

Early Fault Detection with Multi-target Neural Networks

Angela Meyer(✉)

Bern University of Applied Sciences, 2501 Biel, Switzerland
angela.meyer@bfh.ch

Abstract. Wind power is seeing a strong growth around the world. At the same time, shrinking profit margins in the energy markets let wind farm managers explore options for cost reductions in the turbine operation and maintenance. Sensor-based condition monitoring facilitates remote diagnostics of turbine subsystems, enabling faster responses when unforeseen maintenance is required. Condition monitoring with data from the turbines' supervisory control and data acquisition (SCADA) systems was proposed and SCADA-based fault detection and diagnosis approaches introduced based on single-task normal operation models of turbine state variables. As the number of SCADA channels has grown strongly, thousands of independent single-target models are in place today for monitoring a single turbine. Multi-target learning was recently proposed to limit the number of models. This study applied multi-target neural networks to the task of early fault detection in drive-train components. The accuracy and delay of detecting gear bearing faults were compared to state-of-the-art single-target approaches. We found that multi-target multi-layer perceptrons (MLPs) detected faults at least as early and in many cases earlier than single-target MLPs. The multi-target MLPs could detect faults up to several days earlier than the single-target models. This can deliver a significant advantage in the planning and performance of maintenance work. At the same time, the multi-target MLPs achieved the same level of prediction stability.

Keywords: Condition monitoring · Fault detection · Multi-target neural networks · Normal behaviour models · Wind turbines

1 Introduction

The global wind power capacity is growing strongly with a total installed volume of 651 GW in 2019 and an increase of 76 GW in 2020 [1]. The newly installed wind turbines are getting larger and increasingly more complex. At the same time, the operating cost of wind farms still makes up a major fraction, approximately 30%, of their lifetime cost [2]. Major faults can result in days and even weeks of downtime [3, 4]. Therefore, they can substantially reduce the owner's return on investment and pose a considerable economic risk. As a result, many operators want to closely monitor the health state of their turbines in order to be alerted as early as possible of any developing technical problems and to prevent any major damage and downtime. To this end, an automated condition

© Springer Nature Switzerland AG 2021
O. Gervasi et al. (Eds.): ICCSA 2021, LNCS 12951, pp. 429–437, 2021.
https://doi.org/10.1007/978-3-030-86970-0_30

monitoring of wind turbine subsystems provides an essential prerequisite for informed operational decision making and fast responses in case of unforeseen maintenance needs [5, 6].

Data-driven automated monitoring methods have been proposed, amongst others, based on sensor data logged in the turbines' supervisory control and data acquisition (SCADA) systems [7–10]. Temperature can be an important indicator of different types of developing machine problems such as mechanical faults which can give rise to excessive friction generating heat. Therefore, a major focus of the proposed SCADA-based condition monitoring approaches is the temperature-based detection of developing faults in the wind turbine subsystems based on models of the turbine's normal operation behaviour in the absence of operational faults [11–22]. The present study focusses on the gear bearing temperature as an indicator of developing gearbox faults.

The goal of this study is to assess the potential of multi-target regression models for the automated SCADA-based fault detection. Specifically, this work has investigated and compared the delays in detecting gear bearing faults using single-target versus multi-target models of the turbines' normal operation. Moreover, the stability of the alarm signal after the first detection of a developing fault is being assessed.

The remainder of this paper is structured as follows. Section 2 provides a brief overview of previous work in this field. Section 3 describes the data sources and the training and testing of the multi- and the single target regression models. The analysis and results are discussed in Sect. 4. Conclusions and possible future work are proposed in Sect. 5.

2 Related Work

Normal behaviour modelling has become an established technique in wind turbine condition monitoring and fault detection [7, 8]. Normal behaviour models characterize the machine state during normal operation in the absence of faults. They have been in use for monitoring the health state of turbine subsystems such as the gearbox [23, 24] and the generator [25]. Normal behaviour models have also successfully been employed for monitoring the active power generation [15, 26, 27]. We refer to [8] for a comprehensive review of SCADA-based condition monitoring and normal behaviour models of wind turbines.

