Energy and AI 5 (2021) 100085

Contents lists available at ScienceDirect

Energy and AI

journal homepage: www.elsevier.com/locate/egyai

Transfer learning of deep neural networks for predicting thermoacoustic instabilities in combustion systems

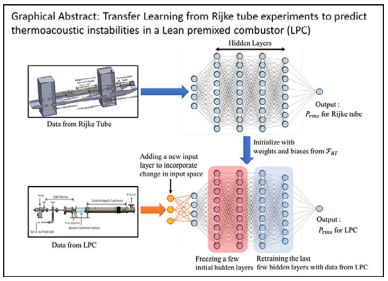
Sudeepta Mondal [a,*], Ashesh Chattopadhyay [b], Achintya Mukhopadhyay [d], Asok Ray [a,c]

[a] Department of Mechanical Engineering, The Pennsylvania State University, University Park, PA 16802, USA
[b] Department of Mechanical Engineering, Rice University, Houston, TX 77005, USA
[c] Department of Mathematics, The Pennsylvania State University, University Park, PA 16802, USA
[d] Department of Mechanical Engineering, Jadavpur University, India

HIGHLIGHTS

- Formulated a machine learning framework for predicting combustion instabilities.
- A transfer learning approach demonstrated using deep neural networks.
- A simple Rijke tube apparatus chosen as a source domain for transfer learning.
- The target domain for transfer learning is a more complex premixed combustor.
- Transfer learning results in better predictions with less training data requirement.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 7 March 2021
Received in revised form 14 April 2021
Accepted 10 May 2021
Available online 12 May 2021

Keywords:
Thermoacoustic instabilities
Deep learning
Transfer learning
Rijke tube
Lean-premixed combustion

ABSTRACT

The intermittent nature of operation and unpredictable availability of renewable sources of energy (e.g., wind and solar) would require the combustors in fossil-fuel power plants, sharing the same grid, to operate with large turn-down ratios. This brings in new challenges of suppressing high-amplitude pressure oscillations (e.g., thermoacoustic instabilities (TAI)) in combustors. These pressure oscillations are usually self-sustained, as they occur within a feedback loop, and may induce severe thermomechanical stresses in structural components of combustors, which often lead to performance degradation and even system failures. Thus, prediction of thermoacoustic instabilities is a critical issue for both design and operation of combustion systems. From this perspective, it is important to identify operating conditions which can potentially lead to thermoacoustic instabilities. In this regard, data-driven approaches have shown considerable success in predicting the instability map as a function of operating conditions. However, often the available data are limited to learn such a relationship efficiently in a data-driven approach for a practical combustion system. In this work, a proof-of-concept demonstration of *transfer learning* is provided, whereby a deep neural network trained on relatively inexpensive experiments in an electrically heated Rijke tube has been adapted to predict the unstable operating conditions for a swirl-stabilized lean-premixed laboratory scaled combustor, for which data are expensive to obtain. The operating spaces and underlying flow physics of these two combustion systems are different, and hence this work presents a strong case of using transfer learning as a potential data-driven solution for transferring knowledge across domains. The results show that the knowledge transfer from the electrically heated Rijke tube apparatus helps in formulating an accurate data-driven surrogate model for predicting the unstable operating conditions in the swirl-stabilized combustor, even though the available data are significantly less for the latter.

* Corresponding author.
E-mail address: sbm5423@psu.edu (S. Mondal).

https://doi.org/10.1016/j.egyai.2021.100085

Nomenclature

TAI	thermoacoustic instabilities
CFD	computational fluid dynamics
DNN	deep neural network
ANN	artificial neural network
AUC	area under the curve
ROC	receiver operating characteristics
SLPM	standard liters per minute
MSE	mean-squared error
RT	the Rijke tube
LPC	lean-premixed combustor
TL	transfer learning
P_{rms}	root-mean-squared pressure fluctuations
L	length of the Rijke tube
x_p	primary heater location in the Rijke tube
x_s	secondary heater location in the Rijke tube
E_s	power input to secondary heater
E_p	power input to primary heater
Q	air flow rate inside the Rijke tube
v_{in}	inlet velocity to the LPC
ϕ	equivalence ratio
l_c	length of the LPC
$\mathcal{F}_{\text{RT}}$	Rijke tube DNN
$\mathcal{F}_{\text{LPC}}$	transfer learned DNN for LPC
z	activation at a layer of a DNN
W	weight matrix connecting 2 layers of a DNN
b	bias for a layer of a DNN
ϵ	normalized L_2 error
Θ	set of trainable parameters for a DNN
λ	regularization parameter
σ	nonlinear activation function
ξ	training data fraction
R^2	coefficient of determination in regression

1. Introduction

The energy portfolio is undergoing a major qualitative change worldwide. International consensus over the need to curb environmental issues like global warming and reducing carbon footprint has led to a decisive shift towards renewable sources of energy. Increasing share of renewables like solar energy and wind energy is redefining the role of thermal power plants to on-demand peak power plants [1]. This intermittent nature of operation and unpredictability of renewable sources like wind or solar energy would require thermal power plants, on the same grid, to be designed to compensate for fluctuations in supply of renewable energy. Consequently, the combustors in fossil-fuel power plants would have to operate with large turn-down ratios. This brings new challenges like thermoacoustic instabilities (TAI) in the design of combustors. Moreover, to achieve carbon neutrality, thermal power plants are increasingly looking for renewable fuels. Renewable fuels, apart from having lower heat contents than common hydrocarbon fuels, suffer from inherent limitations like wide variations in composition due to diversity of energy sources. For example, biogas can contain 40–70% CO_2 depending on its source. Even natural gas, one of the most commonly used gaseous hydrocarbon fuels, shows significant variation in composition, depending on the source of extraction and refining process [2]. Since the goal of the combustor design is to ensure a stable operation in spite of these adverse factors, the design of a modern-day combustor requires high level of flexibility.

One of the crucial issues in lean combustion systems used for low NOx emission is the possibility of TAI. Large turndown ratios and wide variations in fuel composition, leading to variations in flame speed, make the combustion system more vulnerable to TAI. Although TAI have been investigated by a large number of researchers using state of the art computational, experimental and analytical techniques, most of these studies have focused on delineating the stable and unstable regimes or identifying the route to the onset of instability. Stability maps for a fixed design may often fail to clearly demonstrate the extent to which the operating conditions can be modified in order to expect a satisfactory performance in case the design in question fails to produce a stable operation. Moreover, to the best of the authors' knowledge, there does not exist computationally inexpensive analytical or experimental correlations which can provide accurate quantitative estimates of the degree to which TAI can be mitigated as the controller input is varied over a wide range of combustor operating conditions.

The availability of high performance computing systems has led to widespread applications of computational fluid dynamics (CFD) as an efficient design tool for a large variety of thermofluid devices including combustion systems. However, simulation of practical combustion systems under realistic operating conditions is, in general, computationally very expensive owing to the existence of multiple spatial and temporal scales and involvement of complex physics. Often, the physics-based models, themselves, involve significant uncertainties, particularly in estimated values of different parameters. For these reasons, low-order data-driven models have been developed and successfully demonstrated in the investigation of TAI in combustion systems. Kaess et al. [3] have combined a high-fidelity CFD model with a low-order network model to generate a map of combustor operations on the assumption of linear growth rates of both stable and unstable eigen modes, which are then evaluated from a Nyquist plot. A mathematical technique, called Cluster Treatment of Characteristic Roots (CTCR), has been used by Olgac et al. [4] to develop stability charts. Matveev et al. [5] have provided stability maps on a Rijke tube apparatus for various combinations of power and flow-rate based on experimental data and completed the maps by using a reduced-order model of the Rijke tube. Kopitz et al. [6] have performed experimental stability analysis on a model combustor to generate a stability map as a function of equivalence ratio and power output. While using a heat exchanger as a passive controller, Surendran et al. [7] have generated stability maps for a combustor under various configurations of the controller. However, most of these low-order models suffer from a limited range of applicability. Hence these low-order models can only be partially successful because most of them are not based on the actual performance data of the combustion system. To this end, data-driven modeling techniques that use real experimental or system-generated data can be beneficial in predicting the system response under a wide range of operating conditions. Therefore, it is expected that a synergistic combination of model-based and data-driven techniques will provide the designer with a quantitative tool of statistical estimation of a combustion system's probable behavior when a particular control action is taken at operating conditions for which no experimental/simulation data are available. This statistical tool can also be used to learn a stability map for the combustion system, which is given by a data-driven generative model of machine learning.

