AAS 19-103

SENSOR ANALYSIS, MODELING, AND TEST FOR ROBUST PROPULSION SYSTEM AUTONOMY

Jeb S. Orr*

An approach is presented supporting analysis, modeling, and test validation of operational flight instrumentation (OFI) that facilitates critical functions for the Space Launch System (SLS) main propulsion system (MPS). Certain types of OFI sensors were shown to exhibit highly nonlinear and non-gaussian noise characteristics during acceptance testing, motivating the development of advanced modeling and simulation (M&S) capability to support algorithm verification and flight certification. Hardware model and algorithm simulation fidelity was informed by a risk scoring metric; redesign of high-risk algorithms using test-validated sensor models significantly improved their expected performance as evaluated using Monte Carlo acceptance sampling methods. Autonomous functions include closed-loop ullage pressure regulation, pressurant leak detection, and fault isolation for automated safing and crew caution and warning (C&W).

1 INTRODUCTION

The Space Launch System (SLS) is NASA's next-generation exploration-class launch vehicle for large-scale crewed and uncrewed space access, including such objectives as human transit to Mars, rendezvous with near-earth asteroids, and the launch of unmanned probes to distant solar system targets such as Europa. Its design provides for a level of performance and reliability that is unmatched in any existing or planned launch system, including an ability to loft approximately 26 metric tons to trans-lunar injection (TLI) in its initial Block 1 configuration (Figure 1). Its evolved configurations, Block 1B and Block 2, utilize the Exploration Upper Stage (EUS) to substantially increase performance. The Block 1B with EUS has a cargo payload performance capability of approximately 37 metric tons to TLI. The Block 2, more than 115 meters long and using upgraded solid rocket motors (SRMs), is able to loft 130 metric tons to low Earth orbit (LEO) or 45 metric tons to a heliocentric orbit. These capabilities place the SLS in a category of performance commensurate with that of the Saturn V.

The SLS leverages hardware, processes, and design concepts derived from the Space Shuttle program, including an 8.4 m diameter core stage containing more than 2.7 million liters of cryogenic propellants.[1] The core stage is powered by four RS-25E liquid engines derived from the highly successful Space Shuttle Main Engine (SSME), each producing about 2.17 MN of thrust.[2] Additional thrust is provided by two 5-segment Reusable Solid Rocket Motor-V boosters (RSRMVs). Each RSRMV provides a peak sea level thrust of about 14.6 MN and provides primary ascent propulsion during the 126-second boost phase.

*Space Launch System Flight Dynamics and Control Technical Specialist, NASA Marshall Space Flight Center / EV41 Control Systems Design and Analysis Branch (Mclaurin Aerospace - Jacobs ESSCA), Huntsville, AL

