Model-based strategy and surrogate function for health condition assessment of actuation devices

Model-based strategy and surrogate function for health condition assessment of actuation devices

Model-based strategy and surrogate function for health condition assessment of actuation devices

P C Berri[1], M D L Dalla Vedova[1], G Quattrocchi[1] and P Maggiore[1]

[1]Dept. Mechanical & Aerospace Engineering, Politecnico di Torino, 10129 Turin (IT)

Corresponding author: matteo.dallavedova@polito.it

Abstract. Prognostics and Health Monitoring (PHM) is a discipline aiming to determine in advance the Remaining Useful Life (RUL) of a system. To do so, the operation of the system is monitored in search of the early signs of degradation and incipient faults; then, a model for the propagation of faults is employed to estimate the propagation of damages and evaluate the RUL. Usually, a fault threshold is employed as a stopping criterion for the evaluation of damage propagation, but this is not a reliable method when dealing with multiple faults affecting the system at the same time. Specifically, the combined effect of two fault modes can cause the system not to meet its requirements well before the single faults reach their individual thresholds. In this work, we address a model-based strategy to estimate whether the system with incipient faults is still able to meet its performance requirements. The method is applied to aerospace actuators, and performance is evaluated in terms of dynamical response. This model-based algorithm is too slow to be evaluated in real-time, so a Support Vector Machine (SVM) is trained as a surrogate function to speed up the computation. The results and computational times of the full, physics based model and those of its surrogate are compared and discussed.

1. Introduction

Commonly, the PHM process for a system includes three steps, namely (1) Data acquisition and feature extraction, (2) Fault Detection and Identification (FDI), and (3) RUL estimation. In the first step, the behaviour of the system is measured and its informative characteristics (features) are extracted. In the second step, the features are employed to determine the health condition of the equipment: many strategies are available in literature, either physics-based [1-3] or data-driven [4-6]. Eventually, a model for the rate of damage propagation is employed to extrapolate the growth of the faults until a total failure of the system is declared. In common implementations of RUL estimation algorithms, the failure declaration is triggered by the extrapolated fault condition reaching a predetermined threshold: although this strategy may be acceptable in simple cases, it could result in low accuracy when dealing with the combined effects of multiple faults [7, 8]. Indeed, two faults may be individually compatible with the operation of the system, but their superposition could cause the system not to meet its performance requirements. In this work we address an approach initially proposed in [9], and summarised by the flow chart of Figure 1. The propagation of damages is extrapolated through a dynamical model, calibrated on experimental and field data on the wear rate of components. At each integration timestep of the dynamical model, the residual capability of the system is compared to its requirements by an assessment function. Specifically, in [9] the monitored system is an electromechanical actuator, and the assessment function compares the frequency response of the servomechanism under the effect of incipient faults to a set of requirements, and determined whether those incipient faults are compatible or not with the system operation.

The assessment function is model-based and computationally intensive; to achieve real-time computations, it is replaced by a Support Vector Machine (SVM) surrogate.
This study extends that approach to a modified assessment function, increasing the number of considered requirements to achieve a higher representativeness of a real-life application scenario. In particular, the stall load and no-load speed are considered together with the stability margins of the actuator closed loop control. An SVM is trained as a surrogate of the assessment function, and a parametric study on the characteristics of the training algorithm is carried out to highlight the performance and robustness of the method.

2. Application

The application addressed in this paper is an Electromechanical Actuator (EMA) [10-12] for aircraft flight controls. The physics-based numerical model employed to simulate its operation was initially introduced in [13]; its block diagram is reported in Figure 2. The model introduces some significant simplifications to the actual behaviour of the actuator, but has been validated against higher fidelity simulations and experimental data [13, 14]. It is able to simulate with good accuracy the operation of the system with different commands and external loads, either in nominal conditions or under the effect of several incipient fault modes, namely dry friction and backlash of the transmission, partial short circuit of the motor windings, static eccentricity of the rotor and drift of the position feedback sensor.