Multi-target machine learning models [28–32] are regression or classification models which predict multiple target variables simultaneously. It has been demonstrated in other fields that multi-target models hold the potential to enable an increased prediction accuracy compared to single-target models and are less susceptible to overfitting the training data [29, 33–34]. In the field of wind turbine monitoring, we have recently introduced multi-target regression models for simultaneously monitoring the growing number of SCADA channels, and we demonstrated that they can reduce the effort of SCADA-based normal behaviour monitoring in wind turbine condition monitoring [35].

3 Data and Methods

Condition monitoring data from the SCADA system of three commercial onshore turbines was analyzed in this work. The turbines are variable-speed three-bladed horizontal

axis systems with pitch regulation from an onshore wind farm. Their rated power was 3.3 MW, and they operated with a 3-stage planetary/helical gearbox. The turbines' rotors were 112 m in diameter with the hub located at 84 m height above ground. The turbines' cut-in, rated and cut-out wind speeds were specified at 3 m/s, 13 m/s and 25 m/s, respectively.

In this study, fourteen months of ten-minute mean SCADA signals served to train and test the models specified below. The data were anonymized to maintain the privacy of the wind farm operator. We report the results for one of the wind turbines. It was randomly selected and the results were not affected by the choice of turbine. We focus on monitoring the gear bearing condition based on the temperature of the gear bearing. The temperature is an important SCADA-based indicator of incipient fault processes in gearbox components [8, 19, 20, 22, 23]. In the present study, the condition of the gear bearing has been monitored based on two normal operation models of the bearing temperature. Wind speed v_{wind}, wind direction α_{wind} and air temperature T_{air} constitute the models' input variables which were provided as ten-minute averages of measurements from nacelle-mounted anemometers and thermometers. The input variables were selected due to their relevance for explaining and predicting the behaviour of the target variables.

A multi-target fully connected feedforward neural network was designed to predict the gear bearing temperature T_{gear} along with the hydraulic oil temperature T_{oil} and the transformer winding temperature T_{tr} from the input variables at high accuracy, T_{gear}, T_{oil}, $T_{tr} \sim v_{wind} + \alpha_{wind} + T_{air}$. The single-target model estimates the gear bearing temperature only, $T_{gear} \sim v_{wind} + \alpha_{wind} + T_{air}$. In addition, the two fully connected feedforward neural networks (multi-layer perceptrons, MLP) were trained and tested to assess the normal operating behaviour of the gear bearing based on the provided SCADA data. The model architectures were developed to obtain a high predictive accuracy on the training set without overfitting the training data. In this process, the number of neurons and weights to be trained was increased only if this resulted in higher predictive accuracy. The resulting model architectures are detailed in Table 1.

In this work, our goal is to systematically assess the ability of multi-target neural networks to detect developing faults that result in rising component temperatures. We demonstrate this approach by the example of the gear bearing temperature. However, it is equally applicable to faults in other subsystems and other components that result in elevated SCADA-logged temperatures.

A major challenge in data-driven fault detection and isolation is the scarcity of actually observed fault instances. We addressed this point by combining gear bearing temperature measurements with a multitude of synthetic temperature trends in order to mimic the bearing temperature rise induced by a developing fault. To this end, a synthetic temperature trend was overlaid on the normalized gear bearing temperature signal. One of ten different linear temperature trends was added to the normalized bearing temperature. Temperature trends with integer slopes in the range of 1 to 10 were used to simulate slowly and fast evolving fault processes. The temperature trend onset time was randomly sampled from a two-week time window in months 12 and 13 of the 14-months observation period. Fifty different onset times have been randomly sampled from the

two-week window for each of the ten temperature slopes. This ensured that the results did not depend on the choice of the trend onset time.

Table 1. The architectures of the multi-layer perceptrons.

