Autonomous Sensor System for Wind Turbine Blade Collision Detection

Kyle Clocker, *Graduate Student Member, IEEE,* Congcong Hu, Jason Roadman, Roberto Albertani, and Matthew L. Johnston, *Senior Member, IEEE*

***Abstract*—This paper presents an automated blade collision detection system for use on wind turbines, toward the goal of supporting monitoring and quantitative assessment of wind energy impacts on wildlife. A wireless, multisensor module mounted at the blade root measures surface vibrations, and a blade-mounted camera provides image capture of colliding objects. Using sensor data recorded during field testing of the system on an operational wind turbine, we present the development, training, and testing of automated detection algorithms for collision detection using machine-learning approaches. In particular, we compare the use of a new two-step, anomaly-based classification algorithm with conventional adaptive boosting and amplitude-based detection techniques, where the two-step approach improves average precision for the experimental data set. This integrated sensor and classification systems demonstrates a new approach for automated, on-blade collision detection for wind turbines, with broad utility across structural health monitoring applications.**

***Index Terms*—Autonomous sensors, wind energy, sensor systems at the edge**

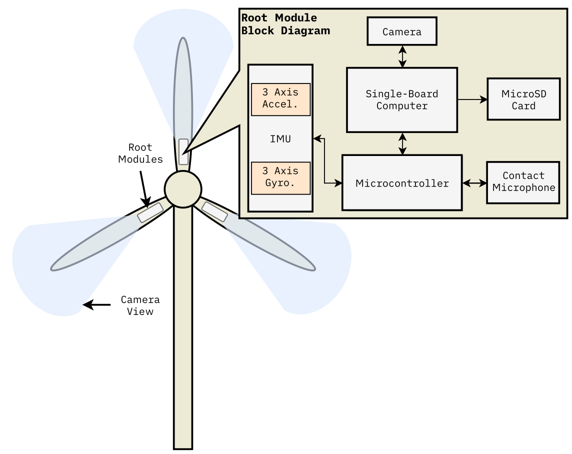

Fig. 1. Overview of a collision detection and imaging system for wind turbine blades. multisensor modules are mounted at the root of each blade to measure vibration, providing input to classification processes used in automated collision detection and image capture of colliding objects.

I. Introduction

Wind turbines serve an increasing proportion of total energy generation, with expanded onshore and offshore installations proceeding worldwide [2], [3]. Continued construction, expansion, and operation of wind energy installations must be managed in conjunction with effects on local and migratory wildlife, specifically bird and bat species that may be affected by wind turbine collisions [4]–[6]. Ongoing efforts to measure and mitigate wind turbine impacts on wildlife include improved preconstruction siting and postconstruction monitoring, development of wildlife deterrent technologies and real-time curtailment strategies, and improved quantitative assessment of wildlife mortality due to wind turbines [7]–[10]. Current automated monitoring solutions leverage cameras on the ground [11] or mounted on the wind turbine tower for recognition of nearby birds [12], and can include audible deterrents or automated curtailment of turbine operation. However, these approaches cannot provide physical verification of blade strikes or images of colliding objects, both of which are critical for quantitative assessment of wildlife impacts. An alternative approach is the use of blade-mounted vibration sensors [13], which have been demonstrated for recording surface vibrations but to date lack the integrated signal processing and automated collision detection required for long-term monitoring.

Toward the goal of long-term, automated blade strike monitoring, we have developed an intelligent, multisensor system designed for installation on a wind turbine blade to provide real-time collision detection along with image capture of colliding objects. The sensor system comprises an accelerometer, gyrometer, and contact microphone to measure vibrations from the blade surface near the hub, and a camera focused down-blade enables video and still image recording along the blade length. Our approach combines physical sensors with trained classification algorithms for automated collision detection.

In this paper, we present an overview of the integrated blade-mounted hardware system, sensor signal processing and feature extraction, development of automated collision detection algorithms using machine-learning approaches, and training and validation using experimental data. The approach is demonstrated using data recorded from multiple blades of an operational wind turbine during field testing of the sensor system. Provided data analysis includes the comparative assessment of multiple classification algorithms, including naive threshold-based classification and two variants of a boosted ensemble classifier, with and without anomaly detection.

The remainder of the paper is organized as follows: Section II provides an overview of the multisensor system designed for real-time, on-blade collision detection monitoring; Section III describes experimental procedures used for data collection on an operational 1.5-MW wind turbine and summarizes recorded data sets; Section IV describes the development, training, validation, and comparison of automated collision detection classifiers; and Section V concludes the paper.

An earlier version of this paper was presented at the 2020 IEEE SENSORS Conference and was published in its Proceedings [1].

K. Clocker and M. Johnston are with the School of Electrical Engineering and Computer Science, Oregon State University, Corvallis, OR 97331 USA (e-mail: clockerk@oregonstate.edu, matthew.johnston@oregonstate.edu).

C. Hu and R. Albertani are with the School of Mechanical, Industrial and Manufacturing Engineering, Oregon State University, Corvallis, OR, 97331.

J. Roadman is with the National Renewable Energy Laboratory, National Wind Technology Center, 15013 Denver W Pkwy, Golden, CO, 80401, USA.

Manuscript submitted March 6, 2021.

Pursuant to the DOE Public Access Plan, this document represents the authors' peer-reviewed, accepted manuscript. The published version of the article is available from the relevant publisher.

1530-437X (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

Authorized licensed use limited to: National Renewable Energy Laboratory. Downloaded on June 02,2021 at 21:42:18 UTC from IEEE Xplore. Restrictions apply.

II. Sensor System Overview

A high-level overview of the automated collision detection system is provided in Fig. 1. A multisensor module is mounted at the root of each wind turbine blade to continuously record local vibrations along multiple axes using accelerometers, gyrometers, and a contact microphone, and a complementary metal-oxide semiconductor (CMOS) imager aligned along the leading edge of the blade towards the tip, is used to capture images of colliding objects. Sensor data can be used for automated collision detection using a trained classifier to trigger recording of the most recent buffered image frames, used for offline analysis, such as species identification.

The primary sensor system is designed for mounting on the surface of each wind turbine blade at the root, near the central hub, where surface vibrations are recorded by multiple transducers. Vibration sensors include a commercial MEMS inertial measurement unit (IMU), comprising three-axis accelerometer and three-axis gyrometer, along with a piezoelectric contact microphone coupled to a custom analog front-end (AFE) input circuit for signal conditioning. A custom printed circuit board (PCB) is used for implementing the AFE, power conditioning for all root module components, IMU interfacing, and a stand-alone microcontroller providing analog-to-digital conversion of the contact microphone signal, as well as all control and digital data flow among sensors.

The system is managed locally by a commercial single board computer (SBC), which interfaces with the custom PCB and with a CMOS imager module; the sensor interface PCB is designed as a daughter board to the SBC to minimize volume. The SBC (Raspberry Pi 3B+) is also used for nonvolatile sensor data logging from the microcontroller, wireless communication, and continuously buffering the CMOS imager data stream to save recordings of detected collisions when triggered by the collision detection algorithm.

Each of three identical wireless sensor systems was mounted in a weather-proof enclosure and validated in a laboratory setting prior to use in the field.

III. System Deployment and Data Collection

The collision detection system was deployed for multi-day field testing in July 2019 on an operational wind turbine at the National Wind Technology Center, located at the National Renewable Energy Laboratory's Flatirons Campus in Boulder, CO. This section details the installation and experimental data collection performed using the developed sensor system, including validation, background vibration recording across multiple operating conditions, and recording during surrogate blade collisions.

