International Journal of Spray and Combustion Dynamics
XX(X):2–20

Forecasting Thermoacoustic Instabilities in Liquid Propellant Rocket Engines Using Multimodal Bayesian Deep Learning

Ushnish Sengupta[1], Günther Waxenegger-Wilfing[2], Jan Martin[2], Justin Hardi[2] and Matthew P. Juniper[1]

Abstract

We present a method that combines multiple sensory modalities in a rocket thrust chamber to predict impending thermoacoustic instabilities with uncertainties. This is accomplished by training an autoregressive Bayesian neural network model that forecasts the future amplitude of the dynamic pressure time series, using multiple sensor measurements (injector pressure/ temperature measurements, static chamber pressure, high-frequency dynamic pressure measurements, high-frequency OH* chemiluminescence measurements) and future flow rate control signals as input. The method is validated using experimental data from a representative cryogenic research thrust chamber. The Bayesian nature of our algorithms allows us to work with a dataset whose size is restricted by the expense of each experimental run, without making overconfident extrapolations. We find that the networks are able to accurately forecast the evolution of the pressure amplitude and anticipate instability events on unseen experimental runs 500 milliseconds in advance. We compare the predictive accuracy of multiple models using different combinations of sensor inputs. We find that the high-frequency dynamic pressure signal is particularly informative. We also use the technique of integrated gradients to interpret the influence of different sensor inputs on the model prediction. The negative log-likelihood of data points in the test dataset indicates that prediction uncertainties are well-characterized by our model and simulating a sensor failure event results in a dramatic increase in the epistemic component of the uncertainty, as would be expected when a Bayesian method encounters unfamiliar, out-of-distribution inputs.

Prepared using sagej.cls

Keywords
thermoacoustic instabilities, multimodal data, bayesian machine learning, interpretable machine learning

Introduction

High-frequency thermoacoustic instabilities are a potentially catastrophic phenomenon in rocket and aircraft engines. They usually arise from the coupling between unsteady heat release rate and acoustic eigenmodes of the combustion chamber. Acoustic waves incident on a flame cause fluctuations in its heat release rate. If a flame releases more (less) heat than average during instants of higher (lower) local pressure, then more work is done by the gas during the acoustic expansion phase than is done on it during the acoustic compression phase. If this work is not dissipated then the oscillation amplitude grows and the system becomes thermoacoustically unstable.

Thermoacoustic instabilities can be challenging to model and predict because they are usually sensitive to small changes in a combustion system (Juniper (2018)). The large pressure fluctuations and accelerated heat transfer caused by these oscillations can lead to structural failure, because rocket engines have high energy densities and low factors of safety in their structural design. From the F1 rocket engine, which underwent costly redesigns in the 1960s (Oefelein and Yang (1993)) to the modern Japanese LE-9 engine (Watanabe et al. (2016)), they continue to plague the development of liquid propellant rocket engines.

Designers can suppress thermoacoustic oscillations in two ways. The first method involves fitting passive devices such as acoustic liners (Oefelein and Yang (1993)), baffles (Eldredge and Dowling (2003)) and resonators (Zinn (1970)) or the use of tuned passive control, where sensor measurements are used to diagnose thermoacoustic oscillations which can then be mitigated by adjusting a passive damping device or avoiding the operating point (Kobayashi et al. (2019a)). The second method is active feedback control (Dowling and Morgans (2005)). The challenging demands of high power output, reliability and robustness placed on an active feedback control system means that feedback control is not used in rocket engines and aircraft engines. This paper proposes the diagnosis of impending instabilities using multiple sensor data-streams as input and Bayesian neural networks for forecasting. The forecast can be used by a rocket engine controller to perform a mass flow rate adjustment and avoid the unstable regions.

[1]Department of Engineering, University of Cambridge, UK [2]Institute of Space Propulsion,German Aerospace Center (DLR), Germany

Corresponding author:
Matthew Juniper, Department of Engineering, University of Cambridge, Cambridge CB2 1PZ, UK.
Email: mpj1001@cam.ac.uk

Prepared using sagej.cls *[Version: 2017/01/17 v1.20]*

The remainder of the paper is paper is structured as follows. We provide an overview of related work in this field and explain the motivation behind the current study. Next, we describe the experimental rocket thrust chamber which generated our dataset. This is followed by details of the statistical tools we used: Bayesian neural network ensembles for forecasting and integrated gradients for interpreting them. We then compare the quality of the forecasts obtained using different model inputs and identify the influence of important input features on predictions using integrated gradients. We conclude with a brief summary of our findings and a discussion of potential avenues for future work.

The construction of instability precursors from pressure measurements and optical measurements has attracted considerable attention from the combustion research community, as reviewed by Juniper and Sujith (2018). Lieuwen (2005) used the autocorrelation decay of combustion noise, filtered around an acoustic eigenfrequency, to obtain an effective damping coefficient for the corresponding instability mode. Other researchers have used tools from nonlinear time series analysis, which capture the transition from the chaotic behaviour displayed by stable turbulent combustors to the deterministic acoustics during instability, such as Gottwald's 0-1 test (Nair et al. (2013)) and the Wayland test for nonlinear determinism (Gotoda et al. (2011)) to assess the stability margin of a combustor. Nair et al. (2014) reported that instability is often presaged by intermittent bursts of high-amplitude oscillations and used recurrence quantification analysis (RQA) to detect these. In a later paper, Nair and Sujith (2014) noted that combustion noise tends to lose its multifractality as the system transitions to instability. They proposed the Hurst exponent as an indicator of impending instability. Measures derived from symbolic time series analysis (STSA) (Sarkar et al. (2016)) and complex networks (Murugesan and Sujith (2016)) are also able to capture the onset of instability.

Recent studies have explored machine learning techniques to learn precursors of instability from data. These promise greater accuracy than physics-based precursors of instability, though at the cost of robustness and generalizability. Hidden Markov models constructed from the output of STSA (Jha et al. (2018)) or directly from pressure measurements (Mondal et al. (2018)) have been used to classify the state of combustors. Hachijo et al. (2019) have projected pressure time series onto the entropy-complexity plane and used support vector machines (SVMs) to predict thermoacoustic instability. SVMs were also employed by Kobayashi et al. (2019b), who used them in combination with principal component analysis and ordinal pattern transition networks to build precursors from simultaneous pressure and chemiluminiscence measurements. Sengupta et al. (2020) showed that the power spectrum of the noise can be used to predict the linear stability of a thermoacoustic eigenmode using Bayesian neural networks. Related work by McCartney et al. (2020) used the detrended fluctuation analysis (DFA) spectrum of the pressure signal as input to a random forest and found that this approach compared favorably to precursors from the literature. A recent study by Gangopadhyay et al. (2021) trained a 3D Convolutional Selective Autoencoder on flame videos to detect instabilities. Nóvoa and Magri (2022) adopt a hybrid approach,

using Bayesian data assimilation to learn the parameters of a low fidelity model and an echo state network to model the forecast bias. Transfer learning is also being explored by Mondal et al. (2021) as a means of efficiently transferring knowledge of precursors across machines.

