# Aircraft Parametric Structural Load Monitoring Using Gaussian Process Regression

Ramon Fuentes, Elizabeth Cross, Andrew Halfpenny, Keith Worden, Robert J. Barthorpe

7th European Workshop on Structural Health Monitoring
July 8-11, 2014. La Cité, Nantes, France

# AIRCRAFT PARAMETRIC STRUCTURAL LOAD MONITORING USING GAUSSIAN PROCESS REGRESSION

**Ramon Fuentes[1, 2], Elizabeth Cross[1], Andrew Halfpenny[2] , Keith Worden[1], Robert J. Barthorpe[1]**

[1] *Dynamics Research Group, Department of Mechanical Engineering, University of Sheffield, Mappin Street, Sheffield, S1 3JD*

[2] *Technology Group, HBM-nCode, United Kingdom, Advanced Manufacturing Park, Brunel Way, Catcliffe, Rotherham, S60 5WG*

ramon.fuentes@sheffield.ac.uk

ABSTRACT

The work presented here demonstrates the capability of Gaussian Process (GP) regression for the prediction of aircraft structural loads based on recorded flight parameters. The objective of monitoring aircraft loads during operation is to develop a better understanding of aircraft usage and thus to provide the operator with an accurate estimate of the remaining useful life of components against their prescribed fatigue life. These loads are often difficult and expensive to measure and this motivates the use of advanced mathematical techniques to estimate them accurately based on other parameters that are typically measured during flight. Gaussian Process regression is a powerful Bayesian machine learning tool whereby predictions and their distributions can be obtained without having to specify a particular model/functional form. Data collected from a military trainer aircraft is used to demonstrate how the mapping of measured strains to basic flight parameters such as airspeed, accelerations, and control surface deflections can be performed using GP regression. It is also shown how these results can be applied in an aircraft usage monitoring context. The GP predictions for strain are compared to the actual measurements for 101 flights, and the results are presented in terms of fatigue life, correlation, and mean-squared error. The results are encouraging, with errors in fatigue life in the range of 5% to 30% in the worst cases.

KEYWORDS : *Structural Health and Usage Monitoring, Gaussian Process Regression, Fatigue,*

## 1 INTRODUCTION

The objective of this work is to illustrate the use of Gaussian Process (GP) regression as a tool for performing parametric load monitoring. The use of the word “parametric” here refers to the basic idea of using data from flight parameters that are typically measured on an aircraft in operation in order to infer structural loads. The motivation for monitoring these loads is twofold. Knowledge of structural loads can be used to accurately track fatigue damage accumulation; this is useful for the operator to be able to track the remaining useful life of structural components. This analysis is currently typically done using simple assumptions about the usage spectrum since loads cannot be directly measured during the life of the aircraft. The problem with this approach is that the actual usage of the aircraft might significantly deviate from the usage spectrum used to guide the design and dictate maintenance. If, for example, the number of service hours was used to predict damage accumulation there could be a significant error in the remaining useful life. If the error is too conservative, then this is potentially something to feed back to the aircraft/component designer, as with the design of any kind of vehicle, a better understanding of its usage will lead to improved and (possibly) optimum design. If the error in life is not conservative this could lead to catastrophic consequences. Figure 1 schematically demonstrates the result of over and under-estimating the life of an aircraft in terms of life consumption, time, risk, and cost. The use of the aircraft fatigue meter goes some way into monitoring actual loads in aircraft during operation but is limited by the simplistic approach of counting exceedances on g-levels and correlating these to fatigue damage accumulation. The results are not always optimum since the structural loads are much more complex than can be correlated directly to accelerations at the centre of gravity (CG). The requirements for a system that replaces the fatigue meter in military aircraft are outlined in [1]. These include the ability to carry out more sophisticated and accurate assessments of fatigue damage based on measured flight parameters using an off-board system that is capable of also performing exploratory data analysis. There are potentially huge cost savings in implementing such a system given that in the UK "the cost of each 1% of fleet fatigue life can be in excess of £100 million" [1].

