

# Applied machine learning model comparison: Predicting offshore platform integrity with gradient boosting algorithms and neural networks

Alec S. Dyer <sup>a,b,\*</sup>, Dakota Zaengle <sup>a,b</sup>, Jake R. Nelson <sup>a,c</sup>, Rodrigo Duran <sup>a,d</sup>, Madison Wenzlick <sup>a,b</sup>, Patrick C. Wingo <sup>a,b</sup>, Jennifer R. Bauer <sup>a</sup>, Kelly Rose <sup>a</sup>, Lucy Romeo <sup>a,b</sup>

<sup>a</sup> National Energy Technology Laboratory, 1450 Queen Avenue SW, Albany, OR, 97321, USA

<sup>b</sup> NETL Support Contractor, 1450 Queen Avenue SW, Albany, OR, 97321, USA

<sup>c</sup> Oak Ridge Institute for Science and Education, 1450 Queen Avenue SW, Albany, OR, 97321, USA

<sup>d</sup> Theiss Research, 7411 Eads Avenue, La Jolla, CA, 92037, USA



## ARTICLE INFO

### Keywords:

Offshore platform  
Remaining useful life prediction  
Risk assessment  
Machine learning  
Gradient boosting  
Artificial neural network

## ABSTRACT

Offshore oil and gas platforms operating past their design life can pose significant risk to operators and the surrounding environment, as the integrity of these structures decreases over time due to a variety of stressors. This has important implications for industry and government, which are seeking to safely extend the life of platforms for continued use or reuse for alternative offshore energy applications. As a result, there is a need to quantify the remaining useful life (RUL) of operating platforms by analyzing the effects that stressors may have on structural integrity. This study provides a platform risk assessment by employing two machine learning models to forecast the removal age of existing platforms in the U.S. federal waters of the Gulf of Mexico (GoM): a gradient boosted regression tree (GBRT) and an artificial neural network (ANN). These data-driven models were applied to a large, extensive dataset representing the natural and engineered offshore system. Both models were found to provide promising predictions, with 95–97% accuracy and predictions within 1.42–2.04 years on average of the observed removal age during validation. These results can be applied to inform life extension opportunities for fixed and mobile offshore platforms, as well as localized maintenance strategies aiming to prevent operational and environmental risk while maintaining energy production.

## 1. Introduction and background

Globally, the spatial footprint of offshore fossil energy infrastructure is relatively small compared to onshore, with only 3.4% of the

; ANN, Artificial Neural Network; CV, Cross-validation; GBRT, Gradient Boosted Regression Tree; GoM, Gulf of Mexico; KNN, K-nearest neighbor; MAE, Mean absolute error; Metocean, Meteorological and oceanographic; ML, Machine learning; MSE, Mean squared error; PFI, Permutation Feature Importance; RFE, Recursive Feature Elimination; RMSE, Root mean squared error; RUL, Remaining useful life; SIM, Structural Integrity Management.

\* Corresponding author. National Energy Technology Laboratory, 1450 Queen Avenue SW, Albany, OR, 97321, USA.

E-mail address: [Alec.Dyer@netl.doe.gov](mailto:Alec.Dyer@netl.doe.gov) (A.S. Dyer).

<https://doi.org/10.1016/j.marstruc.2021.103152>

Received 3 July 2021; Received in revised form 7 December 2021; Accepted 15 December 2021

Available online 10 February 2022

0951-8339/© 2021 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

world's operating wells located offshore [1]. However, offshore energy infrastructure accounts for nearly 30% of oil production [2]. The productivity of offshore reservoirs, coupled with the significant cost of operating in extreme offshore environments, is a driving factor for industry efforts that are aimed at extending the life of platforms in the Gulf of Mexico (GoM) to maintain energy production through continued operations, such as enhanced oil recovery [3,4]. In addition, offshore infrastructure related to hydrocarbon resource extraction presents future potential for alternative offshore energy production [5,6]. This offers opportunities to recoup costly investments involved in the development and deployment of these assets, but also requires appropriate methods for evaluating the structural integrity and remaining useful life (RUL) of existing offshore structures.

Offshore systems are complex, remote, and corrosive environments. Such systems contain hazards and extreme events, such as hurricanes, which have historically damaged and destroyed platforms [7–10]. Damage to platforms can have environmentally hazardous results including well blowouts and leaks. For example, Hurricane Ivan caused significant damage to a Taylor Energy platform in 2004 after being installed 20 years prior [11], starting a long-lasting leak [12,13]. Furthermore, exploration and production associated with deepwater (over 304.8 meters/1,000 feet) and ultra-deepwater (over 1,524 meters/5,000 feet) locations introduce additional hazards such as stress from weather, seafloor hazards, and geophysical processes in the subsurface [14–16]. In conjunction with environmental hazards, infrastructure is susceptible to cumulative impacts as it ages, increasing vulnerability to failure while operating in hazardous conditions [17]. To mitigate the risks associated with aging infrastructure and help identify opportunities to extend and assess their potential for alternative uses, appropriate models and tools are required to evaluate their RUL.

Current platform risk assessment efforts to track and mitigate the potential for failure include site-specific management strategies such as Structural Integrity Management (SIM), which relies on in-situ observational data [18,19]. Various advancements in the field of machine learning (ML) have worked to enhance these structural evaluation strategies by combining ML models with in-situ data for advanced risk assessments [20,21]. Unfortunately, access to data detailing the equipment, structures, and inspection routines is often proprietary and piecemeal, limiting the potential of ML application. This limitation calls to attention the need to incorporate and analyze data spanning the natural offshore environment and the anthropogenically-engineered oil and gas system (henceforth called the natural and engineered offshore system) to identify regional trends that can then be used to derive insights about individual structures.

While previous work has made some progress in this area, efforts have been limited by data on the factors associated with offshore platform integrity and the subsequent application of advanced predictive methods [17,22]. This work fills that gap through a comparative application of two ML models that have been designed to assess the RUL of offshore platforms in the GoM. Specifically, a Gradient Boosted Regression Tree (GBRT) and a feedforward Artificial Neural Network (ANN) were applied to predict the removal age of existing platforms. These models are informed by a unique and extensive dataset containing structural characteristics, incident history, meteorological and oceanographic (metocean) conditions, production information, and geohazard data [23]. Results from this study highlight key factors associated with platform RUL and can be compared to already established markers of platform integrity. By identifying platforms past their original design life, possible areas of investigation can be suggested to decision-makers regarding platform operations, maintenance, and inspections.

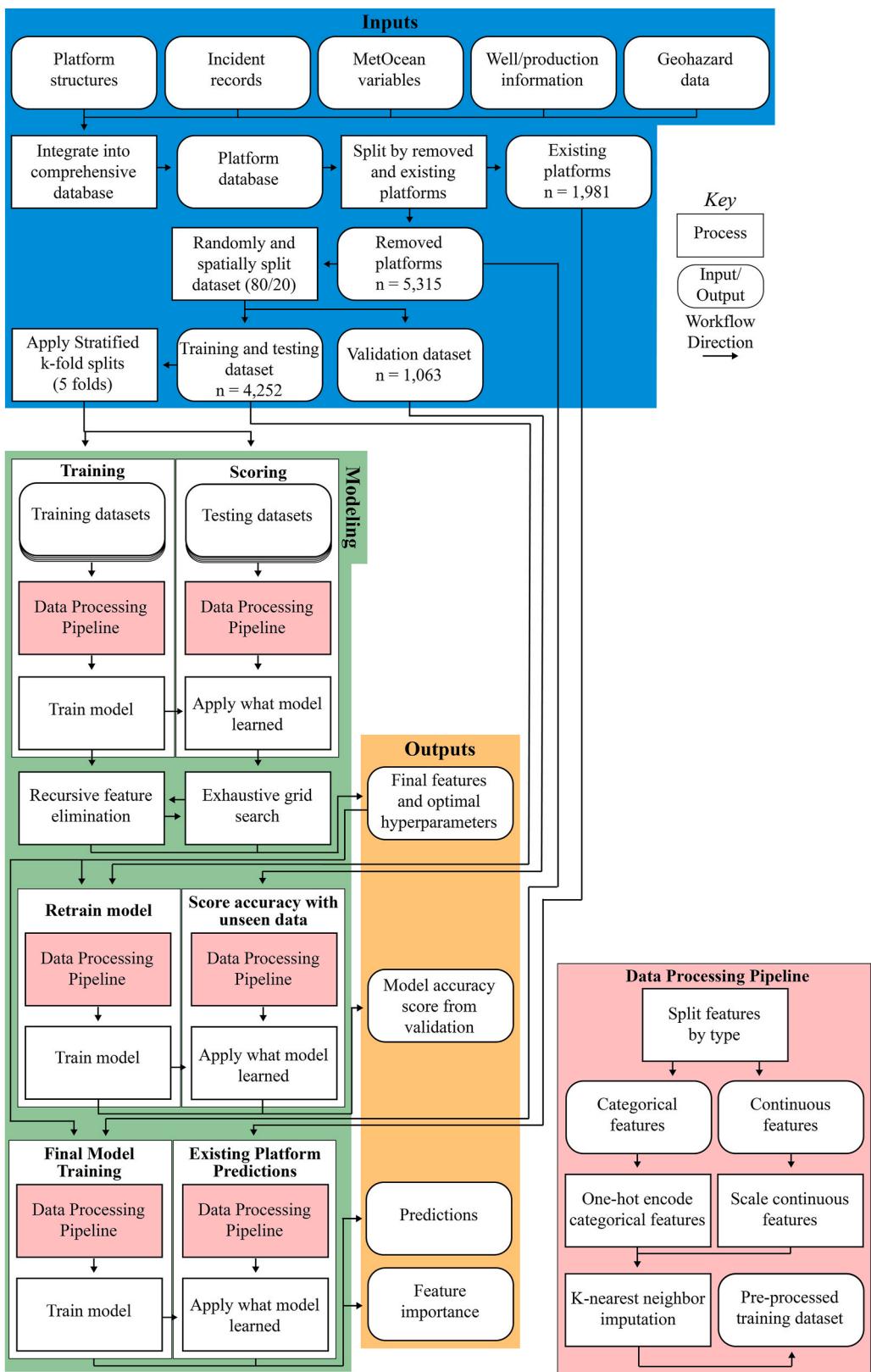
### 1.1. Factors influencing platform lifespan

Platforms are generally designed for a 20- to 30-year operating period, with variations due to structural and environmental factors [24,25]. Historically, platforms have been utilized well past their intended "use-by" date [26,27]. Many offshore conditions are heterogeneous with respect to the following factors: salinity; temperature; oxygen content and nutrients (biological growth); wind, wave, and ocean-current loadings (including storms); seafloor hazards; and subsurface geology. Thus, understanding how these factors influence a platform's lifespan is becoming increasingly important to safely inform life extension decisions, as evidenced by the growing corpus of literature [17,18,25,28–30]. Based on their influence of structural degradation and lifespan, factors of particular interest in this study are corrosion, environmental loadings, structure design, platform incident and production history, and subsurface hazards.