Model	Architecture
Multi-layer perceptron (MLP) with three target variables	Two dense hidden layers with 4 neurons in the first layer and 19 neurons in the second layer, batch normalization, and a 3-neuron output layer. Dropout was applied at a rate of 10% to avoid overfitting
Single-target multi-layer perceptron	Three dense hidden layers with 4 neurons each in the first and second layers and 5 neurons in the third layer, batch normalization and a single node in the output layer. Dropout rate of 10%

4 Results and Discussion

Two common alarm criteria [7] were applied using the residuals of the gear bearing temperature. The residuals were computed as the difference of the observed temperature of the gear bearing and its temperature predicted by the normal operation behaviour models of Table 1. Figure 1 illustrates the residuals and alarms. According to the first criterion, an alarm was raised when the 99.9th percentile of the residuals distribution was exceeded for more than 8 h in the past 24 h. On the other hand, the second criterion triggered an

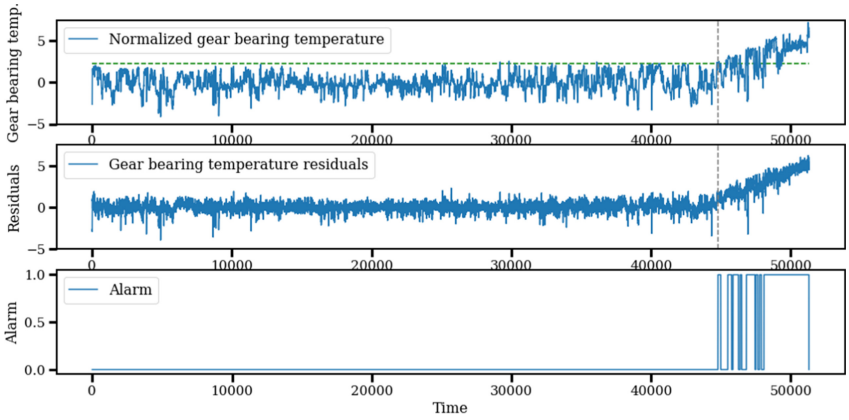

Fig. 1. Normalized gear bearing temperature and residuals. An alarm was raised when the 99.9th percentile of the residuals distribution was exceeded for more than 8 h during the past 24 h (alarm criterion 1). The dashed line indicates the first alarm event. The temperature units are dimensionless due to normalization.

alarm whenever the rolling mean of the residuals computed over the past 8 h exceeded the 99.9th percentile of the residuals distribution. The fault detection capabilities of the multi-target and single-target normal behaviour models were compared with regard to the delay of detecting the induced faults. Moreover, the stability of the triggered alarms was assessed following the detection. The stability is computed as the fraction of true positive alarms after the first detection.

Figure 2 shows the resulting alarms for a fault instance with fast rising bearing temperature and a randomly sampled onset time. The false positive detection rate is zero for both criteria but there are false negatives after the first alarm detection by both alarm criteria. Following the first detection, the alarms are initially unstable and stabilize within three to seven days after the first alarm was triggered, as shown in Fig. 2. We found that the alarm signal stabilized faster in the case of the first criterion, while the second criterion generated a less stable signal that required a week to stabilize in the case shown in Fig. 2. This alarm instability was caused by the signal-to-noise ratio which was relatively small immediately after the first detection and increased over time along with the overlaid increasing temperature.

