Journal of Physics: Conference Series

Identifying evolving leading edge erosion by tracking clusters of lift coefficients

Identifying evolving leading edge erosion by tracking clusters of lift coefficients

Imad Abdallah[1], Gregory Duthé[1], Sarah Barber[2], and Eleni Chatzi[1]

[1] Chair of Structural Mechanics and Monitoring, ETH Zürich
[2] Eastern Switzerland University of Applied Sciences

E-mail: `abdallah@ibk.baug.ethz.ch`

Abstract. This work proposes an approach to identify Leading Edge Erosion (LEE) of a wind turbine blade by tracking evolving and emerging clusters of lift coefficients C_L time-series signals under uncertain inflow conditions. Most diagnostic techniques today rely on direct visual inspection, image processing, and statistical analysis, e.g. data mining or regression on SCADA output signals. We claim that probabilistic multivariate spatio-temporal techniques could play an eminent role in the diagnostics of LEE specifically leveraging C_L time-series signals form multiple sections along the span of the blade. The proposed method extracts clusters' features based on Variational Bayesian Gaussian Mixture Models (VBGMM) and tracks their spatial and temporal changes, as well as interpret the evolution of the clusters through prior physics-based assumptions. The parameters of the VBGMM are the mean, the eigenvalues and eigenvectors of the covariance matrix, and the angle of orientation of the eigenvectors. We show that the distribution of the C_L data may not show statistically separable clusters, however, the parameters of the VBGMM clusters fitted to the C_L data, allows to discriminate moving clusters primarily due to varying inflow and operating conditions, versus emerging clusters primarily due to evolving severity of the blade LEE.

1. Introduction

Leading Edge Erosion (LEE) represents a slow and irreversible accumulation of damage, with effects slow to manifest and hard to detect early on. LEE is caused by variations in the blade surface due to temperature oscillations, moisture, UV radiation, and impacts from raindrops, sand, hailstones or other particles. It could also be initiated by surface cracks due to global strain from blade flexing, errors during manufacturing, inconsistent gelcoat application and bonding strength, or nascent surface damage during blade handling in transport and installation. LEE causes the surface material to be removed from the blade surface, leaving a rough profile that degrades the aerodynamic performance, and, in the long term, if left untreated, will compromise the structural integrity of the blade [6].

Currently, most diagnostic techniques rely on visual inspection, image processing, and statistical analysis via data mining or regression on SCADA output signals [9]. In this work we propose an approach to identify LEE of a wind turbine blade by dynamically tracking moving and emerging clusters [5, 13] of lift coefficients C_L time-series signals form multiple sections along the span of the blade, under uncertain inflow conditions. This work is motivated by current efforts on development of novel micro-electro-mechanical-systems based aerodynamic surface pressure and aero-acoustic measurement systems for Structural Health Monitoring (SHM) of wind turbine blades, as part of the Aerosense project [3].

The remainder of this article is organized as follows. In Section 2, we present the problem setup. In Section 3, we elaborate on the methodological approach to spatio-temporal Variational Bayesian Gaussian Mixture clustering. Finally, we illustrate the novelty and principles of the proposed method, and further provide results and discussion in sections 4 and 5, s.

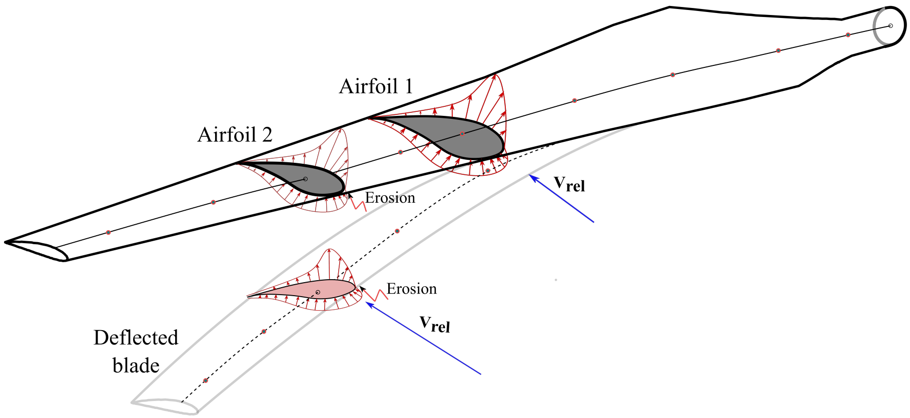

Figure 1. Illustration of a wind turbine blade with one section undergoing leading edge erosion.