It is well established that mechanical properties such as strength, ductility, and toughness of certain offshore structure materials can corrode over time [9,31–33]. The ambient conditions known to be responsible for these corrosive effects include salinity, dissolved oxygen concentration, temperature, pH, carbonate solubility, pollutants and biological growth, bacteria, pressure, wave action, and water velocity [9,14,33–35]. These conditions change spatially and temporally and therefore have varying effects [32].

In addition to corrosion, environmental loading from wind, wave, and currents contributes to structural fatigue [14], particularly in the case of major storms [8,36]. Persistent stress related to waves and wind can weaken critical joint welds over time [37], while combined wave, wind, and current events can increase loadings and impact integrity [14,38]. During major storm events, waves have been known to exceed platform design limits [39]. Taken together, persistent metocean conditions combined with severe storm events can be detrimental to platform integrity and be used to inform platform RUL.

In relation to external stress and environmental loading, an obvious indicator for evaluating platform lifespan is how long it has been installed. With the estimated design life of roughly 20–30 years [24] depending on design criteria and fabrication, a decent understanding of when to consider a mechanical workover or removal of a platform can be achieved given how long it has been in operation. However, when it comes to estimating design life, Guédé [19] also suggested that structural configuration and foundation systems can be critical determinants. Some configurations may be better suited for longer operation, and these configurations can partially be attributed to the year a platform was designed [19]. Indeed, design standards have advanced as industry and regulators have collected knowledge about the effect of environmental conditions [17,29,40], highlighting the fact that extreme weather events differentially impact platforms constructed under older design standards [7].



**Fig. 1.** Workflow diagram for both models detailing input data processing, the modeling framework with a data processing pipeline to minimize data leakage, and periods when key results were produced.

Of course, any structure operating for an extended period will begin to show fatigue due to daily wear and tear, which might materialize in the form of recurring or frequent component failures. For an offshore platform, these failures are reported as “incidents” and could indicate platform dependability or the need for maintenance. These incidents offer insights into fires, explosions, blowouts, or a loss of well control, which operators are required to report to federal oversight organizations [41]. Importantly, Sharp et al. [42] noted that such incidents can add to structural fatigue, crack growth, material degradation and other incremental damages over time, reducing a platform’s structural integrity. Other findings suggest that the complexity of these structures also matter, as platforms associated with more wells and higher production volume often have more reported incidents [43].

Lastly, well conditions and associated geohazards can have a substantial effect on platform integrity [42,44–47]. Mudslides and subsidence threaten seafloor installations; this includes platform foundations, anchors, and wells [46,47]. Meanwhile, problematic geology, reservoir, and production conditions in the subsurface can cause problems for the connected platform or drilling rig [16,48]. These findings suggest that geologic hazards and subsurface characteristics need to be considered during integrity assessments, including differential settlement and instability [42,45], high-pressure high-temperature zones [49], and mudslides [46,47].

By themselves, these factors give only a slight indicator of their influence on offshore platform lifespan. However, together they may be able to accurately represent a platform’s lifespan as it endures adverse conditions throughout its lifetime. Although other methods do exist, factors that may contribute to platform integrity support the use of advanced data science methods, such as ML, that are positioned to handle big data and extract complex patterns unique to the outcome of interest.

## 1.2. Machine learning applications for infrastructure life extension

Supervised ML is the use of observational data to detect patterns in provided input features (i.e., attributes) that help make predictions on a specific target feature [50,51]. The two supervised ML models used in this study are a GBRT and an ANN. These models are based on different approaches with their own strengths and weaknesses. GBRTs use gradient boosting [52] to create a more accurate model by learning from the residuals of many weak predictors and have been widely used for classification and regression predictions [28,53–55]. Decision trees have the benefit of performing well when data is missing by using surrogates [56,57] and being capable of modeling non-linear feature interactions [57,58].

With the ability to generalize and infer unseen relationships on new data, ANNs use fully connected layers of “neurons” that compute mathematical processes to model complex data relationships and make accurate predictions [59]. In addition, ANNs can represent any continuous function [60] and perform better with a larger training dataset and increased network complexity [61]. A notable pitfall is their lack of interpretability, a subject of ongoing research [62]. These fundamental differences between the ML models influence how they are used in research studies.

There are compelling reasons to apply either model to predict offshore infrastructure lifespan. ANNs have already demonstrated success in predicting failures for offshore structures [63–67], but GBRTs have been applied to predict errors in other energy resource contexts [68] and have continually been improved for developing competitive prediction models [69,70]. At least part of the reason for the dearth in models predicting infrastructure lifespan is related to data limitations and uncertainties. While previous studies have used data from structural specifications and sensors to assess the reliability of offshore infrastructure and equipment [20,21,65], these data generally lack important exogenous features associated with platforms across a region. This presents an issue of scalability — existing models only focus on a specific structure or structural component using data that are not readily available to the public.

In this study, we apply a GBRT and ANN model to predict offshore platform lifespan, using data that integrates the complex natural and engineered offshore system across the northern GoM. By evaluating this system using different ML methods, we can build confidence in the model outcomes by comparing and validating predictions as well as identifying key factors that are used to model platform lifespan. This multiple model approach allows us to apply each model’s strengths to extract key insights, which may be unseen when using only a single model.

## 2. Materials and methods

### 2.1. Study area & data

The study focused on platforms in the U.S. federal waters of the GoM, which includes the Exclusive Economic Zone (see Appendix A). Within this area, there are 1981 platforms operated by more than a hundred oil and gas companies [71]. Half of all U.S. offshore hydrocarbon resources are located here [72], and more recently this region has become a focus for potential enhanced oil recovery, carbon storage, and renewable energy possibilities [73,74].

An extensive dataset integrating structural information, incident reports, metocean data, geohazards, and production records was used to model platform lifespan [23]. This dataset represents all model inputs in Fig. 1. This spatio-temporal dataset, composed of public and private data, includes 7,296 existing and removed platform records dating back to 1942. Platform types used in this study include 7,014 fixed platforms (fixed, caisson, compliant tower, well protector), 58 mobile platforms (floating production storage and offloading, mobile offshore production unit, tension leg platform, mini-tension leg platform, semisubmersible, spar, subsea template), and 224 platform records with an unknown type. A total of 2,615 structural and weather-related incident records dating back to 1956 [41] were matched to platforms based on structure name, structure type, location, and date [75]. The incidents that were able to be directly matched to a platform accounted for 58.31% of all incident records, and these were matched to 14.91% of platforms in the platform dataset. To increase the number of incident records for our analysis, incidents were also matched to platforms by OCS lease block to alleviate these gaps. Metocean condition statistics were tied to each platform’s location and lifetime, which was defined from

installation date to removal date or current date (set at Jan. 1, 2021). Metocean conditions included wave height, wind speed, sea surface current speed, tropical storms and hurricanes, and variables influencing external corrosion (i.e., temperature, salinity, dissolved oxygen, and phosphate, nitrate, and silicate as proxies for marine growth) at varying depths [76–92]. Proprietary well data [93] including annual production and casing pressure statistics were spatially aggregated to each platform, accounting for 23,411 wells and 83.2% of all platforms [75]. Geohazards including faults [94,95], landslides [96–99], seeps [100], and high-pressure, high-temperature data [101] were spatially extracted at the well level, then aggregated by platform [75]. Additional information on the dataset can be found in [Appendix B](#).

## 2.2. Target feature

The goal of the models was to predict each platform's *removal age*, and to use this prediction to accurately forecast a platform's RUL by comparing the predicted *removal age* with the current age of existing platforms. The *removal age* feature was created by calculating the difference in years between the installation date and the removal date or the date in which a platform was destroyed based on incident records. Several platforms did not have a removal date but did have a reported incident with the platform noted as destroyed. As shown in the inputs section of [Fig. 1](#), models were trained and validated on platform records with a removal date, hereinafter referred to as removed platforms. Following the training and validation of the models on the removed platforms, existing platforms were evaluated by forecasting the *removal age*. The *removal age* was then translated into RUL by subtracting the predicted *removal age* from the current age of the platform as of Jan. 1, 2021.

## 2.3. Data processing

As outlined in the Data Processing Pipeline of [Fig. 1](#), several steps were taken to prepare the data for analysis. For comparability, this framework was applied identically to both models when appropriate, with variations noted. First, to decrease the possibility of over-fitting, features that were derivatives of *removal age* were removed from the dataset. Then, the 5,315 removed platforms were randomly split into 80%/20% training and validation datasets. Categorical features were reformatted to binary numeric fields for each unique value, commonly referred to as one-hot-encoding (OHE) [102]. Continuous features were standardized differently by each model, where the GBRT model applied a z-score standardization, and the ANN model scaled all values from 0 to 1.

Approximately 10% of the values in the continuous features were missing, which was handled through imputation [103,104] using the k-nearest neighbor (KNN) algorithm [105,106]. Following imputation, the training dataset was split into five equal-sized folds ( $k = 5$ ) composed of training and testing datasets [107]. A constant random state was used to control the randomness of this split and ensure that each split was identical through training iterations. Five folds were used to estimate the prediction error of the models by fitting to each fold's training dataset and validating performance against the grouped testing dataset (k-fold cross-validation). Cross-validation (CV) was later used during feature selection and hyperparameter tuning.

To prepare the validation dataset for model evaluation, the same data processing pipeline executed on the training dataset was applied to the validation dataset, which ensured that both contained the same distribution. The validation dataset was then held aside for model evaluation.

## 2.4. Optimizing performance

With more than 1,000 features going into the models, the dimension size for both models was reduced using recursive feature elimination (RFE). This approach evaluates each of the features in the dataset and removes those that are the weakest predictors, deriving a set of features that is optimal for predicting *removal age* [108].

Since hyperparameter values influence the speed and quality during a ML model's learning process [109,110], additional steps were taken to tune each model properly. Hyperparameter values were optimized using an exhaustive grid search method, which executed all hyperparameter permutations and evaluated each set's performance using k-fold CV. The model's optimized hyperparameters are available in [Appendix C](#).

## 2.5. Evaluating model performance

Both models used regression algorithms to predict *removal age*. The GBRT model used the *eXtreme Gradient Boosting* algorithm provided by Chen and Guestrin [69]. The ANN used a Multilayer Perceptron implemented with the PyTorch library [111]. The output from the ANN was constrained to the range of *removal ages* in the training dataset to adjust for extreme predictions.