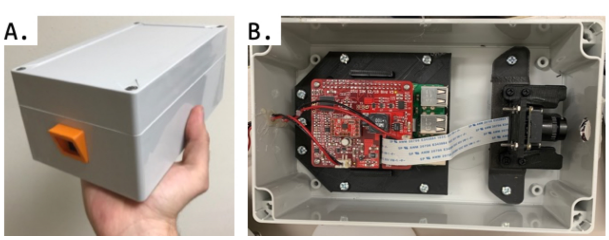

Fig. 2. A multisensor module designed for use on wind turbine blades incorporates a custom printed circuit board (PCB) for data capture from inertial measurement unit and blade-mounted contact microphone, a single-board computer for recording and wireless communication, and a CMOS image sensor for imaging along the length of the wind turbine blade.

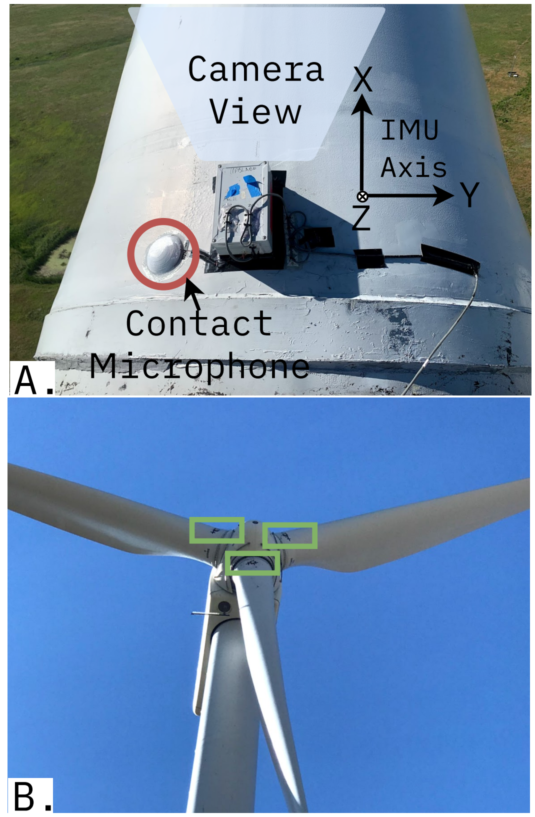

Fig. 3. A) A collision detection module is mounted at the root of the wind turbine blade, near the hub, alongside a blade-mounted contact microphone; the integrated CMOS imager field of view is aligned toward the blade tip; and, B) one module is mounted on each of three blades on a GE 1.5-MW wind turbine for data collection and system validation.

A. System Installation

Sensors were installed on the DOE 1.5 turbine, a GE 1.5-MW three bladed wind turbine, with a rotor diameter of approximately 77 m and a hub height of 80 m. The mounting location for each sensor module was approximately 1.5 m distal to the connection point of the blade to the central hub. A photograph of an installed sensor module is shown in Fig. 3, where each module was installed with its camera view aligned parallel to the leading edge of the blade and focused toward the blade tip. The IMU coordinate system is annotated for use throughout the paper. As shown, the contact microphone was installed next to the root module with direct contact to the blade. A small dome was placed over the contact microphone for wind protection. Sensors were affixed using two-sided adhesive (3M VHB), which provides secure but reversible attachment for multiday testing.

A rail-mounted DC power supply was installed in the rotor hub, on the rotational side of the slip ring. DC power lines (24 V) run from this supply to each root module on the blades.

Pursuant to the DOE Public Access Plan, this document represents the authors' peer-reviewed, accepted manuscript.
The published version of the article is available from the relevant publisher.
1530-437X (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: National Renewable Energy Laboratory. Downloaded on June 02,2021 at 21:42:18 UTC from IEEE Xplore. Restrictions apply.

Fig. 4. Surrogate projectiles, such as tennis balls, were launched into the swept blade path of the operating wind turbine in order to simulate an avian collision with the turbine blades. Projectiles were launched from a compressed air cannon located on a nearby vertical personnel lift.

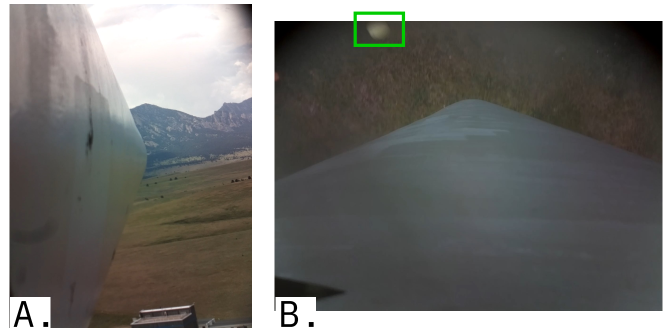

Fig. 5. Images from on-blade CMOS imager located within the root module. Image A. demonstrates the field of view of the on-blade imager. Image B. is a still frame from an automated recording capture triggered by the root module. The surrogate projectile, a potato, is highlighted by the green box annotation.

For sensor module communication and coordination, a wireless access point/router was installed in the internal nacelle housing. Each of the root modules connected to the access point over WiFi. The access point data connection routed to the base of the tower via an Ethernet-over-fiber connection for remote access and control of each of the installed on-blade sensor modules.

B. Baseline Data Collection during Turbine Operation

For data collection and system validation, the sensor modules were installed on a GE 1.5-MW wind turbine, as described previously. Each module was installed at the root of a wind turbine blade, with the camera facing toward the tip, as shown in Fig. 3. All modules were connected to a local, up-turbine wireless network and accessible from the ground using a fiber-optic data connection through the turbine tower.

For initial validation and to provide baseline noise measurements, data were recorded from all on-blade sensor modules in parallel for a variety of turbine operating conditions: this includes both generating and idle rotations (blades rotating with and without the generator engaged); individual blade pitching (rotation about the blade spar axis); and, nacelle rotations (yaw of the nacelle for orientating the turbine blades into the wind). Data were retrieved remotely for review to validate successful sensor module installation and operation prior to projectile testing, including operation of the IMU and contact microphone sensors on each turbine blade, as well as video streams recorded by each on-blade CMOS image sensor.

C. Data Collection Using Surrogate Turbine Blade Collisions

To simulate blade strikes by avian or bat species in a nonlethal manner, surrogate projectiles were launched into the path of the rotating turbine blades. Projectile materials, for these tests tennis balls and small-diameter potatoes, were directed at the wind turbine from a compressed air cannon located on an adjacent vertical personnel lift, approximately 30 m above the ground; this experimental test setup is shown in Fig. 4. For safety, the personnel lift was parked in a location that made it impossible for operators to accidentally extend the basket into the rotor of the operating turbine. Test projectiles were loosed either as individual units or in small batches into the rotational plane of the wind turbine, allowing a collision to occur where the blade strikes the projectile, simulating an avian collision along the leading edge of the turbine blade. Sensor data were recorded by all on-blade modules throughout the testing. In addition to recorded sensor data, additional parameters were recorded, including which blade was struck, approximate blade position at impact, where on the blade impact occurred, current wind turbine operating mode, and projectile type.

D. Automated Camera Capture of Surrogate Blade Collisions

Each blade root sensor module contains a CMOS image sensor aligned to view the blade area distal to the hub, toward the blade tip. The function of this compact camera is to have automated image capture of detected on-blade collisions, which can be later reviewed for collision verification and species identification. The image sensor interfaces with the SBC, which provides a constant, 5 s looping video buffer. If a collision is detected, the 2.5 s buffered frames prior to the collision are automatically saved, along with the 2.5 s of frames following the collision detection; buffer and window sizes are adjustable. Example images provide in Fig. 5 demonstrate the image sensor field of view of along the blade length, and an automatically recorded image of a surrogate collision.