2. Problem Setup

A wind turbine blade operates in unsteady and turbulent conditions. To account for this, unsteady numerical aerodynamic models are introduced in the Blade-Element Momentum (BEM) method to correct the input static lift coefficients (C_L) while accounting for large blade deflections/heaving/twisting, flow hysteresis, including unsteady attached flow, trailing-edge flow separation, dynamic stall, flow reattachment, yaw errors and dynamic wake. The corrected C_L (i.e. Dynamic C_L) via the unsteady aerodynamic models are then used to calculate a turbine's power performance and load effects. Throughout this paper we use the corrected dynamic C_L, which are available as output in OpenFAST. However, in "real-life", the dynamic C_L is calculated as the integral of the aerodynamic pressure distribution (more specifically the pressure coefficient curve) over an aerofoil section. This aerodynamic pressure is measured via sensors such as the ones we are developing in the Aerosense project [3] [1].

Let us consider the case where C_L time-series signals emanate from simulated sensors at two sections along the span of the blade as shown in Figure 1, where one section undergoes a LEE process [6], while the second section remains intact. We set R to be the position along the blade from the rotor centre. Figure 2 shows a scatter plot of time-series of C_L at sections $0.96R$ vs $0.74R$ of the blade for mean wind speeds (U) of 6, 11, and 16 m/s and no LEE (clean blade), obtained through OpenFAST [10] simulations on the NREL 5-MW reference wind turbine [11] with turbulent inflow conditions. Figure 3 shows a similar plot, this time with a simulated LEE severity of 9 (highest severity level in our model). We observe that with increasing LEE severity, the clusters systematically shift below the diagonal line (dashed line), but irrespective of the operating wind speed all clusters maintain a seemingly similar orientation. In Figure 4 we fix the mean wind speed to 11 m/s and evolve the severity of LEE at section $0.96R$. We observe that increasing the LEE severity results in a drop in the clusters' mean but also in the overall clusters' scatter. However, at 16 m/s, in Figure 5, we observe that the C_L clusters perfectly

[1] `https://www.aerosense.ai/`

overlap each other for the three LEE severity levels, i.e. they are statistically inseparable. This is expected as the wind turbine is operating at rated power, and the variable pitch set point is adjusted to maintain constant rotor speed.

We can then deduce that the relational dependency of the C_L at two sections along the blade will change over time either (1) due to heterogenous changes in the severity of the LEE and/or (2) due to short-term variations in inflow and operating conditions. Detecting LEE severity could then be regarded as a problem of determining the optimal discriminants for erosion versus other effects from multivariate time-series signals, namely, C_L signals from multiple locations on the blade. The advantage of looking at this problem from a clustering perspective, is that the inference can be fully explainable and the results fully physically interpretable.

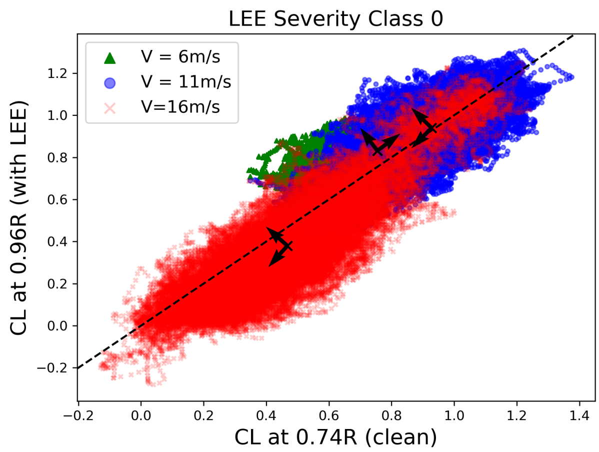

Figure 2. C_L for $U = [6, 11, 16]$ m/s and no blade erosion. (−−): diagonal line. (arrows): eigenvectors.

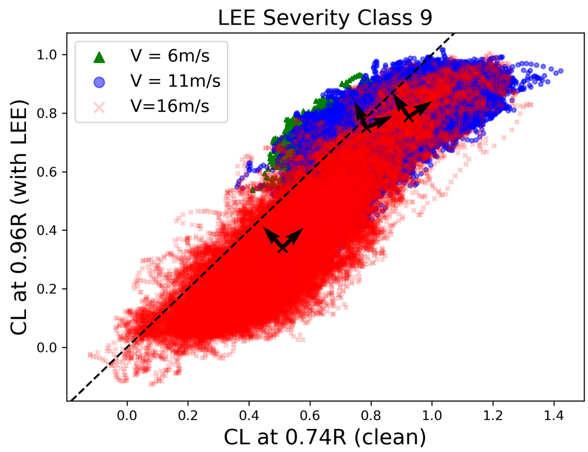

Figure 3. C_L for $U = [6, 11, 16]$ m/s and erosion severity of 9. (−−): diagonal line. (arrows): eigenvectors.