Most papers in the literature identify indicators of approaching instability. In this study, on the other hand, we use nonlinear autoregressive time series modeling with a Bayesian Neural Network to forecast, with uncertainties, the future amplitude of pressure fluctuations given the history of sensor signals and future flow rate control signals. This informs the engine controller of the expected timing and potential severity of a forthcoming instability. This approach also lets us take advantage of the rich instrumentation of our rig by integrating multimodal sensor data into our model: high frequency dynamic pressure ($p'(t)$) and OH* chemiluminescence ($I'(t)$) measurements, injector pressures (p_{H2}, p_{O2}) and temperatures (T_{H2}, T_{O2}), chamber static pressure (p_{cc}) and fuel flow rates ($\dot{m}_{O2}, \dot{m}_{H2}$). We compare models with different combinations of sensors, different sensor-derived features, and different lengths of signal histories in the inputs. We find that, although operating parameters and mass flow rate control signals can anticipate instabilities on their own, inclusion of dynamic pressure signal history, in particular, results in a measurable increase in the forecast accuracy. It is not possible to predict triggered thermoacoustic instabilities from operating parameters and control signals alone, so we need to integrate multimodal sensor data with operating parameters and control signals in an instability forecasting framework.

The size of our dataset is limited by the high cost and effort required for each additional test run on the rocket thrust chamber. Fortunately, the uncertainty-aware nature of the Bayesian Neural Network allows us to work in this medium-data regime without overconfident extrapolation. We compute log-likelihoods on the test dataset and find that the uncertainty is well-quantified. We also simulate a sensor failure to assess the robustness of our Bayesian forecasting tool. As expected, the network exhibits large epistemic uncertainties during the failure, signalling the unfamiliar out-of-distribution nature of the inputs. Finally we use the interpretation technique of integrated gradients (IG) to understand how the different input features in our model influence the predictions.

Experimental Setup

The dataset used in this study is derived from experiments performed on the research combustor BKD (Gröning et al. (2016)) operated at the P8 test facility (Fröhlke et al. (1997), Koschel et al. (1996)) of the DLR Institute of Space Propulsion in Lampoldshausen. It has three main components: an injector head, a cylindrical combustion chamber and a convergent-divergent nozzle, as shown in Figure 1. The combustion chamber is water-cooled and is designed to deal with the high thermal loads that are expected during instability events. The L42 injector head has 42 shear coaxial injectors and is operated with a liquid oxygen (LOX)/ hydrogen (LH2) propellant combination. The cylindrical combustion chamber is

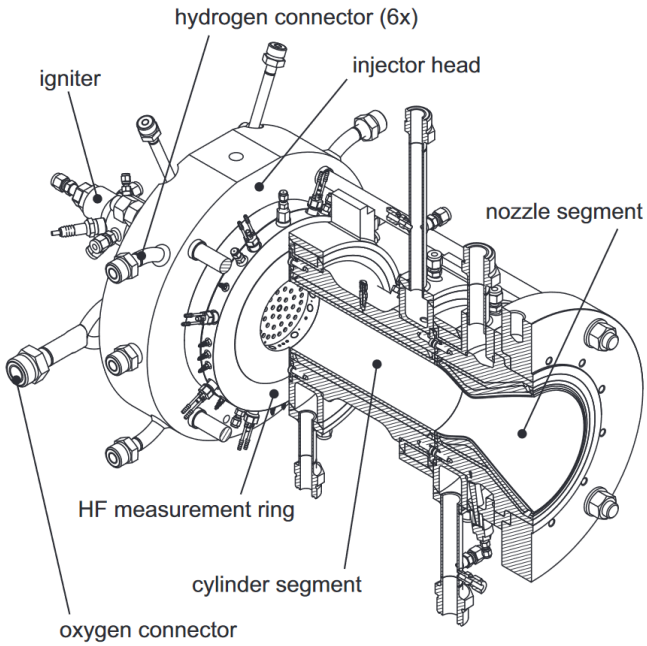

Figure 1. Experimental thrust chamber BKD.

80 mm in diameter and the nozzle throat diameter is 50 mm, resulting in a contraction ratio of 2.56. The chamber static pressure (p_{cc}) is varied between $50 - 80$ bar and mixture ratio of oxidizer to fuel (ROF = $\dot{m}_{O2}/\dot{m}_{H2}$) between $2 - 6$. The experiments considered here have a LOX injection temperature T_{O2} around 110 K and a hydrogen injection temperature T_{H2} around 100 K.

For the operating point with $p_{\mathrm{cc}} = 80$ bar, ROF $= 6$, the total propellant mass flow rate is 6.7 kg/s, the theoretical thermal power is 90 MW and the thrust achieved is about 24 kN. These specifications place BKD at the lower end of small upper stage engines.

A representative BKD test sequence, along with a spectrogram of dynamic pressure oscillations inside the combustion chamber, is shown in Figure 2. Stable and unstable operating conditions can be identified in the spectrogram. Strong high-frequency combustion instabilities of the first tangential (1T) mode at about 10 kHz were excited consistently when approaching the operating point at $p_{\mathrm{cc}} = 80$ bar, ROF $= 6$. The coupling mechanism for this instability is injection-driven (Gröning et al. (2016)). The flame dynamics are modulated by the LOX post acoustics and combustion instabilities emerge when the frequency of the 1T chamber mode matches the second longitudinal eigenmodes of the LOX posts. This mechanism has been confirmed using high-speed flame imaging (Armbruster et al. (2019)).

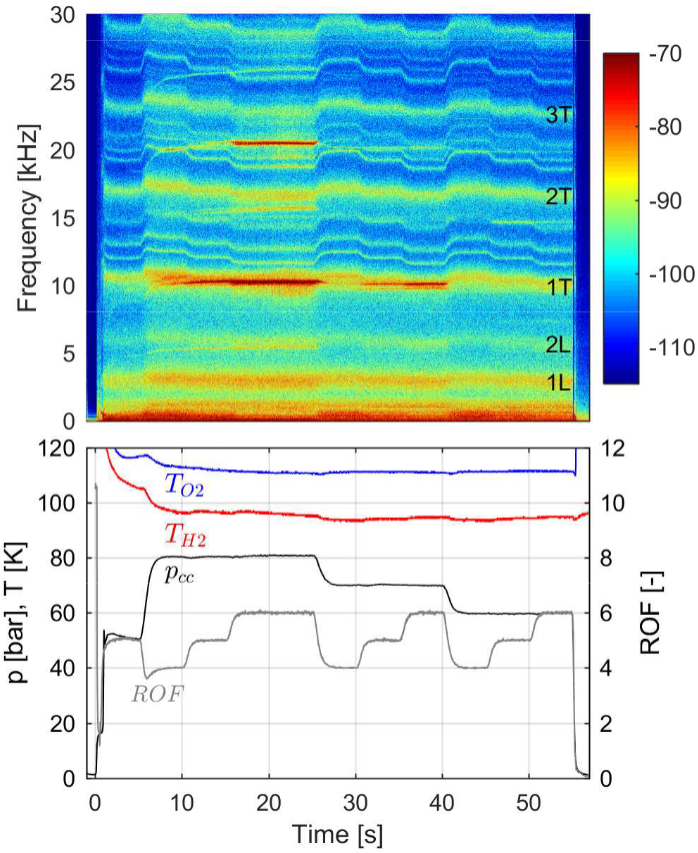

Figure 2. BKD test sequence and p' spectrogram showing self-excited instability of the first tangential mode (Gröning et al. (2016)). p_{cc} is the chamber static pressure, ROF is the mixture ratio of oxidizer to fuel, and T_{H2}, T_{O2} are the LOX and hydrogen injection temperatures respectively.

