Predictive maintenance has been considered as an effective strategy to reduce operations and maintenance (O&M) costs and improve the availability and efficiency of wind turbines through condition monitoring (CM) and fault detection, prediction and diagnosis. Wind turbine monitoring using data collected by the supervisory control and data acquisition (SCADA) systems has been seen as a cost-effective and wide-scale approach. As a result, much research has employed SCADA data to develop reliable, efficient and cost-effective monitoring systems in recent years. Most solutions are based on statistical methods and machine learning (ML) techniques. However, these existing approaches face three major challenges.

The first major challenge involves a common practice of the statistical and ML-based methods regarding the use of normal behaviour models (NBMs) developed for specific turbine components. NBMs are based on the idea of modelling the normal behaviour of critical components using historical SCADA data collected while the wind turbine was operating in healthy condition. Then, the detection of faults are based on the analysis of deviations between the predicted model outputs and actual measured values. However, since each wind turbine has many critical components that need to be monitored, NBM-based approach would require users to build and preserve a large number of NBMs for any wind farm. Moreover, due to many unavoidable reasons, e.g. wind turbine ageing, subsystem replacements, software updates, or sensor recalibration, NBMs are supposed to vary with time. Hence, one would raise many concerns about the updating issues of NBMs, e.g. when or how often a model should be updated to assure the homogeneity between the model and the turbine subsystem. These problems lead to the circumstance that it is not easy for wind farm operators to make use of this monitoring approach.

The second major challenge relates to the fact that many statistical approaches for wind turbine condition monitoring were developed on the basis of parametric models. In addition, it is argued that deep learning models can generally be viewed as parametric methods. When using parametric statistical models, it is often assumed that the analysed SCADA data follow a normal distribution. However, if the data are not normally distributed and not transformed to approximate to a normal distribution, the use of parametric approaches to build NBMs might not be appropriate and could diminish the fault detection process.

The third major challenge links to the ML-based fault detection algorithms, which rely on the turbine's NBMs trained on the SCADA data from healthy wind turbines. Training those models requires large amounts of normal operation data collected over a long period covering a representative range of operating conditions. Certainly, when the data are scarce or when they are no longer representative of the turbine's current operation behaviour, fault detection based on NBMs is usually not feasible because NBMs cannot be trained properly. This is the case for newly installed wind turbines at the initial stage of their operation life when the amount of data accumulated by these turbines is small, which cannot provide sufficient information for training NBMs. Also, regardless of turbine age, major changes in the operation conditions, such as after software updates or hardware replacements, can result in a lack of SCADA data representative of the new operation behaviour.

Aim of this research project is therefore to tackle these challenges through:

- investigating data-driven solutions for condition assessment of wind turbines without using NBMs;
- studying nonparametric statistical approaches for wind turbine condition monitoring and automated fault detection;
- exploring the applicability of transfer learning techniques for transferring (or reusing) NBMs, which are trained for a wind turbine (i.e. the source turbine) with sufficient training data, to other wind turbines (i.e. the target turbines) with scarce or limited operation data.

Specifically, this project aims to develop a non-classical integrated framework, comprising of three new condition monitoring solutions, which can effectively monitor the operation state of wind turbines and reliably detect their faults in the gearbox and generator at the early stage. The development of these solutions is referred to as research tasks:

1. Condition assessment of wind turbines without using NBMs.
2. Nonparametric statistical methods for wind turbine condition monitoring and fault detection.
3. Transfer learning methods for wind turbine monitoring and fault detection using deep learning.

The results of this research project will be a set of procedures and algorithms for condition monitoring and automated fault detection of wind turbines. The focus is on identifying faults in the gearbox and generator. Three real SCADA data sets will be used to validate the proposed methods. The data sets were recorded from operating wind turbines and involve several fault events. The developed procedures and algorithms will improve the performance of condition-based maintenance processes through improving the reliability and performance of wind turbines, so that the efficiency and profitability of a wind farm can be maintained as much as possible throughout its lifetime.