Gaussian processes for forecasting, performance and condition monitoring of windfarms considering uncertainty

I. Antoniadou[1], S. Kiartzis

Hellenic Petroleum Renewable Energy Sources, Chimarras 8A, 15125 Marousi, Athens

[1]IAntoniadou@helpe.gr

Abstract. The focus on renewable energy sources and wind energy in particular has been continuously increasing over the past couple of decades, in an attempt to minimize greenhouse emissions in accordance to ongoing energy policy agreements at an international level. Big part of the current R&D in the area is focused on increasing wind energy competitiveness in the energy market. Some of the challenges that tend to be encountered and need to be further addressed include the necessity to improve wind turbine systems designs by taking into account the harsh environmental and operational conditions commonly encountered in wind farms, especially offshore, but also the need to minimize maintenance and operation costs and to optimize performance based on weather conditions and energy demands. Although there exist different approaches to solving the latter issue, one current trend lately examined is the potential use of recent advances in the area of data analysis (machine learning, statistical and signal processing algorithms) that could be used to effectively analyse Supervisory Control and Data Acquisition (SCADA), vibration and other types of data recorded during the operation of a wind turbine in order to give reliable predictions of performance or condition of the individual wind turbines and/or the wind farm as a whole. Certain challenges arise in this framework, since in most cases lack of reliability of data and complexities in the various processes taking place in wind turbine systems may result in uncertain predictions, that should be taken into account for robust, meaningful, automated inspection. This paper examines, the use of Gaussian processes as a data modelling tool of wind turbine systems that allows the consideration of uncertainty. We demonstrate a couple of examples of how Gaussian process models could be applied in practice on vibration and/or SCADA wind turbine data to achieve non-linear regression by clever handling of the models hyper-parameters.

Keywords. Wind turbine monitoring, forecasting, SCADA data analysis, Gaussian processes, data-driven modelling.

Γκαουσιανές διαδικασίες για πρόβλεψη και επίβλεψη απόδοσης και κατάστασης αιολικών πάρκων λαμβάνοντας υπόψη την αβεβαιότητα

Ι. Αντωνιάδου[1], Σ. Κιαρτζής

ΕΛΠΕ Ανανεώσιμες ΑΕ, Χειμάρρας 8Α, 15125 Μαρούσι, Αθήνα

[1]IAntoniadou@helpe.gr

Περίληψη. Οι ανανεώσιμες πηγές ενέργειας και συγκεκριμένα η αιολική ενέργεια έχει υπάρξει το επίκεντρο ενδιαφέροντος με συνεχώς αυξητικό ρυθμό τις προηγούμενες δύο δεκαετίες, σε μία προσπάθεια για μείωση εκπομπών θερμοκηπίου σύμφωνα με τις υπό εξέλιξη συμφωνίες ενεργειακής πολιτικής σε διεθνές επίπεδο. Μεγάλο μέρος των διαδικασιών έρευνας και ανάπτυξης στον συγκεκριμένο τομέα εστιάζεται στην αύξηση της ανταγωνιστικότητας της αιολικής ενέργειας στην αγορά ενέργειας. Μερικές από τις προκλήσεις που συνήθως συναντώνται και χρειάζονται περεταίρω αναφορά συμπεριλαμβάνουν την ανάγκη για βελτίωση σχεδιασμού των συστημάτων ανεμογεννητριών συμπεριλαμβάνοντας στα σχέδια τα σκληρά καιρικά και λειτουργικά περιβάλλοντα που συνήθως συναντώνται στα αιολικά πάρκα και ειδικά στα υπεράκτια, αλλά επίσης την ανάγκη για μείωση των κοστών λειτουργίας και συντήρησης και για βελτιστοποίηση της απόδοσης με βάση τις καιρικές συνθήκες και ενεργειακή ζήτηση. Παρόλο που υπάρχουν διαφορετικές προσεγγίσεις για την επίλυση του τελευταίου ζητήματος, μία τρέχουσα τάση που εξετάζεται τελευταία είναι η εν δυνάμει χρήση πρόσφατων εξελίξεων στον τομέα της ανάλυσης δεδομένων (μάθηση μηχανών και αλγόριθμοι στατιστικής και ανάλυσης σημάτων) που θα μπορούσαν να χρησιμοποιηθούν για την αποτελεσματική ανάλυση δεδομένων SCADA, ταλαντώσεων ή άλλων τύπων που καταγράφονται κατά τη διάρκεια λειτουργίας μιας ανεμογεννήτριας με στόχο να παραχθούν αξιόπιστες προβλέψεις απόδοσης ή κατάστασης αυτόνομων ανεμογεννητριών ή και ολόκληρου του αιολικού πάρκου. Μέσα σε αυτό το πλαίσιο παρουσιάζονται κάποιες προκλήσεις, αφού στις περισσότερες περιπτώσεις ελλείψεις σε σχέση με την αξιοπιστία των δεδομένων και πολυπλοκότητες των διαφόρων διαδικασιών των συστημάτων ανεμογεννητριών μπορούν να οδηγήσουν σε αβέβαιες προβλέψεις, κάτι που ιδανικά πρέπει να ληφθεί υπόψη για την επίτευξη εύρωστης, σημαίνουσας, αυτόματης επίβλεψης. Το παρόν

