LSTM Based Power Prediction, Torque Reconstruction, and AGMA Gear Stresses Analysis for Wind Turbine Gearbox

El Mehdi Adib Alaoui[a]

[a]*Michael W. Hall School of Mechanical Engineering, Mississippi State University, 479-1 Hardy Road, Mississippi State, 36762, MS, USA*

Abstract

This project develops a methodology for analyzing wind turbine gearbox loading by combining Supervisory Control and Data Acquisition (SCADA) driven deep learning with standards-based gear stress calculations. A Long Short-Term Memory (LSTM) network is trained on time sequences of SCADA measurements to predict the power output of the wind turbine. The predicted power is then converted into mechanical torque at the gearbox input. Using the nominal planetary and parallel helical gear ratios and stage efficiencies, this input torque is propagated through the drivetrain to obtain torque at each intermediate shaft. From these torques, tangential tooth loads are computed and applied to the finalized gears geometry. Tangential, bending, and contact (pitting) stresses are evaluated using American Gear Manufacturers Association (AGMA) rating equations, including the relevant correction factors for overload, dynamics, size, load distribution, rim thickness, hardness ratio, temperature, and reliability. By linking SCADA based LSTM power prediction, torque reconstruction, and AGMA gear stress analysis, this work provides a physically interpretable assessment of actual tooth loadings in an operating wind turbine gearbox. On the test set, the LSTM reaches a mean absolute error of about 79.5 kW (about 3.8% of rated power), a root-mean-square error of about 126.5 kW (about 6.0% of rated power), and a coefficient of determination $R^2 \approx 0.88$. The reconstructed gearbox torques produce tangential tooth loads up to roughly 3.5 MN on the high-speed pinion (H1) and 0.9 MN on the intermediate pinion of the second helical stage, leading to contact stresses on the order of 1–4 GPa and bending stresses mostly below the 379 MPa (55 ksi) AGMA bending limit; these stress time series remain physically plausible and within allowable limits for all gear stages, providing an interpretable assessment of tooth loadings in an operating wind turbine gearbox.

Keywords: Wind power prediction, Gearbox AGMA analysis, Torque reconstruction, LSTM, Gear stress and safety factors.

1. Introduction

In the field of renewable energies, wind power has emerged as a noticeable alternative, mainly due to its sustainable nature, low carbon footprint, and cost advantages [1]. However, the variable nature of wind power generation makes it difficult for the grid to maintain reliable dispatch and stable operation [2, 3, 4]. Consequently, accurately forecasting wind power decreases grid dispatching costs and enhances system performance [2, 3, 4, 5]. The complexity of forecasting wind power originates from various factors that include climate,

seasonal shifts, and the fluctuating patterns of wind [4]. Simultaneously, this variability transfers into fluctuating mechanical loads on the components, specifically the gearbox, which is one of the most vulnerable components to failure in the wind turbine (WT) [6, 7, 8]. Time-varying wind environment causes two major impacts on the degradation,they not only alter the rate of degradation but also introduce significant fluctuations in the amplitude of the condition monitoring signals [9].

As a consequence, failures in this component are a primary cause of extended downtimes and high maintenance costs, making the analysis of gear stresses necessary [7, 8, 10, 11]. In addition to this, reliability studies emphasize that gearbox gear and bearing failures account for nearly 80% of turbine drivetrain faults, underlining the importance of probabilistic reliability modeling that incorporates uncertain wind loads [7]. The gearbox is also one of the most expensive components; a major replacement costs roughly € 230 000, and about 26% of gearbox failures stem from gear failures [11].

Recent studies have extensively worked on wind turbine output prediction methods, producing a wide range of approaches, and the growth of deep learning (DL) techniques has highlighted the effectiveness of the DL models, increasingly adopting advanced machine learning (ML) methods capable of capturing nonlinear patterns and time-series dependencies in wind data [2, 3]. However, modern rotating machinery operates under complex and varied conditions, making traditional methods based on mechanism analysis difficult to apply, and generally inaccurate for Remaining Useful Life (RUL) prediction [5].

Artificial neural networks (ANNs) are a class of predictive models widely used for prediction [12], recurrent neural networks (RNNs), convolutional neural networks (CNNs), and (LSTM) [4, 5, 13, 14].

LSTMs are foundational models of the gated recurrent network family, a class that also includes variants such as the Gated Recurrent Unit (GRU), both of which use specialized gates to manage information flow over long sequences [15].

Conventional forecasting methods, such as physical models (e.g., numerical weather prediction) and statistical models, such as autoregressive (AR), autoregressive moving average (ARMA), a model that combines AR terms with MA error terms for a stationary time-series, and autoregressive integrated moving average (ARIMA), a traditional linear statistical model for time-series forecasting, rely on explicit signal formulations and manually engineered features, which limit their ability to capture nonlinear and nonstationary behavior in wind power data. As opposed to these approaches, LSTMs introduce memory cells and gating mechanisms that allow them to model short-term and long-term dependencies in operational data [16, 17, 18]. In current research, many studies have relied on (SCADA) systems used for monitoring turbine components and logging operational data, and vibration data to predict time to failure [6, 19, 20, 21]. Unplanned downtime and elevated costs, resulting from poor maintenance planning, can be mitigated by accurately predicting the occurrence and timing of fatigue failure of WT components [6, 10]. Data analytics, feature extraction, and ML techniques show potential in achieving accurate prediction [6, 10, 20]. Within the techniques to predict gearbox failure, there are model-based and data-driven approaches [10, 14, 16], or a combination of the two [19]. The quality of failure prediction largely depends on the complexity and accuracy of the model, which can be further adjusted by data-driven components [6]. Recent works validated the ability of DL models to predict degradation and failure of rotating components, such as gears and bearings, using vibration datasets, while other

ones demonstrated SCADA DL approaches that successfully detect gearbox failures while minimizing false alarms [8, 10, 19].

Deep learning-based methods for failure time prediction usually require the precondition that the training and test data obey the same distribution. However, this precondition is often not met in engineering applications, as equipment operating under variable working conditions can lead to significant differences in the distribution of collected data, which degrades model performance[15].

Large-scale industrial studies using SCADA and vibration datasets across thousands of turbines confirmed that ANNs can predict gearbox failures up to six months in advance, substantially reducing downtime and maintenance costs [6]. Approaches based on load history simulations have also been proposed as cost-effective alternatives to direct condition monitoring, demonstrating that gearbox bearing life prediction can be estimated using simplified load-based models with minimal sensor data [20].

Analytical models of gearbox degradation demonstrate that failure probability grows with cumulative thermal and mechanical stresses, reinforcing the role of physics-informed monitoring in supporting data-driven failure prediction [22].

