

ПРОГНОЗИРОВАНИЕ ОСТАТОЧНОГО РЕСУРСА ОБОРУДОВАНИЯ В УСЛОВИЯХ МАЛОЙ ВЫБОРКИ ДАННЫХ

Задиран К. С.¹, Щербаков М. В.², Сай Ван Квонг³
(Волгоградский государственный технический
университет, Волгоград)

Предлагается метод прогнозирования остаточного ресурса оборудования, использующий глубокое обучение и применимый в случаях с малым количеством информации об отказах в данных, где существующие классические методы могут не давать требуемой точности. Процесс поддержания оборудования в рабочем состоянии – один из наиболее важных процессов в эксплуатации оборудования. При этом процесс технического обслуживания зачастую страдает от недостаточной эффективности. Поэтому были разработаны методы прогнозирования, на основе которых была построена концепция проактивного управления процессом техобслуживания, позволяющая оптимизировать структуру и затраты управления оборудованием на протяжении жизненного цикла. Однако данные методы могут показывать недостаточную точность, если для их обучения недостаточно данных, например, в связи с редкостью возникновения отказов в оборудовании. Для решения этой проблемы предлагается новый метод прогнозирования, в основе которого лежит алгоритм, основанный на глубоком обучении и который может улучшить точность прогнозирования. В данном методе произведена замена непрерывного прогнозирования остаточного ресурса оборудования на всем интервале на систему генерации сигналов, содержащих рассчитанный прогноз.

Ключевые слова: машинное обучение, остаточный ресурс, проактивное техническое обслуживание.

1. Введение

В настоящее время в связи с увеличением сложности, разнообразия и количества промышленного оборудования повышается важность проблемы управления данным оборудованием и его жизненным циклом [2]. В течение жизненного цикла оборудования производится техническое обслуживание, эффективность которого в значительной степени зависит от модели

¹ Константин Сергеевич Задиран, аспирант (konstantin.zadiran@gmail.com).

² Максим Владимирович Щербаков, д.т.н., доцент (maxim.shcherbakov@vstu.ru).

³ Ван Квонг Сай, к.т.н. (svcuonghvtqs@gmail.com).

управления данным процессом [2]. Развитие технологий проактивной поддержки принятия решений позволяет перейти к управлению техническим обслуживанием на основе остаточного ресурса [5]. Данные технологии имеют потенциал быть более эффективными в планировании графика технического осмотра и обслуживания оборудования, чем установленные производителем процедуры и рекомендованные интервалы, которые могут не учитывать конкретные условия эксплуатации и нагрузку оборудования [1, 3].

Остаточный ресурс оборудования – величина суммарной наработки оборудования от текущего момента до его отказа. Под отказом понимается предельное состояние оборудования, при котором невозможно проводить дальнейшую эксплуатацию без причинения ущерба оборудованию или необходимости восстановления работоспособного состояния оборудования [3].

Активно развиваются аналитические методы машинного обучения и анализа данных для прогнозирования остаточного ресурса оборудования [4]. Эти методы основаны на прогнозировании временных рядов и требуют для функционирования данные показателей оборудования, записанные во времени. Широкая доступность таких данных в настоящее время обеспечивается развитием и удешевлением технической базы технологий Интернета вещей – сенсоров для считывания данных, технологий и оборудования для передачи и хранения этих данных.

Прогнозирование отказов оборудования сводится к прогнозированию остаточного ресурса оборудования, или RUL – Remaining Useful Life, – подхода, совместимого с применением методов и алгоритмов машинного обучения. Значение параметра RUL определяется как временной интервал, обозначающий остаточный ресурс компонента или системы промышленного оборудования [8, 21]. Если принять t_0 за момент времени возникновения отказа, а t_c – за текущий момент времени, то RUL для момента t_c можно записать в виде

$$(1) \quad RUL_{t_c} = t_0 - t_c.$$

Данная зависимость проиллюстрирована на рис. 1.

