Remaining useful life estimation of HMPE rope during CBOS testing through machine learning

Shaun Falconer a,∗, Ellen Nordgård-Hansen b, Geir Grasmo a

a Department of Engineering Sciences, University of Agder, Grimstad 4876, Norway
b NORCE Norwegian Research Centre AS, Grimstad 4876, Norway

1. Introduction

Fibre ropes are increasingly used for lifting operations, however there are still issues related to the implementation. The material advantages of fibre rope are well documented (Foster, 2002; Faria et al., 2017), as well as issues inhibiting their immediate implementation connected to creep, thermal response during cyclic-bend-over-sheave and lack of available data regarding their implementation in offshore construction cranes. To be able to exploit this potential fully, more advanced maintenance routines must be established to challenge the status quo of manual inspection established by DNVGL (2017). There is significant potential and benefits with regards to avoiding premature retirement and reducing the chance of failure during operation through the development of intelligent maintenance methods for fibre ropes. Automation of manual processes and structured data-driven approaches to quantify historical health data, damage progression and physical measurements can lead to more informed decisions regarding rope condition and remaining useful life (RUL) through more frequent documented state observation. Establishing and verifying these methods would also signal a shift from time-based maintenance and reactive maintenance strategies to condition-based and predictive maintenance approaches. Positive implications of this include: preventing failure of rope during operation from an undetected fault; decreasing operation downtime for routine inspections; and avoiding retiring ropes with substantial remaining useful life.

Other sectors have implemented approaches for RUL estimation based on machine learning, with several reviews available detailing specific implementations (Nguyen et al., 2019; Fink et al., 2020; Lei et al., 2018; Diez-Olivan et al., 2019; Sutharssan et al., 2015; Khan and Yairi, 2018). However, this study will focus on the use of data-driven approaches through machine learning applied to fibre rope condition monitoring data for RUL prediction from cyclic-bend-over-sheave (CBOS) testing.

While the application of machine learning for prognostics and health management in fibre rope condition monitoring is not as established in the publicly accessible research domain, these applications have seen successes in other fields. Much like other engineering components, fibre ropes have damage mechanisms and physical changes that can be detected by sensors, which can be used in machine learning approaches. Offshore fibre rope use in particular is concerned with RUL related to time, tension and temperature as advocated by industry standards by DNVGL (2015) and ABS (2011). Previous studies into rope behaviour detailing CBOS testing also advocated acquiring this data (Törnqvist et al., 2011). Moving rope inspection from manual visual methods to computer vision opens up possibilities in machine learning and intelligent data-driven assessment of fibre rope condition.
There are several examples of RUL prediction methods for fibre ropes. RUL methods for fibre ropes have been mainly based on empirical evidence where a number of studies with CBOS testing have been conducted for both steel and fibre ropes of several different diameters (Vennemann et al., 2008; Davies et al., 2015; Novak et al., 2017; Horigome and Endo, 2018; Schmieder and Golder, 2020). Nuttall (2010) detailed a modified version of the Feyrer model for fibre ropes in lifting operations in conjunction with CBOS testing. There has also been other modelling approaches related to CBOS motion in rope as detailed by Frick et al. (2019) and Sloan (2019). Other model-based and experimental approaches to estimate service life have also been applied to HMPE ropes, under tension–tension regimes as seen in mooring, as shown by Humeau et al. (2018) and Asane et al. (2020). Additionally, Lian et al. (2015) detailed a model based on thermodynamic properties of fibres, yarns and ropes and compared predictions to tensile tests.

Some patents also incorporate discard parameters into a wider fibre rope monitoring system. Examples of this are provided by Mupende and Zerza, who detail similar patents where a discard signal is created from monitoring rope and environmental parameters (Mupende and Zerza, 2014, 2018, 2019)

In the context of lifting operations, machine learning has been applied for prognostics and health management in steel ropes with a focus on mining hoists where data acquisition is more mature and readily available (Oland et al., 2017). Onur et al. (2019) used neural networks to predict RUL of steel wire rope in CBOS testing and compared the performance to Feyrer models at different loads and diameter ratios. Xue et al. (2020) used a form of support vector machines to classify steel wire rope condition based on vibration data and Zhou et al. (2018) used convolutional neural networks (CNN) applied to images for classifying faults in balancing tail ropes for mine shaft hoisting operations. Finally, Huang et al. (2020) also applied a CNN approach and computer vision techniques that detect surface damage in steel wire ropes.