άρθρο εξετάζει τη χρήση Γκαουσιανών διαδικασιών ως εργαλείων μοντελοποίησης δεδομένων συστημάτων ανεμογεννητριών που επιτρέπουν την εκτίμηση υπόθεσης αβεβαιότητας. Δείχνουμε σε παραδείγματα πως μοντέλα Γκαουσιανών διαδικασιών μπορούν πρακτικά να χρησιμοποιηθούν σε μετρήσεις ταλαντώσεων και/ή SCADA ανεμογεννητριών και να έχουν ως αποτέλεσμα αποτέλεσμα την επίτευξη μη γραμμικής ανάλυσης παλινδρόμησης μέσω έξυπνης μεταχείρισης υπερ-παραμέτρων των μοντέλων.

1. Introduction

The current energy policy developments at a global level, e.g. the recent agreements setting the basic goals of various contracting parties on the critical issue of climate change, including the Kyoto Protocol [1] and the Paris Agreement [2], highlight the directives that will be followed until 2100 concerning the development, capitalization and evolution of a variety of diverse energy resources. In order to accomplish the strategic goals agreed, the formation of supportive mechanisms and formal frameworks is deemed necessary and implies that official bodies are committed to ensure funding provision for research development on "clean" technologies and to involve the Green Climate Fund [3] in potential compensations of countries that have been inevitably affected by climate change phenomena. Statistically speaking, according to Wind Europe Organisation [4], in 2017, 51.2 per asset class billion euros have been invested in wind energy. At a national level, respectively, the National Action Plan concerning the Renewable Energy Sources (RES) and their licensing framework, the integration of the internal energy market and the diversification of energy resources portfolios, are some priorities that signal a transition period currently prevailing in the energy sector. All the above are practically encountered in the energy industry as directives and practices that may lead to a faster transition to the use of biofuels and the further development of diverse RES portfolios. The challenges at this stage are not only limited to the processing or handling of the legislative procedures but also to the importance of delivering and applying novel technologies that could improve the competitiveness of RES in the energy market by solving ongoing issues with regards to high maintenance costs and to the need to improve RES systems energy performance through establishing informed technologically RES machines/mechanisms, clever automated monitoring systems and energy storage (IoT, Big Data Analytics, Energy Storage etc.). For wind energy in particular, with the continuous growth of wind turbine designs, majour new challenges that have arisen include mostly their reliability and performance, more so for offshore wind turbines [5]. Improving reliability and performance related factors is one of the main aims of focus, at the moment, in the research and industrial communities since it would give potential to improve incorporation of wind energy in a more competitive and established manner by decreasing operational and maintenance costs and providing higher power outputs. To achieve this, a lot of focus is currently placed on the development of decision making platforms that exploit data and/or models used for accurate predictions of the wind turbines' health/condition and performance [6] but also for forecasting purposes based on upcoming regulatory requirements that current energy policies demand. The above do not only support reduction of costs but also avoidance of unnecessary over-designing, which is undesirable for any cost effective approach to wind energy. For decision making, monitoring systems and data analysis extracted from wind turbine systems is feasible by exploiting current developments in the areas of signal processing and machine learning, although this doesn't come without its own challenges. Many scientific papers that focus on modelling wind turbines or farms using data-driven approaches tend to point towards the main issues faced and attempt to address them by suggesting sophisticated statistical and other approaches [7, 8]. Wind turbines are rather complex systems and for them to be modeled at a satisfying level of accuracy suggests that strategies should ideally consider stochastic phenomena in predictions. Statistics provide in general such frameworks that tend to be used for inference purposes in the field of machine learning by methods that fall into certain categories e.g. predictors based on priors in function spaces rather than parameter spaces. For any data-driven approach to monitoring the basic steps that tend to be followed can be summarized in the following: operational evaluation, data acquisition, feature extraction and statistical modelling for feature discrimination [9]. The first two steps correspond to mainly understanding and formalizing the specific problem to solve as well as hardware installation while the final two steps correspond to the choice and application of appropriate data analysis methods. Briefly, when dealing with large amounts of data that can be potentially corrupted with noise, the feature extraction step is necessary in order to minimize less informative data dimensions and reduce noise levels. It is then relatively easier to proceed to the feature discrimination step which basically involves some form of novelty detection or classification approach, depending on the amount of training data available. Dimensionality and uncertainty in the data is therefore one issue to consider and deal with, while the training or learning step also suggests an additional necessary prerequisite for successful application which is the availability of reliable data in order to train the algorithms and build accurate data-driven models. It is these data problems that this paper tries to address and by using a demonstration of how a specific relatively recent machine learning approach that uses a statistics framework, such as Gaussian Processes, that can be applied in practice for modelling and prediction purposes in wind turbine systems.