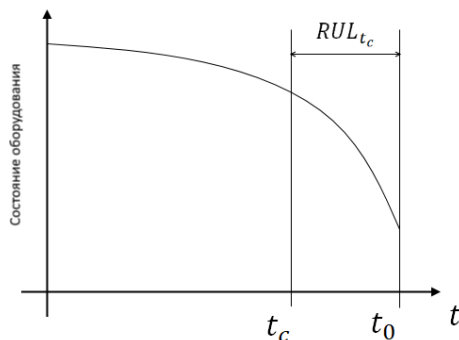


Рис. 1. Зависимость между моментом отказа и остаточным ресурсом

Все методы прогнозирования остаточного ресурса можно разделить на методы, основанные на моделях, и методы, основанные на данных [17]. Методы, основанные на моделях, основываются на применении экспертных знаний для физико-математического моделирования поведения оборудования. Однако недостатком этих методов является сложность построения достоверной модели для сложных или изменяющихся условий эксплуатации, а также при недостатке экспертизы в целевом направлении [13, 15].

Методы, основанные на данных, позволяют уйти от этих недостатков, так как с позиции применения этих методов представление процесса эксплуатации оборудования сводится к временным рядам управляющих параметров и данных о состоянии оборудования [12, 24], а построение прогноза выполняется с применением методов машинного обучения.

Основанные на данных методы, в свою очередь, делятся на классические и методы глубокого обучения [19]. К классическим методам относятся, например, метод опорных векторов (SVM), случайный лес (RF), метод k -ближайших соседей (KNN) [4]. Данные методы требуют меньшего привлечения экспертных знаний для построения моделей, чем методы, основанные на моделях, однако могут зависеть от них в определении значащих

управляющих параметров и параметров эксплуатации оборудования для обучения моделей.

Полностью уйти от зависимости от экспертных знаний позволяют методы глубокого обучения. Данные методы получили широкое развитие в последнее время. В работах [16, 18, 22, 23] рассматриваются методы, основанные на алгоритмах сетей долгой краткосрочной памяти (LSTM), рекуррентных нейронных сетей (RNN), конволюционных нейронных сетей (CNN), SVM, Random Forest, XGBoost, а также их комбинациях. Данные методы показывают большую точность прогнозирования, чем классические, ориентированные на данные методы [7, 9].

Однако для точного прогнозирования RUL методами, использующими алгоритмы глубокого обучения, необходим большой объем данных, что соответствует большому количеству отказов. В случае малого количества отказов данных может быть недостаточно и, следовательно, точность метода может снижаться. В таблице 1 приведен пример сравнения результатов нескольких методов прогнозирования RUL для данных NASA Turbofan Degradation Dataset[4] для случаев, когда доступно 100 и 50 отказов для обучения модели, оцененных с использованием RMSE[20].

Таблица 1. Точность в зависимости от доступности данных

Метод	100 отказов	50 отказов
Random Forest	24,17	28,24
RNN	23,78	37,86
CNN	19,04	25,13
SVM	22,69	29,51

При этом результаты прогнозирования могут содержать в себе полезную информацию, которую можно потенциально из них извлечь. В этой статье предлагается метод прогнозирования RUL, в котором учтен приведенный сценарий использования данных с ограниченным количеством отказов.

2. Метод

Чтобы повысить точность, предлагается прогнозировать RUL не непрерывно для каждого интервала во временном ряду, а в отдельные моменты времени, когда остаточный ресурс оборудования можно определить с наибольшей точностью. В данном подходе производится анализ прогноза, формируемого непрерывно для каждого из измерений, по результатам которого производится генерация сигналов со значением RUL, если входные данные будут соответствовать параметрам, определенным на стадии обучения модели.

Данное решение позволит исключить из результатов прогнозирования отдельные элементы, для которых невозможно определить результат с достаточной точностью, тем самым повысив точность прогноза в целом. Схема предлагаемого метода приведена на рис. 2.

Входными данными являются временные ряды, содержащие данные с датчиков оборудования. Данные группируются по отказам. Для каждого отказа оборудования формируется отдельный временной ряд, в котором последнее измерение совпадает по времени с моментом фиксации отказа.