The extreme conditions within the thrust chamber, where temperatures reach 3600 K and the pressure reaches 80 bar, limit the diagnostic instruments for our instability investigations. A specially designed measurement ring is placed between the injector head and the cylindrical combustion chamber segment, as shown in Figure 1. At this location, the temperatures are moderated by the injection of cryogenic propellants, which allows the mounted sensors to survive several test runs. Eight Kistler type 6043A water-cooled high-frequency piezoelectric pressure sensors are flush-mounted in the ring with an even circumferential distribution, in order to measure the chamber pressure oscillations $p'(t)$. The high-frequency pressure sensors have a measurement range set to ± 30 bar and a sampling rate of 100 kHz. An anti-aliasing filter with a cutoff frequency of 30 kHz is applied. Three fibre-optical probes are used

to record the OH* radiation intensity $I'(t)$ of selected individual flames. The full acceptance angle of the optical probes is approximately 2°. In order to capture the OH* radiation signal, the probes are equipped with inteference filters with a center wavelength of 310 nm. The sampling frequency of the $I'(t)$ signals is also 100 kHz. There are also several low sampling frequency (100 Hz) sensors that measure the chamber static pressure p_{cc}, the mass flow rates of fuel $\dot{m}_{H2}$ and oxidizer $\dot{m}_{O2}$ and injector temperatures T_{H2}, T_{O2}.

Methods

Nonlinear autoregressive models with exogenous variables (NARX) for timeseries modeling

To forecast instabilities, we train NARX-style models (Lin et al. (1996)) to capture the functional relationship between the history of the sensor data, control signals and future pressure fluctuations. These model the evolution of pressure fluctuations as follows:

$$\log(p'_{\text{rms}}(t_m + \Delta t)) = f(\mathbf{x}_m, \mathbf{x}_{m-1}, ... \mathbf{x}_{m-h}, \mathbf{u}_{m+f}) + \sigma_\epsilon \tag{1}$$

Here, $p'_{\text{rms}}(t_m + \Delta t)$ is the root mean square amplitude of the dynamic pressure fluctuation signal in a 50 ms window centered around $t_m + \Delta t$. We choose $\Delta t = 500$ ms for our models because we consider this to be sufficient warning for our rocket engine controller to detect and react to an impending instability with a mass flow rate adjustment. The state variables $\mathbf{x}_m, \mathbf{x}_{m-1}, ... \mathbf{x}_{m-h}$ consist of various sensor data (or transformations thereof) at times $t_m - 50, t_m - 100, ... t_m - 50h$ ms. The order of dependence h determines how far in the past we look. In this study, we explore models with $h = 0$ and $h = 2$. The elements of $\mathbf{u}_{m+f}$ are the two mass flux control signals $\dot{m}_{H2,set}$ and $\dot{m}_{O2,set}$ averaged over the window between t_m and $t_m + \Delta t$. These are the exogenous variables whose future values are known and can therefore be included in the inputs. It is possible to train models that capture higher order dependence on instantaneous values of the control signals instead of reducing them to an average, but we find that this does not improve predictive accuracy in our problem. σ_ϵ is the homoskedastic (constant) observation noise or aleatoric uncertainty term, which is assumed to be Gaussian. We forecast the logarithm of $p'_{\text{rms}}(t_m + \Delta t)$ instead of $p'_{\text{rms}}(t_m + \Delta t)$ itself because the noise in the pressure fluctuations increases with the amplitude, and so the simplifying assumption of constant homoskedastic noise only holds approximately once this logarithmic transform has been applied. It also forces p'_{rms} predictions to be positive.

We model the nonlinear function f in Equation (2) as a feedforward neural network with a simple multi-layer perceptron (MLP) architecture. In a standard neural network, one obtains a point estimate of the network parameters. Here, however, we train a machine learning algorithm on a dataset whose size

is limited by the cost of performing experiments on large rocket thrust chambers. With a standard neural network this size limit could lead to naive and inaccurate extrapolations when the network encounters out-of-distribution inputs that are different from those which were experienced during training. In safety-critical systems such as rockets, control decisions based on overconfident predictions could be catastrophic. Bayesian Neural Networks (BNNs) are an elegant solution to this problem because they provide a confidence interval as well as a prediction.

Bayesian Neural Network Ensembles

The Bayesian framework for training neural networks Neal (1996) involves placing a sensible prior probability distribution $p(\theta)$ over the parameters θ of the network. Observations $\mathcal{D}$ are then used to update our belief about the parameter distribution. The posterior distribution $p(\theta|\mathcal{D})$ is obtained using Bayes' rule.

$$p(\theta|\mathcal{D}) = \frac{p(\mathcal{D}|\theta)p(\theta)}{\int p(\mathcal{D}|\theta)p(\theta)d\theta} \tag{2}$$

Bayesian models are resistant to overfitting on small datasets and are uncertainty aware. They are also useful for continual learning throughout the lifetime of a device in operation (Li et al. (2019)). Exact Bayesian inference, however, is not possible in neural network models because the integral $\int p(\mathcal{D}|\theta)p(\theta)d\theta$ is analytically intractable. The most widely used numerical method to integrate over the posterior, Markov Chain Monte Carlo, is too computationally expensive to be practical. Variational inference is often used (Blundell et al. (2015)), in which the posterior is approximated using a convenient parametrization and the Kullback-Leibler divergence between this variational distribution and the true posterior is minimized using backpropagation. However, although computationally cheap, mean-field variational inference also has drawbacks such as not capturing correlations between parameters. Recently, a different approximate inference method, based on ensembling, has been proposed. This is called the anchored ensembling algorithm. It is cheap, simple and scalable but manages to outperform variational inference in several uncertainty quantification benchmarks (Pearce et al. (2018)).

Consider a dataset of N_D data points $(\mathbf{x}_n, y_n)$, where each data point consists of features $\mathbf{x}_n \in \mathbb{R}^D$ and output $y_n \in \mathbb{R}$. Define the likelihood for each data point as $p(y_n \mid \theta, \mathbf{x}_n, \sigma_\epsilon^2) = \mathcal{N}(y_n \mid \mathrm{NN}(\mathbf{x}_n; \theta), \sigma_\epsilon^2)$, where NN is a neural network whose weights and biases form the latent variables θ while σ_ϵ^2 is the observation noise. Define the prior on the weights and biases θ to be the standard normal $p(\theta) = \mathcal{N}(\theta \mid \mu_{prior}, \Sigma_{prior})$. The anchored ensembling algorithm (Figure 3) then does the following:

1. The parameters $\theta_{anc,j}$ of each j-th member of our neural network ensemble are initialized by drawing from the prior distribution $\mathcal{N}(\mu_{prior}, \Sigma_{prior})$.

