

Review

On the Use of Indirect Measurements in Virtual Sensors for Renewable Energies: A Review

Abderraouf Benabdesselam ^{1,*}, Quentin Dollon ², Ryad Zemouri ², Francis Pelletier ³, Martin Gagnon ²
and Antoine Tahan ¹

¹ Department of Mechanical Engineering, École de Technologie Supérieure, Montreal, QC H3C 1K3, Canada; antoine.tahan@etsmtl.ca

² Institut de Recherche d'Hydro-Québec, Varennes, QC J3X 1S1, Canada; dollon.quentin2@hydroquebec.com (Q.D.); zemouri.mohamedryad@hydroquebec.com (R.Z.); gagnon.martin11@hydroquebec.com (M.G.)

³ Power Factors, Brossard, QC J4Z 1A7, Canada; francis.pelletier@powerfactors.com

* Correspondence: abderraouf-ramdane.benabdesselam.1@ens.etsmtl.ca

Abstract: In the dynamic landscape of renewable energy, the primary goal continues to be the enhancement of competitiveness through the implementation of cutting-edge technologies. This requires a strategic focus on reducing energy costs and maximizing system performance. Within this framework, the continuous online monitoring of assets is essential for efficient operations, by conducting measurements that describe the condition of various components. However, the execution of these measurements can present technical and economic obstacles. To overcome these challenges, the implementation of indirect measurement techniques emerges as a viable solution. By leveraging measurements obtained in easily accessible areas, these methods enable the estimation of quantities in regions that would otherwise be inaccessible. This approach improves the monitoring process's efficiency and provides previously unattainable information. Adopting indirect measurement techniques is also cost-effective, allowing the replacement of expensive sensors with existing infrastructure, which cuts down on installation costs and labor. This paper offers a detailed state-of-the-art review by providing an in-depth examination and classification of indirect measurement techniques and virtual sensing methods applied in the field of renewable energies. It also identifies and discusses the existing challenges and limitations within this topic and explores potential future developments.

Keywords: indirect measurements; virtual sensors; renewable energies; condition monitoring; diagnosis



Citation: Benabdesselam, A.; Dollon, Q.; Zemouri, R.; Pelletier, F.; Gagnon, M.; Tahan, A. On the Use of Indirect Measurements in Virtual Sensors for Renewable Energies: A Review.

Electronics **2024**, *13*, 1545. <https://doi.org/10.3390/electronics13081545>

Academic Editor: Ahmed Abu-Siada

Received: 7 April 2024

Revised: 14 April 2024

Accepted: 16 April 2024

Published: 18 April 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The growing awareness of fossil fuel depletion and its environmental impact underscore the urgent need for a transition to renewable energy sources to ensure a sustainable future. Currently, renewable energies account for approximately 40.3% of Global Installed Electricity Capacity (GIEC) [1]. Among alternatives, which we aim to evaluate in terms of GIEC, an indicator that assesses the current potential of a given energy source, Hydroelectric Power (HP) emerges as the dominant form of non-fossil energy production in the world, with a GIEC estimated at 1391 GW in 2022. It is followed by solar Photovoltaic (PV), which is considered as the fastest-growing renewable energy source, boasting an estimated GIEC of 1055 GW, and slightly surpassing Wind Turbines (WT), with a GIEC of 899 GW. Other forms of renewable energies, such as Bioenergy (BE), Geothermal Energy (GE), Concentrated Solar Power (CSP), and Marine Energy (ME), also contribute to the diverse energy mix. The installed capacities of all the aforementioned technologies from 2013 to 2022 are documented in Table 1. Additionally, while nuclear power is known for its low carbon dioxide emissions, making it a clean energy source, it does not fall into the

renewable energy category because it relies on the use of fuels that are a finite resource on Earth, and can cause potential environmental problems [2].

Table 1. Global Installed Electricity Capacity (GIEC) of different types of renewable energies [1].

Energy	HP	PV	WT	BE	GE	CSP	ME
GIEC in 2013 (GW)	1137	137	300	84	10	3	0.51
GIEC in 2022 (GW)	1392	1055	899	150	14	6	0.52

To make these energy sources more competitive and attractive, their associated Levelized Cost of Electricity (LCOE) must be reduced; this includes initial investment costs, operating and maintenance costs, and financing costs, taking into account the system's service life and power generation [3]. The current challenge lies in making investment costs profitable by minimizing operating costs (mainly from maintenance activities) [3]. As a result, maintenance has evolved from initially having a corrective (reactive) maintenance focus to having one of systematic maintenance, in order to avoid sudden shutdowns that lead to significant operational losses, especially when operations are of a strategic nature [4]. It is in this context that condition-based maintenance, also known as Condition Monitoring (CM), emerged, allowing to monitor the health status of equipment in real time, and intervening only when necessary [4]. This evolved into predictive maintenance, also known as prognosis, involving forecasting the future health status of assets based on their Remaining Useful Life (RUL) [5], allowing for more precise planning of interventions and minimizing economic losses to the greatest extent possible [4,5]. The Prognostic and Health Management (PHM) process has crowned all these advancements by generalizing CM, integrating both prognosis and decision-making to enable proactive management of anomalies and their associated risks, and covering six interdependent phases. The first stage is anomaly detection, which acts as an alarm signal and sends notifications for abnormal events; the second step is the description of the detected anomaly, which provides a detailed understanding of the event to ensure that the anomaly is not due to a sensor failure; the third stage is diagnosis, in which an attempt is made to find the root cause of the anomaly; the fourth stage is the prediction of future events that may be caused by the anomaly; the fifth stage is the prescription of the measures that should be taken and recommendations made; in the final stage, decisions are made to balance asset health, performance, and benefits. Sometimes, short-term operational gains are prioritized over the asset's long-term health [6]. A detailed evaluation of the equipment's condition, fault type, location, severity, remaining useful life projections, and operational guidelines is crucial.

To ensure continuous monitoring of assets within the scope of CM, and to evaluate the condition of their components, it is imperative to conduct measurements of physical quantities. This can be accomplished through traditional sensors, or as has been the case more recently, through virtual sensors, which represent advanced computational tools that can infer unmeasurable states or trigger alerts for anomalies. CM takes on significant importance when the reliability of an asset is strategic, such as with HP, PV, and WT, which are among the largest contributors to renewable energy globally, as shown in Table 1. Concerning bioenergy, which is considered a renewable resource due to its carbon neutrality, it encompasses solid biofuels, liquid biofuels, and biogas. Solid biofuels, also known as biomass, represent the largest segment of bioenergy, with a GIEC of 126 GW, accounting for 85% of bioenergy's GIEC [1]. Biogas, produced through the gasification of biomass, follows with a GIEC of 21 GW, or 14% of bioenergy's GIEC. Liquid biofuels have the smallest share, with a GIEC of 2 GW, constituting 1% of bioenergy's GIEC. Biomass, notable for its significant GIEC, is harnessed for electricity generation through three main methods, pyrolysis, gasification, and direct combustion, with the latter being the most prevalent and efficient method for converting biomass into energy [7,8]. The conversion process involves the use of External Fired Gas Turbine (EFGT) power plants, usually within a cogeneration cycle. This approach is necessitated because the combustion gas produced is considered as

dirty [9], and therefore unsuitable for direct use in the gas turbine. Instead, this gas is used to heat compressed air via a heat exchanger. After this heating process, the air expands within the turbine, enabling energy generation. Power generation efficiency is directly linked to the turbine inlet temperature; however, this efficiency is constrained by the maximum temperature tolerance of the materials used in heat exchangers, which is approximately 1000 °C, making these plants only suitable for micro and mini generation [7–11]. In such contexts, employing CM is not cost-effective, given the sophisticated technology required for relatively low-capacity power plants. Also, the use of GE, CSP, and ME is minimal, as indicated in Table 1. This limited use seems to make CM economically unfeasible due to the low GIEC of these techniques.

Regarding hydraulic turbines, the installations convert the energy of the forced flow of water, which is then used to rotate a runner driving a synchronous alternator. The runner is the critical component of the unit [12], and is susceptible to various forms of damage, including cavitation, corrosion, abrasion, and especially, cracking due to fatigue from vibrations [13]. The blade-crown and blade-band joints are particularly sensitive to strain resulting from the high stress concentration due to the characteristics of water flow [14]. Understanding the runner's degradation mechanisms typically involves two methods: numerical simulations and experimental tests [15]. When dealing with high-dynamic regimes using numerical methods that combine Computational Fluid Dynamics (CFD) and the Finite Element Method (FEM), there are notable challenges to overcome. These regimes are characterized by complex fluid flow patterns, especially in the presence of turbulence and stochastic events [16,17]. The complexity and unpredictability of such phenomena substantially increase the computational requirements, potentially affecting the accuracy of simulations. CFD algorithms must navigate the challenge of precisely capturing the transient and turbulent aspects of fluid flow [16,17]. Furthermore, the incorporation of stochastic elements, marked by random variations and irregularities, introduces an additional layer to the prediction and modeling processes [16,17]. As a result of these limitations, experimental measurements are the only reliable method in this case [15]. This approach involves installing strain gauges in critical areas, although precise placement is challenging due to the complex geometry of the locations (blade-crown and blade-band joints). Consequently, sensors are often not positioned exactly at the critical areas, compromising accuracy. Installing sensors, especially strain gauges, in a submerged, rotating, and corrosive environment is complex, time-consuming, costly, and often requires several weeks to complete. Moreover, these sensors have a limited lifespan, and are prone to dislodgment by the flow, providing limited signals over time which cannot cover all operating modes, and are therefore unsuitable for online monitoring [14,18].

Another critical component of hydraulic turbines is the generator, which is connected to the runner via a shaft supported by bearings [19,20]. The most common mechanical failures in generators include fatigue, eccentricity, rotor or stator ellipticity, and bearing malfunctions [4,19–21]. Electrical faults such as inter-turn short circuits and broken damper bars also pose significant risks, leading to increased vibration levels, higher temperatures, or irregularities in stator currents [19,20]. To effectively monitor generators, it is common practice to place vibration sensors (accelerometers) and acoustic sensors (microphones) on the stator frame [19,20]. Using these vibro-acoustic techniques, valuable information can be obtained, as any alterations detected may indicate potential faults [19,20]. In order to identify faults without the need for machine shutdown, the signals are examined using various analysis techniques. Time domain analysis, through root-mean-square values, is employed, along with frequency domain analysis, which uses Fourier transform. Additionally, time-frequency domain methods such as the short-time Fourier transform and wavelet techniques are used. However, the task of extracting clean signals from the surrounding background noise poses a significant challenge, as the signals can be obscured by ambient noise [19,20]. Data pre-processing is thus necessary to address this issue [19,20]. Other monitoring techniques, such as strain gauges, twist measures, displacement measures for the shaft, along with lubrication analysis for bearings, are also employed [14].

While photovoltaic solar systems are a reliable source of renewable electricity, they are, however, prone to various failures and outages. These systems encounter challenges such as weather-related disruptions (changes in irradiance, cloud cover, wind speed, and temperature), electrical failures, and hardware malfunctions such as panel defects, inverter failures, and grid instability. These issues can significantly affect the system's energy production and efficiency. To address these problems, monitoring methods including infrared thermography and electroluminescence, along with data analytics, are employed. These technologies allow to accurately identify and monitor issues, facilitating quick maintenance interventions and significantly enhancing system reliability and performance. Notwithstanding the advancements in PV systems, they still face inherent challenges related to predicting power generation based on the operating environment. This includes factors such as variable power output and solar irradiance, which necessitate the use of predictive models. These models are crucial for forecasting PV generation, which in turn enables smart demand response, efficient energy management, and ensures an adequate supply [22–28]. Forecasts can be categorized into three types, based on the time horizon: short-term (from a few minutes to several hours), medium-term (from 1 week to 1 year), and long-term (longer than 1 year) [29].

Regarding wind turbines, their critical mechanical components, including the rotor, gearbox (GB), and bearings, are prone to fatigue, which represents the predominant failure mode for these components. Additionally, the blades are subject to erosion [3,4,30–32]. The rotor, comprising the hub and blades, is typically monitored using strain gauges. These installations, when installed on the blades, have a limited lifespan, and require frequent replacement, which poses challenges for online monitoring. Thus, a transmission system, either wired or wireless, is necessary for data transfer from the rotating part to the fixed part, and this introduces reliability issues due to potential interferences. Monitoring of gearboxes and bearings is mainly done through lubrication analysis, acoustic emission analysis, vibration analysis, or using operational data archived in Supervisory Control and Data Acquisition (SCADA) systems. Vibration and acoustic analysis using techniques such as the Fourier transform and Support Vector Machine (SVM) face challenges related to sensor failures (which are also difficult to locate) and interference caused by external sources (noise pollution) [21,30,33]. Lubrication analysis, which includes oil cleanliness, oil viscosity, and temperature assessment, is a low-cost and easy-to-implement anomaly indicator, but it cannot provide a detailed understanding of the origin of anomalies and often produces false positive results [21,33]. SCADA data are typically averaged over a 10-min period, corresponding to low frequencies, and are therefore not suitable for detecting high-frequency anomalies. The type of failures and CM challenges of HP, PV and WT are reported in Table 2.

Table 2. Types of failures and condition monitoring challenges for hydraulic turbines, photovoltaic systems, and wind turbines.

Energy	Types of Failure	CM Challenges
HP	Fatigue, cavitation, corrosion, and abrasion	Placement difficulty and limited lifespan of strain gauges; complex environments for sensor deployment
PV	Weather-related disruptions, electrical issues, hardware failures	Managing variable power output and daylight dependence; the need for advanced predictive models
WT	Fatigue and erosion	Difficulties in sensor placement and durability; complexity associated with gearbox and bearing monitoring

1.1. Problem Statement

The aforementioned energy sectors share some common characteristics: sensors used for asset monitoring often have a limited lifespan, and require frequent replacements [14,15]. In some cases, the situation can be exacerbated by a lack of awareness regarding sensor malfunction, resulting in erroneous data being used for predictive purposes. This leads to

significant downtime and direct economic losses, which are not ideal for online equipment monitoring. Additionally, the geometric complexity sometimes makes it nearly impossible to carry out precise sensor installations at critical locations [14]. Operators are then forced to place the sensors elsewhere, compromising measurement accuracy. Furthermore, some monitoring devices can be expensive, leading to concerns about their economic viability and managers' ability to recoup investments [14].

Because such sensors cannot be used for continuous online monitoring of assets, the indirect measurements approach comes into consideration [14]. It brings about cost reductions with respect to installation and operations, particularly those related to frequent sensor replacements and production losses due to installation [14]. This approach simultaneously addresses technical challenges by enabling data collection in otherwise inaccessible areas from other easily accessible areas, and tackles economic challenges by proposing a model capable of substituting expensive sensors in critical areas while minimizing the costs and efforts required for their deployment [14]. Often integrated into virtual sensor architectures, this approach poses the special challenge of determining the relationship between the required variables of interest and accessible variables. This relationship is called the general transfer function, and can take the form of an analytical function, an artificial intelligence algorithm, or any other form. To the best of our knowledge, no general method for determining this transfer function using indirect measurement methods, especially in the context of renewable energy, has been proposed in the literature. Furthermore, the absence of any review article in this context underscores the need for us to fill this gap. We aim to address this by providing a comprehensive state-of-the-art review, and to proceed by compiling indirect measurement methods used in the renewable energy field.

1.2. Paper Organization

This paper is structured as follows. Section 2 outlines the indirect measurement approach and proposes a classification for methods used in the renewable energy arena. Section 3 offers a literature review of data-based indirect measurement methods. Section 4 reviews the literature on physics-based indirect measurement methods. Section 5 discusses the literature on hybrid indirect measurement methods. Section 6 delves into a discussion on these methods, emphasizing the existing gaps in the literature. Section 7 concludes the paper.

2. Indirect Measurements

An indirect measurement refers to the estimation of a latent physical quantity (or quantities) y by observing another physical quantity (or quantities) x , sometimes of a different physical nature [34,35]. This estimation relies on establishing a relationship between these two quantities, called the transfer function, expressed in the case of a Single-Input Single-Output (SISO) system as $y = f(x)$. This function allows predicting the value of the quantity y being sought itself as a function of the directly measured quantity x . Generally, an indirect measurement is expressed as a Multiple-Input Single-Output (MISO) system, which is the most intuitive approach, but it can also be extended to a Multiple-Input Multiple-Output (MIMO) system [35].