The contribution of the present article can be summarised as follows: machine learning - based methods are presented for RUL prediction of fibre rope during CBOS testing. The methods comprise neural networks (NN), support vector machines (SVM) and random forest (RF), which are applied to data extracted from computer vision and thermal monitoring. All approaches predict a target variable, known as the RUL factor (Rf), which is based on the number of cycles left to failure occurring during testing. The various model performances are then assessed for their effectiveness based on both qualitative and quantitative means. Conditions that are assessed to find the best methods for RUL assessment include: variation of hyperparameters in the models, and variations in the variables included in training the models. The approach of combining data from two rope types under different test conditions for training and RUL prediction is also explored. To the authors’ knowledge, there is no publicly available detailed study into the application of machine learning for RUL prediction specifically for fibre ropes intended for offshore lifting until now. This paper thus presents the application of existing methods for fibre rope condition monitoring and is not intended to develop the algorithms further.

The paper is organised as follows: The machine learning frameworks and how their performance is assessed in this context are detailed in Section 2. The experimental study and data sets are summarised in Section 3 and the results of the various approaches are shown in Section 4. The results are then discussed in Section 5 before further work is considered and conclusions are offered.

2. Applied methods

2.1. Target variable — RUL factor

A target variable is required for regression analysis in ML and will act as the value to be predicted based on training data used in the modelling process. The target variable used in this study is hereby referred to as the RUL factor (Rf), a fraction defined by Eq. (1).

\[ R_f = \frac{CTF}{CTF_{test}} \]  

(1)

Where Rf denotes the RUL factor, CTF, is the number of cycles to failure at the time of measurement and CTFtest is the amount of cycles at failure in each individual test where the measurements are made. This produces a value that starts at 1 representing start of life and ends at 0 representing end of life respectively (i.e 100% and 0%). Fig. 1 shows an example of damage progression in a section from a rope (A5) at various Rf, with the different stages highlighted by decreasing Rf from (a) through to (d). The images show that as Rf decreases the subsection becomes longer and there is more visible wear, as shown by ruptured strands and extruded loops. The errors of the predictions made by the various approaches will be based on comparison to the Rf.

As the CBOS tests are performed at a constant tension, Rf is related to the accumulated damage d from Palmgren-Miner’s rule as shown in Eq. (2).

\[ R_f = 1 - d \]  

(2)

As the CBOS tests progress there is accumulated damage in the form of ruptured stands and compromised sub-ropes. The rope is extended further as the true Rf values drop to lower values. Therefore, the reduction in rope residual strength and thereby the retirement criteria in the experiments can be related to the non-linear progression of the global elongation of the rope.

2.2. Neural networks

The NN structures, designated as NN1 and NN2, used in this study are detailed in Tables 1 and 2, and are implemented using the Keras library (Chollet, 2015) with a Tensorflow backend (Abadi et al., 2016). The number of inputs in the input layer for both architectures correspond to the number of features, Nfeatures used from the data extracted from the ropes during CBOS testing. The hidden layers are of a dense layer type with a specified number of neurons and a Rectified Linear Unit (ReLU) activation function that introduces non-linearity to the data. Both architectures use dropout layers that will randomly prevent 20% of the neuron outputs from proceeding through the network. This is a regularisation technique used to prevent overfitting. The final output layer consists of one neuron coupled with a Sigmoid activation function that produces a value between 0 and 1 as a result. This is done to reflect the previously described Rf in Section 2.1.

The Adam optimisation function was used for both network architectures specified in Tables 1 and 2 to update the weight values associated with the input features. The inputs are fed forward and backpropagated through the networks for 50 epochs, with the model that produces the lower mean square error (MSE) on the test data being saved as the best model used for predictions. Each configuration is simulated 20 times to account for randomness in the weights assigned in the neural network and give a more robust value for model output. The average prediction calculated from these instances is used to compare to the ground truth Rf measured from the CBOS tests and provide a confidence interval.

### Table 1

| NN architecture 1 (NN1) used to predict Rf. |
| Layer | Type |
|-------|------|
| 1 | Input layer, Nfeatures, inputs |
| 2 | Dense layer, 100 neurons, activation function — ‘ReLU’ |
| 3 | Dropout layer — 20% |
| 4 | Dense layer, 50 neurons, activation function — ‘ReLU’ |
| 5 | Dropout layer — 20% |
| 6 | Output, Dense layer, 1 neuron, activation function — ‘Sigmoid’ |
Fig. 1. Example of degradation in rope A5.