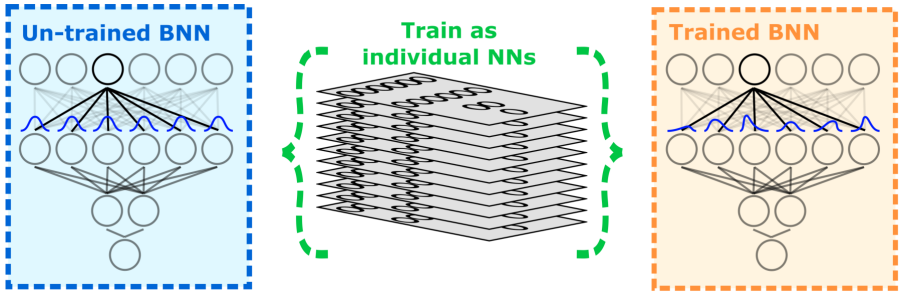

Figure 3. Approximate Bayesian inference using ensembles of neural networks.

2. Each ensemble member is trained using a modified loss function that anchors the parameters to their initial values. The loss function for the j-th ensemble member is given by $Loss_j = \sum_{i=1}^{N_D}(y_i - \hat{y}_j(\mathbf{x}_i, t_i))^2 + \|\mathbf{\Sigma}_{prior}^{-1/2}(\theta_j - \theta_{anc,j})\|_2^2$, where $\hat{y}_j(\mathbf{x}_i, t_i)$ is the neural network output and the i-th diagonal element of Σ is the ratio of data noise to the prior variance of the i-th parameter.

Pearce et al. (2018) prove that this procedure approximates the true posterior distribution for wide neural networks and that the trained neural networks in the ensemble may be treated as samples from an approximate posterior distribution. The resulting ensemble has predictions that converge when they are well-supported by the training data and diverge when they are not. The variance of the ensemble predictions is an estimate of the Bayesian model's epistemic uncertainty. Epistemic or knowledge uncertainty is the reducible portion of uncertainty which stems from uncertainty in model parameters and decreases as more data is added. The total uncertainty of each prediction is obtained by adding the epistemic uncertainty and the irreducible aleatoric uncertainty σ_ϵ which stems from observation noise:

$$\sigma_{\text{total}}^2 = \sigma_\epsilon^2 + \frac{1}{N_e}\sum_j \hat{y}_j^2 - (\frac{1}{N_e}\sum_j \hat{y}_j)^2 \tag{3}$$

Interpretation using Integrated Gradients

Neural networks have a reputation as black boxes, which disincentivizes their application to cases in which practitioners must understand why the algorithm is or is not working. We use the technique of integrated gradients (IG) (Sundararajan et al. (2017)) to attribute the predictions of our network ensemble to the input features. This is a simple scalable method that only requires access to the gradient operation.

IG has several desirable properties that other attribution methods lack, such as implementation invariance, sensitivity, linearity, completeness and preservation of symmetry.

For deep neural networks, the gradients of the output with respect to the input are an analog of the linear regression coefficients. In a linear model, the regression coefficients characterize the relationship between input and output variables globally. In a nonlinear neural network, however, the gradient of the output at a point merely characterizes the local relationship between a predictor variable and the output. IG computes the path integral of the gradient of the outputs with respect to the inputs around a baseline to the input under consideration. For an image recognition algorithm, a completely black image could be a reasonable choice of baseline, while for a regression problem like ours, where the input variables were normalized to have zero mean and unit standard deviation, the average input of all zeros is a sensible baseline. We consider the straight-line path (in the feature space $\mathbb{R}^p$) from the baseline $\boldsymbol{x}'$ to the input $\boldsymbol{x}$ being considered. The IG attribution $attr_j$ for the j-th feature is then defined as follows:

$$attr_j(\boldsymbol{x}) = (\boldsymbol{x}_j - \boldsymbol{x}'_j) \int_{\alpha=0}^{1} \frac{\partial f(\boldsymbol{x}' + \alpha(\boldsymbol{x} - \boldsymbol{x}'))}{\partial \boldsymbol{x}_j} d\alpha \tag{4}$$

This integral is computed numerically by sampling the gradient evenly along the path. For our neural network ensemble, we average the integrated gradient attributions obtained from individual ensemble members to obtain a mean $attr_j$ for each feature.

Transforming sensor signals into neural network inputs

We compare NARX models with different combinations and transformations of the sensor data in their state variables $\mathbf{x}_{m-i}$ (Equation 2) and evaluate their accuracy.

Models that include the low frequency signals $p_{\text{cc}}, \dot{m}_{H2}, \dot{m}_{O2}, T_{H2}$ and T_{O2} in their $\mathbf{x}_{m-i}$ vectors use the average value of the signal between $t_m - 50i$ and $t_m - 50(i+1)$. For the high frequency pressure und OH* signals, p' and I', we use transformations from the literature that are known to capture instability-relevant information as described next. The current p'_{rms} is a natural candidate which is included in all models that incorporate pressure information. Sengupta et al. (2020) and McCartney et al. (2021) have shown that, before a combustor transitions to instability, the power spectral density shows measurable changes which can be used as a precursor. We estimate the spectral density of the p' (p'_{FFT}) and I' (I'_{FFT}) time series by applying Welch's method Welch (1967) to a segment of the signals between $t_m - 50i$ ms and $t_m - 50(i+1)$ ms. In this study, a Hann window of length 30 with 50% overlap is used for both p' and I'.

Kobayashi et al. (2019b) applied machine learning to the the ordinal partition transition network of the p' and I' timeseries to detect instabilities early. The ordinal partition transition networks for an m-dimensional time series are expressed by a weighted adjacency matrix W consisting of $W_{ij} = P(\pi_i \longrightarrow \pi_j)$, $i, j \in [1, 2^m]$, $\sum W_{ij} = 1$, where $P(\pi_i \longrightarrow \pi_j)$ is the observed probability from the ith- to jth-order transition patterns. In this study, the order patterns are π_1, π_2, π_3, and π_4 which account for all combinations of the signs of the increments for our two-dimensional time series $[p'(t), I'(t)]$. The probability distribution of the corresponding 16-dimensional transition patterns W computed from the segment of the signals between $t_m - 50i$ and $t_m - 50(i + 1)$ are used in the state variable vector.

We also include a naive baseline model $p'_{\text{rms}}(t_m + \Delta t) = p'_{\text{rms}}(t_m)$ for comparison. This model, by design, is incapable of forecasting any future instabilities but, because large portions of the run consist of steady operation, its predictions have a reasonably good root mean squared error (RMSE). This model exists to provide a baseline RMSE and any useful forecast needs to be significantly more accurate than this.

Train-test-validation split, hyperparameters, performance metrics

We use a dataset of five experimental runs: three that are 50 seconds long and two that are 90 seconds long. Each run has two occurrences of the 1T instability, where an instability event is defined as being when the peak-to-peak amplitude of the pressure fluctuations exceeds 5% of the chamber static pressure.

