24th INTERNATIONAL CONGRESS ON SOUND AND VIBRATION
23–27 July 2017, London

ICSV24
23-27JULY 2017
LONDON CALLING

USING GAUSSIAN PROCESSES TO MODEL COMBUSTION DYNAMICS

Ioannis Matthaiou, Bhupendra Khandelwal *and* Ifigeneia Antoniadou
Department of Mechanical Engineering, University of Sheffield, Sheffield, UK
email: imatthaiou1@sheffield.ac.uk

Modelling the dynamics of combustion is a challenging task due to the non-linear interaction of many processes involved, including chemical kinetics, flame dynamics and acoustic pressure variations inside the chamber. Given that gas turbine engines are the dominant power generation sources, more sophisticated models that can make accurate and reliable predictions regarding the combustion processes and its efficiency, are always in high demand. This paper discusses the development of a data-driven model that is based purely on experimental data, collected from a combustion test rig. The model has been developed using Gaussian Processes, an advanced Bayesian non-parametric machine learning algorithm. The collected data, including pressure inside the combustion primary zone and structural vibration, were all considered as possible candidates for adapting this algorithm to the dynamical characteristics of the combustion chamber under investigation. Accuracy in prediction using this empirical model was investigated for different combinations of experimental data and fractions of them, using the root mean squared error as performance measure. The covariance function parameters of the Gaussian Process model were optimised using a gradient-based algorithm for the best possible adaptation to the experimental dataset.

Keywords: Gaussian Processes, modelling of combustion dynamics, supervised learning, combustion stability and control.

1. Introduction

Over the past three decades the gas turbine engine has dominated the power generation industry, while it has been the only source for powering aircraft. Therefore, improving the overall efficiency and safety of gas turbine engines, while complying with environmental regulations, has always been one of the key areas that engineering manufacturers have invested heavily on, for further research and development. One way to move towards higher performing engines, is to come up with innovative solutions that can advance the way combustion systems work for a large range of operating conditions. To test hypotheses, engineers have been using high fidelity computational fluid dynamics models to predict, for instance, reaction kinetics during combustion and formations of several pollutants for different types of fuels. At the same time, experimental work has also been used together with these modelling and simulation activities to investigate the problem further and achieve progress in optimising combustion. An alternative way to those approaches is to use machine learning, in parallel with experiments. These models, being completely data-driven, make no assumptions regarding the system being analysed as they directly model the data. In case of combustion, this is extremely valuable because the processes involved are characterised as being highly stochastic, having great complexity and non-linear interaction between them. In addition, these data-driven models offer the flexibility of conventional computational approaches, in terms of conducting analysis in regions that

we may wish to investigate. In that way, we can make predictions in regions where there are no experimental data, which may be valuable for optimisation, analysis and/or control.

Recently, in [1] the authors modelled experimental measurements of exhaust emissions and fuel consumption for calibrating an engine using Gaussian Processes (GP). GP is a non-parametric Bayesian machine learning method that is particularly suited to regression problems. Also, in [2] a GP model was developed from digital flame images in a combustion chamber, in order to predict its oxygen flow content. The prediction was used to control the air flow, to maintain energy requirements and complete combustion. Moreover, in [3] another data-driven model was developed that can predict fuel consumption and exhaust emissions from certain percentages of sustainable fuel. The model was then used in a multi-objective optimisation routine to compute optimal fuel percentage. Apart from performance requirements, a combustion device must operate safely at all times. The problem of thermoacoustic instability has been known for a long time, and studies such as in [4], have demonstrated that unstable combustion favours lean (high air-to-fuel flow rates) operating conditions. Due to the complexity of the problem, a data-driven model developed from stable combustion can be used to derive an accurate and reliable reference signal for control, which can be used to eliminate the possibility of thermoacoustic instability.