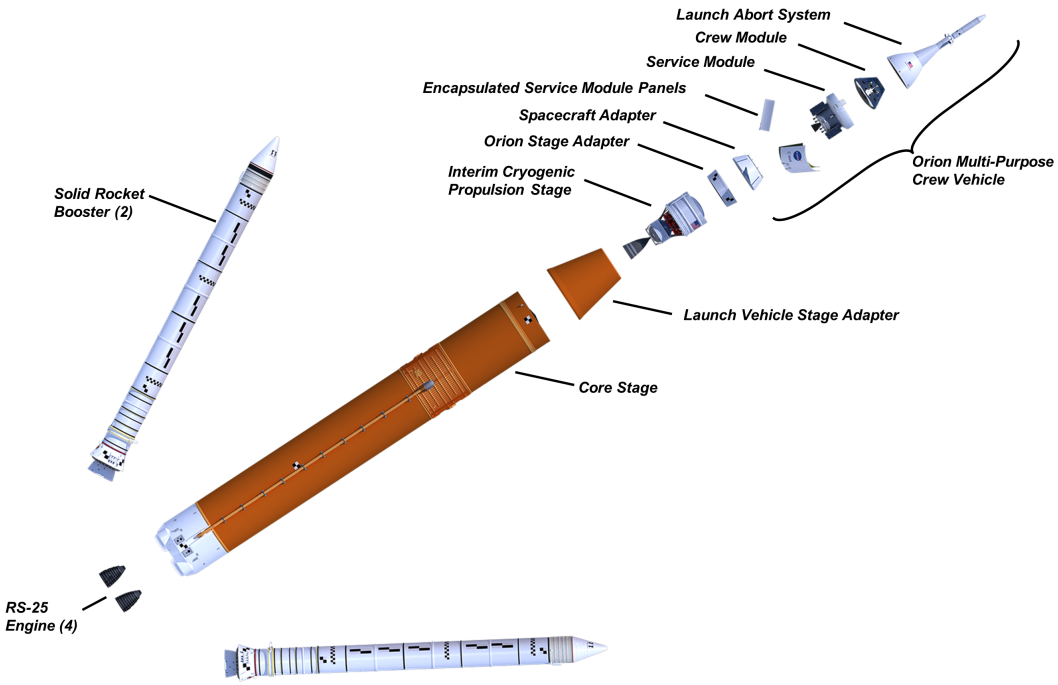

Figure 1. Space Launch System Block 1 Configuration (NASA)

Importantly, SLS is a NASA human-rated system, with a commensurate level of reliability and attendant subsystem complexity. The SLS Main Propulsion System (MPS) includes the RS-25 engines, the liquid hydrogen (LH2) and liquid oxygen (LO2) propellant tanks, the gaseous helium (GHe) pressurization system, and numerous propellant flow paths. Engine operation, including mixture ratio and throttle level, is regulated by the RS-25 engine controller.[3]

In contrast, fault tolerance, inlet condition regulation, and supervisory functions (such as caution and warning) are performed in flight software (FSW) by the Vehicle Management (VM) subsystem of the SLS avionics. Redundancy sets of Combined Control System Electronics (CCSE) and Flight Computers (FCs) rely on MPS Operational Flight Instrumentation (OFI) in order to determine the state of the propellant tanks and in turn open/close pressurization valves, identify faulted sensors or functions, and alert crew and mission operations personnel in the event of an anomaly. Importantly, these FSW functions are also used for largely autonomous execution of the upcoming green run hot-fire test of the first integrated SLS core stage at the Stennis Space Center (SSC) B test complex. These critical measurements must be available and reliable in order to achieve nominal MPS operation and ensure successful mission execution.

2 RISK PRIORITIZATION

The sensors supporting the SLS MPS are the subject of this paper. Early in the integration of the MPS hardware with the SLS avionics, it was recognized that limitations in the fidelity or availability of detailed MPS OFI sensor models could increase risk of unmodeled or unexpected interactions or failure modes that could not be identified via simulation-based analysis and verification. More than 100 OFI sensors are used to support the SLS MPS operation and mission management functions, but not all sensor systems are involved in critical functions.

Sensor Model Fidelity	Description	Model Score s_m
High Fidelity	Existing contractor-vetted DM supported by test data delivered as software module with defined APIs	0.10
Medium Fidelity	Contractor-vetted DM is based only on analysis and not supported by test	0.25
Limited Fidelity	NASA internally developed model based on data derived from engineering specifications and other sources	0.50
Low Fidelity	No model implemented but sufficient data and/or experience exists to produce one internally or via action to contractor	0.75
No Model	Insufficient data exists to characterize sensor performance and no model is available ($s_a = 1$) by default for this case)	1.00

Table 1. Sensor Model Fidelity Subcategories

In the evaluation of MPS subsystem risk, a scoring metric was developed to prioritize the identification of gaps in the understanding of MPS sensor interactions with flight software and its consequences on other systems and functions. The majority of critical sensors used for SLS MPS VM functions are pressure transducers, including the Active Electronics Pressure Transducers (AEPTs), the Passive Electronics Pressure Transducers (PEPTs), Differential Pressure Transducers (DPTs), and Ambient Pressure Transducers (APTs). In addition, MPS OFI includes the Cryogenic Level Sensor System (CLSS) and various temperature transducers such as the Immersion Temperature Probe Assemblies (ITPAs) and Resistance Temperature Detectors (RTDs).

In the development of the scoring metric, each device was assigned a model fidelity score s_m, a model availability score s_a, and a model criticality score s_c. Risk quantification was calculated as the product of the scoring factors, $r = s_c \times s_a \times s_m$. Categorical scores were assigned *ad hoc* to multiple subcategories of each scoring metric. While the allocation of risk factors to each scoring subcategory was empirical and based upon past program experience, the allocation of specific sensors and sensor functions to those subcategories was based upon strictly defined evaluation criteria, enabling a relatively objective assessment across the entire suite of MPS sensor functions. It is recognized that there are potentially infinite methods to weight some combination of scoring factors s_i. It was determined that this simple multiplicative model was adequate to assess risk and agreed well with subject matter expert (SME) insight.