Table 2
NN architecture 2 (NN2) used to predict $R_f$.

| Layer | Type |
|-------|------|
| 1     | Input layer, $N_{f_{\text{feat}}}$ inputs |
| 2     | Dense layer, 100 neurons, activation function — ‘ReLU’ |
| 3     | Dropout layer — 20% |
| 4     | Dense layer, 100 neurons, activation function — ‘ReLU’ |
| 5     | Dropout layer — 20% |
| 6     | Dense layer, 50 neurons, activation function — ‘ReLU’ |
| 7     | Dropout layer — 20% |
| 8     | Output, Dense layer, 1 neuron, activation function — ‘Sigmoid’ |

2.3. Support vector machine

SVM was first used for classification by Cortes and Vapnik (1995) and was later adapted for regression problems by Vapnik et al. (1997). To predict the $R_f$ in this context, the latter approach is adopted which has also been applied in other studies related to RUL prediction, such as García Nieto et al. (2015). Essentially, the data is separated by a hyperplane in a higher vector space. This plane can be formed by use of a kernel. In a classification implementation, this line is used to separate the measured vectors into classes, but for the regression analysis it will be used to predict a continuous variable for the other instances in the data set.

To compare to the performance of the NN, it is chosen to use an SVM framework adapted for regression analysis from scikit-learn (Pedregosa et al., 2011). Linear and Gaussian kernels are used to form the fit to the data and to compare their relative accuracy to $R_f$ are compared to other methods.

2.4. Random forest

RF is an example of an ensemble method which utilises a user-specified number of decision trees created by bootstrapping data from features and data available from a training pool Breiman (2001) and Kundu et al. (2020). The models created will assign RUL values to the test samples in each individual tree and an average RUL value will be calculated.

The RF algorithm for regression analysis from scikit-learn (Pedregosa et al., 2011) is used for $R_f$ prediction. The configurations for RF implementation in each data set are specified in Table 3 for data sets A and B. Different tree depths are chosen due to the difference in number of measurements available between the different data sets and to prevent overfitting. The differences between the data sets are highlighted in Table 5.

3. Experimental study

Fig. 2 details the flowchart of operations implemented in the ML process from beginning to final output and performance assessment.

Table 3
Configurations used for RF to predict $R_f$ for data sets A and B.

| Data set A | Data set B |
|------------|------------|
| Tree depth  | 1–4        | Tree depth  | 1–2        |
| Number of trees | 50, 100, 200 | Number of trees | 50, 100, 200 |

Specific details related to (a) test methods and (b) data acquisition are summarised in Section 3.1 and the subsequent pre-processing techniques applied to the data for ML application in Section 3.2. Then in Sections 3.3 and 3.4 the (c) training and RUL estimation stages and (d) model assessment are detailed.
3.1. Test methods and data acquisition

The data sets were recorded from CBOS experiments performed at the Mechatronics Innovation Lab in Grimstad, Norway. The fibre ropes used in testing were 12-strand braided ropes with nominal diameter of 28 mm. The two types used were five samples of Dyneema DM20 XBO and four samples of Samson AmSteel Blue which are designated as separate campaigns A and B, respectively. Additionally, A is tested at safety factor (SF) 11 and B is tested at SF 8. This equates to average tensions of 1/11 and 1/8 of each rope type’s max tensile capacity.

Features are engineered from data recorded during CBOS testing and the ropes were separated into different zones that reflect the distinct bending regimes during CBOS testing, as highlighted in Fig. 3. The acronyms SZ, SBZ and DBZ refer to straight zone, single bend zone and double bend zone, respectively. These are the positions in the rope where the rope is bending and unbending due to the cyclic movement of the sheaves. No bending occurs in section SZ, a single bend occurs during each cycle in SBZ and two bends occur in the DBZ during one cycle.

The computer vision setup takes account of the whole rope surface, with four Edmund Optics 13122C cameras placed around the rope as shown in Fig. 4. This takes into account that degradation will not be uniform around the rope and the different camera perspectives can view this. Computer vision data is recorded for 2000 images, corresponding to 13–15 complete cycles, every 1000 cycles.

A FLIR A6753sc thermal camera is placed next to the test sheave of the CBOS machine, as shown in Fig. 5. It was set to sample at 100 Hz for 2000 images, resulting in a 20 s video for each period. This was sufficient to record at least one full cycle in the CBOS test. Temperatures are only available for the lumped zones SZ, SBZ and DBZ.

