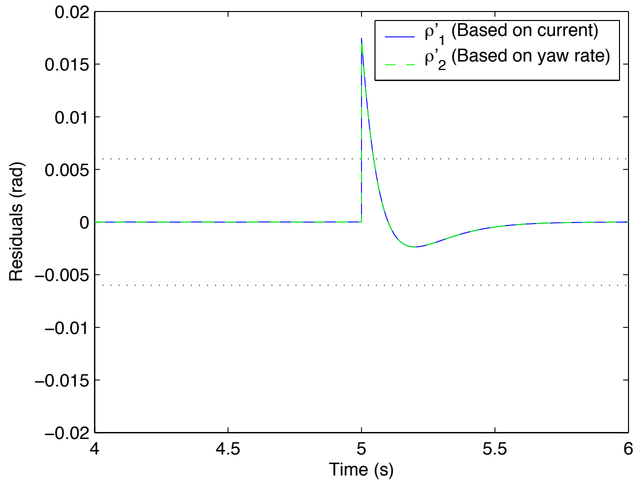

Fig. 4: Simulated residuals for 1° steer angle bias shift

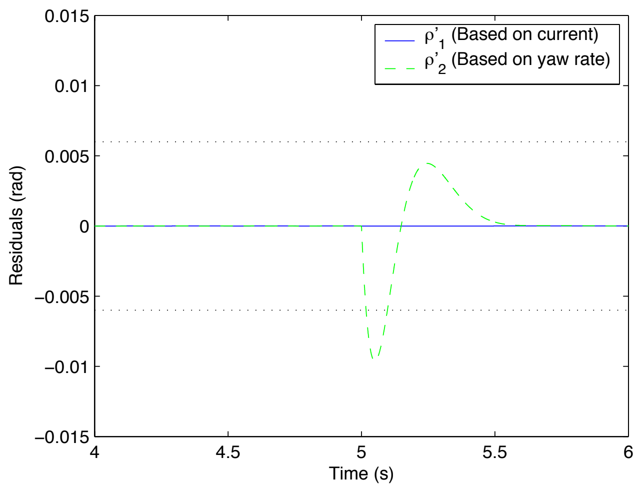

Fig. 5: Simulated residuals for 2.8°/s yaw rate bias shift

fault occurs at $t = 5$ seconds, but in this case only the yaw rate-based residual, ρ'_2 responds. In both cases, the affected residuals return to zero within roughly five time constants of the high-pass filter, even though the simulated faults persist.

4. EXPERIMENT

Using the steer-by-wire vehicle and software simulation of fault conditions, we collected experimental data demonstrating the techniques discussed here. All of the tests were performed with the vehicle moving at a roughly constant speed of approximately 6.5 m/s, while being piloted through a slalom course. The peak lateral acceleration during these tests was kept under 5 m/s^2, so the linear tire model approximation is valid. Two types of faults were simulated by modifying the software controlling the steer-by-wire system: a 1° steering angle sensor bias shift and a 2.8°/s yaw rate sensor bias shift.

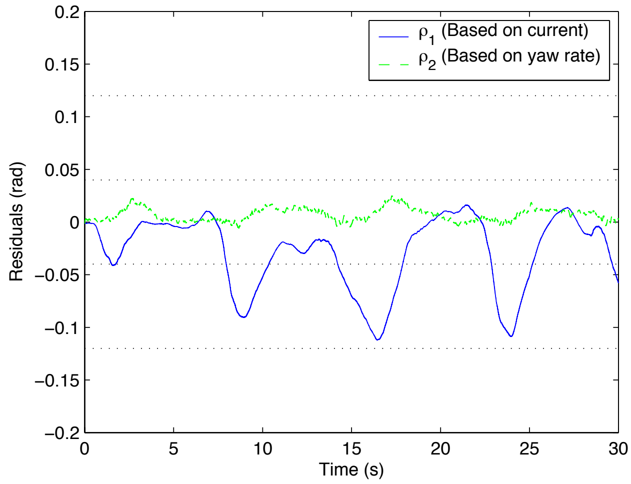

Fig. 6: Residuals for un-faulted system

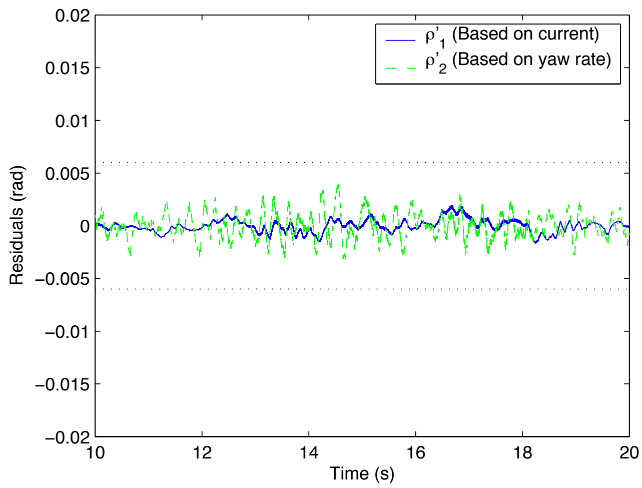

Fig. 7: Filtered residuals for un-faulted system

Figure 6 shows the unfiltered residuals collected during an experiment where no simulated faults occurred. In both residuals, the driver command signal can be seen; it is particularly noticeable in the current-based residual, ρ_1. The horizontal dotted lines represent threshold levels for each residual, chosen such that no false positives occurred during any of our experiments. In figure 7 the filtered residuals for the same experiment are plotted, showing how effective the high-pass filtering is at eliminating the effects of model inaccuracies. In this plot a new threshold level is shown, also chosen such that no false positives were recorded. This threshold is an order of magnitude lower than the thresholds used in the unfiltered case.