E. Overview of Collected Data

The complete set of recorded on-turbine sensor data was collected, saved, and used to develop a collision detection algorithm (Section IV). Here we summarize the data set generated during field testing of the integrated sensor system.

1) Wind Turbine Operational Modes: The data collected can be categorically labeled by three different wind turbine operational modes used during testing: stopped, idle spinning, and spinning with generator engaged. A stopped turbine occurs during instances with very low wind velocity, and the turbine blade is not rotating; this can be observed as a near-zero slope in y- and z-axis accelerometer data. Idle operation occurs when there exists sufficient wind velocity for rotation, but insufficient wind velocity and aerodynamic torque for the

Pursuant to the DOE Public Access Plan, this document represents the authors' peer-reviewed, accepted manuscript.
The published version of the article is available from the relevant publisher.
1530-437X (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: National Renewable Energy Laboratory. Downloaded on June 02,2021 at 21:42:18 UTC from IEEE Xplore. Restrictions apply.

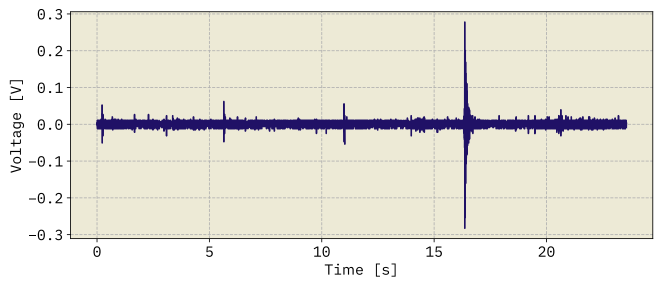

Fig. 6. Typical transient data recorded from a blade-mounted contact microphone using the installed sensor module.

generator to engage and start producing power. Spinning with generation occurs when the turbine is rotating at sufficient rate of rotation to produce power and can be characterized by a sine wave signal in the accelerometer y- and z-axes. All operational mode data sets are useful for baseline noise analysis, and possibly for future applications to structural health monitoring; however, for developing blade collision detection algorithms, analysis was focused on the spinning-mode data sets, which represent the greatest threat for avian collisions.

2) Contact Microphone Data: Each sensor module contact microphone records vibrations on the blade surface near the blade root. These data were sampled at 2 kS/s with a 10-bit sample depth, and an example of typical data recorded during a surrogate blade strike is shown in Fig. 6. Collisions can be visually identified as a decaying spike in the recording. However, other turbine motions can present similar spikes, appearing as artifacts in the contact microphone data, precluding the implementation of a simple, single-sensor, amplitude-based detection algorithm for reliable classification.

3) Accelerometer Data: The on-blade accelerometer measures acceleration on the root module unit in each of the x-, y-, and z-axis directions, as pictured in Fig. 3; these data were recorded at 200 S/s with a 16-bit sample depth; a typical three-axis recording during blade rotation is shown in Fig. 7. During operation, the primary observable feature on each axis is a sine wave, with a base frequency of the wind turbine operational rate of rotation, representing the blades rotation and change of position relative to gravity. This primarily presents in the y- and z-axis recordings, but due both to box alignment on the blade surface and to relative blade pitch of the wind turbine, the peaks and offsets of these periodic signals vary among recordings and from unit to unit. Collisions can be identified as abrupt perturbations from the baseline rotational sine wave. As with the contact microphone, other turbine motions unrelated to collisions can also cause abrupt changes to blade position and can provide similar changes in the accelerometer signals.

4) Gyrometer Data: The three-axis gyrometer measures rotation of installed sensor module about each of the x, y, and z axes Fig. 3, and it shares the same axes orientation and sample rate of 200 S/s as the accelerometer; a typical recording during wind turbine operation is shown in Fig. 8. During wind turbine operation, each axis is relatively constant, with larger changes happening as the wind turbine rotates to match wind direction, or as directed blade movements, such as pitching; these can be identified through characteristic changes to the

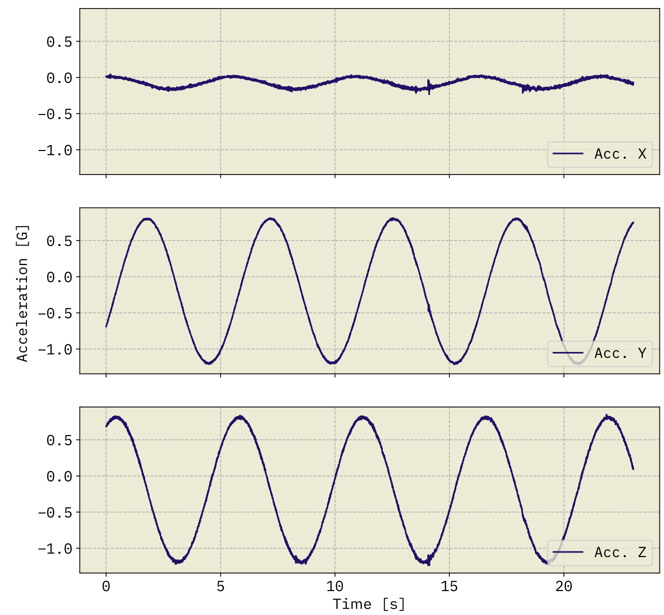

Fig. 7. Typical transient data recorded from the three-axis accelerometer integrated into the installed sensor module.

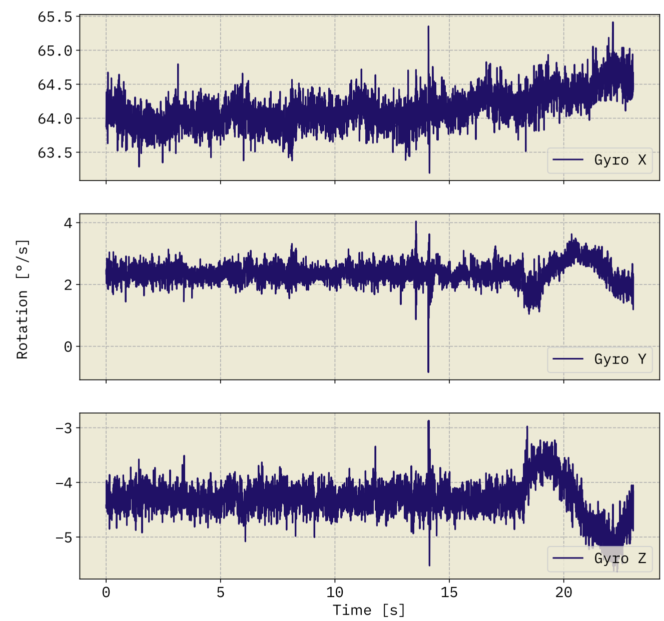

Fig. 8. Typical transient data recorded from the three-axis gyrometer integrated into the installed sensor module.

y-, and z-axis rotation measurements. Blade collisions can be visually identified as transient spikes superimposed on top of the gyrometer baseline.

IV. Automated Collision Detection Methodology

In this section, we present our methodology for developing and evaluating wind turbine blade collision detection algorithms using recorded wind turbine vibration data, gathered as described in Section III. The goal of the collision detection algorithm is to provide an automated system-level trigger that saves a looped video buffer from the on-blade camera,

Pursuant to the DOE Public Access Plan, this document represents the authors' peer-reviewed, accepted manuscript.
The published version of the article is available from the relevant publisher.
1530-437X (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: National Renewable Energy Laboratory. Downloaded on June 02,2021 at 21:42:18 UTC from IEEE Xplore. Restrictions apply.

recording image frames shortly before and after the detected collision; these images and data can be reviewed later to determine the collision source, magnitude, and species.