With the continued growth of computational power over the last decade, deep neural networks (DNNs) have been successfully used in a wide variety of application domains [8]. While data-driven models (mostly based on deep learning) have seen a lot of success in fluid dynamics [8,9], atmospheric science [10], and dynamical systems [11], a major shortcoming of these models is the inability to generalize beyond training distributions (the configuration of the system on which they are trained). For example, a data-driven model trained on turbulent flow with a particular Reynolds number or a particular level of chaoticity cannot *extrapolate* to flows with higher Reynolds number or to systems that are more chaotic than the one they are trained on. This limits the practical usage of these data-driven models for physical systems since it is intractable to re-train the model on copious amounts of data every time one needs to apply them to newer system configurations. In the last few years, we have seen that such shortcomings can

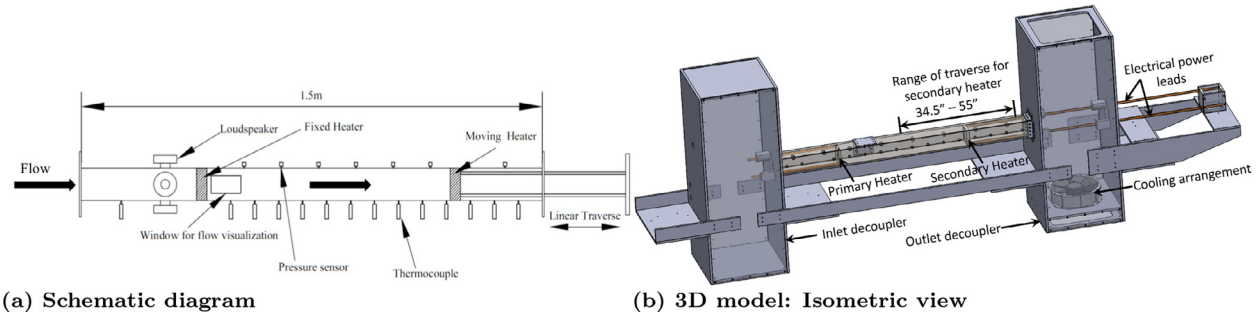


Fig. 1. Schematic diagram and isometric view of the experimental apparatus.

be addressed robustly and efficiently using transfer learning [12–14]. In transfer learning, we re-train (a few of) the layers of a pre-trained deep neural network using a relatively small amount of data (typically in the range of 1% to 10% of the original training size) from the new distribution (new system configuration), but by initializing the network parameters with the same weights as had been obtained during training on data from the initial distribution (old system configuration). In this way, transfer learning allows deep neural networks to generalize to systems involving different physics than what they have been trained with, for example, to more chaotic systems [15], or to turbulent flow with higher Reynolds number [16] with little training data from the new applications and at low computational costs.

Most practical combustors are extremely complex large-scale systems and operate under very demanding conditions like high pressure and temperature. So, it is generally not only difficult but unsafe to run them under abnormal conditions to generate a large volume of training data. Also, in case of practical operating systems, it is mostly not feasible to employ a large number of sensors to characterise their behavior fully. On the other hand, it is possible to develop smaller laboratory-scale systems with requisite instrumentation and operate them under less severe conditions like moderate pressures and temperatures to generate a large volume of training data for abnormal behavior.

The above discussion motivates the theme of the paper, where we seek an answer to the question: *Can we learn a stability map with data from a relatively simpler and easily accessible combustion apparatus, and hope to use it to achieve generalizability in predicting the stability map for a more complex combustion system where we have limited data?* The technique of transfer learning, discussed herein, can prove to be a very useful tool to generate a large volume of data on abnormal behavior like thermoacoustic instability under laboratory conditions, train a predictive model and translate the model to give predictions for full-scale industrial systems. As a proof-of-concept, we have chosen an electrically heated Rijke tube apparatus [17] as the baseline combustion system which is easy and inexpensive to operate and generate data to learn a DNN for predicting the stability map. Then, through transfer learning, the DNN trained on the Rijke tube has been extended to predict the stability map for a much more complex setup: a swirl-stabilized lean-premixed combusting system [18], for which the flow physics is significantly different from that in the Rijke tube. The results have shown that transfer learning allows for better predictions of the steady state root-mean-squared pressure (P_{rms}) values as a function of the operating conditions with limited training data as compared to learning a new network altogether for the swirl-stabilized combustion system.

The manuscript is structured as follows. First, the experimental apparatuses of the electrically heated Rijke tube and the lean-premixed combustor are described (Section 2). Then, the statement of the problem, addressed in this paper, is highlighted in Section 3. Next, the solution methodology and transfer learning architecture is presented in Section 4. Section 5 presents the key results of this work. The paper ends with a summary of the important findings of this work, including directions for future research (Section 6).

2. Description of experimental apparatuses

This section describes the two experimental apparatuses from which data have been procured for demonstration of the transfer learning approach.

2.1. Electrically heated Rijke tube

The data for training the source model for predicting P_{rms} have been generated from a laboratory-scale electrically heated Rijke tube apparatus as shown in Fig. 1. The nature of the acoustic signature obtained from resonating Rijke tubes is very similar to the combustion instabilities obtained in real-life combustors, making them useful laboratory-scale apparatuses for studying thermoacoustics. Electrically heated Rijke tubes possess some of the important characteristics of (fuel-burning) combustion systems like gas turbine combustors (i.e., acoustically compact localized heat sources). The simplicity of the experimentation and their capability of generating clean signals have made electrically heated Rijke tubes popular for investigations on fundamental aspects of thermoacoustic instabilities and their control [19]. Moreover, it is much safer to operate them as compared to their counterparts involving flames from burning fuels.

The apparatus in Fig. 1 comprises a 1.50 m long horizontal Rijke tube with an external (square) cross-section of 0.1 m × 0.1 m and a wall thickness of approximately 6.35 mm. The inlet air flow is supplied by a compressor, which is prone to pressure fluctuations and is also expected to have impurities like moisture. A Parker $P32E$ series air filter-regulator is used to suppress the pressure fluctuations and to filter the impurities. The mean air flow is then controlled through a 0–1000 standard liters per minute (SLPM) Alicat mass flow controller. The horizontal arrangement of the tube allows an independent control of the flow rate and the heater power. Two damping chambers are provided at the two ends of the tube in order to ensure that the tube ends are maintained at pressure nodes under open-open boundary conditions. The damping chambers also serve in decoupling the acoustics of the tube from those of the main air line. The chamber dimensions are 0.45 m ×0.45 m ×1.14 m. The primary and secondary heating elements in the Rijke tube are a couple of square weave 40 × 40 nichrome wire meshes which are brazed to two copper strips each on machinable ceramic frames [5]. The primary heater is placed at about quarter length (L) of the tube from the air input end (i.e., $x_p/L = 0.25$), and the secondary (control) heater is mounted to a sliding arrangement that can move it from $x_s/L = 0.58$ to $x_s/L = 0.93$. This movable arrangement of the secondary heater is aimed at modeling a distributed arrangement of pilot fuel injection as a control measure that interferes with the thermoacoustics from the primary flame to dampen the oscillations. The heating elements are powered by TDK Lambda programmable (0–8 V DC, 0–400 Amps) DC power sources [20]. The downstream half of the tube is insulated to reduce heat loss and to guard against any accidental physical contact with the hot metal surface.

Eight (8) PCB-116B03 pressure sensors and fifteen (15) K-type Omega thermocouple probes have been used for acquiring the acoustic pressure and temperature data from the experiments. The first and the eighth pressure sensors are placed at a distance of 125 mm from the tube ends, while the distance between two consecutive sensors is approximately 180 mm. The sensors are powered through an 8-channel unity gain signal conditioner. The pressure sensors have a high sensitivity of 6 pC/psi and are coupled with in-line charge converters having an amplification of 100 mV/pC. The 13 out of the 15 thermocouple probes are placed downstream of the heater with an intermediate spacing of 90 mm, while the remaining two are placed upstream starting with a distance of approximately 63 mm from the upstream tube end. The thermocouple probes are capable of measurements up to 1360 °C. All sensors are flush-mounted with the inner walls of the Rijke tube to reduce friction in air flow path. The sensor data acquisition and the voltage input to the programmable power supply unit are automated using data acquisition devices from National Instruments (NI) in conjunction with NI LabVIEW 2016. The pressure sensor data is acquired through an NI-9205 (C Series Voltage Input Module) and the thermocouple data through an NI-9213 (C Series Temperature Input Module). The DC voltage supply is controlled through an NI-9264 (C Series Voltage Output Module).

The pressure sensor data are sampled at a rate of 8192 Hz and the acquired data are filtered with a 4th order Butterworth high-pass filter having a cutoff frequency of 40 Hz in order to eliminate low-frequency noises and acoustics from the damping chambers along with other environmental effects. Experiments have been conducted by varying the secondary heater power input (E_s) and air flow rate (Q), for different combinations of primary heater power input (E_p) and secondary heater location (x_s). The ranges of the parameter variations are as follows:

- E_p from 1400 W to 2000 W in steps of 200 W.
- E_s from 200 W to 1400 W in steps of 400 W, along with $E_s = 0$ W (no excitation)
- Q from 140 LPM to 250 LPM in steps of 10 LPM.
- x_s from 35 inches to 55 inches in steps of 5 inches.

A technique, similar to the one demonstrated by Rigas et al. [21], has been used in this paper for data collection:

1. For every run, the air flow-rate (Q) is set at a constant value, with a fixed primary heater power and secondary heater location.
2. First the Rijke tube system is heated to a steady state with the primary heater power input (E_p) of ~ 200 W.
3. Then power inputs to both primary and secondary heaters are abruptly increased to their respective set point values. For some of the parameter combinations, there could be occurrence of TAI even in the steady state, while most of the parameter combinations should remain stable.
4. Pressure data are procured by acoustic sensors for 30 s of operation at a sampling rate of 8192 Hz.
5. The acquired data are filtered with a 10th order Butterworth high-pass filter having a cutoff frequency of 40 Hz in order to eliminate low-frequency noise and acoustics from the damping chambers along with other environmental effects.

