Dynamic Data-Driven Combustor Design for Mitigation of Thermoacoustic Instabilities

Pritthi Chattopadhyay
Mechanical Engineering Department,
Pennsylvania State University,
University Park, PA 16802

Sudeepta Mondal
Mechanical Engineering Department,
Pennsylvania State University,
University Park, PA 16802

Asok Ray[1]
Fellow ASME
Mechanical Engineering Department,
Pennsylvania State University,
University Park, PA 16802
e-mail: axr2@psu.edu

Achintya Mukhopadhyay
Mechanical Engineering Department,
Jadavpur University,
Kolkata 700 032, India

A critical issue in design and operation of combustors in gas turbine engines is mitigation of thermoacoustic instabilities, because such instabilities may cause severe damage to the mechanical structure of the combustor. Hence, it is important to quantitatively assimilate the knowledge of the system conditions that would potentially lead to these instabilities. This technical brief proposes a dynamic data-driven technique for design of combustion systems by taking stability of pressure oscillations into consideration. Given appropriate experimental data at selected operating conditions, the proposed design methodology determines a mapping from a set of operating conditions to a set of quantified stability conditions for pressure oscillations. This mapping is then used as an extrapolation tool for predicting the system stability for other conditions for which experiments have not been conducted. Salient properties of the proposed design methodology are: (1) It is dynamic in the sense that no fixed model structure needs to be assumed, and a suboptimal model (under specified user-selected constraints) is identified for each operating condition. An information-theoretic measure is then used for performance comparison among different models of varying structures and/or parameters and (2) It quantifies a (statistical) confidence level in the estimate of system stability for an unobserved operating condition by using a Bayesian nonparametric technique. The proposed design methodology has been validated with experimental data of pressure time-series, acquired from a laboratory-scale lean-premixed swirl-stabilized combustor. [DOI: 10.1115/1.4040210]

1 Introduction

Instabilities in combustion systems of gas turbine engines are usually related to the spontaneous excitation of one or more natural acoustic modes of the combustor. These phenomena are typically manifested by large-amplitude self-sustained oscillations in the combustion chamber, which results from a feedback loop established between the heat release rate from the flame and the combustion chamber acoustics [1]. The problem is aggravated with the implementation of low emission technologies like lean-premixed combustion, which is susceptible to thermoacoustic instabilities [2]. The detrimental effects of such instabilities include generation of externally audible tones at intolerable levels and sustained high-amplitude pressure oscillations that cause mechanical stresses in the structural components of the combustor, leading to thermomechanical damage. Hence, mitigation of thermoacoustic instabilities is a critical issue for both design and operation of combustion systems.

The dynamics of combustion are described by coupled nonlinear partial differential equations, which cause difficulties in analytically modeling these instabilities. Hence, solely model-based design optimization strategies may not be reliable for combustion systems that involve several input parameters [3,4]. From the perspectives of instrumentation and control (I&C), several active and passive measures have been investigated by researchers (e.g., Ref. [5]). These measures involve additional hardware and/or software to be integrated with the combustor system in order to reduce instabilities.

From the design perspectives, it is necessary to identify an optimum (or a near-optimum) combustor geometry, which is expected to produce stable operation of the system. The design procedure can be realized in several ways. One of the common methods is to use network-based models for studying the collective response of different components in the combustor [6], where high-fidelity numerical simulations can be used for predicting the dynamic behavior at different regimes of operation [7]. Such simulations are often computationally expensive as a design tool, and the networked platforms may lack the memory and accuracy requirements for different dynamic regimes of the combustor. Due to these limitations, data-driven methods have been used by several researchers, and such endeavours are well established particularly with respect to characterization and control of combustion dynamics [8,9]. However, from the viewpoint of developing a design methodology for thermoacoustically stable operations, these technologies are rather uncommon.

This technical brief proposes a dynamic data-driven methodology [10] for design of combustors by utilizing the domain knowledge based on experimental observations at different operating conditions. To this end, a Bayesian nonparametric statistical algorithm has been developed to predict the system response at unknown operating conditions, thus enabling an informed choice of conditions at which the combustion system can be reliably operated.

The work reported here is an extension of the authors' recent work [11], where the underlying concepts were introduced for prediction of the system response with a fixed-structure data-driven model. In this work, experimental data from a laboratory-scale swirl-premixed combustor apparatus have been used to generate a stability map in the parameter space with a variable-structure dynamic data-driven model that is capable of effectively capturing the history of process dynamics. To this end, advanced modeling tools [12] have been used for analyzing time series of pressure oscillations under different combinations of combustor length (l), inlet air velocity (u), and equivalence ratio (ϕ). A stability map is constructed from the (limited-amount) available data, which is then used to predict the system response by extrapolating the aforesaid combustor parameters. By taking advantage of the information embedded in the time series, the reported work yields more representative features as compared to those extracted in Ref. [11]; in addition, the performance is validated for 18 data sets, instead of just one [11]. The designer is thus provided with a quantitative tool of statistical estimation of system stability at new parameters without the need for additional (and possibly expensive and time-consuming) experimentation.

Copyright © 2019 by ASME

2 Background and Mathematical Preliminaries

This section presents pertinent background and mathematical preliminaries regarding construction of information-theoretic measures and D-Markov machines that form the core concept of the proposed design methodology; further details are reported in Ref. [12].

2.1 Probabilistic Finite State Automaton.
A probabilistic finite state automaton (PFSA) is constructed by symbolization of time series of measured signals [10,12], which requires partitioning (also known as quantization) of the data. The signal space is thus partitioned into a finite number of cells that are labeled as symbols, where the number of cells is identically equal to the cardinality $|\Sigma|$ of the (symbol) alphabet Σ. If the value of a data point at a given instant is located in a particular cell, then it is coded with the symbol associated with that cell. As such, a symbol from the alphabet Σ is assigned to each (signal) value corresponding to the cell where it belongs; details are reported in Refs. [13] and [14]. Thus, a (finite) array of symbols, called a symbol string (or symbol block), is generated from the (finite-length) time series. The following definitions, which are available in the standard literature, are recalled for the sake of completeness.