Wind speed fluctuations cause uncertainty in wind power prediction and, by extension, gearbox time to failure prediction, which complicates grid operation and limits higher penetration of renewable energy [1, 4, 12]. Beyond system-level challenges, this variability translates into fluctuating mechanical loads on turbine components, specifically the gearbox [20]. Therefore, if gearbox failures can be detected or predicted in advance, operators can reduce unplanned maintenance costs and improve turbine availability [5, 8, 23]. However, a significant challenge under these conditions is the risk of "false alarms," where signal fluctuations caused by high speeds can cross a fixed failure threshold prematurely, leading to an underestimation of the true predicted time to failure and unnecessary maintenance costs [9].

ANNs and CNNs are less appropriate for time-series forecasting and nonlinear data, such as wind power [8]. As these approaches are limited by their focus on extracting degradation features from data in a single time frame, neglecting to capture valuable degradation information contained across multiple time frames, which consequently degrades estimation performance [24]. LSTM addresses this limitation through gating mechanisms that regulate information flow, allowing them to model both short-term and long-term dependencies more effectively [10, 13, 25, 26]. Studies consistently show that LSTM models outperform classical methods, such as ARIMA, in time-series forecasting, and have been successfully implemented in wind power prediction as well as machinery life estimation [10, 27].

By accurately predicting power output, it becomes possible to estimate the operational torque applied to the gearbox, which prior studies have used as an input to time-to-failure assessments [28]. Recent works provide evidence in this direction, demonstrating that DL can predict the service life of gears and bearings using vibration data [19], while other studies proposed DL frameworks based on SCADA data that detect gearbox failures with high accuracy and reduce false alarms [25].

However, most of these studies treat wind speed forecasting, power prediction, torque reconstruction, and gearbox stress analysis as separate problems. In contrast, the present work links a multivariate LSTM power–prediction model with a physics–based drivetrain and AGMA gear–rating analysis, using SCADA data to reconstruct gearbox loads and evaluate bending and contact safety factors over the operating duty cycle.

Various approaches have been proposed for forecasting wind power output, ranging from physical and statistical models to advanced DL techniques [5]. To overcome limitations faced by traditional models, researchers have increasingly turned to ML and DL approaches. Empirical results confirmed that LSTM consistently outperformed other architectures in forecasting accuracy, reinforcing its suitability for sequential data [4, 18]. Beyond forecasting, it is emphasized that model-based approaches, such as fatigue crack growth and stochastic degradation models, are valuable for integrating physical knowledge but are limited in scalability and adaptability to operational data [29].

This limitation has further motivated the adoption of data-driven methods [30]. Further explored transfer learning strategies for failure time prediction, demonstrating that feature-transferred networks can adapt knowledge across machinery domains, reducing the dependency on large training datasets. DL frameworks have demonstrated that deep networks can extract features from vibration signals and achieve higher life prediction accuracy than shallow methods [19].

Even within advanced deep learning techniques, a common limitation is that many multiscale models extract features from only a single time frame, failing to capture the crucial correlations of how degradation features evolve over time[24] . The polyvalence of the LSTM framework is demonstrated by its integration with attention mechanisms, which have been shown to improve feature selection under noisy environments and improve robustness in estimation of component lifetime[16].

Extending beyond forecasting power output, a growing body of research has focused on gearbox reliability and failure time prediction. Reliability studies have primarily concentrated on modeling gear and bearing degradation under variable wind loads [3]. Reliability models have been used to capture stochastic stress effects [6]. Building on this, studies have proposed approaches for determining the failure time estimation of WT based on real operating data, providing practical methods for assessing gearbox reliability [20]. In industrial settings where historical failure data is often scarce, "self-data-driven" methods have emerged as a practical strategy; these approaches predict the service life of a machine based on its own operational data, without relying on examples from previously failed units [9]. More recently, SCADA-based data-driven approaches have been applied to identify degradation patterns and estimate the service life prediction [6, 27, 28]. Data-driven methods have been shown to effectively predict gearbox failures and support condition-based maintenance [6]. While supervised learning models have also shown their effectiveness with vibration measurements to assess gearbox bearing life [23]. Building on this, DL frameworks have demonstrated the ability to predict gear and bearing failure time using vibration datasets. More recently, SCADA-based approaches have been validated for accurate detection of gearbox failures, reducing false alarms [8]. A comprehensive review of life estimation approaches for rotating machinery has also emphasized the growing role of hybrid and DL frameworks in addressing variability and uncertainty [5].

Recent advances in prognostics and health management (PHM) have highlighted the importance of accurate service life prediction for rotating machinery [25]. Traditional reliability models for gearboxes often fail to take into consideration real-world variability, as noisy SCADA signals and vibration data form the basis for most life prediction efforts [11, 23].

ML has been applied to address gearbox failures in wind turbines, showing that accurate failure time prediction directly supports condition-based maintenance and reduces costly

downtime[6].To better handle nonstationary and noisy vibration signals, time-frequency analysis is often used as an effective preprocessing step to reveal clearer degradation patterns.[24] Extending this data driven approach, the combination of vibration measurements with supervised learning to assess gearbox bearing life further demonstrates the feasibility of data-driven approaches for turbine drivetrains [23].

DL methods, particularly recurrent architectures, have shown promise in capturing temporal degradation patterns. For example, studies have demonstrated that DL can extract features from vibration signals and predict the degradation and failure of rotating components with higher accuracy than shallow methods [19]. At the same time, review studies emphasize the diversity of life prediction methodologies and provide a comprehensive survey of life estimation approaches for rotating machinery, highlighting the growing role of RNNs and other deep learning frameworks in addressing data variability and uncertainty [5]. Model-based approaches remain valuable for integrating physical knowledge, but their limitations in scalability and adaptability to operational data reinforce the appeal of deep architectures [29].

Despite extensive work on wind turbine power prediction, SCADA-driven condition monitoring, and gear failure modeling, only few studies combine data-driven power forecasting with physics-based torque reconstruction and AGMA stress analysis. As a result, existing approaches do not provide a unified framework that links SCADA measurements, LSTM-based prediction, and standards-based gear stress evaluation for an operating wind turbine gearbox. This gap is particularly important because the resulting tangential loads, bending stresses, and contact (pitting) stresses are fundamental failure-driving variables used in machinery degradation modeling and life estimation.

Taken together, these findings suggest that while DL models, such as LSTMs offer a strong framework for sequence modeling, their value appreciates when they are coupled with physics based models of the drivetrain. Accordingly, this paper combines a multivariate LSTM power prediction model with a physics consistent speed/torque ladder and an AGMA-sized planetary–helical gearbox. SCADA measurements are used both to train the LSTM and to reconstruct gearbox torque histories, from which tangential tooth loads, bending stresses, and contact (pitting) stresses are evaluated for an operating wind turbine gearbox.

2. Methodology

2.1. Multivariate LSTM model

LSTM is a type of RNN designed to learn long and short-range temporal dependencies in sequential data while avoiding the vanishing and exploding gradient problems. LSTM is composed of memory cells that use input, forget, and output gates to regulate information flow, enabling the network to selectively store, update, and retrieve information over long time horizons.