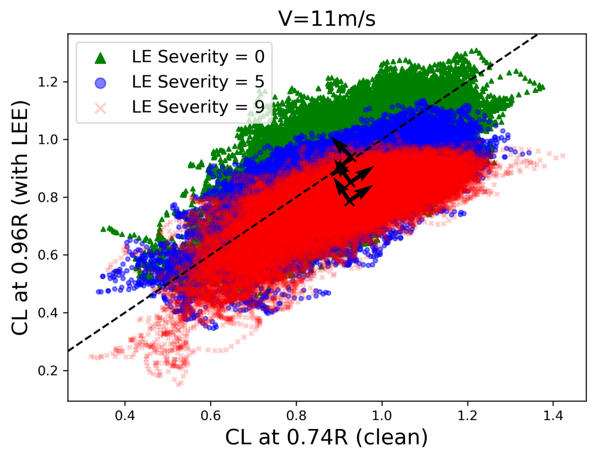

Figure 4. C_L for fixed $U = 11$ m/s and evolving severity of LEE. (−−): diagonal line. (arrows): eigenvectors.

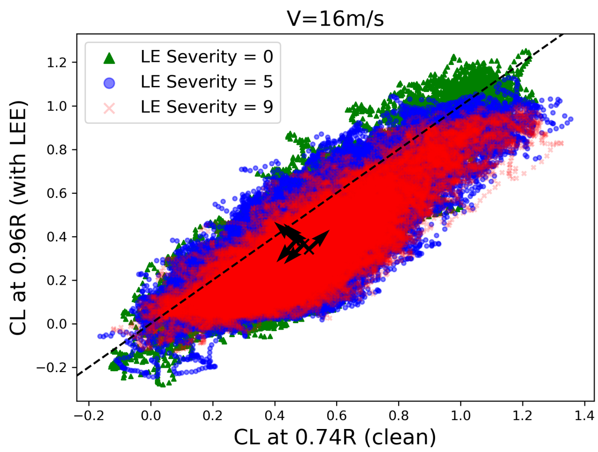

Figure 5. C_L for fixed $U = 16$ m/s and evolving severity of LEE. (−−): diagonal line. (arrows): eigenvectors.

3. Methodology

We propose a data-driven approach for tracking clusters of C_L time-series signals with changing parameters over time and space along the span of the blade. The proposed method extracts clusters' features based on Variational Bayesian Gaussian Mixture Models (VBGMM) and tracks their spatial and temporal changes, as well as interprets the evolution of the clusters through prior physics-based assumptions.

Finite Gaussian mixtures are a flexible probabilistic model for irregularly shaped densities and data samples from heterogeneous stochastic populations generated from a mixture of a finite number K of Gaussian distributions with unknown parameters:

$$p\left(\boldsymbol{\mathcal{Y}} \mid \boldsymbol{\pi}, \boldsymbol{\mu}, \boldsymbol{\Sigma}\right) = \sum_{k=1}^{K} \pi_k \mathcal{N}\left(\boldsymbol{\mathcal{Y}} \mid \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k\right) \tag{1}$$

where π_k is the mixing coefficient, and each Gaussian density $\mathcal{N}\left(\boldsymbol{\mathcal{Y}} \mid \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k\right)$ is a component of the mixture comprising its own parameters: mean $\boldsymbol{\mu}_k$ and covariance $\boldsymbol{\Sigma}_k$ [4]. The classical approach to estimating the parameters is by maximizing the likelihood function of the Gaussian mixture:

$$p\left(\boldsymbol{\mathcal{Y}} \mid \boldsymbol{\pi}, \boldsymbol{\mu}, \boldsymbol{\Sigma}\right) = \prod_{i=1}^{N} \left[\sum_{k=1}^{K} \pi_k \mathcal{N}\left(\boldsymbol{\mathcal{Y}}^{(i)} \mid \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k\right)\right] \tag{2}$$