The potential of using such advanced data-driven models for the purpose of analysing, optimising and controlling systems has just started to be realised. In general, a good data-driven model must capture the system characteristics very accurately without over-fitting [5], while its construction should be simple enough, so that it can be easily amenable to analysis and interpretation. In addition, it must be able to quantify for uncertainty (this is especially valuable in stochastic systems like combustion). In this paper, the method of GP is used to develop a data-driven model from experimental measurements acquired from a combustion test rig. The model can be used for making predictions at any location within the multi-dimensional input space boundaries.

2. Gaussian Processes for modelling combustion dynamics

The aim of this section is to provide a basic background in GP for regression, following closely the explanation and notation as found in the canonical text on the subject in [6]. Regression is one type of supervised learning (the other is classification), in which a dataset of n experimental observations is provided $\mathcal{D}_n = \{(\mathbf{x}_i, y_i | i = 1, ..., n)\}$, where $\mathbf{x}$ is a d-dimensional input vector and y is a scalar target value, and the task is to determine the relationship between them, represented by the function $f(\mathbf{x})$. By finding the functional form $f(\mathbf{x})$ that relates these pairs of data, it is possible to make predictions at any unseen data point $\mathbf{x}_*$, coming from a separate dataset $\mathcal{D}_m = \{(\mathbf{x}_{*i}, y_{*i} | i = 1, ..., m)\}$. In the context of combustion, we may wish to determine, for instance, the functional form of the pressure variation inside the combustion chamber (i.e. the target) from the air-to-fuel flow ratio that flows through (i.e. the input), such that we can predict complex behaviour, e.g. combustion instabilities. This implies that such characteristics must be included in $\mathcal{D}_n$, otherwise, the problem will be cast as part of an unsupervised learning framework for detecting novel scenarios.

In Bayesian modelling, apart from obtaining $f(\mathbf{x})$ from a set of noisy targets $y = f(\mathbf{x}) + \varepsilon$, where $\varepsilon \sim \mathcal{N}(0, \sigma_n^2)$, i.e. an identically independent normally distributed noise with mean zero and variance σ_n^2, it is also possible to calculate the probability $p(f_*|\mathbf{x}_*)$. The latter, corresponds to the quantification of uncertainty for predicting a new target value f_*, given $\mathbf{x}_*$. This may guide us in performing future experiments around regions that we have high uncertainty regarding the model predictions. In parametric modelling, the functional form of this pair-wise relationship is given as a linear combination of its coefficient vector $\mathbf{w}$, so that $f(\mathbf{x}) = \mathbf{x}^T\mathbf{w}$. By re-arranging the sum and product rules of probability the maximum a posteriori estimate of the coefficients, which is the mean value of the following conditional distribution, can be obtained as,

$$p(\mathbf{w}|\mathbf{y}, X) = \frac{p(\mathbf{y}|X, \mathbf{w})p(\mathbf{w})}{\int p(\mathbf{y}|X, \mathbf{w})p(\mathbf{w})d\mathbf{w}} \tag{1}$$

where, X is the matrix including n input vectors and $\mathbf{y}$ its corresponding vector of target values. The above equation combines all information about the coefficients; the prior belief $\mathbf{w} \sim \mathcal{N}(0, \Sigma_p)$, where Σ_p is the covariance matrix of the multivariate normal distribution, and the likelihood of every observation $\mathbf{y}$ given the coefficients $p(\mathbf{y}|X, \mathbf{w})$. The likelihood function is a multivariate Gaussian with its mean and covariance directly dependent on the coefficients, thus, maximising it the values that are most suitable for this dataset can be found. The denominator in Eq. 1 is a normalising constant that is independent on the coefficients (since we are marginalising over $\mathbf{w}$), so it does not affect this inference procedure.