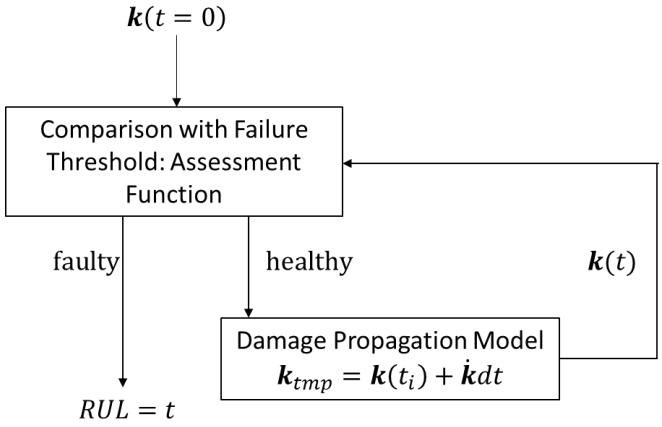

Figure 1. Flow chart of the RUL estimation algorithm [9]

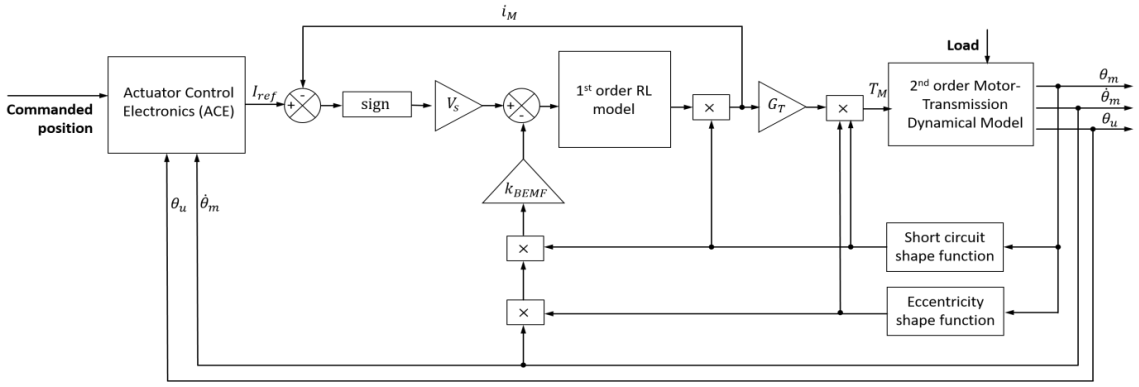

Figure 2. Block diagram of the EMA dynamical model

3. Model-based assessment function

The assessment function is a routine that aims to associate a given set of faults to a "healthy" or "faulty" classification: that is, the performance of the system under the effect of that specific set of faults is assessed and compared to the requirements; if those are still met by the equipment, it is considered as "healthy", otherwise it is "faulty". The assessment function employed in this work is an extension of that presented in [9], and introduces the evaluation of no-load speed and stall-load as an addition to the stability characteristics of the control loop. These features are compared to the corresponding requirements; the system is considered "healthy" as long as all of the requirements are met. This approach differs from the traditional methods for RUL estimation, as threshold is defined on the actual performance delivered from the system, as opposed to on the fault parameters. The accuracy is potentially very high, only limited by the applicability of the employed dynamical model.

3.1. No-load speed

No-load speed is the maximum speed reached by a servomechanism in an actuation without external load. To evaluate the no-load speed from the dynamical model of Figure 2, a large amplitude step command is employed to saturate the position and speed control loops of the actuator. The controller gives the maximum voltage to the motor; the resulting steady state velocity is the no-load speed of the actuator. Figure 3 shows the response of the system to this command. After a short transient the motor reaches a steady state velocity; the maximum slope of the actual position curve is taken as the no-load speed of the actuator

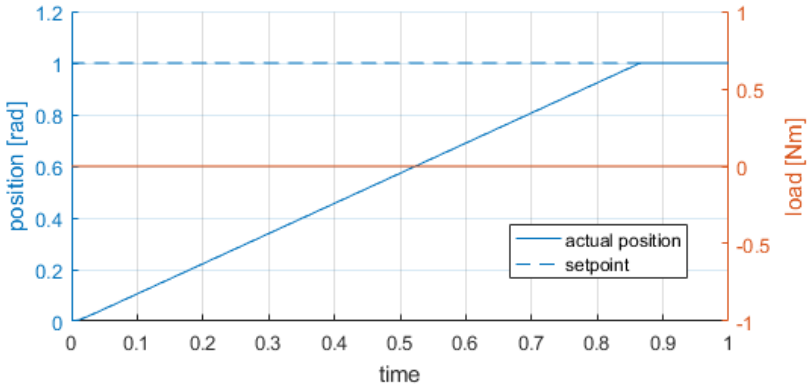

Figure 3. Dynamical response to a large amplitude step command: the position and speed controllers are saturated, enabling the measurement of maximum no-load speed.