The global length elongation was measured by a Fluke 414D distance measuring laser that tracked the extension of the CBOS machine hydraulic cylinder as the tests progressed. The sensor has an accuracy of +/−2.0 mm.

Computer vision data is extracted via algorithms developed in OpenCV (Bradski, 2000). The “geo” data represents the local length and width measurements and values for each recording are thus aggregated to give median, maximum, minimum and standard deviations for these geometric features. The “therm” data represents thermal measurements taken from the thermal camera and extracted using FLIR ResearchIR 4 software (FLIR, 2015). These are the temperature values within the rope part of each relevant image are aggregated as average, maximum, minimum and standard deviation. A list of features and their respective bending zones are summarised in Table 4.

Specific details of the various operations the algorithms used to extract data are summarised in previous work detailing the various CBOS testing campaigns by Falconer et al. (2020).

3.2. Data pre-processing

The data sets are separated into “geo” and “geo therm” feature sets based on data type to assess the effect of adding thermal features to the ML model. The data set compositions are summarised in Table 5. The records are the number of data samples recorded by the data acquisition set up mentioned in Section 3.1. The ropes used in campaign B were tested at a higher tension and failed sooner than those in campaign A, hence the lower number of records in B compared to A. The number of features is found by multiplying the number of zones by the number of statistical parameters for each type of data (e.g. 8 × 4 + 8 × 4 + 1 = 65 for geometric features and 3 × 5 + 1 = 16 for thermal features).

The raw measurements from the data acquisition phase are subject to pre-processing, where the data from the various features is standardised. The data was scaled using the approach detailed in Eq. (3):

\[
z_i = \frac{y_i - \bar{y}}{\sigma_y}
\]
where $z_i$ is the individual standardised value, $y_j$ is the individual raw value, $\bar{y}$ is the average of all readings for the specific feature in the rope sample and $\sigma_y$ is the standard deviation of all readings for the specific feature in the rope sample. This is done for each rope in the data set to increase comparability between the different samples and is a standard pre-processing step to prepare data for ML application.

### 3.3. Training and RUL estimation

The models are tested through leave one out cross validation (LOOCV), where $RF$ predictions made on a single rope will be made using models trained with the remaining ropes in the data set. This process is done separately for both ropes in campaign A and campaign B and as a combined data set with both A and B together.

In addition to investigating the effect of only geometric and combined geometric and thermal measurements as outlined in Table 5, different combinations of features from the various bending zones detailed in Fig. 3 are also trained and tested. This includes: testing using all zones; using only SBZ and DBZ measurements; and finally, using only DBZ measurements.

### 3.4. Model assessment

The metrics used for model assessment are root mean square error (RMSE) and $R^2$ score ($R^2$):

\[
\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{y}_i)^2} \quad (4)
\]

\[
R^2 = 1 - \frac{SS_{res}}{SS_{tot}} \quad (5)
\]

Where $N$ is the number of measurements made for each rope sample, $y_i$ is the observed $RF$ at instance $i$, $\hat{y}_i$ is the predicted $RF$ at instance $i$, $SS_{res}$ is the sum of squares of residuals and $SS_{tot}$ is the total sum of squares.

RMSE in this context will give an insight into how concentrated the $RF$ predictions are around the $RF$ ground truth using each algorithm configuration outlined in Section 2. Additionally, the $R^2$ score gauges the correlation $RF$ predictions have with the ground truth $RF$ using the input variables outlined in Table 4.

The performance of the models is assessed through the use of $RF$ graphs and residual analysis. The $RF$ graphs will include the ground truth from the observed experimental measurements and tolerance bounds at ±20% based on this data. The average $RF$ prediction and a ±95% confidence interval of the repeated simulations of each distinct configuration are compared to the ground truth and tolerance bounds.

The residual values, $\epsilon$ are calculated using Eq. (6):

\[
\epsilon = y_i - \hat{y}_i \quad (6)
\]

where $y_i$ are the RUL values from CBOS testing and $\hat{y}_i$ are the predictions made by each respective model. The analysis includes plotting the residual values against the predicted $RF$ values, the actual $RF$ value against the predicted $RF$ value and finally a histogram analysis accounting for the numerical spread of the residual values. This provides further information about the model ability to predict $RF$, potential model bias and where overestimation and underestimation in rope health occurs.