Alternatively, if we perform inference directly from an infinite space of functions, we can incorporate maximum flexibility in the modelling phase. As there are no coefficients to calculate, this approach is defined as Bayesian non-parametric modelling, and any function that models the data is a valid candidate. The observed values indexed at any input $\mathbf{x}_1, ..., \mathbf{x}_n$, are realisations of random variables $f(\mathbf{x}_1), ..., f(\mathbf{x}_n)$. For reasons of computational convenience, we may specify a joint Gaussian distribution between these variables, which gives rise to a GP. So for any two inputs $\mathbf{x}$ and $\mathbf{x}'$, a GP can be specified by its mean $m(\mathbf{x})$ and covariance function $k(\mathbf{x}, \mathbf{x}')$,

$$f(\mathbf{x}) \sim \mathcal{GP}(m(\mathbf{x}), k(\mathbf{x}, \mathbf{x}')) \tag{2}$$

In principle, if we were to include all points in the real line, the infinite joint Gaussian between them would correspond to a distribution over functions. Now, assuming that $m(\mathbf{x}) = 0$, we can evaluate the covariance function at discrete input locations generating values that are solely determined by the characteristics implied by the covariance. For this paper, the squared exponential is used, to specify the pair-wise covariances of random variables. Its form for a d-dimensional input, which might include several variables from the combustion experiments, e.g. pressure and vibration, is given as,

$$cov(y(\mathbf{x}), y(\mathbf{x}')) = k(\mathbf{x}, \mathbf{x}') + \sigma_n^2 \delta_{x,x'} = \sigma_f^2 \prod_i^d \exp\Big(-0.5\gamma_i (x_i - x_i')^2 \Big) + \sigma_n^2 \delta_{x,x'}, \quad \forall \gamma_i > 0 \tag{3}$$

where, σ_f controls the vertical scale of the output and $\delta_{x,x'}$ is a Kronecker delta function. The longitudinal or input scale is controlled by $\gamma (= \gamma_1, ..., \gamma_d)$, which measures the rate of decrease in target values covariances as the distance between two locations in the same input space increases. Therefore, the squared exponential covariance function models the concept that closely spaced inputs give highly correlated target values, which is a very reasonable assumption to make. It is, therefore, clear that the $(d + 2)$-dimensional parameters vector of the squared exponential kernel $\mathbf{q} = [\gamma, \sigma_f, \sigma_n]$ defines the distribution of functions that we can use to explain our data. Therefore, given the dataset $\mathcal{D}_n$, the following conditional probability, which is called the marginal likelihood function, can be used to find the optimal $\mathbf{q}$,

$$p(\mathbf{y}|X, \mathbf{z}) = -\Big((2\pi)^{(n/2)} |K(X, X) + \sigma_n^2 I|\Big)^{-1} \exp\Big(-0.5\mathbf{y}^T (K(X, X) + \sigma_n^2 I)^{-1} \mathbf{y}\Big) \tag{4}$$

where for convenience, all covariance computations between the data points were assembled on an $n \times n$ matrix of covariances $cov(\mathbf{y}) = K(X, X) + \sigma_n^2 I$. Taking the log of Eq. 4 (for practical reasons, as explained in [5]), and then deriving the expression of partial derivatives analytically with respect to $\mathbf{q}$, a standard gradient-based optimisation routine can be used to calculate the parameters that maximise the marginal log likelihood. In this paper, the scaled conjugate gradient method, as implemented by the Matlab toolbox in [7], was used to obtain the optimal $\mathbf{q}$.

Given the optimised covariance and a matrix of unseen data X_* from the dataset $\mathcal{D}_m$, we can generate function values $\mathbf{f}_*$ at any location $\mathbf{x}_*$, by sampling from the zero-mean prior joint distribution,

$$\begin{bmatrix} \mathbf{y} \\ \mathbf{f}_* \end{bmatrix} \sim \mathcal{N}\left(\mathbf{0}, \begin{bmatrix} K(X, X) + \sigma_n^2 I & K(X, X_*) \\ K(X_*, X) & K(X_*, X_*) \end{bmatrix}\right) \tag{5}$$

where we have assumed that the mean vector is $\mathbf{0}$. By conditioning this prior on the observations in our matrix $\mathcal{D}_n$, we can obtain the posterior mean $\bar{\mathbf{f}}_*$ and covariance $cov(\mathbf{f}_*)$ using standard formulae,

$$\begin{aligned} \bar{\mathbf{f}}_* &= K(X_*, X)[K(X, X) + \sigma_n^2 I]^{-1}\mathbf{y} \\ cov(\mathbf{f}_*) &= K(X_*, X_*) - K(X_*, X)[K(X, X) + \sigma_n^2 I]^{-1}K(X, X_*) \end{aligned} \tag{6}$$