The evaluation metrics used to assess the prediction accuracy of the supervised ML models include correlation coefficient ( $R^2$ ) [55, 66], mean absolute error (MAE) [66], mean squared error (MSE) [112], and root mean squared error (RMSE) [53,66,113]. There is no "single best metric" to evaluate ML models [114,115] and some researchers dispute the utility of some of these metrics [114–118]. All metrics were calculated for each of the folds of training and testing datasets and were averaged to get general performance metrics. Model accuracy against unseen data was estimated by training each model on the full training dataset and evaluating the predictions on the validation dataset.

Models trained with all the removed platform data were used to predict the *removal age* for existing platforms and determine feature importance. To identify features that had the highest impacts on predicting *removal age*, feature importance was calculated using permutation feature importance (PFI), which is the difference between the  $R^2$  score before and after a feature is permuted from the

dataset [119]. Feature permutation depends on the model's learned feature weights; consequently, a feature's importance may change with model variations.

### 3. Results

Model results during training and testing, as well as hyperparameter tuning results, can be found in [Appendix C](#).

#### 3.1. Accuracy assessment

The performance metrics of the models when predicting on the held-out validation dataset is shown in [Table 1](#). By these metrics, the GBRT model outperformed the ANN model, with MAE, MSE and RMSE values about 70–80% of the values from the ANN model; the GBRT  $R^2$  score was also slightly higher. When comparing the metric scores using the training and testing datasets and the validation dataset, there was minimal overfitting. When scoring on the training datasets, the  $R^2$  was 1.00 for the GBRT model and 0.97 for the ANN model, which was within 0.03 and 0.02, respectively, of their validation scores shown in [Table 1](#). The GBRT model predicted an average *removal age* of 21.19 years on the validation dataset, with a standard deviation of 13.61 years. The ANN model generated similar results with an average *removal age* of 21.23 years and a standard deviation of 12.98 years on the validation dataset.

Prediction error range and magnitude were calculated by taking the difference between the predicted and observed *removal age* on the validation dataset ([Fig. 2](#)). The GBRT model predicted within one year of the observed *removal age* for 49.01% ( $n = 521$ ) of the validation dataset and underestimated by one year or greater for 26.34% ( $n = 280$ ), meaning that the predicted *removal age* is less than the observed *removal age*. This model also overestimated by one year or greater for 24.65% ( $n = 262$ ). The largest prediction outlier was observed where a platform was predicted to be removed at 51.23 years of age but was destroyed during a blowout event at only 3.79 years of age. Comparatively, the ANN model predicted within one year of the observed *removal age* for 33.21% ( $n = 353$ ) of the validation dataset. In addition, the ANN model overestimated by one year or greater for 35.46% ( $n = 377$ ) and underestimated by one year or greater for 31.33% ( $n = 333$ ). Three prediction outliers were present outside of the standard deviation for this model, but no incident history of being destroyed was found for those records.

#### 3.2. Feature importance

Through RFE, the GBRT model yielded the optimal number of features to be 23, of which 4.35% were structural information, 86.95% were metocean statistics, 4.35% were related to external corrosion, and 4.35% were production statistics. The feature selected as most important to the GBRT model was the mean annual number of Category 1 hurricanes (C1) that hit a platform's location, with a 93.22% loss of explained variance when removed from the model. Noticeably, 60.87% of the selected features relate to tropical storms and hurricanes, including mean annual occurrences and duration (in days) of tropical storms and C1 through C4 hurricanes, and various lifespan statistics for maximum reported wind gusts and minimum central pressure during storm events.

For the ANN model, through RFE the optimal number of features was found to be 792. Of the features selected, 11.99% were structural information, 17.93% were metocean statistics, 6.57% were incident features, 3.03% related to corrosion or marine growth, and 60.48% were production statistics. Of these categories, corrosion or marine growth had the lowest percentage of original features remaining in the ANN model (75.00%), while 98.61% of metocean features, 91.24% of production statistics, 89.62% of structural features, and 78.79% of incident features were retained in the model. [Appendix D](#) provides further information on the top 10 features and their importance in both models.

#### 3.3. Existing platform removal age and RUL predictions

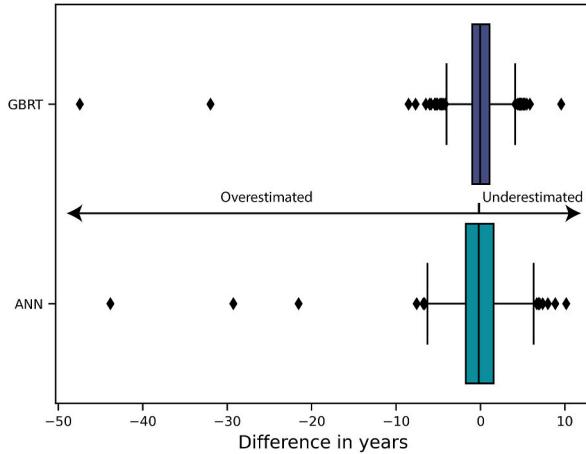
*Removal ages* of the 1,981 existing platforms were predicted to be between 1.05 – 53.67 years and 0 – 71.64 years for the GBRT and ANN models, respectively. As shown in [Fig. 3](#), the distribution of the GBRT models aligns most with the removed platforms data, as both are left skewed at ages under 20 years. In addition, the mean predicted *removal age* from the GBRT model predictions is within a month of the mean from the removed platforms data. The mean *removal age* from the ANN model was predicted as 14.34 years higher than the GBRT model and the removed platforms removal ages. Though not as evident, the distribution of the ANN model aligns more with the existing platform current age distribution, with a mean age difference of less than six months.

There was a considerable amount of variance between the model predictions with an average difference in *removal age* of 14.72 years and a standard deviation of 8.67 years. The current ages of the existing platforms ranged from six months to 73.05 years, 19.93 years wider than the GBRT model predictions. The GBRT model was not given a boundary condition for age prediction, whereas the ANN model, which is capable of extrapolation, was explicitly limited to the removal ages of the observed removed platforms. As a

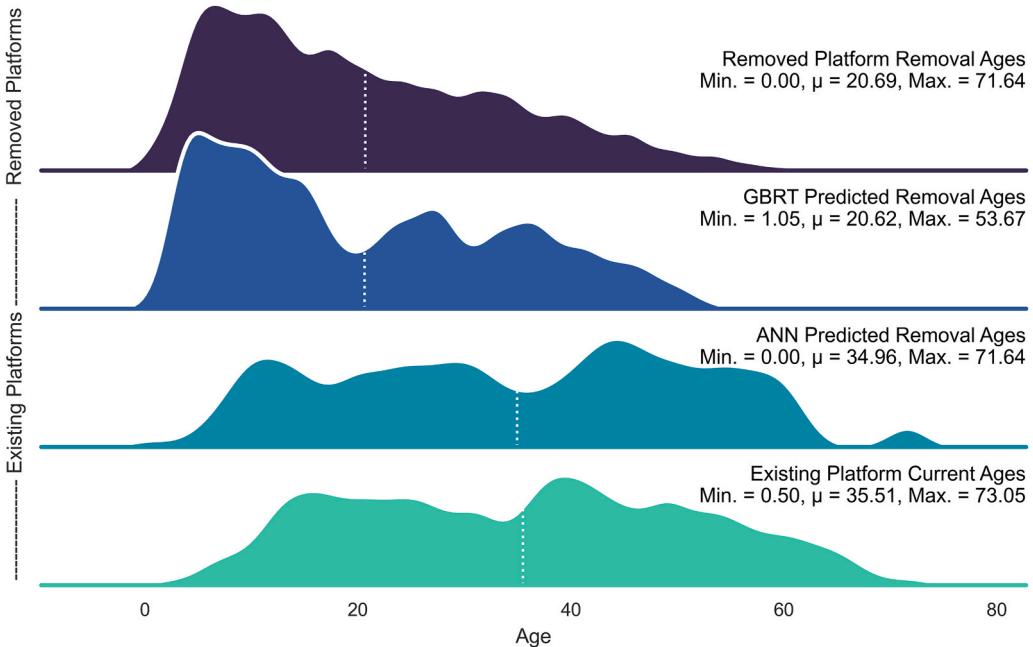
**Table 1**

Evaluation metric scores of model predictions on the validation set.

Model	$R^2$ Score	MAE	MSE	RMSE
GBRT	0.97	1.42	6.33	2.52
ANN	0.95	2.04	9.20	3.03



**Fig. 2.** Boxplot visualizing the distributions of the differences between the observed and predicted removal age in years for the GBRT (A) and ANN (B) models during validation. Negative values indicate overestimation and positive values indicate underestimation.