### 4. Results

#### 4.1. Experimental results

The amount of cycles at failure for all ropes in campaigns A and B are summarised in Table 6. It is noted that there is a wider spread of values in A than B. Ropes A1 and A4 were shown to fail at lower amounts of cycles compared to the other ropes in campaign A. This is owed to the rope failing at the driving sheave, rather than the test sheave. Due the method used, parts of the splice were in contact with the sheave during testing and lead to premature failure of the rope samples. The features monitored at the test sheave in ropes A1 and A4 did not develop as much as those in A2, A3 and A5 but all ropes showed the same patterns in progression of global elongation through out every test. This included a period of accelerated elongation of global length towards the failure of the rope.

Moreover, rope A5 completed a greater number of cycles than the rest of the ropes in campaign A. This particular sample included an attempt to embed thermocouples within the strands of the rope and therefore could have contributed to slight discrepancy in the results.

Some halts in logging occurred sporadically during CBOS testing for campaign A, meaning data acquisition equipment had to be restarted occasionally. Furthermore, features for Rope B1 were recorded every 1000 cycles and has less data compared to the other campaign B ropes, which were recorded every 500 cycles. This change was made as a reaction to the comparatively shorter test times for campaign B than campaign A.

#### 4.2. Average metrics

A quantitative assessment of $RF$ prediction is performed by calculating the average $RMSE$ and $R^2$ scores in each data set using different feature sets. Better performance is reflected by lower and higher values for $RMSE$ and $R^2$, respectively. The results for different feature combinations are compared for both A and B rope datasets and ranked by performance of the algorithm on only using geometric features (i.e A geo with A model and B geo with B model). The results using the various feature combinations cross validated on a combined A and B rope data set are also presented.

Fig. 6 gives an overview of each algorithm performance based on different feature sets and training model composition. It is seen that both configurations of BF performed best, followed by NN and finally SVM based on producing the lowest $RMSE$ scores. Introducing thermal features generally either changes nothing or leads to detriment in performance when only A ropes are used as training data, which is particularly noticeable in both NN and SVM. When the combined A and B model is used there is shown to be a slight improvement in performance when thermal features are introduced for NN when the combined SF+DBZ and DBZ feature sets are considered. There is also shown to be a general reduction of $RMSE$ as the number of zones used in the training are decreased.
This performance is also reflected in the $R^2$ values as shown in Fig. 7. As seen with the $RMSE$, data types, feature sets and model compositions influenced the algorithms differently. The $R^2$ values for both RF configurations are shown to be the highest, followed by NN and SVM. The high values recorded for $R^2$ indicate that the features used to create the RF models better explain the changes in $R_f$. It is seen that by using both A and B data to train the NN configurations, that a higher $R^2$ is achieved than by only using A, indicating that the same feature with enhanced data from different ropes better explain the changes in $R_f$. This is also seen for both SVM configurations but had little impact on RF.

The $RMSE$ values for data set B are shown in Fig. 8. Both configurations of NN were shown to perform best when only B data is considered. A noticeable difference in the B data is that for both NN and SVM, the thermal data led to significantly better predictive performance as indicated by lower $RMSE$. Using the combined A and B data set also improved performance of NN with the exception of the NN2 configuration in the SBZ+DBZ feature set. The combined A and B data set has a detrimental impact on the RF configuration with depth one.

The $R^2$ scores for the same algorithms using the B rope data set are shown in Fig. 9. The highest $R^2$ values are achieved with NN, with improvements being shown when thermal data is introduced. It is noted that higher $R^2$ is recorded using A and B data for training when only geometric features are considered but similar values are seen for B only and A and B combined when thermal features are introduced. RF is shown to have similar $R^2$ values no matter what features are used.
There is also a noticeable increase in $R^2$ for SVM Gaussian when the combined A and B data set is considered.

### 4.3. RUL graphs

A qualitative assessment of the algorithm performance during cross validation is given through plotting the predicted $R_f$ at various cycles throughout each test. It is possible to gain a general idea of algorithm performance from the graphs, but they can also reveal at which times in the test both the best and worst predictions are made. The results of cross validation predictions made only using the single and combined data sets is also considered. Selected RUL prediction results from both A and B data sets are shown in Figs. 10, 11, 12 and 13.