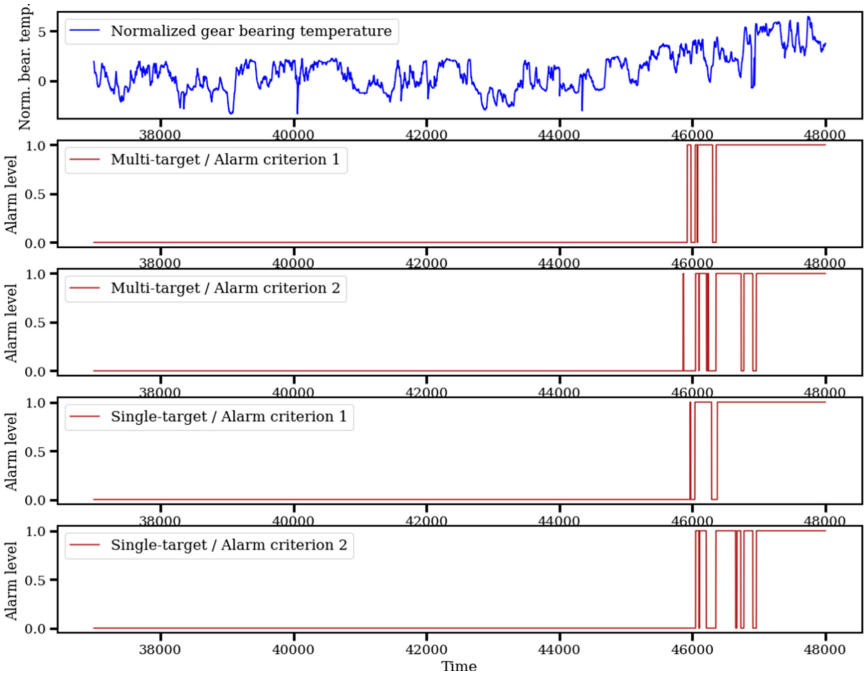

Fig. 2. A fast evolving gear bearing fault with trend slope 10 is shown at one of the 50 randomly sampled onset times. The gear bearing temperature starts to rise near time step 44000 in this case. The alarm criteria are compared based on the multi-target and the single-target normal behaviour models of the gear bearing temperature. Alarm criterion 2 detects the trend earlier but is less stable than alarm criterion 1.

The detection delays and detection stabilities are reported in Figs. 3 and 4 in terms of the means and the standard deviations computed over the 50 randomly sampled onset times. The fault detection delays could be quantified precisely as the faults were induced at known times in terms of temperature trend onset times.

We found that the multi-target model could detect the induced faults at least as fast as the single-target model. In the majority of the fault instances in this case study, the multi-target MLP even enabled a somewhat faster detection of the gear bearing fault than the single-target network. As shown in Figs. 3 and 4, the multi-target MLP facilitated shorter detection delays for both slow and fast trends and regardless of the chosen alarm criterion.

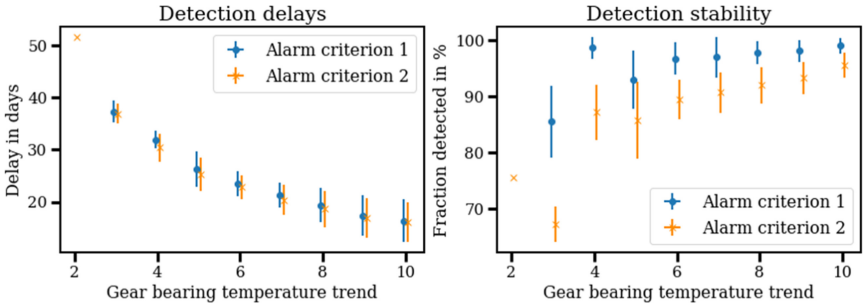

Fig. 3. Fault detection delay and the detection stability are reported based on *multi-target* normal behavior MLP of the bearing temperature. The stability is the fraction of true positives from the first alarm until the end of the observation period, i.e. till the end of month 14. Mean and standard deviation of the detection delay and stability are reported for ten gear temperature trends. For each trend, the mean and standard deviations of detection delay and stability were computed from 50 randomly sampled temperature trend onset times in months 12 and 13. Mean detection delay and mean detection stability were computed based on randomized temperature trend onsets times and for 10 trend progression velocities, assuming a linear temperature trend. Fifty randomly selected onset times were used to simulate the bearing temperature increase.