Выходными данными работы метода являются сигналы, для которых $T_{\text{сигн}} + \Delta t = T_{\text{отк}}$, где $T_{\text{сигн}}$ – время возникновения сигнала, $T_{\text{отк}}$ – спрогнозированный момент наступления отказа оборудования, Δt – интервал времени, между возникновением сигнала и наступлением отказа оборудования, определенный на стадии обучения модели, построенной с использованием предлагаемого метода.

Предлагаемый метод включает два режима работы: режим «Оффлайн» и режим «Онлайн», подразумевающий работу в режиме реального времени. В первом режиме производится подготовка исторических данных, модели прогнозирования и правил для генерации сигналов. Во втором режиме на вход подаются данные с оборудования в реальном времени, производится их обработка и вычисление прогноза RUL и на основании прогноза при совпадении с одним из правил производится генерация сигнала.

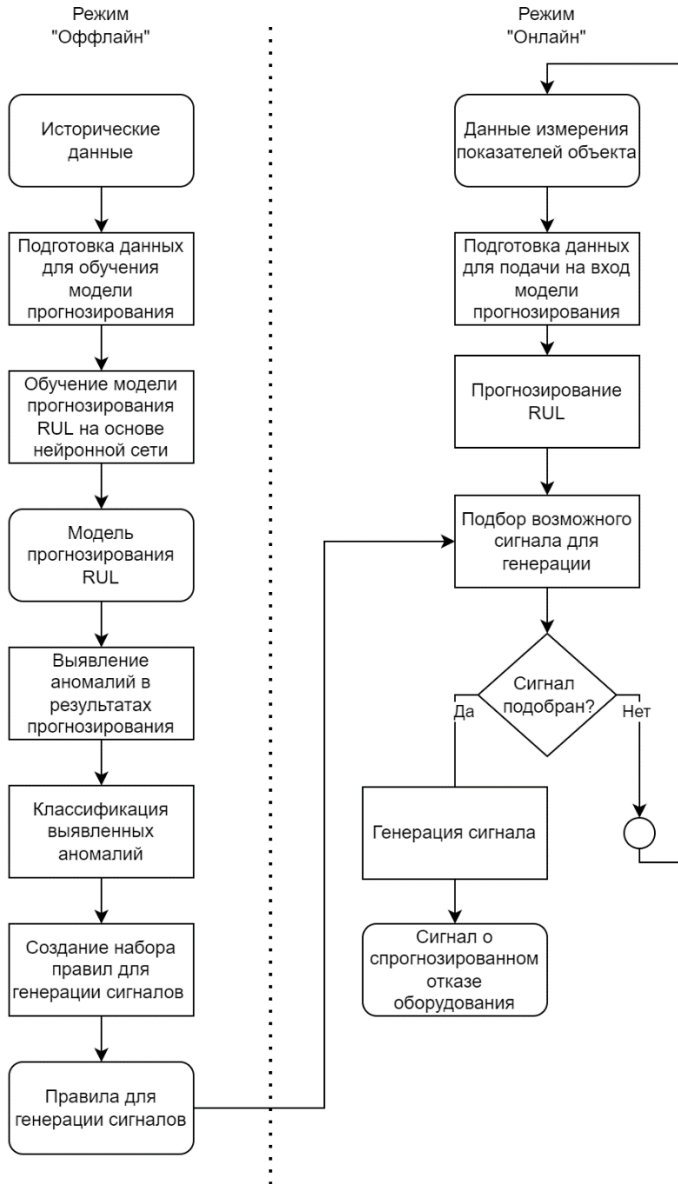


Рис. 2. Схема предлагаемого метода

Предлагаемый метод имеет два гиперпараметра, влияющих на процесс обучения: M – количество записей перед точкой наступления отказа $T_{отк}$, которые используются в процессе обучения модели, образуя диапазон $(T_{отк} - M; T_{отк}]$; w – количество записей перед точкой $t \in (T_{отк} - M; T_{отк}]$, из которых формируются входные данные.

Алгоритм режима «Оффлайн» состоит из двух крупных шагов, которые подразделяются на более мелкие операции. На первом шаге производится расчет F^{RUL} – спрогнозированного значения RUL для входных данных.