which is not always a well posed problem because of the singularities that will occur whenever one of the Gaussian components "collapses" onto a specific data point in the dataset (see [4] for proof). Training of Gaussian mixture models suffers the problem of Knowing which component a sample belongs to, which can be solved by using the Maximum Likelihood Estimation (MLE) estimates to update the component. Conversely, knowing the parameters of the components we can predict where each sample lies within each component. This is solved using the expectation-maximization (EM) algorithm, which iterates between the two until convergence. EM is a well-founded statistical algorithm which approaches these problems by means of an iterative process in order to determine maximum likelihood solutions (for details see [4]). A limitation of EM is that it requires a proper initialization(s) in order to consistently find good maximum likelihood solutions. Another limitation is that there is no clear guidance on the choice of the number of Gaussian mixtures (components) K to be used. The Bayesian variational treatment of the Gaussian mixture model circumvents this by automatically providing the number of clusters determined by model selection. The objective is to cluster the data into K components, each of which comprises a mixing coefficient $\{\pi_k, k = 1, \ldots, K\}$. The end goal is to evaluate the posterior distribution:

$$P\left(\boldsymbol{\pi}, \boldsymbol{\mu}, \boldsymbol{\Sigma} \mid \boldsymbol{Y}\right) = \frac{P\left(\boldsymbol{Y}, \boldsymbol{\pi}, \boldsymbol{\mu}, \boldsymbol{\Sigma}\right)}{P\left(\boldsymbol{Y}\right)} = P\left(\boldsymbol{\pi}, \boldsymbol{\mu}, \boldsymbol{\Sigma}\right) \frac{P\left(\boldsymbol{Y} \mid \boldsymbol{\pi}, \boldsymbol{\mu}, \boldsymbol{\Sigma}\right)}{P\left(\boldsymbol{Y}\right)} \tag{3}$$

which is generally intractable. Variational methods are instead used in order to define a tractable lower bound on $P\left(\boldsymbol{Y}\right)$, by minimizing the Kullback-Leibler divergence $KL\left(q(\boldsymbol{\Theta}) \mid\mid P\left(\boldsymbol{\pi}, \boldsymbol{\mu}, \boldsymbol{\Sigma} \mid \boldsymbol{Y}\right)\right)$. Replacing the posterior distribution, it turns out that the following approximation holds true [14]:

$$\begin{aligned} ln\, P\left(\boldsymbol{Y}\right) &= ln\, \frac{P\left(\boldsymbol{Y}, \boldsymbol{\pi}, \boldsymbol{\mu}, \boldsymbol{\Sigma}\right)}{P\left(\boldsymbol{\pi}, \boldsymbol{\mu}, \boldsymbol{\Sigma} \mid \boldsymbol{Y}\right)} \\ &= \underbrace{\mathcal{F}\left(q\left(\boldsymbol{\Theta}\right)\right)}_{\text{Neg. free energy}} + \underbrace{KL\left(q\left(\boldsymbol{\Theta}\right) \mid\mid P\left(\boldsymbol{\pi}, \boldsymbol{\mu}, \boldsymbol{\Sigma} \mid \boldsymbol{Y}\right)\right)}_{\text{KL divergence}} \end{aligned} \tag{4}$$

where $\mathcal{F}\left(q\left(\mathbf{\Theta}\right)\right)$ is the so-called negative free energy, which is a lower bound approximation of $P\left(\boldsymbol{Y}\right)$. The variational method introduces a distribution $q(\mathbf{\Theta})$, which provides an approximation to the true posterior distribution, where $\mathbf{\Theta}$ is a vector collecting all parameters. $q(\mathbf{\Theta})$ generally corresponds to a simple parametric family of the posterior probability density. In this regard, the Kullback-Leibler divergence is chosen as a relative measure of the dissimilarity of the two probability densities $p\left(\boldsymbol{\pi}, \boldsymbol{\mu}, \mathbf{\Sigma} \mid \boldsymbol{Y}\right)$ and $q(\mathbf{\Theta})$. Minimization of $KL\left(q(\mathbf{\Theta}) \mid\mid P\left(\boldsymbol{\pi}, \boldsymbol{\mu}, \mathbf{\Sigma} \mid \boldsymbol{Y}\right)\right)$ reduces the divergence between the true posterior $p\left(\boldsymbol{\pi}, \boldsymbol{\mu}, \mathbf{\Sigma} \mid \boldsymbol{Y}\right)$ and its approximation $q\left(\mathbf{\Theta}\right)$. The problem then reduces to finding the set of probability densities $q(\mathbf{\Theta})$ that maximize the lower bound $\mathcal{F}\left(q\left(\mathbf{\Theta}\right)\right)$, which is equivalent to minimizing the $KL\left(q(\mathbf{\Theta}) \mid\mid P\left(\boldsymbol{\pi}, \boldsymbol{\mu}, \mathbf{\Sigma} \mid \boldsymbol{Y}\right)\right)$ or tightening $\mathcal{F}$ as a lower bound to the log model evidence $ln P\left(\boldsymbol{Y}\right)$. In this work, the algorithm in [12] is adopted to solve the Variational Bayesian Gaussian Mixture inference in order to identify the posterior distribution of the clusters.