Model fidelity subcategories are shown in Table 1. Model fidelity scoring is contingent primarily on whether the sensor model in question has achieved Design Model (DM) certification as per the Space Launch System Program (SLSP) Design Model Delivery Standard, SLS-STD-038, and has been supported by test data. Also important for consideration is whether the model is delivered in a format having defined, documented, and compatible application programming interfaces (APIs) for integration with real-time (e.g., hardware-in-the-loop) avionics simulation facilities.

Most models evaluated for MPS functions were determined to be in the Limited Fidelity category. In this case, and in part owing to NASA's responsibility for VM flight software, NASA leveraged reasonable engineering assumptions in the development of models for algorithm design based on sensor hardware performance requirements that were flowed to vendors for parallel hardware development. While designers endeavor to produce algorithms that are robust to variations

Sensor Model Availability	Description	Model Score s_a
Production Ready	Model implemented in all relevant design and verification simulations with V&V process completed and documented and can support all SLS operations	0.10
Limited	Model not implemented in all simulations or only for algorithm design; V&V not completed or documented; does not support all operational use cases	0.50
Not Implemented	Model is not implemented or tested for design or verification	1.00

Table 2. Sensor Model Availability Subcategories

in actual sensor performance, some risk is accepted that final acceptance and/or qualification tests of actual sensor performance will not meet requirements, and either costly hardware redesign or algorithm modifications will be necessary to ensure performance. More commonly, early subsystem requirements have insufficient detail to capture the necessary performance specifications, and vendors proceed with designs isolated from the actual engineering needs of the system.

Model availability subcategories are shown in Table 2. Model availability scoring is a measure of the relative maturity of the integration of a hardware performance model into simulation environments that are used for design verification and flight certification. A Production Ready model supports design simulations, verification simulations, multiple hardware-in-the-loop environments, and supports all use cases including its use for ground systems functions and off-nominal scenarios. Models seldom reach this level of maturity prior to the commencement of flight operations when flight data can be used to resolve observed anomalies and validate preflight math models. Most models of MPS OFI evaluated for SLS were in the Limited category.

Finally, the criticality of each sensor function is scored in Table 3. Criticality scores range from inconsequential ($s_c = 0.1$) to safety critical ($s_c = 1.00$). Due to the redundancy and fault detection design paradigm that satisfies the human rating requirements of the SLS vehicle, no single vehicle management MPS sensors were classified as safety critical since typically at least two sensor faults in the same subsystem can be tolerated before a functional impact results. In contrast to sensors used directly by the RS-25 engine controller, MPS OFI are associated with typically low-bandwidth processes (such as ullage pressure regulation) that have long times-to-criticality and are backed up by redundant hardware fail-safes (such as pressure relief valves). However, the ability to safely continue the mission may be impacted in the event that an incorrect automated or human decision is made, or contributed to, on the basis of faulty sensor data. These factors contributed to many OFI functions being placed in the Mission Impact or Mission Critical categories.

3 MODELING AND ANALYSIS

The development of advanced sensor models and additional statistical assessment of certain functions was informed by the risk analysis described above. The Active Electronic Pressure Transducer (AEPT) device is of particular significance due to its use for closed-loop ullage pressure regulation, engine helium supply leak detection, and other caution and warning functions. Early assessments of the software design indicated that a false positive He leak detection could inadvertently trigger an automated advance-to-shutdown safing action during the full-scale green run hot fire test, and while not a threat to safety, this automated safing action was to be avoided due to the high operational

Sensor Criticality	Description	Model Score s_c
Inconsequential	Sensor classified as OFI but not used for any software functions other than logging or telemetry	0.10
Operational Concern	Used in software functions but only for generating C&W and no automated actions	0.25
Mission Impact	Incorrect interpretation of sensor data may affect level of redundancy or operability of non-mission critical components, but will not compromise performance	0.50
Mission Critical	Incorrect interpretation of sensor data will result in actions that may result in loss of mission or abort, e.g., Launch Control Center (LCC) scrub or early RS-25 engine shutdown	0.75
Safety Critical	Incorrect interpretation of data poses imminent threat to vehicle integrity and/or crew safety, e.g., un-contained engine failure	1.00

Table 3. Functional Criticality Subcategories

costs of a test recycle. In contrast, low-fidelity sensor models used in design simulations suggested that small amounts of noise would lead to a large bias toward false negatives.