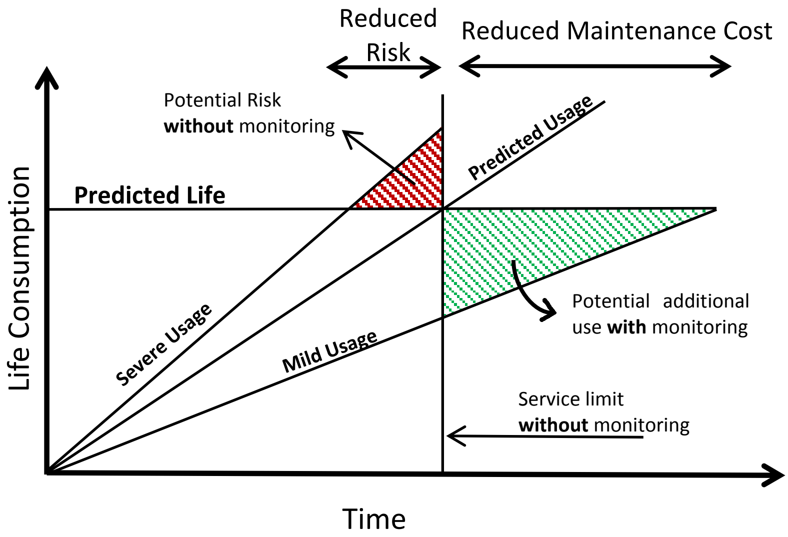


Figure 1 – Advantages of a Structural Health Monitoring System over the life cycle of an aircraft in terms of risk and maintenance cost [1]

A feasibility study is presented here into the application of Gaussian Process (GP) regression as a tool to perform parametric load monitoring. To investigate its potential for this type of application, a Gaussian Process regression model has been built in order to map strains at the wing from flight parameters. This has been done on a Short Tucano, an RAF military trainer aircraft, for which 101 flights were available from two different aircraft with the required measured parameters. The dataset consisted of measured time histories for flight parameters such as airspeed, normal and lateral accelerations at the centre of gravity, altitude, fuel, control surface deflections, undercarriage position, Boolean values for weight-on-wheels and airbrake position. Strain measurements were available at various locations on the aircraft. For the purpose of demonstrating GP regression, only two strain channels were considered here; from a bending bridge at the port inner wing and port outer wing. The strains were recorded in micro strain units ($\mu\varepsilon$).

The current work builds on the ideas presented in [2, 3, 4, 5] which developed a similar parametric-based structural monitoring system using Artificial Neural Networks (ANNs) to map the input flight parameters to the output strain measurements. Here, Gaussian Processes Regression (GPR) is studied for this task. GPR has previously been introduced for inferential monitoring of loads in undercarriage systems in [6]. An overview of regression using Gaussian Processes is presented in the next section.


Figure 2: Short Tucano used during OLM. The locations used for monitoring strains on the port inner and outer wing are shown. [5]

## 2 STRAIN PREDICTION USING GAUSSIAN PROCESS REGRESSION

### 2.1 Background on Gaussian Processes

Gaussian Process Regression (GPR) is closely related to other methods including linear regression, spline fitting, and neural networks. The objective of this section is to give the reader a conceptual understanding of GPR, what is involved in the modelling decisions, and how it differs fundamentally from other regression modelling techniques. For precise mathematical definitions and further reading and information the reader is referred to [7]. Machine Learning is a branch of artificial intelligence in which mathematical models are built from observed data. The modelling procedure is thus separated into training and prediction phases. During the training phase example data is provided to the model for it to learn the relationships in the data. The prediction phase uses this learned model to make predictions on new unseen data.

Machine Learning algorithms largely separate into two groups; supervised and unsupervised learning. GPR is a supervised learning model. The distinction between supervised and unsupervised learning is that in supervised methods, examples of both inputs and outputs are required in order to train the model parameters. These parameters will depend on the model chosen. If, for example, a linear regression model is being used to fit a function $f(x) = mx + b$, from a training input set $\mathbf{x}$

and output **y,** the parameters needed to be found would be the slope and intercept $m$ and $b$. A linear regression model is of course restrictive since the relationship between $\mathbf{x}$ and $\mathbf{y}$ might not necessarily be linear. In this case, a different model needs to be chosen, but the question is then of deciding which functional form should $f(x)$ take; one could choose to model it as a quadratic, cubic, periodic, etc. The choices are literally infinite. The task of choosing a particular functional form for the model is usually called 'model selection', and is an important part of the training procedure in any supervised learning algorithm.