This transfer function, which can take various forms such as an analytical function, a system of differential equations, or even an artificial intelligence model, must be capable of connecting the quantity of interest $y \in \mathbb{R}$ to the vector of measured quantities $x \in \mathbb{R}^n$ (in the case of a MISO system), where n represents the number of established measurements [34,35]. Additionally, it should quantify the uncertainty associated with this method. When an estimated uncertainty u_x exists in the measurement of x , determining the uncertainty in the indirect measurement of y , denoted as u_y , becomes essential to ensure reliable and accurate results [35]. Furthermore, it is critical to also consider epistemic uncertainties, which do not originate from the measurement data themselves, but rather, from inaccuracies in the modeling of the transfer function. These uncertainties reflect our limited knowledge

about the model and its parameters, highlighting the importance of acknowledging and quantifying both types of uncertainties to achieve more robust and reliable outcomes [35].

Considering the case where the relationship between the output quantity of interest y and the input quantity x can be explicitly formulated as an analytical equation, this relationship is expressed in the Laplace domain in this case of a SISO system by Equation (1).

$$Y(s) = H(s) \cdot X(s), \quad s \in \mathbb{C} \quad (1)$$

where $H(s)$, $X(s)$, and $Y(s)$, respectively, denote the transfer function, the input function, and the output function, all expressed in the Laplace domain, where s is a complex variable that belongs to the set of complex numbers \mathbb{C} . The past relationship (Equation (1)) can be written in the time domain by Equation (2), where t is the time variable that belongs to the set of real numbers \mathbb{R} , and τ is a real-parameter.

$$y(t) = h(t) * x(t) = \int_0^t h(\tau) \cdot x(t - \tau) d\tau, \quad t \in \mathbb{R}_+ \quad \tau \in \mathbb{R}_+ \quad (2)$$

This approach requires a thorough exploration of the physical behavior of the structure under examination. The selection of essential and relevant variables of the structure, understanding the level of correlation between these variables, as well as identifying phenomena or variables that could compromise the accuracy of the model, are all critical aspects. An indirect measurement offers several advantages in various contexts, particularly when direct measurement is inaccessible or complex, when cost reduction is necessary, and when non-intrusive measurements are preferred. This type of measurement provides significant flexibility and versatility by allowing the measurement of different quantities using the same set of instruments.

The principle of indirect measurements is widely applied across various industries beyond renewable energies, often without explicit mention of the principle, as seen in manufacturing processes. On the one hand, Hamada et al. [36] proposed a method for assessing the quality of spherical containers produced with inner liners, where the equators do not coincide due to manufacturing differences. In this method, the main goal is to quantify the maximum distance between these equators, which is considered to be an indirect quality indicator. Using the concept of equivariance, it models the distribution of the distances, which follows an exponential distribution, and then determines the upper tolerance limit based on scale equivalence statistics. On the other hand, Moricz et al. [37] addressed the problem of milling tools wear used for drilling ceramic coatings. They presented an indirect method aimed at quantifying tool wear by analyzing geometric variations during machining operations and establishing correlations between these variations and tool degradation. This approach provides a precise and reliable alternative for monitoring tool wear, thus contributing to the optimization of machining processes in the industrial domain. Furthermore, in the same manufacturing context, De Paiva Silva et al. [38] used the same indirect measurement approach as [37] through correlations to evaluate the wear of tungsten carbide micro-milling cutters during the machining of Inconel 718 by measuring the diameter loss on micro-slots.

When it comes to renewable energies, indirect measurements are often integrated into virtual sensor architectures. While one of the primary functions of virtual sensors is to enable indirect measurements, their utility extends beyond this specific application, which represents only one aspect of their versatile capabilities [39]. In the field of virtual sensors mainly used for indirect measurements, three main categories stand out: data-driven sensors, physically modeled sensors, and finally, sensors that use a hybrid approach combining data and physical models.

Physics-based virtual sensors are founded on first-principles models, which can provide high accuracy but present challenges with respect to adaptation to real-world changes. Their modeling usually approaches the stable ideal case of the structure. In contrast, data-driven virtual sensors are increasingly popular in process industries and more accurately

reflect actual process conditions. While close to reality, these virtual sensors face challenges related to their so-called black box nature and to a lack of interpretability of their results as well as a lack of generalizability when data-based models attempt to extrapolate beyond their trained range. However, ongoing technological advancements are promising when it comes to continued improvements of the field of data-based virtual sensors and of the accuracy and reliability of industrial systems [40].

In the following, we will examine existing works in the literature, classifying them according to the categories to which they belong (i.e., data-driven methods, physical models, or hybrid methods). It is worth noting that the boundaries between these categories are subtle, and that this classification may vary from one researcher to the next, which could potentially be confusing for readers based on their specific areas of expertise. Therefore, it is essential to clearly define the criteria that guide our classification from the outset. We provide a distinction between data-driven methods, physics-based methods, and hybrid approaches.

2.1. Definition of a Data-Based Method

Data-driven approaches, especially those used in virtual sensor environments, are characterized by being based on measured data, which distinguishes them from traditional physical system modeling. These methods use statistical correlations and/or machine learning algorithms to predict quantities of interest; further, they do not require direct consideration of the underlying physical laws that govern system behavior, and therefore represent a major shift in the modeling paradigm [34].

This paradigm allows the creation of models that are flexible and adaptable, properties that are particularly advantageous in situations where the physical processes are intricate and not well understood, or are too complex to be accurately represented by traditional equations [34]. In this case, traditional physics-based models may not be enough, and data-driven models can fill the gap. The power of data-driven models is a function of their ability to learn and adapt to training data [34].

However, despite the advantages of such models, they also face significant challenges, including their explainability and generalization [41]. Generally, data-based models struggle to predict situations that have not been encountered during training, which can be difficult and requires large amounts of data and computational power [34,39]. Their accuracy depends on feature engineering, particularly on the judicious choice of variables, as adding irrelevant variables to the model would introduce noise and complicate training, whereas reducing the inputs necessary for prediction would decrease the accuracy. Furthermore, in models in which temporal dependence is considered, the accuracy would depend on the choice of sampling frequency. Another important aspect is the propagation of uncertainties, which can lead to overconfidence in predictions. Artificial Intelligence (AI) appears to be very promising for solving these issues. Bayesian neural networks, which integrate Bayesian inference principles to quantify uncertainty, could help manage the propagation of uncertainties [42,43]. Contextualized AI aims to understand and take into account the context in which data is generated and used, potentially improving the generalization of data-based models. Moreover, causal AI focuses on understanding causal relationships rather than merely correlational ones, which could lead to better accuracy and the ability to generalize the model [42,43].

An example of a data-based measurement method within the renewable energy context is the prediction of wind turbine blade strain based on tower strain. In this case, a data-driven approach involves collecting strain measurements from both the blades and the tower. Subsequently, a statistical model is formulated to link these two quantities and identify appropriate correlations.

2.2. Definition of a Physics-Based Method

Unlike data-driven approaches, those that are physics-based rely on pre-established representations of the physical model to forecast the system's state. The latter approaches

are rooted in the comprehension and the use of the governing physical laws. The models are formulated based on these laws, serving as a theoretical structure for predicting the system's behavior [34].

When using physics-based methods, data is used only to quantify the inputs of a system, such as in finite element models, or for validation purposes. This use of data does not have any impact on the outcomes of the predictions and does not involve any statistical modeling. However, when data is used to enhance the accuracy of predictions made by a physics-based model through the integration of a statistical model, such as in state-space models, this approach is then categorized as a hybrid method.

An example of a physics-based method within the context of renewable energy is the prediction of wind turbine blade strain. This method involves modeling the wind turbine structure as a mass-spring-damper system with multiple degrees of freedom, which is then discretized using the finite element method. By incorporating measured data, inputs to the structure, such as wind loads on the blades and the tower, are quantified. Ultimately, blade strain is determined by solving the system of equations derived from the finite element model.

Physics-based models have the ability to provide an in-depth understanding of how a system behaves and can be extremely precise when the underlying physical laws are well-established and the system can be accurately represented [34]. However, these models may face challenges when dealing with complex or poorly understood physical processes that cannot be effectively captured using traditional equations [34]. In these situations, data-driven models, with their flexibility and adaptability, can provide substantial benefits.

2.3. Definition of a Hybrid Approach

Hybrid models represent a fusion of physical and data-driven approaches to system modeling and prediction. These models feature integrated physical representations of systems, and include statistical models for predictive purposes.

In hybrid models, the physical representation provides a theoretical framework based on the known physical laws of the control system. This framework serves as the basis for the model, providing a priori insights into the system's behavior. However, unlike purely physics-based models, this physical representation is not the only source of predictive power in hybrid models [13,44].

The physical representation is supplemented by statistical models created based on measured data collected by the system. The data-driven component of the hybrid model uses statistical correlation and machine learning algorithms to refine the predictions of the physical model. Statistical models are able to capture complex patterns and relationships in data that may not be easily obvious or easily modeled using only the laws of physics [44].

An example of a hybrid method within the context of renewable energy is the prediction of wind turbine blade strain. Initially, a physics-based approach using the finite element method is employed, as detailed in Section 2.2. Once the blade strain is analytically determined, it can be further refined by employing a statistical model that correlates the strain of the tower and the blade, as discussed in Section 2.1.

The goal of this hybrid approach is to leverage the strengths of physics- and data-based models to optimize prediction accuracy and robustness. By leveraging the physical understanding of the system and the statistical correlation of available data, hybrid models aim to provide the best of both worlds: the explainability of physics-based models and the adaptability and flexibility of data-driven models [44].

In essence, hybrid models represent a balanced approach to system modeling and prediction. They recognize the value of physical laws in providing a solid theoretical foundation, while recognizing the power of data in capturing the complexity of real systems. This balance enables hybrid models to adapt to new data and evolving system dynamics, making them powerful and versatile tools in the field of systems modeling and prediction. The proposed classification is shown in Figure 1.

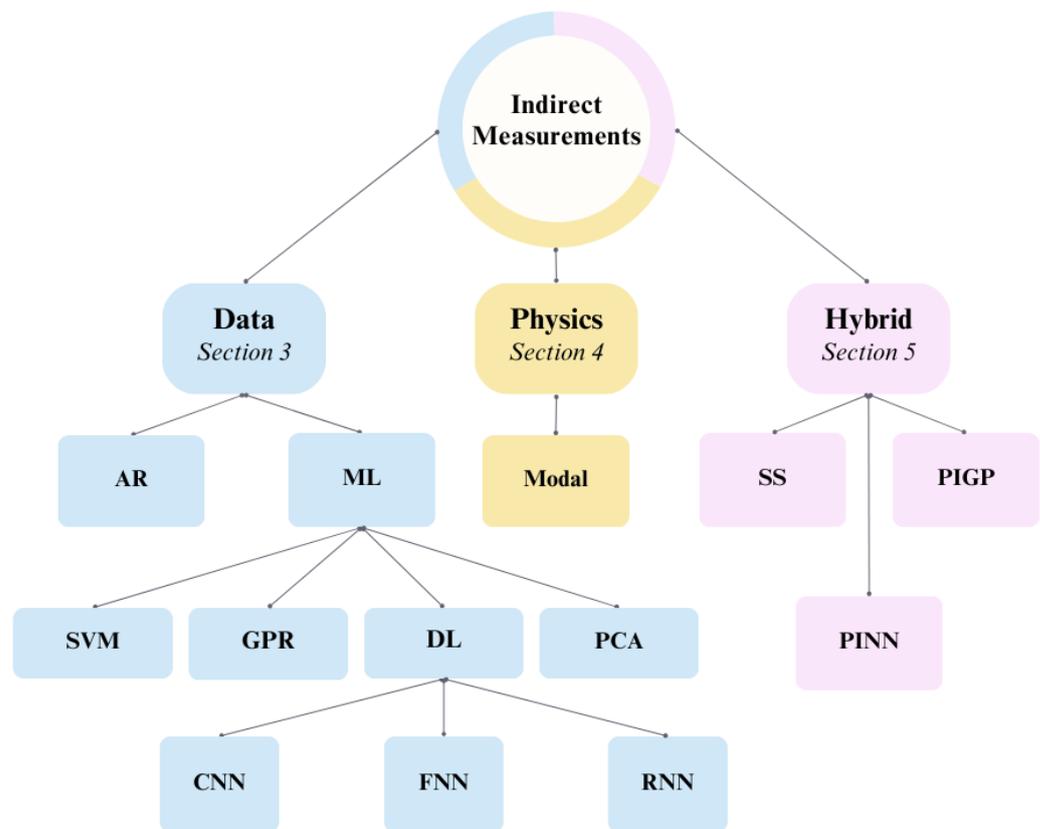


Figure 1. Indirect measurement methods classification.

3. Data-Based Methods

In the context of data-driven approaches, TayebiHaghighi et al. [45] suggest a classification into two main categories: classical statistical methods on the one hand, and those based on artificial intelligence, on the other. In this document, we introduce a classification based on applications identified in the literature on indirect measurements in the energy domain. We specifically examine methods using autoregressive models, as well as those using Machine Learning (ML), including Deep Learning (DL), SVM, Principal Component Analysis (PCA) and Gaussian Process Regression (GPR).

3.1. Autoregressive Models

For dynamic systems where the mathematical model is not known, the system identification method is used to build the model from observed data by monitoring input disturbances and analyzing the output. Among the models used in system identification, Autoregressive (AR) models are considered one of the pillars of system identification [46]. In an AR model, it is assumed that the current value of a time series is a linear combination of its past values from the same time series. In other words, the model is described as autoregressive because it regresses on itself. Specifically, an autoregressive model of order p , denoted as $AR(p)$, uses the previous p -values to predict the current value. The basic form of a linear $AR(p)$ model is described in Equation (3).

$$y(t) = \phi_0 + \sum_{i=1}^p \phi_i \cdot y(t-i) + \epsilon(t) \quad (3)$$

where $y(t)$ represents the value of the time series at time t ; ϕ_0 and ϕ_i are the coefficients of the model (autoregressive coefficients), and $\epsilon(t)$ is a random error term at time t (a white noise following a normal distribution). The model coefficients are generally obtained using Yule-Walker equations.

Diagne [14] explored the application of purely autoregressive indirect measurement methods for predicting strain on the runner of a hydroelectric turbine using data from the turbine shaft (torsion, flexion, etc.). The study initially demonstrated the existence of linear correlations between sensors on the shaft (particularly torsion) and those on the runner (strain). However, it highlighted the insufficiency of these correlations for precise prediction. This led to the adoption of a dynamic model, specifically, an AR model, to enhance prediction accuracy. After conducting a comparative study of different AR models such as Autoregressive with exogenous Inputs (ARX), Autoregressive Integrated with exogenous Inputs (ARIX), Autoregressive Moving Average with exogenous Inputs (ARMAX), and Autoregressive Integrated Moving Average with exogenous Inputs (ARIMAX) models, the ARMAX model was identified as optimal. Model validation confirmed its consistency with real-world measurements, allowing for optimized startup sequences and reduced stress levels.

The Autoregressive Integrated Moving Average (ARIMA) model, as employed by Pasari et al. [47], was used to predict daily and monthly wind speeds and temperatures, without exogenous variables, by analyzing 15 years of hourly data. This method achieved Root-Mean-Square Error (RMSE) of 0.893 and 0.659 for temperature and wind speed, respectively, showing a good fit with actual data. Najdawi et al. [48] adopted a Vector Autoregression (VAR) model for forecasting short-term solar irradiance, using weather conditions (atmospheric pressure, temperature, and relative humidity) and solar irradiance, but faced limitations due to the model's low lag order, capping the forecast at four hours. Zhang et al. [49] developed an Autoregressive Dynamic Adaptive (ARDA) model with no exogenous inputs, for real-time wind power forecasting, which, when compared with ARIMA and Long Short-Term Memory (LSTM) models, showed superior performance in accuracy, speed, and adaptability to wind data fluctuations.