3.1 Helium Leak Algorithm

Detection of a high rate of consumption of gaseous helium is determined by polling each of four AEPTs every 20 ms, each associated with the supply manifold of one of four heated 750 liter pressurized helium storage bottles. Each pressure system supplies GHe to one of four RS-25 engines and is used for functions such as propellant valve actuation. The nominal storage pressure is approximately 17.2 MPa and is consumed at a nominal rate of mass flow such that $\frac{dP}{dt}$ ranges from approximately -15 to -20 $^{\mathrm{kPa}}/_{\mathrm{s}}$. Pressure rates exceeding -31 $^{\mathrm{kPa}}/_{\mathrm{s}}$ are considered indicative of a leak and are used to trigger a caution and warning action.

The software detection algorithm operates by buffering samples and continually finite differencing samples separated by j counts; $\frac{\hat{dP}}{dt} = (P_k - P_{k-j})/\Delta T_j$. A persistence counter is used; if n consecutive checks are flagged, a warning is latched.

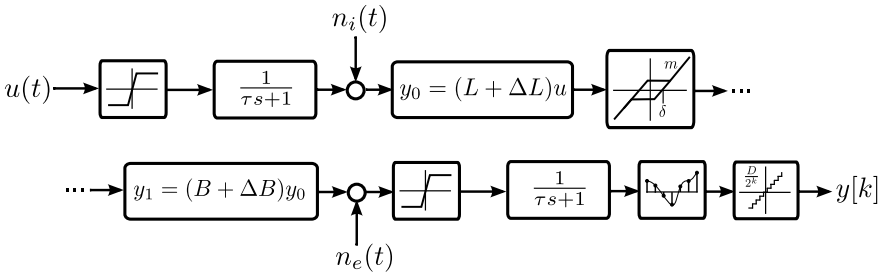

Figure 2. Simplified Generic Transducer Model

AEPTs are a family of state-of-the-art active analog transducers with a bandwidth of more than 200 Hz, providing an analog output with a range of 0.5-5.5V corresponding linearly to each sensor's design dynamic range. Dynamic range is tailored to each installation and depends on the physical characteristics of the sense diaphragm. In the present application, the AEPT has a dynamic range of 34.5 MPa absolute. Initial analog to digital conversion of the sensor output occurs in the Combined Control System Electronics (CCSE) avionics at approximately 150 Hz, where it is conditioned using a passive analog single-pole input filter with a bandwidth of 25 ± 5 Hz. Conversion of the pressure value, following internal calibration, uses 12-bit linear quantization yielding a least significant bit (LSB) of 8.4 kPa. The resultant value is then decimated with no further antialias filtering to the avionics bus rate of 50 Hz. Narrowband aliasing risk in the decimation operation was assumed to be negligible since the input signal is dominated by aeroacoustic noise with a broad, random spectral component.

Design simulations using a simplified sensor model with additive Gaussian noise (Figure 2) indicated that the algorithm was generally biased in the false negative direction. In this model, bias and linearity (scale factor) errors are captured in addition to transducer dynamics, hysteresis, and sampling/quantization effects. "Internal" and "external" noise sources (n_i and n_e) are modeled using appropriately scaled discrete-time Gaussian noise approximations.

A false negative bias was not unexpected, as the noise levels were derived from early hardware acceptance test data and were generally larger than allowed for in the engineering requirements. Combined with a relative absence of filtering in the signal path and an LSB having a value of about 27% of the target threshold rate, the positive tail of the noise distribution could easily reset the aforementioned persistence counter and would seldom latch a warning with the true leak rate at least 1-2 LSBs above the threshold.

3.2 Monte Carlo Assessment

Statistical analysis using Monte Carlo sampling was performed to quantify the operating characteristic curves and error probabilities for the threshold detection at a fixed confidence level. In this process, a sample of simulation runs of size N over a $\frac{dP}{dt}$ ranging from below nominal (-13.8 $\mathrm{kPa/s}$) to above nominal (-48.3 $\mathrm{kPa/s}$) is generated from a uniform distribution, such that the mean consumption rate is equal to the target threshold of -31 $\mathrm{kPa/s}$. Over each simulation run of a few seconds, the consumption rate and sensor error parameters are held constant. Since the distribution is uniform, approximately $N/2$ samples have actual consumption rates less than the threshold rate, and vice-versa.