To avoid data leakage, one experimental run is exclusively reserved for hyperparameter tuning. The timeseries is split into contiguous 250 millisecond blocks and every fifth block is used to evaluate the negative log-likelihood, which is the metric we use for tuning the BNN hyperparameters. The hyperparameters are the network depth, layer width, noise σ_ϵ, the learning rate for the ADAM optimizer, the number of epochs and the number of neural networks in the ensemble. For simplicity, and to facilitate comparison between models, the same hyperparameters are chosen for all models. This does not compromise the performance of any particular model, since increasing the key hyperparameters, width and depth, benefits all models. We choose a 3 hidden layer MLP with 100 units in each layer because larger networks show only minimal gains in performance. A learning rate of 2×10^{-4} is found to be optimal for convergence. The aleatoric noise parameter σ_ϵ is set to 0.07. In anchored ensembling, we train all members of the ensemble until convergence so the number of epochs was set to 512, which is found sufficient for all models. Converged estimates of the test data log-likelihood are obtained using 25 neural networks per ensemble.

To evaluate model performance, leave-one-out cross validation (LOOCV) is performed on the remaining four runs. Time series data involves strong temporal correlations, so assigning training and test data points randomly is dishonest because then the two sets would be highly similar. The test set must

therefore be a completely independent experimental run. Additionally, LOOCV is necessary because with only four runs, firm conclusions cannot be drawn based on performance metrics for one particular test-train split, since any claimed efficacy or differences between models could be caused by chance. We report average performance metrics across the four possible splits.

To evaluate model accuracy, we compute the root mean squared error (RMSE) of our pressure amplitude forecast for the entire test run (full RMSE). We also calculate this on a subset of the test run comprising the two segments of the timeseries one second before and after a transition to instability event (transition RMSE). The RMSE is computed separately on this subset because our primary interest is in forecasting instabilities and the relative merits of different models become clearer on this subset where there is a dramatic change in the amplitude. As a measure of the quality of uncertainty, we report mean negative log-likelihood per test data point.

Results

The models were trained on a laptop with 16 GB RAM, an Intel i7-10870H CPU and an NVIDIA RTX 2070 GPU. Each NARX model took ~ 2.5 hours to train. The inference time of the ensemble averaged around 10 milliseconds per sample, indicating that the algorithm is suitable as a real-time diagnostic tool, once it has been trained on experimental data.

Table 1 shows the average prediction RMSEs for the NARX models computed during cross-validation across the four runs. We note, firstly, that the NARX models outperform the naive baseline, which means that the data contains a predictable signal that can be extracted. We also observe that the simple models, which use only the two operating point parameters p_{cc} and ROF as inputs, perform well. This is unsurprising because we know that the 1T instabilities are consistently triggered when the operating point at $p_{\text{cc}} = 80$ bar and ROF $= 6$ is approached. This means that knowing the current operating point and future flow rate control signals is informative. All transitions to instability in the test runs differ from each other, however, even though though they have identical control trajectories. The cross-validation shows unambiguously that the data from the additional instrumentation improves the prediction of future pressure fluctuations. Models that include the dynamic pressure spectrum p'_{FFT}, radiation intensity spectrum I'_{FFT}, current amplitude p'_{rms}, ordinal network transition probabilities, $W, T_{O2}, T_{H2}, p_{O2}$ or p_{H2}, have better prediction accuracies, especially on the transition subset. Including higher order dependencies on state variables also provides a slight boost in accuracy. Interestingly, the dynamic pressure spectrum appears more informative than the spectrum of the OH* radiation intensity. This could be due to the fact that a single optical probe observes only a small volume of the chamber, while the pressure field integrates information across the whole chamber.

State variable inputs $\mathbf{x}$	Order	Full RMSE	Transition RMSE	NLL
Baseline	–	0.1552	0.2134	–
p_{cc}, ROF	0	0.1289	0.1630	-1.622
p_{cc}, ROF	2	0.1202	0.1576	-1.631
p'_{rms}, p'_{FFT}	0	0.1197	0.1546	-1.5752
p'_{rms}, p'_{FFT}	2	0.1169	0.1508	-1.5467
p'_{rms}, I'_{FFT}	0	0.1255	0.1612	-1.5154
p'_{rms}, W	0	0.1241	0.1566	-1.5801
$T_{O2}, T_{H2}, p_{O2}, p_{H2}$	0	0.1243	0.1578	-1.5948
p_{cc}, ROF, p'_{rms}, p'_{FFT}, $T_{O2}, T_{H2}, p_{O2}, p_{H2}$	0	0.1048	0.1407	-1.5632
p_{cc}, ROF, p'_{rms}, p'_{FFT}, $T_{O2}, T_{H2}, p_{O2}, p_{H2}$	2	0.1005	0.1311	-1.5350

Table 1. Cross-validation RMSEs (in bar) for entire runs (Full RMSE) and the transition subset of the run (Transition RMSE) and mean negative log-likelihoods (NLL) for different NARX models. Lower RMSEs and NLLs are better.

Table 1 also reports the mean negative log-likelihood (NLL) per datapoint, averaged across the four cross-validations. The mean negative log-likelihoods are fairly low for most of our models, implying that very few observations lie far outside the uncertainty bounds of the forecast (Figure 4). This is also confirmed by the percentage of data points in the test set within the ± 1 standard deviation (s.d.) and ± 3 s.d. bounds, which are roughly in line with our Gaussian assumptions (68% within 1 s.d., 99.7% within 3 s.d.). Additionally, though models using more features have higher prediction accuracies, they have slightly worse NLLs in some cases because the presence of additional features results in higher epistemic uncertainties.

To test the robustness of the Bayesian model against out-of-distribution inputs, we simulate a sensor failure event. This is shown in Figure 5. To simulate an optical probe icing event between 10 sec and 40 sec, the original photomultiplier signal is replaced by white noise with mean 0.05 and amplitude 0.015. During this simulated probe failure, the epistemic uncertainty of the Bayesian neural network greatly increases, indicating that the inputs are highly dissimilar to conditions experienced during training. Predictions remain reasonable despite the uninformative optical signal, although they become much less accurate.

We use the technique of integrated gradients to produce feature-level attribution plots, which can tell us how much a particular predictor influenced the prediction for a particular input. For example, the attribution plots in Figure 6 shows this technique being applied to two data points. The first one is a data point just prior to the first instability event in a run, where a large jump in the dynamic pressure fluctuation amplitude is forecast by the model. This model uses p'_{FFT}, p'_{rms} and the flowrate control signals as input. The most prominent feature is the large positive attribution for the oxygen flowrate control signal