The above corresponds to restricting our analysis to those functions that agree with the observations. As shown, the posterior mean is a linear combination of the observed target values, which implies that GP can be thought of as linear smoothing technique. Moreover, the model presented here is stationary as decided by the covariance function selected, because of its dependence on the geometric distances between data points.

3. Combustion experiments

In order to collect data for modelling the dynamics of combustion using the method outlined in Section 2, a number of experiments were performed on a combustion test rig. This includes an annular combustion chamber, which was originally part of a turbofan engine manufactured by Rolls Royce for aircraft propulsion. In the tests that will be described, the air was supplied from an industrial fan at a rate of $0.3kg/s$, while the mass flow rate of fuel was kept fixed at about $1.35g/s$, i.e. a lean mixture.

In these experiments, the fuel is injected in a fine atomized form through a spray nozzle into the combustion chamber to be mixed and ignited with the air supplied by the industrial fan, producing constant heat release. After igniting the air-to-fuel mixture to initiate combustion, the flame is then self-sustained. In examining the process of combustion more closely, there are two main regions or zones of interest: the primary and dilution zones, related to the utilisation and distribution of fresh air coming from the compressor. In a typical gas turbine engine, since air mass flow rates are normally too high for combustion, the flow is first decelerated, while its static pressure is raised. In more detail, around 20% of the air passes through swirl vanes that cause its recirculation, while a further 20% of air passes through the flame tube passages into the primary zone creating a toroidal vortex as it mixes with the air coming from the swirl vanes. This effect results in stabilisation and anchoring of the flame, while better air-fuel mixing takes place, due to the fuel being injected to the centre of vortex and the general turbulence that exists inside the primary zone. Normally, temperature of combustion gasses can reach $2000°C$, which cannot be tolerated by the material properties of the turbine nozzle guide vanes. Inside the dilution zone, the remaining air flows through passages to cool the gasses and reduce the flame tube temperature. Complete combustion of the air-fuel mixture is expected to be made prior to entering the dilution zone. The combustion test rig has been manufactured accordingly,

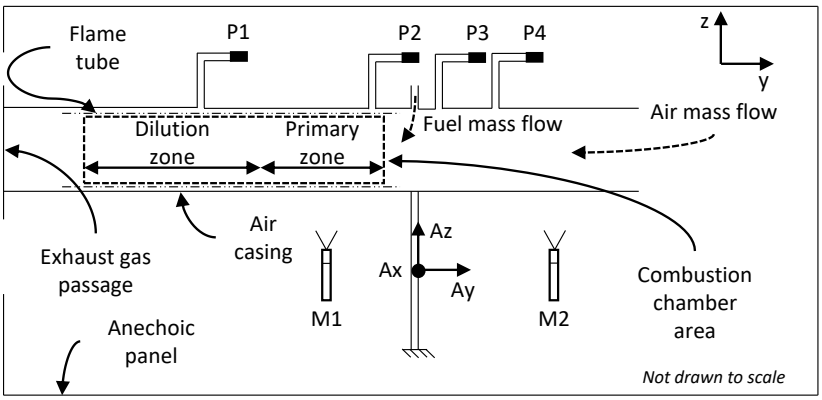

Figure 1: Schematic diagram of the combustion test rig. The following sensors are shown: pressure P1-P4, microphones M1 and M2 and accelerometers in $x - y - z$ directions Ax, Ay and Az.