After each experiment, a cool-off settling period is maintained to ensure similar steady-state initial temperature fields in the Rijke tube for each experiment. Maintaining similar initial temperatures ensures that the mean velocity of the sound waves in the air remain approximately constant for each experiment, which in turn ensures that the fundamental frequency of the tube is kept constant.

Typically, in a real-life combustor, the location of the primary fuel injector (main burner) is fixed by design, and the demand to be satisfied by a combustor is a particular value of power or thrust output, which in turn dictates the fuel flow-rate into the combustion system (i.e., the chemical energy input into the system). In the Rijke tube, the respective analogous parameters are the position of the primary heater (x_p) and the power into the system via the primary heater power (E_p). Given these two parameters as fixed or demanded, the role of the (secondary heater) controller is to prevent or quench as best as possible any possible thermoacoustic instabilites. In a real-life combustor, having an array of pilot fuel injectors, the movement of the secondary heater would be equivalent to activation and deactivation of selected pilot fuel injectors. Of late, several researchers have looked into mitigating thermoacoustic oscillations using different open loop control strategies [22–26]. While Huhn and Magri [22] have proposed an optimization strategy of minimizing the acoustic energy of chaotic oscillaitons as means to controlling TAI, Deshmukh and Sharma [23] have proposed a passive control strategy of air injection into the unstable flame through radial microjets from the combustor tube wall. Other passive control strategies include the use of perforated liners [26], placing an electrically heated wire at strategic locations in the Rijke tube [25], and using multiple Helmholtz resonators as passive dampers [24]. Rigas et al. [21] and Jamieson et al. [27] studied the stability characteristics of Rijke tubes quantified by measurements of growth and decay rates when a secondary heat source is introduced in the tube. The Rijke tube apparatus which has been considered in this work is a modification of the setup of Jamieson et al. [27], whereby a horizontal Rijke tube with a movable secondary heater has been constructed.

Keeping the primary heater power (E_p) fixed, the variation of the air flow-rate (Q) would be similar to varying the equivalence ratio in a real-life (e.g., gas turbine) combustor, where a higher value of Q implies a leaner fuel-air mixture. It is noted that, in a real-life combustor, Q cannot be directly controlled as it is a function of the demanded thrust or power as well as the ambient operating conditions. However, by exploring the parameter space of Q, it is possible to determine the location-power combination of the secondary source (e.g., pilot fuel) to suppress TAI in the combustor. Although Q is an indirectly controlled parameter, it can be used as a design parameter to quantify the control actions to be undertaken at different equivalence ratios, if the TAI reach undesirable proportions. Therefore, in a real-life combustor, the directly controllable parameters, akin to the location-power combination of the secondary heater in the Rijke tube apparatus, are the location of a pilot fuel injector and flow-rate of the injected pilot fuel. Typically, it is undesirable to have too high a flow rate of the pilot fuel injector because, being located downstream in the flow path, the pilot fuel will tend to shift the flame further downstream and this may cause damage to the blades of the high-pressure turbine.

2.2. Lean-premixed combustor (LPC)

A swirl-stabilized, lean-premixed, laboratory-scale combustor has been used for validation of the proposed algorithm with experimental data. Fig. 2 depicts a schematic diagram of the variable-length combustor apparatus [18], 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. High pressure air is delivered to the apparatus from a compressor system after passing through filters to remove any liquid or particles that might be present. The air supply pressure is set to 180 psig (1.338 MPa) using a dome pressure regulator. The air is pre-heated to a maximum temperature of 250 °C by an 88 kW electric heater. The fuel for this study is natural gas (approximately 95% methane) which is supplied to the system at a pressure of 200 psig (1.475 MPa). The flow rates of the air and natural gas are measured by thermal mass flow meters. The desired equivalence ratio and mean inlet velocity are set by adjusting the flow rates. Further details about the experimental apparatus can be found in the work by Kim et al [18]. Synchronized time series data of pressure oscillations have been collected under different operating conditions, by varying the following parameters:

1. Inlet velocity (v_{in}) from 25 to 50 m/s in steps of 5 m/s.
2. Equivalence ratio (ϕ) as 0.525, 0.550, 0.600 and 0.650.

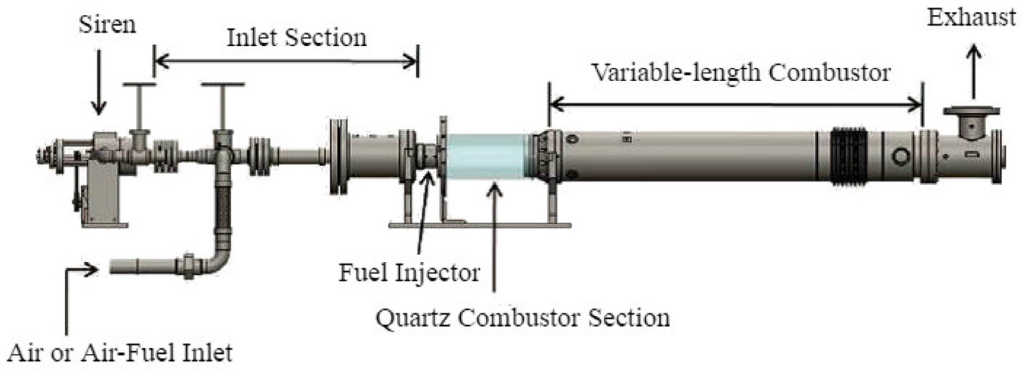


Fig. 2. Schematic diagram of swirl-stabilized combustion apparatus.

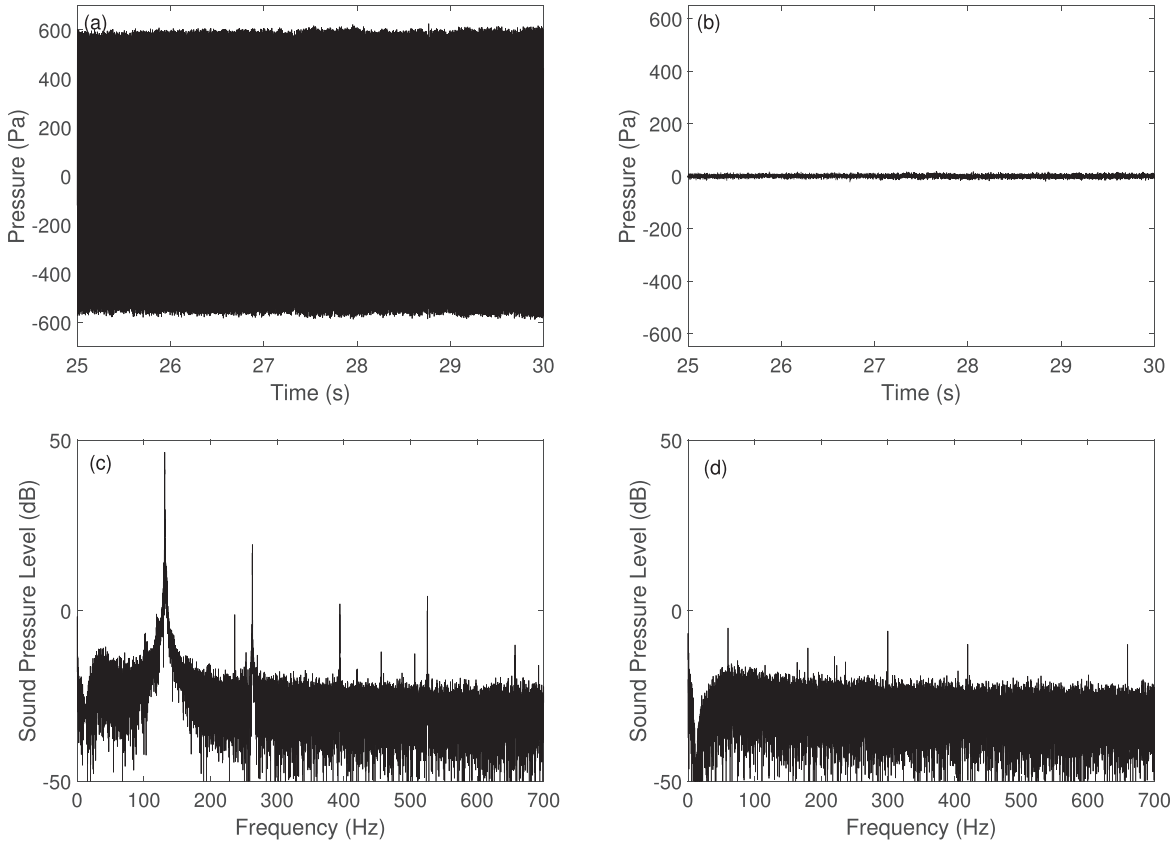


Fig. 3. Representative plot of fluctuating pressure amplitudes (top) and their respective power spectral densities (bottom) at unstable (a and c) and stable (b and d) operating conditions of the Rijke tube. For (a) and (c): E_p = 2000 W, E_s = 0 W, x_s = 34.5 inches, Q = 220 LPM ; and for (b) and (d): E_p = 1600 W, E_s = 200 W, x_s = 50 inches, Q = 170 LPM.