This study demonstrates that multi-target models can describe the temperature behaviour of wind turbine gear bearings in normal operation at least as accurately as corresponding single-target models. In the present work, the multi-target MLP provided a test-set accuracy of 0.49 for predictions of the gear bearing temperature, while the test-set accuracy of the single-target MLP was 0.51 in normalized temperature. As the multi-target MLP produced smaller prediction residuals in many cases, temperature trends became visible earlier and thus the detection delay was shorter. In addition, a paired sample t-test was performed to test the null hypothesis that there are no systematic differences in the detection delays based on the normal behaviour descriptions of the multi-target MLP versus the single-target MLP. The null hypothesis was clearly rejected ($p < 10^{-21}$) in favor of the alternative hypothesis of shorter detection delays based on the multi-target MLP. The test result was confirmed for both alarm criteria.

Similarly, a paired sample t-test was performed to test for systematic differences of the detection stability. This provided a more ambiguous picture. Based on alarm criterion 2, the multi-target MLP enabled a more stable alarm after the first detection (p

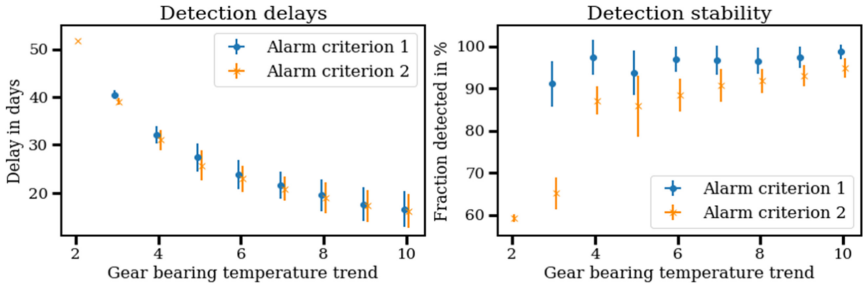

Fig. 4. Fault detection delay and stability as in Fig. 3 but based on the *single-target* normal operation MLP of the gear bearing temperature.

$< 10^{-14}$). However, no systematic difference in the detection stability was found in the case of alarm criterion 1 ($p = 0.22$) Comparing both alarm criteria, it is also found that the second criterion detects faults significantly earlier than the first criterion, but this comes at the cost of significantly reduced detection stabilities (Figs. 3 and 4).

In summary, this study demonstrated that multi-target MLPs can detect faults as fast as and in some cases even earlier than single-target MLPs, and at the same time achieve the same level of detection stability. Comparing Figs. 3 and 4, the detection speed-up observed in this study ranged from several hours to several days. The earlier detection enabled by the proposed multi-target approach can deliver a significant advantage in the planning and performance of maintenance activities. If wind farm operators learn about a developing fault hours or up to several days ahead, they have more time to respond and schedule inspections and adjustment work that may prevent more serious damage and component replacement.

5 Conclusions

This study investigated how multi-target neural networks compare to single-target models with regard to the speed and accuracy of detecting incipient faults in wind turbines from SCADA data. We analyzed the delays in detecting gear bearing faults in onshore wind turbines. Gear bearing faults can result in anomalous temperature increases that are detectable from SCADA data. In this work, synthetic temperature trends were overlaid on the gear bearing temperature SCADA signals in order to facilitate a systematic analysis despite the scarcity of fault observations. We compared the detection delay and the detection stability of the alarm signals among the multi- and single-target models based on different alarm criteria. In this study, we found that multi-target neural networks can meet and even go below the detection delays achieved with single-target MLPs of the turbine normal operation. At the same time, the multi-target MLPs achieved the same level of prediction stability. We demonstrated that the detection of temperature-related faults could be accelerated by up to several days compared to the state-of-the-art fault detection with single-target models. With regard to future studies, we propose to also investigate the potential of multi-target models for normal behaviour modelling and fault detection tasks based on high-frequency data, in particular from vibration measurements in the drive train, which did not form part of this paper.