The traditional AR models, while effective, are linear and may not accurately predict outcomes in highly nonlinear scenarios. To address this, combining autoregressive models with machine learning techniques, such as neural networks or SVM, has been proposed to capture the nonlinear aspects of predictions. Shukur et al. [50] suggest a hybrid approach, which will be discussed in Section 5, integrating neural networks and Kalman filtering with ARIMA to overcome the limitations of the latter with nonlinear wind speed data. Similarly, Babu et al. [51] employ a data-hybrid strategy for wind speed prediction, initially applying an ARIMA model for its linear prediction capabilities, and then enhancing it with a neural network to capture nonlinear dynamics. Wang et al. [52] introduced a forecasting model combining a Seasonal Autoregressive Integrated Moving Average (SARIMA) and SVM for ultra-short-term photovoltaic power generation prediction, using historical PV generation data as inputs of the model, which divides time series data into linear and nonlinear components, using SARIMA to model the linear part and SVM for the nonlinear part. Their findings show that this combined method outperforms traditional SVM and SARIMA models in forecasting accuracy, offering a solid foundation for safe PV grid operation and scheduling.

A second approach to account for the non-linearity in predictions uses the Nonlinear Autoregressive (NAR) model, expressed by Equation (4) [53].

$$y(t) = F(y(t-1), y(t-2), \dots, y(t-p)) + \epsilon_t \quad (4)$$

Unlike linear models (as shown in Equation (3)), the above approach employs a nonlinear function F , often a neural network in virtual sensors applications [53]. This method retains the autoregressive model's framework, but applies it through deep learning techniques, indicating a shift towards a purely deep learning application, which will be further discussed in part Section 3.2.1.

3.2. Machine Learning-Based Methods

There are four main categories of machine learning methods used for virtual sensing in the context of renewable energies: one that uses DL exclusively, and three approaches

that incorporate techniques adapted from general machine learning, such as SVM, PCA and GPR.

3.2.1. Deep Learning-Based Methods

In the following, we present DL methods used for virtual sensing applied to renewable energies. Generally, these DL methods are either Feed-forward Neural Networks (FNN), Recurrent Neural Networks (RNN) or Convolutional Neural Networks (CNN).

FNN, a foundational neural network model using Multi-Layer Perceptron (MLP), is adept at pattern recognition. Presas et al. [54] applied an FNN model to predict static and dynamic strains of HP runner using vibrations, displacements, pressure and strain data from the stationary part [55], and, through the Levenberg-Marquardt learning method, found that a two-hidden-layer, 25-neuron configuration is optimal for minimizing the Mean Absolute Error (MAE). This setup achieved a high strain prediction accuracy, with an MAE of less than 1% for mean strain and around 3% for dynamic strain. Azzam et al. [56] designed a virtual sensor for 6 Degree of Freedom (DOF) load measurements on a wind turbine gearbox using FNN and a multi-body simulation model used to produce time-series data to train the model by generating a simulated wind field, demonstrating the virtual sensor's effectiveness in accurately capturing gearbox loads.

Regarding photovoltaic power generation prediction models, Das et al. [57] concluded in a 2018 review that FNN models and SVM models perform well under rapidly changing environmental conditions. Refs. [58–64] all proposed FNN methods that use various inputs such as index data, aerosol data, and traditional weather parameters such as temperature, humidity, and wind speed, to predict a PV system's 24-h power output. In particular, Chen et al. [61] introduced previous power measures to examine temporal dependencies. Martin et al. [65] and Pedro et al. [66] compared FNN with AR methods and highlighted the performance of FNN relative to AR prediction methods in the 1 h to 3 day forecast range.

Nevertheless, FNN models, which do not explicitly take into account temporal dependence, treat all inputs discretely and equivalently, making them less suitable for time series prediction. In contrast, RNN are more suitable for sequence models [67], where input features represent sequences and weighted connections between cells enable the learning of temporal properties of the signal. Dimitrov et al. [68] conducted a comparison between these FNN models and an LSTM model (which is RNN) for the estimation of the bending moment at the root of the blade, the wake center position, and the blade tip deflection of wind turbines using SCADA data. The study results highlighted the importance of memory in predicting time series signals of wind turbines. Using at least one step memory, particularly with LSTM models, showed improvements over using only the current state, demonstrating the superiority of LSTM over FNN in time series prediction. Also, the relative errors of the models ranged from 10% to 50% of the standard deviations of the estimated signal. Gulgec et al. [69] used the same LSTM cells approach to estimate the strain responses in a mechanical structure using acceleration sequences as inputs, demonstrating the high accuracy of the model even for unmeasured locations and varying loads. Park et al. [70] used an LSTM model to forecast the natural frequency degradation of offshore WT. Initially, they constructed a 3D finite element model that took into account aero/hydrodynamic loads to predict the turbine's dynamic response. This response, along with vibration and strain data, served as inputs to train the model. Validation was performed using various wind distributions, demonstrating the model's ability to successfully predict the natural frequency degradation trends over three years, based on input data from the previous six years. Balluff et al. [71] similarly developed an RNN model for short-term wind speed and energy production forecasts for offshore wind farms. They used the wind speed, wind direction, temperature, and surface pressure data as inputs, achieving acceptable prediction results. Han et al. [72] employed a similar approach for mid-to-long-term wind and PV power prediction. They used historical data on wind and PV power generation, along with key meteorological factors, as inputs, demonstrating the effectiveness of the model employing data from various countries.

Regarding the application of RNN in the PV sector, particularly for short-term solar irradiance forecasting, Yu et al. [73] first employed LSTM for this task by incorporating historical meteorological data and the clearness index as inputs, thus enhancing the model’s ability to handle weather variability. Their research highlighted notable advancements in forecasting accuracy, especially under cloudy and mixed weather conditions, across diverse locations. This approach was further refined through the integration of a Variational Quantum Circuit (VQC) with LSTM, as detailed by Yu et al. [74], to tap into the potential of quantum computing to improve the LSTM’s predictive accuracy by optimizing the weight parameters of the LSTM gates. The inputs for this enhanced model included meteorological and solar radiation data, with the outputs being the forecasted solar irradiance values. Both models exhibited significant performance improvements, as evidenced by metrics such as RMSE and MAE, with the hybrid model combining quantum computing and deep learning demonstrating superior results. Wang et al. [75], Konstantinou et al. [76], and Gao et al. [77] each proposed a model for predicting PV power generation. Wang et al. [75] suggested a one-day in advance prediction model by using irradiance and temperature data adjusted to a 15-min temporal resolution for training. Additionally, they enhanced the accuracy of predictions by integrating temporal correlations, yielding significant improvements over conventional methods. Additionally, Konstantinou et al. [76] focused on predicting PV power output accurately 1.5 h ahead, leveraging historical data from a PV plant. Their validation revealed the model’s effectiveness, achieving an RMSE of 0.11 and a standard deviation of 0.016. Despite acknowledging limitations stemming from the inherent unpredictability of weather conditions and solar irradiance, their approach showed promising results. Furthermore, Gao et al. [77] developed a one-hour PV power production prediction model using LSTM, incorporating historical PV power data and meteorological conditions as inputs. Their model demonstrated improved prediction accuracy, with RMSE values for forecasting using LSTM in the four seasons being 5.34%, 9.57%, 13.86%, and 9.26%, respectively.

NAR models, which are introduced in Equation (4), use a non-linear function F that is often implemented as a neural network. This function takes into account temporal dependencies and is considered a type of RNN because its output is recycled as its input. Given that NAR models are a form of neural network, they are referred to as Non-Linear Autoregressive Neural Networks (NARNN). When the system is supplemented with exogenous inputs, it evolves into a Non-Linear Autoregressive with exogenous inputs Neural Network (NARXNN). A unique feature of this network is the inclusion of tapped delay lines that store previous values of input and output sequences. The output of the NARXNN is then fed back into the network input with a delay, as shown in Figure 2 [28].

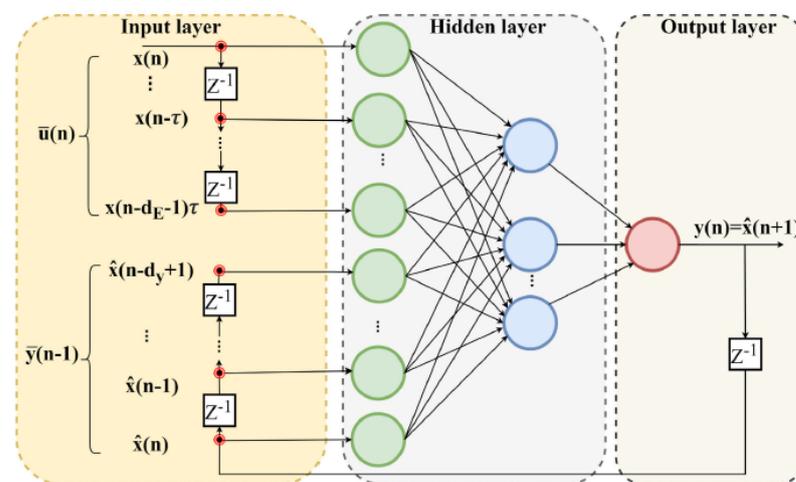


Figure 2. NARXNN model architecture [28].

Bimenyimana et al. [78] predicted the output power of PV modules based on meteorological data, using the NARXNN model for monthly and annual power prediction. Inputs include temperature data and other environmental parameters. The model achieved good accuracy under different temperature conditions, with regression values exceeding 70%. On the other hand, Wang et al. [79] and Garcia et al. [80] used NARXNN for WT forecasting. Wang et al. [79] used a NARXNN model fed by 10-min SCADA data to predict gear oil temperature, in a bid to detect anomalies in WT. The inputs included power, nacelle temperature, and 2 previous gear oil temperature values (AR-terms). When the model is used without the autoregressive terms, it does not follow the data well at all, indicating that the autoregressive terms are dominant and that the exogenous inputs contribute little. Additionally, Garcia et al. [80] combined a SARIMA model with a NARNN to forecast the daily energy output of a wind turbine. The SARIMA model was used to forecast high-power samples with seasonal patterns while the NARNN was used to forecast low-power samples without clear seasonal patterns. The performance of these models was compared against a persistence model, and showed improvements in forecasting accuracy with them.

As an RNN subset, NARXNN faces the problem of a vanishing gradient to some degree. After a certain number of inputs, RNN stops learning and negatively affects the prediction accuracy. This challenge occurs when the gradient descent is reduced by long-range dependencies, leading to memory reduction issues. The LSTM network was introduced to resolve the vanishing gradient problem in the RNN architectures. A combination of NARXNN and LSTM, called the NARX-LSTM model, which incorporates LSTM's capabilities to tackle the vanishing gradient issue, and takes into account the NARX architecture, is proposed to yield better results [28].

Massaoudi et al. [28] investigated the prediction of PV power from meteorological data using a NARX-LSTM approach. First, the NARXNN model acquires data to generate residual vectors, which were then implemented by LSTM, together with the original data. The aim here was to reduce the sensitivity of the network to time dependence, allowing to predict the PV performance and the associated uncertainty. The performance of the model was compared with that of various methods, including LSTM, NARXNN and FNN. The NARX-LSTM method outperformed the baseline method under various weather conditions, with an overall Normalized Root Mean Square Error (NRMSE) of 1.98% and 1.33% for the Australian and US datasets, respectively. Inspired by the work of [28], Gagnon et al. [81] proposed three architectures for strain prediction on hydraulic turbine blades based on shaft displacements. The proposed methodology aimed to replace missing or erroneous values in blade strain measurements, and comprises the following architectures: NARX-LSTM, Injector MA-Net, and a combination of both. The authors evaluated their method in four data loss cases, using virtual sensors. They showed that the Injector MA-Net and the combined architecture are capable of capturing the dynamic behavior of the blades, even with poorly correlated inputs, whereas the NARXNN, which performs well in cases where inputs are strongly correlated with outputs, struggles to capture variations between blades or instabilities of the draft tube. They emphasize the need to quantify the uncertainty, which is not addressed.

RNN can also be used for prediction through its combination with Autoencoders (AE), which are especially adapted for anomaly detection. Zhang et al. [82] proposed a method to estimate daily photovoltaic energy production at 15-minute intervals under changing weather conditions. Input parameters included temperature, relative humidity, wind speed, total cloud cover, solar radiation, and derived features such as solar zenith angle and time series attributes. Gensler et al. [83] proposed an LSTM-AE model for photovoltaic power prediction using meteorological data. The model's performance (as evidenced by the RMSE and MAE metrics) was significantly better than that of FNN, achieving the best average RMSE of 0.07 and MAE of 0.03, with excellent correlation with real data. Chen et al. [84] proposed an LSTM augmented model using Variational Auto-Encoders (VAE) for anomaly detection of WT gearboxes. The inputs were SCADA parameters such as wind speed, power, generator voltage, rotor speed, blade angle and ambient temperature, and the

outputs were gearbox oil and shaft temperature parameters. This model has been validated by actual cases and can effectively alert and locate transmission anomalies, having a lower false alarm rate than other models.

Regarding the use of CNN, Huang et al. [85] proposed a multiple-input deep CNN model for short-term PV power forecasting over the next 24 h. This model leverages meteorological data such as temperature, solar radiation, and historical PV system output data. The model's performance shines through its lower MAE of 109.4845 compared to alternative models such as LSTM (MAE = 124.0625), MLP (MAE = 196.6818), and SVM (MAE = 147.3763). Furthermore, Li et al. [86] introduced a combination of CNN and LSTM for wind speed prediction aimed at addressing inaccuracies and missing measurements from physical wind speed sensors. This approach, which is somewhat similar to an AR-model, combines the CNN's ability to extract spatial features from wind speed data across multiple turbines with the strength of the LSTM in capturing temporal dependencies. The method demonstrated an enhancement in sensor reliability in reconstructing missing or inaccurate wind speed data, achieving an RMSE of 0.45.

3.2.2. SVM-Based Methods

SVM is a supervised machine learning method used for classification and regression. Its goal is to find a higher or equal dimension data space that optimally separates object classes. Support Vector Regression (SVR), a variant of SVM, is used exclusively for regression by seeking a function that optimally predicts the real variable of a continuous variable based on training data while minimizing the error margin. SVM offers a highly flexible non-linear model for prediction due to its ability to learn without prior knowledge, unlike neural networks. Additionally, SVM does not encounter issues with local minima, as do neural networks. It facilitates the simplification of complex mathematical problems that may be associated with PV forecasting. However, SVM is highly sensitive to the parameters used in developing a forecasting model, making the appropriate selection of these parameters one of its challenges [87].

To improve wind speed (U) estimation for WT, Ji et al. [88] developed a virtual sensor using SVR enhanced by Sequential Minimal Optimization (SMO) to overcome the SVR's inherent challenges, particularly the extensive time required to process large datasets. They used readily measurable variables, such as the rotor rotation speed (Ω), generator power (P), and blade pitch angle (β), expressing their model as a nonlinear function $U = f(\Omega, P, \beta)$. The SVR algorithm, which addresses a quadratic programming problem, benefits from the efficiency of SMO. This method accelerates the optimization process by adjusting only two Lagrange multipliers simultaneously. In comparison with an FNN model using simulation data, their findings show that the SVR model, when integrated with SMO, surpasses the neural network both in predictive accuracy and generalization capabilities. Also, Yamin et al. [89] used the SVR approach to enhance the wind speed prediction spanning from 1 to 24 h in a wind energy production forecasting context. The SVR model integrated wind speed alongside other atmospheric conditions as input parameters, resulting in an 8 to 9% enhancement in forecasting accuracy compared to persistence models. Further contributions to WT forecasting include proposals by Wang et al. [90] and Shamshirband et al. [91]. Wang et al. [90] introduced a strategy for optimal power point tracking to maximize energy capture in WT through SVR. This approach dynamically adjusts system parameters for optimal wind energy capture, significantly enhancing the dynamic performance of generator speed control with inputs such as wind speed and turbine parameters (pitch angle, etc.). Meanwhile, Shamshirband et al. [91] investigated the use of SVR with polynomial and Radial Basis Function (RBF) kernels for predicting the WT reaction torque, incorporating inputs such as wind speed and rotor speed. Their methodology exhibits a superior predictive accuracy and generalization ability over conventional models.