One of the key advantages of GPR is that no specific functional form needs to be assumed. In contrast, GPR lets the training data describe the input-output relationships more directly without assuming that the data is generated by any one particular $f(x)$. Gaussian Processes define a distribution over functions, and they do so by extending the concept of a multivariate Gaussian distribution to infinite dimensions. What this means in practice is that observed data sets can be thought of as being "generated" by a Gaussian Process which is specified by a mean and a covariance function. This is a key practical point for the modeller, as there needs to be a decision about the form of both the mean and covariance functions. There exist several choices for these functions. The choice of covariance will describe in broad terms what the data "look like". A popular and flexible choice is the squared exponential:

$$k(x,x') = \sigma_f^2 \exp\left[\frac{-(x-x')^2}{2l^2}\right] \tag{1}$$

where $x$ is a sample input point and $x'$ is some other sample input point within the dataset, $\sigma_f^2$ is the maximum allowed covariance within the data, and $l$ is a parameter of the covariance function. For this covariance function, $l$ controls the smoothness. These parameters are termed hyper-parameters since they are not used directly when making predictions from new inputs, but rather define the distribution that generates the output data. In order to make a prediction for an output, whatever the functional form of (1) is, $k(x,x')$ needs to be defined for every single pair of input points, and collected into a covariance matrix $K$. Thus for a dataset containing $n$ data-points $K$ will be an $n \times n$ matrix. This matrix needs to be inverted to predict the mean and variance of the outputs. The mean of the outputs is essentially the prediction conditioned on the training data and the variance is the uncertainty or confidence interval around it.

### 2.2 Mapping flight parameters to strains

This section describes the procedure carried out for modelling the mapping between flight parameters and strains at the wing. The flight parameters used as inputs to the regression model were: indicated airspeed, CG normal acceleration, CG lateral acceleration, weight-on-wheels, flaps position, undercarriage position and fuel. Before passing this to the GPR, some domain knowledge was used to generate inputs that more closely represent the physics of the underlying process, following some of the ideas presented by Reed [3]. Namely, the accelerations were multiplied by the total mass of the aircraft, and the indicated airspeed was squared to give a measure of dynamic pressure, since loads in air surfaces scale accordingly. Hence a total of 6 inputs were given to the regressor. The objective for the regressor was to correctly predict two strains on the port inner wing and port outer wing based on these inputs. Data from 101 sorties were used, comprising of 16 sorties from aircraft ZF512 and the rest from aircraft ZF406. Training was done on 5 sorties from ZF406 which contained a high level of aerobatic manoeuvres, which cover most of the flight envelope. The chosen flights to include in the training set were the same as in other studies [5], [2] in order to assess the ability of GPR to generalise the results to unseen data against other methodologies.

Although no functional form needs to be chosen a-priori and the training data defines the Gaussian Process through the use of a covariance function, there is still the need for a training phase in order to provide a sensible choice for the hyper-parameters. This can be done using optimisation. In this case a gradient descent optimisation technique was used. For the application of mapping flight parameters to strains, it was found that using a linear mean function and a linear covariance function captured most of the trend in the strains. A squared exponential covariance function was later summed to the linear function which improved the quality of the predictions further. Thus the GPR model implemented here uses a linear+squared exponential covariance function. Furthermore, it is well known that exact prediction using Gaussian Processes can be computationally costly since it requires inversion of an $n \times n$ matrix ($n$ being the number of data points), this becomes computationally expensive when the number of data points grows beyond approximately 2000. In this work, an approximation was implemented using the "fully independent training conditional" (FITC) method [8] for both the training and prediction phases. The idea behind the FITC method is to use a number of "inducing points", where the covariance function is computed for the inducing points only. This provides a low rank plus diagonal approximation to $\mathbf{K}$, the covariance matrix, which is then easier to invert. The use of this approximation was key to the practical implementation of GPR to load monitoring due to the large datasets involved. Training (hyper-parameter optimisation) for the 5 flights totalling 1.96 million points took under 15 minutes. Predictions alone took an average of 3 to 5 seconds per flight. All GP computations presented here were carried out using the Gaussian Processes for Machine Learning (GPML) Toolbox [9] using an Intel Core i5 at 2.5GHz per processor on a 64-bit Windows operating system.