The forget gate determines which information from the cell state shall be discarded and is represented as

$$\mathbf{f}_t = \sigma(\mathbf{W}_f \mathbf{x}_t + \mathbf{U}_f \mathbf{h}_{t-1} + \mathbf{b}_f), \tag{1}$$

where $\mathbf{x}_t \in \mathbb{R}^F$ is the input vector at time t, $\mathbf{h}_{t-1} \in \mathbb{R}^H$ is the previous hidden state, $\sigma(\cdot)$ is the logistic function, $\mathbf{W}_f \in \mathbb{R}^{H \times F}$ is the input to hidden weight matrix for the forget gate, $\mathbf{U}_f \in \mathbb{R}^{H \times H}$ is the recurrent weight matrix for the forget gate, and $\mathbf{b}_f \in \mathbb{R}^H$ is the

corresponding bias vector. Here F denotes the number of input features and H is the number of hidden units in the LSTM layer.

The input gate determines what information will be stored in the cell state, which is represented as

$$\mathbf{i}_t = \sigma(\mathbf{W}_i \mathbf{x}_t + \mathbf{U}_i \mathbf{h}_{t-1} + \mathbf{b}_i), \tag{2}$$

where $\mathbf{W}_i$, $\mathbf{U}_i$, and $\mathbf{b}_i$ are the input-to-hidden, recurrent, and bias parameters for the input gate, with the same dimensions as the corresponding forget-gate matrices and vectors.

The candidate cell update computes the candidate content to be written to the cell state, which is represented as

$$\tilde{\mathbf{c}}_t = \tanh(\mathbf{W}_c \mathbf{x}_t + \mathbf{U}_c \mathbf{h}_{t-1} + \mathbf{b}_c), \tag{3}$$

where $\tanh(\cdot)$ is the hyperbolic tangent, and $\mathbf{W}_c$, $\mathbf{U}_c$, and $\mathbf{b}_c$ are the input, recurrent, and bias parameters for the candidate cell.

The cell state acts as a long-term memory conveyor that allows information to flow largely intact through time, and is represented as

$$\mathbf{c}_t = \mathbf{f}_t \odot \mathbf{c}_{t-1} + \mathbf{i}_t \odot \tilde{\mathbf{c}}_t, \tag{4}$$

where $\mathbf{c}_{t-1} \in \mathbb{R}^H$ is the previous cell state and $\odot$ denotes element-wise multiplication.

The output gate controls what information from the cell state is passed to the output, and is represented as

$$\mathbf{o}_t = \sigma(\mathbf{W}_o \mathbf{x}_t + \mathbf{U}_o \mathbf{h}_{t-1} + \mathbf{b}_o), \tag{5}$$

where $\mathbf{W}_o$, $\mathbf{U}_o$, and $\mathbf{b}_o$ are the input, recurrent, and bias parameters for the output gate.

Finally, the hidden state produces the exposed state at time t, which is represented as

$$\mathbf{h}_t = \mathbf{o}_t \odot \tanh(\mathbf{c}_t). \tag{6}$$

The additive update of $\mathbf{c}_t$ creates a constant error pathway through time, enabling the network to retain or discard information dynamically according to its relevance for future predictions, an ability that distinguishes LSTMs from conventional RNNs. In wind turbine SCADA datasets, the forget and input gates learn to preserve slowly varying operating conditions, such as mean wind speed or rotor speed. This stable memory mechanism allows the LSTM to capture long and short-range temporal dependencies more effectively than a standard RNN. The three gates regulate how information flows into, through, and out of the cell state. The forget and input gates use logistic sigmoid activations, producing values in $(0, 1)$ that act as smooth multiplicative filters on the previous cell state and the candidate update. The candidate update employs a tanh nonlinearity to keep the cell state bounded and to provide expressive nonlinear representations. Finally, the output gate determines how much of the internal cell state is revealed through the hidden state $\mathbf{h}_t$ for subsequent layers.

2.2. Gearbox design

The gearbox is designed as a three-stage planetary–helical transmission that converts the low-speed rotor torque to the high-speed generator shaft while satisfying the required overall speed ratio. In this section, we select the stage architecture, ratio split, and key geometric parameters, and then define the AGMA factors used for bending and contact stress calculations. Based on the operating data, the median generator/rotor speed ratio is ≈ 118.8, so the design target is set to a clean $i_{\text{tot}} \approx 120$. This is realized as a planetary first stage followed by two parallel helical stages as shown in Table 1. A planetary stage at the low–speed input is suitable for high torque because its load sharing across multiple planets decreases the individual tooth loading prior to the parallel stage. The remaining

Table 1: Gearbox stage ratio breakdown

Stage	Ratio
Planetary (P)	6
Helical 1 (H1)	4
Helical 2 (H2)	5
Product	**120**

reduction ratio after the planetary stage was divided into two moderate helical stages with ratios of 4:1 and 5:1 as shown in Table 2. This selection ensures that the pinions can employ practical tooth counts (24 and 26 teeth) while maintaining favorable geometric properties. A normal pressure angle of $\phi_n = 20^\circ$ and a helix angle of $\psi = 25^\circ$ were chosen because they are widely used standard values. The material used is carburized and hardened steel, as shown in Table 3, which presents the material properties.

Table 2: Key gear geometry used in the gearbox design.

Item	Symbol	Value
Normal pressure angle	ϕ_n	20°
Helix angle	ψ	25°
Hand	–	H1: RH, H2: LH (axial thrusts opposed)
Tooth counts (H1)	z_{p1}/z_{g1}	$24/96 \Rightarrow i_{H1} = 4.00$
Tooth counts (H2)	z_{p2}/z_{g2}	$26/130 \Rightarrow i_{H2} = 5.00$

Table 3: AGMA material data used in the gearbox design for steel, carburized and hardened, Grade 1 (adapted from ANSI/AGMA 2001–D04 [31]).

Item	Symbol / Description	Value
Material	Steel, carburized and hardened	–
Grade	AGMA Grade 1 (bending & contact)	–
Modulus of elasticity	$E_p = E_g$	30×10^6 psi
Poisson's ratio	ν	0.30
Bending strength	S_t (repeated load, 10^7 cycles)	55 000 psi
Contact strength	S_c (repeated load, 10^7 cycles)	180 000 psi
Elastic coefficient	C_p (steel–steel mesh)	2 300 $\sqrt{\text{psi}}$

To reduce the AGMA stresses, this gear geometry was defined for stages H1 and H2. The nominal dimensions of this revised design are summarized in Table 4. The AGMA

Table 4: Geometry table for the Gears design.