Acknowledgments. The author thanks Bernhard Brodbeck, Janine Maron, Dimitrios Anagnostos of WinJi AG, Switzerland, and Kaan Duran of Energie Baden-Wuerttemberg EnBW, Germany, for valuable discussions.

References

1. Global Wind Energy Council GWEC: Global Wind Report 2019 (2019). https://gwec.net/global-wind-report-2019. Accessed 5 June 2021
2. International Renewable Energy Agency: Renewable power generation costs in 2019 (2019). https://www.irena.org. Accessed 5 June 2021
3. Faulstich, S., Hahn, B., Tavner, P.: Wind turbine downtime and its importance for offshore deployment (2011). https://doi.org/10.1002/we.421
4. Pfaffel, S., Faulstich, S., Rohrig, K.: Performance and reliability of wind turbines: a review. Energies (2017). https://doi.org/10.3390/en10111904
5. Garcia Marquez, F., Tobias, A., Pinar Perez, J., Papaelias, M.: Condition monitoring of wind turbines: techniques and methods. Renew. Energy **46**, 169–178 (2012). https://doi.org/10.1016/j.renene.2012.03.003
6. Fischer, K., Coronado, D.: Condition monitoring of wind turbines: state of the art, user experience and recommendations. VGB PowerTech **7**, 51–56 (2015)
7. Stetco, A., et al.: Machine learning methods for wind turbine condition monitoring: a review. Renew. Energy **133**, 620–635 (2019). https://doi.org/10.1016/j.renene.2018.10.047
8. Tautz-Weinert, J., Watson, S.: Using SCADA data for wind turbine condition monitoring - a review. IET Renew. Pow. Gener. **11**(4), 382–394 (2017). https://doi.org/10.1049/iet-rpg.2016.0248
9. Helbing, G., Ritter, M.: Deep learning for fault detection in wind turbines. Renew. Sustain. Energy Rev. **98**, 189–198 (2018). https://doi.org/10.1016/j.rser.2018.09.012
10. Qiao, W., Lu, D.: A survey on wind turbine condition monitoring and fault diagnosis—part II: signals and signal processing methods. IEEE Trans. Industr. Electron. **62**, 6546–6557 (2015). https://doi.org/10.1109/TIE.2015.2422394
11. Zaher, A., McArthur, S., Infield, D., Patel, Y.: Online wind turbine fault detection through automated SCADA data analysis. Wind Energy **12**, 574–593 (2009). https://doi.org/10.1002/we.319
12. Kusiak, A., Verma, A.: Analyzing bearing faults in wind turbines: a data-mining approach. Renew. Energy **48**, 110–116 (2012). https://doi.org/10.1016/j.renene.2012.04.020
13. Salameh, J., Cauet, S., Etien, E., Sakout, A., Rambault, L.: Gearbox condition monitoring in wind turbines: a review. Mech. Syst. Sig. Process. **111**, 251–264 (2018). https://doi.org/10.1016/j.ymssp.2018.03.052
14. McKinnon, C., Turnbull, A., Koukoura, S., Carroll, J., McDonald, A.: Effect of time history on normal behaviour modelling using SCADA data to predict wind turbine failures. Energies (2020). https://doi.org/10.3390/en13184745
15. Meyer, A., Brodbeck, B.: Data-driven performance fault detection in commercial wind turbines. In: Proceedings of the 5th European Conference of the Prognostics and Health Management Society (PHME20) (2020). ISBN 978-1-93-626332-5
16. Liu, Y., Wu, Z., Wang, X.: Research on fault diagnosis of wind turbine based on SCADA data. IEEE Access (2020). https://doi.org/10.1109/ACCESS.2020.3029435
17. Zhang, S., Lang, Z.: SCADA-data-based wind turbine fault detection: a dynamic model sensor method. Control. Eng. Pract. (2020). https://doi.org/10.1016/j.conengprac.2020.104546