2. Some basic theory on Gaussian Processes

Gaussian processes are a relatively recent development when it comes to modelling nonlinear systems [10]. Their record though could be considered to be rather significant in the field of geostatistics or spatial statistics, where interpolation methods corresponding to Gaussian process regression are known as “kriging” and the first research papers in this area date back to 1960 [11]. Gaussian processes belong to the category of kernel methods and while one is still dealing with a heuristics problem, when trying to create data-driven models using Gaussian processes, their main advantage as a kernel method over neural networks is the potential of exact precision in the optimization of the model produced by clever manipulation of their hyperparameters (such as the spread of the Gaussian kernel). So, compared to classic neural networks models that most times are limited by the training parameters used and the convergence affected by their input function, Gaussian processes might be more tractable and precise in certain applications. In regression problems one can perform the first level of Bayesian inference (computing in posterior distribution of the parameters) analytically [12] using Gaussian processes that are more suited to such problems rather than classification. In neural network models, on the other hand, approximations or sampling are used to evaluate the acquired integrals. A Gaussian process predictor is similar to a Bayesian generalized linear regression that is based on priors in function space rather than parameters space. The above will be described more mathematically in the next paragraphs, although by no means this is a full mathematical description of the method. For anyone interested, the authors advise reading reference [13], which is probably the most cited book on Gaussian processes, and part of the theory described in the book is provided in the following subsections.

2.1 Bayesian Regression

With regards to Bayesian regression then, if one has a training dataset D with a one-dimensional target given by input-output pairs $(x_1, t_1), (x_2, t_2), \dots, (x_n, t_n)$, to motivate the introduction of Gaussian processes, generalized linear regression can and will be used here to model this data. In this case, linear regression using a fixed set of M basis functions $\varphi_1(x) \dots, \varphi_M(x)$ can be performed. The radial basis with fixed (trained) basis functions is a special case of this model, as in linear regression (with the basis functions being the identity function). The model's functional form is given by:

$$Y(x) = \sum_{j=1}^{M} w_j \varphi_j(x) \qquad (equation\ 1)$$

where w is the vector of weights. The assumption made in the above is that of a zero mean Gaussian prior on the weights and zero mean Gaussian noise on the outputs. In this framework it is possible to derive Bayesian predictions from two different viewpoints: the weight-space viewpoint and the function-space viewpoint. The differences between these two viewpoints are described theoretically in [12], and will not be reiterated in such a detailed manner here, but the main point that can be derived is that the function-space allows for one to define the covariance function for a Gaussian process directly without having to first define basis functions and weights. Based on this conclusion one can proceed to the definition of a Gaussian process model which is given in the following.

A stochastic process is a generalization of a probability distribution (which describes a finite-dimensional random variable) to functions. As already implied previously, neural networks are not so easy to apply in practice due to the many decisions which are needed to be made: architecture, activation functions, learning rate, no principled framework etc. One of the reasons they became popular is because they allowed the use of adaptive basis functions as opposed to well known linear models (kernels) [13]. Kernel era showed though that the limitation of fixed basis functions is not a big restriction if one has enough of them (infinitely many). Modern variants of neural networks, such as for example Deep Learning approaches, as a matter of fact, address this issue and might be a quite attractive option due to their ability to scale to large data sets, and describe underlining functions or patterns in the data non-locally.