Quantity	Symbol	H1	H2
Face width	F	300 mm (11.81 in)	300 mm (11.81 in)
Normal module	m_n	26 mm	26 mm
Transverse module	m_t	28.69 mm (1.13 in)	28.69 mm (1.13 in)
Helix angle	ψ	25°	25°
Pinion pitch diameter	d_p	688.5 mm (27.11 in)	745.9 mm (29.37 in)

method converts the nominal load into a working stress using service factors that capture real operating and manufacturing effects, directly influencing S_F, S_H, and the predicted fatigue life. The overload factor K_o of *1.25* accounts for externally imposed load excursions and duty irregularities that are not captured by the nominal transmitted tangential load W_t. In the AGMA contact–strength model, the hardness ratio factor C_H accounts for a harder pinion running against a softer gear (ANSI/AGMA 2001–D04). When the pinion is not harder than the gear, $C_H = 1$. From the specified pinion and gear Brinell hardnesses, the present design uses the constants in Table 5: the pinions are taken with $C_H = 1.00$ in both stages, while the gears use $C_H = 1.06$ (H1) and $C_H = 1.07$ (H2).

Table 5: Hardness ratio factors C_H used for each mesh.

Component	Stage H1	Stage H2
Pinion	1.00	1.00
Gear	1.06	1.07

The reliability factor K_R modifies the allowable AGMA strength to meet a specified reliability R (probability of survival). Higher required reliability yields a larger K_R, i.e., a stricter allowable.

A design reliability of $R = 0.99$ (99% survival) is adopted using the AGMA reliability factor K_R values listed in Table 6.

Table 6: AGMA reliability factor K_R as a function of design reliability R (ANSI/AGMA 2001–D04; Shigley's *Mechanical Engineering Design* [31]).

Reliability R	Reliability factor K_R
0.9999	1.50
0.999	1.25
0.99	1.00
0.90	0.85
0.50	0.70

$$K_R = 0.50 - 0.109\,\ln(1 - R) \quad (0.99 \le R \le 0.9999), \tag{7}$$

The load–distribution factor K_m (face load–distribution factor) increases the working stress to account for non-uniform load across the face width caused by mounting deflection,

tooth lead errors, misalignment, and housing/shaft flexibility.

$$K_m = C_{mf} = 1 + C_{mc}\big(C_{pf}C_{pm} + C_{ma}C_e\big) \tag{8}$$

The gearbox is classified as precision enclosed units. It is a fully enclosed, high-duty wind-turbine drivetrain with controlled mounting and alignment.

The resulting face load–distribution factors K_m for H1 and H2 are summarized in Table 7.

Table 7: Face load–distribution factor K_m for the precision–enclosed gearbox stages (values from AGMA 2001–D04).

Factor	Stage H1	Stage H2
K_m (face load–distribution)	1.177	1.170

The size factor K_s accounts for the size effect in bending fatigue: wider/heavier sections tend to experience higher effective stress, so the working stress is multiplied by K_s.

$$K_s = \frac{1}{k_b} = 1.192\left(\frac{F\sqrt{Y}}{P}\right)^{0.0535} \tag{9}$$

where F is face width (in), P is diametral pitch (1/in), and Y is the Lewis form factor.

The rim–thickness factor K_B is the AGMA modifier that increases bending root stress when the gear rim (backing) is thin. with t_r the minimum rim thickness beneath the root fillet and h_t the whole tooth height. following *Shigley's Mechanical Engineering Design* [31].

$$K_B = \begin{cases} 1.6\,\ln\!\left(\dfrac{2.242}{m_B}\right), & m_B < 1.2, \\ 1, & m_B \geq 1.2\ , \end{cases} \tag{10}$$

The temperature factor K_T is the AGMA temperature factor that modifies the allowable bending and contact strengths to account for elevated operating temperature. For gearboxes whose bulk oil/gear temperature does not exceed 120°C (250°F), the book specifies $K_T = 1.0$; for higher temperatures, $K_T > 1$ is used.

The dynamic factor K_v amplifies the working stress to account for dynamic effects in mesh (velocity error, tooth spacing/lead error, flexibility, runout).

$$K_v = \begin{cases} \left(\dfrac{A+\sqrt{V}}{A}\right)^B, & V \text{ in ft/min}, \\ \left(\dfrac{A+\sqrt{200\,V}}{A}\right)^B, & V \text{ in m/s}, \end{cases} \qquad A = 50{+}56(1{-}B), \qquad B = 0.25\,(12{-}Q_v)^{2/3}. \tag{11}$$

The choice of Q_v came from the accuracy/quality class of the gearing, computed A, B, then evaluated K_v.

Table 8: Geometric quantities for helical stages H1 and H2.

Quantity	H1	H2
Z (mm)	28.097248	28.561666
m_N	1.1049	1.0870
I	0.0995988424	0.1020618184

The geometry factors for contact (pitting) are represented by (AGMA) pitting-resistance geometry factor I, which accounts for the local Hertzian contact geometry of a meshing pair. For the helical stages defined in the baseline, the derived quantities are shown in Table 8.

The bending strength geometry factor J replaces the classical Lewis form factor in AGMA bending fatigue; it embeds the root–fillet stress concentration and the load sharing on the line of action.

$$J = \frac{Y}{K_f \, m_N}, \tag{12}$$

where K_f is the fillet (root) stress–concentration factor and m_N is the face–contact (load–sharing) ratio term.

The resulting bending geometry factors J for the helical pinions, after applying the modifier to J', are summarized in Table 9.

Table 9: Final bending geometry factor $J = (\text{modifier}) \times J'$. Values shown to 4 s.f.; H2 rounds to $J \approx 0.570$.

Stage	Gear ratio i	Mating gear teeth N_G	$m_G = N_G/N_P$	Modifier	J'	J
H1	4	96	4.00	0.98	0.5408	0.5300
H2	5	130	5.00	1.0053	0.5670	0.5700

All AGMA modifying factors used in the stress and safety–factor calculations for the helical stages are collected in Table 10.

Table 10: Collected AGMA factors and selections used in the stress and safety calculations.

Symbol	Value	Source
K_o	1.25	Service class (overload).
Q_v	10	Quality level (precision enclosed).
K_v (H1/H2)	1.092843 / 1.202365	Computed from $Q_v = 10$ with stage pitch line speeds.
K_s (H1/H2)	1.150659 / 1.151928	Size factor (AGMA guidance; pinions).
K_m (H1/H2)	1.176904 / 1.169501	Face load distribution (straddle mount, crowned).
K_B	1.00	Rim thickness factor (no thin–rim penalty).
K_t	1.00	Temperature factor (120 °C).
C_f	1.00	Surface finish factor.
K_r (Y_z)	0.85	Reliability factor (90% target).
C_p (U.S.)	2300.243837 $\sqrt{\text{psi}}$	Elastic coefficient (steel/steel).
Z_E (SI)	191 $\sqrt{\text{MPa}}$	Elastic coefficient (SI form).
C_H (pinion H1/H2)	1.00 / 1.00	Hardness ratio factor.
C_H (gear H1/H2)	1.06 / 1.07	Hardness ratio factor.
J (pinion H1/H2)	0.53 / 0.57	Helical J via virtual tooth number N_v.
I (H1/H2)	0.099599 / 0.102062	With transverse $\phi_t \approx 21.88°$, $\psi = 25°$.
Y_N, Z_N	per duty cycles	Life factors to be applied with the duty spectrum.