A. Overview of Collision Detection Methodology

To achieve high-precision automated collision detection, we utilized on-turbine sensor recordings to train and test multiple algorithms, with a goal of improving algorithm sensitivity while reducing the false positive rate.

Specifics of classifier development, training, and testing are described here in more detail. Following a brief description of data sets and preprocessing, we first start by describing the naive detection algorithm for amplitude-based collision detection, which was running on installed sensor modules during field testing. To further improve sensitivity and precision, we then describe a conventional machine-learning approach using an adaptive boost, AdaBoost, classifier. Finally, we present a custom, two-step classification approach using an anomaly detector, further improving the precision of the detection algorithm.

B. Data Sets

Two broad sets of data acquired from on-turbine sensor modules (Section III) were selected for use for algorithm development: the first is a baseline set comprising sensor data recorded over approximately 42 minutes of combined operation, where the turbine was in operational mode, but where no surrogate collisions were recorded and no objects struck the blades; the second data set comprises an additional 25 multisensor stream recordings, each containing a confirmed surrogate blade collision with a launched projectile.

The two data sets can be therefore described by two categories, data that contain a blade collision and data that do not, and further categorical classification is provided by describing wind turbine operational behavior at the time of recording. We have observed three coded categories of wind turbine behavior: stopped, idle, and generating, as described in detail in Section III. Briefly, a stopped turbine occurs during instances with very low wind velocity, and the turbine blade is not rotating; spinning occurs when the turbine is spinning at its operational rate of rotation; and, idle operation occurs with sufficient wind velocity for rotation, but insufficient for generator operation. For developing the detection algorithm, analysis was focused on the spinning mode data sets, which present the greatest threat for avian collisions. Additionally, only recordings from the sensor module mounted to the struck blade were considered; recordings from the two nonstruck blades during each surrogate collision are available for future development, as are other operational modes.

C. Data Preprocessing

Prior to use in any collision detection algorithm training or testing, the data was first subjected to review and preprocessing. The first of these steps was a manual review of the recordings and identification and labeling/annotation of each surrogate collision based on recorded field notes. This labeling process sets the ground truth for further algorithm evaluation; for each recording, a window of time of approximately 2-3 s, depending on the collision and signal decay, was identified and labeled as when the collision is determined to have occurred; an example is shown in Fig. 13.

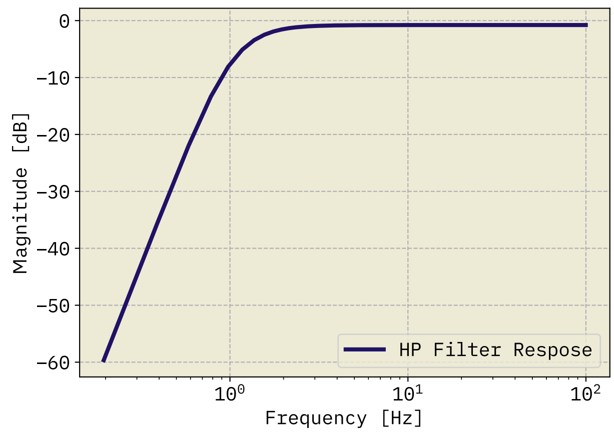

Fig. 9. High-pass frequency response of the digital filter used to remove DC offset and baseline rotational artifacts from recorded IMU sensor signals.

As an additional step, all recorded sensor signals were filtered using a digital high-pass filter to remove sensor DC offsets and low-frequency baseline sensor signal variation that is solely associated with the rotation of the wind turbine blades. The high-pass filter, plotted in Fig. 9, is designed to have a flat gain and 0 dB ripple through the passband, with a 3 dB corner frequency placed at 5 Hz.

D. Naive Collision Detection Algorithm

The naive collision detection algorithm is designed to identify when the peak-to-peak amplitude in a frame exceeds a threshold in the accelerometer data. The complete process is illustrated in Fig. 10. First, all signals are filtered (Fig. 9), after which they are windowed using a sliding window (0.5 s). Within a window, the maximum and minimum amplitude values are determined, and their difference is taken. This process is repeated across all windows and for all sensor axes. The maximum peak-to-peak amplitude differences from each axis are combined via a weighted sum, and the sum is compared against a preset threshold to determine if a collision has likely occurred. Through tuning of the per-axis weighted sum multipliers, this algorithm can detect most surrogate projectile collisions, with a resulting trade-off to the precision of the system, as quantified in Section V.

While straightforward, the naive detection algorithm is limited to information encoded in the sensor signal amplitude alone. Collisions can stimulate vibrations in the turbine blade differently, depending where along the chord or span of the blade the object struck; frequency domain analysis demonstrates this effect and highlights the difficulty of a detection algorithm base purely on time-domain amplitude thresholding [14]. In addition, other wind-turbine-related events such as blade pitching or nacelle yaw can induce vibrations that have similar frame peak values to collision events.

Pursuant to the DOE Public Access Plan, this document represents the authors' peer-reviewed, accepted manuscript.
The published version of the article is available from the relevant publisher.
1530-437X (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: National Renewable Energy Laboratory. Downloaded on June 02,2021 at 21:42:18 UTC from IEEE Xplore. Restrictions apply.

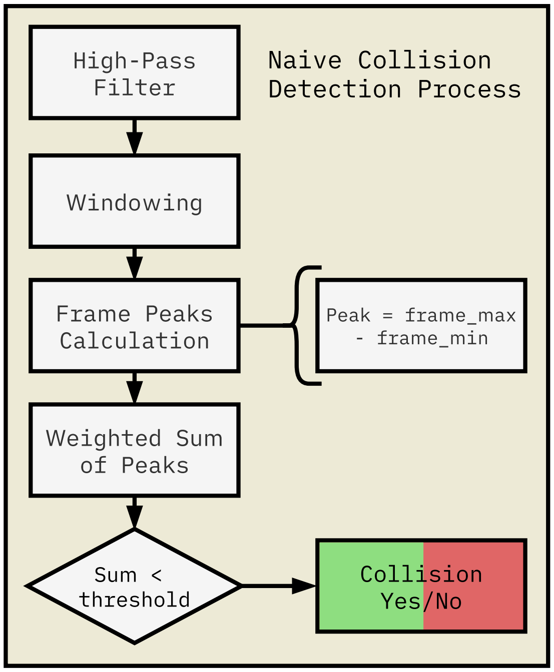

Fig. 10. Signal flow diagram illustrating a naive, amplitude-based collision detection algorithm. Accelerometer signals are first digitally filtered, and then windowed using a sliding window. The maximum peak-to-peak value in each frame is calculated for each axis, and then all axes values are combined using a weighted sum. This sum is compared to a preset threshold to determine if a collision has occurred along the blade.

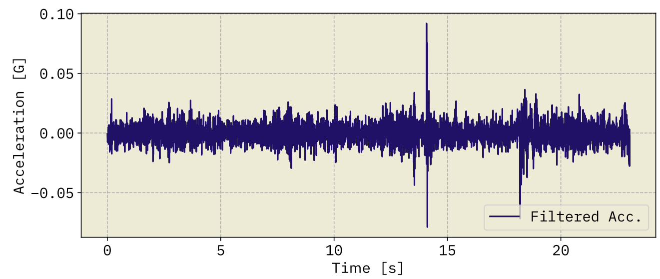

Fig. 11. Typical filtered accelerometer signal recorded from the installed sensor module during a surrogate blade collision includes a collision, visible at ∼14 s, as well as a turbine motion artifact visible at ∼18 s; such artifacts are common due to motor locks and motors, making difficult accurate, automated collision detection using naive amplitude-based thresholding approaches.