Fig. 10 shows the results of $R_f$ prediction using NN2 for geometric features from all bending zones for rope A4. By using the combined A and B data set, the predictions in the first half of the test are shown to be closer to the ground truth, as well as showing a reduction in the overestimation “peak” seen in the middle of the A ropes model. The method was still shown to overestimate from 50,000 cycles until failure, but was slightly improved with the addition of all ropes to the model.

Fig. 11 presents the results for $R_f$ prediction for rope A3 using RF at tree depth 4 using geometric features from all bend zones. As shown in the previous $RMSE$ results for RF, the combined A and B training data performed either the same or to slight decrease in performance, which is reflected in the $R_f$ graphs. There is a noticeable period of overestimation in the combined A and B model between 40,000 and
60,000 cycles but returns to being closer to RUL just after. Additionally, both models have very small confidence intervals compared to the NN configurations.

Fig. 12 considers the predictions by NN2 on rope B2 when both geometric and thermal features from the DBZ are considered. The confidence interval in the prediction is significantly reduced and almost totally confined to the tolerance bounds. This is seen at the majority of test times with the exception of between 0 to 40,000 test cycles. The overestimation in the B model from 4000 cycles until failure is also corrected to being closer to the true RUL values by introducing the A rope data.

Fig. 13 shows the $R_f$ prediction results of using SVM with a linear kernel when using geometric and thermal features from the DBZ. The model trained using the B ropes data is shown to significantly overestimate rope health after around 11,000 cycles but when the A and B data is included, this period of $R_f$ predictions is corrected, albeit with some underestimation after 12,000 cycles.

4.4. Residual analysis

A residual analysis is performed to further investigate $R_f$ predictions and to compare the differences in results depending on algorithm, features and training data used.

Fig. 14 shows the residual analysis comparison of using RF with tree depth 4 and NN2 predictions for rope A2 using only geometric features from all bend zones. In this specific case, both models were trained using only the A rope data. The RF model produces low residual values indicating closer agreement with the ground truth, while NN2 prediction are shown to both overestimate and underestimate as indicated by the spread of residual values.

Fig. 15 presents the residual analysis comparison of using geometric and the combined geometric and thermal features. This is done for rope B3 predictions made using NN2 with features from the DBZ. By including thermal data, the magnitude of all residuals is reduced to be within 0.1 of the $R_f$ ground truth. When comparing predicted and actual $R_f$ values, predictions are more effective generally but there is little effect on improving the predictions at the end of the test.

Fig. 16 provides analysis of improvements that are possible by combining the A and B rope data sets in model training. This is presented for rope A1 with NN2 using combined geometric and thermal data from all bend zones. An improvement with the combined training set is observed, with magnitude of the outermost outliers being reduced. Rope A1 completed the least amount of cycles at failure when compared to the other ropes in the A data set and showed poor prediction results on models trained on this data set.

5. Discussion

Model performance varied depending on the algorithm and the data set used. A major difference shown was that RF performs better than NN for campaign A but NN performs better than RF in campaign B. Differences in performance can be explained by the training data used, the size of the data set, and algorithm mechanisms.

5.1. Recorded data and availability

Two of the ropes in campaign A, ropes A1 and A4, broke where the rope splice was bent over the driving sheave. This led to earlier failure than for the other ropes of the campaign, and to failure which was not located where the cameras were, at the test sheave. Still, the failure mechanism was very similar to what caused failure in the other ropes tested in this study. RF frequently identified the global elongation as the most important split variable. In line with this overall sameness of failure mechanism of all ropes tested, this variable showed a similar development for all ropes, including ropes A1 and A4.

It was chosen to include the ropes in the analysis, a choice that has two positive effects. Excluding them would have lead to even less data for training the machine learning algorithm, and as discussed earlier, CBOS test data are time consuming and expensive to obtain. Secondly, in real life, ropes will not always break where expected. Including ropes A1 and A4 in the data set makes the models able to handle some such cases, where the deviation from the expectation is not too large.

The model stability is addressed through the 95% confidence interval for the predictions. It was found that combining campaigns A and B in all cases resulted in a narrower confidence interval than from each of the campaigns alone. For unstable models, adding new samples is expected to change the model qualitatively, resulting in wider confidence intervals. Conversely, the present study’s narrowing of the confidence intervals indicates a stable model.
5.2. Random forest

RF is formed of several individual decision trees, where data is separated based on feature values that give the purest split. The global elongation shows a steady increase during testing, and it will therefore create one of the best splits in the decision trees of the RFs. However, this feature will not be available to all trees in the forest, due to feature bagging. Still, the averaging over 200 trees will ensure its contribution to the overall forests. This is shown in the residual analysis performed in Fig. 14, where RF has outperformed NN2. RF was also shown to have a steady balance between slight overestimation and underestimation of predicted $R_f$ compared to the true value.