Regarding the application of SVR in PV power prediction, Shi et al. [92] focused on enhancing the accuracy of PV system power output predictions by integrating weather classification with SVM. Their methodology involves classifying weather conditions and

then using SVM for power output forecasting, thereby improving the prediction accuracy by adapting forecasts to specific weather conditions. Essential inputs for this model include weather data and historical power output. Its accuracy is contingent upon the quality of weather classification and the input data. On the other hand, Wolff et al. [93] aimed at forecasting PV power output using SVR, incorporating inputs from PV power measurements, numerical weather predictions, and cloud motion vector irradiance forecasts. Their approach concentrated on optimizing the SVR parameters and training sizes to accommodate data under various cloud conditions. The study highlighted the effectiveness of SVR in scenarios where detailed specifications of the PV system are unavailable. Furthermore, Pan et al. [94] proposed an ultra-short-term PV power prediction method using an Ant Colony Optimization (ACO)-optimized SVM. The model predicts active power and measures received-delivered active energy, with input parameters including atmospheric temperature, relative humidity, global horizontal radiation, diffuse horizontal radiation, wind direction, and sampling time. The ACO-SVM model exhibited a regression coefficient (R^2) of 0.997, demonstrating high PV power prediction accuracy and indicating a significant performance for ultra-short-term forecasting periods.

3.2.3. PCA-Based Methods

PCA is a statistical technique used for dimensionality reduction in data analysis. It transforms a set of possibly correlated variables into a smaller set of uncorrelated variables, called principal components. These components are ordered such that the first few retain most of the variation present in all of the original variables. PCA simplifies the complexity in high-dimensional data while retaining trends and patterns.

Tarpø et al. [95] aimed to enhance structural health monitoring in offshore wind turbines by estimating dynamic strain. They employed the PCA method on vibration data to reduce the dimensionality of these inputs, retaining only relevant elements for deformation prediction. Using the eigenvector obtained through this process, they constructed a transformation matrix to map the inputs into the domain of principal components. This transformation allowed them to estimate the deformation response. The PCA-based virtual sensing technique was validated through a two-month monitoring campaign, demonstrating precise strain estimation over extended periods, with minimal training data, and a normalized error of 2.95%, thus offering a robust alternative to traditional monitoring methods.

3.2.4. Gaussian Process-Based Methods

Gaussian Process (GP) is a non-parametric supervised machine learning method, widely used in regression and classification tasks. Especially in indirect measurement contexts, GP is used to solve regression problems [96,97].

Consider a MISO regression problem, where we have a function f that takes an input vector $\mathbf{x} \in \mathcal{R}^D$ and produces a scalar output y , as shown by Equation (5). The output is subject to additional noise ϵ , which is assumed to be a Gaussian white noise in basic GPR, i.e., $\epsilon \sim \mathcal{N}(0, \sigma_n^2)$. Equation (5) can be represented graphically by Figure 3.

$$y = f(\mathbf{x}) + \epsilon \quad (5)$$

To build a GPR model following the common practices in machine learning, a training dataset $\mathcal{D} = \{\mathbf{x}_i, y_i\}_{i=1}^n$ is used, consisting of n pairs of input vectors and the corresponding observed outputs. It is hypothesized that a latent function f correlates with the training data, as demonstrated by Equation (6).

$$y_i = f(\mathbf{x}_i) + \epsilon_i \quad (6)$$

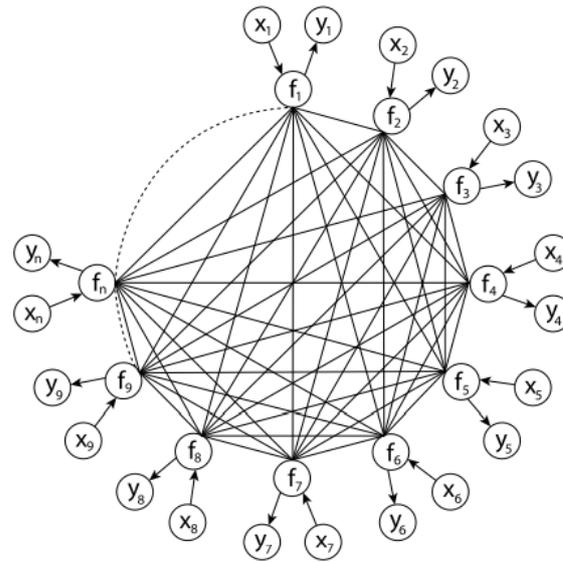


Figure 3. Graphical model of the Gaussian process [96].

The primary assumption of GPR is based on assuming a priori that the function values behave according to a multivariate Gaussian distribution, where $f = [f(x_1), f(x_2), \dots, f(x_n)]^T$ is a vector of latent function values, as shown by Equation (7).

$$p(f|x_1, x_2, \dots, x_n) = \mathcal{N}(0, \mathbf{K}) \quad (7)$$

where \mathbf{K} is a covariance matrix, and elements $K_{i,j}$ are given by the kernel function k evaluated at each pair of training inputs, as stated by Equation (8).

$$K_{i,j} = k(x_i, x_j) \quad (8)$$

This kernel function governs the family of functions which comprise the prior knowledge, and the choice of this function is critical to the performance of the GPR [96,97]. An instance of typically adopted kernel function is the squared exponential kernel, or the basic linear function as shown by Equation (9).

$$k(x_i, x_j) = \sigma_f^2 x_i x_j^T \quad (9)$$

where σ_f^2 is a hyperparameter. It should be noted that selecting the best kernel is a challenging task, and there are no one-size-fits-all solutions. Practitioners should prioritize finding the right covariance function before starting any modeling with GPR [96,97].

Pandit et al. [98] used GPR models incorporating the rotor speed, blade pitch angle, wind speed and power output to improve WT power curve predictions, demonstrating high accuracy with their approach, with an R^2 coefficient of 0.9960. Similarly, Paiva et al. [99] applied GPR to estimate wind speed from SCADA measurements such as the active power, rotor speed, and pitch angle of the blades, achieving remarkable accuracy with a median errors of less than 1% compared to traditional methods such as LiDAR measurements. Rajiv et al. [100] developed Gaussian process metamodels, including GP, Variational Heteroscedastic Gaussian Process (VHGP), and Sparse Spectrum Gaussian Process (SSGP), to simulate the dynamics of floating offshore WT, using the wave force and moment data to predict the platform movements (displacement and rotation). The model demonstrated improved prediction accuracy with lower computational demands. Moreover, Haghi et al. [101] explored the use of GPR to forecast WT Damage Equivalent Load (DEL) across different wind speeds. They combined SCADA measurements with simulated data as inputs and validated the model's accuracy through comparison with actual measurements. The aforementioned studies emphasized GPR's capability in enhancing WT

performance assessment and health management, despite facing challenges related to the necessity for precise input data and overcoming computational complexity.

Regarding the use of GPR for PV power plant predictions, Kanwal et al. [102] conducted a comparative analysis between GPR and SVM for PV power output forecasting. They used irradiance and temperature data as inputs to predict the power generation of the PV system. The results showed that SVM surpassed GPR in terms of solar power prediction accuracy, as indicated by the lower RMSE values associated with SVM. This superiority suggests that SVM is more able of handling the variability in PV power output under diverse weather conditions.

4. Physics-Based Methods

Physical models use mathematical equations to simulate the dynamics and physics of structures, typically described by partial differential equations. Due to their non-linear nature, these equations often lack a unique solution and are challenging to solve analytically. Their derivation is a complex process, influenced by the chosen model and assumptions, which usually deviate from real-world scenarios. In practice, data-based or hybrid models are commonly employed, although purely physical models remain of significant interest for their precise representations.

Regarding the PV sector, Dolara et al. [103] conducted an evaluation of various physical models in predicting the PV power output. The central idea revolves around using equivalent electric circuits to represent PV cells, capturing their electrical behavior under different conditions. These circuits use parameters such as series resistance, shunt resistance, and diode characteristics to simulate how PV cells convert radiation into electricity. The complexity of these models varies: three-parameter models are simpler, employing fewer variables, while five-parameter models provide more detailed representations for improved accuracy. By closely mimicking the physical and electrical properties of PV cells, these models aid in understanding and predicting PV power output. Environmental and operational data serve as inputs, while the predicted power output of PV modules constitutes the output. Notably, the study reveals that the model performance depends on the calibration data used, and that complex models do not consistently outperform simpler ones. Additionally, limitations include reliance on accurate calibration data and varying effectiveness across different PV module types.

In their study, Ibrahim et al. [104] introduced a digital twin technology as a virtual sensor for WT. The purpose was to address inaccuracies and failures associated with physical wind speed sensors. To this end, they employed a test bench featuring a direct-drive WT equipped with a permanent magnet synchronous generator for the physical model. The virtual model was constructed by proposing an equivalent electrical circuit for the generator and an equivalent mechanical model for the drive train. Firstly, the electrical block of the virtual model receives input parameters such as current, voltage from the stator, and rotor rotation speed. Using these inputs, it calculates the power produced by the generator. After that, the mechanical block of the virtual model combines the generated power with the rotor speed and torque input. This combination allows to estimate of the wind speed required to produce a specific mechanical power output. The assumption here is that the wind turbine system operates within certain efficiency and power capture characteristics, which are integrated into the mathematical model. The researchers validated their model using real-world data, demonstrating its significant potential in estimating wind speeds with notable accuracy. However, they acknowledge limitations related to external environmental factors that were not fully taken into account in the simulation.

In the literature on asset monitoring within the energy sector via physical modeling-based methods, modal analysis emerges as the predominant approach. This method involves examining the natural modes of a structure to predict its dynamic behavior. Researchers at Vrije Universiteit Brussel in Belgium have applied this approach with

the goal of predicting vibrations and strains in areas of offshore wind turbines that are inaccessible (submerged) [105–109].

Iliopoulos et al. [105] used the modal expansion method on WT to predict acceleration responses at locations where measurements were unavailable, relying on sparse acceleration data from accessible points. They first developed a finite element model to numerically estimate the natural modes of the structure composed of n -DOF, with calibration performed using experimentally obtained natural modes. The Modal Assurance Criterion (MAC) is used to assess the agreement between numerical and experimental results. When the agreement is sufficiently good in accessible areas, corresponding to a -DOF, it can be concluded that the numerical model accurately represents the areas where measurements are unavailable, corresponding to b -DOF, and giving rise to the global modal matrix $\Phi_{n \times m}$, where m denotes the number of considered modes. The global modal matrix is composed of two sub-matrices: one for the measurable a -DOF $\Phi_{a \times m}^{mes}$ and the other for the unmeasurable b -DOF $\Phi_{b \times m}^{um}$, where (mes) indicates measurable quantities and (um) refers to the unmeasurable DOF. Similarly, the acceleration vector $a(t)$ of the total n -DOF, where $n = a + b$, consists of two sub-vectors: one measurable a_{mes} , and the other unmeasurable a_{um} . If $m = a$, then the matrix of measurable natural modes is square and therefore invertible. Using modal expansion, the acceleration of the n -DOF of the structure is given by Equation (10).

$$a(t) = \Phi \cdot q(t) \quad (10)$$

where $a(t)_{1 \times n}$ is the acceleration vector at the DOF of the structure. Φ is the matrix of the structure's natural modes, composed of n natural modes corresponding to n -DOF. However, it is possible to consider a reduced number of modes, denoted by m (i.e., opting for an approximation rather than a complete discrete analysis) with $m < n$, and in this case, the matrix is of dimension $\Phi_{n \times m}$. Furthermore, $q(t)_{1 \times m}$ corresponds to the modal coordinates.

The measurable accelerations a_{mes} at the a -DOF can be conveniently expressed as given by Equation (11).

$$a_{mes}(t) = \Phi_{mes} \cdot q(t) \quad (11)$$

As for the b unmeasurable DOF, the relation between their accelerations and the modal (natural) coordinates is expressed in Equation (12).

$$a_{um}(t) = \Phi_{um} \cdot q(t) \quad (12)$$

The modal coordinates $q(t)$ can then be obtained from the following expression if the number of modes is equal to the number of measurable DOF ($m = a$), i.e., in the case where the matrix Φ_{mes} is square and invertible, as shown in Equation (13).

$$q(t) = \Phi_{mes}^{-1} \cdot a_{mes}(t) \quad (13)$$

And in this case, the unmeasurable accelerations are expressed by Equation (14).

$$a_{um}(t) = \Phi_{um} \cdot \Phi_{mes}^{-1} \cdot a_{mes}(t) \quad (14)$$

If the matrix Φ_{mes} is not square and thus not invertible (if $m < a$), the expression of $q(t)$ is established using the Moore-Penrose pseudo-inverse, and given by Equation (15).

$$q(t) = (\Phi_{mes}^T \cdot \Phi_{mes})^{-1} \cdot \Phi_{mes}^T \cdot a_{mes}(t) = \Phi_{mes}^\dagger \cdot a_{mes}(t) \quad (15)$$

And in this case, the unmeasured accelerations are expressed by Equation (16).

$$a_{um}(t) = \Phi_{um} \cdot \Phi_{mes}^\dagger \cdot a_{mes}(t) \quad (16)$$

This approach allows predicting acceleration in inaccessible areas of wind turbines using modal expansion. The method is validated by comparing the predictions obtained with on-site measurements under various loading conditions, using the Time Response

Assurance Criterion (*TRAC*), which, along with the analysis of measured and predicted signals, demonstrates excellent agreement between predictions and measurements. The Fourier amplitude spectrum of modal accelerations, obtained by double differentiation of modal displacements in the frequency domain, reveals the distinct influence of each mode. Three regions are identified, indicating signal decomposition into specific frequencies, thus leading to the prediction of strains in critical areas [106] with limited acceleration and strain measurements, employing the same modal expansion approach established by Equation (17), where $\epsilon(t)$ represents the strain and Ψ is the modal shape matrix of strains.

$$\epsilon(t) = \Psi \cdot q(t) \quad (17)$$

The results of the prediction of dynamic strains under shutdown conditions show excellent agreement between measured and predicted signals. However, although the accelerometers used perform well for predicting dynamic strains, they have limitations in the case of quasi-static strains, as they are insensitive to low-frequency vibrations corresponding to this type of strain.

To overcome this limitation, Iliopoulos et al. [107] introduced the principle of multi-band modal decomposition, which consists in dividing the prediction regions into 3. The first is the near-static region for frequencies $f < 0.05$ Hz corresponding to aero and hydrodynamic load perturbations, captured by strain sensors. The second region is between 0.05 and 0.5 Hz, corresponding to the fundamental mode, known as the low-frequency dynamic region where one accelerometer is used to capture the dynamic response. Finally, the third region goes from 0.5 to 5 Hz, and is called the high-frequency dynamic region, and includes harmonic modes, where multiple accelerometers are used. The strain corresponding to the high-frequency zone was estimated through a double integration of acceleration modal coordinates. Furthermore, the authors estimated the quasi-static strain, attributed to low-frequency wind and wave loads, through static analysis employing a tuned finite element model of the WT and load approximation at the tower's top, derived from measured bending moments in the pile [108,109]. The validation of the proposed approach was conducted using strain data obtained from an offshore WT situated in the Belgian North Sea.

5. Hybrid Models

In addition to data-driven methods and those based on physical models, hybrid approaches combining these two methods also exist. Such methods generally start with a State-Space (SS) modeling approach, followed, in most cases, by the use of a Sequential Bayesian Updating (SBU), such as Kalman Filter (KF). Nonetheless, there have been recent advances in hybrid methods such as Physics-Informed Gaussian Processes (PIGP), which include Gaussian Process Latent Force Models (GPLFM). Furthermore, Physics-Informed Neural Networks (PINN) have emerged as another noteworthy hybrid approach.