## 3 RESULTS

After optimisation of the hyper-parameters, predictions were made for the strains on all 101 flights. The performance of the strain predictions is assessed using three measures: a mean squared error normalised by the standard deviation (MSE), the Pearson's correlation coefficient, and the accumulated fatigue damage. Table 1 shows a summary of these results

Table 1: Summary of results showing the average of the performance indicators for strain predictions. P3B and P9B indicate port inner and outer wing bending strains correspondingly.

| | **ZF406** | | **ZF512** | | **Training Set** | |
|---|---|---|---|---|---|---|
| **Performance Indicator** | P3B | P9B | P3B | P9B | P3B | P9B |
| Mean Squared Error | 2.308 | 1.074 | 7.365 | 2.714 | 0.754 | 0.765 |
| Pearson's Correlation | 0.992 | 0.996 | 0.977 | 0.988 | 0.997 | 0.997 |
| Relative Damage (Predicted/Measured) | 1.091 | 0.967 | 1.372 | 0.939 | 1.016 | 0.911 |

The MSE and the correlation coefficient give an overall measure of the prediction accuracy. For the normalisation applied to the MSE, a figure under 100% is indicative of captured correlation of the model. The Pearson's correlation coefficient gives an indication of linear correlation between measurements and prediction and is 1 for perfect correlation and 0 for no correlation. These summarise the model error but do not tell the full story. Ultimately it is the accumulation of fatigue damage that will dictate whether the predictions are useful for a monitoring application or not.

The relative damage is the ratio of the damage obtained using the predictions to the damage obtained using the actual strain measurements. In order to derive fatigue damage, the strains were processed using Rainflow cycle counting to extract cycle range-mean pairs, damage was then calculated using Miner's rule for fatigue damage accumulation using an S-N curve with a slope of -4. This synthetic S-N curve ensures all of the cycles are included in the calculation. Although it does not represent any particular material, it serves the purpose of estimating relative fatigue damage. All time series manipulations and fatigue damage calculations were carried out using the nCode Glyphworks package [10]

Figure 3 shows an illustration of typical predictions on the outer wing strain. Figures 3a) and 3b) show a result from a typical good prediction, while Figure 3c) shows a zoom-in of the worst prediction on the outer wing strain (P9B). Finally, Figure 4 compares measured against predicted fatigue damage for both strain channels, highlighting results for both aircraft (ZF406 and ZF512) as well as those flights that were included on the training set (all from ZF406).