DEFINITION 2.1. *A finite state automaton (FSA) G, having a deterministic algebraic structure, is a triple (Σ, Q, δ) where:*
Σ is a (nonempty) finite alphabet with cardinality $|\Sigma|$.
Q is a (nonempty) finite set of states with cardinality $|Q|$.
$\delta : Q \times \Sigma \to Q$ is a state transition map.

DEFINITION 2.2. *A symbol block, also called a word, is a finite-length string of symbols belonging to the alphabet Σ, where the length of a word $w \triangleq s_1 s_2 \cdots s_\ell$ with $s_i \in \Sigma$ is $|w| = \ell$, and the length of the empty word ϵ is $|\epsilon| = 0$. The parameters of FSA are extended as:*
The set of all words constructed from symbols in Σ, including the empty word ϵ, is denoted as $\Sigma^\star$.
The set of all words, whose suffix (respectively, prefix) is the word w, is denoted as $\Sigma^\star w$ (respectively, $w\Sigma^\star$).
The set of all words of (finite) length ℓ, where $\ell > 0$, is denoted as Σ^ℓ.

DEFINITION 2.3. *A PFSA K is a pair (G, π), where:*
The deterministic FSA G is called the underlying FSA of the PFSA K.
The probability map $\pi : Q \times \Sigma \to [0, 1]$ is called the morph function (also known as symbol generation probability function) that satisfies the condition: $\sum_{\sigma \in \Sigma} \pi(q, \sigma) = 1$ for all $q \in Q$.
Equivalently, a PFSA is a quadruple $K = (\Sigma, Q, \delta, \pi)$, where
The alphabet Σ of symbols is a (nonempty) finite set, i.e., $0 < |\Sigma| < \infty$, where $|\Sigma|$ is the cardinality of Σ.
The set Q of automaton states is (nonempty) finite, i.e., $0 < |Q| < \infty$, where $|Q|$ is the cardinality of Q.
The state transition function $\delta : Q \times \Sigma \to Q$.
The morph function $\pi : Q \times \Sigma \to [0, 1]$, where $\sum_{\sigma \in \Sigma} \pi(q, \sigma) = 1$ for all $q \in Q$. The morph function π generates the $(|Q| \times |\Sigma|)$ morph matrix Π.

2.2 Entropy Rate.
This section introduces the notion of entropy rate that, given the current state, represents the predictability of PFSA. Details are reported in Ref. [12].

DEFINITION 2.4. *The entropy rate of a PFSA (Σ, Q, δ, π) is defined as follows:*

$$H(\Sigma|Q) \triangleq - \sum_{q \in Q} \sum_{\sigma \in \Sigma} P(q)P(\sigma|q)\log P(\sigma|q) \tag{1}$$

where $P(q)$ is the (unconditional) probability of a PFSA state $q \in Q$, and $P(\sigma|q)$ is the (conditional) probability of a symbol $\sigma \in \Sigma$ emanating from the PFSA state $q \in Q$.

Next the notion of a metric is introduced to quantify the distance between two PFSA.

DEFINITION 2.5. *Let $K_1 = (\Sigma, Q_1, \delta_1, \pi_1)$ and $K_2 = (\Sigma, Q_2, \delta_2, \pi_2)$ be two PFSA with a common alphabet Σ. Let $P_1(\Sigma^j)$ and $P_2(\Sigma^j)$ be the steady-state probability vectors of generating words of length j from the PFSA K_1 and K_2, respectively, i.e., $P_1(\Sigma^j) \triangleq [P(w)]_{w \in \Sigma^j}$ for K_1 and $P_2(\Sigma^j) \triangleq [P(w)]_{w \in \Sigma^j}$ for K_2. Then, the metric as the distance between the PFSA K_1 and K_2 is defined as*

$$\Phi(K_1, K_2) \triangleq \lim_{n \to \infty} \sum_{j=1}^{n} \frac{\|P_1(\Sigma^j) - P_2(\Sigma^j)\|_{\ell_1}}{2^{j+1}} \tag{2}$$

where the norm $\|\star\|_{\ell_1}$ indicates the sum of absolute values of the elements in the vector $\star$.

2.3 D-Markov Machines.
This section introduces the pertinent concepts and definitions that are necessary to construct a D-Markov machine. The PFSA model of a D-Markov machine generates symbol strings $\{s_1 s_2 \cdots s_\ell : \ell \in \mathbb{N}, \forall s_j \in \Sigma\}$ on the underlying Markov process. The morph function π implicitly alludes to the fact that the PFSA satisfies the Markov condition, where generation of a symbol *only* depends on this state. However, from the perspectives of PFSA construction from a symbol sequence, the states are implicit and generation of the next symbol may depend on the complete history of the symbol sequence. In the construction of a D-Markov machine [10], generation of the next symbol depends only on a *finite* history of at most D consecutive symbols, i.e., a symbol block of length not exceeding D. A formal definition of the D-Markov machine follows.

DEFINITION 2.6. *A D-Markov machine [10] is a PFSA in the sense of Definition 2.3, and it generates symbols that solely depend on the (most recent) history of at most D last symbols, where the positive integer D is called the depth of the machine. Equivalently, a D-Markov machine is a statistically stationary stochastic process $S = \cdots s_{-1} s_0 s_1 \cdots$, where the probability of occurrence of a new symbol depends only on the last D symbols, i.e.,*

$$P[s_n| \cdots s_{n-D} \cdots s_{n-1}] = P[s_n|s_{n-D} \cdots s_{n-1}]$$

Consequently, for $w \in \Sigma^D$ (see Definition 2.2), the equivalence class $\Sigma^\star w$ of all (finite-length) words, whose suffix is w, is qualified to be a D-Markov state that is denoted as w.