in order to allow for several measurements to be made by incorporating sensors at various locations, some of which are shown by the diagram in Fig. 1. Two free-field microphones (PCB 378B02) with sensitivity of $50mV/Pa$, were used to measure the sound radiated during the combustion process, inside a room with minimal sound reflection from the walls. Four pressure sensors (PCB 106B52) with sensitivity of $725mV/kPa$ measured sound pressure fluctuations at different locations inside the combustion chamber and at the air stream. These sensors were mounted on the test rig via tubes. Passages were manufactured at the air casing, just above the primary and dilution zones, to mount two sensors to measure the variation in combustion pressure. Structural vibrations of the process were measured at $x-y-z$ directions by attaching three accelerometers (Dytran 3225F1) with $10mV/g$ of sensitivity, on a support column fixed to the test rig. These 9 sensors were sending signals on National Instruments 9234 modules directly, which were then recorded through the LabVIEW software interface. The 9234 modules were configured to acquire data at their maximum sampling rates, i.e. $f_s = 51.2kHz$, in addition to providing the required $20mA$ of current excitation for each sensor. The 9234 modules have built-in anti-aliasing filters with a passband and stopband frequencies at $0.45f_s$ and $0.55f_s$, respectively. The data acquisition time was set to 10 seconds, for all the experiments. Thermocouple sensors mounted on the air casing near the primary zone (not shown in the diagram), measured maximum temperatures of about $500°C$, which is much lower than when in service.

Prior to conducting any analysis, each raw time-series measurement, say $\mathbf{x}_{raw}$, acquired from the 9 sensors was normalized within the range [-1, +1], using Eq. 7. This step is necessary when working with machine learning algorithms as data can have different value ranges, e.g. pressure and vibration, as in this study. Scaling the data within a given range, ensures that they will all receive similar importance in further computations, e.g. in optimisation.

$$\mathbf{x} = \frac{\mathbf{x}_{raw} - \min(\mathbf{x}_{raw})}{\max(\mathbf{x}_{raw}) - \min(\mathbf{x}_{raw})} \tag{7}$$

where, $\mathbf{x}$ is the scaled time-series measurement, $\max(\mathbf{x}_{raw})$ and $\min(\mathbf{x}_{raw})$ represent the computation of upper and lower scalar values, respectively, of $\mathbf{x}_{raw}$.

As an example, the plots in Fig. 2, show the scaled (according to Eq. 7) amplitudes of four sensor measurements for the whole test duration: structural acceleration in y-direction α_y, pressure at dilution and primary combustion zones, p_1 and p_2, respectively, and the sound captured by the microphone located next to the combustion chamber ψ_1. Visually, these time-series show that the processes are non-stationary, which can be particularly challenging to model with the assumed form of GP considered here. In the next section, all nine sensor measurements are considered as candidates for the modelling task. In addition, the time is also considered as plausible input to the model.

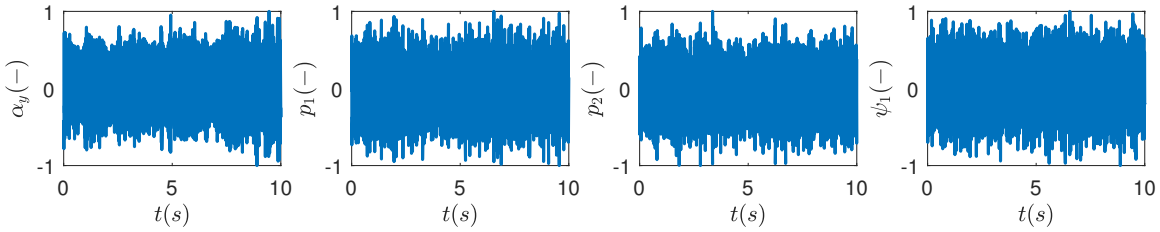

Figure 2: Time-domain plots of scaled experimental data, recorded from 2 pressure sensors (p_1 and p_2), 1 accelerometer (α_y) and 1 microphone (ψ_1).