Visual analysis of the recorded wind turbine vibration data can highlight some of the challenges in designing an amplitude-based algorithm for collision detection. For example, examining Fig. 11, which plots a filtered accelerometer signal recording during a surrogate collision, the blade strike is visible as the largest peak just before the 15 s mark. However, other perturbations in the accelerometer signal contain similar distinctive spikes and peak-to-peak amplitudes; the perturbations around 18 s, for example, are caused by an artifact from turbine motion rather than a collision with the blade. In this example, a threshold could be set such that the collision could be uniquely identified, but the absolute signal level varies greatly with collision position and attitude as an object strikes the blade, limiting the value of amplitude-based thresholding.

E. Collision Detection Algorithm Using Supervised Learning

To address the challenges presented by the use of a naive classifier, more sophisticated learning-based approaches can introduce analysis of additional signal features and to systematically find a more robust separating boundary between instances of time in which a collision occurs and does not occur. Prior work has demonstrated the potential for using a supervised machine-learning classifier to separate blade collisions from turbine operations [15]. Building on this general approach, we present the design, implementation, training, and testing of two automated classifiers for detecting on-blade collisions.

1) Classifier-Based Collision Detection Algorithm: A classifier is a machine-learning algorithm which attempts to separate data into two more classes. Classification algorithms, generally, represent supervised machine-learning approaches; all data is labeled with a selected class prior to training, such that the ground truth is known during training.

Selecting an appropriate classifier for detection wind turbine blade collisions presents a unique challenge, as most classification algorithms rely on training data that is balanced among the classes, where each class should have approximately the same number of training samples. However, natural collisions with wind turbine blades are rare, and collecting sufficient surrogate collisions to have a balance between collision and non-collision data points is infeasible. In addition, even within a surrogate collision recording, a majority of the recording is non-collision data, as the collision itself is of short duration.

Given these considerations, this application requires a classifier that can handle a highly imbalanced data set. For this work we chose to employ an AdaBoost ensemble classifier [16]–[18]. An ensemble classifier comprises of numerous small models, implemented as decision trees in this work. Each of these smaller models attempt to classify each input and provide a vote toward a selected class.

Implementing AdaBoost can further improve the performance of an ensemble of classifiers. Boosting works by using the error of each classifier to weight the training set and sequentially train additional weak learning models; each of the weak classifiers is then weighted based on its respective error. When a sample is to be classified, every learner votes and their weight is applied to the vote, and the outcome of this voting determines the class of the sample. The AdaBoost algorithm has been broadly demonstrated as effective in classifying in scenarios with imbalanced data [19], [20]. The AdaBoost algorithm has also been demonstrated to be resistant to overfitting [21], which occurs when a model is trained such that it only reflects the data used in training and generalizes poorly, which is also a concern when training with smaller, unbalanced datasets.

A high-level illustration of the classifier-based collision detection system is illustrated in Fig. 12, including filtering, windowing, standardization, and the use of a trained AdaBoost ensemble classifier. This approach was implemented using Scikit-learn [22], and comparative results are provided in Section V.

Pursuant to the DOE Public Access Plan, this document represents the authors' peer-reviewed, accepted manuscript.
The published version of the article is available from the relevant publisher.
1530-437X (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: National Renewable Energy Laboratory. Downloaded on June 02,2021 at 21:42:18 UTC from IEEE Xplore. Restrictions apply.

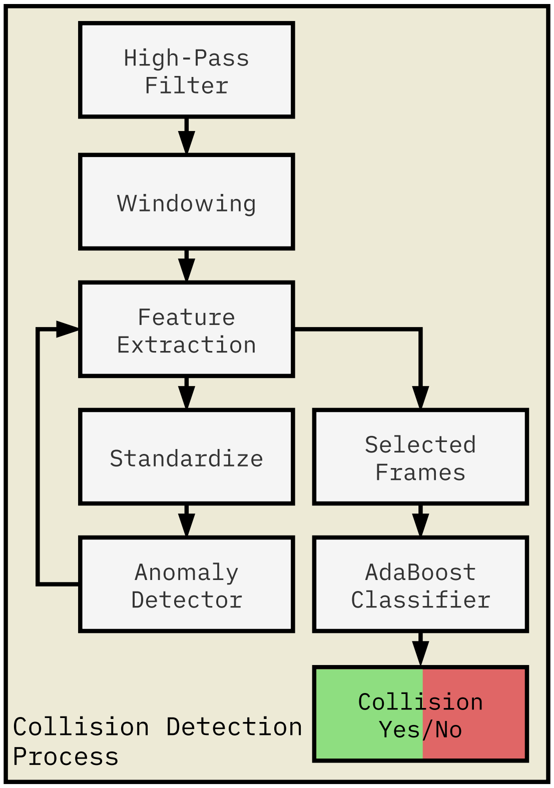

Fig. 12. Signal flow overview of two-step classification process used for classifying collisions with the wind turbine blades. An SVM-based anomaly detector is used in both training and inference operations to improve AdaBoost performance on an inherently imbalanced experimental data set.

2) Feature Selection: Classifiers are trained and tested using calculated features of a given data sample, as opposed to the data itself. Signal features typically comprise statistical measures of each sample window, such as mean, standard deviation, peak amplitude, and more complex quantitative metrics. The selection of proper features is an essential step in designing the collision detection algorithm and extending the performance beyond that of the naive algorithm.

A strong feature set is one that maximizes information from recorded signals while reducing the number of false positive classifications. Feature selection was guided, in part by prior work on wind turbine collision detection classification [15], in part by direct observation of the signals, and in part by machine-learning approaches applied to structural health monitoring for wind turbines and wind turbine blades [23], [24]. For this work, 10 features are calculated for each of the accelerometer and gyrometer axes, yielding a total of 60 features from each sample for training and inference. Several of these features represent conventional signal statistics, including root-mean-square (RMS) value, peak value, standard deviation, variance, kurtosis, and skewness. Additional features were included that measure waveform parameters, identified as previously useful for structural health monitoring on wind turbines and effective in separating anomalous vibration data from background noise [23]; these include energy, crest factor, impulse factor, and RMS entropy estimator.

To generate features for classifier training, the signals were first separated into overlapping 1.5 s windows with a sliding window time step of 0.5 s; the width and step of the time frame were chosen to balance sufficient width to produce relevant statistical features, computational complexity, and system latency from collision to notification. For this system, 1.5 s is long enough to yield relevant statistical features, but short enough that collision data still stands out from background noise. The 0.5 s time step is small enough when combined with the looping video buffer to capture the collision inside the recording, rather than only at a frame edge. Each of these sample windows was labeled to contain either a blade collision or no collision. Features are calculated for all axes of the accelerometer and gyrometer data.

3) Anomaly Detection System: As described, a blade collision is a sparse event during the operation of a wind turbine, leading to an unbalanced training data set. While the AdaBoost algorithm is able to train and generalize an inference model despite this imbalance, other applications classifying sparse signals have demonstrated the use of an anomaly detector to down-sample the majority class, improving classifier performance and a reduction in false positive errors [25]. Anomaly detection is an unsupervised learning approach, where training data is assumed to represent ”normal” behavior, providing a model that can then detect ”abnormal” or anomalous data, without requiring a labeled or balanced training set.