2.3.1 State Splitting in D-Markov Machines. The process of splitting a state q is executed by replacing the symbol block q by its *branches* as described by the set $\{\sigma q : \sigma \in \Sigma\}$ of words. Maximum reduction of the entropy rate is the governing criterion for selecting the state to split. In addition, the generated set of states must satisfy the self-consistency criterion, which only permits a unique transition to emanate from a state for a given symbol. If $\delta(q, \sigma)$ is not unique for each $\sigma \in \Sigma$, then the state q is split further. In the state splitting algorithm, a stopping rule is constructed by specifying the threshold parameter as a maximal number of states $N_{\max}$. The operation of state splitting is described in Ref. [12] as Algorithm 1.

2.3.2 State Merging in D-Markov Machines. Once state splitting is performed, the resulting D-Markov machine is a statistical representation of the symbol string under consideration. Depending on the choice of alphabet size $|\Sigma|$ and depth D, the number of states after splitting may run into hundreds. The motivation behind the state merging is to reduce the number of states, while preserving the D-Markov structure of the *PFSA*. A stopping rule is constructed by specifying an acceptable threshold η on the distance $\Phi(\cdot, \cdot)$ between the merged PFSA and the PFSA generated from the original time series (see Eq. (2)); details are provided in Ref. [12].

The next task is to decide which states have to be merged. States that behave similarly (i.e., have similar morph probabilities) have a higher priority for merging. The similarity of two

states, $q, q' \in Q$, is measured in terms of morph functions (i.e., conditional probabilities) of future symbol generation as the distance between the two rows of the estimated morph function $\hat{\pi}$ corresponding to the states q and q'. The ℓ_1-norm (i.e., the sum of absolute values of the vector components) has been adopted to be the distance function as seen below:

$$\mathcal{M}(q, q') \triangleq \|\hat{\pi}(q, \cdot) - \hat{\pi}(q', \cdot)\|_{\ell_1} = \sum_{\sigma \in \Sigma} |\hat{\pi}(q, \sigma) - \hat{\pi}(q', \sigma)| \tag{3}$$

A small value of $\mathcal{M}(q, q')$ indicates that the two states have close probabilities of generating each symbol. Note that this measure is bounded above as $\mathcal{M}(q, q') \leq 2 \, \forall q, q' \in Q$, because $0 \leq \sum_{\sigma \in \Sigma} \hat{\pi}(q, \cdot) \leq 1$ and $0 \leq \sum_{\sigma \in \Sigma} \hat{\pi}(q', \cdot) \leq 1$. Now the procedure of state merging is briefly described below.

First, the two closest states (i.e., the pair of states $q, q' \in Q$ having the smallest value of $\mathcal{M}(q, q')$) are merged [12]. Subsequently, distance $\Phi(\cdot, \cdot)$ (see Eq. (2)) of the merged PFSA from the initial symbol string is evaluated. If $\Phi < \eta$, where η is a specified threshold, then, the machine structure is retained, and the states next on the priority list are merged. On the other hand, if $\Phi > \eta$, then, the process of merging the given pair of states is aborted, and another pair of states with the next smallest value of $\mathcal{M}(q, q')$ is selected for merging. This procedure is terminated if no such pair of states exist, for which $\Phi < \eta$. The state merging procedure is described in Ref. [12] as Algorithm 4.

2.4 Gaussian Process Regression. Gaussian process (GP) regression [15] is a nonparametric tool that can model a wide range of relations between an operating condition and its system response without making any other specific assumptions on the relation. To this end, a Gaussian process is modeled as a finite linear combination of random variables with multivariate jointly Gaussian distributions. In particular, a collection of random variables $\{\xi(t) : t \in T\}$, where T is an index set, is said to be generated from a Gaussian process with mean function $m(\cdot)$ and covariance function $k(\cdot, \cdot)$ if, for any finite set of elements $t_1, .., t_l \in T$, the corresponding random variables $\xi(t_1), .., \xi(t_l)$ have multivariate jointly Gaussian distribution as

$$\begin{bmatrix} \xi(t_1) \\ .. \\ \xi(t_l) \end{bmatrix} \sim N\left(\begin{bmatrix} mm(t_1) \\ .. \\ m(t_l) \end{bmatrix}, \begin{bmatrix} kk(t_1, t_1) ... k(t_1, t_l) \\ .. \\ k(t_l, t_1) ... k(t_l, t_l) \end{bmatrix} \right) \tag{4}$$

where $m(t) \triangleq E[\xi(t)]$ is the mean function, and $k(t, t') \triangleq E[(\xi(t) - m(t))(\xi(t') - m(t'))]$ is the covariance function.

Let $X = \{x_i\}$ and $Y = \{y_i\}$, $i = 1, \ldots, n$, be the training data set, where X denotes a set of operating conditions, and Y denotes the corresponding set of system responses. In the GP regression algorithm, it is assumed that $y = \xi(x) + \varepsilon$, where ε is independent and identically distributed (iid) additive noise, $N(0, \sigma^2)$. Then, a zero-mean Gaussian process prior $GP(0, K)$ is assumed for the function ξ. By concatenating the training and testing sets of operating conditions as $[X, X^{\text{test}}]$, the marginal distribution of the respective responses $[\xi(X), \xi(X^{\text{test}})]$ is also multivariate jointly Gaussian
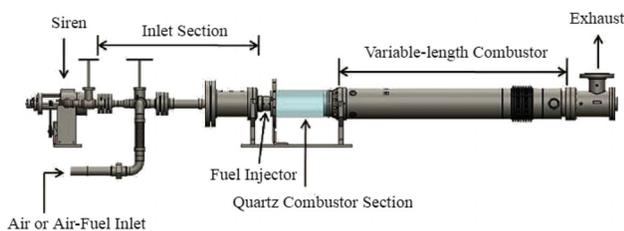
Fig. 1 Schematic diagram of the combustion apparatus

$$\begin{bmatrix} Y \\ Y^{\text{test}} \end{bmatrix} \sim N\left(0, \begin{bmatrix} K(X, X) + \sigma^2 I & K(X, X^{\text{test}}) \\ K(X^{\text{test}}, X) & K(X^{\text{test}}, X^{\text{test}}) + \sigma^2 I \end{bmatrix} \right) \tag{5}$$