The AGMA bending stress is defined below for both U.S. customary and SI unit systems:

$$\sigma = \begin{cases} W_t K_o K_v K_s \dfrac{P_t}{F} \dfrac{K_m K_B}{J}, & \text{U.S. customary units,} \\ W_t K_o K_v K_s \dfrac{1}{b\, m_t} \dfrac{K_m K_B}{J}, & \text{SI units.} \end{cases} \tag{13}$$

The corresponding AGMA contact stress expressions are:

$$\sigma_c = \begin{cases} C_P \sqrt{W_t K_o K_v K_s \dfrac{K_m}{d_p F} \dfrac{C_f}{I}}, & \text{U.S. customary units,} \\ Z_E \sqrt{W_t K_o K_v K_s \dfrac{K_H}{d_p b} \dfrac{Z_R}{Z_I}}, & \text{SI units.} \end{cases} \tag{14}$$

2.3. Data Exploration

The turbine's supervisory system records 10-minute sequences for electrical output and drivetrain process, which correspond to 26064 time-ordered samples. Values in a row are synchronous at the timestamp and represent averaged behaviour rather than instantaneous samples, which compensates for sensor noise but also smooths fast transients. The SCADA extract contains 26 064 ten-minute records (4 344 h, 181 days), after removing rows with missing predictor or target values, 23 821 records remain, corresponding to about 3 970 hours ($\approx$ 165 days) of data used for wind power prediction as shown in Table 11.

Table 11: Raw SCADA data with training features for power forecasting of a wind turbine (Kelmarsh Turbine, 2021).

Date & time	Wind speed (m/s)	Energy theor. (kWh)	Rotor speed (RPM)	Wind dir. (°)	Yaw angle (°)	Power (kW)
1/1/2021 00:00	6.2047	84.5564	11.7642	304.8331	313.3654	-1.3381
1/1/2021 00:10	5.8598	71.0691	12.4192	298.8192	313.3654	-2.5933
1/1/2021 00:20	7.5828	94.7227	9.8220	305.2070	313.3654	-1.3194
1/1/2021 00:30	7.5198	96.5188	7.5376	297.4892	313.3654	-2.1187
1/1/2021 00:40	8.5198	93.4287	8.9572	287.7796	312.8034	-2.8043
1/1/2021 00:50	8.2528	89.2892	9.4874	289.5720	312.8034	-2.6143
1/1/2021 01:00	8.2058	92.5203	8.9460	289.5720	312.8034	-1.9694
			... ...			
12/31/2021 23:10	6.5872	82.9531	12.4460	301.5734	312.9420	1482.3756
12/31/2021 23:20	8.5672	94.2365	12.1184	303.2874	312.9420	1334.9642
12/31/2021 23:30	8.5672	94.2365	12.1184	303.2874	312.9420	1334.9642
12/31/2021 23:40	8.5672	94.2365	12.1184	303.2874	312.9420	1334.9642

Power (kW) is the supervised target for the power model and an intermediate quantity used to compute torque; it reflects both aerodynamic input and drivetrain conversion. The gearbox speed (RPM) encodes the operating regime and is required to convert power to torque. The transmitted torque is represented as

$$\tau_t = \frac{9549.3\,P}{n}, \tag{15}$$

where P is the mechanical power in kilowatts and n is the shaft speed in revolutions per minute. Drivetrain acceleration captures dynamic loading and transient events that often precede thermal and mechanical changes. Gear oil temperature (°C) tracks sustained loading and friction with a thermal lag, providing a slower but robust indicator of operating stress. Gear-oil pump pressure (bar) reflects lubrication flow and back-pressure; deviations may indicate wear, viscosity changes, or pump malfunctions. Torque (Nm) is used directly when measured; otherwise it is reconstructed from power and RPM using (15) and serves as the mechanical input to the AGMA gear-rating analysis, where it is converted to tangential tooth loads and, ultimately, to pinion bending and contact stresses at each gearbox stage.

A concise summary of the processed SCADA dataset and the LSTM input–output configuration used in this study is given in Table 12.

Table 12: Summary of the SCADA dataset and LSTM input-window configuration

Item	Value	Notes
Rows (samples)	26064	26015
Sampling period	10 min	fixed cadence
Coverage	4,500 h	$\approx$ 187.5 days
Features used	6	listed above
LSTM window L	96	$\approx$ 16 h
Overlapping windows	25920	$N - L + 1$

2.4. Performance metrics

Generated Power forecasting is trained with Mean Absolute Error (MAE) loss and Adam. The evaluation metrics computed after training are the Root Mean Square Error (RMSE), the (MAE), and the coefficient of determination R^2.

$$\mathrm{RMSE}(\hat{y}, y) = \sqrt{\frac{1}{N}\sum_{n=1}^{N}|\hat{y}_n - y_n|^2}, \tag{16}$$

where N is the number of samples, y_n is the true (measured) target at index n, $\hat{y}_n$ is the model prediction, and $|\cdot|$ denotes absolute value.

$$\mathrm{MAE}(\hat{y}, y) = \frac{1}{N}\sum_{n=1}^{N}|\hat{y}_n - y_n|, \tag{17}$$

$$R^2(\hat{y}, y) = 1 - \frac{\sum_{n=1}^{N}(y_n - \hat{y}_n)^2}{\sum_{n=1}^{N}(y_n - \bar{y})^2}, \tag{18}$$

where $\bar{y} = \frac{1}{N}\sum_{n=1}^{N} y_n$ is the sample mean of the targets.

2.5. Hyperparameters for model training and forecast horizon

The LSTM model is configured for short-horizon power prediction using a small number of manual trials on the train/validation split, selecting the configuration that produced the lowest validation error and smooth learning curves. The tests varied the number of stacked LSTM layers, the number of hidden units per layer, the dropout rate, and the batch size. The lookback window length was fixed to $L = 10$ time steps (approximately 100 minutes of past data at a 10-minute sampling time), with a stride of 1 and a one-step-ahead forecast horizon, i.e., the power at time t is predicted from the previous ten samples. The dataset was divided chronologically into 70% of the samples for training, 15% for validation, and 15% for testing, with no shuffling in order to preserve the temporal dependencies between consecutive measurements. All input features and the target were first scaled to the interval [0,1] using a MinMaxScaler fitted on the full dataset before the split. The network architecture consists of three stacked LSTM layers with a progressively decreasing number of neurons, with 256 units in the first layer, 128 units in the second layer, and 64 units in the third layer, each followed by a dropout layer with a rate of 0.3 to reduce overfitting, and a final dense layer with a linear activation function to output continuous power values. The sequence length was set to 10 time steps, corresponding to approximately 100 minutes of history at a 10-minute sampling interval, and the model performs a one-step-ahead forecast of the next power value. Model optimization was performed using the Adam optimizer with a learning rate of 0.001, and the loss function during training was the Mean Squared Error (MSE). The batch size was set to 16, and training was allowed to continue for a maximum of 50 epochs as shown in Table 13, while a ModelCheckpoint callback stored the set of weights that achieved the lowest validation loss.