Для обучающей выборки определяется RUL по формуле

$$(2) \quad RUL_i = T_{отк} - T_i,$$

где RUL_i – RUL для момента времени T_i . После определения RUL производится отсеивание измерений, не входящих в интервал $(T_{отк} - M; T_{отк}]$.

Далее факторы входных данных преобразуются методом экспоненциального сглаживания.

Затем производится нормализация данных

$$(3) \quad x_{norm} = \frac{x_i - x_{min}}{x_{max} - x_{min}}.$$

Далее производится построение выборки, используемой для обучения искусственной нейронной сети CNN. Для этого применяется метод скользящего окна, при котором для выходного параметра RUL_i входные параметры определяются как матрица A_i вида

$$(4) \quad A_i = \begin{pmatrix} x_{(i-w;1)} & - & x_{(i;1)} \\ | & \backslash & | \\ x_{(i-w;1)} & - & x_{(i;n)} \end{pmatrix},$$

где w – количество измерений до момента времени t_i , n – количество факторов во входных данных.

Далее производится построение и обучение искусственной нейронной сети CNN. Обученная нейронная сеть используется для прогнозирования F^{RUL} – прогнозного значения RUL.

На втором шаге производится построение модели (или набора правил) для генерации сигналов.

Для спрогнозированного F^{RUL} вычисляется ряд изменений значения:

$$(5) \quad \Delta F_i^{RUL} = F_i^{RUL} - F_{i-1}^{RUL}.$$

Далее, с использованием метода IQR [14] производится выявление аномальных значений в ряду ΔF^{RUL} . В рамках разработанного метода эти аномальные значения могут быть соотнесены с конкретным значением RUL.

Чтобы соотнести эти аномальные значения с RUL, производится их классификация с использованием алгоритма Random Forest [10]. Классификация производится в два этапа. На первом в режиме обучения без учителя строится матрица расстояний для аномальных значений. На втором этапе на основании этой матрицы производится распределение значений на классы.

Для каждого класса вычисляется две характеристики: диапазон изменения значений, соответствующих этому классу, и RUL, вычисляемый как среднее значение RUL для всех элементов класса. Так как диапазоны изменения значений классов могут пересекаться, принадлежность точки к классу в режиме тестирования/использования обученной модели производится в порядке уменьшения значения RUL для класса, и если для какого-то класса получено совпадение, то новая точка классифицируется как элемент данного класса и дальнейшая проверка на принадлежность не производится.

По завершении процесса обучения становится доступен переход к работе в режиме «Онлайн». В данном режиме используется модель прогнозирования RUL на основе нейронной сети CNN, подготовленная в режиме «Оффлайн».

Данные о состоянии оборудования поступают и обрабатываются в реальном времени. Для этого выполняется их предобработка, состоящая из аналогичных предыдущему режиму шагов. Обработанные данные подаются на вход модели прогнозирования, которая строит прогноз RUL. Вычисляется разница между текущим и предыдущим значением RUL.

Для полученной разницы производится поиск подходящего правила для генерации сигнала по попаданию в диапазон значений изменения RUL. Если правило найдено, генерируется сиг-

нал, содержащий значение остаточного ресурса оборудования для данного правила.

3. Реализация

Алгоритм прогнозирования был реализован с использованием языка программирования R, среды разработки R-Studio.

При реализации использовались следующие пакеты R: 1) keras – для построения и обучения искусственной нейронной сети CNN [10]; 2) readr, dplyr – для работы с данными [6]; 3) forecast – для выявления аномалий во временном ряду [11]; 4) randomForest, cluster – для кластеризации [10].

4. Результаты

Для оценки точности предложенного метода использовались данные с эксплуатации турбинных электрогенерирующих установок. Эти данные содержат 9 факторов, соответствующих наиболее важным показателям состояния оборудования. Записи во временных рядах произведены с интервалом в 1 минуту. Всего в данных насчитывается 5 отказов, что делает их подходящими для проверки эффективности предложенного метода.