This leads to: $Y^{\text{test}} | Y \sim N(\mu^{\text{test}}, \Sigma^{\text{test}})$ where

$$\mu^{\text{test}} = K(X^{\text{test}}, X)(K(X, X) + \sigma^2 I)^{-1} Y \tag{6}$$

$$\Sigma^{\text{test}} = K(X^{\text{test}}, X^{\text{test}}) + \sigma^2 I$$
$$- K(X^{\text{test}}, X)(K(X, X) + \sigma^2 I)^{-1} K(X, X^{\text{test}}) \tag{7}$$

Thus, the algorithm predicts the mean and variance of the system response for every test condition. Instead of a zero-mean prior (i.e., $E[\xi(x)] = 0$), a mean function $m(x)$ could also be incorporated into the prior.

3 Description of the Experimental Apparatus

This section describes the laboratory-scale experimental apparatus, which is a swirl-stabilized, lean-premixed combustor that has been used for collection of time series data to validate the proposed design methodology. Figure 1 depicts a schematic diagram of the variable-length combustor apparatus [11,16], consisting of an inlet section, an injector, a combustion chamber, and an exhaust section. There is an optically accessible quartz section followed by a variable-length steel section. This laboratory-scale combustor has been used to generate the experimental data. Tests were conducted at a nominal combustor pressure of 1 atm over a range of operating conditions, as listed in Table 1. In each test, pressure dynamics in the combustion chamber were measured to study the mechanisms of combustion instability. The measurements were collected at a sampling rate of 8192 Hz, and the data were collected for 8 s. The results were classified as unstable when limit cycle oscillations set in. On the other hand, stable cases were characterized by irregular fluctuations, caused by turbulence in the combustor. These irregular fluctuations, which have traditionally been referred to as combustion noise, have recently been shown to be high-dimensional deterministic chaos [17].

4 Technical Approach

This section describes the technical details of the proposed combustor design methodology. It starts with the pertinent steps for construction of the design methodology as an extension of the work reported in Ref. [11] that also provides a block diagram of the design concept. The underlying algorithm has been developed in the following two steps:

Step 1. Feature extraction from pressure time series: the features characterize the combustion dynamics.

Step 2. GP regression to identify the relation between the operating condition and the corresponding system response as a function of the features extracted in step 1. This inferred relation is then used for predicting response for each unobserved operating condition.

For specified values of inlet air velocity (u) and equivalence ratio (ϕ)), the problem under consideration is to determine a mapping from combustor length (l) to the degree of stability of the combustion system; similar exercises can be carried out with respect to other design variables as well. To achieve this goal, the entire dataset has been divided into 18 subsets of constant u and ϕ. The data in each subset thus consist of a set of combustor

Table 1 Operating conditions of the combustion apparatus

Parameters	Values
Equivalence ratio (ϕ)	0.525; 0.55; 0.60; 0.65
Inlet velocity (u)	25–50 m/s in 5 m/s increments
Combustor length (l)	25–59 in in 1 in increments

lengths (*l*) and the corresponding pressure time series. There is a single time series for each operating condition (i.e., each value of *l*). The data in each subset are first randomly divided into training and testing sets in the proportion of 80% and 20%, respectively. The root-mean-square (rms) value, P_{rms}, of pressure is then calculated for each time series. The combustor length in the training set with the lowest P_{rms} is taken to be the one representing the nominal state of the combustion system for the given values of *u* and ϕ.

The next step involves symbolization of each time series to obtain a symbol string. In this paper, maximum entropy partitioning [13] has been adopted to discretize the range space of the time series. The state-splitting and merging algorithms, are then applied on the symbol string to determine the *D*-Markov structure best describing the given time series. This requires finding the number and identity of states that provides a good model for the given time series under the chosen values of the following parameters: alphabet size ($|\Sigma|$), threshold (η) for the upper bound on the distance between merged PFSA and the symbol string, and maximum number (N_{max}) of states after state splitting (see Sec. 2). Here, $|\Sigma|$ represents the resolution of (time series) data partitioning, η determines accuracy of the PFSA state representation, and N_{max} reflects the importance given by the user to computational complexity.

After the suboptimal *D*-Markov structure is determined, its morph matrix (see Definition 2.3) is calculated. The above procedure yields a *D*-Markov model for each time series, which might be different in the number and identity of states. Higher the number of states in the PFSA of a *D*-Markov machine, lower is the predictability of the symbol sequence, as expected. Hence, in order to compare between two PFSAs with different number of states, the unnormalized entropy rate (see Definition 2.4) would not be an adequate measure; the rationale is that PFSA with a higher number of states would always have a higher entropy rate. If instead, the entropy rate is divided by the maximum entropy rate for that *D*-Markov model (which is a function of the number of states in the PFSA), the measure would be a reflection of the proportion of the system behavior, which can be explained by the best model under the chosen complexity; the measure would always lie between 0 and 1. Higher values would mean that the best model at that complexity is not able to encode the dynamics of the sequence properly. The PFSAs are chosen as the features here. The difference in normalized entropy rates of two PFSAs, has been used as the divergence measure F_{div}, for comparison of two PFSAs.