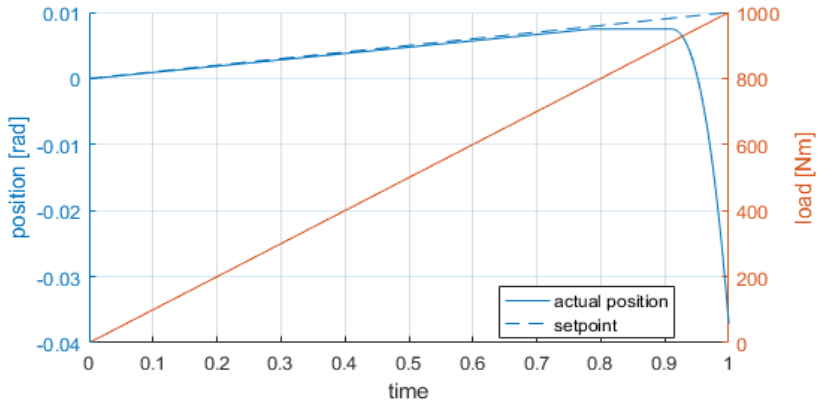

Figure 4. Dynamical response to a low slope ramp command under linearly increasing load.

3.2. Stall load

Stall load is the maximum external load that can be countered by an actuator without stopping or receding. It is evaluated with a similar procedure as the no-load speed. In this case, a ramp command with a low slope is given to the actuator, while the external load increases linearly with time. Figure 4 shows the effect of this operating condition on the behaviour of the system. At the beginning, the actuator follows the position setpoint with high accuracy; as the external load rises, the position error increases until the system stops moving. Since the slope of the ramp command is low, the actuation speed is slow and the inertial contributions to the dynamics of the system are negligible. Then, the external load applied when the movement stops is approximately equal to the stall load of the actuator.

3.3. Dynamical response

The open-loop dynamical response of the actuator is computed to determine the stability characteristics of the control system, in terms of gain margin and phase margin. The numerical model of the EMA contains some significant nonlinearities that cannot be neglected without compromising its accuracy. For example, dry friction, backlash, endstops and saturations operate in a discontinuous manner, and cannot be linearized locally. As a result, the Bode diagram of the EMA cannot be computed with common analytical methods, but shall be determined point by point with direct simulations of the numerical model. To do so, a sine wave input with variable frequency and amplitude is fed to the input of the model; the phase and amplitude of the output are extracted for each frequency. The feedback line is disconnected to obtain the open-loop Bode diagram, which is useful to determine the stability of the system. Different command amplitudes are tested; the corresponding gain and phase are different because the model is strongly nonlinear.

The gain G and phase ϕ of the model response for a given input frequency $\omega/2\pi$ are computed as:

$$G = \frac{\sqrt{Re_y^2+Im_y^2}}{\sqrt{Re_x^2+Im_x^2}} \tag{1}$$

$$\phi = \operatorname{atan}\frac{Im_y}{Re_y} - \operatorname{atan}\frac{Im_x}{Re_x} \tag{2}$$

where x is the input to the system, y is the open-loop output, and:

$$Re_x = \frac{1}{t}\int x \sin\omega t\, dt \tag{3}$$

$$Im_x = \frac{1}{t}\int x \cos\omega t\, dt \tag{4}$$

$$Re_y = \frac{1}{t}\int y \sin\omega t\, dt \tag{5}$$

$$Im_y = \frac{1}{t}\int y \cos\omega t\, dt \tag{6}$$

Then, the gain and phase margins are determined: as shown in Figure 5, by definition, the gain margin is the open-loop attenuation when the phase is -180°. The phase margin is the distance from the actual phase to -180°, when the open-loop gain is 0dB. The Bode diagram is computed for three different amplitudes, as the system is highly nonlinear; the response, as well as the stability margins, vary significantly with the input amplitude.

4. Machine Learning approach

The model-based approach discussed in Section 3 is relatively computationally intensive. A single evaluation of the assessment function requires tenths of seconds to one minute on an average laptop.