The simulation outputs are then separated into four populations in a contingency table as shown in Table 4. The actual sample sizes of the populations possessing or not possessing leaks are separated into two groups of size N_1 and N_2. The corresponding errored cases (i.e., classifier made an incorrect decision) are counted as k_1 and k_2, respectively. Using these data, the expected error rates p_{f1} and p_{f2} and total failure rate p_f can be constructed from the actual (arbitrary) sample sizes using binomial statistics. These methods are commonly used in industrial acceptance sampling and are applied extensively by NASA for requirements verification.[4]

The failure rate p_f is derived from the cumulative binomial distribution,

$$F_{\mathrm{BIN}}(k, p_f, N) = \sum_{j=0}^{k} \left(\begin{array}{c} N \\ j \end{array} \right) p_f^j \left(1 - p_f\right)^{N-j} \tag{1}$$

e.g., the probability that the sampling process can generate k or fewer failures when the underlying failure rate is p_f. The value of this distribution is the *consumer risk* (CR) or Type II error probability β; the probability that the actual population failure rate is greater than the predicted failure rate. The conservative estimate of the actual failure rate at a specified consumer risk β is given by the solution of $p_f^\star \mid F_{\mathrm{BIN}}(k, p_f^\star, N) = \beta$. The value of $p_f^\star > p_f$ can be found via the inverse of the F cumulative distribution, or via numerical iteration. Of course, high reliability systems are hard to verify at high confidence with small numbers of Monte Carlo runs; in fact, the minimum number of runs required to verify a given success probability p_s is $N_0 = \left\lceil \frac{ln\beta}{\ln p_s} \right\rceil$. Thus, for "one-sided 3σ" equivalent failure rates of 0.135%, sample sizes on the order of $N_0 = 1800$ are required. In the present two-outcome problem, $N = 4000$ such that $N_i \approx 2000$.

		Classifier Output	
		P	N
True Value	T	TP $(N_1 - k_1, N_1)$	FN (k_1, N_1)
	F	FP (k_2, N_2)	TN $(N_2 - k_2, N_2)$

Table 4. Leak Detection Contingency Table

Initial Monte Carlo assessments predicted that the reliability of the leak detection algorithm was as low as 7.5%, with a FN rate of 92.5% (10% CR) and a mean time-to-warning of 37 seconds.

3.3 Analysis and Mitigation

The oversights in the initial algorithm design were a combined effect of insufficient requirements that did not consider the tradeoff in algorithm performance as a function of time-to-detection in the presence of noise and quantization error, and an insufficient focus on the effects of the real transducer dynamics and signal flow paths in the underlying avionics hardware. In particular, the sensor hardware was made more susceptible to noise due to the vendor's implementation of a very fast response time requirement that drove the sensor bandwidth to an extremely high value. This requirement was partly to establish the capability for rapid detection of *catastrophic* failures, e.g., pressure bottle ruptures.

An ideal binary classifier in the presence of zero mean Gaussian noise can theoretically achieve perfect performance in the limit as the decision time goes to infinity. However, in the present application, leak detection times longer than about 30 seconds can potentially impact mission objectives. Since the detection algorithm in question did not perform any stateful averaging of measurements, its performance was unacceptably poor. Performance was significantly improved by the implementation of a single-parameter first-order discrete-time software filter of the form $y_n = (1-\alpha)y_{n-1} + \alpha u_n$ where α is the filter parameter and is related to the equivalent continuous-time time constant τ by $\alpha = T_s/\left(\tau + T_s\right)$, where T_s is the sample time. A filter time constant of $\tau = 0.66\,\mathrm{s}$ was determined to provide an optimal balance of error rates, yielding a predicted reliability of 99.5% (10% CR) with an intentional FP bias of 3.1%. The time-to-detection was also improved to an average of 12 seconds, only 4 seconds longer than the minimum imposed by the length of the software buffers used for finite differencing. The nearly ideal distribution of the operating characteristic curves is achieved, where the probability of detection is uniformly distributed until reaching a region very near the threshold.

4 TEST VALIDATION

Following the analysis and software modifications conducted to improve the performance of the GHe leak detection and other algorithms using OFI MPS pressure sensors, a data analysis campaign was initiated to perform a detailed assessment of high-fidelity AEPT test data that was collected during the initial hardware acceptance test program. Since acceptance testing revealed higher-than anticipated noise levels in the time domain, it was desired that the program fully quantify its effects and the adequacy of models previously developed to support the aforementioned statistical performance assessment.