As implemented here, the sample windows are first processed through an anomaly detection algorithm for preselection. Then, the anomalous frames are then, fed into an AdaBoost classifier to determine if they contain a blade collision. The overall process is described by Fig. 12 and further illustrated in Fig. 14. This approach extends the work shown in [1] by including a new, two-step classification process, as well as through the inclusion of additional baseline recordings in the training and testing data set to reflect more realistic wind turbine operating conditions. By operating on only anomalous frames, the system precision can be improved from our prior results.

This work uses a one-class support vector machine (SVM) as a tuned anomaly detector to both down-sample majority frames for training the classifier, and for use as a preselection stage to boost classifier performance during inference (Fig. 14). A one-class SVM is used here as an unsupervised anomaly detection algorithm. It does this by iteratively determining a multi-dimensional boundary around the normal data, and then evaluating if a new point is within the boundary of normal observations; points outside this boundary are considered anomalous. For wind turbine operation, a majority of time is within a normal set of operation, but blade collisions and other acute vibration events can be filtered out as anomalous to that normal operation. The one-class SVM was trained on the set of data that does not contain collisions, and its hyperparameters were then adjusted iteratively such that when the collision data was put through the detector, nearly all labeled collisions were flagged as anomalous.

Finally, the tuned one-class SVM was then used to filter out normal frames and build a testing and training set of only anomalous frames for the AdaBoost classifier to act on. The full signal flow for the two-stage classification system is shown

Pursuant to the DOE Public Access Plan, this document represents the authors' peer-reviewed, accepted manuscript.
The published version of the article is available from the relevant publisher.
1530-437X (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: National Renewable Energy Laboratory. Downloaded on June 02,2021 at 21:42:18 UTC from IEEE Xplore. Restrictions apply.

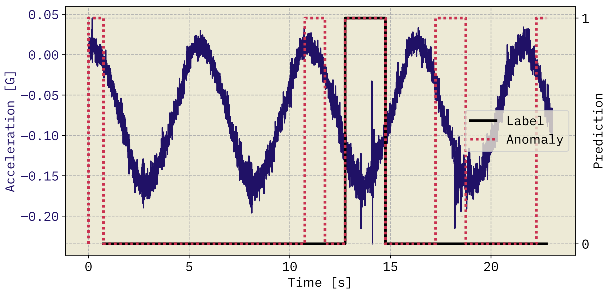

Fig. 13. Comparison of the output of the anomaly detector with the labeled ground truth for accelerometer recording during a surrogate blade collision; while imperfect, as part of a two-step algorithm, the anomaly detector down-selects possible collision frames prior to final classification, decreasing overall false positive rates.

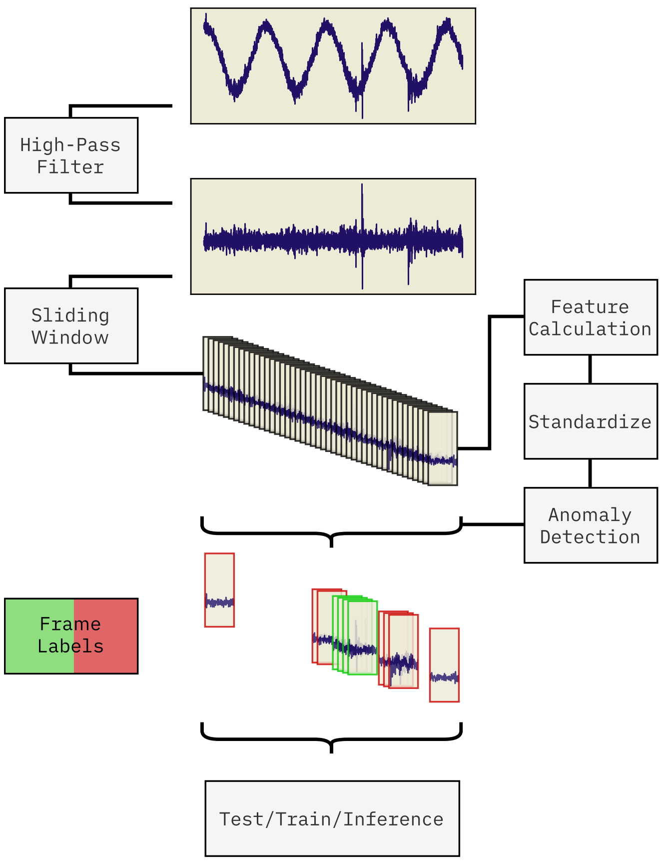

Fig. 14. Illustrated process flow for training and inference data set generation and selection. The anomaly detector is trained on features extracted from the complete set of sensor data windows, and the subset of frames identified as anomalous are labeled with ground truth and randomly split for training, testing, and inference validation.

in Fig. 14.

V. Results and Discussion

A. Classifier Training and Tuning

For comparison, we provide side-by-side performance results from the naive algorithm, an AdaBoost classifier, and the proposed two-step anomaly-based boosted classifier. All

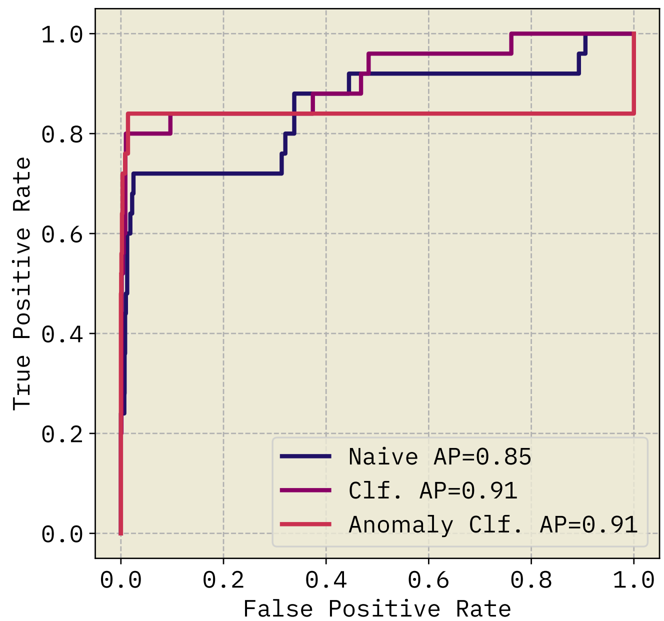

Fig. 15. Receiver operating characteristic curves comparing the performance of the naive, algorithm, an AdaBoost classifier, and the two-step anomaly classifier. The two-step anomaly classifier shows strong performance at low false positive rate, with the inherent trade-off that some test collisions are not classified as anomalous.

algorithms were implemented using the SciKit-Learn package in Python [22]. The classifiers were trained using a combination of recordings with and without collisions, as described in Section III. The full data set was then randomly split such that 75% of the data was used for training and 25% of the data was reserved for testing and evaluation. All results shown from both classifier-based algorithms and the naive algorithm are generated based on the reserved 25% of testing data. Both classification model hyperparameters were tuned using a cross-validated grid search over the training data, where the best model was selected based on maximizing the cross-validated average precision score.

B. Overview of Performance Metrics

The performance of the proposed two-step anomaly classification system for blade collision detection was evaluated using two metrics. The first metric is a receiver operating characteristic curve (ROC), shown in Fig. 15, which illustrates the trade-offs between true positive rate, accuracy, and false positive rate across tunable thresholds for a classification system. Area under curve (AUC) is a measure of the overall performance of a ROC, with a perfect system having an AUC of 1.0, and random chance in a binary classification system having an AUC of 0.5.

While ROC and AUC are useful for understanding how a system performs at various thresholds, it does not take into account any imbalance in the two classifier outcomes. For wind turbine blade collision detection, for example, collisions are rare, and a low false positive rate can still result in a large absolute number of false positive notifications compared to true collision strikes.