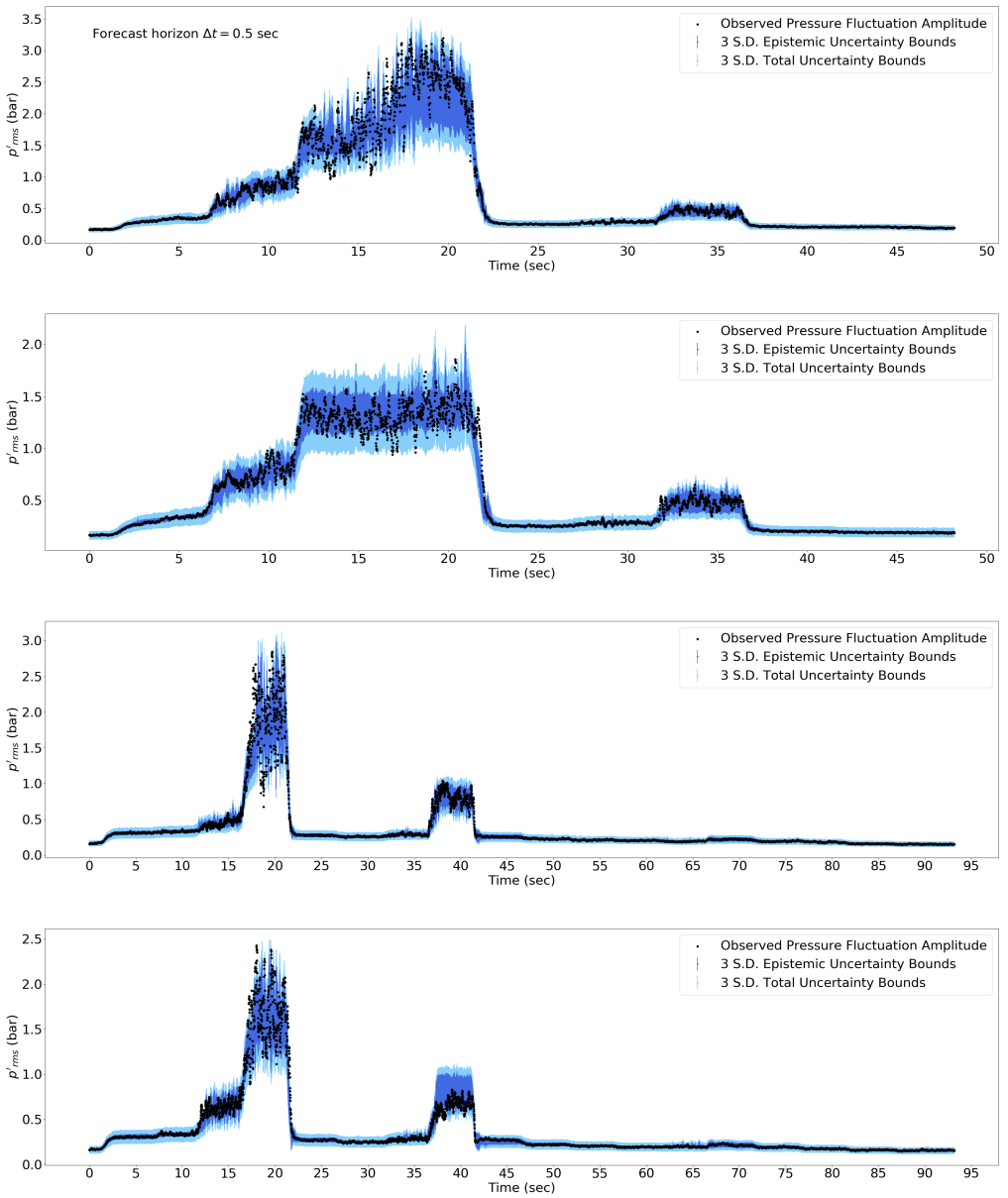

Figure 4. Forecast p'_{rms} values for the four experimental runs based on sensor data prior to $t - 0.5$ sec, superimposed on observed p'_{rms} values at time t. The model was trained on the other three runs in each case. This model used p'_{FFT} and p'_{rms} as state variables and order $h = 2$. Actual observations are shown as black dots, 3 standard deviation (S.D.) total uncertainty bounds for the forecast are in light blue and 3 standard deviation (S.D.) epistemic uncertainty bounds are in deep blue.

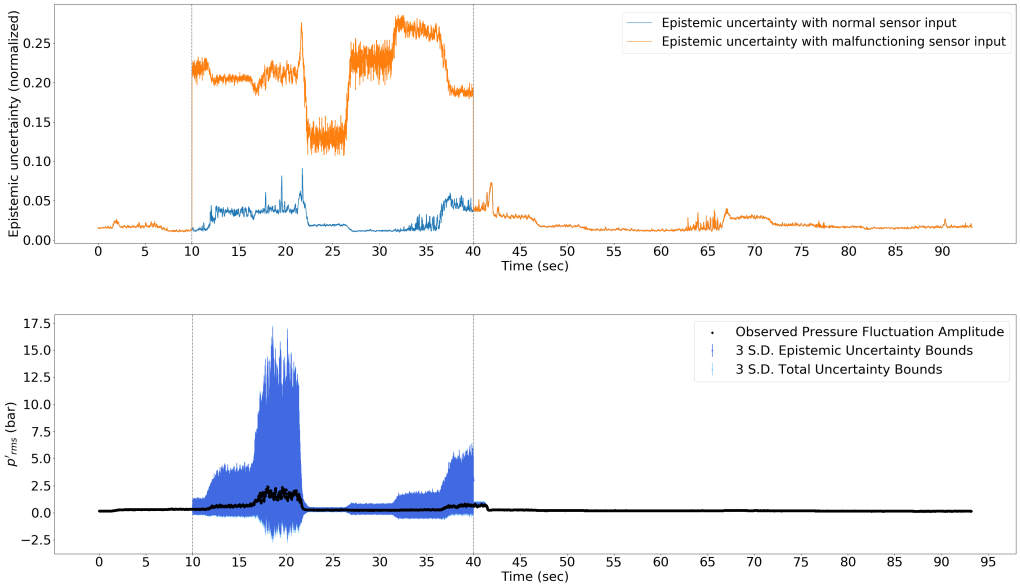

Figure 5. Bayesian Neural Network epistemic uncertainty (top) and p'_{rms} forecasts (bottom) for experimental run no. 4 where a fibre optical probe failure was simulated between $t = 10$ sec and $t = 40$ sec (marked by two grey dotted lines). The model used $I'_{\text{FFT}}, p'_{\text{rms}}$ as input state variables and order = 0. Actual observations are shown as black dots, 3 standard deviation (S.D.) total uncertainty bounds for the forecast are represented by light blue bars and 3 standard deviation (S.D.) epistemic uncertainty bounds are in deep blue. The light blue bars are barely discernible from the deep blue bars in the failure region because the total uncertainty is almost entirely dominated by the large epistemic uncertainty.

$\dot{m}_{O2}$, whose feature value is also large for this datapoint. This is expected, because a large $\dot{m}_{O2}$ would increase the p_{cc} and ROF, bringing the combustor closer to instability. The prediction is also modulated by contributions from the p'_{FFT} features. This datapoint has a concentration of power around the 10 kHz spectrum and low power in the mid to higher frequencies. There is a positive contribution from the 10 kHz frequency component of the spectrum, which is the frequency of the 1T mode, as well as from the 16.66k Hz frequency. This shows that the model has learned that an increased concentration of acoustic power around the frequency of the 1T mode and away from the mid-frequencies is an indication that the system is close to instability. The second data point is from stable operation. The attributions in this case have mostly negative values because the prediction of the model is strongly negative. The current pressure fluctuation amplitude p'_{rms} is close to -1.0 in this case and naturally has a large negative attribution. The $\dot{m}_{H2}$ input (here -0.78) also contributes negatively because low ROFs are stable. Finally, the high concentration of power in the mid-frequencies also tells the network that this operating point will stay stable.