5.1. SS Hybrid-Based Methods

The SS modeling approach, also known as an SS representation, is used to describe dynamic systems through state and measurement equations, expressed in terms of state variables. This approach allows predicting the temporal evolution of a dynamic system. For a linear, time-invariant system with m -inputs, q -outputs, and n -state variables, the state-space representation is given by Equation (18).

$$\begin{cases} \dot{x} = Ax + Bu \\ y = Cx + Du \end{cases} \quad (18)$$

Here, $x(t)$ is the state vector of dimension n , $u(t)$ is the vector of deterministic inputs of dimension m , and $y(t)$ is the measurement vector of dimension q . The matrices $A_{n \times n}$, $B_{n \times m}$, $C_{n \times n}$, and $D_{n \times q}$ represent, respectively, the state, input, output, and transfer

matrices of the system. The first equation is the state equation, and the second is the measurement equation.

The transition from the SS model to the use of SBU allows for a more accurate and robust estimation of a system's state by integrating new measurement data. SBU is a statistical method based on Bayesian inference, which updates the probability distribution of a system's state with each new set of data. This approach relies on Bayes' theorem to merge new evidence (the likelihood of observing new data, given the state) with prior beliefs (the prior probability of the state), thus forming an updated belief (the posterior probability of the state). This recursive process is particularly useful for handling uncertainties and continuously integrating new information, making state estimates more precise and robust against noise and various uncertainties. In practice, SBU enables the sequential updating of state estimates as new data becomes available, provides predictive estimates of future states, combines data from multiple sources, and estimates non-observable states by linking observable outputs to hidden states.

The methodology introduced by Nabiyan et al. [110] involves a sequential Bayesian approach for FEM updating in the context of digital twins for offshore WT. The approach aims to enhance the accuracy of digital twin models by incorporating sparse real-world measurement data, such as strain responses from the turbine structure, as inputs. By using a time-domain framework, the method allows for the simultaneous estimation of unknown model parameters and the history of input forces, such as the wind load, acting on the turbine. The outputs of this methodology include not only the updated digital twin model, which reflects the real condition of the wind turbine more accurately, but also, detailed information on the system's unmeasured responses. This means the method can predict how the WT will respond to various conditions, even in areas where sensors may not be present. Nabiyan et al. [110] compared the proposed methodology with a classical modal-based updating method. Performance was evaluated on a 2 MW offshore WT at the Blyth wind farm, demonstrating both methods' effectiveness in accurate response prediction, with the Bayesian approach showing slightly better results and additional capabilities in terms of input load identification and uncertainty quantification. In addition, Teymouri et al. [111] developed a Bayesian Expectation-Maximization (BEM) methodology for coupled input-state-parameter-noise identification in dynamical systems, incorporating dummy observations to stabilize estimations. This advanced approach aims to enhance the accuracy of system identification and damage detection by quantifying uncertainties and propagating them through model parameters and input forces. Using vibration data, the inputs are operational and environmental conditions, while outputs include the estimated states, parameters, and noise characteristics. The BEM model demonstrates improved accuracy in state and parameter estimations across numerical and experimental examples, highlighting its potential for real-time applications, with limitations including challenges in handling partial knowledge systems and sparse measurements.

Also, the Kalman filter, in all its form (linear, augmented, etc.), is a specific application of sequential Bayesian updating, tailored to the context of dynamic systems. This filter allows estimating the state of a dynamic system from measurements which may be incomplete or noisy. In a discrete environment, the Kalman filter is a recursive Bayesian estimator that relies solely on previous state estimates and current measurements, without requiring a complete history of observations and estimates. The process of the Kalman filter is divided into two main phases: prediction, where the future state is estimated using the system model, and update, where actual measurements adjust the initial prediction. The linear Kalman filter assumes that the actual discrete process (x_t) follows a linear evolution defined by Equation (19) below, where w_t is the process noise, following a centered normal distribution.

$$x_t = Ax_{t-1} + Bu_t + w_t \quad (19)$$

To deal with nonlinear systems, variants of the Kalman filter have been developed. The Extended Kalman Filter (EKF) linearizes the model equations around the current estimate, while the Unscented Kalman Filter (UKF) employs a nonlinear approach without

explicit linearization, based on a set of sigma points, to estimate the probability distribution of the system state. The Unscented filter provides better performance in highly nonlinear situations.

Perisic et al. [112] introduced a cost-effective approach for CM in wind turbines by modeling their mechanical transmission as a 2-DOF dynamic system. Their research aimed to enhance wind turbine maintenance strategies without the high costs associated with traditional measurement equipment. They applied an Augmented Kalman Filter with Fading Memory (AKFF) to predict the shaft torque from typical wind turbine data. Inputs for their model were the generator torque and the angular velocities of both high- and low-speed shafts, with the output being the estimated shaft torque. Their approach outperformed existing models such as the standard Augmented Kalman Filter (AKF), demonstrating better accuracy across various load conditions and noise levels. This indicates its practical utility in predictive maintenance for wind turbines. In the same context of WT gearbox torque prediction, Cappelle et al. [113] presented an approach based on an Augmented Extended Kalman Filter (AEKF). The model inputs are vibration measurements on high and low-speed shafts, as well as strain gauges positioned on the gearbox casing. The performance is evaluated through numerical validation, achieving high accuracy in torque estimation with an *NRMSE* of 1.27% and 0.02% for the low- and high-speed shaft torque, respectively, under specific sensor configurations.

In another context, where offshore WT face technical measurement constraints in the submerged part of the structure, Maes et al. [114] aimed to predict a continuous strain for fatigue assessment in WT using Kalman filtering techniques. They developed a response estimation technique to estimate the strain at critical, inaccessible locations of a WT, using a limited set of response measurements and a system model. Their method was validated with data from a monitoring campaign on an offshore Vestas V90 3 MW wind turbine. The inputs for the model included acceleration and strain measurements, while the outputs were the estimated strains at unmeasured locations. The Kalman filter successfully predicted strains at the minimum measurement height available ($h = 19m$), demonstrating excellent agreement with measured strains. Noppe et al. [115] explored fatigue stress estimation for offshore WT using a Kalman filter and accelerometers. They compared the modal decomposition and expansion (presented in section 4) using KF techniques, using data from an offshore WT. The model inputs included accelerations and a measured thrust load signal, aimed at reconstructing the strain history at critical points. Their findings showed that both techniques performed well in time domain analysis, with modal decomposition and expansion offering slightly better results in the frequency domain. Shukur et al. [50] presented a hybrid Kalman Filter-Artificial Neural Network (KF-ANN) model, grounded in the ARIMA model, to enhance wind speed forecasting accuracy. This methodology combines ARIMA's effectiveness in initial data structuring with the Kalman filter stochastic uncertainty management and neural network nonlinearity handling. Inputs include daily wind speed data, processed to predict wind speeds accurately. The hybrid KF-ANN model outperforms its individual components and other models such as ARIMA in forecasting accuracy, demonstrating its potential for reliable wind speed prediction, albeit with the inherent limitations of computational complexity and data quality dependence.

Branlard et al. [116] developed a digital twin model for real-time load and fatigue estimation of WT. This model combines OpenFAST linearizations with a Kalman filter algorithm to estimate WT loads and fatigue states from SCADA measurements, including tower-top acceleration, pitch, generator torque, and rotational speed, without extensive physical instrumentation, enhancing wind turbine monitoring and maintenance. Validation against GE 1.5 MW turbine measurements showed an average accuracy of about 10% in estimating tower bottom moment loads, demonstrating the model's potential for practical application in wind energy operations. Mehlan et al. [117] also developed a digital twin framework for online fatigue damage monitoring of wind turbine gearboxes, using vibration and SCADA data for dynamic state estimation. The methodology involves inverse methods for load estimation, specifically the Kalman filter, least squares, and quasi-static

approaches. Inputs include the generator torque and vibration data, with outputs being the estimated gearbox bearing loads and fatigue damage. The performance evaluation showed a moderate-to-high correlation ($R = 0.50 - 0.96$) of estimated loads with actual measurements, and fatigue damage estimation errors ranged between 5 and 15%. The study highlighted the approach's potential for enhancing maintenance efficiency through improved monitoring capabilities. The same methodology was extended to include the estimation of drivetrain loads and accumulated fatigue damage [118], calculated using the Palmgren–Miner model. The performance evaluation in a numerical case study showed moderate to high correlations between estimated and actual measurements, with fatigue damage estimation errors ranging from 5% to 15%. Kamel et al. [119] developed a data-driven virtual sensor for estimating the internal loads on the drivetrain bearings of WT using SCADA data (generator rotation and power, gear displacements, and bearing acceleration). They first identified a linear state-space model, and instead of applying a Kalman filter as is usually done, they attempted to explore a linear state-space model augmented by a FNN to predict the error between the dynamic response of the linear model and the actual system response. This approach achieved a 98% correlation coefficient in the time domain and accurate frequency signature matching, indicating a high performance in online load estimation for wind turbine monitoring and predictive control applications.

Regarding the use of SS-SBU methods in the PV sector, Tuyishimire et al. [120] developed a Kalman predictor approach for forecasting solar PV power generation, targeting real-time market scenarios with 15-minute-ahead predictions based on one-minute interval data. It contrasts two designs for a steady-state variance versus a transient response, using data from a 13 kW PV array. The model aims to improve the forecasting accuracy for solar power generation, with inputs being the high-resolution operational data of PV systems. Outputs include forecast power generation values. The study demonstrates the potential of the KF in enhancing solar PV forecasting, although it acknowledges trade-offs in design choices affecting the performance.

5.2. PIGP-Based Methods

Unlike traditional GPR, which relies on observed data for prior statistics, PIGP models use stochastic differential equations derived from physical models.

Tartakovsky et al. [121] introduced a PIGP model adapted for real-time forecasting and state estimation within WT power grids. This method enhances standard GPR by integrating physical laws, through stochastic differential equations, and thus offers improved prediction of the phase angles, angular speeds, and wind mechanical power. The model used inputs that include sparse measurements of observed states (e.g., phase angles, angular speed) and unobserved states (e.g., wind mechanical power) which are treated as stochastic processes. The PIGP model addressed both observed and unobserved grid states, leveraging stochastic processes for inputs and showing significant advancements, in terms of accuracy and computational efficiency, over conventional forecasting techniques, such as the Kalman filter and ARIMA.

Additionally, GPLFM integrates Bayesian machine learning techniques with physical modeling frameworks to analyze complex systems. Using Gaussian processes, the GPLFM infers latent or unidentified forces acting upon a system, offering a dynamic and robust methodology for internal dynamics simulation. This approach is notably effective in embedding prior physical laws into statistical models, thus improving the accuracy of predictions and enriching the comprehension of fundamental processes.

Bilbao et al. [122] explored the application of GPLFM for the monitoring of fatigue loads on wind turbine towers through virtual sensing, by employing acceleration measurements to forecast the dynamic strain. This model facilitates the estimation of both the dynamic loading applied and the resultant strain responses, with its validity is confirmed by comparison with actual strain gauge measurements. The evaluation of the model's performance reveals its proficiency in precise strain response estimation, suggesting a pathway to more efficient structural health monitoring for wind turbines. Nevertheless, the

model's effectiveness is contingent upon the availability of accurate acceleration data and the intricacy involved in optimizing its hyperparameters. Further, Zou et al. [123] investigated the use of GPLFM for virtual sensing in offshore WT to estimate the dynamic strain under the mudline. This approach addressed the challenges associated with direct strain measurement in such complex environments. The model's performance was confirmed by actual strain data, demonstrating its potential for continuous monitoring. However, its application is limited by the necessity for precise prior knowledge and the difficulties in adapting to varied operational contexts.

5.3. PINN-Based Methods

PINN combine deep learning with physical laws to tackle both forward and inverse problems in nonlinear partial differential equations. They incorporate physical constraints into the training process, providing a strong framework for finding solutions to complex equations and even identifying the equations themselves. This breakthrough in computational science enables accurate predictions and a deeper understanding of complex system behaviors. It represents a major advancement in applying machine learning to the physical sciences [124].

Li et al. [125] introduced a physics-informed deep learning approach that combines RNN and Deep Residual RNN (DR-RNN) architectures with the structural dynamics of wind turbines to forecast their responses under dynamic operational conditions. This model used inputs such as the wind speed and turbine control settings, in order to predict structural outputs such as displacements and forces, showcasing enhanced prediction accuracy and robustness compared to traditional methods, particularly for long-term forecasts. However, it encounters challenges in accurately predicting the high-frequency responses, underscoring the necessity for further development to integrate system SS equations across various operational scenarios. Further, Cobelli et al. [126] investigated the application of PINN in modeling the wind fields within wind farms, employing sparse velocity data from numerical simulations for training. Their findings illustrate PINNs' capability to accurately simulate the wind fields around WT, presenting a computationally efficient alternative to conventional simulation techniques. This study validates the effectiveness of PINN in wind field modeling, though it suggests the need for model refinement and broader application to complex wind farm layouts.

6. Discussion

After reviewing various methods of indirect measurement applied to the field of renewable energies, it is clear that data-based methods are the most prevalent. Deep learning-based methods are currently the most used. The simplest approach in deep learning involves the use of FNN, which represents the foundation of basic deep learning, and demonstrates remarkable efficiency. However, the main limitation of FNN models is their inability to take into account temporal dependencies, rendering them less suitable for time series prediction, when time dependence is expected. Most monitored physical quantities are dynamic in nature, meaning that the temporal effect is present, and sometimes even predominant, which accounts for the effectiveness of AR models in cases where the system is linear. For this reason, RNN can excel in time series prediction due to their recurrent memory effect, which takes into account the temporal effect. It is important to highlight that the effectiveness of RNN models is strongly influenced by the nature of the data sequence being processed. When attempting to capture long temporal dependencies in data characterized by high-frequency phenomena, which exhibit weak temporal dependencies, there's a risk of introducing noise into the model. This approach can, therefore, decrease the accuracy of the predictions. Conversely, overlooking long temporal dependencies in situations where they are relevant can also reduce the model's effectiveness. It should be noted that models purely based on data may not be sufficient to answer all questions, highlighting their intrinsic limitations. Furthermore, it's crucial to consider the high computational cost associated with the use of data-based methods, particularly for RNN. This

factor represents a significant challenge in the implementation and optimization of these models, requiring special attention during their design and use.

Another data-driven approach involves the use of SVM, which offer high flexibility, and unlike neural networks, do not have to deal with the local minima problem. Several applications of SVM have been proposed for monitoring wind turbines and photovoltaic systems and have been shown to be as efficient as neural networks, although they are still not as powerful as neural networks in terms of adaptability and overall generalization capabilities. Meanwhile, the PCA approach remains underused in the renewable energy sector, making it difficult to assess its effectiveness. Additionally, over the past five years, the adoption of GP, representing advanced methodologies in probabilistic machine learning, has marked a significant evolution, particularly in the wind energy sector. These techniques have proven effective for predicting and optimizing the performance of wind turbines, exhibiting remarkable accuracy. The flexibility of GP and their inherent ability to model uncertainty provide a robust framework for predictive modeling and analysis. However, their application in the renewable energy domain remains less widespread compared to that of deep learning. In this context, models combining Gaussian processes and neural networks, known as Neural Network Gaussian Processes (NNGP), represent a promising avenue. Nonetheless, to our knowledge, this approach has not yet been explored in the renewable energy sector.

Overall, there is a trend towards the use of RNN models in virtual sensors for predicting various physical quantities across the three types of energy sources. However, it must be acknowledged that the choice of neural network architecture and training approach significantly influences the accuracy of predictions. Moreover, machine learning methods generally create a virtual sensor which operates within a fixed system health state [41,42]. Once degradation occurs and the behavior changes, thus the model encounters data not seen during training, the entire model becomes obsolete (due to out-of-distribution issues), often without any warning [40–42]. Furthermore, while deep learning has made significant strides, there is a pressing need for further improvements, especially in terms of interpretability [42]. Additionally, the current lack of focus on uncertainty estimation in studies highlights an area that requires more comprehensive exploration in future research to ensure the accuracy of predictions.

Physics-based methods continue to garner increasing interest, particularly in the context of wind turbines, but they present either a very complex mathematical formulation or sometimes an overly simplistic one. When the mathematical formulation is complex, its implementation in computational models becomes difficult, and the computational cost becomes even higher [15,34]. On the other hand, using simplifying assumptions reduces computation time at the expense of accuracy. It is also noted that the use of physics-based models is less prevalent in virtual sensor structures as compared to their counterparts (data-based and hybrid approaches), but they are instead used indirectly in research proposing models that could be extended to a virtual sensor structure. In the context of offshore wind turbines, modal expansion is widely used. However, the accuracy of this method heavily depends on the defined model. To achieve significant accuracy, the analytical formulation of the model becomes very complex, leading to numerous challenges in computation, and in the modeling of certain phenomena that are still poorly understood [13,15,34,44].