4. Results and discussion

This section presents some of the main results from the modelling procedure. For reasons of computational efficiency, only a subsection of 5120 data points from the whole 10 seconds of each

time-series was considered for modelling. As it was shown in Eq. 4, the marginal likelihood requires the inverse of $K(X, X)$, which is of size $n \times n$. The computational complexity of standard methods for matrix inversion scales according to $\mathcal{O}(n^3)$, which can render the problem infeasible to solve if the number of data points n is too large [6]. Nevertheless, there is sufficient variation within this subsection of data to demonstrate the effectiveness of GP for this specific task.

The variation in pressure inside the primary combustion zone p_2 is an important parameter that can provide reliable and direct information regarding the stability of combustion. In general, there are limitations in air-to-fuel ratios: the lean and rich limits beyond which the flame can become extinct. The phenomenon when combustion becomes unstable because of excessive air entering the chamber, as it may be the issue in these tests, is called "lean blow-out" [8]. Hence, it may be of interest, e.g. for a control system design point-of-view, to predict p_2 from other measurements. Now, the task is to determine which input parameter, or a combination of them, can be used to predict p_2.

A useful performance metric that can assess how close the mean predictions (model outputs) $\bar{\mathbf{f}}_*$ are, to the actual observed data included in $\mathcal{D}_m$, is the root mean square error ($RMSE$) value,

$$RMSE = \sqrt{\frac{1}{m}\sum_{k=1}^{m}(y_k - \bar{f}_{*k})^2} \tag{8}$$

From the 5120 data points considered for modelling, a certain percentage was used to infer the parameters vector $\mathbf{q}$. This is called the "training" data, which are all the $\mathbf{x}$ and $\mathbf{y}$ pairs included in $\mathcal{D}_n$. All the data pairs $\mathbf{x}_*$ and $\mathbf{y}_*$ included in $\mathcal{D}_m$, i.e. the remaining portion of data that is called "testing" data, were used to validate the mean prediction $\bar{\mathbf{f}}_*$ accuracy.

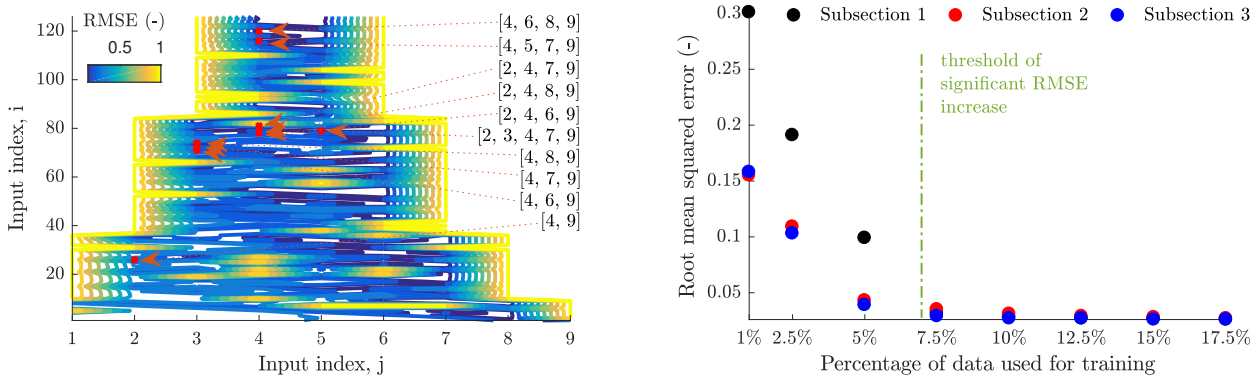

Figure 3: On the left, is a contour plot of $RMSE$ (normalised) values calculated for 511 GP model input combinations. On the right, $RMSE$ values were calculated for different percentages of training data, on three arbitrarily chosen subsections.