3. Combustor length (l_c) from 25 to 59 inches in steps of 1 inch.

Time series data of pressure oscillations have been collected at a sampling rate of 8192 Hz. The time span of data collection has been 8 s (i.e., 65,536 measurement data per channel) after the system reaches a steady state for each operating condition, which is within the safe limit of operation of the combustor apparatus and which is long enough to provide sufficient information for statistical analysis.

3. Problem statement

Fig. 3 shows a representative plot of the acoustic time series and their respective frequency signatures for an unstable and a stable operating condition of the Rijke tube at the steady state (between 25 s and 30 s of operation). The unstable condition is characterized by a sharp thermoacoustic excitation of the Rijke tube's natural frequency (~ 131 Hz) and its harmonics, which is missing in the stable operation condition. Fig. 4 shows the corresponding plots for a stable and unstable operation of the LPC. An important thing to note is the difference in magnitude of the amplitude of unsteady pressure oscillations in the LPC, which are 2 orders of magnitude higher than that of the Rijke tube. Also there exist high frequency modes which are harmonics of the fundamental frequency (~ 500 Hz) for the depicted unstable operating condition : l_c = 30 inches, ϕ = 0.65 and v_{in} = 50 m/s.

As an operator/designer of a combustor, it is desirable to have an estimate of the P_{rms} value of the acoustic signature in the steady date for an unknown operating condition. This is because P_{rms} is strongly correlated the level of limit cycle instability - higher P_{rms} indicates a more unstable

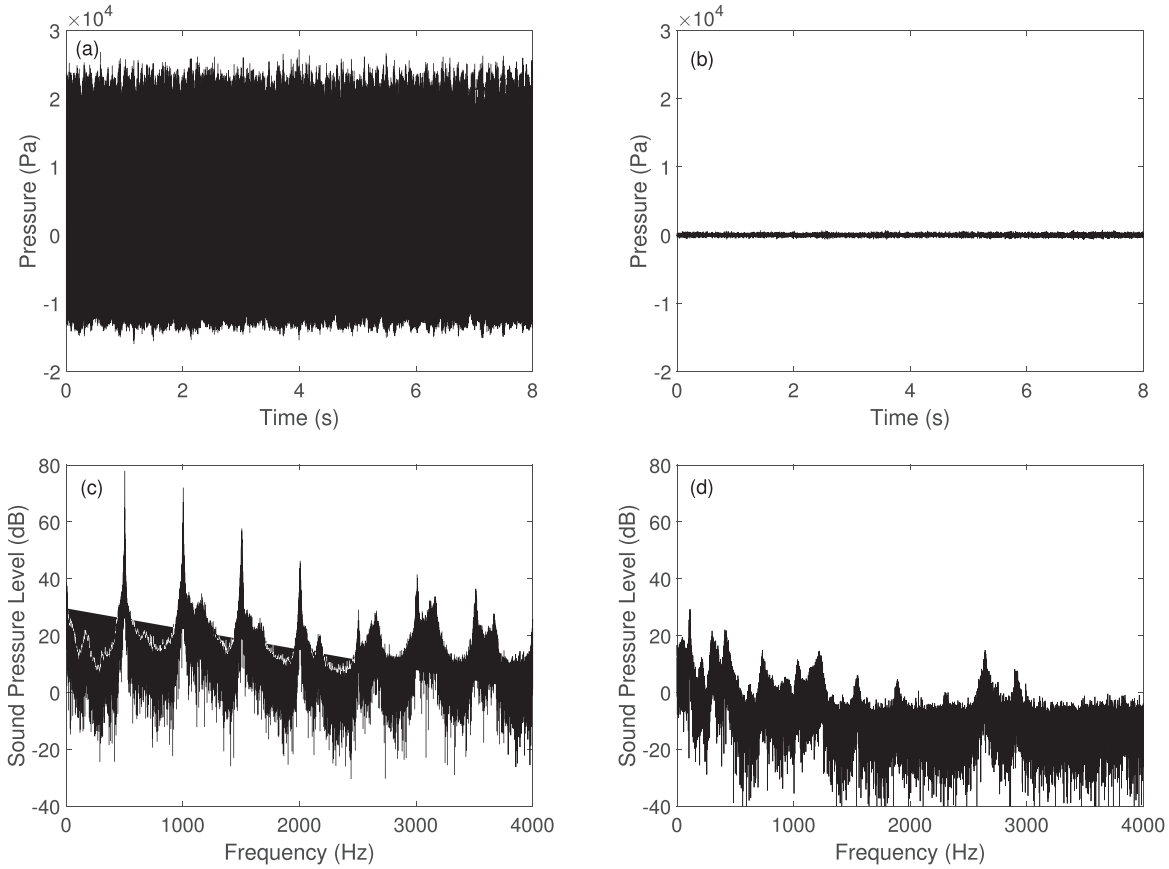


Fig. 4. Representative plot of fluctuating pressure amplitudes (top) and their respective power spectral densities (bottom) at unstable (a and c) and stable (b and d) operating conditions of the lean-premixed combustor. For (a) and (c): $l_c = 30$ inches, $\phi = 0.65$, $v_{in} = 50$ m/s ; and for (b) and (d): $l_c = 38$ inches, $\phi = 0.525$, $v_{in} = 25$ m/s.

system. From this perspective, a stability map can be learned, which is a regression model that predicts P_{rms} as a function of the operating conditions of a combustor. In this paper, first a stability map is learned for the Rijke tube, which is a DNN, mathematically denoted by $\mathcal{F}_{\text{RT}} : \mathbb{R}^4 \to \mathbb{R}$ (since operating conditions of the Rijke tube lie in the 4-dimensional space of (E_p, E_s, x_s, Q)). Then a transfer learning methodology is applied to extend $\mathcal{F}_{\text{RT}}$ to predict the stability map $\mathcal{F}_{\text{LPC}} : \mathbb{R}^3 \to \mathbb{R}$ for the LPC (inputs to LPC lie in the 3-dimensional space of (l_c, ϕ, v_{in})).

3.1. Comparison metrics

The metrics chosen for evaluating the performance of the regression predictions are:

(a) Normalized L_2 error (ϵ): It is defined as the ratio of Euclidean norms of the error of the predicted output ($\hat{y}$) and the true value of the output (y^*), i.e. $\epsilon = \frac{||\hat{y}-y^*||_2}{||y^*||_2}$, calculated on the test set. A low value of ϵ is desired for a good regressor.

(b) R^2 score: It is the Coefficient of Determination assessing the degree to which the predicted output compares with the respective true values, calculated on the test set. A high R^2 score is desired for a good regressor, with $R^2 = 1$ indicating that the regression predictions fit the true data perfectly.

First, the regressor $\mathcal{F}_{\text{RT}}$ is learned till we have a satisfactory regression performance as gauged by ϵ and R^2 score on the Rijke tube dataset. Then, the regressor $\mathcal{F}_{\text{LPC}}$ is transfer learned from $\mathcal{F}_{\text{RT}}$, and the regression performance of $\mathcal{F}_{\text{LPC}}$ is gauged by using the aforementioned metrics on the test dataset from the LPC. Fig. 5 shows a schematic of the transfer learning approach proposed in this paper. More details of the transfer learning framework is provided in the subsequent Section 4.

4. Methodology

This section briefly describes the DNN-based regression framework and transfer learning methodology adopted in this paper. For further details regarding transfer learning and the applicability in different application domains, an interested reader is referred to a comprehensive survey of transfer learning applied to DNNs [28].

4.1. Deep learning with artificial neural networks

An artificial neural network (ANN) [29], is a directed acyclic graph of vertices and edges that propagates an input through successive layers of affine transformation (typically through weight matrices that are optimized through a training process) and non-linear activations to produce an output, so as to learn a functional relationship between the input and output. Typically, these networks have multiple layers of weights which are optimized over many samples of input-output data pairs and are thus very expressive in terms of learning complicated functional relationships. Generally we write neural network functions in linear-algebraic forms as:

$$z^l = \sigma\left(W^l z^{l-1} + b_l\right) \tag{1}$$

where z^l is the activation at layer l, W^l is the weight matrix connecting layer l and layer $l-1$, b^l is the bias at each layer l, and σ is the non-linear activation function. In this work, the network that is trained on experimental data from the Rijke tube ($\mathcal{F}_{\text{RT}}$) has a total of 9 hidden layers. So, l ranges from 1 to 10. The number of neurons in the hidden layers are 4, 8, 16, 64, 128, 64, 32, 16, and 4, respectively. Each layer has an exponential linear unit (ELU) activation function [30] except for the last layer which has a linear activation. To avoid overfitting of