During this test activity, instrumented flight-qualified devices were attached to a test fixture and their output was recorded by a calibrated data acquisition system (DAQ) at a rate of 100 kHz with a fixed pressure input. A simulated vibration profile matching the predicted flight environment spectrum at maximum dynamic pressure was ramped to the full acceptance power level over approximately 60 seconds. While data was collected and analyzed in all three axes, the driving case is derived from vibration input in the sense axis as the motion is normal to the transducer diaphragm.

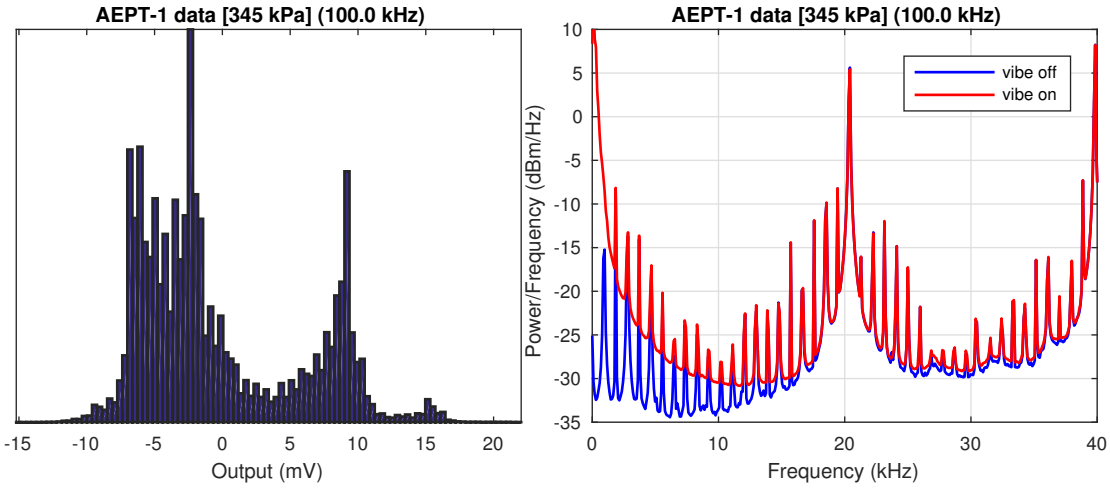

Figure 3. Power Spectrum and Distribution of Representative Sensor Output in Test

Spectral analysis of the test data revealed an unexpected characteristic of the sensor output: a relatively high energy, broadband, biased and non-Gaussian noise spectrum centered at about 40 kHz with harmonics as low as 900 Hz. The effect of the vibration input on the noise response, in fact, was shown to be comparable to the overwhelming harmonic content at higher frequencies. The spectrum in question has the telltale signature of a switched mode power supply, with the added concern that the 100 kHz DAQ bandwidth was insufficient to prevent aliasing of a probable 60 kHz sideband (Figure 3). It was determined that lacking any specific requirement to the contrary, the vendor's efforts to satisfy the response time specification resulted in an omission of effective filtering and isolation from the signal lines, which coupled the power supply harmonics directly into the measurement.

The result of this test complicated the verification and validation effort. First, it was immediately apparent that a simplified sensor model with zero mean Gaussian noise would not be sufficient to capture the observed dynamics of the sensor, which in addition to noise effects, included some random bias components and other nonlinearities. In addition, the intermediate analog antialias