The feature, corresponding to the nominal combustor length, is chosen to be the nominal feature for the particular subset of *u* and ϕ. For every combustor length in the training set, the divergence of its corresponding feature from the nominal feature F_{div} is calculated. Hence, the system response in the proposed approach is taken to be the feature divergence. The inputs to the GP regression algorithm thus consist of the pair, namely, combustor length *l* and feature divergence F_{div}, where the objective is to predict F_{div} for different values of *l*. For the GP regression algorithm [15], a variety of mean and covariance functions can be used. The following mean and covariance functions are compared to determine the best combination for a dataset.

(1) Mean function: (i) constant $m(x) = c$, (ii) linear $m(x) = \sum_{i=1}^{J} a^i x^i$, and (iii) sum of the constant and linear terms yields: $m(x) = c + \sum_{i=1}^{J} a^i x^i$, where *J* is the dimension of the input space.
(2) Covariance function: (i) linear $k(x^p, x^q) = x^p * (x^q)'$, where $*$ is the operation of matrix multiplication and (ii) squared exponential automatic relevance determination $k(x^p, x^q) = sf^2 * \exp(-(x^p - x^q)' * P^{-1} * (x^p - x^q)/2)$, where the matrix *P* is diagonal with automatic relevance determination parameters $\ell_1^2, \ldots, \ell_J^2$, and sf^2 is the signal variance.

The loglikelihood of training data has been compared under all combinations of mean and covariance functions. The combination of constant mean function and squared exponential automatic relevance determination covariance function is observed to yield highest likelihood for the selected dataset; hence, the analysis has been performed under this combination. In addition, by virtue of GP regression being a Bayesian algorithm, it is not necessary to know the optimal values of the hyper-parameters (i.e., *c*, $\{\ell_i^2\}$ and sf^2) in the mean and covariance function a priori. The algorithm itself identifies optimal values of these hyper-parameters, which lead to highest log likelihood of the training data. For each subset of constant inlet velocity (*u*) and equivalence ratio (ϕ), a GP regression algorithm is used to determine the mapping from the combustor length *l* to the system response. Using this mapping, the algorithm thus predicts the mean μ and variance σ^2 of the distribution of F_{div} for each combustor length in the testing set.

One of the major reasons for using GP regression for combustor design is its ability to quantify the uncertainty in the estimate of the system response for unobserved combustor lengths. The proposed design methodology estimates the most likely system response (i.e., mean) together with possible variations about the mean, which may result from different sources of uncertainties (e.g., measurement noise and insufficient training data, and the assumptions of GP not holding strictly because of usage of the estimated parameters).
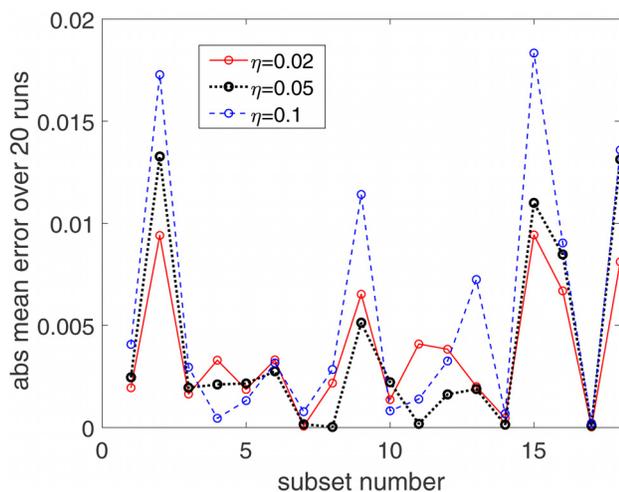
Fig. 2 Mean error in prediction over 20 runs
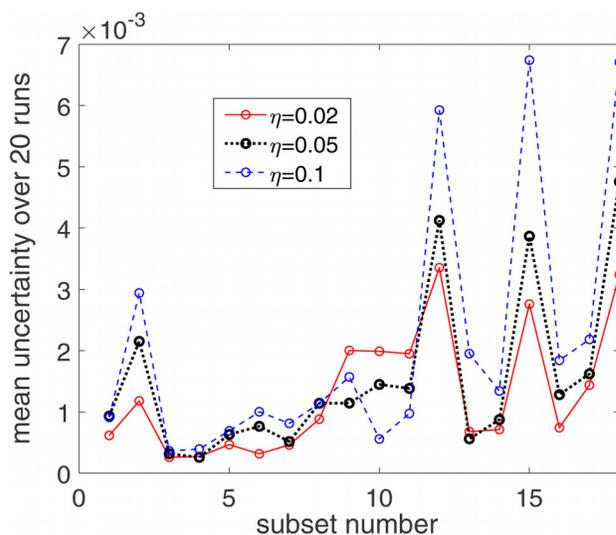
Fig. 3 Mean uncertainty in estimation over 20 runs
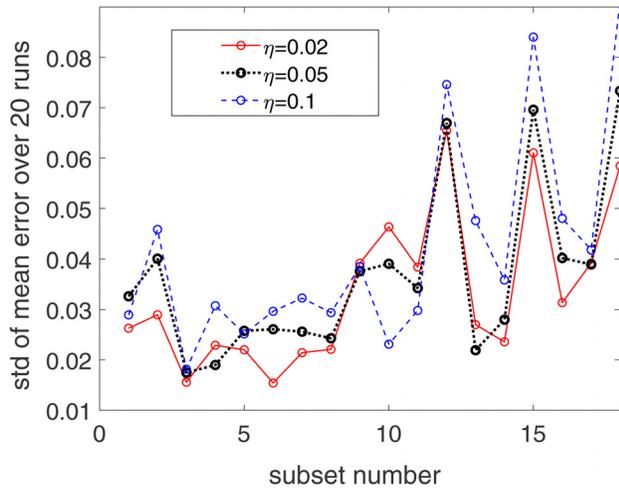
Fig. 4 Standard deviation of mean error over 20 runs

5 Results and Discussion

Given the 18 subsets of constant inlet velocity (u) and equivalence ratio (ϕ), this section analyzes the effects of $|\Sigma|$, N_{max}, and η (see Sec. 2) on the performance of the proposed stability map prediction with the following values:

(1) $|\Sigma| = 3$ and 5.
(2) $N_{max} = 10$, 20, and 30.
(3) $\eta = 0.02$, 0.05, and 0.1.