Table 13: Hyperparameters used for the LSTM power prediction model in this work.

Item	Hyperparameter
Train/Val/Test split	70% / 15% / 15%
Scaling	MinMaxScaler
LSTM layers	3
Units per LSTM layer	256, 128, 64
Neurons (total)	256, 128, 64
Dropout	0.3
Sequence length L	10 time steps
Forecast horizon	1 step ahead (10 min)
Stride	1
Batch size	16
Optimizer	Adam
Learning rate	0.001
Loss	Mean Squared Error (MSE)
Epochs	50

The chosen hyperparameters represent a compromise between model capacity and robustness, given the limited size of the SCADA dataset. The stacked LSTM architecture captures nonlinear temporal dependencies in wind speed, direction, yaw angle, and rotor speed, while a relatively short lookback window focuses the model on the most informative recent history. Regularization (dropout and early stopping) mitigates overfitting, and the selected batch size and optimization settings provide stable gradient updates and reasonable training time

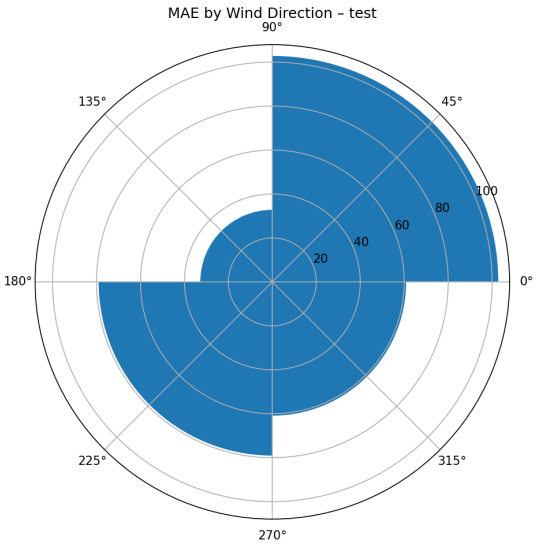
3. Results and discussion

3.1. LSTM power prediction

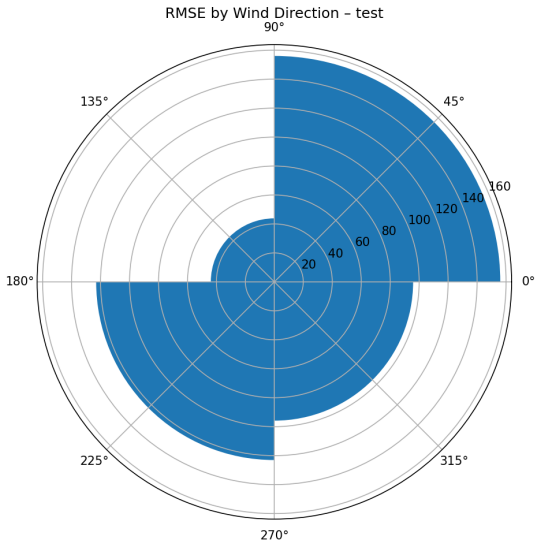
Table 14 shows that train and validation metrics are very close, indicating that the model experiences no obvious overfitting or data leakage. The test set coefficient of determination is $R^2 \approx 0.88$, which, although lower than the train and validation values, is still strong and promising for noisy SCADA win-power data. Additionally, the MAE and RMSE values are of the same order of magnitude across all three splits, confirming that the model performance is stable and consistent.

Table 14: Performance metrics of the LSTM model for power prediction.

Dataset	**MAE (kW)**	**RMSE (kW)**	$\boldsymbol{R^2}$
Train Power	105.358	160.951	0.945
Validation Power	101.197	156.577	0.950
Test Power	79.461	126.544	0.878



(a) Mean Absolute Error (MAE) of test-set power predictions as a function of wind direction.



(b) Root Mean Squared Error (RMSE) of test-set power predictions as a function of wind direction.

Figure 1: Test-set power prediction errors as a function of wind direction.

Figures 1a and 1b summarize the model accuracy as a function of wind direction on the test set. The MAE polar plot shows that the largest errors occur for inflow directions around the north–east and south–west sectors, where the mean absolute error reaches values close to 100 kW, while the north–west sector exhibits the lowest MAE, below about 40 kW. A similar pattern appears in the RMSE plot, with peak values of roughly 160 kW for the most challenging sectors and lower RMSE (around 40–80 kW) elsewhere. All directional errors remain within approximately 5–8% of the 2.1 MW rated power, indicating that the LSTM model maintains reasonable accuracy across the full range of wind directions. The higher errors in specific sectors suggest that certain inflow angles are more difficult to predict, possibly due to site-specific wake effects, local terrain, or yaw misalignment, and they highlight directions where additional model refinement or physical correction could further improve performance.

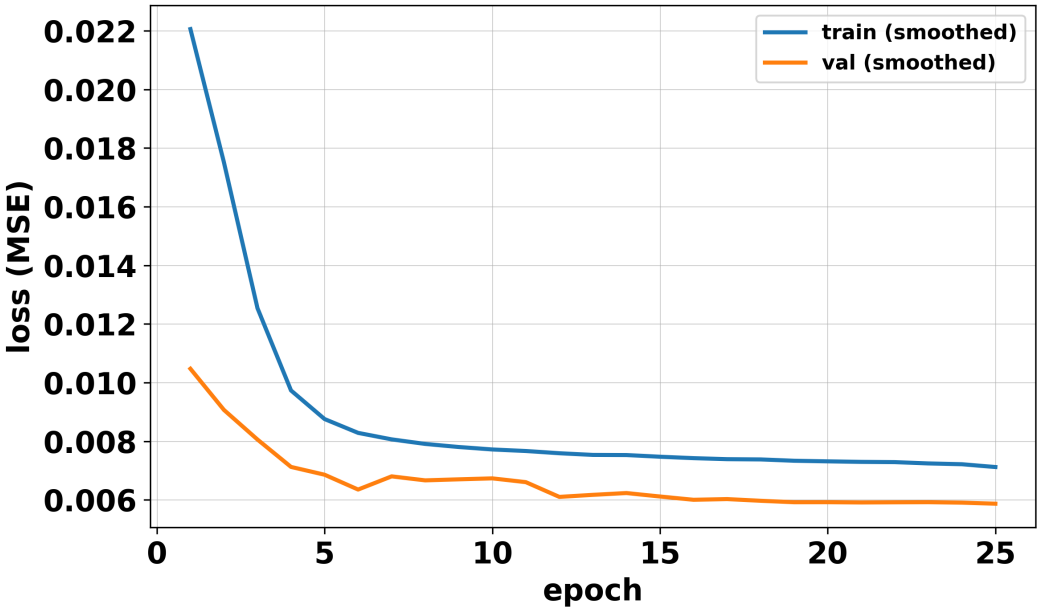


Figure 2: Training and validation loss compared epoch for the LSTM power prediction model.