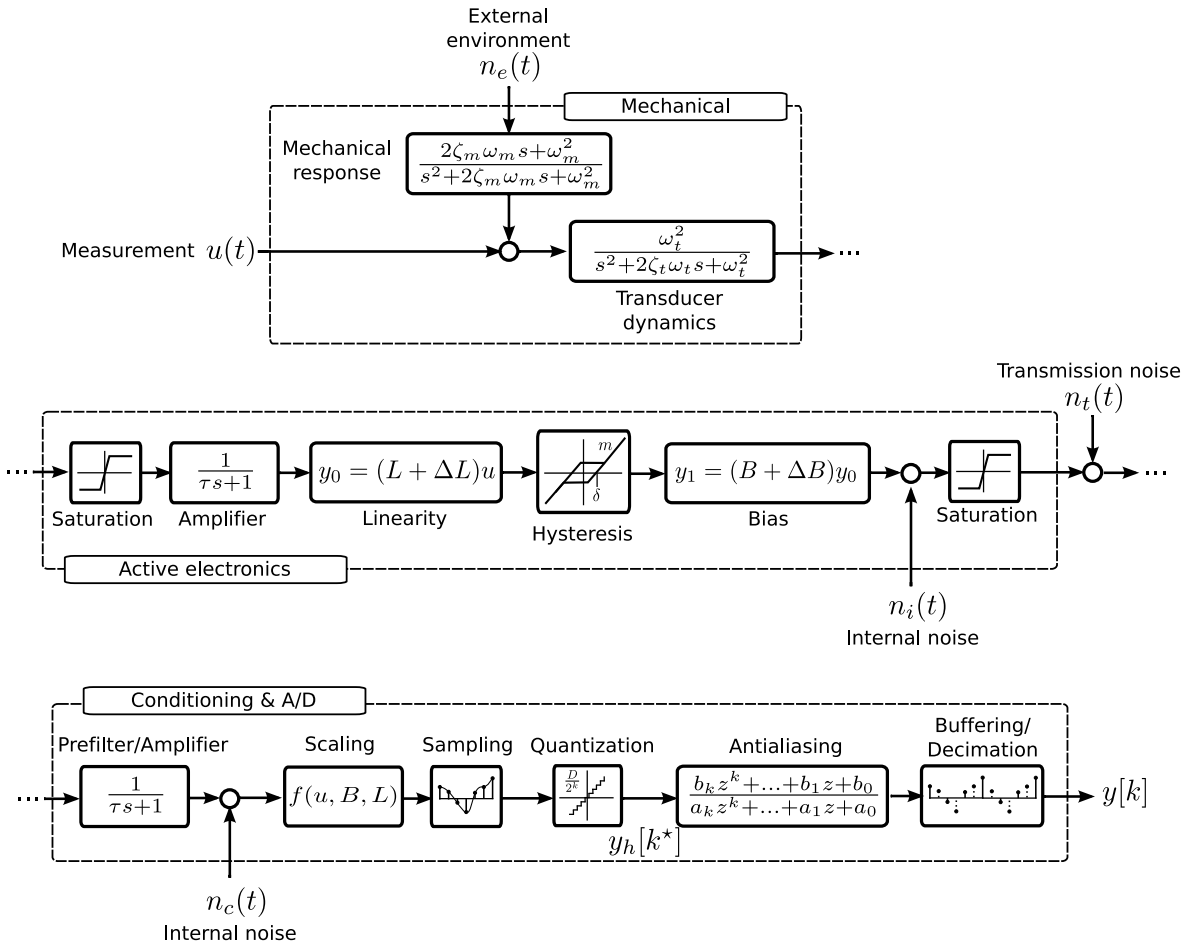

Figure 4. High Fidelity Transducer Model

filtering, sampling, and decimation operations including the analog to digital conversion (ADC) polling mechanism were modeled in order to mitigate the risk of aliasing of the troublesome harmonics below the 25 Hz Nyquist bandwidth of the flight computer. The simplified simulation model previously described was replaced by a high-fidelity model as shown in Figure 4.

Since it would be necessary to verify the mission software in the presence of the problematic harmonics, the use of an enveloping Gaussian noise PSD was considered but abandoned as it would induce excess conservatism into the noise model. Instead, a specialized noise model was constructed that uses a discrete noise sequence followed by a bilinear transformation of the following second-order biquadratic transfer function,

$$T(s) = \frac{as^2 + s\omega_0(k-b) + c\omega_0^2}{s^2 + \frac{\omega_0}{Q}s + \omega_0^2} \tag{2}$$

where a, b, k, ω_0, and Q are parameterized to produce an "inverse notch" resonance with a low damping ratio ($\zeta \approx 0.0005$) such that the filter, driven by discrete noise, produces an output whose power spectrum closely matches that of the observed test data (Figure 5, left). In addition, the large quantity of high-rate data was used to implement background denoising using a spectral subtraction

method, and anomalies attributed to structural resonances in the test fixture were removed from the model (Figure 5, right).

Cascaded filters were used to capture the primary harmonics below 2 kHz, above which the CCSE ADC input filter provided sufficient attenuation. These models were then implemented for real-time simulation to support hardware-in-the-loop verification, albeit at high computational cost compared with the simplified model in Figure 2. Fortunately, the robust filtering designed for the low-fidelity model was found to be adequate, with no measurable change to the leak detection algorithm operating characteristics discussed in Section 3.3.