Pursuant to the DOE Public Access Plan, this document represents the authors' peer-reviewed, accepted manuscript.
The published version of the article is available from the relevant publisher.
1530-437X (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: National Renewable Energy Laboratory. Downloaded on June 02,2021 at 21:42:18 UTC from IEEE Xplore. Restrictions apply.

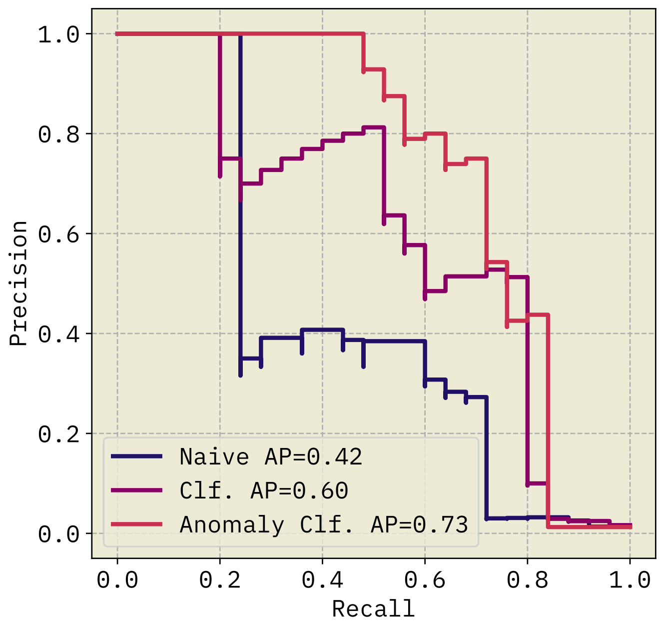

Fig. 16. Precision recall curves comparing the performance of the naive algorithm, an AdaBoost classifier, and the two-step anomaly-based classifier. By focusing the inference on classifying only the anomalous frames, precision is improved compared to both the conventional classifier and the naive algorithm.

TABLE I
COMPARISON OF ROC AND PRC PERFORMANCE

Summary of Results		
Collision Detection Method	AUC	AP
Naive	0.85	0.42
AdaBoost Clf.	0.91	0.60
Anomaly Clf.	0.91	0.73

To help capture this imbalance during our evaluation, a precision-recall curve (PRC) is also presented in Fig. 16, including a measure of average precision (AP). A PRC curve illustrates the trade-offs between precision, the likelihood that a sample classified as a collision was a true collision, with recall, the ratio of total collision events classified to the number of collision events in the test data, across different thresholds. For a system with a precision of 50% and a recall of 70%, for example, the system is capable of detecting 70% of the collision events, and when the system identifies a collision event there is a 50% chance it was an actual collision.

C. Discussion of Results

The results from our evaluation are shown for the ROC, Fig. 15, and for the PRC, Fig. 16, for each of the collision detection algorithms. As demonstrated by the ROC results, all three methods perform relatively well, with the naive algorithm doing slightly worse at more stringent thresholds. The proposed two-step anomaly classifier performs the best at the stringent thresholds, but its performance is limited by the anomaly detection algorithm failing to detect a few collision frames in the initial testing dataset.

Analysis of the PRC results provides a clearer representation of the performance difference among the three presented collision detection approaches. The naive algorithm is limited due to its poor precision, driven by a high false positive rate, and thus performs poorly compared to the two learning-based classification algorithms. By introducing anomaly detection and transforming the broad classifier into a classifier of anomalous data points using our two-step approach, we have improved its selectivity and reduced the false positive rate, boosting the average precision in this test data set to 0.73, compared to 0.60 for the one-step AdaBoost classification-based system without in-built anomaly detection.

VI. CONCLUSION

In this article, we demonstrate the design and implementation of an on-blade sensor system for automated detection of blade collisions on wind turbines, with particular application to monitoring wind energy impacts on wildlife. Through this work, we have designed and deployed a multisensor system on an operational wind turbine and recorded both baseline vibration data and vibration during surrogate blade collisions. Using this field data, preprocessed and randomly split into training/testing data sets, we have developed, trained, and tested a two-step machine-learning algorithm for automated collision detection.

The presented two-step algorithm makes use of a initial SVM-based anomaly detection stage, preselecting anomalous frames from all frames of sensor data recorded by the installed sensor module during normal wind turbine operation. The anomalous frames are provided to an AdaBoost classifier trained on all anomalous frames to provide automated blade collision detection. This new two-step approach was trained and tested using cross-validation, and then evaluated and compared using the same data set with conventional classification approaches, including AdaBoost-only and naive amplitude-based threshold detection. Our two-step approach improves upon both prior methods, with an average precision of 0.73 over the complete recorded data test set.

By developing and validating our automated blade collision detection algorithm using field test data from an operational wind turbine, we demonstrate real-world utility of this approach for future application to the monitoring and mitigation of wind turbine impacts on wildlife. In addition to monitoring blade collisions, the system and algorithmic approach may also be extended for broader structural health monitoring (SHM) of wind turbines, including ice accumulation on blades and lighting strikes, as well as other non-turbine SHM applications across multiple industries.

ACKNOWLEDGMENT

This work was supported in part by the Department of Energy Office of Energy Efficiency & Renewable Energy through Award Number DE-EE0007885.

This work was authored in part by the National Renewable Energy Laboratory, operated by Alliance for Sustainable Energy, LLC, for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. Funding provided by the U.S. Department of Energy Office of Energy Efficiency and Renewable Energy Wind Energy Technologies Office. The

Pursuant to the DOE Public Access Plan, this document represents the authors' peer-reviewed, accepted manuscript.
The published version of the article is available from the relevant publisher.
1530-437X (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: National Renewable Energy Laboratory. Downloaded on June 02,2021 at 21:42:18 UTC from IEEE Xplore. Restrictions apply.

views expressed in the article do not necessarily represent the views of the DOE or the U.S. Government. The U.S. Government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this work, or allow others to do so, for U.S. Government purposes.

REFERENCES

[1] K. Clocker, C. Hu, R. Albertani, and M. L. Johnston, "Sensor system and signal processing for automated blade collision detection on wind turbines," in *2020 IEEE Sensors*. IEEE, 2020, pp. 1–4.

[2] P. Beiter, J. Heeter, P. Spitsen, and D. Riley, "Comparing offshore wind energy procurement and project revenue sources across US states," National Renewable Energy Lab.(NREL), Golden, CO (United States), Tech. Rep., 2020.

[3] R. Lacal-Arántegui, "Globalization in the wind energy industry: contribution and economic impact of european companies," *Renewable Energy*, vol. 134, pp. 612–628, 2019.

[4] T. D. Allison, J. E. Diffendorfer, E. F. Baerwald, J. A. Beston, D. Drake, A. M. Hale, C. D. Hein, M. M. Huso, S. R. Loss, J. E. Lovich *et al.*, "Impacts to wildlife of wind energy siting and operation in the united states," *Issues in Ecology*, vol. 21, pp. 1–24, 2019.

[5] J. E. Diffendorfer, J. Beston, M. D. Merrill, J. Stanton, M. Corum, S. Loss, W. Thogmartin, D. Johnson, R. Erickson, and K. Heist, "A methodology to assess the national and regional impacts of us wind energy development on birds and bats," US Geological Survey, Tech. Rep., 2019.

[6] A. Bose, T. Dürr, R. A. Klenke, and K. Henle, "Predicting strike susceptibility and collision patterns of the common buzzard at wind turbine structures in the federal state of brandenburg, germany," *PloS one*, vol. 15, no. 1, p. e0227698, 2020.