In general, increased tree depth increases random forest accuracy and improvements by increasing forest size tend to plateau. In campaign B, each rope had fewer records than in campaign A. To avoid overfitting, the result was shallower trees for this campaign, giving less accurate RF results. The data is split by random feature sampling in each tree, there would be slightly more variation in the quality of fits due to data splits made on less suitable features. This leads to less accurate RUL predictions as indicated by the higher RMSE and lower $R^2$ values.
5.4. Support vector machine

The SVM algorithms performed poorly compared to the other algorithms applied in this study. If the hyperplane fit to the data is poor and non-representative, it will in turn have a detrimental impact on $R_f$ prediction using the algorithm. As with NN, the presence of noise via lack of variation in certain features will negatively influence this fit. This is particularly prevalent in campaign A, however it is noticed that SVM performed slightly better in campaign B due to the greater variation in values measured. It is also shown that the fit created by the linear kernel suits these rope test data set better than the Gaussian kernel. The SVM is also shown to overestimate rope health towards the later stages of the CBOS tests.

Potential improvements for SVM could be reducing the number of features used to create a simpler model. There is also further potential in hyperparameter optimisation for both linear and Gaussian kernels that can contribute to improve models.

5.5. Feature selection

Creating models from different feature sets consisting of the different bending zones features also influences algorithm performance. During CBOS testing, the SBZ and DBZ are subject to substantially more bending, whereas there is none present in the SZ. Therefore, in line with what was previously stated about NN, measurements from the SZ essentially contribute noise in the modelling process. As only the features that give the best splits in the data are considered from the random subsets in RF, the likelihood that features from SZ will be consistently picked as splitting criteria in individual trees is extremely low, hence minimal effect on both $RMSE$ and $R^2$ values.

However, limiting the features used to only those from the DBZ is shown to improve algorithm performance for both NN and SVM. There was very little difference between the results from models that used all features and the combination of SBZ and DBZ features but the most accurate results for SVM were achieved by reducing to only training with DBZ features. More frequent repeated bending will cause more accumulated damage and variation in these features will relate better to the associated RUL value.

The residual analysis performed in Fig. 15 also shows the advantage of combining both geometric and thermal features as training data. The higher temperatures in campaign B in this example contribute features with greater variation and therefore improve the prediction capabilities of the NN2 model used in this example.

5.6. Combining data sets

Combining the two data sets for predictions based on cross validation had both a positive and negative impact depending on the algorithm considered. The main benefactors of this approach were the NN configurations, as shown by improved results in average $RMSE$ and $R^2$ for both the A and B ropes. Introducing thermal features for A rope predictions led to decreased performance when only considering the A data set, but improved for the cases where SBZ+DBZ and DBZ features are used with models trained using the A and B ropes. This can be attributed to the different thermal behaviour in each data set, with the thermal information from the B ropes contributing to better predictions. From a B ropes prediction perspective, combining the A and B data sets creates a larger training data set and improves estimations from both NN and SVM. This suggests that simply increasing the amount of training data with slightly different feature behaviours will benefit model fitting in these cases. The observations support the finding from this study that unique rope types subject to different relative test tensions can be combined to produce more accurate results.

The residual analysis performed in Fig. 16 also shows the advantage of combining the separate test campaigns. In campaign A, the thermal data varied very little due to the lower test tension applied. Therefore the little variation in the thermal data would contribute noise to the modelling process. The introduction of campaign B, improved the...
models due to both the presence of similar global elongation trends and thermal features that showed more variation due to the higher test tension.

6. Future work and adaptation for field deployment

From the perspective of RUL prediction in CBOS testing, the effectiveness of ML approaches is influenced by the quality and quantity of data recorded. Though good prediction results are demonstrated with a small pool of data at two safety factors, further ML application to CBOS test sets performed at a wider range of lower safety factors is needed. These further tests would also benefit from more frequent data recording, that would give a more rounded picture of the performance of a specific rope type. This would be beneficial for methods such as RF, which gave excellent results in campaign A but had its performance inhibited by the limited tree depths in campaign B. Furthermore, it is shown that these further enhanced data sets can be combined into a common training data pool that can be used to make predictions on new test data from distinct rope types from various test conditions.