The underlying algorithm has been executed for different combinations of the following parameters: $|\Sigma|$, η, and N_{max}. Figures 2 and 3, respectively, show the mean error in prediction and mean predicted uncertainty in estimation (i.e., square root of the predicted variance which is calculated by Eq. (7)). These results have been obtained by averaging over 20 runs (i.e., random combinations of training and testing sets, by randomly dividing each subset in the ratio of 80% and 20%, respectively) for different values of η while keeping the other parameters fixed at $|\Sigma| = 5$ and $N_{max} = 20$. The entire procedure of system response prediction has been repeated for all 18 subsets, and the results are reported here. It is observed that, for a majority of the subsets, both mean error in the prediction and uncertainty in the estimate increase as
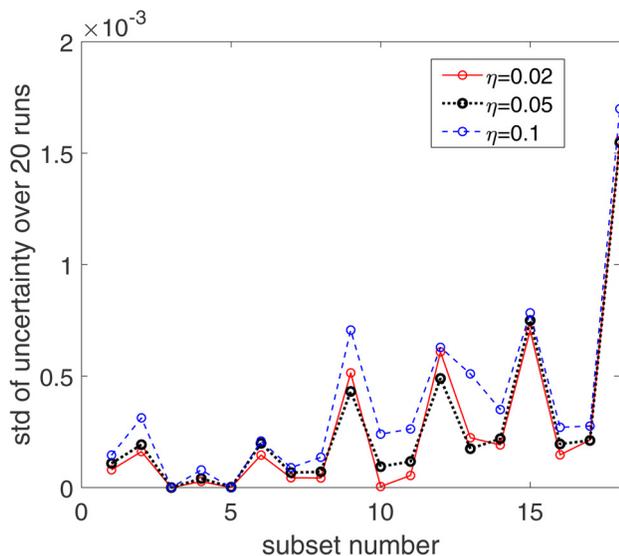
η is increased. Figures 4 and 5, respectively, show the standard deviations of the mean error and predicted uncertainty over 20 runs for different values of η's while keeping the other parameters fixed, for each of the 18 subsets. Again, as η is increased, the generalization ability of the algorithm worsens. Thus, for a majority of the subsets, the performance of the design methodology improves as η is decreased. This suggests that larger values of D are necessary to construct the D-Markov models that encode the process dynamics, implying the need for more memory in Markov models of the combustion process, which is in agreement with the recent findings of Sarkar et al. [8]. The rationale is that a high value of η might result in dissimilar states being merged, leading to inadequacy of the extracted feature which ultimately results in poor prediction of instabilities. Similar behavior has also been observed for other combinations of $|\Sigma|$ and N_{max}. For the best case (i.e., $\eta = 0.02$), the average error in prediction over all 18 subsets is 0.0037, and the average uncertainty in the prediction over all subsets is 0.0013, where the system response (i.e., the difference in normalized entropy rate F_{div}) can vary between 0 and 1.

The system behavior has also been analyzed at stable and unstable conditions of combustion. It is observed that the entropy rate is consistently lower for unstable conditions as compared to stable conditions. Figures 6 and 7, respectively, show the profiles of P_{rms} and entropy rate for one out of the 18 subsets. A reasonable interpretation for this physical phenomenon, is that the system becomes less chaotic as limit cycle oscillations set in, i.e., at the onset of instability. It has also been observed that, even for a set of operating conditions that yield D-Markov models with the same number of states, the entropy rate for the unstable conditions is always lower than that for the stable conditions.

5.1 Feature Extraction: Comparison of $D = 1$ and $D \geq 1$.
The profile of P_{rms}, observed on the experimental apparatus (see Fig. 1) over a sufficiently long time window, has been found to be good indicator of combustion stability, where the system is found to be unstable for $P_{rms} \geq 0.07$ psi. Based on this knowledge, the design algorithm for $D \geq 1$ is now compared with that for $D = 1$ as explained below.

Reiterating again, there are 18 subsets of constant inlet velocity and equivalence ratio. In each subset, for each 8 s time series associated with a unique combustor length, the corresponding P_{rms} is computed. The combustor length associated with the lowest P_{rms} value (P_{rms}^{stable}) is designated as the most stable operating condition (CL_{stable}). For each combustor length, the divergence of P_{rms} from the lowest value (i.e., $P_{rms} - P_{rms}^{stable}$) is computed. Then, from each time series in the subset, features are extracted using symbolic time series analysis, with $D = 1$ and $|\Sigma| = 8$ (see Ref. [11]). The feature associated with CL_{stable} is designated as the nominal feature, and the divergence of all other features from the nominal