18. Liu, X., Du, J., Ye, Z.: A condition monitoring and fault isolation system for wind turbine based on SCADA data. IEEE Trans. Industr. Inf. (2021). https://doi.org/10.1109/TII.2021.3075239
19. Yang, Y., Liu, A., Wang, J.: Fault early warning of wind turbine gearbox based on multi-input support vector regression and improved ant lion optimization. Wind Energy (2021). https://doi.org/10.1002/we.2604
20. Astolfi, D., Scappaticci, L., Terzi, L.: Fault diagnosis of wind turbine gearboxes through temperature and vibration data. Int. J. Renew. Energy Res. (2017)
21. Leahy, K., Hu, R., O'Sullivan, D.: Diagnosing and predicting wind turbine faults from SCADA data using support vector machines. Int. J. Prognost. Health Manag. (2018)
22. Zeng, X., Yang, M., Bo, Y.: Gearbox oil temperature anomaly detection for wind turbine based on sparse Bayesian probability estimation. Int. J. Electr. Pow. Energy Syst. (2020)
23. Wang, Y., Infield, D.: Supervisory control and data acquisition data-based non-linear state estimation technique for wind turbine gearbox condition monitoring. IET Renew. Pow. Gener. **7**, 350–358 (2013). https://doi.org/10.1049/iet-rpg.2012.0215
24. Wang, L., Zhang, Z., Long, H., Xu, J., Liu, R.: Wind turbine gearbox failure identification with deep neural networks. IEEE Trans. Industr. Inf. **13**, 1360–1368 (2017). https://doi.org/10.1109/TII.2016.2607179
25. Guo, P., Infield, D.: Wind turbine generator condition monitoring using temperature trend analysis (2012). https://doi.org/10.1109/TSTE.2011.2163430
26. Kusiak, A., Zheng, H., Song, Z.: Online monitoring of power curves. Renew. Energy **34**, 1487–1493 (2009). https://doi.org/10.1016/j.renene.2008.10.022
27. Schlechtingen, M., Santos, F., Achiche, S.: Using data-mining approaches for wind turbine power curve monitoring: a comparative study. IEEE Trans. Sustain. Energy **4**, 671–679 (2013). https://doi.org/10.1109/TSTE.2013.2241797
28. Caruana, R.: Multitask learning. Mach. Learn. **28**, 41–75 (1997). https://doi.org/10.1023/A:1007379606734
29. Piccart, B.: Algorithms for multi-target learning. Doctoral thesis, KU Leuven (2012)
30. Borchani, H., Varando, G., Bielza, C., Larranaga, P.: A survey on multi-output regression. Wiley Interdiscipl. Rev. Data Min. Knowl. Discov. (2015). https://doi.org/10.1002/widm.1157
31. Waegeman, W., Dembczyński, K., Hüllermeier, E.: Multi-target prediction: a unifying view on problems and methods. Data Min. Knowl. Disc. **33**(2), 293–324 (2018). https://doi.org/10.1007/s10618-018-0595-5
32. Xu, D., Shi, Y., Tsang, I., Ong, Y., Gong, C., Shen, X.: Survey on multi-output learning. IEEE Trans. Neural Netw. Learn. Syst. **31**, 2409–2429 (2020). https://doi.org/10.1109/TNNLS.2019.2945133
33. Kocev, D., Dzeroski, S., White, M., Newell, G., Griffioen, P.: Using single- and multi-target regression trees and ensembles to model a compound index of vegetation condition. Ecol. Model. **220**(8), 1159–1168 (2009)
34. Segal, M., Xiao, Y.: Multivariate random forests. Wiley Interdiscipl. Rev. Data Min. Knowl. Discov. (2011). https://doi.org/10.1002/widm.12
35. Maron, J., Anagnostos, D., Brodbeck, B., Meyer, A.: Gear bearing fault detection using multi-target neural networks. In: Wind Energy Science Conference WESC 2021. Hanover, Germany, May 2021