Regarding hybrid methods, we have observed a growing interest among researchers in hybrid approaches that demonstrate excellent performance by combining the advantages of data-based methods with a foundational layer of physics. The vast majority of these methods rely on sequential Bayesian updating, although recent attempts have been made to integrate state-space models with neural networks, as demonstrated by the application mentioned in the work of [119]. Other methods that blend physics with machine learning exist but have not yet been exploited enough in the context of virtual sensors applied to renewable energies, such as PIGP and PINN, which have the ability to incorporate the knowledge of physical laws governing a dataset into the learning process.

This analysis also underscores that research on indirect measurements in the area of renewable energy is relatively nascent. Studies concentrate mainly on the implementation of virtual sensors, which have yet to achieve technological maturity. The proposed models in these studies are highly diverse, pointing to the need to develop a comprehensive and tailored methodology—a gap that none of the reviewed studies have yet addressed. Furthermore, it must be acknowledged that most research efforts are confined to a single dimension or attempt to estimate just one relevant quantity, leaving many potential avenues unexplored. This comprehensive examination reveals the intricate dynamics of indirect measurement methods in renewable energy, highlighting the critical need for enhanced methodologies, accuracy, and understanding to advance the field.

Furthermore, it would be relevant to develop a general methodology for conducting indirect measurements in the renewable energy arena. Adopting a strictly physics-based approach as a general method seems impractical, given the complexity of modeling both structural dynamics and fluid dynamics of an energy asset. As previously highlighted, the mathematical formulation is extremely challenging, and certain phenomena remain poorly understood to this day. Moreover, a complex model induces a significant computational load, while an overly simplified model can significantly reduce the accuracy of the results. By choosing a data-driven approach, one might consider using RNN or GP, which seem to better adapt to asset variations. However, these models suffer from a lack of interpretability, a common limitation of data-based methods, thus complicating diagnosis. Additionally, the uncertainty of predictions, often neglected in the literature, and the risk of model obsolescence when encountering new, unseen data during training, remain major challenges, although these approaches outperform physical models in terms of flexibility. Kalman filters, mainly used in the wind turbine sector, represent a promising alternative that, in some cases, may surpass RNN in accuracy [127]. However, their effectiveness heavily depends on the underlying physical model, making it difficult to generalize their performance. Approaches such as PIGP and PINN show significant potential for creating a general model of indirect measurements but remain underexploited in the renewable energy field. In conclusion, to develop a generic model applicable to renewable energies, capable of conducting indirect measurements through a virtual sensor structure, viable paths include FNN, RNN, KF, PINN, GP, PIGP, and NNGP.

A summary of all the methods discussed is provided in Table 3 below.

Table 3. Summary of methods used for indirect measurements

Method	Energy	Year	Reference	Input Data	Output Data
ARMAX	HP	2016	[14]	Shaft torsion	Blade strain
ARIMA	PV-WT	2020	[47]	Wind speed, temperature	Wind speed, temperature
VAR	PV	2023	[48]	Weather conditions, solar irradiance	Solar irradiance
ARDA	WT	2021	[49]	Wind power	Wind power
ARIMA-FNN	WT	2023	[51]	Wind speed	Wind speed
SARIMA-SVM	PV	2023	[52]	PV power	PV power
FNN	HP	2019–2021	[54,55]	Shaft vibrations and displacement	Blade strain
FNN	WT	2021	[56]	MBS simulations	Gearbox loads
FNN	PV	2008	[58]	Weather parameters	24 h-PV power
FNN	PV	2010	[59]	Weather parameters	24 h-PV power
FNN	PV	2011	[60]	Weather parameters	24 h-PV power

Table 3. Cont.

Method	Energy	Year	Reference	Input Data	Output Data
FNN	PV	2011	[61]	Weather parameters, PV power	24 h-PV power
FNN	PV	2014	[62]	Weather parameters	24 h-PV power
FNN	PV	2015	[63]	Weather parameters	24 h-PV power
FNN	PV	2015	[64]	Weather parameters	24 h-PV power
FNN	PV	2010	[65]	Weather parameters	PV power
FNN	PV	2012	[66]	Weather parameters	PV power
LSTM	WT	2022	[68]	SCADA	Blade bending moment
LSTM	WT	2022	[70]	Vibrations, strain, FEM simulation	Gearbox loads
LSTM	WT	2015	[71]	Wind speed, temperature, surface pressure	Wind speed
LSTM	WT-PV	2019	[72]	Wind/PV power, weather parameters	Wind / PV power
LSTM	PV	2019	[73]	Meteorological data, clearness index	Solar irradiance
LSTM-VQC	PV	2023	[74]	Meteorological data, solar radiation	Solar irradiance
LSTM	PV	2020	[75]	Irradiance, temperature	Power output
LSTM	PV	2021	[76]	Historical power data	Power output
LSTM	PV	2019	[77]	Historical power data, meteorological data	Power output
NARXNN	PV	2017	[78]	Meteorological data	Power output
NARXNN	WT	2012	[79]	SCADA	Gear oil temperature
NARNN-SARIMA	WT	2019	[80]	Meteorological data	Wind power
NARX-LSTM	PV	2021	[28]	Meteorological data	PV power
NARX-LSTM	HP	2022	[81]	Shaft displacement	Blade strain
LSTM-AE	PV	2020	[82]	Meteorological data	PV power
LSTM-AE	PV	2016	[83]	Meteorological data	PV power
LSTM-VAE	WT	2023	[84]	SCADA	Gearbox oil temperature
CNN	PV	2019	[85]	Temperature, solar radiation, historical power output	Power output
CNN-LSTM	WT	2022	[86]	Wind speed from physical sensors	Accurate wind speed
SVR	WT	2008	[88]	Rotor rotation speed, generator power, blade pitch angle	Wind speed
SVR	WT	2022	[89]	Wind speed, atmospheric conditions	Wind speed
SVR	WT	2022	[90]	Wind speed, WT parameters	Optimal power point tracking
SVR	WT	2014	[91]	Wind speed, rotor speed	Reaction torque
SVM	PV	2012	[92]	Meteorological data	PV power
SVR	PV	2016	[93]	Historical power data, weather, irradiance	PV power

Table 3. Cont.

Method	Energy	Year	Reference	Input Data	Output Data
SVM-ACO	WT	2020	[94]	Meteorological data	Power output
PCA	WT	2022	[95]	Vibrations	Dynamic strain
GPR	WT	2020	[98]	Rotor speed, blade pitch angle, wind speed, power output	Power curve
GPR	WT	2021	[99]	SCADA : active power, rotor speed, pitch angle	Wind speed
GPR	WT	2023	[100]	Wave force, moment data	Offshore WT platform movements
GPR	WT	2024	[101]	SCADA	Damage equivalent load
GPR/SVM	PV	2018	[102]	Irradiance, temperature	PV power
Physic model	PV	2015	[103]	Environmental and operational data	PV power
Physic model	WT	2023	[104]	Current, voltage, torque, rotor rotation speed	Wind power
Modal	WT	2014–2017	[105–109]	Vibrations and strains from accessible areas	Strain of inaccessible areas
SS-SBE	WT	2021	[110]	Strain responses	Wind load
AKFF	WT	2015	[112]	Generator torque and angular velocities of shaft	Shaft torque
AEKF	WT	2021	[113]	Vibrations and strains of shafts	Shaft torque
KF	WT	2015	[114]	Vibrations and strains of accessible areas	Strain of inaccessible areas
KF-Modal	WT	2018	[115]	Vibrations and strains of accessible areas	Strain of inaccessible areas
ARIMA-KF-ANN	WT	2015	[50]	Daily wind speed	Accurate wind speed
KF	WT	2020	[116]	SCADA	External loads and fatigue
KF	WT	2022–2023	[117,118]	Generator torque and vibrations	Gearbox bearing loads
SS-FNN	WT	2023	[119]	SCADA	Internal loads on bearings
PIGP	WT	2023	[121]	Sparse measurements of the outputs	Phase angles, angular speeds, power
GPLFM	WT	2022	[122]	Vibrations	Dynamic strain
GPLFM	WT	2023	[123]	Vibrations	Dynamic strain
PINN	WT	2022	[125]	Wind speed, turbine control settings	Dynamic response
PINN	WT	2023	[126]	Sparse velocity data	Wind field

7. Conclusions

In conclusion, it can be stated that the use of indirect measurements represents a beneficial method for monitoring assets. This approach simplifies installation, reduces expenses, and minimizes the need for frequent sensor replacements. Moreover, it effectively addresses technical obstacles by allowing data collection in areas that would otherwise be inaccessible, using data from easily reachable locations. Furthermore, it tackles economic challenges by proposing a model that can replace costly sensors in crucial areas, thereby minimizing deployment costs and efforts.

Although there are clear advantages to consider concerning the use of indirect measurements, the task of establishing relationships between measurable quantities and quantities-of-interest is a significant obstacle, and a general method for determining such relationships is currently unavailable. This highlights the need for more research and progress in virtual sensing field, as significant gaps remain that have not yet been addressed. The purpose of this review was to take a step towards closing this gap by examining the current literature in a bid to motivate researchers to develop a comprehensive indirect measurement approach.

Our review highlights the significant potential and challenges of indirect measurement methods in the renewable energy arena. Data-based methods, especially those using deep learning, have shown promise in capturing the dynamic behavior of renewable systems, but nevertheless face challenges related to model obsolescence and the need for improved interpretability. Physics-based methods remain interesting, notwithstanding their complexity and accuracy limitations. Emerging hybrid approaches, which combine data-driven insights with physical principles, offer exciting prospects for overcoming existing limitations. However, the field is still in its early stages, and there is a clear need for more comprehensive methodologies and a broader exploration of predictive models. Future research must focus on developing sophisticated and interpretable models that can navigate the complexities of renewable energy systems, marking a path toward optimizing and sustaining renewable energy resources.

Author Contributions: Conceptualization, A.B., Q.D., R.Z., F.P., M.G. and A.T.; validation, A.B., Q.D., R.Z., F.P., M.G. and A.T.; writing—original draft preparation, A.B.; review and editing, A.B., Q.D., R.Z., F.P., M.G. and A.T.; project administration, A.T., M.G. and F.P. All authors have read and agreed to the published version of the manuscript.

Funding: This work is supported in part by the CRSNG under grant ALLRP 580853-22 and MITACS under grant IT34429.

Acknowledgments: The authors would like to thank the Natural Sciences and Engineering Research Council of Canada (NSERC), MITACS, Hydro-Québec, Power Factors, the industrial partners of the project, and École de Technologie Supérieure (ETS) for the financial support.

Conflicts of Interest: Author Francis Pelletier was employed by the company Power Factors. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest

Abbreviations

The following abbreviations are used in this manuscript:

ACO	Ant Colony Optimization
AE	Autoencoders
AEKF	Augmented Extended Kalman Filter
AKF	Augmented Kalman Filter
AKFF	Augmented Kalman Filter with Fading Memory
ANN	Artificial Neural Networks
AR	Autoregressive
ARDA	Autoregressive Dynamic Adaptive
ARIMA	Autoregressive Integrated Moving Average
ARIMAX	Autoregressive Integrated Moving Average with eXogenous Inputs
ARIX	Autoregressive Integrated with eXogenous Inputs
ARMAX	Autoregressive Moving Average with eXogenous Inputs
ARX	Autoregressive with eXogenous Inputs
CNN	Convolutional Neural Network
DL	Deep Learning
EKF	Extended Kalman Filter
FNN	Feedforward Neural Network
GPLFM	Gaussian Process Latent Force Model
GPR	Gaussian Process Regression
KF	Kalman Filter

LSTM	Long Short-Term Memory
ML	Machine Learning
NAR	Nonlinear Autoregressive
NARNN	Nonlinear Autoregressive Neural Network
NARX	Nonlinear Autoregressive with eXogenous Inputs
NARXNN	Nonlinear Autoregressive with eXogenous Inputs Neural Network
PCA	Principal Component Analysis
PIGP	Physics Informed Gaussian Process
PINN	Physics Informed Neural Network
RNN	Recurrent Neural Network
SARIMA	Seasonal Autoregressive Integrated Moving Average
SBU	Sequential Bayesian Updating
SS	State-Space
SVM	Support Vector Machine
SVR	Support Vector Regression
UKF	Unscented Kalman Filter
VAE	Variational Autoencoders
VAR	Vector Autoregression
VQC	Variational Quantum Circuit

References

- IRENA (International Renewable Energy Agency)—Renewable Energy Statistics 2023. Available online: <https://www.irena.org/Publications/2023/Jul/Renewable-energy-statistics-2023> (accessed on 16 February 2024).
- Qu, C.; Bang, R.N. Metals At The Nexus: Renewable vs. Nuclear Energy Systems, metal Import Requirements, and Energy Security in the European Union. *Miner. Econ.* **2023**, *37*, 101–119. [\[CrossRef\]](#)
- Dao, C.; Kazemtabrizi, B.; Crabtree, C. Wind Turbine Reliability Data Review and Impacts on Levelised Cost of Energy. *Wind Energy* **2019**, *22*, 1848–1871. [\[CrossRef\]](#)
- Hyers, R.W.; MCGowan, J.G.; Sullivan, K.L.; Manwell, J.F.; Syrett, B.C. Condition monitoring and prognosis of utility scale wind turbines. *Energy Mater.* **2006**, *3*, 187–203. [\[CrossRef\]](#)
- Qi, J.; Zhu, R.; Liu, C.; Mauricio, A.; Gryllias, K. Anomaly detection and multi-step estimation based remaining useful life prediction for rolling element bearings. *Mech. Syst. Signal Process.* **2024**, *206*, 110910. [\[CrossRef\]](#)
- Liu, X.T.; Feng, F.Z.; Si, A.W. Condition based monitoring, diagnosis and maintenance on operating equipments of a hydraulic generator unit. *Iop Conf. Ser. Earth Environ. Sci.* **2012**, *15*, 042014. [\[CrossRef\]](#)
- Benabdesselam, A. Etude et Conception d'une Turbine à Air Chaud à Énergie Totale pour une Huilerie Moderne. Master's Thesis, Ecole Nationale Polytechnique, El Harrach, Algiers, 2023.
- Riviere, G. Conception d'un Micro-Cogénérateur aux Granulés de Bois. Ph.D. Thesis, Université de Lorraine, Nancy, France, 2018.
- Vera, D.; Jurado, F.; de Mena, B.; Schories, G. Comparison between externally fired gas turbine and gasifier-gas turbine system for the olive oil industry. *Energy* **2011**, *36*, 6720–6730. [\[CrossRef\]](#)
- Balanescu, D.T.; Homutescu, V.M. Straw Energy saving Solution: Power Plant Based on a Hot Air Turbine. *Procedia Eng.* **2017**, *181*, 698–705. [\[CrossRef\]](#)
- Traverso, A.; Massardo, A.F.; Scarpellini, R. Externally Fired micro-Gas Turbine: Modelling and experimental performance. *Appl. Therm. Eng.* **2006**, *26*, 1935–1941. [\[CrossRef\]](#)
- Presas, A.; Luo, Y.; Wang, Z.; Guo, B. Fatigue life estimation of Francis turbines based on experimental strain measurements: Review of the actual data and future trends. *Renew. Sustain. Energy Rev.* **2019**, *102*, 96–110. [\[CrossRef\]](#)
- Dollon, Q. Analyse Modale Opérationnelle des Roues de Turbines Hydroélectriques par L'étude de Régimes Transitoires. Ph.D. Thesis, Ecole de Technologie Supérieure, Montréal, QC, Canada, 2021.
- Diagne, A.I. Optimisation des Contraintes Mécaniques au Démarrage des Turbines Hydroélectriques à L'aide de Mesures Indirectes. Master's Thesis, Ecole de Technologie supérieure, Montréal, QC, Canada, 2016.
- Pham, Q.H. Développement d'un Outil D'interpolation Inter-Régimes d'une Turbine Hydroélectrique. Ph.D. Thesis, Ecole de Technologie Supérieure, Montréal, QC, Canada, 2022.
- Bouhleb, M.A.; He, S.; Martins, J. Scalable Gradient-enhanced Artificial Neural Networks for Airfoil Shape Design in the Subsonic and Transonic Regimes. *Struct. Multidiscip. Optim.* **2020**, *61*, 1363–1376. [\[CrossRef\]](#)
- Tezzele, M.; Demo, N.; Stabile, G.; Mola, A.; Rozza, G. Enhancing CFD predictions in shape design problems by model and parameter space reduction. *Adv. Model. Simul. Eng. Sci.* **2020**, *7*, 40. [\[CrossRef\]](#)
- Doujak, E.; Maly, A.; Unterluggauer, J.; Haller, F.; Maier, M.; Blasbichler, C.; Stadler, S. Fatigue Strength Analysis of a Prototype Francis Turbine in a Multilevel Lifetime Assessment Procedure Part III: Instrumentation and Prototype Site Measurement. *Energies* **2023**, *16*, 6072. [\[CrossRef\]](#)