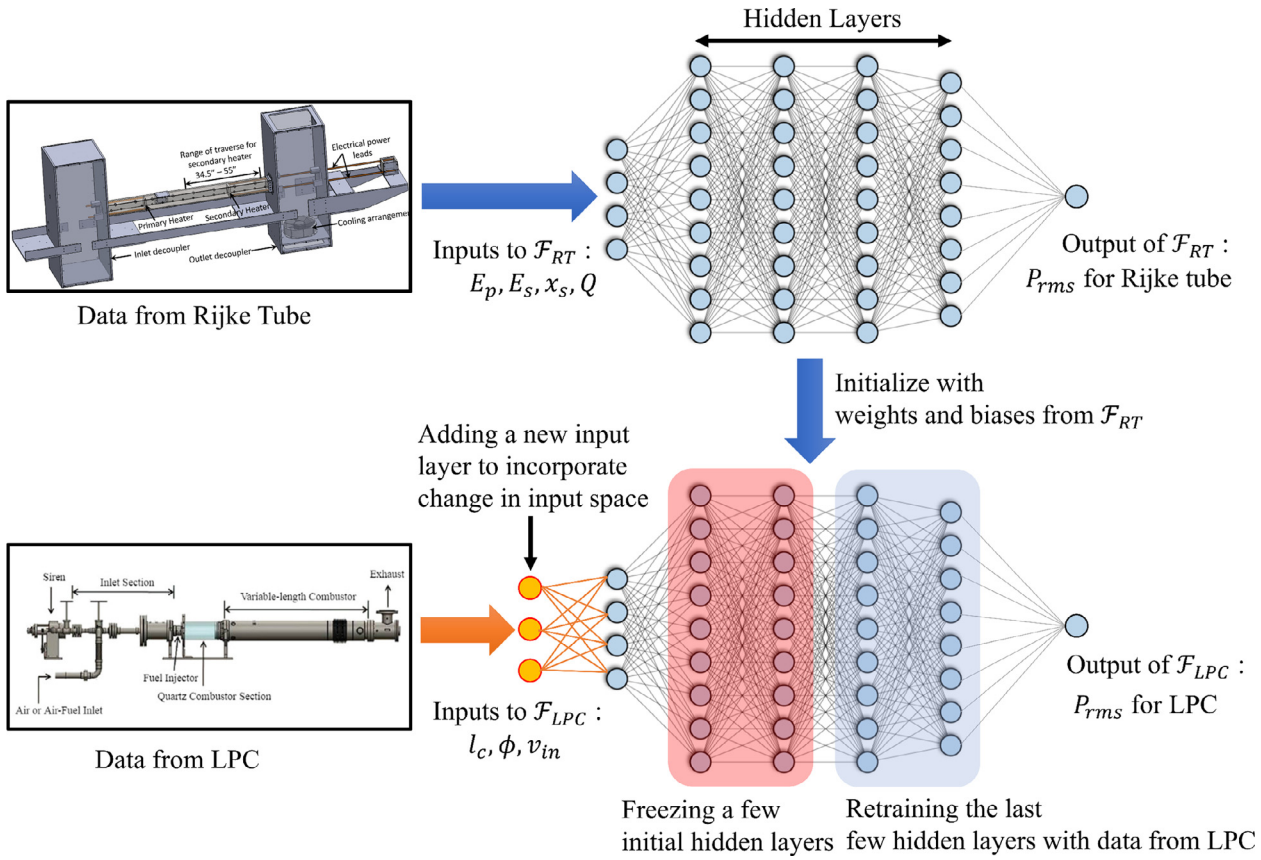


Fig. 5. Schematic overview of the proposed transfer learning approach.

the network, we apply L_2- regularization in each of the layers except for the last one, with a regularizing coefficient 1×10^{-5}. Although 700 epochs are chosen for the training routine, we employ early stopping criterion [31] during the training phase to track the validation score and stop the training process if overfitting is detected. The hyperparameters such as the size of the layers, activation function, and the learning rate (chosen to be 1×10^{-4}) has been chosen after exhaustive trial and error. During training, W^l and b^l, where $l \in [1, 2, \ldots\ 10]$ are obtained using ADAM optimizer [32] through backpropagation [33], to minimize the regluarized mean-squared error (MSE) loss function:

$$\mathcal{L}(y, \hat{y}; \mathbf{\Theta}) = ||y - \hat{y}||_{MSE} + \lambda \sum_{i=1}^{n} \theta_i^2 \tag{2}$$

Here $\mathbf{\Theta} = \{\theta_i\}_{i=1\ldots\ n}$ denotes the set of trainable parameters of the regressor (with total number of parameters $= n$). We refer to the trained parameters, which are the weights and biases of the trained model $\mathcal{F}_{\text{RT}}$ as $W^l_{trained}$ and $b^l_{trained}$, respectively. λ is the regularization parameter, which controls the weights and biases and reduces the likelihood of overfitting. The DNN framework is implemented via Python using Tensorflow [34] - an open source library for large-scale machine learning.

4.2. Transfer learning

As discussed in Section 1, transfer learning has been shown [15,16] to be a robust technique to generalize trained models beyond the training distribution from a simpler system (in this case, data obtained from the Rijke tube experiment) to other, often more complicated systems (in this case, the LPC). When re-training the trained model ($\mathcal{F}_{\text{RT}}$) using data from the LPC, we need to account for the fact that the input dimension of the data from this system has changed from 4 (in the Rijke tube experiment) to 3 (in the LPC). In order to do that, we introduce a new layer with weights and biases W^0 and b^0 in the network with random numbers generated from a Gaussian distribution and re-train this layer and the last three layers of $\mathcal{F}_{\text{RT}}$, i.e., we re-estimate W^l and b^l for $l \in [7, 8, 9, 10]$ using ADAM optimizer with a learning rate of 5×10^{-4}. The weights and biases of this new network, $\mathcal{F}_{\text{LPC}}$ are $\hat{W}^l$ and $\hat{b}^l$ respectively, where $l \in [2, \ldots\ 10]$ are initialized to $W^l_{trained}$ and $b^l_{trained}$. Moreover, only weights and biases corresponding to $l = [7, 8, 9, 10]$ are re-trained, while weights and biases corresponding to $l = [1, 2, \ldots\ 6]$ are not updated (often referred to as *freezing* these layers). Similar to the Rijke tube, overfitting is avoided by employing L_2 regularizer in all layers but the last, with a regularizing coefficient of 1×10^{-6}. Moreover, early stopping is employed as a means to prevent overfitting, along with setting the maximum number of training epochs to 500 and batch size to 16. The batch size for the network is set to 16. The transfer learned model, $\mathcal{F}_{\text{LPC}}$, is then used to make predictions on the data obtained from the LPC.

4.3. Selection of model hyperparameters

Hyperparameters are model parameters which are not learned during the model training process; these parameters are chosen by the user beforehand. In this paper, the number of hidden layers, number of neurons in each layer, the type of activation function in each layer, L_2-regularization coefficient, maximum number of training epochs, batch size, and early stopping scheme are all hyperparameters in the setting of neural networks. In order to select the hyperparameters for the DNNs, a pre-defined grid search strategy has been employed, which is a commonly followed strategy for choosing neural network architectures in practice. The values of hyperparameters which resulted in the best performance (lowest MSE loss for each network) without overfitting have been chosen for each network.

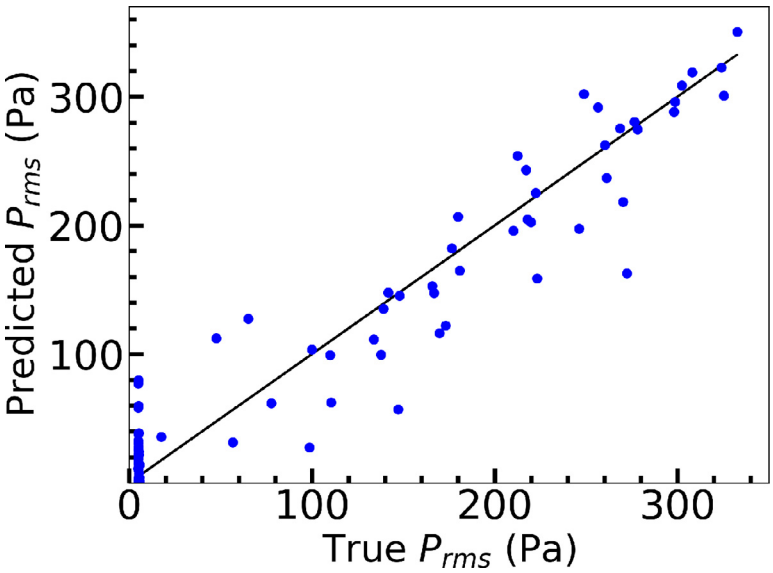


Fig. 6. Parity plot comparing the predictions of $\mathcal{F}_{\mathrm{RT}}$ on the test dataset of the Rijke tube.