Once the clusters have been fitted and identified, and assuming that Gaussian mixture models can correctly model C_L clusters, we proceed to extract the fundamental clusters' features, namely: the mean $\boldsymbol{\mu}_k$, the eigenvalues $\boldsymbol{\lambda}_k$ and eigenvectors $\boldsymbol{v}_k$ (main axis of the cluster) of the covariance matrix $\mathbf{\Sigma}_k$ from which we are able to compute an additional parameter, the angle of orientation of individual C_L clusters [7, 13].

4. Computational Experiment

The numerical simulations are based on a stochastic LEE model coupled to an aeroelastic simulator, for which we provide a short summary here but the reader is referred to [6] for further details. The stochastic spatio-temporal erosion model of the leading edge of wind turbine blades is characterized by a non-homogeneous compound Poisson Process (NHCPP) across discrete LEE severity states. The output of the NHCPP is the stochastic degradation paths describing the erosion process at various sections along the span of the rotor blades as depicted in Figure 6 with severity ranging from 0 (clean zone) to 9 (fully degraded) over a 20 year period. Each arriving shock at a given time along these degradation paths corresponds to a jump of a certain magnitude in the LEE severity, which corresponds to a deterioration of the aerodynamic properties of the affected airfoils' sections. Furthermore, the LEE model takes into account the inherent uncertainty in airfoil static lift and drag coefficients during the erosion period via a stochastic model of static airfoil lift and drag polar curves, the details of which can be found in [1]. The generated stochastic ensemble of degrading airfoil aerodynamic polars are used as input to the OpenFAST aeroelastic simulations. Finally, the coupled model computes the aeroelastic non-stationary response of the wind turbine due to LEE and varying turbulent input inflow conditions over a long period of degradation.

We chose turbulence, σ_U, mean wind speed, U, wind shear exponent, α, horizontal inflow skewness, Ψ, and the vertical inflow skewness, Σ, as described in Table 1, as the main variables for the inflow in the aeroelastic simulations. For each realization of the turbulent inflow wind field, we sample one instance of the stochastic lift and drag coefficients, which form the main input to the OpenFAST aeroelastic simulations. We use the three-bladed up-wind horizontal-axis NREL 5 MW reference wind turbine, with a rotor diameter of 126 m and a 90 m hub height. With this setup, we generate 1200 10-minute aeroelastic response time-series of the wind turbine structure under continuous erosion at two sections along the span of the blade ($0.96R$ and $0.74R$).

5. Results and Discussions

Figure 7 depicts the evolution of the mean (center, given by μ_{CL}) of identified C_L clusters for only LEE severity level zero corresponding to a clean airfoil. In the absence of erosion, the movement

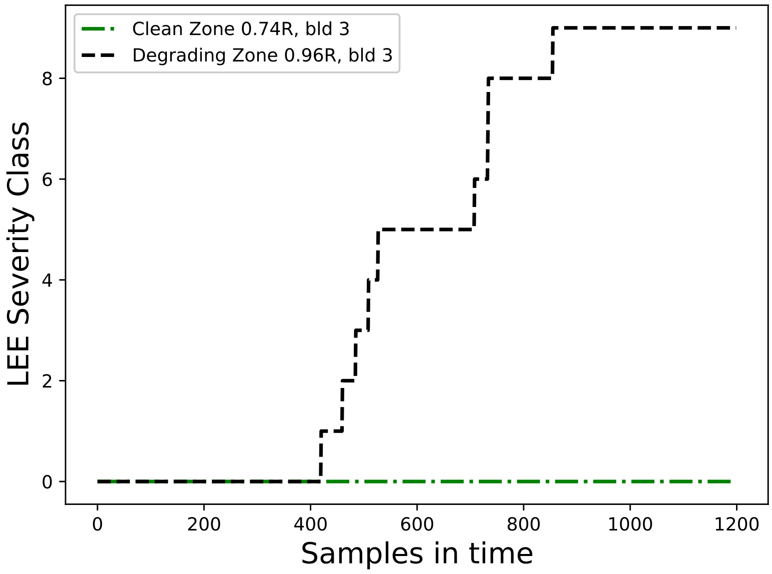

Figure 6. Stochastic LEE degradation paths.