Data availability is another limiting factor to further development of ML algorithms for application to fibre ropes. Other engineering components such as bearings and motors benefit from publicly available datasets and can therefore specifically focus on further development of prognostics methods, rather than the condition monitoring framework for data acquisition.

The data acquisition methods can also be further developed to improve the presented models. The current method does not account for the internal degradation for the rope, which would provide another perspective to assess damage in fibre rope. This could be done through embedded sensors that could account for changes in internal temperature or further exploration of localised strain. Additionally, as the structure of the rope fuses with the test progression, acoustic sensors could be placed on the sheave or internally to monitor changes in the micro-structure of the rope.

This study provides an overview of the main ML techniques used for RUL prediction, but each algorithm mentioned has potential for further exploration. This was done on tabular data that dealt with measurable physical quantities such as width, length and temperatures but there are still further opportunities for enhancement by using CNN to quantify damage in the rope condition directly from visual and thermal images. This could be extended to automatically detect visible abrasion, fluffiness, polishing and broken strands from the same images used in this study. ML as a method combined with current inspection methods would further facilitate its application in real time inspection of fibre ropes in cranes and would provide a useful additional source for decision making on rope retirement.

However, there is further work required to fully adapt the ML approach usage in offshore construction cranes. Firstly, CBOS testing is an example of a run-to-failure test where a constant tension is applied at the same rope sections. In real practice, it will be different sections of the rope that will be subject to bending depending on lift depth and sea state during operations. Other factors such as payload size, the temperature at lift location and the operation time will also influence rope longevity. Data from potential lifting campaigns would have to be recorded as part of a wider rope condition monitoring system, where data is continually added to form a model with historical data. This would create a model that specifically pertains to fibre rope use for offshore lifting, rather than relying on more commonly used S-N curves for these specific types of ropes. The findings from this study of different ropes at different relative test tensions show there is potential in using data from a fleet of fibre rope cranes performing different lifting operations to validate the condition of an individual fibre rope crane.

Intelligent prognosis techniques would then be applied to assess this data and establish what are the most relevant features to be monitored in practice and further develop a retirement criteria based on informed data-driven methods for condition based maintenance, rather than manual inspection geared towards reactive and time-based maintenance practices. In this study it is shown that there are benefits from monitoring the length, both locally and globally, and the temperature in different bend zones. Such measurements would be made from sensors collecting visual and IR data, ideally from a location near the main sheave, from a rope that has clearly defined sections. The global rope length measurement would also have to be incorporated through some form of embedded sensor measurements within the rope structure to be monitored continually during use. This could be done in the form of fibre optic threads, allowing the changes in the light propagation to be continually measured and estimates to be made on the global length changes.

7. Conclusion

Several approaches for RUL prediction in fibre ropes during CBOS testing are discussed in this work. The algorithms in this study are capable of predicting a continuous target variable, known as $R_f$, for ropes using features derived from an experimental set-up that uses computer vision and thermal monitoring.

In this investigation machine learning methods, such as neural networks, random forest and support vector machine were applied for prognostics for two sets of CBOS test data at different safety factors. For data set A, random forest showed the most promise as a RUL prediction method, while NN was the best performing algorithm in data set B. The benefit of combining data from different types of ropes for training data for RUL prediction is also demonstrated for NN and SVM. RF has been shown to be the most effective in this study, particularly in cases where larger amounts of data are available, allowing a suitably complex model to be developed based on features selected by the algorithm. NN is also shown to be useful, but slightly less effective compared to RF. If large amounts of data are not available, then NN application with more focus in including only relevant features is a useful fall-back solution.

CRediT authorship contribution statement

Shaun Falconer: Conceptualisation, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing – original draft, Writing – review & editing, Visualisation. Ellen Nordgård-Hansen: Conceptualisation, Methodology, Data curation, Resources, Writing – review & editing, Supervision. Geir Grasmo: Conceptualisation, Resources, Methodology, Writing – review & editing, Supervision.

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 research presented in this paper has received funding from the Norwegian Research Council, SFI Center for Offshore Mechatronics, project number 237896.

Special thanks are given to Nadia Saad Noori for valuable discussions and advice related to the application of neural networks in this context, Benyamin Akdemir for help in setting up the equipment and the monitoring software, and to Yannick Bafanga for making Matlab code for pre-processing the thermal data.