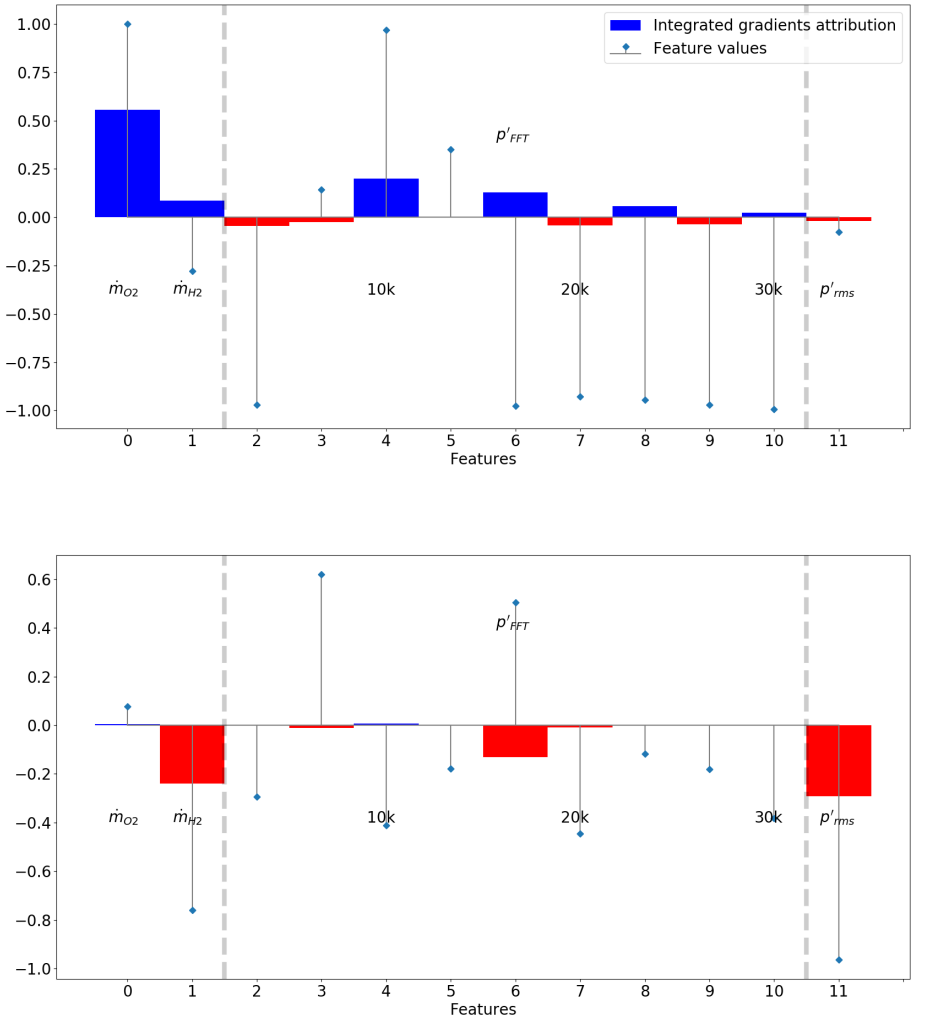

Figure 6. Integrated gradient attributions and input feature values (normalized to have zero mean and unit standard deviation) for two datapoints from experimental run no. 4 at t = 16.0 seconds (top, immediately preceding an instability) and t = 60.0 seconds (bottom, stable). The model used p'_{FFT} and p'_{rms} as state variables and order = 0. The blue dots show the normalized values of the features for this particular datapoint and the bars represent the integrated gradient attributions for each feature. For example, feature 6 (spectral power around the 16.66 kHz frequency) in the top plot has a large negative value and a positive integrated gradients attribution, meaning that the network has learned that low power in this frequency band is a potential indicator of instability.

Conclusions

In this paper, we present the application of an uncertainty-aware machine learning algorithm to forecast instabilities using multimodal sensor data in a realistic, thermoacoustically unstable rocket thrust chamber. We predict the amplitude of dynamic pressure fluctuations 500 ms in advance using temperature, static pressure, high-frequency dynamic pressure and OH* radiation intensity data recorded during test runs performed with the cryogenic rocket thrust chamber BKD. We perform a rigorous cross-validation of the autoregressive Bayesian neural network and find that our models are able to anticipate all the instability events with varying accuracy ($25 - 40\%$ lower RMSEs on the transition subset compared to baseline), depending on the combination of inputs chosen. The inclusion of longer histories and additional sensor information, in general, boosts predictive accuracy. Models that use the dynamic pressure signal, in particular, show a measurable improvement in the forecast, confirming the findings of the literature on this topic. We gain insight into the predictions made by the neural network by using Integrated Gradients, which reveals the contribution made by each input feature to the final prediction.

A key benefit of our Bayesian approach is the ability to quantify the uncertainty in our predictions. In safety-critical machines such as rocket engines, a diagnostic tool must be robust to unexpected events such as sensor failures. We compute the log-likelihood of the test data given model predictions and find that uncertainties are well-characterized. We also simulate a sensor failure event in a model whose inputs included the spectrum of the fibre-optical probe signal and demonstrate the robustness of Bayesian algorithms. Integrated Gradients helps us identify important features and how they influence a particular prediction of the neural network.

Future work will focus on exploring the effectiveness of these tools on DLR's other experimental datasets, such as the LOX-methane tests, which also displayed injection-coupled acoustic instabilities (Klein et al. (2020)). We will also explore the generalizability of data-driven diagnostics across multiple datasets and combustors, using approaches that utilise transfer learning or domain invariance. Hardware upgrades are currently being undertaken at the test facility that will allow the implementation of these methods into the engine control algorithm (Waxenegger-Wilfing et al. (2020)).

Acknowledgements

The authors would like to thank the crew of the P8 test bench and also Stefan Gröning for preparing and performing the test runs. Furthermore, we thank Wolfgang Armbruster for helping us process and understand the data. Ushnish Sengupta is an Early Stage Researcher within the MAGISTER consortium which receives funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 766264.

References

Armbruster W, Hardi JS, Suslov D and Oschwald M (2019) Injector-driven flame dynamics in a high-pressure multi-element oxygen-hydrogen rocket thrust chamber. *Journal of Propulsion and Power* 35: 632–644. DOI: 10.2514/1.B37406.

Blundell C, Cornebise J, Kavukcuoglu K and Wierstra D (2015) Weight Uncertainty in Neural Networks URL http://arxiv.org/abs/1505.05424.

Dowling AP and Morgans AS (2005) Feedback control of combustion oscillations. *Annu. Rev. Fluid Mech.* 37: 151–182.

Eldredge JD and Dowling AP (2003) The absorption of axial acoustic waves by a perforated liner with bias flow. *Journal of Fluid Mechanics* 485: 307.

Fröhlke K, Haberzettl A, Haidn O, Heinrich S, Sion M, Vuillermoz P, Froehlke K, Haberzettl A, Haidn O, Heinrich S et al. (1997) First hot fire test campaign at the french/german research facility p8. In: *33rd Joint Propulsion Conference and Exhibit*. p. 2929.

Gangopadhyay T, Ramanan V, Akintayo A, K Boor P, Sarkar S, Chakravarthy SR and Sarkar S (2021) 3d convolutional selective autoencoder for instability detection in combustion systems. *Energy and AI* 4: 100067. DOI:https://doi.org/10.1016/j.egyai.2021.100067. URL https://www.sciencedirect.com/science/article/pii/S2666546821000215.