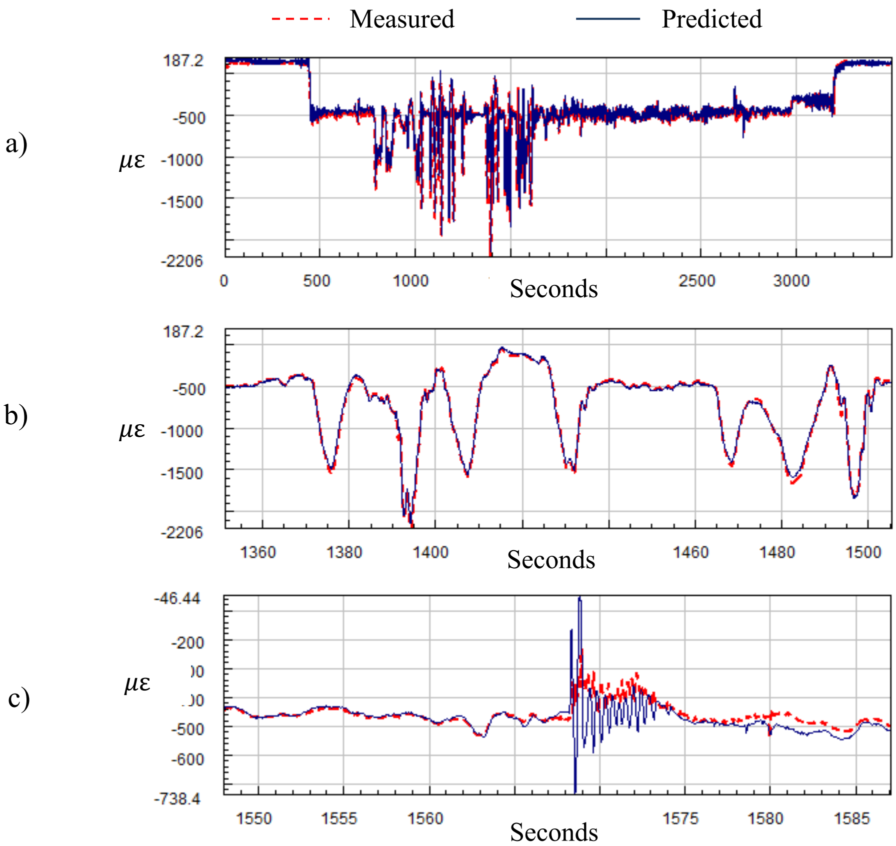


Figure 3 - Example results for the predictions of the Gaussian Process regression model on the outer wing. a) shows a full view of a typical prediction from aircraft ZF406. b) shows a zoomed-in version of the same flight. c) shows a zoom-in during touch-down during the sortie with the worst outer wing strain prediction, from aircraft ZF512.

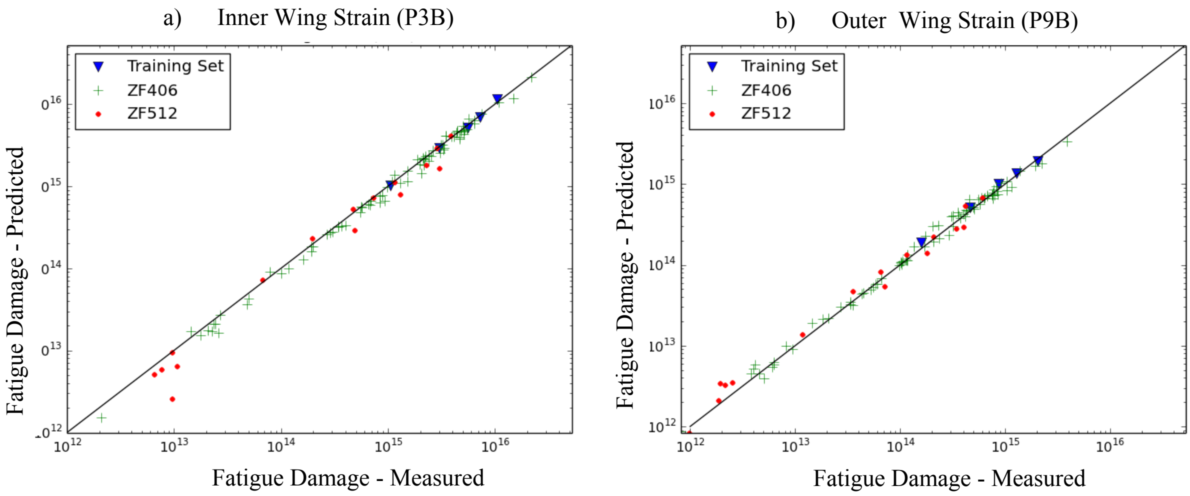


Figure 4 – Comparison of accumulated fatigue damage for measured against predicted strains. for a) Port Inner Wing (P3B) and b) Port Outer Wing (P9B). The markers on each plot highlight whether the prediction was made on aircraft ZF406, ZF512 or from the training set.