Fig. 5 Standard deviation of uncertainty over 20 runs
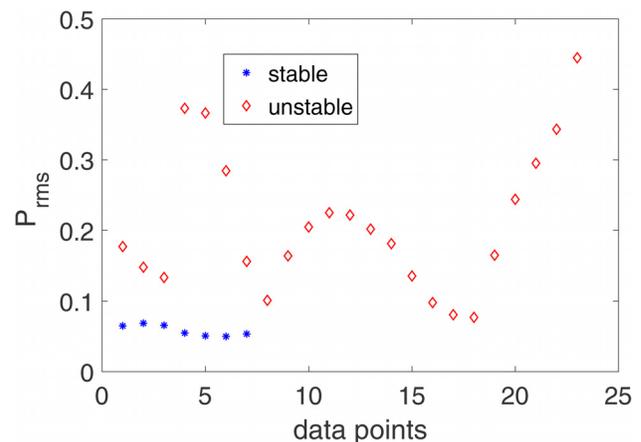
Fig. 6 Profile of P_{rms} plot for a typical subset
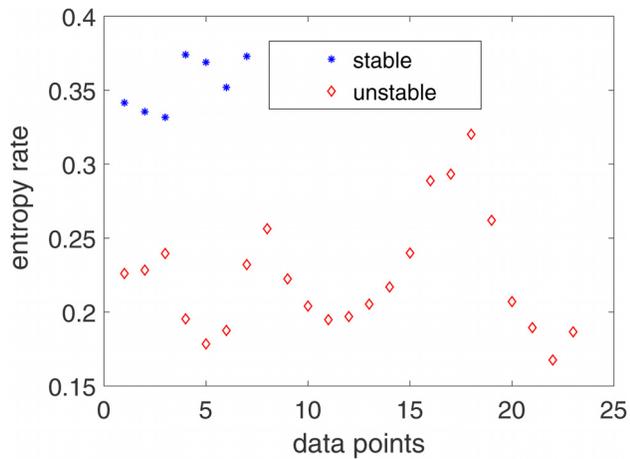
Fig. 7 Entropy rate for the same subset as in Fig. 6

feature is computed. In the second scenario, features are extracted from each time series in the subset by using symbolic time series analysis with $D \geq 1$ (see Sec. 4) for two combinations: (i) $|\Sigma| = 3$, $N_{max} = 20$, $\eta = 0.02$ and (ii) $|\Sigma| = 5$, $N_{max} = 20$, $\eta = 0.02$. Since $\eta = 0.02$ results in the best performance, this η value was chosen for the comparison. The rationale behind choosing $N_{max} = 20$ for this comparison, is that they represent a reasonable computational complexity, and sufficient flexibility given to the algorithm to find a good model fitting the data. Again, the feature associated with CL_{stable} is designated as the nominal feature, and the divergence of all other features from the nominal feature is calculated. Thus, for each subset, a set of P_{rms} divergence is calculated, along with the corresponding feature divergences for both $D = 1$ and $D \geq 1$. These divergence sets are individually normalized (i.e., divided by the maximum value in the set and thus making [0,1] the range of the divergence). Next, the correlations are computed between the P_{rms} divergence, and the feature divergence for $D = 1$, where the correlation coefficient of two random variables is a measure of their linear dependence. Similarly, the correlations between P_{rms} divergence and the feature divergence for $D \geq 1$ are computed. The entire procedure is repeated for all 18 subsets.

Figure 8 shows the correlations of P_{rms} divergence with those of feature divergences for $D = 1$ and $D \geq 1$ with $|\Sigma| = 3$, $N_{max} = 20$, $\eta = 0.02$. The correlation of $D = 1$ feature divergence is consistently lower than that of $D \geq 1$ for 15 out of the 18 subsets. For the remaining three subsets, the correlation value is almost the same for $D = 1$ and $D \geq 1$; the rationale for this behavior is that, for these three subsets, the simplest model (i.e., with $D = 1$) is possibly the best fitting model even after state splitting
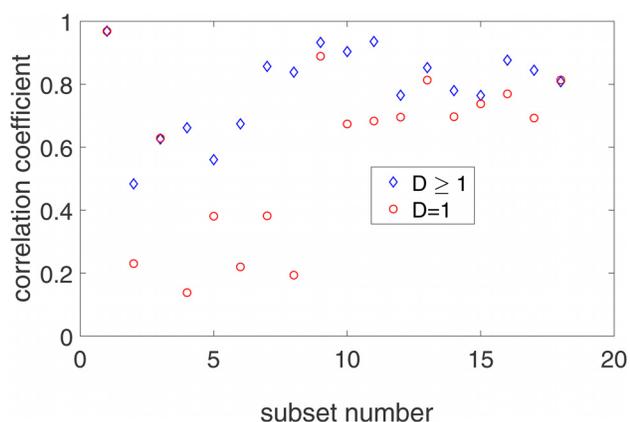
Fig. 8 Correlations of P_{rms} divergence with $D = 1$ and $D \geq 1$ feature divergence for $|\Sigma| = 3$, $N_{max} = 20$, $\eta = 0.02$
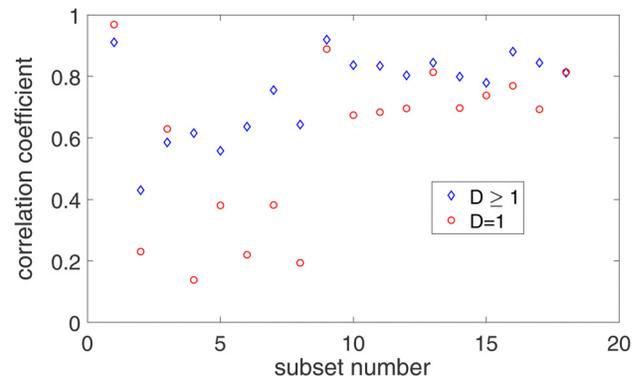
Fig. 9 Correlations of P_{rms} divergence with $D = 1$ and $D \geq 1$ feature divergences for $|\Sigma| = 5$, $N_{max} = 20$, $\eta = 0.02$

and state merging, and it is also possible that the simplest model is much less prone to data overfitting.