In order to make accurate predictions, there will need to be a direct relationship between the model inputs chosen (i.e. the independent variables) and the output (i.e. the dependent variable). With this in mind, the $RMSE$ for nine candidate input variables was computed. The smallest $RMSE$ value will give us the best combination of model inputs and output. The nine candidate input variables, together with their respective indices considered in the input matrix X are: $x-y-z$ accelerations (1, 2, 3), pressures at dilution zone and air streams (4, 5, 6), sound recorded next to combustion chamber and air streams (7, 8) and time (9). By arranging these input variables according to their indices in X, we computed the predictions of 511 different GP models, i.e. $9!/((9-k)!k!)$, for $k = \{1, ..., 9\}$. The resulting $RMSE$ (normalised by the maximum value) contour plot for all combinations of model inputs is shown on the left of Fig. 3. The ten combinations that result in the smallest $RMSE$ value are also shown. Note that $n = 0.1 \times 5120$ (so that $m = 5120 - n$) data points for this case. Clearly, the time input (index 9), together with the pressure at dilution zone (index 4) play an important role

in the predictions, since they are always included in each of the 10 "optimum" combinations shown. However, a closer look can reveal that adding the dilution zone pressure as an input, together with the time, only increases prediction accuracy by a marginal amount $\approx 5\%$. A similar observation is made for the other nine combinations identified here. Given that we require the construction of a model that is easily interpretable, the additional 5%, or so, in prediction accuracy does not justify the extra complexity induced in the model. Hence, only time is considered as input variable, i.e. single input (time) and output (pressure at primary combustion zone).

To examine the model accuracy further, the $RMSE$ was calculated for different percentages of training samples taken from the 0.1 seconds of data. Three arbitrarily chosen subsections of equal length n from the whole time-series of p_2, were considered for modelling. On the right of Fig. 3, the $RMSE$ is shown to decrease exponentially up to about 5% of training data percentage, for the three subsections of the time-series considered. Therefore, it can be said that any training data percentage above 5%, i.e. $n > 256$ data points (randomly chosen from the time-series), can result in a good model. This is because $\mathbf{q}$ can be better adapted to the given dataset, as more information becomes available.

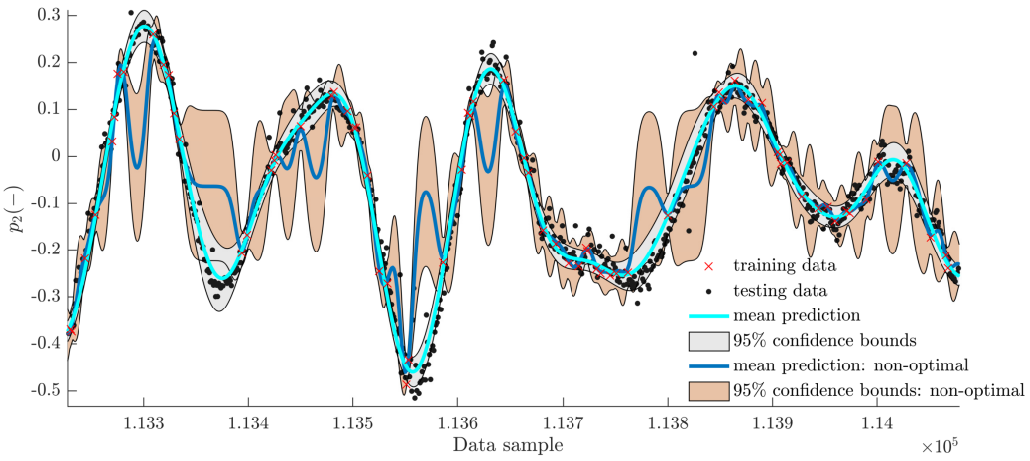

Figure 4: Mean prediction and variance computed on a subsection of the dataset, for the optimised and non-optimised cases, using 10% of the total data for training.

The resulting prediction mean and its variance calculated for the optimised $\mathbf{q}$ was also compared to the non-optimised case. It was found that a decrease in $RMSE$ value from 0.089 to 0.028 can result when $\mathbf{q}$ is optimised, i.e. adapted to the dataset instead of using fixed parameter values. For illustration purposes, the prediction on a small subsection of around 900 data points (17.6 milliseconds), can be shown in Fig. 4. The 95% confidence bounds were computed as $\bar{\mathbf{f}}_* \pm 2\sqrt{cov(\mathbf{f}_*)}$, while also the optimised and non-optimised mean predictions are shown in the plot. More specifically, we see that in regions where there are no training data available, the mean prediction for the non-optimised case deviates significantly from the testing data. Moreover, as indicated also by its confidence regions, there is much higher uncertainty in predictions as compared to the optimised case. As shown here, when $\mathbf{q}$ is adapted to the dataset it can capture very well the underlying dynamics of p_2, using only 10% as the training data.