Table 1. Random variables of the turbulent inflow.

Random Variable	Distribution	Parameters
Inflow		
Mean wind speed	$U \sim \mathcal{WBL}\left(A_U, K_U\right)$	$\mathbb{E}\left(U\right) = 8.5,\ where\ A_U = \frac{2\times\mathbb{E}(U)}{\sqrt{\pi}}$ $K_U = 2.0$
Turbulence	$\sigma_U \sim \mathcal{LN}\left(\mu_{\sigma_U}, \sigma^2_{\sigma_U}\right)$	$\mathbb{E}\left(\sigma_U \mid U\right) = I_{ref}\left(0.75u + 3.8\right)$ $\mathbb{V}\left(\sigma_U \mid U\right) = \left(1.4 I_{ref}\right)^2$
Wind shear exponent	$\alpha \sim \mathcal{N}\left(\mu_\alpha, \sigma^2_\alpha\right)$	$\mathbb{E}\left(\alpha \mid U\right) = 0.088\left(ln(u) - 1\right)$ $\mathbb{V}\left(\alpha \mid U\right) = \left(\frac{1}{u}\right)^2$
Horizontal inflow skew	$\psi \sim \mathcal{N}\left(\mu_\psi, \sigma^2_\psi\right)$	$\mathbb{E}\left(\psi \mid U\right) = ln(u) - 3$ $\mathbb{V}\left(\psi \mid U\right) = \left(\frac{15}{u}\right)^2$
Vertical inflow skew	$\Sigma \sim \mathcal{N}\left(\mu_\Sigma, \sigma^2_\Sigma\right)$	$\mathbb{E}\left(\Sigma \mid U\right) = 1.5$ $\mathbb{V}\left(\Sigma \mid U\right) = 1$

of μ_{CL} along a "v"-shaped path simply reflects natural variations due to environmental and operating conditions. Similarly, in Figure 8 we observe that the angle of orientation of the main Eigenvector of the identified C_L clusters are stable over time, and the scatter again simply reflect variations in environmental and operating conditions.

In Figure 9, as the LEE evolves to higher severity levels the clusters' mean μ_{CL} show a systematic and gradual long term shift to lower and non-overlapping operating "v"-shaped paths. We interpret this as follows. The centre of an individual C_L cluster represents the mean operating lift coefficient at two separate airfoil sections along the span of the blade over a 10-min

Journal of Physics: Conference Series **2265** (2022) 032089 doi:10.1088/1742-6596/2265/3/032089

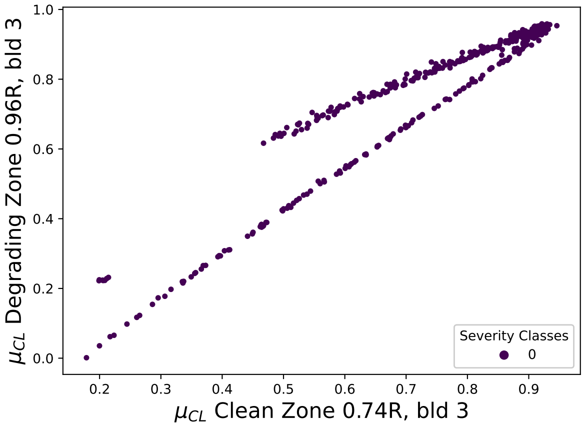

Figure 7. Evolving clusters' means for LEE severity=0 (no erosion).

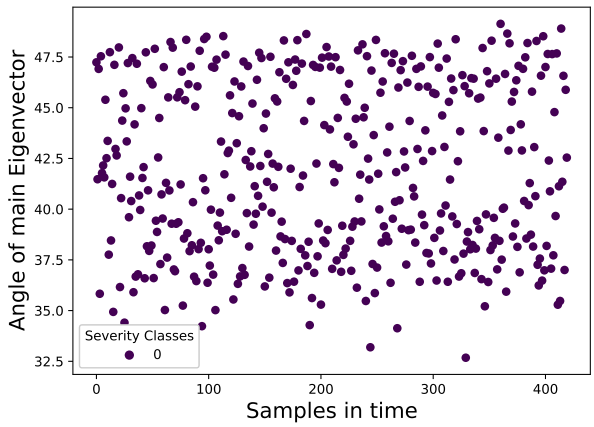

Figure 8. Evolving angle of eigenvectors for LEE severity=0 (no erosion).