Figure 2 shows the smoothed training and validation loss, measured as the mean squared error (MSE) on the MinMax normalized target power (i.e., the loss values are dimensionless), as a function of epoch for the LSTM power prediction model. Both curves exhibit a rapid decrease during the first few epochs, with the training loss dropping from approximately 2.2×10^{-2} to below 8×10^{-3} and the validation loss decreasing from about 1.0×10^{-2} to around 6×10^{-3}. After roughly 10–15 epochs, the losses gradually level off and no further significant improvement is observed, indicating that the network has effectively converged. The validation loss remains slightly lower than the training loss over almost the entire training process, which is consistent with the use of dropout (active only during training) and with the moving-average smoothing of the curves, and does not indicate data leakage. Importantly, there is no late-epoch increase or divergence of the validation curve, so there is no visible sign of overfitting. The small and stable gap between the two curves suggests that the chosen architecture and regularisation provide a good trade-off between bias and variance for this dataset, and that training for about 25–30 epochs is sufficient to obtain a well-trained model.

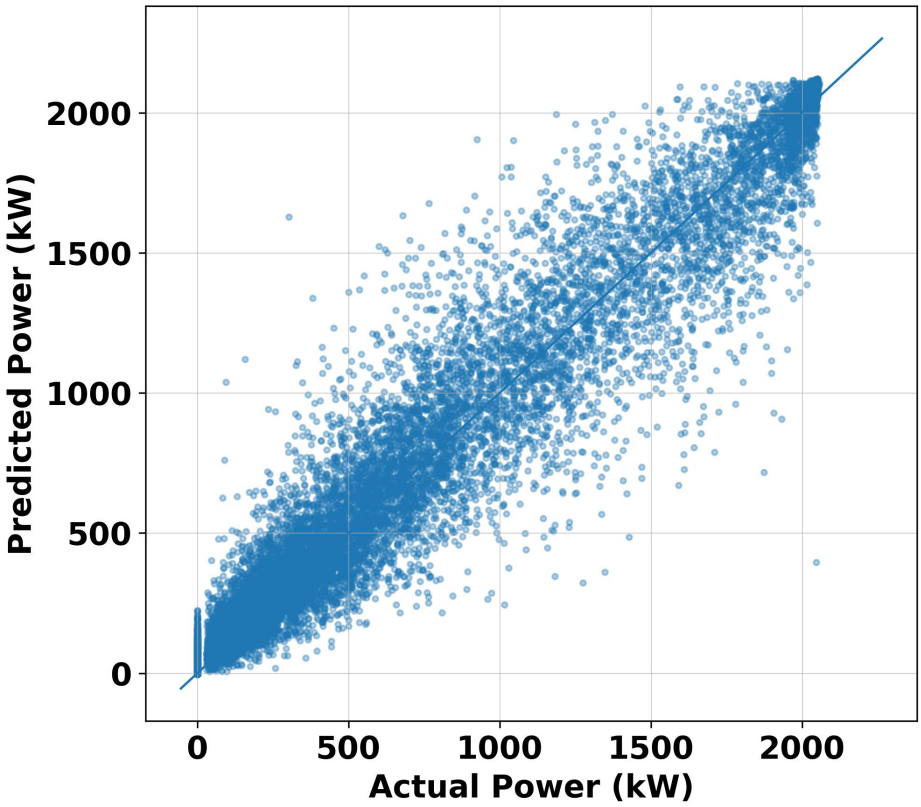


Figure 3: Predicted Power compared to Actual Power in kW(TEST)

Figure 3 shows that most points are concentrated around the 45° diagonal, indicating that the LSTM captures the main non linear relationship between wind conditions and power output with no obvious systematic bias.

The test coefficient of determination is $R^2 \approx 0.945$, while the mean absolute error and root-mean-square error are approximately 79.5 kW ($\approx 3.8\%$ of the 2.1 MW rated power) and 126.5 kW ($\approx 6.0\%$), respectively. The slightly increased spread at higher power levels reflects the reduced number of high-wind samples and the inherently higher variability of turbine operation near rated conditions, but the overall alignment with the diagonal confirms that the model provides a reliable short-horizon power prediction.

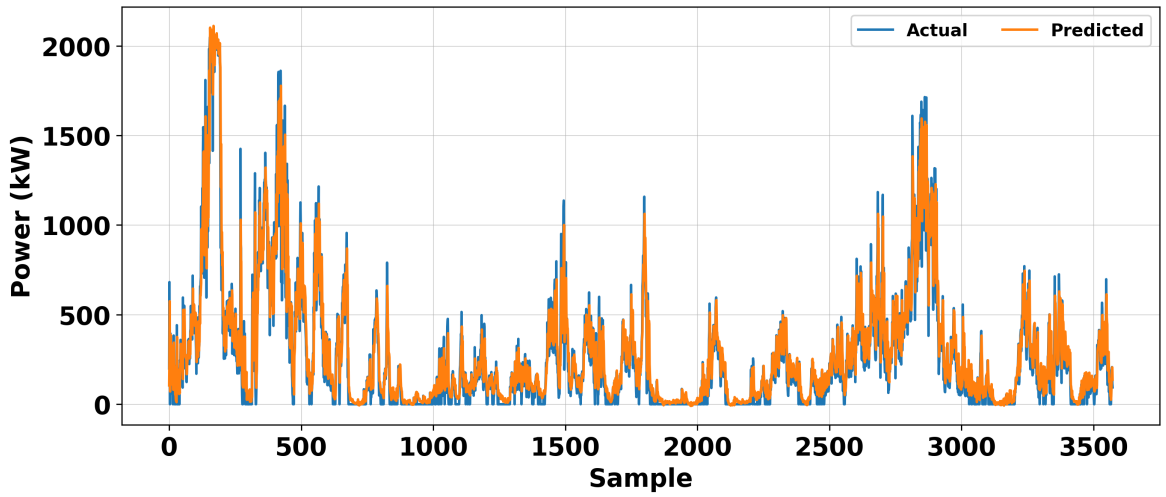


Figure 4: Actual and predicted power time series on the test set for the LSTM model.