Gotoda H, Nikimoto H, Miyano T and Tachibana S (2011) Dynamic properties of combustion instability in a lean premixed gas-turbine combustor. *Chaos* DOI:10.1063/1.3563577.

Gröning S, Hardi JS, Suslov D and Oschwald M (2016) Injector-driven combustion instabilities in a Hydrogen/Oxygen rocket combustor. *Journal of Propulsion and Power* 32(3): 560–573. DOI:10.2514/1.B35768.

Hachijo T, Masuda S, Kurosaka T and Gotoda H (2019) Early detection of thermoacoustic combustion oscillations using a methodology combining statistical complexity and machine learning. *Chaos: An Interdisciplinary Journal of Nonlinear Science* 29(10): 103123. DOI:10.1063/1.5120815. URL https://doi.org/10.1063/1.5120815.

Jha DK, Virani N and Ray A (2018) Markov modeling of time series via spectral analysis for detection of combustion instabilities. In: *Handbook of Dynamic Data Driven Applications Systems*. Springer, pp. 123–138.

Juniper MP (2018) Sensitivity analysis of thermoacoustic instability with adjoint helmholtz solvers. *Physical Review Fluids* 3(11): 110509.

Juniper MP and Sujith RI (2018) Sensitivity and nonlinearity of thermoacoustic oscillations. *Annual Review of Fluid Mechanics* DOI:10.1146/annurev-fluid-122316-045125.

Klein S, Börner M, Hardi JS, Suslov D and Oschwald M (2020) Injector-coupled thermoacoustic instabilities in an experimental lox-methane rocket combustor during start-up. *CEAS Space Journal* : 1–13.

Kobayashi T, Murayama S, Hachijo T and Gotoda H (2019a) Early detection of thermoacoustic combustion instability using a methodology combining complex networks and machine learning. *Physical Review Applied* 11(6): 064034.

Kobayashi T, Murayama S, Hachijo T and Gotoda H (2019b) Early detection of thermoacoustic combustion instability using a methodology combining complex networks and machine learning. *Phys. Rev. Applied* 11: 064034. DOI:10.1103/PhysRevApplied.11.064034. URL https://link.aps.org/doi/10.1103/PhysRevApplied.11.064034.

Koschel W, Haidn J and Krülle G (1996) P8-the new french/german test facility for h2/o2-high pressure rocket engine combustion research. In: *11th World Hydrogen Energy Conference, Stuttgart, June 23-28, 1996*. DECHEMA, Frankfurt, Germany, pp. 1815–1825.

Li H, Barnaghi P, Enshaeifar S and Ganz F (2019) Continual learning using bayesian neural networks.

Lieuwen T (2005) Online Combustor Stability Margin Assessment Using Dynamic Pressure Data. *Journal of Engineering for Gas Turbines and Power* DOI:10.1115/1.1850493.

Lin T, Horne B, Tino P and Giles C (1996) Learning long-term dependencies in narx recurrent neural networks. *IEEE Transactions on Neural Networks* 7(6): 1329–1338. DOI:10.1109/72.548162.

McCartney M, Indlekofer T and Polifke W (2020) Online prediction of combustion instabilities using machine learning. *Journal of Engineering for Gas Turbines and Power* .

McCartney M, Sengupta U and Juniper M (2021) Reducing uncertainty in the onset of combustion instabilities using dynamic pressure information and bayesian neural networks. In: *Proceedings of ASME Turbo Expo 2021*. ASME.

Mondal S, Chattopadhyay A, Mukhopadhyay A and Ray A (2021) Transfer learning of deep neural networks for predicting thermoacoustic instabilities in combustion systems. *Energy and AI* : 100085DOI:https://doi.org/10.1016/j.egyai.2021.100085. URL https://www.sciencedirect.com/science/article/pii/S2666546821000392.

Mondal S, Ghalyan NF, Ray A and Mukhopadhyay A (2018) Early detection of thermoacoustic instabilities using hidden markov models. *Combustion Science and Technology* .

Murugesan M and Sujith R (2016) Detecting the onset of an impending thermoacoustic instability using complex networks. *Journal of Propulsion and Power* 32(3): 707–712.

Nair V and Sujith RI (2014) Multifractality in combustion noise: predicting an impending combustion instability. *Journal of Fluid Mechanics* 747: 635–655. DOI:10.1017/jfm.2014.171. URL https://www.cambridge.org/core/product/identifier/S0022112014001712/type/journal{_}article.

Nair V, Thampi G, Karuppusamy S, Gopalan S and Sujith RI (2013) Loss of chaos in combustion noise as a precursor of impending combustion instability. *International Journal of Spray and Combustion Dynamics* 5(4): 273–290. DOI:10.1260/1756-8277.5.4.273. URL https://doi.org/10.1260/1756-8277.5.4.273.

Nair V, Thampi G and Sujith R (2014) Intermittency route to thermoacoustic instability in turbulent combustors. *Journal of Fluid Mechanics* 756: 470.

Neal RM (1996) *Bayesian Learning for Neural Networks*, *Lecture Notes in Statistics*, volume 118. New York, NY: Springer New York. ISBN 978-0-387-94724-2. DOI:10.1007/978-1-4612-0745-0. URL http://link.springer.com/10.1007/978-1-4612-0745-0.

Nóvoa A and Magri L (2022) Real-time thermoacoustic data assimilation. *Journal of Fluid Mechanics* 948: A35.

Oefelein JC and Yang V (1993) Comprehensive review of liquid-propellant combustion instabilities in f-1 engines. *Journal of Propulsion and Power* 9(5): 657–677.

Pearce T, Zaki M, Brintrup A, Anastassacos N and Neely A (2018) Uncertainty in Neural Networks: Bayesian Ensembling URL http://arxiv.org/abs/1810.05546.

Sarkar S, Chakravarthy SR, Ramanan V and Ray A (2016) Dynamic data-driven prediction of instability in a swirl-stabilized combustor. *International Journal of Spray and Combustion Dynamics* 8(4): 235–253.

Sengupta U, Rasmussen C and Juniper M (2020) Bayesian machine learning for the prognosis of combustion instabilities from noise. *Journal of Engineering for Gas Turbines and Power* .

Sundararajan M, Taly A and Yan Q (2017) Axiomatic attribution for deep networks. In: *International Conference on Machine Learning*. PMLR, pp. 3319–3328.

Watanabe D, Manako H, Onga T, Tamura T, Ikeda K and Isono M (2016) Combustion stability improvement of le-9 engine for booster stage of h3 launch vehicle. *Mitsubishi Heavy Industries Technical Review* 53(4): 28.

Waxenegger-Wilfing G, Dresia K, Oschwald M and Schilling K (2020) Hardware-in-the-loop tests of complex control software for rocket propulsion systems. In: *71st International Astronautical Congress, Virtual*. IAC.

Welch P (1967) The use of fast fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms. *IEEE Transactions on audio and electroacoustics* 15(2): 70–73.

Zinn BT (1970) A theoretical study of non-linear damping by helmholtz resonators. *Journal of Sound and Vibration* 13(3): 347–356.