Формирование выборки для каждого из отказов производилось следующим образом: взято M измерений перед отказом в момент $T_{\text{отк}}$, составляющих интервал $(T_{\text{отк}} - M; T_{\text{отк}}]$. Для каждой точки $i \in (T_{\text{отк}} - M; T_{\text{отк}}]$ был рассчитан RUL по формуле (2) в качестве выходных значений. В качестве входных значений для каждой точки i был сформирован по формуле (4) двумерный массив данных размерностью $w \times 9$. В результате был подготовлен набор, содержащий $5M$ (по количеству отказов в выборке) или 750 образцов исторических данных.

Для сравнения были взяты примеры классических алгоритмов и алгоритмов глубокого обучения, которые были определены как наиболее точные [4, 16, 23] для прогнозирования остаточного ресурса оборудования.

В результате экспериментов были определены параметры, при которых алгоритмы дают наибольшую точность. Параметры

$M = 150$, $w = 25$, задающие размерность и форму входных данных, а также структура нейронной сети, приведены на рис. 3.

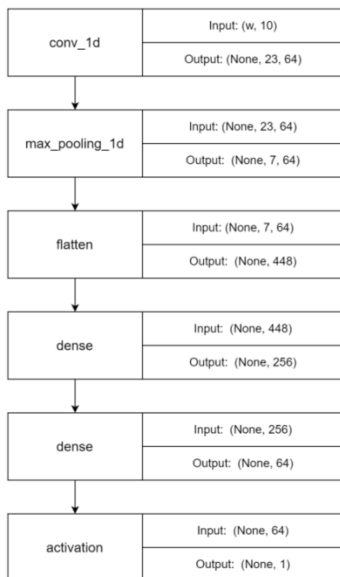


Рис. 3. Структура нейронной сети CNN

Для оценки эффективности работы методов использовалась метрика RMSE[20]. Для расчета среднеквадратичной ошибки используется следующая формула

$$(6) \quad RMSE = \sqrt{\frac{\sum_{t=1}^n (\bar{y}_t - y_t)^2}{n}},$$

где n – размер выборки, \bar{y}_t – спрогнозированное значение, y_t – реальное значение, полученное в результате наблюдений. Метод считается тем точнее, чем меньше для результата прогнозирования модели ошибка RMSE.

Для тестирования использовалась техника кросс-валидации, с разделением данных в соотношении 3/1/1 для обучающей, тестовой и верификационной выборки. Полученные результаты были усреднены. Результаты тестирования приведены в таблице 2.

Таблица 2. Результаты тестирования

Метод	Значение RMSE
Random Forest	30,54
KNN	44,91
SVM	39,85
LSTM	33,74
RNN	36,02
CNN	28,67
Предлагаемый метод	23,79

На рис. 4 представлен пример прогноза остаточного ресурса оборудования, построенного с использованием предложенного метода. Пропуски в данных означают, что для данной точки не удалось спрогнозировать RUL с достаточной точностью.

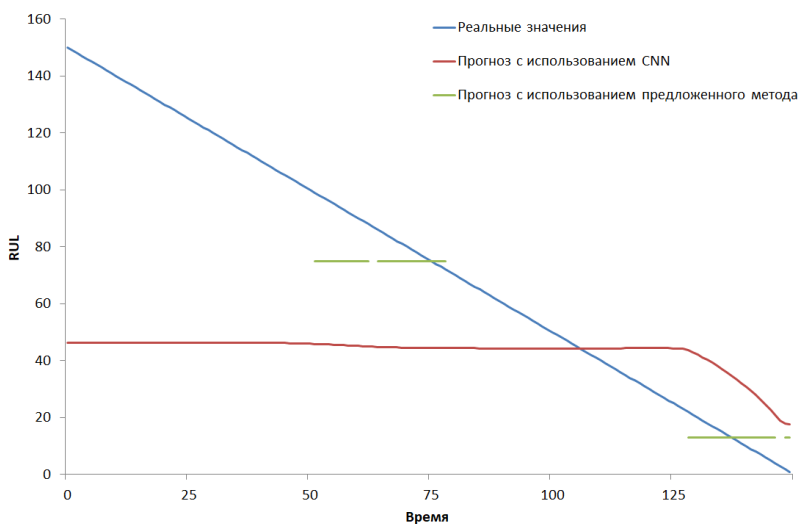


Рис. 4. Пример прогноза, построенного с использованием предложенного метода

5. Заключение

В данной работе предложен и рассмотрен новый метод прогнозирования остаточного ресурса оборудования. Данный метод