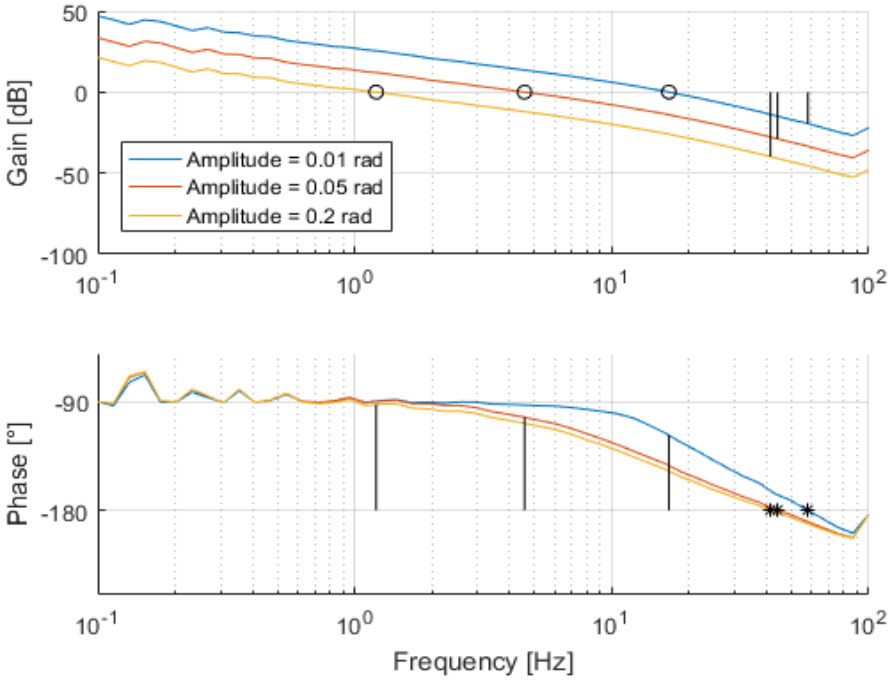

Figure 5. Open-loop bode diagram of the EMA for different input amplitudes. The gain and phase margins are highlighted by the black vertical segments. The black circles on the gain curves highlight the 0dB gains, while the stars on the phase curves are the 180° delays

Although this may be acceptable for some applications, it leads to a prohibitive computational time if the assessment function has to be evaluated iteratively within the procedure of Figure 1, with strict hardware and time constraints: this is the case, for example, of onboard real-time PHM.
To reduce the computational burden and achieve real-time evaluations, we followed the approach proposed in [9]. As the assessment function behaves as a binary classifier, it can be replaced with a Support Vector Machine (SVM) surrogate model. SVMs [15, 16] are supervised learning algorithms that find the best separating hyperplane between two classes of training data points. If the training points are not linearly separable, a kernel projection is employed [17, 18]: the input of the SVM is projected nonlinearly to an auxiliary space, where the points become linearly separable.
For the implementation discussed in this work, we use a standard SVM with polynomial kernel. A training set of 1000 fault combinations is collected with a particular importance sampling [9, 19, 20] and the corresponding "healthy" or "faulty" value is computed with the model-based assessment function. The particular sampling technique allowed to obtain a meaningful representation of both classes, while a sampling with uniform distribution would have collected almost only "faulty" points. Additionally, a validation set of 100 data points was collected.

4.1. Results
A parametric study was performed on the SVM to determine the best settings for the training algorithm. The effect of kernel scale k_s was considered as this hyperparameter scales the input before applying the kernel function, discriminating the variance due to noise from the global shape of the input data, and providing a handle to tradeoff between generalization (i.e. to avoid overfitting) and accuracy of the SVM. For these reasons, k_s is the most influential hyperparameter in this specific SVM implementation.