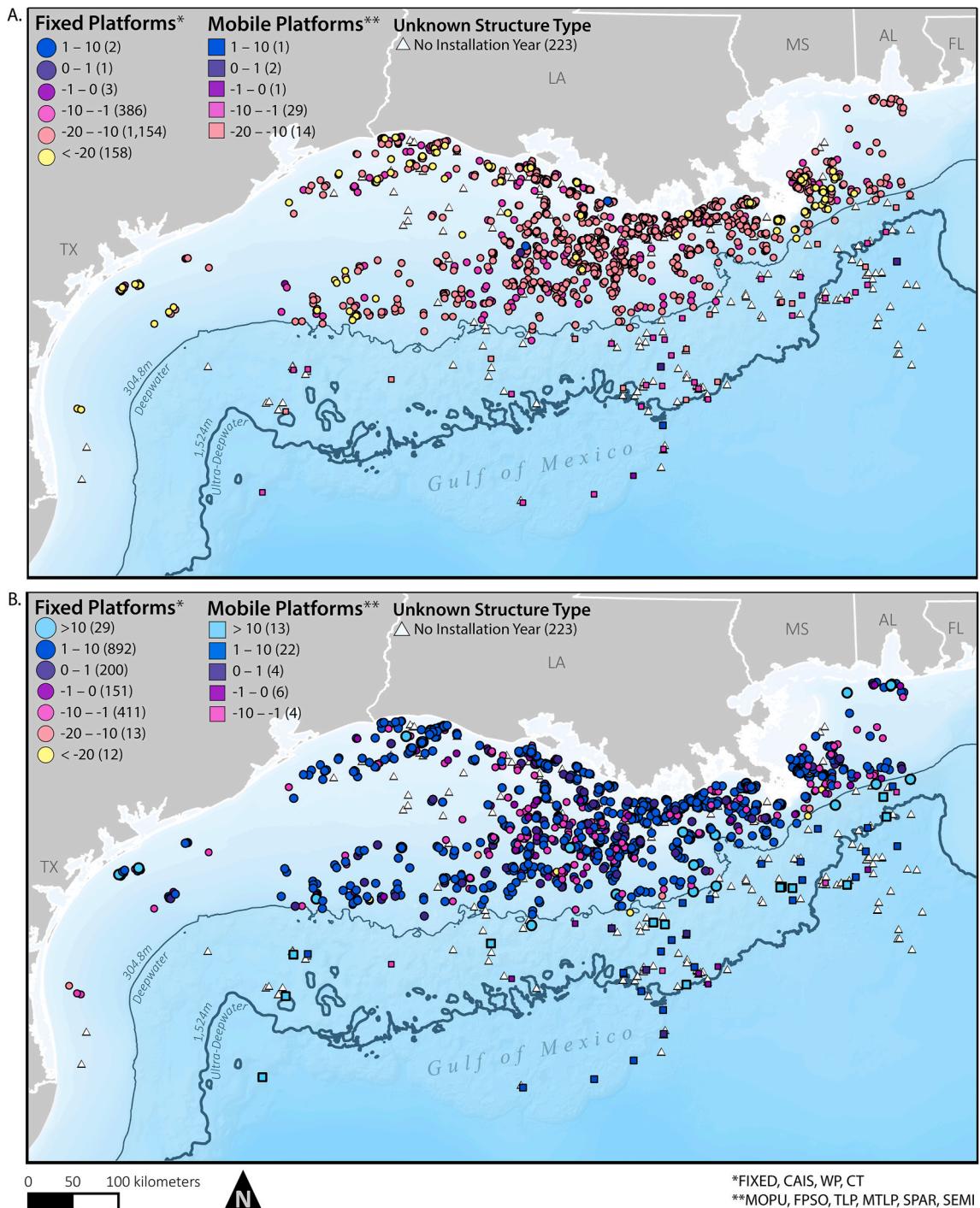
**Fig. 3.** Distributions of the age statistics in years on known removal ages of removed platforms, predicted removal ages on existing platforms from the GBRT and ANN models, and the current ages of existing platforms. White dotted lines indicate the mean of each distribution, and distributions are given as kernel densities for visualization purposes.

result, the range of the ANN predictions were identical to the removed platform removal ages, and about 11 months short of the range of the existing platform current ages.

RUL was calculated as the difference between the predicted *removal age* and the current age. Current age is the difference between the current date (set as Jan. 1, 2021) and the platform installation date. Negative RUL values indicate that the predicted removal age is less than the current age of the platform. RUL could not be calculated for 223 of the existing platform records, as they did not have an installation date. The GBRT model predicted RUL ranging from -30.88 to 2.71 years, with a mean of -13.35 years. Of the 1,757 existing platforms with installation dates, the GBRT model found that 99.66% ( $n = 1751$ ) of existing platform current ages were past the predicted *removal age*. From these predictions, of the six platforms with positive RUL, there were three fixed platforms and three mobile platforms (two tension leg platforms and one semisubmersible platform). Alternatively, the ANN model predicted RUL ranging from -72.04 to 66.53 years, with a mean of 1.57 years. The ANN model found that 34.32% ( $n = 603$ ) of existing platform current ages were past the predicted removal age, while a majority (65.68%) had positive RUL. Of the 1,152 platforms with positive RUL, 96.44% were labeled as fixed, caisson, compliant towers, and well protectors, while the remaining 3.56% were mobile platforms (tension leg, mini

tension leg, spar, semisubmersible, mobile offshore production units, or floating production storage and offloading platforms).

Fig. 4 shows the spatial distribution of the existing platforms symbolized by RUL (color) and classified structure type (shape). From the GBRT model, all six platforms predicted to have positive RUL are south of Louisiana, four are along the shelf in shallow waters, one is in deepwater, and one is in ultra-deepwater. From the ANN model, existing platforms with predicted positive RUL appear randomly distributed when compared to all existing platforms, with the majority (96.00%) along the shelf, 2.86% in deepwater, and 1.14% in ultra-deepwater, with the deepest depth at approximately 2,743.2 meters (9,000 feet).



**Fig. 4.** Spatial distribution of existing platforms symbolized by structure type and predicted RUL, calculated as predicted removal age minus current age for A) the GBRT model and B) the ANN model. Positive values indicate that a platform is predicted to have a remaining lifespan, whereas negative values indicate that a platform is predicted to be past the original design life.

#### 4. Discussion

A unique and extensive infrastructure dataset integrating data representing the natural and engineered offshore system [23] was assessed using supervised ML models to evaluate the relationships between factors known to affect structural integrity and the age at which a platform is removed. To aid in building a better understanding of the integrity of platforms in the federal GoM, this work applied two different ML models, yielding several results worth further discussion.

The accuracy of both models was found to be statistically significant, with average  $R^2$  scores of 0.97 and 0.95 for removed platforms by the GBRT and ANN models, respectively, during validation. When comparing these validation metrics to the training results, both models showed similar accuracy scores which suggests minimal overfitting. The high level of accuracy indicates that the models were able to identify relationships among the input features and *removal age*, as well as to operate effectively with outliers present in the training and validation datasets. Predictions from the GBRT model found that most platforms are past their RUL, which aligns with several past sources [25,26]. Although the ANN model resulted in most platforms having positive RUL, many platforms were predicted to be near the end of their life, as seen by the average RUL prediction being less than two years. While both models achieved a similar level of accuracy in evaluating *removal age* for the validation set, the difference in predicted RUL between the two models indicates a variation in the model algorithms and ability to generalize from the removed platform dataset to the existing platform dataset. Additionally, considering that RFE selected more than 34 times as many features for the ANN model than the GBRT model, differences in predicted RUL may depend on the magnitude and variety of features selected during RFE to predict platform *removal age*. Furthermore, the spatial distribution of existing and removed platforms (Figure A1.) varies. Generally speaking, older platforms tend to be more nearshore than newer platforms, which are being developed in deepwater and ultra-deepwater environments. These newer platforms were fabricated using more stringent design standards [40] to account for extreme conditions, such as increased deck height to account for higher wave heights, as statistics for extreme events become more reliable.

The top features for both models (Appendix D) revealed that the yearly number of hurricanes has a strong association with a platform's removal age, agreeing with several past studies [8,10,36]. The GBRT model, which selected 23 features, mostly relied on storm, wind, and wave features to predict platform RUL. This may show that the GBRT model is influenced by the inherent multicollinearity in the input data because the metocean variables have a high degree of correlation [75]. Additional factors can be traced back to structural properties, including water or oil injection flags, biological factors which affect external corrosion rates [9], and annual production rates. In contrast, the ANN model performed optimally with 792 features. The ANN model selected features representing almost all key stressors. Like the GBRT model, the ANN model found metocean to be important for predicting *removal age*, likely indicating that variation in metocean conditions at a platform's location can support RUL calculations. Unlike the GBRT model, the ANN model also utilized features representing past incidents and geohazards. This distinction could be attributed to the algorithmic differences between the models. One model does not necessarily select more suitable features than the other; rather, the method of predicting the target feature is algorithm-dependent with each model discovering different relationships between predictor and target variables while, in this case, still performing at the same approximate accuracy. So, while certain features may be unimportant in one model for predicting RUL, those features should not necessarily be disregarded. These differences demonstrate the strength of using a multiple model approach as additional relationships between input features and *removal age* were identified that might otherwise have gone unnoticed.

Whereas many platform risk assessments typically rely on in situ observational data, the methodology used in this study was successful at modeling general system reliability by utilizing spatial and temporal factors that are known to cause structural degradation over time. This analytical approach provides information on regional trends, which can be used to inform local insights as to the effect that potential degradational factors might have on platform longevity. With this understanding, additional maintenance for those platforms that are closer to or past their predicted lifespan could be made, especially if the risk of enduring a hurricane is higher. In general, the platforms that were identified as significantly beyond their predicted *removal age* might be at a higher risk of failure, which could result in oil spills, leaks, or blowout events.

Lastly, there are limitations of this work that future studies could address. The data driving this study relied on the recorded removal date to estimate the remaining lifespan and did not distinguish between the date of structure decommission and physical removal. The absence of this distinction can potentially introduce errors in this study (i.e., a platform is decommissioned and there is a significant lag time before actual removal). Additionally, this dataset does not contain records of maintenance, which might contribute to the extended lifespan of these platforms [18,120]. Another limitation arises with the transition from drilling in shallow, coastal waters to deepwater and ultra-deepwater, which requires the use of mobile platforms or rigs. The ML models are trained on the dataset of removed platforms, where mobile platforms are significantly under-represented, as the dataset is composed of 99.19% fixed platforms and 0.81% mobile platforms. Therefore, it is likely that the predictions from these models will perform better for fixed platforms, but there are not enough mobile records in the removed platform dataset to confirm this.

#### 5. Conclusion

Using big data and two supervised ML models, we were able to successfully forecast platform RUL in the U.S. federal GoM. Moreover, both the GBRT and ANN model provided evidence that environmental loading and extreme weather events are significant in predicting removal age, as measured by RUL [37,39]. To the best of our knowledge, this data-driven approach is the first of its kind, which means that there are no other results to compare model accuracy to. As an alternative, the application and results of the ML models were compared to provide discussion on the best practices for predicting offshore platform lifespan.

The variability of the modeled RUL of existing platforms demonstrates the benefits of taking a multiple model approach. This

analytical framework captured a possible distribution of life expectancy; the discrepancies should be further researched. What the models agreed on was that many platforms operating in the GoM today are past or close to their modeled design life, and therefore, the use and reuse of these platforms should be carefully considered along with an interpretation of the relative risk factors [3].

Future offshore infrastructure RUL prediction and modeling can be improved by incorporating additional information such as decommissioning and maintenance data coupled with improved variable selection methods to reduce multicollinearity in the input data. In addition, with access to more mobile platform data, the models developed for this study could more confidently forecast the integrity of platforms in deepwater and ultra-deepwater areas. Moreover, with the appropriate input data and hyperparameters, these models could be adapted to predict other target features, such as decommissioning date or ship collision risk. Furthermore, these models can be used to evaluate alternative offshore structures, such as pipelines and renewable energy infrastructure.

### Funding source statement

This manuscript was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference therein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed therein do not necessarily state or reflect those of the United States Government or any agency thereof.

### Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

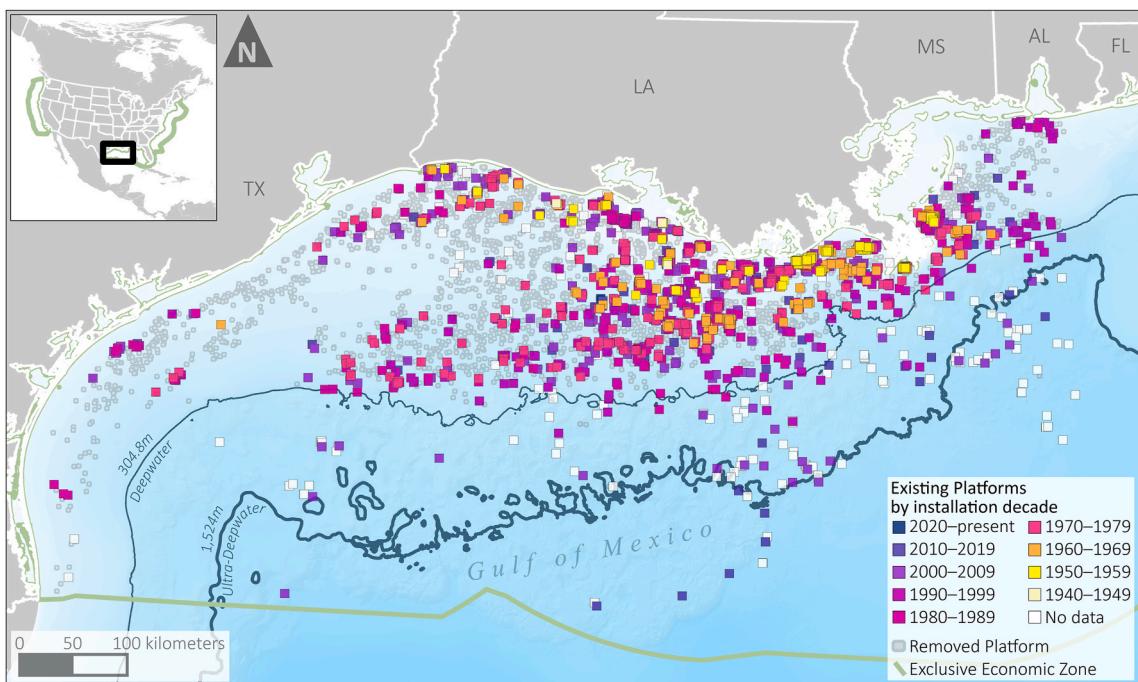
### Acknowledgements

The authors would like to express their appreciation to Michael Sabbatino for his valuable work with processing our dataset, critical for running our models. They would also like to thank Anuj Suhag for his advisement in machine learning expertise.

This work was performed in support of the U.S. Department of Energy's Fossil Energy Offshore Unconventional Resources – DE FE-1022409. This research was supported in part by an appointment to the U.S. Department of Energy (DOE) Postgraduate Research Program at the National Energy Technology Laboratory (NETL) administered by the Oak Ridge Institute for Science and Education (ORISE).

This project was funded by the U.S. Department of Energy, National Energy Technology Laboratory, in part, through a site support contract. Neither the United States Government nor any agency thereof, nor any of their employees, nor the support contractor, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

### Appendix A



**Fig. A.3.** Map of platforms symbolized by structural status (existing or removed) and installation decade for those currently existing in the northern GoM with demarcations for the Exclusive Economic Zone boundary and contour lines for deepwater and ultra-deepwater depths.

## Appendix B

**Table B.1**  
Summary table of the data and supporting information.

Data	Dataset coverage	Example Fields	Data Type	Date range	Source(s)
Platform data	5.6%	<i>Location information</i> Area code, block number, latitude, longitude, distance to shore, water depth, district code <i>Structure specifications</i> Structure type code, rig count, crane count, deck count, slant slot count, slot count, slant slot count, satellite completion count, underwater completion count, power source type, bed count <i>Structure flags</i> Major structure, abandon, fired vessel, gas production, gas flaring, manned 24 h, major complex, Liquid Automatic Custody Transfer (LACT) meter, heliport, workover, water production, tank gauge, sulfur production, store tank, production equipment, production, power generation equipment, oil production, gas sale meter, compressors, commingling production, meter prover <i>Other</i> Authority status, injection code	Numerical, categorical	1942–2020	[71]
Reported incidents	15.4%	<i>Counts</i> Total occurrences, fatalities, injuries, human-related, weather-related, structure-related <i>Counts of incident types</i> Collision, crane, damaged/disabled safety system, fire, explosion, fire-explosion category (minor, major, incidental, or catastrophic), loss of well control, major and minor property damage, H <sub>2</sub> S release, required muster, facility shut-in, gas release, + 145 more <i>Other</i> Damage value, platform destroyed, incident severity sum, incident severity sum minus incident count, sum of ages at incident, average age at incident, age at last incident	Numerical, categorical	1956–2018	[41]

(continued on next page)

**Table B.1** (continued)

Data	Dataset coverage	Example Fields	Data Type	Date range	Source(s)
Metocean conditions	17.8%	<i>Surface current speed (<math>u, v, magnitude</math>)</i> Min, max, mean, median, and 25th, 75th, and 90th percentile	Numerical	1/1/2003–4/30/2020	[77–80]
		<i>Wind speed (<math>u, v, magnitude</math>)</i> Min, max, mean, median, and 25th, 7th, and 90th percentile	Numerical	1/1/1979–5/31/2019	[76,81]
		<i>Wave height, period, direction, and power</i> Min, max, mean, median, and 25th, 7th, and 90th percentile			
		<i>Storms – tropical, category 1–5</i> Yearly max occurrences, yearly mean occurrences, yearly max days	Numerical	1/1/1842–12/31/2019	[82–84]
		<i>Storms – maximum sustained wind speed, maximum reported wind gust</i> Max, average, standard deviation, median, and 25th, 75th, and 90th percentile			
		<i>Storms – minimum central pressure</i> Min, max, average, standard deviation, median, and 25th, 75th, and 90th percentile			
		<i>Nitrate, oxygen, phosphorus, salinity, silicate, temperature</i> Mean and standard deviation at the surface,	Numerical	1/1/1900–12/31/2017	[85–92]
		Mean and standard deviation at 25% water depth,			
		Mean and standard deviation at 65% water depth			
		Well count per platform, spud and completion dates, well status, production details, annual production, injection pressure, depths	Numerical, categorical	1947–2020	[93]
Geohazard data	2.6%	Mean, minimum, and maximum subsurface pressure and temperature gradients	Numerical	N/A	[94–101]
		Mean, minimum, and maximum values representing geologic domain-based risk for seep anomalies, landslides, H2S occurrences, and plumes			

## Appendix C

The evaluation metrics on model performance for both the training and testing datasets of each of the five folds were averaged to get an overall score (Table C1). On average, the GBRT model explained 97.4% of the variance in *removal age* in the testing dataset. The amount of overfitting for each of the folds was low with an average absolute difference of the R<sup>2</sup> score between the training and testing set of 0.02. Similarly, the ANN model on average explained 95.1% of the variance in *removal age* in the testing dataset with an average absolute difference of the R<sup>2</sup> score between the training and testing set of 0.016.

Through the exhaustive grid search, the best set of hyperparameter values for the GBRT and ANN models were learned based on the training data and can be found in Tables C2 and C3. The ANN grid search found a single small hidden layer size of 21 neurons to perform best. Training over 400 iterations with a learning rate of 0.06 and dropout set to 0.48 resulted in the highest R<sup>2</sup> score on training folds during the grid search for the ANN model.

**Table C.1**

Evaluation metric scores on each of the 5 k-fold training and testing sets.

MODEL	Dataset	R <sup>2</sup> Score		MAE		MSE		RMSE	
		Avg.	Std. Dev.	Avg.	Std. Dev.	Avg.	Std. Dev.	Avg.	Std. Dev.
<b>GBRT</b>	Training	0.997	0.0003	0.523	0.020	0.512	0.044	0.715	0.030
	Testing	0.974	0.008	1.434	0.042	4.680	1.400	2.141	0.309
<b>ANN</b>	Training	0.967	0.003	1.90	0.868	6.042	0.560	2.456	0.113
	Testing	0.951	0.02	2.038	0.133	9.201	3.447	3.033	0.541

**Table C.2**

Values used in exhaustive grid search for the GBRT model and the optimal values based on the training data.

HYPERPARAMETER	Grid Search Values	Optimal Value
L1 regularization	0, 1, 3, 5	1
L2 regularization	0, 1, 3, 5	5
Learning rate	0.01, 0.05, 0.1, 0.3	0.05
Number of estimators	1, 5, 10, 20, 50, 100, 500, 1000	1000
Maximum depth	1, 2, 3, 4, 5, 6	6
Subsample	0.1, 0.3, 0.5, 1.0	0.3

**Table C.3**

Values used in exhaustive grid search for the ANN model and optimal values based on training data.

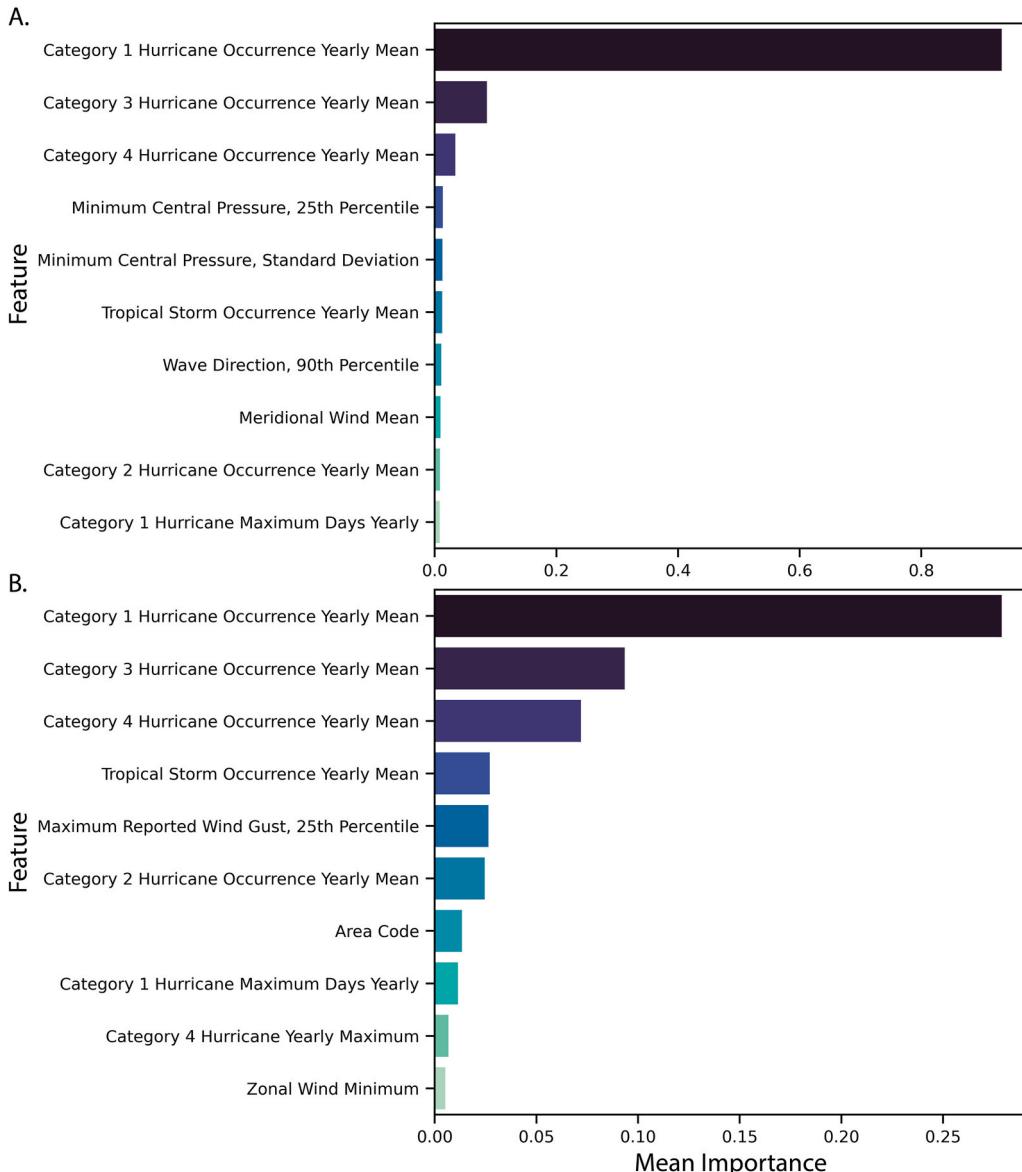
Hyperparameter	Grid Search Values	Optimal Value
----------------	--------------------	---------------