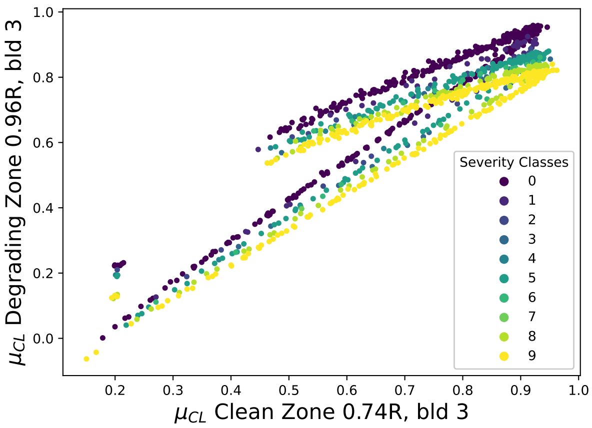

Figure 9. Evolving clusters' means.

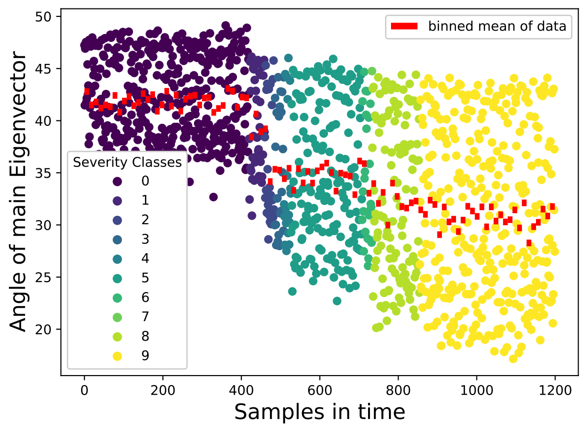

Figure 10. Evolving angle of eigenvectors.

period. The evolution of LEE at one airfoil section is reflected in either a systematic vertical or horizontal translation (shift) of the cluster's mean over time. Consequently, the long term μ_{CL} paths will also shift with increasing LEE severity.

Figure 10 shows the slow but systematic and gradual decrease of the angle of orientation of the main eigenvector of the identified C_L clusters, inversely emulating the LEE degradation paths in Figure 6. In particular we point to the large shift that is observed between samples 400-600. The LEE process is modelled according to a Non-Homogeneous Compound Poisson Process, which essentially compounds the damage due to the stochastic arrival of shocks (as shown in Figure 6). The damage is related to the magnitude of the shock. For instance, a sudden and severe hailstorm accompanied with high wind speeds and large shift in temperature is a shock that will induce a much larger erosion damage compared to, say, a sudden sea wave surge resulting in spraying of salt water on the blades (offshore). So, looking at Figure 6, we observe that between samples 400-600 a quick succession of shocks has brought the blade erosion severity from class 0 to 5. As a result, the large shift in the angle of orientation of the main eigenvector of the identified C_L clusters that we refer to in Figure 10 indicates that our proposed method works and is able to capture the rapid degradation that occurred between samples 400-600.

An additional indicator of LEE in Figure 10 is the larger scatter in the angle of orientation of clusters from newer samples, which can be primarily attributed to elevated LEE severity classes. We interpret this result as follows. The evolution of LEE at one airfoil section is reflected in a reduction in the stall angle of attack and an early onset of flow separation (starting at the trailing edge), which reduces the slope of the C_L curve as a function of the angle of attack. We then expect the orientation of the C_L clusters to decrease as LEE advances at one section vs another along the span of the blade, especially at higher operating angles of attack. Furthermore, lower maximum C_L and lower stall angle of attack, result in the eroding sections entering in and out of stall more frequently, which in turn changes the clusters' orientation more frequently and more severely, yielding larger scatter in the angle of orientation.