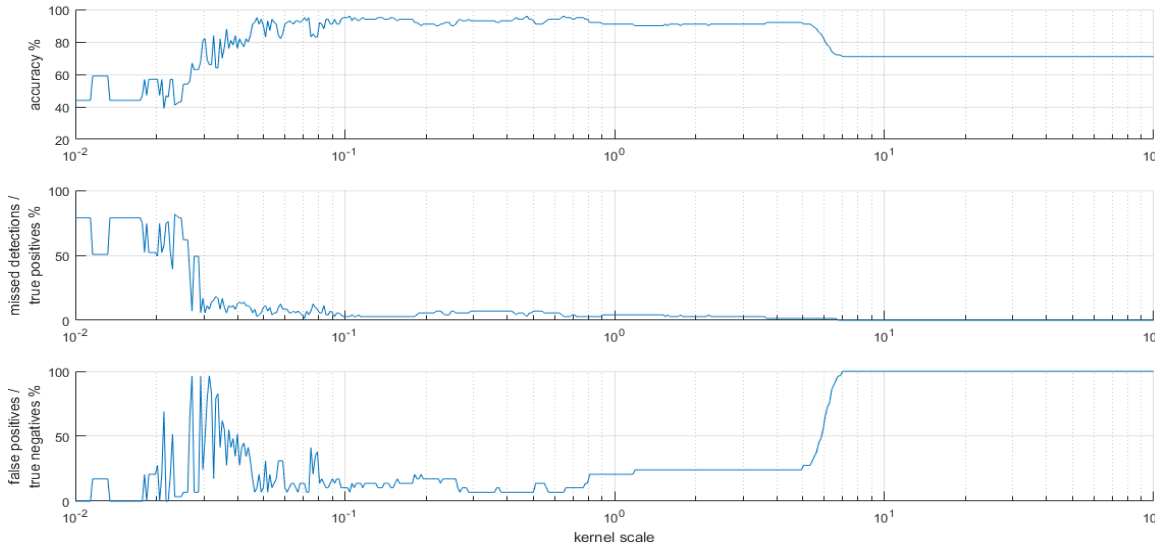

Figure 6. Performance of the SVM (accuracy, missed detections, false positives) versus kernel scale

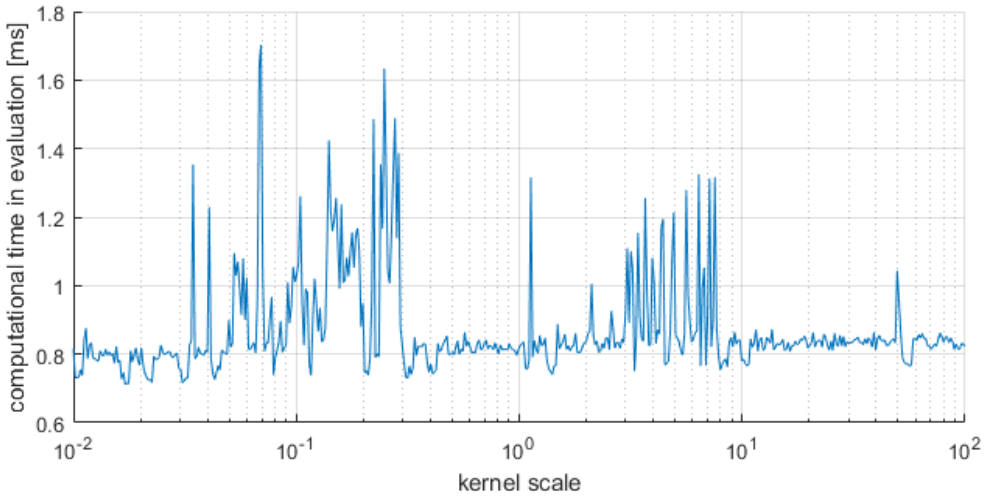

Figure 7. Computational time in evaluation of the SVM versus kernel scale

Figure 6 shows the performance of the SVM on the validation set, for varying k_s. A relatively stable performance is achieved for k_s ranging from 10^{-1} to $5 \cdot 10^{1}$. This behaviour suggests a good robustness of the method with respect to the particular training data and settings employed. The best performance is achieved for $k_s = 0.4735$, with a 96% accuracy; among the 100 validation samples, only 2 false positives and 2 missed detections are found.

Figure 7 shows the computational time in evaluation of the SVM. Apart from a large dispersion in correspondence to the transition of the accuracy performance (i.e. around $k_s = 10^{-1}$ and around $k_s = 5 \cdot 10^{1}$), the evaluation requires a stable computational time of about 0.8 milliseconds. This translates in an improvement of more than 4 orders of magnitude from the model-based approach, and is compatible with real-time, on-board RUL estimation.

5. Conclusions

A methodology for model-based health condition assessment was tested and discussed. The strategy leverages a physics-based model of a dynamical system in order to evaluate the effect of incipient faults on the performance, and to determine whether or not the requirements are still met. This model-based approach can be highly accurate, but requires long computational times, not compatible with real-time evaluation and possibly impractical for RUL estimation during scheduled maintenance.

A parametric study was conducted on a Support Vector Machine to be employed as a surrogate of the model-based approach; this method showed a good accuracy and robustness, while achieving a reduction of computational time by more than four orders of magnitude.

Future work will include the application of this methodology to other case studies, and the integration within a complete RUL estimation framework. Additionally, the RUL estimation algorithm will be tested on resource constrained embedded hardware, in order to prove its capability to run in real-time on-board applications.