To investigate the effects of the alphabet size $|\Sigma|$, which is the most critical design constraint, Fig. 9 shows a similar comparison with $|\Sigma| = 5$, while $N_{max} = 20$ and $\eta = 0.02$ are unchanged. The results are clearly in favor of $D \geq 1$ for 16 out of 18 subsets. In the remaining two subsets, $D = 1$ appears to perform slightly better than $D \geq 1$ possibly due to the effects of data overfitting. Another possible reason for the seemingly better performance of $D = 1$ is that the frequency counting estimation of probability parameters is better due to lesser number of states (i.e., more visits per state and hence more accurate estimation), because of the finite data length [12].

6 Summary, Conclusions, and Future Work

This technical brief has proposed and validated a dynamic data-driven methodology, based on the Bayesian nonparametric theory, as a tool of combustor design for gas turbine engines. The underlying concept is dynamic, because no fixed structure has been assumed for the model that captures the information embedded in the time series. Instead, a suboptimal variable structure of the model is identified for every time series, corresponding to a specific operating condition, under user-selected constraints of complexity and accuracy. An information-theoretic measure is then used for comparative evaluation of different D-Markov model structures. Using experimental data from a laboratory-scale combustion apparatus, it is observed that the proposed design algorithm yields, on the average, more accurate estimates of the system response with high confidence. In addition, the results have been validated with 18 sets in contrast to only one set [11], which confirms the generalization ability of the proposed design algorithm.

The advantages of using higher depth D-Markov models (i.e., $D \geq 1$ instead of $D = 1$) have been investigated. By making the choice of $D \geq 1$, the proposed algorithm is shown to yield features consistently more correlated with P_{rms}, which is used as a good indicator of system stability, as compared to the recently reported work [11] with $D = 1$.

While there are many areas of research that could be pursued, the authors suggest two topics of future research as delineated below:

(1) Theoretical research on integrating dynamic data-driven methods with the current state-of-the-art tools of model-based combustor design.
(2) Evaluating the proposed design methodology on combustors of different geometries and input parameters.

Acknowledgment

Funding Data

References

[1] Matveev, K., 2003, "Thermoacoustic Instabilities in the Rijke Tube: Experiments and Modeling," Ph.D. thesis, California Institute of Technology, Pasadena, CA.

[2] Lieuwen, T. C., and Yang, V., 2005, "Combustion Instabilities in Gas Turbine Engines: Operational Experience, Fundamental Mechanisms, and Modeling," American Institute of Aeronautics and Astronautics, Reston, VA, pp. 3–26.

[3] Lefebvre, A. H., 1999, Gas Turbine Combustion, 2nd ed., CRC Press, Boca Raton, FL.

[4] O'Connor, J., Acharya, V., and Lieuwen, T., 2015, "Transverse Combustion Instabilities: Acoustic, Fluid Mechanic, and Flame Processes," Prog. Energy Combust. Sci., 49, pp. 1–39.

[5] McManus, K. R., Poinsot, T., and Candel, S. M., 1993, "A Review of Active Control of Combustion Instabilities," Prog. Energy Combust. Sci., 19(1), pp. 1–29.

[6] Bauerheim, M., Parmentier, J.-F., Salas, P., Nicoud, F., and Poinsot, T., 2014, "An Analytical Model for Azimuthal Thermo-Acoustic Modes in Annular Chamber Fed by an Annular Plenum," Combust. Flame, 161(5), pp. 1374–1389.

[7] Innocenti, A., Andreini, A., Facchini, B., and Cerutti, M., 2016, "Numerical Analysis of the Dynamic Flame Response and Thermo-Acoustoc Stability of a Full Annular Combustor," ASME Paper No. GT2016-57182.

[8] Sarkar, S., Ray, A., Mukhopadhyay, A., and Sen, S., 2015, "Dynamic Data-Driven Prediction of Lean Blowout in a Swirl-Stabilized Combustor," Int. J. Spray Combust. Dyn., 7(3), pp. 209–242.

[9] Sarkar, S., Chakravarthy, S. R., Ramanan, V., and Ray, A., 2016, "Dynamic Data-Driven Prediction of Instability in a Swirl-Stabilized Combustor," Int. J. Spray Combust. Dyn., 8(4), pp. 235–253.

[10] Ray, A., 2004, "Symbolic Dynamic Analysis of Complex Systems for Anomaly Detection," Signal Process., 84(7), pp. 1115–1130.

[11] Chattopadhyay, P., Mondal, S., Bhattacharya, C., Mukhopadhyay, A., and Ray, A., 2017, "Dynamic Data-Driven Design of Lean Premixed Combustors for Thermoacoustically Stable Operations," ASME J. Mech. Des., 139(11), p. 111419.

[12] Mukherjee, K., and Ray, A., 2014, "State Splitting and Merging in Probabilistic Finite State Automata for Signal Representation and Analysis," Signal Process., 104, pp. 105–119.

[13] Rajagopalan, V., and Ray, A., 2006, "Symbolic Time Series Analysis Via Wavelet-Based Partitioning," Signal Process., 86(11), pp. 3309–3320.

[14] Subbu, A., and Ray, A., 2008, "Space Partitioning Via Hilbert Transform for Symbolic Time Series Analysis," Appl. Phys. Lett., 92(8), p. 084107.

[15] Rasmussen, C. E., and Williams, C. K. I., 2005, Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning), The MIT Press, Cambridge, MA.

[16] Kim, K. T., Lee, J. G., Quay, B. D., and Santavicca, D. A., 2010, "Response of Partially Premixed Flames to Acoustic Velocity and Equivalence Ratio Perturbations," Combust. Flame, 157(9), pp. 1731–1744.

[17] Nair, V., and Sujith, R. I., 2014, "Multifractality in Combustion Noise: Predicting an Impending Combustion Instability," J. Fluid Mech., 747, pp. 635–655.