[7] E. B. Arnett and R. F. May, "Mitigating wind energy impacts on wildlife: approaches for multiple taxa," *Human–Wildlife Interactions*, vol. 10, no. 1, p. 5, 2016.

[8] R. May, A. B. Gill, J. Köppel, R. H. Langston, M. Reichenbach, M. Scheidat, S. Smallwood, C. C. Voigt, O. Hüppop, and M. Portman, "Future research directions to reconcile wind turbine–wildlife interactions," in *Wind Energy and Wildlife Interactions*. Springer, 2017, pp. 255–276.

[9] K. C. Sinclair, "Approaches to addressing environmental challenges with wind energy in the united states," National Renewable Energy Lab.(NREL), Golden, CO (United States), Tech. Rep., 2017.

[10] A. E. Copping, A. M. Gorton, R. May, F. Bennet, E. DeGeorge, M. Repas Goncalves, and B. Rumes, "Enabling renewable energy while protecting wildlife: An ecological risk-based approach to wind energy development using ecosystem-based management values," *Sustainability*, vol. 12, no. 22, p. 9352, 2020.

[11] J. Verhoef, C. Westra, P. Eecen, R. Nijdam, and H. Korterink, "Development and first results of a bird impact detection system for wind turbines," *Proc. EWEC 2003, Policies and Programmes*, 2003.

[12] R. F. May, Ø. Hamre, R. Vang, and T. Nygård, "Evaluation of the dtbird video-system at the smøla wind-power plant. detection capabilities for capturing near-turbine avian behaviour," *NINA rapport*, 2012.

[13] C. Hu, R. Albertani, and R. M. Suryan, "Wind turbine sensor array for monitoring avian and bat collisions," *Wind Energy*, vol. 21, no. 4, pp. 255–263, 2018.

[14] L. Rademakers, E. Wiggelinkhuizen, S. Barhorst, and H. de Boon, "Bird collision monitoring system for multi megawatt wind turbines wt bird prototype development and testing," ECN, Tech. Rep., 2006.

[15] C. Hu and R. Albertani, "Machine learning applied to wind turbine blades impact detection," *Wind Engineering*, p. 0309524X19849859, 2019.

[16] Y. Freund, R. Schapire, and N. Abe, "A short introduction to boosting," *Journal-Japanese Society For Artificial Intelligence*, vol. 14, no. 771-780, p. 1612, 1999.

[17] Y. Freund and R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," *Journal of Computer and System Sciences*, vol. 55, no. 1, pp. 119–139, 1997. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S002200009791504X

[18] T. Hastie, S. Rosset, J. Zhu, and H. Zou, "Multi-class adaboost," *Statistics and its Interface*, vol. 2, no. 3, pp. 349–360, 2009.

[19] R. E. Schapire and Y. Singer, "Improved boosting algorithms using confidence-rated predictions," *Machine learning*, vol. 37, no. 3, pp. 297–336, 1999.

[20] Y. Sun, M. S. Kamel, and Y. Wang, "Boosting for learning multiple classes with imbalanced class distribution," 12 2006, pp. 592–602.

[21] J. Friedman, T. Hastie, and R. Tibshirani, "Additive logistic regression: a statistical view of boosting (With discussion and a rejoinder by the authors)," *The Annals of Statistics*, vol. 28, no. 2, pp. 337 – 407, 2000. [Online]. Available: https://doi.org/10.1214/aos/1016218223

[22] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[23] J. Solimine, C. Niezrecki, and M. Inalpolat, "An experimental investigation into passive acoustic damage detection for structural health monitoring of wind turbine blades," *Structural Health Monitoring*, vol. 19, no. 6, pp. 1711–1725, 2020.

[24] T. Regan, C. Beale, and M. Inalpolat, "Wind turbine blade damage detection using supervised machine learning algorithms," *Journal of Vibration and Acoustics*, vol. 139, no. 6, 2017.

[25] N. Islah, J. Koerner, R. Genov, T. A. Valiante, and G. O'Leary, "Machine learning with imbalanced eeg datasets using outlier-based sampling," in *2020 42nd Annual International Conference of the IEEE Engineering in Medicine Biology Society (EMBC)*, 2020, pp. 112–115.

Kyle Clocker (S'12) received the B.S degree in electrical engineering from the University of North Texas, Denton, TX, USA in 2015, and the M.S degree in electrical and computer engineering from Oregon State University, Corvallis, OR, in 2018. He is currently pursuing the Ph.D. degree at Oregon State University. His research interests include developing novel sensing systems at the edge and algorithm development for edge systems.

Concgcong Hu, biography not available at the time of publication.

Jason Roadman received the B.S. degree in aerospace engineering from Virginia Tech, Blacksburg, VA, USA in 2006 and the M.S. degree in aerospace engineering from the University of Colorado, Boulder, CO, USA in 2010. He is a Field Test Engineer with the National Renewable Energy Laboratory National Wind Technology Center.

Roberto Albertani received the M.S. degree in aeronautical engineering from the Polytechnic University of Milan, Italy, in 1980, and the Ph.D. degree in mechanical and aerospace engineering from the University of Florida, Gainesville, FL, USA, in 2005. He joined Oregon State University in 2011, where he is currently an Associate Professor in the School of Mechanical, Industrial and Manufacturing Engineering.

Dr. Albertani's research interests include aerodynamics and stress analysis of flexible structures, environmental impact of wind energy, high-performance sailboat testing techniques, fiber composites technology, micro air vehicles, and biological flight mechanics.

Pursuant to the DOE Public Access Plan, this document represents the authors' peer-reviewed, accepted manuscript. The published version of the article is available from the relevant publisher.

1530-437X (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

Authorized licensed use limited to: National Renewable Energy Laboratory. Downloaded on June 02,2021 at 21:42:18 UTC from IEEE Xplore. Restrictions apply.

Matthew L. Johnston (S'03–M'12–SM'21) received the B.S. degree in electrical engineering from the California Institute of Technology, Pasadena, CA, USA, in 2005, and the M.S. and Ph.D. degrees in electrical engineering from Columbia University, New York, NY, USA, in 2006 and 2012.

He was a Co-Founder and the Manager of Research with Helixis, Carlsbad, CA, USA, a Caltech-based spinout developing instrumentation for real-time DNA amplification, from 2007 to its acquisition by Illumina, San Diego, CA, USA, in 2010. From 2012 to 2013, he was a Post-Doctoral Scholar with the Bioelectronic Systems Lab, Columbia University. In 2014, he joined Oregon State University, Corvallis, OR, USA, where he is currently an Associate Professor with the School of Electrical Engineering and Computer Science. His current research interests include the integration of sensors and transducers with silicon integrated circuits, stretchable circuits and sensor systems, lab-on-CMOS platforms, bio-energy harvesting, and low-power distributed sensing.

Dr. Johnston was the recipient of the 2020 Semiconductor Research Corporation (SRC) Young Faculty Award. He is currently an Associate Editor of the IEEE Transactions on Biomedical Circuits and Systems and the IEEE Open Journal of Circuits and Systems. He also serves on the Biomedical and Life Science Circuits and Systems Technical Committee and the Analog Signal Processing Technical Committee of the IEEE Circuits and Systems Society.

Pursuant to the DOE Public Access Plan, this document represents the authors' peer-reviewed, accepted manuscript.
The published version of the article is available from the relevant publisher.
1530-437X (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: National Renewable Energy Laboratory. Downloaded on June 02,2021 at 21:42:18 UTC from IEEE Xplore. Restrictions apply.