$Definition.$ A Gaussian process is a collection of random variables, any finite number of which have a joint Gaussian distribution.

A Gaussian process is completely specified by its mean function and covariance function. We define mean function $m(x)$ and the covariance function $k(x, x')$ of a real process $f(x)$ as:

$$m(x) = E[f(x)] \qquad (equation\ 2)$$

$$k(x, x') = E[(f(x) - m(x))\big(f(x') - m(x')\big)] \qquad (equation\ 3)$$

And will write the Gaussian process as:

$$f(x) \sim GP(m(x), k(x, x')) \qquad (equation\ 4)$$

Usually for notational simplicity the mean function is assumed to be zero, although this does not necessarily need to be done. Here the random variables represent the value of the function $f(x)$ at location x. Gaussian processes are defined over time, i.e. where the index set of the random variables is time. This is not (normally) the case in our use of Gaussian processes; here the index set X is the set of possible inputs, which could be more general, e.g. R^D. A Gaussian process is defined as a collection of random variables. Therefore, the definition automatically implies a consistency requirement, which is also sometimes known as the marginalization property. In other words, examination of a larger set of variables does not change the distribution of the smaller set. One should notice that the consistency requirement is automatically fulfilled if the covariance function specifies entries of the covariance matrix. The definition does not exclude Gaussian processes with finite index sets (which would be simply Gaussian distributions), but these are not particularly interesting for our purposes.

3. Example

In order to give a practical example of how the above machine learning method can be applied to a specific problem of data coming from a wind farm for analysis and modelling purposes, one can assume the case of wind farm SCADA data from an offshore windfarm. For full details on the data and part of the results provided here, one can read references [6,7,8] that describe the research developed using a variety of different sophisticated machine learning methods for power curve modelling. Figure 1, demonstrates an example of the application of the method on wind turbine SCADA data, and how the application of the method can improve predictions in some cases by providing knowledge based on a probabilistic viewpoint: marginal boundaries are provided with the predictions and this can be significantly helpful when we are dealing with stochasticity. The figure also shows how the machine learning method performs with the use of a confusion matrix that presents the results of the models for each individual wind turbine of the wind farm. The aim here is to show how such research developments in the data analysis field will be in the recent future a key aspect for wind turbine reliability due to the need to solve such problems and to start working on and exploiting more any available industrial data. Analysis of this type of data could aim to predict power output (e.g. power curve modeling) by performing some form of system modeling in which the system (wind turbines) could be treated as a black box in the case that no knowledge of the physical parameters or equations taking place is available or easy to reproduce. The latter is, at most times, common since any approach to modelling based on physics of real life complex systems such as wind turbines would be quite a hard problem and in practice some form of validation and in some cases model updating of the model based on the system's measurements would be necessary for the physics based model to produce reliable results at any type of environment it might be operating (including unexpected events). The alternative approach, of using purely data to create clever models of individual wind turbines or the wind farm in total, might seem more appealing based on the above. Nevertheless, it doesn't necessarily come with no cost or challenges either. It requires reliable data for training, in-depth understanding of the type of data being analysed in order to pick the most appropriate methods available, and in most cases consideration of uncertainty and similar related concepts/conditions stemming from environmental and/or operational variability and system complexities. Data reliability is an issue to address on its own and lately considerable research is being conducted on developing methods in the machine learning field mostly that can potentially solve some issues when it comes to lack of or missing data for training.