## 4 DISCUSSION AND CONCLUSIONS

It has been demonstrated that Gaussian Process regression is a potentially useful tool for predicting loads in aero-structures during operation by showing how it can be used to predict strains from flight parameters. The results are presented in Figure 4 which compares predicted against measured accumulated fatigue damage for each of the 101 flights analysed. A key result is the generalisation ability of the GPR predictions. Generalisation is the ability of a machine learning algorithm to correctly make predictions given inputs from data outside the training set. In this case the training set consisted of five different sorties carried out in a single aircraft (ZF406). The model was able to accurately predict strains on unseen data from the same aircraft with an average error of under 10% in fatigue damage for both inner and outer wing strain channels. The error in fatigue damage on a different aircraft (ZF512) was higher in some cases, and it is useful and constructive to consider the reason for this by looking at the worst results. The worst predictions on this aircraft came from types of sorties not covered within the training set. Typical flights with high model errors include very short flights with long periods on the ground, high altitude flights and flights that involved circuits around an airfield. Figure 3 shows a comparison of a typical good prediction against the worst prediction (in terms of fatigue damage). This particular flight, for example, consisted of circuits with continuous “touch-and-go’s”. These are landings followed by immediate acceleration on the runway and take-off, typical of training sorties. Figure 3c) highlights the touch-down during one of these touch-and-go’s, where the model has incorrectly predicted the strains. Unlike regular landings, during these events, the weight-on-wheels channel did not indicate that the wheels had touched the ground, which may have been the cause for the prediction error. Predictions on regular landings for the rest of the sorties are generally good, with the weight-on-wheels channel indicating the wheels are touching the ground. Increasing the accuracy on these specific manoeuvres might require the addition of more input parameters in order to capture all the factors affecting the structural loads at these type of manoeuvres.

High altitude flying also caused the GPR model to produce poor predictions. As the aircraft climbed higher than the altitude range covered in the training set, the strain predictions became biased. This

is reasonable given that in this study, no means were given to the model to account for the decrease in air density which means a lower dynamic pressure (and thus load) at a given airspeed.
Lastly, it should be mentioned that although the main input flight parameters used in this study followed those used in other studies [3, 4, 5], some of the inputs used here were different. In general, less input parameters were used in this study. A key difference is for example that in [3] an Angle of Attack (AoA) channel is calculated and conditioned by the weight-on-wheels channel, which is then used as an input. In this study the weight-on-wheels discrete channel is used directly as an input and no angle-of-attack was calculated. Another key difference is that no Mach number was used here as input and the flaps position was not conditioned by dynamic pressure.

## 5 ACKNOWLEDGEMENTS

The authors would like to thank Steve Reed from the Defence Science and Technology Laboratory (DSTL) for the advice and information that has enabled this study.

## 6 REFERENCES

[1] J. D. Cronkhite and L. Gill, "Technical Evaluation Report on 1998 Specialists' Meeting on Exploitation of Structural Loads/Health Data for Reduced Life Cycle Costs," in *NATO Research and Technology Organisation Procedings 7*, Brussels, Belgium.

[2] S. Reed, "A parametric-based empennage fatigue monitoring system using artificial neural networks," in *Proceedings of the 23rd Symposium of the International Committee on Aeronautical Fatigue*, Hamburg, Germany, 2005.

[3] S. Reed, "Development of a Parametric-Based Indirect Aircraft Structural Usage Monitoring System using Artificial Neural Networks," *Aeronautical Journal,* 2006.

[4] S. Reed, Indirect Aircraft Structural Monitoring Using Artificial Neural Networks, Sheffield: Dynamics Research Group, Mechanical Engineering Department, University of Sheffield, 2006.

[5] S. Reed and D. Cole, "Development of a Parametric Aircraft Fatigue Monitoring System using Artificial Neural Networks," in *Proceedings of the 22nd Symposium of the International Committee on Aeronautical Fatigue*, Lucerne, Switzerland, 2003.

[6] E. J. Cross, P. Sartor, K. Worden and P. Southern, "Prediction of Landing Gear Loads from Flight Test Data using Gaussian Process Regression," in *Proceedings of the International Workshop in Structural Health Monitoring*, Stanford, CA, 2013.

[7] C. E. Rasmussen and C. K. I. Williams, Gaussian Processes for Machine Learning, Cambridge, Massachusetts: The MIT Press, 2006.

[8] A. Naish-Guzman and S. Holden, "The Generalized FITC Approximation," in *Advances in Neural Information Processing Systems 20 (NIPS)*, 2007.

[9] C. Rasmussen and H. Nickisch, "Gaussian Processes for Machine Learning (GPML) Toolbox," *Journal of Machine Learning Research,* vol. 11, pp. 3011-3015, 2010.

[10] *nCode Glyphworks, nCode Products,* HBM United Kingdom Limited, AMP Technology Centre, Brunel Way, Catcliffe, Rotherham, S60 5WG. UK..