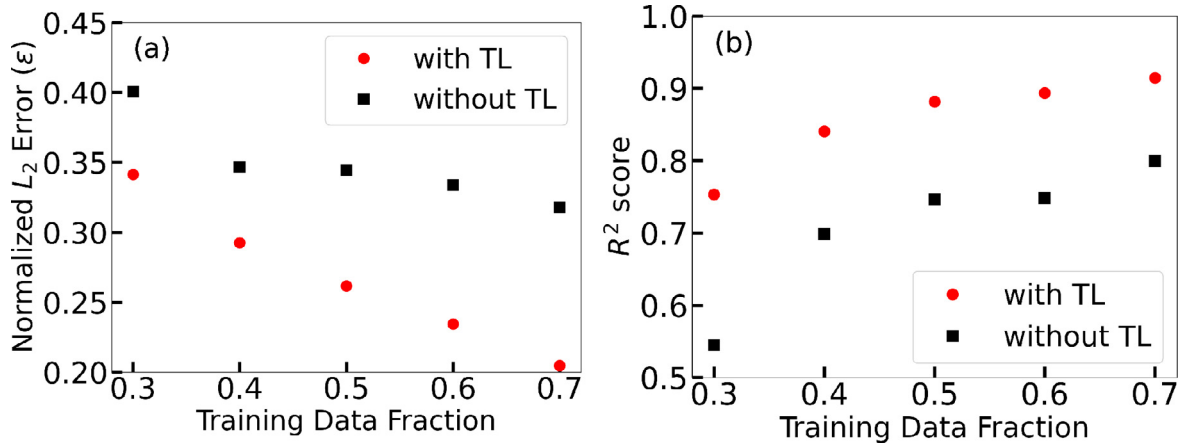


Fig. 7. Comparison of (a) ϵ and (b) R^2 score of predicting the $Prms$ values for unknown operating conditions of the lean premixed combustor, as a function of training data fraction. Red markers indicate prediction metrics when DNN transfer learned from Rijke tube data has been employed, and black markers indicate corresponding prediction performance without transfer learning. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

5. Results and discussions

5.1. Regression performance on Rijke tube data

The Rijke tube data had in total 764 operating conditions for which the steady state P_{rms} values were recorded. 90% of the available data from the Rijke tube was randomly selected and used for training and validation of $\mathcal{F}_{\mathrm{RT}}$, while remaining 10% data was used for testing. Since a well-trained baseline regression model from the source domain is assumed for the transfer learning task, 90% of the data from the Rijke tube has been used to train $\mathcal{F}_{\mathrm{RT}}$ in this paper. ϵ in the training and test sets are 0.20 and 0.19, respectively, which shows satisfactory prediction performance without evidence of overfitting. Fig. 6 shows the parity plot comparing the predicted P_{rms} outputs with the true outputs in the test dataset of the Rijke tube. An R^2 score of 0.90 was achieved by the predictions on the test dataset.

5.2. Transfer learning for LPC

Fig. 7 shows the performance of transfer learning whereby $\mathcal{F}_{\mathrm{RT}}$ is suitably adapted to predict the stability map for LPC, as described in Section 4. The effect of transfer learning (red markers, labelled as 'with TL') has been highlighted by comparing the predictions with DNNs learned on the LPC data without any transfer learning (black markers, labelled as 'without TL'). For networks learned without any transfer learning, the same architecture as $\mathcal{F}_{\mathrm{LPC}}$ is maintained, except that the weights and nodes for all the layers are randomly initialized from a Gaussian distribution and updated through backpropagation, i.e., amounting to training a fresh network.

There were 780 datapoints in total for the LPC for which P_{rms} values were known for the corresponding operating conditions. Hence, transfer learning is performed over a range of training data sizes to investigate its effect as the training data size varies from low to high. It is seen that transfer learning improves the prediction performance, as indicated by lower ϵ and higher R^2 score than when a fresh network is trained (i.e. without transfer learning). The fraction of training data from the LPC has been varied from 0.3 to 0.7, with the transfer learned networks performing better for all training sizes. Care is taken through regularization and early stopping for each network to ensure that overfitting is prevented during the training phase. It is observed that with increasing training data size there is a greater reduction in ϵ with transfer learning. The absolute reduction in ϵ is ~15% when the training size = 0.3, and ~35% when the training size = 0.7. This is possibly due to the fact that with an informed initialization from $\mathcal{F}_{\mathrm{RT}}$, the trainable layers of $\mathcal{F}_{\mathrm{LPC}}$ can

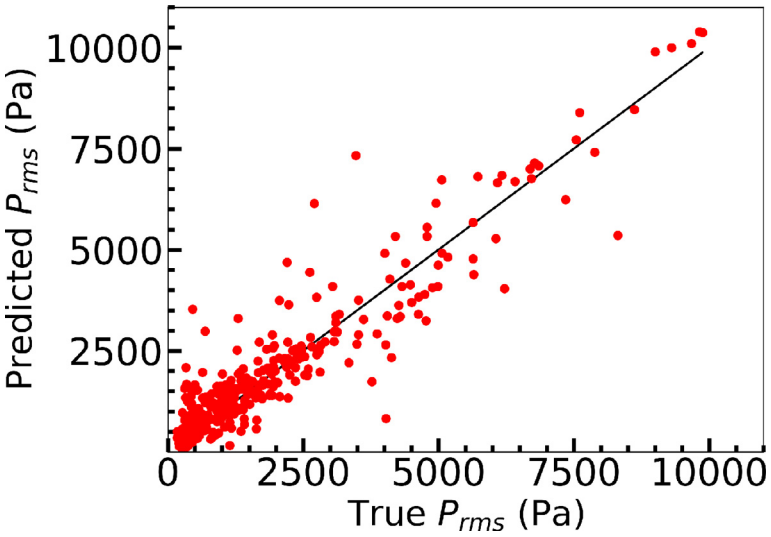


Fig. 8. Parity plot comparing the predictions of $\mathcal{F}_{\text{LPC}}$ on the test dataset of LPC with training size = 0.5.

capture the input-output relationship between the operating conditions and P_{rms} of LPC more efficiently when there is higher data availability, and this makes sense because ϵ is directly related to the MSE error loss function which is minimized during the training phase of the regressor 2. The corresponding increase in R^2 score is ∼36% at training size = 0.3 and ∼16% at training size = 0.7. This result highlights that transfer learning can be effectively employed for training a regression model with satisfactory accuracy for a different application where there might be a paucity of training data to effectively train a good model. Moreover, this application indicates that although the physics inside a Rijke tube is fundamentally very different from that inside a swirl-stabilized lean-premixed combustor, still there is transferable knowledge which can be extracted in the hidden layers of a neural network, and can be gainfully utilized to predict the level of instability for a combustion system.

Fig. 8 shows the predictions of $\mathcal{F}_{\text{LPC}}$ when the training data size = 0.5. The predictions provide the operator with an a priori estimate of the expected level of pressure oscillations in the steady state for a design set of unknown operating conditions, which can be extremely beneficial for the design process. For example, based on the prediction of the P_{rms} a designer can be conservative and rule out the design choices that do not meet the requirements such as an acceptable threshold of limit-cycle oscillations.

5.3. Response surface projections of predicted P_{rms}

$\mathcal{F}_{\text{LPC}}$ serves as an inexpensive surrogate model for predicting P_{rms} over the space of operating conditions (l_c, ϕ, v_{in}). Fig. 9 shows the projections of the predicted response surface $P_{rms}(l_c, \phi, v_{in})$ from the transfer-learned DNN ($\mathcal{F}_{\text{LPC}}$) on different 2-D slices. It is seen in Fig. 9(a) that high values of P_{rms} ($P_{rms} > 8000$ Pa) are predicted for $0.55 < \phi < 0.65$ at the highest velocity level (i.e., $v_{in} = 50$ m/s). Fig. 9(e) and (f) show similar high values of P_{rms} at higher velocity ranges ($40 < v_{in} < 50$ m/s), but at low values of l_c ($30 < l_c < 35$ inches). Such projection maps can be useful for an operator/designer to estimate how the combustor would behave for a range of operating conditions. Moreover, TL provides an efficient way of learning a model that can provide such response surface estimates almost instantaneously, thereby saving the expense of generating sufficiently large amount of data in order to learn an accurate data-driven model.

5.4. Binary classification: detection of instabilities

The P_{rms} predictions from the ANNs can be utilized for detection of thermoacoustic instabilities in the LPC, based on a user-defined threshold (τ). For a binary classification problem, where the classes 'A' and 'B' denote stable and unstable operating conditions, respectively, an algorithm can be designed to classify each input combination of (l_c, ϕ, v_{in}) into either class based on the following hypothesis:

$$P_{rms}(l_c, \phi, v_{in}) \underset{A}{\overset{B}{\gtrless}} \tau \tag{3}$$

The true labels of instability (i.e., binary labels for the input operating conditions based on whether they result in an unstable operation or not) are obtained based on a threshold of 480 Pa [35], which has been provided by the experiments on the LPC. A popularly used criterion for choosing τ for a classifier is the Receiver Operating Characteristics (ROC) curve [36] that is obtained by varying τ so as to reach a trade-off between the probability of successful detection $\left(p_D \triangleq Prob[\text{Decided Class} = B|\text{True Class} = B]\right)$ and the probability of false alarms $\left(p_F \triangleq Prob[\text{Decided Class} = B|\text{True Class} = A]\right)$. Fig. 10 shows the ROC curves using DNN regressor predictions of P_{rms} for the LPC, with and without TL. The range of training data fraction (ξ), as explored in Fig. 7, is also studied in Fig. 10 to understand how the classificaiton performance behaves as a function of the amount of training data from the LPC. In an ROC curve, it is typically desired to have a sharp increase in p_D to a value close to 1 at low values of p_F (ideally, it is desired to have $p_D = 1$ for $p_F = 0$, which results in a perfect classification). It is seen that the classifier using the predictions from transfer-learned networks (e.g., $\mathcal{F}_{\text{LPC}}$) reaches close to $p_D = 0.90$ along with having $p_F < 0.10$ for $\xi = 0.7$. But, even with the highest amount of training data ($\xi = 0.7$), $p_D < 0.70$ with $p_F = 0.10$ for the classifier that relies on the DNN predictions without TL.