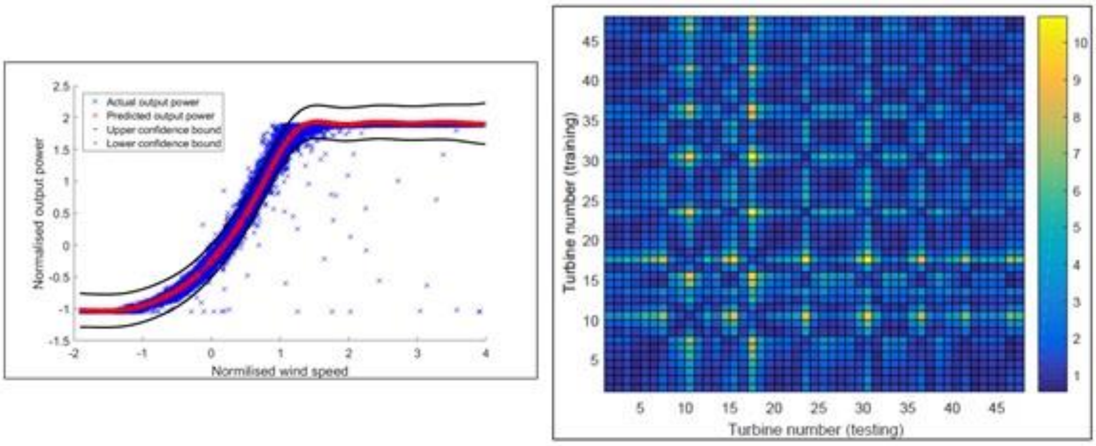

Figure 1. Demonstration of the use of Gaussian processes for power curve modelling and their performance as a machine learning method with the use of a confusion matrix (based on references [6,7,8]).

4. Conclusions

Current energy policies require additional growth of green energy and, as a consequence, wind energy particularly, while continuous developments in data analysis and the need for further digitalization of industrial environment suggest that analysis and exploitation of data coming from wind farms and other green systems for better decision making will be a core need in the industrial sector especially due to the fact that it is going to be very soon supported by regulatory policies. The paper discusses how data analysis can be successfully applied to wind turbine monitoring systems and forecasting, if it is not done in an ad hoc manner. This is clear due to the challenges involving data-driven modelling approaches: for real life industrial applications environmental and operational variations are important factors to be considered since extraction of meaningful data features and reliable data discrimination become a difficult task to achieve. Lack of sufficient data is an issue, but new methodologies are being developed. Gaussian processes may be one the methods solving some of these issues.

Acknowledgements. The authors would like to acknowledge Prof. Keith Worden and Dr. Nikolaos Dervilis from the Dynamics Research Group at the Mechanical Engineering Department, The University of Sheffield and Dr. Evangelos Papatheou from the Engineering Department at the University of Exeter for big part on the development of the work based on years of collaboration.

References

[1] UNFCCC (1997) Kyoto Protocol to the United Nations Framework Convention on Climate Change adopted at COP3 in Kyoto, Japan, on 11 December 1997.

[2] UNFCCC (2018) Paris Agreement to the United Nations Framework Convention on Climate Change adopted at Paris, France, on 22 October 2018.

[3] https://www.greenclimate.fund/home

[4] Wind Europe Organisation, Financing and investment trends: The European wind industry in 2017, published in April 2018, Brussels, Belgium.

[5] I. Antoniadou, N. Dervilis, E. Papatheou, A.E. Maguire, K. Worden,"Aspects of structureal health and condition monitoring of offshore wind turbines", Philosophical Transactions of The Royal Society A, 373 (2035), 20140075, 2015.

[6] E. Papatheou, N. Dervilis, A.E. Maguire, I. Antoniadou, K. Worden, "A performance monitoring approach for the novel Lillgrund offshore wind farm", IEEE Transactions on Industrial Electronics 62 (10), 6636-6644, 2015.

[7] E. Papatheou, N. Dervilis, E.A. Maguire, C. Campos, I. Antoniadou, K. Worden, Extreme function theory for SHM: a case study for wind turbines, Structural Health Monitoring, System Reliability for Verification and Implementation, 2015.

[8] E. Papatheou, N. Dervilis, A.E. Maguire, C. Campos, I. Antoniadou, K. Worden, Performance monitoring of a wind turbine using extreme function theory, Renewable Energy, 113, 1490-1502, December 2017.

[9] C. R. Farrar, K. Worden, Structural Health Monitoring: A Machine Learning Perspective, John Wiley & Sons, November 2012.

[10] C. K. Williams, D. Barber, Bayesian classification with Gaussian processes, IEEE Transactions on Pattern Analysis and Machine Intelligence, 20(12):1342-1351, 1998.

[11] G. Matheron, Kringeage d'un panneau rectangulaire par sa périphérie [Kringing of a rectangular panel via its peripherie], Note géostatistique No.28 Fontainebleau, Centre de Géostatistique, École des Mines de Paris, 1960.

[12] I. T. Nabney, NETLAB Algorithms for Pattern Recognition, Chapter 10: Gaussian Processes, Advances in Pattern Recognition, Springer, January 2001.

[13] C. E. Rasmussen, C. K. Williams, Gaussian Processes for Machine Learning, the MIT Press, 2006.