From the above results, we claim that we can indeed link the learned clusters' parameters to environmental and operating conditions versus the LEE process. Due to LEE being a very slow deterioration process, in real world SHM applications, the high frequency C_L time-series signals would arrive in streams (sequences/batches) over pre-defined intermittent cycles. The method could thus predict clusters assignment online, in real-time, for each new data batch without accessing the whole historical dataset. An extension to our proposed method is to assume each measurement batch as a sub-sequence of a long time-series allowing us to utilize a graph-based structure of each sub-sequence [15]. This should provide further interpretability to the identified clusters and allow us to discover types of patterns that our approach may be unable to find. Clustering in latent space using variational autoencoders and contrastive learning [8, 2, 16] will also be evaluated in future work.

6. Conclusions

We proposed a method to track the parameters of Variational Bayesian Gaussian Mixtures of evolving spatio-temporal clusters of C_L data from two sections along the span of a blade to identify the occurrence and severity of leading edge erosion. The centre of individual C_L clusters and the angle of orientation of clusters proved to be solid predictors of evolving leading edge erosion severity versus natural variations due to environmental and operational conditions.

Acknowledgement

This work is funded by the BRIDGE Discovery Programm of the Swiss National Science Foundation and Innosuisse, project number $40B2-0\_187087$.

References

[1] I. Abdallah, A. Natarajan, and J.D. Sørensen. Impact of uncertainty in airfoil characteristics on wind turbine extreme loads. *Renewable Energy*, 75:283 – 300, 2015.

[2] Junwen Bai, Shufeng Kong, and Carla P. Gomes. Gaussian mixture variational autoencoder with contrastive learning for multi-label classification, 2021.

[3] S. Barber, J. Deparday, Y. Marykovskiy, E. Chatzi, I. Abdallah, G. Duthé, M. Magno, T. Polonelli, R. Fischer, and H. Müller. Development of a wireless, non-intrusive, mems-based pressure and acoustic measurement system for large-scale operating wind turbine blades. *Wind Energy Science Discussions*, 2022:1–25, 2022.

[4] C Bishop. *Pattern recognition and machine learning.* Springer, ISBN 978-0-387-31073-2, 2006.

[5] David Corne, Julia Handl, and Joshua Knowles. *Evolutionary Clustering*, pages 332–337. Springer US, Boston, MA, 2010.

[6] Gregory Duthé, Imad Abdallah, Sarah Barber, and Eleni Chatzi. Modeling and monitoring erosion of the leading edge of wind turbine blades. *Energies*, 14(21), 2021.

[7] Benjamin Ertl, Jörg Meyer, Achim Streit, and Matthias Schneider. Application of mixtures of gaussians for tracking clusters in spatio-temporal data. In *Proceedings of the 11th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management*, volume 1, pages 45–54. SciTePress, 2019.

[8] Jean-Yves Franceschi, Aymeric Dieuleveut, and Martin Jaggi. Unsupervised scalable representation learning for multivariate time series. In *NeurIPS 2019*, volume 32, 2019.

[9] Robbie Herring, Kirsten Dyer, Ffion Martin, and Carwyn Ward. The increasing importance of leading edge erosion and a review of existing protection solutions. *Renewable and Sustainable Energy Reviews*, 115:109382, 2019.

[10] J. Jonkman and M. Buhl. FAST user's guide. Technical Report NREL/EL-500-38230, National Renewable Energy Laboratory, Golden, CO., USA, 2005.

[11] J. Jonkman, S. Butterfield, W. Musial, , and G. Scott. Definition of a 5-MW Reference Wind Turbine for Offshore System Development. Technical Report NREL/TP-500-38060, National Renewable Energy Laboratory, Golden, CO., USA, 2009.

[12] W.D. Penny. Variational bayes for d-dimensional gaussian mixture models. Technical report, University College London, 2001.

[13] Solange Pruilh, Anne-Sophie Jannot, and Stéphanie Allassonnière. Spatio-temporal mixture process estimation to detect population dynamical changes. *HAL Open Science*, hal-02933217v2, 2021.

[14] S. Sun. A review of deterministic approximate inference techniques for bayesian machine learning. *Neural Comput & Applic*, 23:2039–2050, 2013.

[15] Linxiao Yang, Ngai-Man Cheung, Jiaying Li, and Jun Fang. Deep clustering by gaussian mixture variational autoencoders with graph embedding. In *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 6439–6448, 2019.

[16] Dejiao Zhang, Feng Nan, Xiaokai Wei, Shangwen Li, Henghui Zhu, Kathleen McKeown, Ramesh Nallapati, Andrew Arnold, and Bing Xiang. Supporting clustering with contrastive learning, 2021.