19. Ibrahim, R.; Zemouri, R.; Tahan, A.; Lafleur, F.; Kedjar, B.; Merkhouf, A.; Al-Haddad, K. Anomaly Detection for Large Hydrogenerators Using the Variational Autoencoder Based on Vibration Signals. In Proceedings of the 2022 International Conference on Electrical Machines (ICEM), Valencia, Spain, 5–8 September 2022; pp. 1609–1615.
20. Ibrahim, R.; Tahan, A.; Kedjar, B.; Al-Haddad, K.; Zemouri, R.; Merkhouf, A. Vibroacoustic Signal-Based Diagnosis for Rotor Faults in Large Synchronous Machines—An Overview. In Proceedings of the 2023 6th International Conference on Renewable Energy for Developing Countries (REDEC), Zouk Mosbeh, Lebanon, 5–7 July 2023; pp. 72–77.
21. Lu, B.; Li, Y.; Wu, X.; Yang, Z. A Review of Recent Advances in Wind Turbine Condition Monitoring and Fault Diagnosis. In Proceedings of the 2009 IEEE Power Electronics and Machines in Wind Applications, Lincoln, NE, USA, 24–26 June 2009; pp. 1–7.
22. Triki, Y.; Bechouche, A.; Seddiki, H.; Ould Abdeslam, D. Sensorless predictive control of voltage source inverters for renewable energies integration under unbalanced and distorted grid conditions. *Electr. Eng.* **2022**, *104*, 1781–1796. [[CrossRef](#)]
23. Tyagi, S.; Dhingra, A.; Tomar, A. Condition Monitoring & Fault Detection in Photovoltaic Modules Using Machine Learning. In Proceedings of the 2022 1st International Conference on Sustainable Technology for Power and Energy Systems (STPES), Srinagar, India, 4–6 July 2022; pp. 1–6.
24. Nebras, M.; Sobahi, N.M.; Haque, A.; Kurukuru, B.V.S.; Alam, M.M.; Khan, I.A. Data-Driven Approach for Condition Monitoring and Improving Power Output of Photovoltaic Systems. *Comput. Mater. Contin.* **2023**, *74*, 5757–5776.
25. Høiaas, I.; Grujic, K.; Imenes, A.G.; Burud, I.; Olsen, E.; Belbachir, N. Inspection and condition monitoring of large-scale photovoltaic power plants: A review of imaging technologies. *Renew. Sustain. Energy Rev.* **2022**, *161*, 112353. [[CrossRef](#)]
26. Jaen-Cuellar, A.Y.; Elvira-Ortiz, D.A.; Osornio-Rios, R.A.; Antonino-Daviu, J.A. Advances in Fault Condition Monitoring for Solar Photovoltaic and Wind Turbine Energy Generation: A Review. *Energies* **2022**, *15*, 5404. [[CrossRef](#)]
27. Afrasiabi, S.; Allahmoradi, S.; Salimi, M.; Liang, X.; Chung, C.Y. Machine Learning-Based Condition Monitoring of Solar Photovoltaic Systems: A Review. In Proceedings of the 2022 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE), Halifax, NS, Canada, 18–20 September 2022; pp. 49–54.
28. Massaoudi, M.; Chihi, I.; Sidhom, L.; Trabelsi, M.; Refaat, S.S.; Abu-Rub, H.; Oueslati, F.S. An Effective Hybrid NARX-LSTM Model for Point and Interval PV Power Forecasting. *IEEE Access* **2021**, *9*, 36571–36588. [[CrossRef](#)]
29. Chow, S.K.H.; Lee, E.W.M.; Li, D.H.W. Short-term prediction of photovoltaic energy generation by intelligent approach. *Energy Build.* **2012**, *55*, 660–667. [[CrossRef](#)]
30. Salameh, J.; Cauet, S.; Etien, E.; Sakout, A.; Rambault, L. Gearbox condition monitoring in wind turbines: A review. *Mech. Syst. Signal Process.* **2018**, *111*, 251–264. [[CrossRef](#)]
31. Hahn, B.; Durstewitz, M.; Rohrig, K. Reliability of Wind Turbines: Experiences of 15 years with 1500 WTs. In *Wind Energy—Proceedings of the Euromech Colloquium*; Springer: Berlin/Heidelberg, Germany, 2007; pp. 329–332.
32. Wilkinson, M.R.; Spinato, F.; Tavner, P.J. Condition Monitoring of Generators & Other Subassemblies in Wind Turbine Drive Trains. In Proceedings of the 2007 IEEE International Symposium on Diagnostics for Electric Machines, Power Electronics and Drives, Cracow, Poland, 6–8 September 2007; pp. 388–392.
33. Gbashi, S.; Madushele, N.; Olatunji, O.; Adedeii, P.; Jen, T. Wind Turbine Main Bearing: A Mini Review of Its Failure Modes and Condition Monitoring Techniques. In Proceedings of the 2022 IEEE 13th International Conference on Mechanical and Intelligent Manufacturing Technologies (ICMIMT), Cape Town, South Africa, 25–27 May 2022; pp. 127–134.
34. Fan, Z.; Hu, X.; Gao, R.X. Indirect Measurement Methods for Quality and Process Control in Nanomanufacturing. *Nanomanuf. Metrol.* **2022**, *5*, 209–229. [[CrossRef](#)]
35. Rabinovich, S.G. *Evaluating Measurement Accuracy*, 1st ed.; Springer Series in Measurement Science and Technology; Springer: Cham, Switzerland, 2017.
36. Graves, T.; Hamada, M. Making Inferences with Indirect Measurements. *Qual. Eng.* **2005**, *17*, 555–559. [[CrossRef](#)]
37. Moricz, L.; Viharos, Z. J.; Németh, A.; Szépligeti, A. Indirect Measurement and Diagnostics of the Tool Wear for Ceramics Micro-milling Optimisation. *J. Physics Conf. Ser.* **2018**, *10*, 102003. [[CrossRef](#)]
38. De Paiva Silva, G.; Bacci da Silva, M.; De Oliveira, D. Influence of abrasive deburring in indirect tool wear measurement in micromilling of Inconel 718. *J. Braz. Soc. Mech. Sci. Eng.* **2023**, *45*, 262. [[CrossRef](#)]
39. Martin, D.; Kühl, N.; Satzger, G. Virtual Sensors. *Bus. Inf. Syst. Eng.* **2021**, *63*, 315–323. [[CrossRef](#)]
40. Kadlec, P.; Gabrys, B.; Strandt, S. Data-driven Soft Sensors in the process industry. *Comput. Chem. Eng. Comput. Chem. Eng.* **2009**, *33*, 315–323. [[CrossRef](#)]
41. Dollon, Q.; Labbé, P.; Léonard, F. Promoting Explainability in Data-Driven Models for Anomaly Detection: A Step Toward Diagnosis. *Annu. Conf. Phm Soc.* **2023**, *15*. [[CrossRef](#)]
42. Li, Z.; Zhu, Y.; Van Leeuwen, M. A Survey on Explainable Anomaly Detection. *ACM Trans. Knowl. Discov. Data* **2023**, *18*, 23. [[CrossRef](#)]
43. Bellucci, M.; Delestre, N.; Malandrin, N. Une terminologie pour une IA explicable contextualisée. In Proceedings of the EXPLAIN'AI Workshop EGC 2022, Blois, France, 24 January 2022.
44. Blakseth, S.S.; Rasheed, A.; Kvamsdal, T.; San, O. Combining physics-based and data-driven techniques for reliable hybrid analysis and modeling using the corrective source term approach. *Appl. Soft Comput.* **2022**, *128*, 109533. [[CrossRef](#)]
45. TayebiHaghighi, S.; Koo, I. Fault Diagnosis of Rotating Machine Using an Indirect Observer and Machine Learning. In Proceedings of the 2020 International Conference on Information and Communication Technology Convergence (ICTC), Jeju, Republic of Korea, 21–23 October 2020; pp. 277–282.

46. Ljung, L. *System Identification: Theory for the User*, 2nd ed.; Prentice Hall PTR: Upper Saddle River, NJ, USA, 1999.
47. Pasari, S.; Shah, A. Time Series Auto-Regressive Integrated Moving Average Model for Renewable Energy Forecasting. In *Enhancing Future Skills and Entrepreneurship. Sustainable Production, Life Cycle Engineering and Management*; Sangwan, K., Herrmann, C., Eds.; Springer: Cham, Switzerland, 2020.
48. Najwadi, F.; Villarreal, R. Utilizing the Vector Autoregression Model (VAR) for Short-Term Solar Irradiance Forecasting. *Energy Power Eng.* **2023**, *15*, 353–362.
49. Zhang, F.; Li, P.C.; Gao, L.; Liu, Y.Q.; Ren, X.Y. Application of autoregressive dynamic adaptive (ARDA) model in realtime wind power forecasting. *Renew. Energy* **2021**, *169*, 129–143. [[CrossRef](#)]
50. Shukur, O.B.; Lee, M.H. Daily wind speed forecasting through hybrid KF-ANN model based on ARIMA. *Renew. Energy* **2015**, *76*, 637–647. [[CrossRef](#)]
51. Babu, M.K.; Ray, P. Sensitivity analysis, optimal design, cost and energy efficiency study of a hybrid forecast model using HOMER pro. *J. Eng. Res.* **2023**, *11*, 100033. [[CrossRef](#)]
52. Wang, X.; Gao, Y.; Long, X. A Novel Ultra-short-term Photovoltaic Power Generation Forecasting Method Based on Seasonal Autoregressive Integrated Moving Average. *J. Phys. Conf. Ser.* **2023**, *2427*, 012006. [[CrossRef](#)]
53. Guo, W.; Xue, H. Crop Yield Forecasting Using Artificial Neural Networks: A Comparison between Spatial and Temporal Models. *Math. Probl. Eng.* **2014**, *2014*, 857865. [[CrossRef](#)]
54. Presas, A.; Valentin, D.; Zhao, W.; Valero, C.; Egusquiza, M.; Egusquiza, E. Strain Prediction in Francis Runners by Means of Stationary Sensors. In Proceedings of the IOP Conference Series: Earth and Environmental Science, 30th IAHR Symposium on Hydraulic Machinery and Systems (IAHR 2020), Lausanne, Switzerland, 21–26 March 2021; Volume 774, p. 012084.
55. Presas, A.; Valentin, D.; Zhao, W.; Egusquiza, M.; Valero, C.; Egusquiza, E. On the use of neural networks for dynamic stress prediction in Francis turbines by means of stationary sensors. *Renew. Energy* **2021**, *170*, 652–660. [[CrossRef](#)]
56. Azzam, B.; Schelenz, R.; Roscher, B.; Bassir, A.; Jacobs, G. Development of a wind turbine gearbox virtual load sensor using multibody simulation and artificial neural networks. *Forsch Ingenieurwes* **2021**, *85*, 241–250. [[CrossRef](#)]
57. Das, U.K.; Tey, K.S.; Seyedmahmoudian, M.; Mekhilef, S.; Idris, M.Y.I.; Van Deventer, W.; Horan, B.; Stojcevski, A. Forecasting of photovoltaic power generation and model optimization: A review. *Renew. Sustain. Energy Rev.* **2018**, *81*, 912–928. [[CrossRef](#)]
58. Yona, A.; Senjyu, T.; Saber, A.Y.; Funabashi, T.; Sekine, H.; Kim, C.-H. Application of neural network to 24-hour-ahead generating power forecasting for PV system. In Proceedings of the 2008 IEEE Power and Energy Society General Meeting—Conversion and Delivery of Electrical Energy in the 21st Century, Pittsburgh, PA, USA, 20–24 July 2008; pp. 1–6.
59. Mellit, A.; Pavan, A.M. A 24-h forecast of solar irradiance using artificial neural network: Application for performance prediction of a grid-connected PV plant at Trieste, Italy. *Sol. Energy* **2010**, *84*, 807–821. [[CrossRef](#)]
60. Ding, M.; Wang, L.; Bi, R. An ANN-based Approach for Forecasting the Power Output of Photovoltaic System. *Procedia Environ. Sci.* **2011**, *11*, 1308–1315. [[CrossRef](#)]
61. Chen, C.; Duan, S.; Cai, T.; Liu, B. Online 24-h solar power forecasting based on weather type classification using artificial neural network. *Solar Energy* **2011**, *85*, 2856–2870. [[CrossRef](#)]
62. De Giorgi, M.G.; Congedo, P.M.; Malvoni, M. Photovoltaic power forecasting using statistical methods: impact of weather data. *IET Sci. Meas. Technol.* **2014**, *8*, 90–97. [[CrossRef](#)]
63. Ramsami, P.; Oree, V. A hybrid method for forecasting the energy output of photovoltaic systems. *Energy Convers. Manag.* **2015**, *95*, 406–413. [[CrossRef](#)]
64. Liu, J.; Fang, W.; Zhang, X.; Yang, C. An Improved Photovoltaic Power Forecasting Model With the Assistance of Aerosol Index Data. *IEEE Trans. Sustain. Energy* **2015**, *6*, 434–442. [[CrossRef](#)]
65. Martin, L.; Zarzalejo, L.F.; Polo, J.; Navarro, A.; Marchante, R.; Cony, M. Prediction of global solar irradiance based on time series analysis: Application to solar thermal power plants energy production planning. *Solar Energy* **2010**, *84*, 1772–1781. [[CrossRef](#)]
66. Pedro, H.T.C.; Coimbra, C.F.M. Assessment of forecasting techniques for solar power production with no exogenous inputs. *Solar Energy* **2012**, *86*, 2017–2028. [[CrossRef](#)]
67. Abdel-Nasser, M.; Mahmoud, K. Accurate photovoltaic power forecasting models using deep LSTM-RNN. *Neural Comput. Appl.* **2019**, *31*, 2727–2740. [[CrossRef](#)]
68. Dimitrov, N.; Göçmen, T. Virtual Sensors for Wind Turbines with Machine Learning-based Time Series Models. *Wind Energy* **2022**, *25*, 1626–1645. [[CrossRef](#)]
69. Gulgec, N.S.; Takac, M.; Pakzad, S. Structural Sensing with Deep Learning: Strain Estimation from Acceleration Data for Fatigue Assessment. *Comput.-Aided Civ. Infrastruct. Eng.* **2020**, *35*, 277–282. [[CrossRef](#)]
70. Park, G.; You, D.; Oh, K.-Y.; Nam, W. Natural Frequency Degradation Prediction for Offshore Wind Turbine Structures. *Machines* **2022**, *10*, 356. [[CrossRef](#)]
71. Balluff, S.; Bendfeld, J.; Krauter, S. Short term wind and energy prediction for offshore wind farms using neural networks. In Proceedings of the 2015 International Conference on Renewable Energy Research and Applications ICRERA, Palermo, Italy, 22–25 November 2015; pp. 379–382.
72. Han, S.; Qiao, Y.; Yan, J.; Liu, Y.; Li, L.; Wang, Z. Mid-to-long term wind and photovoltaic power generation prediction based on copula function and long short term memory network. *Appl. Energy* **2019**, *239*, 181–191. [[CrossRef](#)]
73. Yu, Y.; Cao, J.; Zhu, J. An LSTM Short-Term Solar Irradiance Forecasting Under Complicated Weather Conditions. *IEEE Access* **2019**, *7*, 145651–145666. [[CrossRef](#)]