(continued on next page)

**Table C.3 (continued)**

Hyperparameter	Grid Search Values	Optimal Value
Hidden layer size	30, 26, 22, 21, 20, 19, 18, 17	21
Dropout	0.48, 0.5, 0.52	0.48
Learning rate	0.05, 0.06, 0.07	0.06
Number of training epochs	100, 200, 300, 400, 500	400

## Appendix D



**Fig. D.1.** Top 10 feature importance values using permutation feature importance (PFI) for the A) GBRT model, and B) ANN model. Importance mean represents the mean of all PFI values over the 5 k-folds during cross validation of the entire removed platform dataset.

## References

- [1] IHS. IHS Enerdeq® well header database. IHS; 2015.

- [2] Manning M. Offshore production nearly 30% of global crude oil output in. U.S. Energy Information Administration; 2015–16. 12 January 2021, <https://www.eia.gov/todayinenergy/detail.php?id=28492#>.
- [3] Kolian SR, Godec M, Sammarco PW. Alternate uses of retired oil and gas platforms in the Gulf of Mexico. *Ocean Coast Manag* 2019;167:52–9. <https://doi.org/10.1016/j.ocecoaman.2018.10.002>.
- [4] Beaubouef B. Gulf operators continue to advance platform life extension projects. *Offshore Magazine*; 2020 [accessed 11 May 2021], <https://www.offshore-mag.com/production/article/14181754/gulf-operators-continue-to-advance-platform-life-extension-projects#:~:text=Home.,Gulf%20operators%20continue%20to%20advantage%20platform%20life%20extension%20projects,structures%20and%20two%20compliant%20towers>.
- [5] Leporini M, Marchetti B, Corvaro F, Polonara F. Reconversion of offshore oil and gas platforms into renewable energy sites production: assessment of different scenarios. *Renew Energy* 2019;135:1121–32. <https://doi.org/10.1016/j.renene.2018.12.073>.
- [6] Minerals Management Service. Programmatic environmental impact statement for alternative energy development and production and alternate use of facilities on the Outer Continental Shelf (2007). U.S. Dept. of the Interior, Minerals Management Service; 2007.
- [7] ABS Consulting. ABS Consulting project No.: 1200044: Hurricane Lili's impact on fixed platforms. MMS order No: 0103P072325. ABS Consulting; 2004. San Antonio, TX.
- [8] Cruz AM, Krausmann E. Damage to offshore oil and gas facilities following hurricanes Katrina and Rita: an overview. *J Loss Prevent Proc* 2008;21:620–6. <https://doi.org/10.1016/j.jlp.2008.04.008>.
- [9] Bhandari J, Khan F, Abbassi R, Garaniya V, Ojeda R. Modelling of pitting corrosion in marine and offshore steel structures – a technical review. *J Loss Prevent Proc* 2015;37:39–62. <https://doi.org/10.1016/j.jlp.2015.06.008>.
- [10] Puskar F, Spong R, Ku A, Gilbert R, Choi Y. In: Assessment of fixed offshore platform performance in Hurricane Ivan. Houston, Texas, USA: Offshore Technology Conference; 2006. <https://doi.org/10.4043/18325-MS>.
- [11] Bureau of Safety and Environmental Enforcement. Incident archive - Taylor Energy (Mississippi canyon) oil spill. Bureau of Safety and Environmental Enforcement; 2015. Nov. 5th, 2020, <https://www.bsee.gov/newsroom/library/incident-archive/taylor-energy-mississippi-canyon/ongoing-response-efforts>.
- [12] Sun S, Hu C, Garcia-Pineda O, Kourafalou V, Le Hénaff M, Androulidakis Y. Remote sensing assessment of oil spills near a damaged platform in the Gulf of Mexico. *Mar Pollut Bull* 2018;136:141–51. <https://doi.org/10.1016/j.marpolbul.2018.09.004>.
- [13] Mason AL, Taylor JC, MacDonald IR. An integrated assessment of oil and gas release into the marine environment at the former Taylor energy MC20 site. NOAA Technical Memorandum NOS NCCOS 260. Silver Spring, MD; 2019. p. 147.
- [14] Dhanak MR, Xiros N I. *Springer handbook of ocean engineering*. Cham: Springer; 2016.
- [15] Close F, McCavitt RD, Smith B. Deepwater Gulf of Mexico development challenges overview. In: SPE North Africa Technical Conference & Exhibition. Marrakech, Morocco: Society of Petroleum Engineers; 2008. <https://doi.org/10.2118/113011-MS>.
- [16] Shaughnessy JM, Romo LA, Soza RL. Problems of ultra-deep high-temperature, high-pressure drilling. In: SPE Annual Technical Conference and Exhibition. Denver, Colorado, USA: Society of Petroleum Engineers; 2003. <https://doi.org/10.2118/84555-MS>.
- [17] Aeraan A, Siriwardane SC, Mikkelsen O, Langen I. A framework to assess structural integrity of ageing offshore jacket structures for life extension. *Mar Struct* 2017;56:237–59. <https://doi.org/10.1016/j.marstruc.2017.08.002>.
- [18] Solland G, Sigurdsson G, Ghosal A. Life extension and assessment of existing offshore structures. In: SPE Project and Facilities Challenges Conference at METS. Doha, Qatar: Society of Petroleum Engineers; 2011. <https://doi.org/10.2118/142858-MS>.
- [19] Guédé F. Risk-based structural integrity management for offshore jacket platforms. *Mar Struct* 2019;63:444–61. <https://doi.org/10.1016/j.marstruc.2018.04.004>.
- [20] Tygesen UT, Worden K, Rogers T, Manson G, Cross EJ. State-of-the-art and future directions for predictive modelling of offshore structure dynamics using machine learning. In: Pakzad S, editor. *Dynamics of civil structures*. Conference Proceedings of the Society for Experimental Mechanics Series2. Springer, Cham; 2019. p. 223–33. [https://doi.org/10.1007/978-3-319-74421-6\\_30](https://doi.org/10.1007/978-3-319-74421-6_30).
- [21] Wen K, He L, Liu J, Gong J. An optimization of artificial neural network modeling methodology for the reliability assessment of corroding natural gas pipelines. *J Loss Prevent Proc* 2019;60:1–8. <https://doi.org/10.1016/j.jlp.2019.03.010>.
- [22] Sharma P, Knezevic D, Huynh P, Malinowski G. In: RB-FEA based digital twin for structural integrity assessment of offshore structures. Houston, Texas, USA: Offshore Technology Conference; 2018. <https://doi.org/10.4043/29005-MS>.
- [23] Romeo L, Dyer A, Wenzlick M, Duran R, Nelson J, Sabbatino M, et al. Comprehensive GOM federal waters platform, incident, metocean, and geohazard dataset. *Energy Data Exchange*; 2021. <https://doi.org/10.18141/1779221>.
- [24] Ersdal G, Selnes PO. Life extension of aging petroleum production facilities offshore. In: Rio de Janeiro, Brazil: SPE International Conference on Health, Safety and Environment in Oil and Gas Exploration and Production; 2010. <https://doi.org/10.2118/127134-MS>.
- [25] Animah I, Shafee M. Condition assessment, remaining useful life prediction and life extension decision making for offshore oil and gas assets. *J Loss Prevent Proc* 2018;53:17–28. <https://doi.org/10.1016/j.jlp.2017.04.030>.
- [26] Smyth L. Using an offshore platform beyond its expected lifespan. *Engineer Live*; 2018 [accessed, <https://www.engineerlive.com/content/using-offshore-platform-beyond-its-expected-lifespan>. Accessed 11 May 2021].
- [27] Stacey A, Birkinshaw M, Sharp J. Life extension issues for ageing offshore installations. In: Proceedings of the ASME 2008 27th International Conference on Offshore Mechanics and Arctic Engineering. Volume 5: Materials Technology; CFD and VIV. Estoril, Portugal: ASME; 2008. p. 199–215. <https://doi.org/10.1115/OMAE2008-57411>.
- [28] Lu P, Liu H, Serratella C, Wang X. Assessment of data-driven, machine learning techniques for machinery prognostics of offshore assets. In: Houston, Texas, USA: Offshore Technology Conference; 2017. <https://doi.org/10.4043/27577-MS>.
- [29] Moan T. Life cycle structural integrity management of offshore structures. *Struct Infrastruct E* 2018;14:911–27. <https://doi.org/10.1080/15732479.2018.1438478>.
- [30] Ferreira NN, Martins MR, de Figueiredo MAG, Gagnon VH. Guidelines for life extension process management in oil and gas facilities. *J Loss Prevent Proc* 2020; 68:104290. <https://doi.org/10.1016/j.jlp.2020.104290>.
- [31] Al-Fozan SA, Malik AU. Effect of seawater level on corrosion behavior of different alloys. *Desalination* 2008;228:61–7. <https://doi.org/10.1016/j.desal.2007.08.007>.
- [32] Guedes Soares C, Garbatov Y, Zayed A. Effect of environmental factors on steel plate corrosion under marine immersion conditions. *Corrosion Eng Sci Technol* 2011;46:524–41. <https://doi.org/10.1179/147842209X12559428167841>.
- [33] Melchers RE. A review of trends for corrosion loss and pit depth in longer-term exposures. *Corros Mater Degrad* 2018;1:42–58. <https://doi.org/10.3390/cmd101004>.
- [34] Revie WR. *Uhlig's corrosion handbook*. 3rd ed. Volume 51. Wiley; 2011.
- [35] Nunez M. *Prevention of metal corrosion: new research*. Nova Publishers; 2007.
- [36] Kaiser MJ. The impact of extreme weather on offshore production in the Gulf of Mexico. *Appl Math Model* 2008;32:1996–2018. <https://doi.org/10.1016/j.apm.2007.06.031>.
- [37] Holmes R. The fatigue behaviour of welded joints under north sea environmental and random loading conditions. In: Houston, Texas, USA: Offshore Technology Conference; 1980. <https://doi.org/10.4043/3700-MS>.
- [38] Bruserud K, Haver S, Myrhaug D. Joint description of waves and currents applied in a simplified load case. *Mar Struct* 2018;58:416–33. <https://doi.org/10.1016/j.marstruc.2017.12.010>.
- [39] Bea RG, Ramos R, Valle O, Valdes V, Maya R. Risk assessment & management based hurricane wave criteria for design and requalification of platforms in the Bay of Campeche. In: Houston, Texas, USA: Offshore Technology Conference; 1998. <https://doi.org/10.4043/8692-MS>.
- [40] API RP2-SIM. *Structural Integrity Management of fixed offshore structures*. American Petroleum Institute; 2014.
- [41] Bureau of Safety and Environmental Enforcement. Offshore incident statistics. BSEE; 2020. <https://www.bsee.gov/stats-facts/offshore-incident-statistics>.