A commonly used metric for gauging the performance of classifiers is the area under the curve (AUC) of receiver operating characteristics (ROC) for each classifier [37]. Higher AUC is generally associated with a better overall performance of a classifier, with AUC = 1 for a perfect classifier. Table 1 enlists the AUC for the ROC curves at

Table 1
Comparison of AUCs for binary classification using P_{rms} predictions for the LPC from regressors with and without TL at different values of training data fraction (ξ).

ξ	AUC without TL	AUC with TL
0.3	0.65	0.85
0.4	0.83	0.86
0.5	0.85	0.90
0.6	0.83	0.94
0.7	0.85	0.95

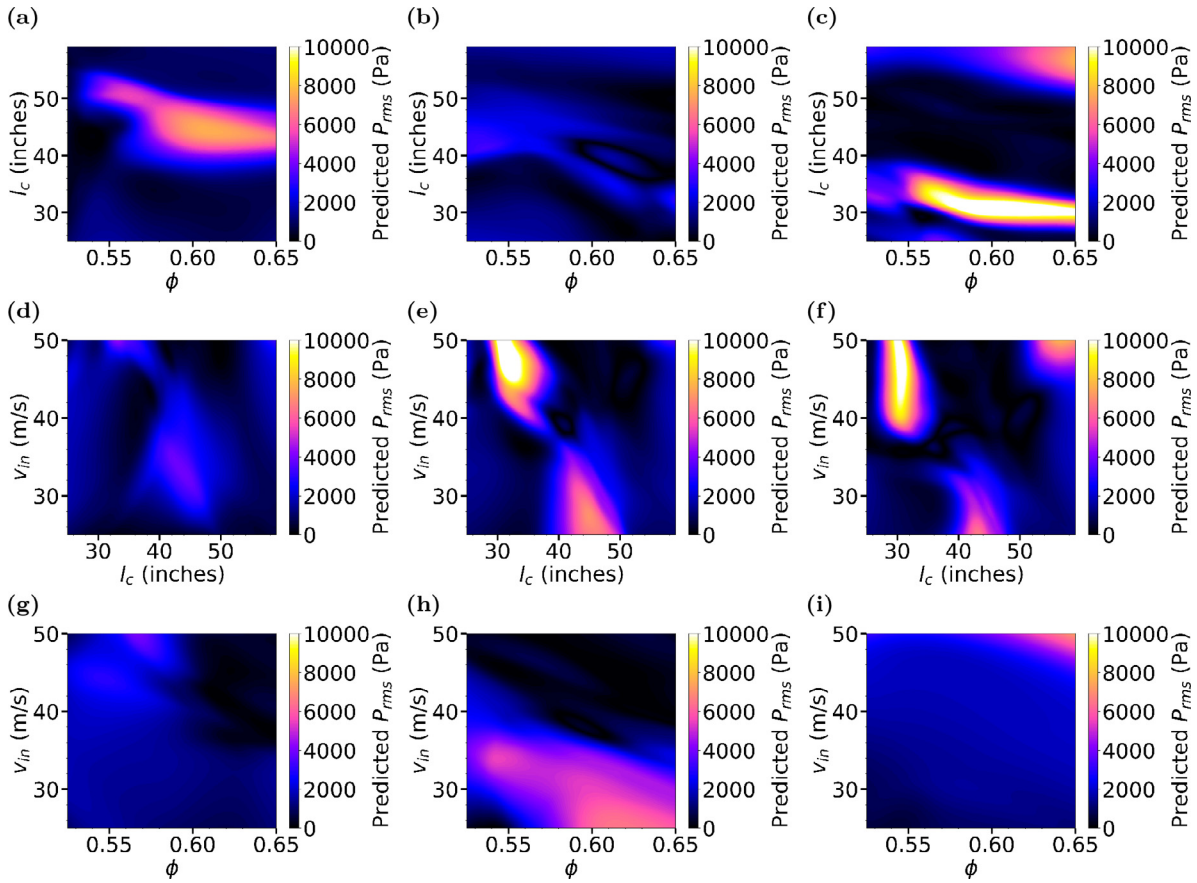


Fig. 9. Projection plots of the $\mathcal{F}_{\text{LPC}}$ predicted $P_{rms}(l_c, \phi, v_{in})$ on different 2-D slices, with training size = 0.5. First row (a–c): $P_{rms}(\phi, l_c)$ when v_{in} = 25, 37.5 and 50 m/s, respectively. Second row (d–f): $P_{rms}(l_c, v_{in})$ when ϕ = 0.525, 0.5875 and 0.650, respectively. Third row (g–i): $P_{rms}(\phi, v_{in})$ when l_c = 25, 42 and 59 inches, respectively.

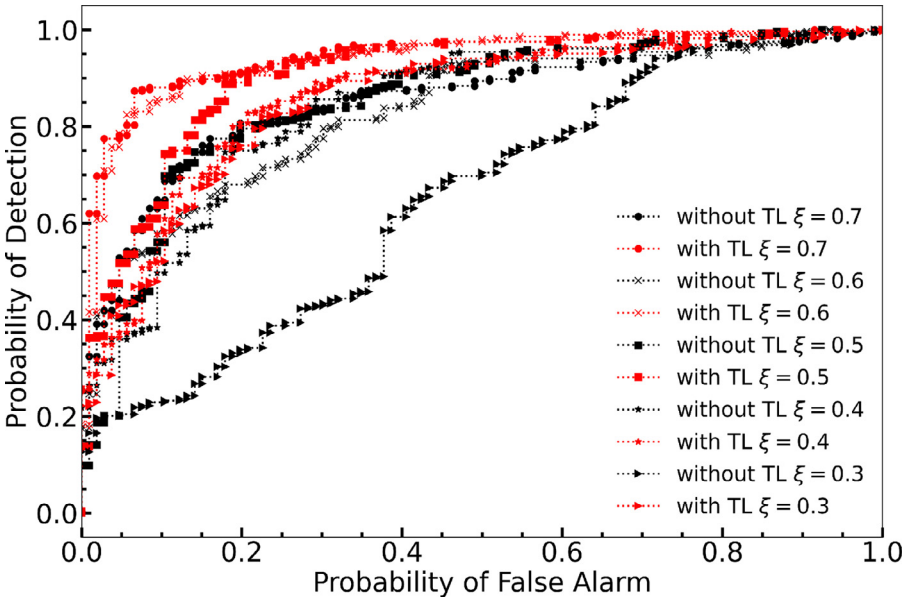


Fig. 10. ROC curves for binary classification problem of predicting instabilities in the LPC based on P_{rms} predictions using different values of the training data fraction (ξ). Red markers indicate ROC curves obtained using P_{rms} predictions from transfer learned regressor for LPC. Black markers indicate corresponding ROC curves using the regressor for the LPC without transfer learning. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

different values of ξ. As can be evinced from Fig. 10, it is also seen from Table 1 that when TL is utilized, even with lower training sizes (e.g., $\xi = 0.4$ and 0.5) it is possible to have classification performance similar to/better than that using high training sizes (e.g., $\xi = 0.6$ or 0.7) in the ANNs without TL. Moreover, the classifier with the lowest training size ($\xi = 0.3$) performs poorly when TL is not utilized (AUC = 0.65), but has a significantly improved performance (AUC = 0.85) when TL is utilized. These observations clearly highlight the importance of TL in providing better predictions which can serve as a critical tool in improving the performance of combustion systems by accurately classifying which design conditions can lead to potential instabilities. With such a tool, an operator can have quick and accurate estimates of the of the safe running conditions of the combustor over a wide range of design conditions, which can serve as a key resource

in defining the operational regimes for efficient performance of the combustor.

6. Summary, conclusions and future work

This paper addresses the issue of suppressing high-amplitude pressure oscillations (e.g., those due to thermoacoustic instabilities (TAI)) in power plants and gas turbine engines. A transfer learning methodology is proposed, by virtue of which regression models predicting stability maps in a relatively simple and easy to operate electrically heated Rijke tube apparatus is efficiently adapted to predict the level of instability for unknown operating conditions in a swirl-stabilized lean-premixed combustion (LPC) system. The average reduction in normalized L_2-error in P_{rms} prediction is found to be $\sim 25\%$, and the average increase in the R^2 score $\sim 22\%$ using transfer learning on the combustion system, across different training sizes of the datasets investigated. A classifier based on the P_{rms} predictions for the LPC is formulated, and it is found that transfer learning significantly improves the accuracy of classification, even with low training data sizes.