Figure 8 shows the residuals during a simulated 1° steering angle bias shift. Just as in simulation both residuals have an immediate response, then quickly return to zero. In figure 9 a simulated 2.8°/s yaw rate bias shift is shown only affecting ρ'_2, which also returns quickly to zero.

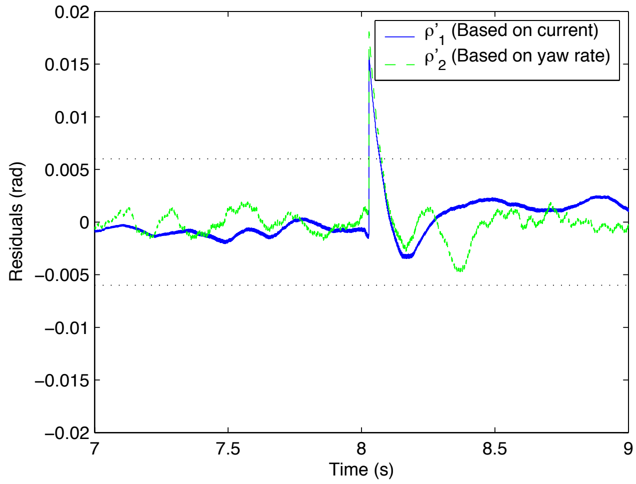

Fig. 8: Filtered residuals for 1° steer angle bias shift

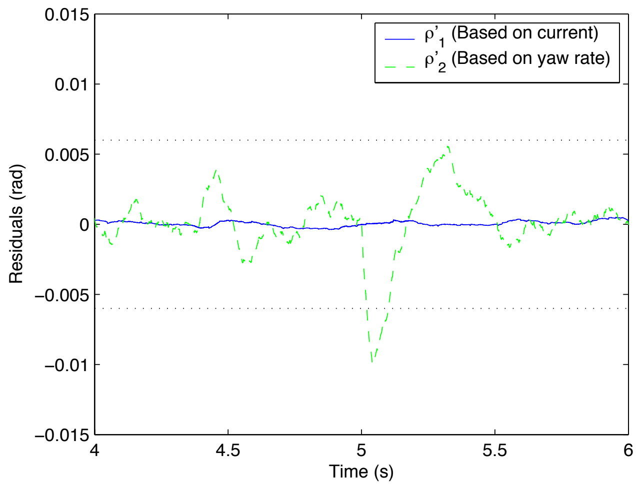

Fig. 9: Filtered residuals for 2.8°/s yaw rate bias shift

5. CONCLUSION

These results illustrate the effectiveness of a diagnostic system based on vehicle dynamics models and residual filtering in detecting suddenly occurring faults. The experiments show how modeling inaccuracies can be readily accommodated when the spectral content of a fault differs significantly from that of normal system responses. This represents only a component of a complete diagnostic system, as it is designed to work in parallel with a system which can diagnose a gradually worsening fault condition. The successful combination of diagnostic techniques which can accurately isolate sudden faults quickly and gradual faults before they become serious enough to affect vehicle handling is an avenue for further research.

6. ACKNOWLEDGMENTS

The authors would like to acknowledge Nissan Motor Corporation for sponsoring research in diagnostic systems for steer-by-wire vehicles, with special thanks to Toshimi Abo, Kazutaka Adachi, Takeshi Mitamura, Dr. Kimio Kanai, and Masaharu Asano for their support of this project.

REFERENCES

[1]. Hammett, R. C., Babcock, P. S.: "Achieving 10^{-9} Dependability with Drive-by-Wire Systems", Society of Automotive Engineers (SAE) World Congress. Detroit, MI (2003).

[2]. Isermann, R., and Ballé, P.: "Trends in the Application of Model Based Fault Detection and Diagnosis of Technical Processes", In Proc. of IFAC 13th Triennial World Conference (1996).

[3]. Isermann, R.: "Process fault detection based on modeling and estimation methods-a survey", Automatica; July 1984; vol.20, no.4, pp. 387-404 (1984).

[4]. Isermann, R.: "Diagnosis Methods for Electronic Controlled Vehicles", 5th International Symposium on Advanced Vehicle Control, Ann Arbor, Michigan (2000).

[5]. Park, T. G., Ryu, J.S., Lee, K. S.: "Actuator fault estimation with disturbance decoupling", IEE Proceedings-Control Theory and Applications; Sept. 2000; vol.147, no.5, pp. 501-8 (2000).

[6]. Patton, R. J., Chen, J.: "Observer-based fault detection and isolation: robustness and applications", Control Engineering Practice; May 1997; vol.5, no.5, pp. 671-82 (1997).

[7]. Schwall, M. L., and Gerdes, J. C.: "Multi-Modal Diagnostics for Vehicle Fault Detection", In Proceedings of IMECE 2001, DSC−24600 (2001).

[8]. Schwall, M. L., and Gerdes, J. C. "A Probabilistic Approach to Residual Processing for Vehicle Fault Detection", Proceedings of the 2002 ACC, pp. 2552-7 (2002).

[9]. Willsky, A. S.: "A Survey of Design Methods for Failure Detection in Dynamic Systems" Automatica, Vol. 12, pp. 601-11 (1976).

[10]. Yih, P., Gerdes, J. C.: "Modification of Vehicle Handling Characteristics via Steer-by-Wire" Proceedings of the 2003 ACC, Denver, CO (2003).