- [42] Sharp J, Ersdal G, Galbraith D. Meaningful and leading structural integrity KPIs. SPE offshore Europe conference and exhibition. In: Aberdeen, Scotland, UK: SPE Offshore Europe Conference and Exhibition; 2015. <https://doi.org/10.2118/175519-MS>.
- [43] Muehlenbachs L, Cohen MA, Gerarden T. The impact of water depth on safety and environmental performance in offshore oil and gas production. Energy Pol 2013;55:699–705. <https://doi.org/10.1016/j.enpol.2012.12.074>.
- [44] Nezamian A, Nicolson RJ, Iosif D. State of art in life extension of existing offshore structures. In: Proceedings of the ASME 2012 31st International Conference on Ocean, Offshore and Arctic Engineering, Volume 2: Structures, Safety, and Reliability. Rio de Janeiro, Brazil: American Society of Mechanical Engineers; 2013. p. 165–74. <https://doi.org/10.1115/OMAE2012-83302>.
- [45] Sharp JV, Ersdal G, Galbraith D. Development of key performance indicators for offshore structural integrity. Proceedings of the ASME 2008 27th International Conference on Offshore Mechanics and Arctic Engineering, Volume 5: Materials Technology; CFD and VIV. Estoril, Portugal: American Society of Mechanical Engineers Digital Collection; 2009. p. 123–30. <https://doi.org/10.1115/OMAE2008-57203>.
- [46] Kaiser MJ, Pulsipher AG. Loss categories, hazard types in marine operations describe. Oil Gas J 2007;105:39–44.
- [47] Lacasse S, Nadine F, Vanneste M, L'Heureux JS, Forsberg CF, Kvalstad TJ. Case studies of offshore slope stability. Stability and Performance of Slopes and Embankments III. Geo-Congress 2013; 2013. p. 2369–408. <https://doi.org/10.1061/9780784412787.228>.
- [48] Skogdalen JE, Utne IB, Vinmenn JE. Developing safety indicators for preventing offshore oil and gas deepwater drilling blowouts. Saf Sci 2011;49:1187–99. <https://doi.org/10.1016/j.ssci.2011.03.012>.
- [49] Nguyen TH, Bae W, Hoang NT. Effect of high pressure high temperature condition on well design development in offshore Vietnam. In: Kuala Lumpur, Malaysia: Offshore Technology Conference Asia; 2016. <https://doi.org/10.4043/26374-MS>.
- [50] Jiang T, Gradus JL, Rosellini AJ. Supervised machine learning: a brief primer. Behav Ther 2020;51:675–87. <https://doi.org/10.1016/j.beth.2020.05.002>.
- [51] Jordan MI, Mitchell TM. Machine learning: trends, perspectives, and prospects. Science 2015;349:255–60. <https://doi.org/10.1126/science.aaa8415>.
- [52] Friedman JH. Greedy function approximation: a gradient boosting machine. Ann Stat 2001;29:1189–232.
- [53] Yin X, Zhao X. Big data driven multi-objective predictions for offshore wind farm based on machine learning algorithms. Energy 2019;186:115704. <https://doi.org/10.1016/j.energy.2019.07.034>.
- [54] Gu J, Zhang H, Chen L, Lian S. The application of the big data algorithm for pipeline lifetime analysis. 2019 Chinese Automation Congress (CAC); 2019. p. 824–9. <https://doi.org/10.1109/CAC48633.2019.8996228>.
- [55] Pathy A, Meher S, B P. Predicting algal biochar yield using eXtreme Gradient Boosting (XGB) algorithm of machine learning methods. Algal Res 2020;50: 102006. <https://doi.org/10.1016/j.algal.2020.102006>.
- [56] Breiman L, Friedman J, Stone CJ, Olshen RA. Classification and regression trees. CRC press; 1984.
- [57] Elith J, Leathwick JR, Hastie T. A working guide to boosted regression trees. J Anim Ecol 2008;77:802–13. <https://doi.org/10.1111/j.1365-2656.2008.01390.x>.
- [58] Friedman JH. Stochastic gradient boosting. Comput Stat Data Anal 2002;38:367–78. [https://doi.org/10.1016/S0167-9473\(01\)00065-2](https://doi.org/10.1016/S0167-9473(01)00065-2).
- [59] Zou J, Han Y, So S-S. Overview of artificial neural networks. Methods in Molecular BiologyTM. Humana Press; 2008.
- [60] Leshno M, Lin VY, Pinkus A, Schocken S. Multilayer feedforward networks with a nonpolynomial activation function can approximate any function. Neural Networks 1993;6:861–7. [https://doi.org/10.1016/S0893-6080\(05\)80131-5](https://doi.org/10.1016/S0893-6080(05)80131-5).
- [61] Sug H. The effect of training set size for the performance of neural networks of classification. WSEAS Trans Comput 2010;9:1297–306. <https://doi.org/10.5555/1973282.1973288>.
- [62] Gilpin LH, Bau D, Yuan BZ, Bajwa A, Specter M, Kagal L. Explaining explanations: an overview of interpretability of machine learning. In: Turin, Italy: 2018 IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA); 2018. p. 80–9. <https://doi.org/10.1109/DSAA.2018.00018>.
- [63] Shi J, Zhu Y, Kong D, Khan F, Li J, Chen G. Stochastic analysis of explosion risk for ultra-deep-water semi-submersible offshore platforms. Ocean Eng 2019;172: 844–56. <https://doi.org/10.1016/j.oceaneng.2018.12.045>.
- [64] Sidarta DE, O'Sullivan J, Lim H-J. Damage detection of offshore platform mooring line using artificial neural network. 51203. In: ASME 2018 37th International Conference on Ocean, Offshore and Arctic Engineering; 2018, V001T01A058. <https://doi.org/10.1115/OMAE2018-77084>.
- [65] Bhowmik S. Life extension of offshore structure using machine learning. In: Rio de Janeiro, Brazil: Offshore Technology Conference Brasil; 2019. <https://doi.org/10.4043/29759-MS>.
- [66] El-Abbasy MS, Senouci A, Zayed T, Mirahadi F, Parvizsedghy L. Artificial neural network models for predicting condition of offshore oil and gas pipelines. Autom ConStruct 2014;45:50–65. <https://doi.org/10.1016/j.autcon.2014.05.003>.
- [67] de Pina AA, Monteiro BdF, Albrecht CH, de Lima BSLP, Jacob BP. Artificial neural networks for the analysis of spread-mooring configurations for floating production systems. Appl Ocean Res 2016;59:254–64. <https://doi.org/10.1016/j.apor.2016.06.010>.
- [68] Zhang D, Qian L, Mao B, Huang C, Huang B, Si Y. A data-driven design for fault detection of wind turbines using random forests and XGBoost. IEEE Access 2018;6:21020–31. <https://doi.org/10.1109/ACCESS.2018.2818678>.
- [69] Chen T, Guestrin C. XGBoost: scalable tree boosting system. In: 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; 2016. p. 785–94. <https://doi.org/10.1145/2939672.2939785>.
- [70] Ke G, Meng Q, Finley T, Wang T, Chen W, Ma W, et al. LightGBM: a highly efficient gradient boosting decision tree. Adv Neural Inf Process Syst 2017;30: 3146–54.
- [71] Bureau of Safety and Environmental Enforcement. Platform/rig information. BOEM Data Center; 2020. <https://www.data.boem.gov/Main/Platform.aspx>.
- [72] Bureau of Safety and Environmental Enforcement. 2016a national assessment of undiscovered oil and gas resources of the U.S. Outer Continental Shelf. OCS Report. U.S. Department of the Interior, Bureau of Ocean Energy Management; 2017.
- [73] Pastor J, Liu Y. Hydrokinetic energy: overview and it's renewable energy potential for the Gulf of Mexico. In: 2012 IEEE Green Technologies Conference; 2012. p. 1–3. <https://doi.org/10.1109/GREEN.2012.6200995>.
- [74] Eccles JK, Pratson L. Economic evaluation of offshore storage potential in the US Exclusive Economic Zone. Greenh Gases 2013;3:84–95. <https://doi.org/10.1002/ghg.1308>.
- [75] Nelson J, Dyer A, Romeo L, Wenzlick M, Zaegle D, Duran R, et al. Evaluating offshore infrastructure integrity. U.S. Department of Energy - National Energy Technology Laboratory: Albany OR; 2021. p. 70. <https://doi.org/10.2172/1780656>.
- [76] The WaveWATCH III R Development Group (WW3DG). User manual and system documentation of WAVEWATCH III R version 6.07. College Park, MD, USA: NOAA/NWS/NCEP/MMAB; 2019. p. 465.
- [77] Cummings JA. Operational multivariate ocean data assimilation. Q J Roy Meteorol Soc 2005;131:3583–604. <https://doi.org/10.1256/qj.05.105>.
- [78] Cummings JA, Smedstad OM. Variational data assimilation for the global ocean. In: Park SK, Xu L, editors. Data assimilation for atmospheric, oceanic and hydrologic applications (vol II). Berlin, Heidelberg: Springer; 2013. p. 303–43.
- [79] Helber RW, Townsend TL, Barron CN, Dastugue JM, Carnes MR. Validation test report for the Improved Synthetic Ocean Profile (ISOP) system, Part I: synthetic profile methods and algorithm. NRL Memo. Naval Research Laboratory; 2013.
- [80] Hogan T, Liu M, Ridout J, Peng M, Whitcomb T, Ruston B, et al. The navy global environmental model. Oceanography 2014;27:116–25. <https://doi.org/10.5670/oceanog.2014.73>.
- [81] Kalnay E, Kanamitsu M, Kistler R, Collins W, Deaven D, Gandin L, et al. The NCEP/NCAR 40-year reanalysis project. Bull Am Meteorol Soc 1996;77:437–72. [https://doi.org/10.1175/1520-0477\(1996\)077<437:TNYRP>2.0.CO;2](https://doi.org/10.1175/1520-0477(1996)077<437:TNYRP>2.0.CO;2).
- [82] Knapp KR, Kruck MC, Levinson DH, Diamond HJ, Neumann CJ. The international best track archive for climate stewardship (IBTrACS) unifying tropical cyclone data. Bull Am Meteorol Soc 2010;91:363–76. <https://doi.org/10.1175/2009BAMS2755.1>.
- [83] Kruck MC, Knapp KR, Levinson DH. A technique for combining global tropical cyclone best track data. J Atmos Ocean Technol 2010;27:680–92. <https://doi.org/10.1175/2009JTECHA1267.1>.
- [84] Levinson DH, Diamond HJ, Knapp KR, Kruck MC, Gibney EJ. Toward a homogenous global tropical cyclone best-track dataset. Bull Am Meteorol Soc 2010;91: 377–80.