74. Yu, Y.; Hu, G.; Liu, C.; Xiong, J.; Wu, Z. Prediction of Solar Irradiance One Hour Ahead Based on Quantum Long Short-Term Memory Network. *IEEE Trans. Quantum Eng.* **2023**, *4*, 3100815. [[CrossRef](#)]
75. Wang, F.; Xuan, Z.; Zhen, Z.; Li, K.; Wang, T.; Shie, M. A day-ahead PV power forecasting method based on LSTM-RNN model and time correlation modification under partial daily pattern prediction framework. *Energy Convers. Manag.* **2020**, *212*, 112766. [[CrossRef](#)]
76. Konstantinou, M.; Peratikou, S.; Charalambides, A.G. Solar Photovoltaic Forecasting of Power Output Using LSTM Networks. *Atmosphere* **2021**, *12*, 124. [[CrossRef](#)]
77. Gao, M.; Li, J.; Hong, F.; Long, D. Short-Term Forecasting of Power Production in a Large-Scale Photovoltaic Plant Based on LSTM. *Appl. Sci.* **2019**, *9*, 3192. [[CrossRef](#)]
78. Bimenyimana, S.; Asemota, G.N.O.; Lingling, L. Output Power Prediction of Photovoltaic Module Using Nonlinear Autoregressive Neural Network. *J. Energy Environ. Chem. Eng.* **2017**, *2*, 32–40.
79. Wang, Y.; Infield, D. Neural network modelling with autoregressive inputs for wind turbine condition monitoring. In Proceedings of the International Conference on Sustainable Power Generation and Supply (SUPERGEN), Hangzhou, China, 8–9 September 2012; pp. 1–6.
80. Garcia, J.L.T.; Calderon, E.C.; Avalos, G.G.; Heras, E.R.; Tshikala, A.M. Forecast of daily output energy of wind turbine using sARIMA and nonlinear autoregressive models. *Adv. Mech. Eng.* **2019**, *11*, 1–15.
81. Gagnon, M.; Vouligny, L.; Cauchon, L.; Giroux, A.M. Virtual Sensors to Generate Turbine Runner Blade Strains from Indirect Measurements. *Iop Conf. Ser. Earth Environ. Sci.* **2022**, *1079*, 012067. [[CrossRef](#)]
82. Zhang, Y.; Qin, C.; Srivastava, A.K.; Jin, C.; Sharma, R.K. Data-Driven Day-Ahead PV Estimation Using Autoencoder-LSTM and Persistence Model. *IEEE Trans. Ind. Appl.* **2020**, *56*, 7185–7192. [[CrossRef](#)]
83. Gensler, A.; Henze, J.; Sick, B.; Raabe, N. Deep Learning for solar power forecasting—An approach using AutoEncoder and LSTM Neural Networks. In Proceedings of the 2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Budapest, Hungary, 9–12 October 2016; pp. 2858–2865.
84. Chen, Y.; Liu, Y.; Han, S.; Qiao, Y. Multi-component condition monitoring method for wind turbine gearbox based on adaptive noise reduction. *Iet Renew. Power Gener.* **2023**, *17*, 2613–2624. [[CrossRef](#)]
85. Huang, C.J.; Kuo, P.H. Multiple-Input Deep Convolutional Neural Network Model for Short-Term Photovoltaic Power Forecasting. *IEEE Access* **2019**, *7*, 74822–74834. [[CrossRef](#)]
86. Li, Y.; Shen, X. A Novel Wind Speed-Sensing Methodology for Wind Turbines Based on Digital Twin Technology. *IEEE Trans. Instrum. Meas.* **2022**, *71*, 2503213. [[CrossRef](#)]
87. Akhter, M.N.; Mekhilef, S.; Mokhlis, H.; Shah, M.N. Review on forecasting of photovoltaic power generation based on machine learning and metaheuristic techniques. *IET Renew. Power Gener.* **2019**, *13*, 1009–1023. [[CrossRef](#)]
88. Ji, G.; Dong, Z.; Qiao, H.; Xu, D. SVR-Based Soft Sensor for Effective Wind Speed of Large-Scale Variable Speed Wind Turbine. In Proceedings of the 2008 Fourth International Conference on Natural Computation, Jinan, China, 18–20 October 2008; Volume 2, pp. 193–196.
89. Yamin, M.; Giyats, A.F. Support Vector Regression Approach for Wind Forecasting. *Int. J. Adv. Sci. Comput. Eng.* **2022**, *4*, 95–101. [[CrossRef](#)]
90. Wang, H.; Zhang, Z.; Zhang, W.; Li, M.; Zhang, Y. Support Vector Regression Inverse System Control for Small Wind Turbine MPPT with Parameters' Robustness Improvement Support. *J. Control. Sci. Eng.* **2022**, *2022*, 2978380. [[CrossRef](#)]
91. Shamshirband, S.; Petkovi, D.; Amini, A.; Anuar, N.B.; Nikoli, V.; Cojbasic, Z.; Kiah, M.L.M.; Gani, A. Support vector regression methodology for wind turbine reaction torque prediction with power-split hydrostatic continuous variable transmission. *Energy* **2014**, *67*, 623–630. [[CrossRef](#)]
92. Shi, J.; Lee, W.J.; Liu, Y.; Yang, Y.; Wang, P. Forecasting Power Output of Photovoltaic Systems Based on Weather Classification and Support Vector Machines. *IEEE Trans. Ind. Appl.* **2012**, *48*, 1064. [[CrossRef](#)]
93. Wolff, B.; Kühnert, J.; Lorenz, E.; Kramer, O.; Heinemann, D. Comparing support vector regression for PV power forecasting to a physical modeling approach using measurement, numerical weather prediction, and cloud motion data. *Solar Energy* **2016**, *135*, 197–208. [[CrossRef](#)]
94. Pan, M.; Li, C.; Gao, R.; Huang, Y.; You, H.; Gu, T.; Qin, F. Photovoltaic power forecasting based on a support vector machine with improved ant colony optimization. *J. Clean. Prod.* **2020**, *277*, 123948. [[CrossRef](#)]
95. Tarpø, M.; Amador, S.; Katsanos, E.; Skog, M.; Gjørdvad, J.; Brincker, R. Data-driven virtual sensing and dynamic strain estimation for fatigue analysis of offshore wind turbine using principal component analysis. *Wind Energy* **2022**, *25*, 505–516. [[CrossRef](#)]
96. Rogers, T.J.; Mclean, J.; Cross, E.J.; Worden, K. Gaussian Processes. In *Machine Learning in Modeling and Simulation. Computational Methods in Engineering & the Sciences*; Rabczuk, T., Bathe, K.J., Eds.; Springer: Cham, Switzerland, 2023.
97. Sanz-Alcaine, J.M.; Sebastián, E.; Sanz-Gorrachategui, I.; Bernal-Ruiz, C.; Bono-Nuez, A.; Pajovic, M.; Orlik, P.V. Online voltage prediction using gaussian process regression for fault-tolerant photovoltaic standalone applications. *Neural Comput. Appl.* **2021**, *33*, 16577–16590. [[CrossRef](#)]
98. Pandit, R.K.; Infield, D.; Kolios, A. Gaussian process power curve models incorporating wind turbine operational variables. *Energy Rep.* **2020**, *6*, 1658–1669. [[CrossRef](#)]
99. Paiva, E.B.R.F.; Nguyen, H.-N.; Lepreux, O.; Bresch-Pietri, D. A Gaussian Process Based Approach to Estimate Wind Speed Using SCADA Measurements from a Wind Turbine. *IFAC-PapersOnLine* **2021**, *54*, 65–71. [[CrossRef](#)]

100. Rajiv, G.; Verma, M.; Subbulakshmi, A. Gaussian process metamodels for floating offshore wind turbine platforms. *Ocean. Eng.* **2023**, *267*, 113206. [[CrossRef](#)]
101. Haghi, R.; Stagg, C.; Crawford, C. Wind Turbine Damage Equivalent Load Assessment Using Gaussian Process Regression Combining Measurement and Synthetic Data. *Energies* **2024**, *17*, 346. [[CrossRef](#)]
102. Kanwal, S.; Khan, B.; Ali, S.M.; Mehmood, C.A.; Rauf, M.Q. Support Vector Machine and Gaussian Process Regression Based Modeling for Photovoltaic Power Prediction. In Proceedings of the 2018 International Conference on Frontiers of Information Technology (FIT), Islamabad, Pakistan, 17–19 December 2018; pp. 117–122.
103. Dolar, A.; Leva, S.; Manzolini, G. Comparison of different physical models for PV power output prediction. *Solar Energy* **2015**, *115*, 83–99. [[CrossRef](#)]
104. Ibrahim, M.; Rassölkin, A.; Vaimann, T.; Kallaste, A.; Zakis, J.; Hyunh, V.K.; Pomarnacki, R. Digital Twin as a Virtual Sensor for Wind Turbine Applications. *Energies* **2023**, *16*, 6246. [[CrossRef](#)]
105. Iliopoulos, A.; Devriendt, C.; Guillaume, P.; Van Hemelrijck, D. Continuous Fatigue Assessment of an Offshore Wind Turbine Using a Limited Number of Vibration Sensors. In Proceedings of the EWSHM—7th European Workshop on Structural Health Monitoring, Nantes, France, 8–11 July 2014.
106. Iliopoulos, A.; Weijtjens, W.; Van Hemelrijck, D.; Devriendt, C. Prediction of Dynamic Strains on a Monopile Offshore Wind Turbine Using Virtual Sensors. *J. Phys. Conf. Ser.* **2015**, *628*, 012108. [[CrossRef](#)]
107. Iliopoulos, A.; Weijtjens, W.; Van Hemelrijck, D.; Devriendt, C. Full-Field Strain Prediction Applied to an Offshore Wind Turbine. *Model Valid. Uncertain. Quantif.* **2016**, *3*, 349–357.
108. Noppe, N.; Iliopoulos, A.; Weijtjens, W.; Devriendt, C. Full Load Estimation of an Offshore Wind Turbine Based on SCADA and Accelerometer Data. *J. Phys. Conf. Ser.* **2016**, *753*, 072025. [[CrossRef](#)]
109. Iliopoulos, A.; Weijtjens, W.; Van Hemelrijck, D.; Devriendt, C. Fatigue Assessment of Offshore Wind Turbines on Monopile Foundations using Multi-band Modal Expansion. *Wind Energy* **2017**, *20*, 1463–1479. [[CrossRef](#)]
110. Nabian, M.S.; Khoshnoudian, F.; Moaveni, B.; Ebrahimi, H. Mechanics-based Model Updating for Identification and Virtual Sensing of an Offshore Wind Turbine Using Sparse Measurements. *Struct. Control. Health Monit.* **2021**, *28*, e2647. [[CrossRef](#)]
111. Teymouri, D.; Sedehi, O.; Katafygiotis, L.S.; Papadimitriou, C. Input-state-parameter-noise identification and virtual sensing in dynamical systems: A Bayesian expectation-maximization (BEM) perspective. *Mech. Syst. Signal Process.* **2023**, *185*, 109758. [[CrossRef](#)]
112. Perišić, N.; Kirkegaard, P.H.; Pedersen, B.J. Cost-effective Shaft Torque Observer for Condition Monitoring of Wind Turbines. *Wind Energy* **2015**, *18*, 1–19. [[CrossRef](#)]
113. Cappelle, C.; Cattebeke, M.; Bosmans, J.; Kirchner, M.; Croes, J.; Desmet, W. Sensor Selection for Cost-effective Virtual Torque Measurements on a Wind Turbine Gearbox. *Forsch Ingenieurwes* **2021**, *85*, 325–334. [[CrossRef](#)]
114. Maes, K.; De Roeck, G.; Lombaert, G.; Iliopoulos, A.; Van Hemelrijck, D.; Devriendt, C.; Guillaume, P. Continuous Strain Prediction for Fatigue Assessment of an Offshore Wind Turbine using Kalman Filtering Techniques. In Proceedings of the 2015 IEEE Workshop on Environmental, Energy, and Structural Monitoring Systems (EESMS) Proceedings, Trento, Italy, 9–10 July 2015; pp. 44–49.
115. Noppe, N.; Tatsis, K.; Chatzi, E.; Devriendt, C.; Weijtjens, W. Fatigue Stress Estimation of Offshore Wind Turbine using a Kalman Filter in Combination with Accelerometers. In Proceedings of the ISMA 2018—International Conference on Noise and Vibration Engineering and USD 2018—International Conference on Uncertainty in Structural Dynamics, Leuven, Belgium, 17–19 September 2018; pp. 4847–4855.
116. Branlard, E.; Jonkman, J.; Dana, S.; Doubrawa, P. A digital twin based on OpenFAST linearizations for real-time load and fatigue estimation of land-based turbines. *J. Phys. Conf. Ser.* **2020**, *1618*, 022030. [[CrossRef](#)]
117. Mehlan, F.C.; Nejad, A.R.; Gao, Z. Estimation of Wind Turbine Gearbox Loads for Online Fatigue Monitoring Using Inverse Methods. In Proceedings of the ASME 2021 40th International Conference on Ocean, Offshore and Arctic Engineering, Online, 21–30 June 2021; Volume 9: Ocean Renewable Energy.
118. Mehlan, F.C.; Nejad, A.R.; Gao, Z. Digital Twin Based Virtual Sensor for Online Fatigue Damage Monitoring in Offshore Wind Turbine Drivetrains. *ASME J. Offshore Mech. Arct. Eng.* **2022**, *144*, 060901. [[CrossRef](#)]
119. Kamel, O.; Kretschmer, M.; Pfeifer, S.; Luhmann, B.; Hauptmann, S. Data-driven Virtual Sensor for Online Loads Estimation of Drivetrain of Wind Turbines. *Forsch Ingenieurwes* **2023**, *87*, 31–38. [[CrossRef](#)]
120. Tuyishimire, B.; McCann, R.; Bute, J. Evaluation of a Kalman predictor approach in forecasting PV solar power generation. In Proceedings of the 2013 4th IEEE International Symposium on Power Electronics for Distributed Generation Systems (PEDG), Rogers, AR, USA, 8–11 July 2013; pp. 1–6.
121. Tartakovsky, A.M.; Ma, T.; Barajas-Solano, D.A.; Tipireddy, R. Physics-informed Gaussian process regression for states estimation and forecasting in power grids. *Int. J. Forecast.* **2023**, *39*, 967–980. [[CrossRef](#)]
122. Bilbao, J.; Lourens, E.-M.; Schulze, A.; Ziegler, L.; Virtual sensing in an onshore wind turbine tower using a Gaussian process latent force model. *Data-Centric Eng.* **2022**, *3*, 35. [[CrossRef](#)]
123. Zou, J.; Cicirello, A.; Iliopoulos, A.; Lourens, E.M. Gaussian Process Latent Force Models for Virtual Sensing in a Monopile-Based Offshore Wind Turbine. In *European Workshop on Structural Health Monitoring, EWSHM 2022, Lecture Notes in Civil Engineering*; Rizzo, P., Milazzo, A., Eds.; Springer: Cham, Switzerland, 2023; Volume 253.

124. Raissi, M.; Perdikaris, P.; Karniadakis, G.E. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J. Comput. Phys.* **2019**, *378*, 686–707. [[CrossRef](#)]
125. Li, X.; Zhang, W. Physics-informed deep learning model in wind turbine response prediction. *Renew. Energy* **2022**, *185*, 932–944. [[CrossRef](#)]
126. Cobelli, P.; Shukla, P.K.; Nasmachnow, S.; Draper, M. Physics informed neural networks for wind field modeling in wind farms. *J. Phys. Conf. Ser.* **2023**, *2505*, 012051. [[CrossRef](#)]
127. Yunus, A.P.; Shirai, N.C.; Morita, K.; Wakabayashi, T. Comparison of RNN-LSTM and Kalman Filter Based Time Series Human Motion Prediction. *J. Phys. Conf. Ser.* **2022**, *2319*, 012034. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.