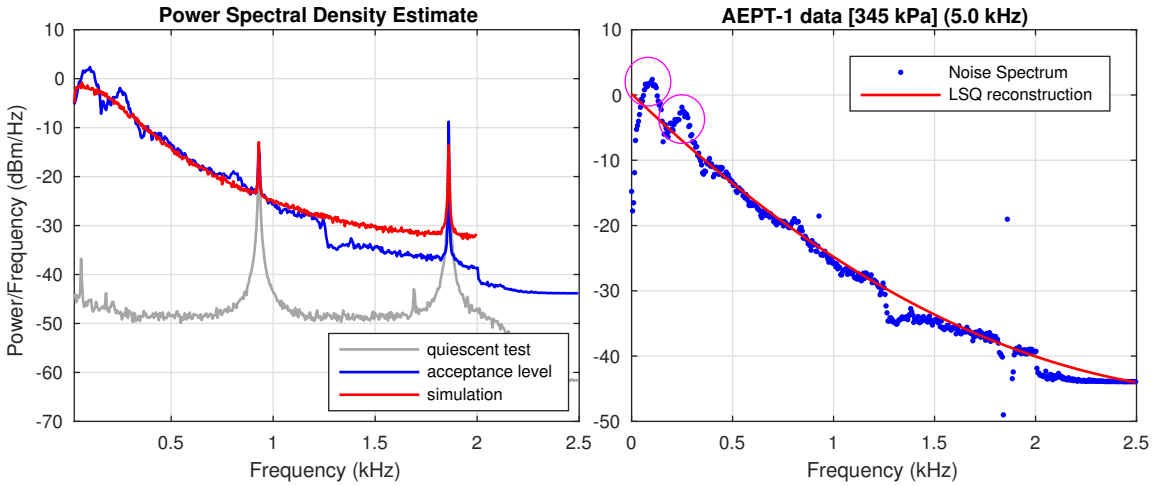

Figure 5. Comparison of Model Versus Test Data With Test Fixture Structural Modes

5 CONCLUSIONS

For the SLS program and all human-rated flight operations, careful management of risk is paramount to ensure that critical systems such as mainstage propulsion are able to execute autonomous functions without undue threats to the mission, the crew, or the public. Careful engineering integration and attention to real hardware performance must be considered early in the development lifecycle. In the case of the subject transducers and associated algorithms, the parallel development of hardware and software to support autonomy functions introduced interactions that were not anticipated early in the software development lifecycle. In part, the issues with noise were a result of incomplete requirements that drove the development of advanced sensor hardware whose performance capabilities exceeded actual functional needs.

The incorporation of autonomy into complex systems is a challenging exercise that must consider all factors from the physics of the transducer element to the results of autonomous actions. The present case study highlights the value of Model-Based Engineering (MBE) in the development of complex autonomous systems, although it does suggest that the MBE process needs to be applied much earlier in the development lifecycle. It also clarifies that gaps in subsystem requirements can readily manifest as unmodeled interactions; that is, the rigor and benefit of MBE is only as good as the models employed. Importantly, element engineers must be guided by expert insight that reaches across subsystem boundaries. For example, software engineers may lack hardware experience, electronics engineers may not be familiar with the physics of flight vibration environments, and systems engineers may formulate specifications that are disjoint with functional goals.

MBE helps identify gaps and connect interfaces, but test-based validation of the underlying models is crucial. In the case of the MPS pressure transducers, the rigorous verification activities allowed the identification and resolution of a potential performance issue in software, whose mitigation subsequently alleviated a hardware problem uncovered during acceptance testing.

REFERENCES

[1] Donahue, B., "The Space Launch System: Development Progress," AIAA SPACE 2016, AIAA SPACE Forum, Long Beach, CA, September 2016 (AIAA 2016-5415).

[2] Ballard, R., "SSME to RS-25: Challenges of Adapting a Heritage Engine to a New Vehicle Architecture," Proc. Sixth European Conference for Aeronautics and Space Sciences (EUCASS), Krakow, Poland, June 2015.

[3] Vetcha, N. et al., "Overview of RS-25 Adaptation Hot-Fire Test Series for SLS, Status and Lessons Learned," 2018 Joint Propulsion Conference, AIAA Propulsion and Energy Forum, Cincinnati, OH, July 2018 (AIAA 2018-4459).

[4] Hanson, J., and Beard, B., "Applying Monte Carlo Simulation to Launch Vehicle Design and Requirements Analysis," NASA TP-2010-216447, 2010.