- [85] Boyer T, Garcia H, Locarnini R, Zweng M, Mishonov A, Reagan J, et al. World ocean atlas 2018. Dissolved oxygen, apparent oxygen utilization, and dissolved oxygen saturation, Volume 3. NOAA National Centers for Environmental Information; 2018.
- [86] Garcia H, Weathers K, Paver C, Smolyar I, Boyer T, Locarnini M, et al. World Ocean Atlas 2018. Vol. 4: Dissolved Inorganic Nutrients (phosphate, nitrate and nitrate+ nitrite, silicate). In: NOAA Atlas NESDIS. A. Mishonov Technical Editor; 2019. p. 35. <https://archimer.ifremer.fr/doc/00651/76336/>.
- [87] Locarnini M, Mishonov A, Baranova O, Boyer T, Zweng M, Garcia H, et al. World Ocean Atlas 2018, Volume 1: Temperature. In: NOAA Atlas NESDIS. A. Mishonov Technical Editor; 2018. p. 52. <https://archimer.ifremer.fr/doc/00651/76338/>.
- [88] Zweng M, Seidov D, Boyer T, Locarnini M, Garcia H, Mishonov A, et al. World ocean atlas 2018, volume 2: salinity. A. Mishonov technical editor. NOAA Atlas NESDIS vol. 82. 2019:50 pp.
- [89] Boyer T, Garcia H, Locarnini R, Zweng M, Mishonov A, Reagan J, et al. World ocean atlas 2018. In: Temperature. NOAA National Centers for Environmental Information; 2018.
- [90] Boyer T, Garcia H, Locarnini R, Zweng M, Mishonov A, Reagan J, et al. In: World ocean atlas 2018 Salinityume 2. NOAA National Centers for Environmental Information; 2018.
- [91] Boyer T, Garcia H, Locarnini R, Zweng M, Mishonov A, Reagan J, et al. World Ocean Atlas 2018. Vol. 4: Dissolved Inorganic Nutrients (phosphate, nitrate and nitrate+ nitrite, silicate). NOAA National Centers for Environmental Information; 2018.
- [92] Garcia H, Weathers K, Paver C, Smolyar I, Boyer T, Locarnini M, et al. World Ocean Atlas 2018, Volume 3: Dissolved Oxygen, Apparent Oxygen Utilization, and Dissolved Oxygen Saturation. In: NOAA Atlas NESDIS. A. Mishonov Technical Editor; 2019. p. 38. <https://archimer.ifremer.fr/doc/00651/76337/>.
- [93] Enverus. Drillinginfo. 2020. Dec. 9th, 2020, <https://www.enverus.com/>.
- [94] United States Geological Survey. Faults in the Gulf coast. Data.gov; 2004.
- [95] Eikrem V, Li R, Medeiros M, McKee BJ, Boswell BL, Shumilak EE, et al. SS: Perdido development project: great White WM12 Reservoir and Silvertip M. Frio Field development plans and comparison of recent well results with pre-drill models. In: Houston, Texas, USA: Offshore Technology Conference; 2010. <https://doi.org/10.4043/20879-MS>.
- [96] McAdoo BG, Pratson LF, Orange DL. Submarine landslide geomorphology, US continental slope. *Mar Geol* 2000;169:103–36.
- [97] Sawyer DE, Mason RA, Cook AE, Portnov A. Submarine landslides induce massive waves in subsea brine pools. *Sci Rep* 2019;9:1–9. <https://doi.org/10.1038/s41598-018-36781-7>.
- [98] Twichell D. MASSWASTING.SHP - mass-wasting deposits within the GLORIA survey area. In: Gulf of Mexico: open-file report 2005-1071, U.S. Geological survey, coastal and marine geology Program. MA: Woods Hole Science Center, Woods Hole; 2005.
- [99] Nodine MC, Cheon JY, Wright SG, Gilbert RB. Mudslides during Hurricane Ivan and an assessment of the potential for future mudslides in the Gulf of Mexico. In: Mineral Management Service Project 552. Offshore Technology Research Center; 2007.
- [100] Bureau of Ocean and Energy Management. Seismic water bottom anomalies - Gulf of Mexico. BOEM; 2019. <https://www.boem.gov/oil-gas-energy/mapping-and-data/map-gallery/seismic-water-bottom-anomalies-map-gallery>.
- [101] Rose K, Bauer JR, Mark-Moser M. A systematic, science-driven approach for predicting subsurface properties. *Interpretation* 2020;8:T167–81.
- [102] Garavaglia S, Sharma A. A smart guide to dummy variables: four applications and a macro. Proceedings of the Northeast SAS Users Group Conference; 1998.
- [103] Batista GEA, Monard MC. An analysis of four missing data treatment methods for supervised learning. *Appl Artif Intell* 2003;17:519–33. <https://doi.org/10.1080/713827181>.
- [104] Lakshminarayana K, Harp SA, Samad T. Imputation of missing data in industrial databases. *Appl Intell* 1999;11:259–75. <https://doi.org/10.1023/A:1008334909089>.
- [105] Cover T, Hart P. Nearest neighbor pattern classification. *IEEE Trans Inf Theor* 1967;13:21–7. <https://doi.org/10.1109/TIT.1967.1053964>.
- [106] Triguero I, García-Gil D, Maillo J, Luengo J, García S, Herrera F. Transforming big data into smart data: an insight on the use of the k-nearest neighbors algorithm to obtain quality data. *WIREs Data Min Knowl* 2019;9:e1289. <https://doi.org/10.1002/widm.1289>.
- [107] Bengio Y, Grandvalet Y. No unbiased estimator of the variance of k-fold cross-validation. *J Mach Learn Res* 2004;5:1089–105. <https://doi.org/10.5555/1005332.1044695>.
- [108] Guyon I, Weston J, Barnhill S, Vapnik V. Gene selection for cancer classification using support vector machines. *Mach Learn* 2002;46:389–422. <https://doi.org/10.1023/A:1012487302797>.
- [109] Wang B, Gong NZ. Stealing hyperparameters in machine learning. In: 2018 IEEE Symposium on Security and Privacy; 2018. p. 36–52. <https://doi.org/10.1109/SP.2018.00038>.
- [110] Feurer M, Hutter F. Hyperparameter optimization. In: Hutter F, Kotthoff L, Vanschoren J, editors. *Automated machine learning*. Cham: Springer; 2019. p. 3–33.
- [111] Paszke A, Gross S, Massa F, Lerer A, Bradbury J, Chanan G, et al. Pytorch: An imperative style, high-performance deep learning library. *Adv Neur In* 2019;32: 8026–37.
- [112] Borchani H, Varando G, Bielza C, Larrañaga P. A survey on multi-output regression. *WIREs Data Min Knowl* 2015;5:216–33. <https://doi.org/10.1002/widm.1157>.
- [113] Natekin A, Knoll A. Gradient boosting machines, a tutorial. *Front Neurorob* 2013;7. <https://doi.org/10.3389/fnbot.2013.00021>.
- [114] Chai T, Draxler RR. Root mean square error (RMSE) or mean absolute error (MAE)? – arguments against avoiding RMSE in the literature. *Geosci Model Dev* 2014;7:1247–50. <https://doi.org/10.5194/gmd-7-1247-2014>.
- [115] Botchkarev A. Performance metrics (error measures) in machine learning regression, forecasting and prognostics: properties and typology. *Interdiscipl J Inf Knowl Manag* 2019;14. <https://doi.org/10.28945/4184>.
- [116] Armstrong JS, Collopy F. Error measures for generalizing about forecasting methods: empirical comparisons. *Int J Forecast* 1992;8:69–80. [https://doi.org/10.1016/0169-2070\(92\)90008-W](https://doi.org/10.1016/0169-2070(92)90008-W).
- [117] Makridakis S. Accuracy measures: theoretical and practical concerns. *Int J Forecast* 1993;9:527–9. [https://doi.org/10.1016/0169-2070\(93\)90079-3](https://doi.org/10.1016/0169-2070(93)90079-3).
- [118] Li J. Assessing the accuracy of predictive models for numerical data: not r nor r<sup>2</sup>, why not? Then what? *PLoS One* 2017;12(8):e0183250. <https://doi.org/10.1371/journal.pone.0183250>.
- [119] Breiman L. Random forests. *Mach Learn* 2001;45:5–32. <https://doi.org/10.1023/A:1010933404324>.
- [120] O'Connor PE, Bucknell JR, DeFrance SJ, Westlake HS, Puskar FJ. Structural integrity management (SIM) of offshore facilities. In: Houston, Texas, USA: Offshore Technology Conference; 2005. <https://doi.org/10.4043/17545-MS>.