At the same time, because the marginal likelihood is maximised via optimising the covariance function parameters vector $\mathbf{q}$, the confidence bounds are much more closely spaced. Note also, because of the noise term σ_n, the two confidence bounds will always have non-zero spacing between them even at the locations of training data. Generally, when moving away from the training data, the uncertainty in model prediction rises significantly. Thus, the developed model is limited to interpolating within the training data regions. The mean prediction value when extrapolating far away from the training data, tends to its zero value.

5. Conclusions and future work

The preliminary investigation presented in this paper illustrated the use of Gaussian Processes, a non-parametric Bayesian machine learning tool, for modelling the dynamics of combustion systems. Several sensor measurements were collected from an experimental test rig, housing a gas turbine engine combustion chamber, including pressures inside the hot and cold sections of the device and structural vibrations. A procedure that was followed to develop a data-driven model from those measurements was described in some detail. Firstly, all possible combinations of inputs were considered to predict pressure variations inside the primary combustion zone, a critical parameter when examining the air-to-fuel ratio limits of combustion. Using a performance metric, it was found that there was a more direct relationship between time and pressure than any other input variable considered here. Hence, a single input and output data-driven model was developed, where it can be used to predict pressure at any time interval specified, given that it lies within the bounds of training data, i.e. interpolating. The prediction using Gaussian Processes provides us with an uncertainty measure, which is very important that can lead us to perform more targeted experiments in the future, in regions where the uncertainty is high. It was also demonstrated that when the covariance function (squared exponential) parameters are optimised, the predictions can be very accurate, while the uncertainty reduces significantly.

To investigate the accuracy in predicting the pressure variation, when the combustion is unstable, future experiments will need to be performed where instability can be initiated manually by modifying the amount of air-to-fuel ratio. The developed data-driven model will be used with a control system, which is designed to mitigate the possibility of flame extinction. In addition, further experimental measurements collected during these tests, including temperature variations at several locations around the combustion chamber, will be used as possible candidate input variables to the existing model presented here. At the same time, the possibility of a multi-output Gaussian Process for the control scheme will also be explored.

REFERENCES

1. Berger, B., Rauscher, F., Lohmann, B. Analysing Gaussian Processes for Stationary Black-Box Combustion Engine Modelling, *Proceedings of the 18th World Congress, The International Federation of Automatic Control*, Milano, Italy, 28 August–2 September, (2011).

2. Chen J., Chan, L.T., Cheng, Y. Gaussian process regression based optimal design of combustion systems using flame images, *Applied Energy*, **111**, 153–160, (2013).

3. Wong, K.I., Wong, P.K., Cheung, C.S., Vong C.M. Modeling and optimization of biodiesel engine performance using advanced machine learning methods, *Energy*, **55**, 519–528, (2013).

4. Nair, V., Sujith, R.I. Multifractality in combustion noise: predicting an impending combustion instability, *Journal of Fluid Mechanics*, **747**, 635–655, (2014).

5. Bishop, C. M., *Pattern Recognition and Machine Learning*, Springer, New York, NY (2006).

6. Rasmussen, C. E., Williams, C. K. I., *Gaussian Processes for Machine Learning*, MIT Press, Cambridge, MA (2006).

7. Nabney, I.T., *NETLAB: Algorithms for Pattern Recognition*, Springer, Berlin, Germany (2004).

8. Lefebvre, H. A., Ballal, R. D., *Gas Turbine Combustion: Alternative Fuels and Emissions, 3rd Edition*, CRC Press, Boca Raton, FL (2010).