предложен для ситуаций с небольшим количеством отказов, где нет возможности построить точную модель существующими способами. Для решения этой проблемы предложено заменить выходные данные алгоритма прогнозирования сигнальной системой, при которой в зависимости от результатов анализа входных данных и при срабатывании проверок будет выдаваться один из predetermined сигналов с оцениваемым остаточным ресурсом. Данные изменения позволяют в отдельных сценариях повысить точность прогнозирования остаточного ресурса оборудования.

Литература

1. ГУНИНА И.А., ШКАРУПЕТА Е.В., РЕШЕТОВ В.В. *Прорывное технологическое развитие промышленных комплексов в условиях цифровой трансформации* // Инновационные кластеры цифровой экономики: теория и практика. – 2018. – С. 535–554.
2. КУПРИЯНОВСКИЙ В.П., НАМИОТ Д.Е., ДРОЖЖИНОВ В.И., КУПРИЯНОВСКАЯ Ю.В., ИВАНОВ М.О. *Интернет вещей на промышленных предприятиях* // International Journal of Open Information Technologies. – 2016. – №12. – С. 156–161.
3. ЛЫСЕНКО С.В., ТЕН Э.В. *Об оценке остаточного ресурса башенных кранов* // Проблемы современной науки и образования. – 2016. – №1. – С. 98–102.
4. САЙ ВАН КВОНГ, ЩЕРБАКОВ М.В. *Метод прогнозирования остаточного ресурса на основе обработки данных многообъектных сложных систем* // Прикаспийский журнал: управление и высокие технологии. – 2019. – №1(45). – С. 33–44.
5. ШИЛОВА Е.В., ДЬЯКОВ А.Р. *О феномене четвертой промышленной революции и его влиянии на экономику и управление* // Вестник Прикамского социального института. – 2018. – №3(81). – С. 86–95.
6. *A Grammar Of Data Manipulation – Dplyr.* – URL: <https://dplyr.tidyverse.org> (дата обращения: 03.04.2022).

7. CERNUDA C. *On the relevance of preprocessing in predictive maintenance for dynamic systems* // In: Lughofer E., Sayed-Mouchaweh M. (Eds.). *Predictive Maintenance in Dynamic Systems*. Cham, Springer. – 2019. – P. 53–93.
8. CHENG J.C., CHEN W., CHEN K., WANG Q. *Data-driven predictive maintenance planning framework for MEP components based on BIM and IoT using machine learning algorithms* // *Automation in Construction*. – 2020. – Vol. 112. – P. 1– 21.
9. CHE-SHENG H., JEHN-RUEY J. *Remaining useful life estimation using long short-term memory deep learning* // *IEEE Int. Conf. on Applied System Invention (ICASI-2018)*, Chiba, Japan. – 2018. – P. 58–61.
10. CRAN – *Packages*. – URL: <https://cran.r-project.org/web/packages/> (дата обращения: 03.04.2022).
11. *Forecast Package – RDocumentation*. – URL: <https://www.rdocumentation.org/packages/forecast> (дата обращения: 03.04.2022).
12. GOEBEL K., SAHA B., SAXENA A., CELAYA J.R., CHRISTOPHERSEN J.P. *Prognostics in battery health management* // *IEEE Instrumentation & Measurement Magazine*. – 2008. – Vol. 8. – P. 33–40.
13. HUANG Z.Y., XU Z.G., WANG W.H., SUN Y.X. *Remaining usefullife prediction for a nonlinear heterogeneous Wiener process model withan adaptive drift* // *IEEE Trans. Rel.* – June, 2015. – Vol. 64, No. 2. – P. 687–700.
14. *Identify and replace outliers in time series*. – URL: <https://pkg.robjhyndman.com/forecast/reference/tsoutliers.html> (дата обращения: 03.04.2022).
15. LEI Y., LIN., GONTARZ S., LIN J., RADKOWSKI S., DYBALA J. *A Model-Based Method for Remaining Useful Life Prediction of Machinery* // *IEEE Trans. on Reliability*. – 2017. – Vol. 65. – P. 1314–1326.
16. LI X., DING Q., SUN J. *Remaining useful life estimation in prognostics using deep convolution neural networks* // *Reliab. Eng. Syst. Saf.* – 2018. – Vol. 172. – P. 1–11.
17. LIU J., WANG W., MA F., YANG Y.B., YANG C.S. *A data-model-fusionprognostic framework for dynamic system state*