6. References

[1] Isermann R 2005 Model-based fault-detection and diagnosis – status and applications *Annual Reviews in Control* vol 29 issue 1 pp. 71-85 ISSN 1367-5788.

[2] Abidin M S Z, Yusof R, Kahlid M and Amin S M 2002 Application of a model-based fault detection and diagnosis using parameter estimation and fuzzy inference to a DC-servomotor *Proc. IEEE Internatinal Symposium on Intelligent Control (Vancouver, BC, Canada)* pp. 783-788 doi: 10.1109/ISIC.2002.1157861.

[3] Ding S X 2013 *Model-Based Fault Diagnosis Techniques - Design Schemes, Algorithms and Tools* (London: Springer-Verlag)

[4] Chen Z 2017 *Data-Driven Fault Detection for Industrial Processes - Canonical Correlation Analysis and Projection Based Methods* (Springer Vieweg).

[5] Qin S J 2009 Data-driven Fault Detection and Diagnosis for Complex Industrial Processes *IFAC Proceedings Volumes* vol 42 issue 8 pp. 1115-1125 ISSN 1474-6670.

[6] Md Nor N, Che Hassan C and Hussain M 2020 A review of data-driven fault detection and diagnosis methods: applications in chemical process systems *Reviews in Chemical Engineering* vol 36 issue 4 pp. 513-553 doi: https://doi.org/10.1515/revce-2017-0069.

[7] Vachtsevanos G, Lewis F L, Roemer M, Hess A and Wu B 2006 *Intelligent Fault Diagnosis and Prognosis for Engineering System*s (Wiley).

[8] Actis Grosso L, De Martin A, Jacazio G and Sorli M 2020 Development of data-driven PHM solutions for robot hemming in automotive production lines *International Journal of Prognostics and Health Management* ISSN 2153-2648.

[9] Berri P C, Dalla Vedova M D L and Mainini L 2019 Real-time Fault Detection and Prognostics for Aircraft Actuation Systems *AIAA Scitech 2019 Forum (San Diego, CA, USA)* AIAA 2019-2210.

[10] Berri P C, Dalla Vedova M D L, Maggiore P and Riva G 2020 Design and Development of a Planetary Gearbox for Electromechanical Actuator Test Bench through Additive Manufacturing *MDPI Actuators* vol 9 35 doi: https://doi.org/10.3390/act9020035.

[11] Qiao G, Liu G, Shi Z, Wang Y, Ma S and Lim T C 2018 A review of electromechanical actuators for More/All Electric aircraft systems *Proc. Inst. of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science* vol. 232 issue 22 pp. 4128-4151.

[12] Fu J, Maré J C and Fu Y 2017 Modelling and simulation of flight control electromechanical actuators with special focus on model architecting, multidisciplinary effects and power flows, *Chinese Journal of Aeronautics*, vol 30 issue 1 pp. 47-65 ISSN 1000-9361.

[13] Berri P C, Dalla Vedova M D L and Maggiore P. 2016 A Smart Electromechanical Actuator Monitor for New Model-Based Prognostic Algorithms *International Journal of Mechanics and Control* vol 17 issue 2 pp. 19-25.

[14] Berri P C, Dalla Vedova M D L and Maggiore P 2018 A Simplified Monitor Model for EMA Prognostics *EASN-CEAS 2018 Matec Web of Conferences 233, 00016.*

[15] Hastie T, Tibshirani R and Friedman J 2008 *The Elements of Statistical Learning, second edition* (New York: Springer).
[16] Fan R E, Chen P H and Lin C J 2005 Working set selection using second order information for training support vector machines *Journal of Machine Learning Research* vol 6 pp. 1889–1918.
[17] Christianini N and Shawe-Taylor J 2000 *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods* (Cambridge, UK: Cambridge University Press).
[18] Kecman V, Huang T M, and Vogt M 2005 Iterative Single Data Algorithm for Training Kernel Machines from Huge Data Sets: Theory and Performance *In Support Vector Machines: Theory and Applications* 255–274 (Berlin: Springer-Verlag).
[19] Bucklew J A 2004 *Introduction to Rare Event Simulation* (New York: Springer-Verlag)
[20] Press W H, Teukolsky S A, Vetterling W T and Flannery B P 2007 Importance Sampling *In: Numerical Recipes: The Art of Scientific Computing (3rd ed.)* (New York: Cambridge University Press) ISBN 978-0-521-88068-8.