The results show that transfer learning can be a potential methodology for transferring knowledge from a simpler to a more complex combustion system. It also highlights the convenience of using such a learning methodology for having reliable estimates of the stability maps over a range of operating conditions, which has a potential of accelerating and improving engine design and performance. A future study can focus on determining the optimum combination of simpler systems and volume of training data needed for satisfactory results on a complex system. For example, training with relatively smaller volume of data on a simple laboratory combustor may give similar accuracy as training with a larger volume of data generated on an electrically heated Rijke tube. A challenging future step would be to combine training data from multiple laboratory-scale systems of increasing complexities like electrically heated Rijke tubes and laboratory-scale turbulent combustors and use them to train a full-scale industrial system for which limited training data would be available.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The work reported in this paper has been supported in part by the U.S. Air Force Office of Scientific Research (AFOSR) under Grant nos. FA9550-15-1-0400 and FA9550-18-1-0135 in the area of dynamic data-driven application systems (DDDAS). The authors are grateful to Professor Domenic Santavicca at Penn State, who kindly provided the experimental data on the combustor apparatus. Any opinions, findings and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the sponsoring agencies.

References

[1] Baumgärtner M, Sattelmayer T. Improvement of the turn-down ratio of gas turbines by autothermal on board syngas generation. J Global Power Propuls Soc 2017;1:55–70. doi:10.22261/JGPPS.D0HPA5.

[2] Lowry W, de Vries J, Krejci M, Petersen E, Serinyel Z, Metcalfe W, et al. Laminar flame speed measurements and modeling of pure alkanes and alkane blends at elevated pressures. J Eng Gas Turbine Power 2011;133(9):091501–091501–9. doi:10.1115/1.4002809.

[3] Kaess R, Poinsot T, Polifke W. Determination of the stability map of a premix burner based on flame transfer functions computed with transient CFD. 4th European combustion meeting, Vienna, Austria. The Combustion Institute; 2009. URL http://www.combustion.org.uk/ECM_2009/P810304.pdf;

[4] Olgac N, Cepeda-Gomez R, Zalluhoglu U, Kammer A. Parametric investigation of thermoacoustic instability (TAI) in a Rijke tube: a time-delay perspective. Int J Spray Combust Dyn 2015;7(1):39–68. doi:10.1260/1756-8277.7.1.39.

[5] Matveev KI. Thermoacoustic Instabilities in the Rijke Tube: Experiments and Modeling. California Institute of Technology; 2003. Ph.D. thesis.

[6] Kopitz J, Huber A, Sattelmayer T, Polifke W. Thermoacoustic stability analysis of an annular combustion chamber with acoustic low order modeling and validation against experiment. In: Turbo expo: power for land, sea, and air, 2; 2005. p. 583–93. doi:10.1115/GT2005-68797.

[7] Surendran A, Heckl M. Passive instability control by a heat exchanger in a combustor with nonuniform temperature. Int J Spray Combust Dyn 2017;9(4):380–93. doi:10.1177/1756827717695282.

[8] Kutz JN. Deep learning in fluid dynamics. J Fluid Mech 2017;814:1–4. doi:10.1017/jfm.2016.803. URL https://www.cambridge.org/core/article/deep-learning-in-fluid-dynamics/F2EDDAB89563DE5157FC4B8342AD9C70.

[9] Brunton SL, Noack BR, Koumoutsakos P. Machine learning for fluid mechanics. Annu Rev Fluid Mech 2020;52:477–508.

[10] Kashinath K, Mustafa M, Albert A, Wu J, Jiang C, Esmaeilzadeh S, et al. Physics-informed machine learning: case studies for weather and climate modelling. Philos Trans R Soc A 2021;379(2194):20200093.

[11] Chattopadhyay A, Hassanzadeh P, Subramanian D. Data-driven predictions of a multiscale lorenz 96 chaotic system using machine-learning methods: reservoir computing, artificial neural network, and long short-term memory network. Nonlinear Process Geophys 2020;27(3):373–89.

[12] Yosinski J, Clune J, Bengio Y, Lipson H. How transferable are features in deep neural networks?. Advances in neural information processing systems (NIPS '14); 2014.

[13] Pan SJ, Yang Q. A survey on transfer learning. IEEE Trans Knowl Data Eng 2009;22(10):1345–59.

[14] Zhuang F, Qi Z, Duan K, Xi D, Zhu Y, Zhu H, et al. A comprehensive survey on transfer learning. In: Proceedings of the IEEE; 2020 arXiv:1911.02685v3]. [cs.LG, 23 Jun 2020

[15] Chattopadhyay A, Subel A, Hassanzadeh P. Data-driven super-parameterization using deep learning: experimentation with multiscale lorenz 96 systems and transfer learning. J Adv Model Earth Syst 2020;12(11). e2020MS002084.

[16] Subel A., Chattopadhyay A., Guan Y., Hassanzadeh P.. Data-driven subgrid-scale modeling of forced burgers turbulence using deep learning with generalization to higher Reynolds numbers via transfer learning. arXiv preprint arXiv:2012.06664

[17] Mondal S, Ghalyan NF, Ray A, Mukhopadhyay A. Early detection of thermoacoustic instabilities using hidden Markov models. Combust Sci Technol 2019;191(8):1309–36. doi:10.1080/00102202.2018.1523900.

[18] Kim K, Lee J, Quay B, Santavicca D. Response of partially premixed flames to acoustic velocity and equivalence ratio perturbations. Combust Flame 2010;157(9):1731–44.

[19] Gelbert G, Moeck J, Paschereit C, King R. Feedback control of unstable thermoacoustic modes in an annular Rijke tube. Control Eng Pract 2012;20(8):770–82.

[20] Gopalakrishnan EA, Sujith RI. Influence of system parameters on the hysteresis characteristics of a horizontal Rijke tube. Int J Spray Combust Dyn 2014;6(3):293–316.

[21] Rigas G, Jamieson NP, Li LKB, Juniper MP. Experimental sensitivity analysis and control of thermoacoustic systems. J Fluid Mech 2016;787. doi:10.1017/jfm.2015.715.

[22] Huhn F, Magri L. Stability, sensitivity and optimisation of chaotic acoustic oscillations. J Fluid Mech 2020;882:A24. doi:10.1017/jfm.2019.828.

[23] Deshmukh NN, Sharma S. Suppression of thermo-acoustic instability using air injection in horizontal Rijke tube. J Energy Inst 2017;90(3):485–95. doi:10.1016/j.joei.2016.03.001. URL https://www.sciencedirect.com/science/article/pii/S1743967116300447.

[24] Zhao D, Morgans AS. Tuned passive control of combustion instabilities using multiple Helmholtz resonators. J Sound Vib 2009;320(4):744–57. doi:10.1016/j.jsv.2008.09.006. URL https://www.sciencedirect.com/science/article/pii/S0022460X08007530.

[25] Zhao D, Ji C, Li X, Li S. Mitigation of premixed flame-sustained thermoacoustic oscillations using an electrical heater. Int J Heat Mass Transf 2015;86:309–18. doi:10.1016/j.ijheatmasstransfer.2015.03.012. URL https://www.sciencedirect.com/science/article/pii/S0017931015002707.

[26] Zhao D, Gutmark E, Reinecke A. Mitigating self-excited flame pulsating and thermoacoustic oscillations using perforated liners. Sci Bull 2019;64(13):941–52. doi:10.1016/j.scib.2019.05.004. URL https://www.sciencedirect.com/science/article/pii/S2095927319302889

[27] Jamieson N, Rigas G, Juniper M. Experimental sensitivity analysis via a secondary heat source in an oscillating thermoacoustic system. Int J Spray Combust Dyn 2017;9(4):230–40. doi:10.1177/1756827717696325.

[28] Tan C., Sun F., Kong T., Zhang W., Yang C., Liu C. A survey on deep transfer learning. arXiv:1808.01974.

[29] Jain AK, Jianchang Mao, Mohiuddin KM. Artificial neural networks: a tutorial. Computer 1996;29(3):31–44. doi:10.1109/2.485891.

[30] Pedamonti D. Comparison of non-linear activation functions for deep neural networks on mnist classification task. arXiv preprint arXiv:1804.02763.

[31] Caruana R, Lawrence S, Giles L. Overfitting in neural nets: backpropagation, conjugate gradient, and early stopping. In: Proceedings of the 13th international conference on neural information processing systems, NIPS'00. Cambridge, MA, USA: MIT Press; 2000. p. 381–7.

[32] Kingma D.P., Ba J. Adam: A method for stochastic optimization. arXiv:1412.6980.

[33] Rumelhart DE, Hinton GE, Williams RJ. Learning representations by back-propagating errors. Nature 1986;323(6088):533–6. doi:10.1038/323533a0.

[34] Abadi M., Agarwal A., Barham P., Brevdo E., Chen Z., Citro C., et al. Tensorflow: large-scale machine learning on heterogeneous distributed systems. 2016; CoRR arXiv:1603.04467

[35] Chattopadhyay P, Mondal S, Bhattacharya C, Mukhopadhyay A, Ray A. Dynamic data-driven design of lean premixed combustors for thermoacoustically stable operations. J Mech Des 2017;139(11):111419. doi:10.1115/1.4037307. https://asmedigitalcollection.asme.org/mechanicaldesign/article-pdf/139/11/111419/6232627/md_139_11_111419.pdf.

[36] Poor HV. An introduction to signal detection and estimation. Springer Science & Business Media; 2013.

[37] Fawcett T. An introduction to ROC analysis. Pattern Recognit Lett 2006;27(8):861–74. ROC Analysis in Pattern Recognition.