- forecasting* // Eng. Appl. Artif. Intell. – June, 2012. – Vol. 25, No. 4. – P. 814–823.
18. PATIL S., PATIL A., HANDIKHERKAR V., DESAI S., PHALLE V. M., KAZI F. S. *Remaining Useful Life (RUL) Prediction of Rolling Element Bearing Using Random Forest and Gradient Boosting Technique* // ASME 2018 International Mechanical Engineering Congress and Exposition. – 2018. – P. 1–7.
 19. RAN Y., ZHOU X., LIN P., WEN Y., DENG R. *A Survey of Predictive Maintenance: Systems, Purposes and Approaches* // IEEE communications Surveys & Tutorials. – 2019. – P. 1–36.
 20. SHCHERBAKOV M.V. *A Survey of Forecast Error Measures* // World Applied Sciences Journal (WASJ). – 2013. – No. 24. – P. 171–176.
 21. XIONGZI C., JINSONG Y., DIYIN T. YINGXUN W. *Remaining useful life prognostic estimation for aircraft subsystems or components: A review* // 10th IEEE Int. Conf. on Electronic Measurement & Instruments (ICEMI). – 2011. – Vol. 2. – P. 94.
 22. YAN M., WANG X., WANG B., CHANG M., MUHAMMAD I. *Bearing remaining useful life prediction using support vector machine and hybrid degradation tracking model* // ISA Trans. – 2020. – Vol. 98. – P. 471–482.
 23. YU J. *Remaining useful life prediction for lithium-ion batteries using a quantum particle swarm optimization-based particle filter* // Quality Engineering. – 2017. – Vol 29. – P. 536–546.
 24. ZHANG Z., SI X., HU C., LEY Y. *Degradation data analysis and remaining useful life estimation: A review on Wiener-process-based methods* // European Journal of Operational Research. – 2018. – Vol. 271(3). – P. 775–796.

FORECASTING OF THE REMAINING USEFUL LIFE IN CONDITIONS OF SMALL DATA SAMPLE

Konstantin Zadiran, Volgograd state technical university, Volgograd, postgraduate student (konstantin.zadiran@gmail.com).

Maxim Shcherbakov, Volgograd state technical university, Volgograd, Doctor of Science, assistant professor (maxim.shcherbakov@vstu.ru).

Cuong Sai, Volgograd state technical university, Volgograd, Candidate of Science (svcuonghvtqs@gmail.com).

Abstract: In the article a method for forecasting the residual life of equipment using deep learning is proposed. The method is applicable in cases with a small amount of information about data failures, where existing classical methods may not provide the required accuracy. The process of maintaining the equipment in working condition is one of the most important processes in the operation of the equipment. At the same time, the maintenance process often suffers from inefficiency. Therefore, forecasting methods were developed, on the basis of which the concept of proactive maintenance process management was built, which allows optimizing the structure and costs of equipment management throughout the life cycle. However, these methods may show insufficient accuracy if there is not enough data to train them, for example, due to the rarity of equipment failures. To solve this problem, a new prediction method based on deep learning is proposed that can improve the prediction accuracy. In this method, the continuous prediction of the remaining service life over the entire interval is replaced by a model for generating signals containing the calculated prediction.

Keywords: machine learning, remaining useful life, proactive maintenance.

УДК 004.9

ББК 32.972

DOI: 10.25728/ubs.2023.102.6

*Статья представлена к публикации
членом редакционной коллегии Я.И. Квинто.*

Поступила в редакцию 20.07.2022.

Опубликована 31.03.2023.