Figure 4 compares the actual and predicted power time series on the test set. The two curves almost overlap, including during high-power episodes close to the 2.1 MW rated power, indicating that the LSTM is able to track both low-power periods and sharp ramps with only small phase and amplitude errors. The remaining discrepancies are mainly short under or over-predictions around fast transients, which is expected given the 10-minute SCADA sampling and the stochastic nature of wind inflow. With a test MAE of about 79.5 kW ($\approx 3.8\%$ of rated) and an RMSE of roughly 126.5 kW ($\approx 6.0\%$ of rated), the magnitude of the errors is physically plausible and consistent with typical uncertainty levels reported for short-horizon turbine power forecasting.

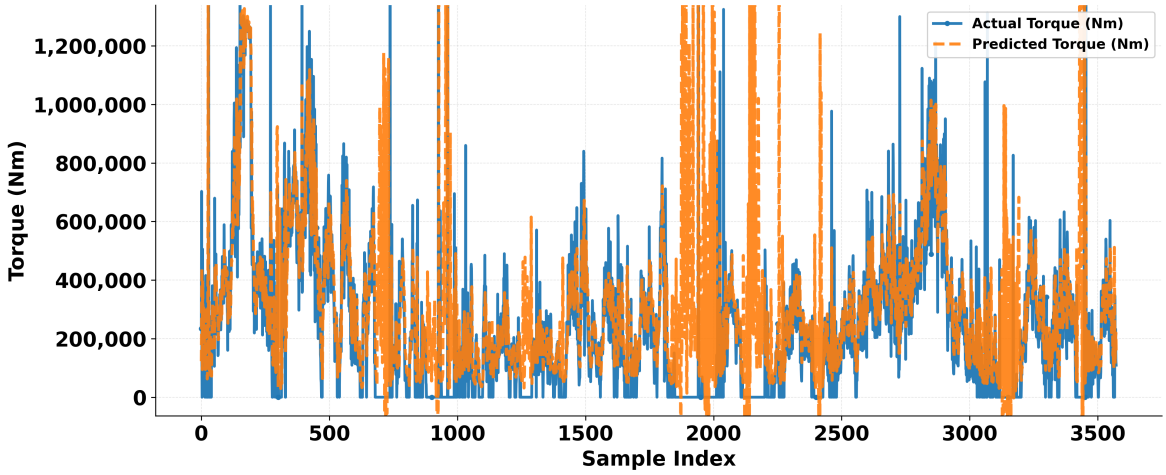


Figure 5: Actual and predicted drivetrain torque on the test set (full time series).

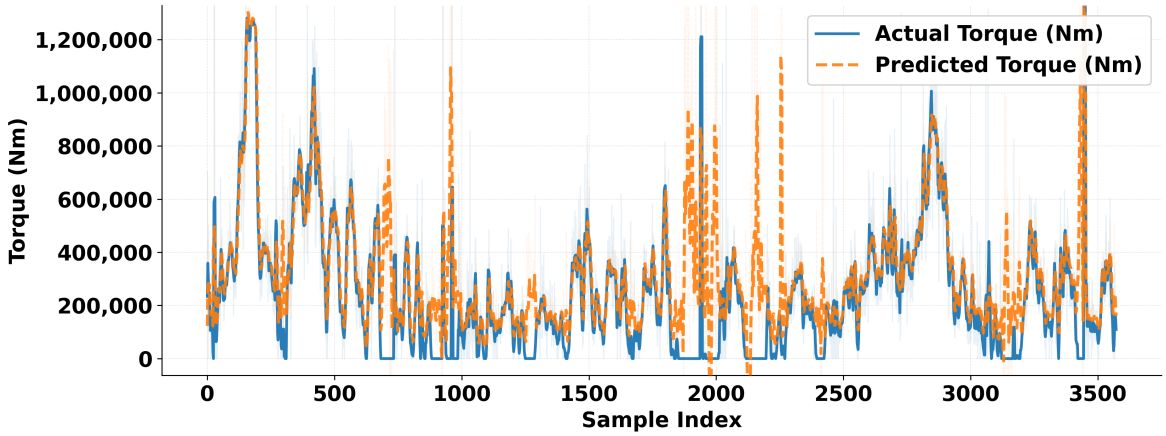


Figure 6: Smoothed and downsampled actual and predicted drivetrain torque on the test set, with percentile-based clipping for readability.

Figures 5 and 6 show time series of actual and predicted drivetrain torque on the test set. After moving–average filtering, the envelopes of the two curves almost coincide, indicating that the LSTM–based power prediction combined with the P–to–torque conversion reproduces both the magnitude and temporal pattern of the torque loads. The model tracks the main peaks and valleys of the torque history, with slight underestimation of the most extreme spikes but good agreement in the duration and frequency of high–load events. The torque range remains on the order of 10^6 Nm, which is physically plausible and realistic for a 2.1 MW turbine operating at the measured rotor speeds, confirming that the predicted torque signals are consistent with the underlying SCADA physics rather than numerical artefacts.

3.2. Gearbox Design and AGMA Specifications

The reconstructed torque histories from Section 3.1 are then used as inputs to the gearbox rating analysis. For the specified three stage planetary-helical gearbox geometry and materials, the AGMA procedure is followed to compute the tangential tooth load W_t, the bending (root) stress σ_b, and the contact (pitting) stress σ_c for each helical pinion. The stresses are evaluated using the standard AGMA bending and surface durability forms, including all prescribed modifying factors (overload, dynamic, size, load distribution, rim-thickness, temperature, and reliability) together with the geometry factors J for bending and I for contact. Allowable bending and contact strengths are obtained by applying the AGMA stress cycle (life) factors Y_N and Z_N along with corrections of temperature and reliability.The time varying safety factors $S_F(t)$ and $S_H(t)$ are evaluated over the reconstructed duty cycle to monitor how close the mesh operates to the allowable bending and contact strengths.

Figures 7 and 8 show the resulting tangential load time series $W_t(t)$ for the Helical Stage 1 (H1) and Stage 2 (H2) pinions on the test set, plotted at the 10-minute SCADA resolution along with a 6-point moving average. The H1 loads are consistently higher than H2, as expected from the drivetrain ratios and efficiencies. The moving average tracks the operating envelope, while short-duration transients appear as spikes in the raw series. The 99.5th-percentile clipping used in subsequent plots is purely a visualization choice and is not applied in any stress or life calculation.

Figures 9-12 present the corresponding bending and contact stress histories for the two helical stages. Because both σ_b and σ_c are driven by W_t, their temporal patterns closely follow the load curves. Bending stress is approximately proportional to W_t, whereas contact stress varies roughly with $\sqrt{W_t}$, so peaks in σ_c are visibly compressed relative to σ_b. Differences in absolute level between H1 and H2 arise from stage-specific geometry and factors such as pitch diameter d, face width F, and the AGMA factors J, I, K_v, and K_m. For H1, the 6-point average of σ_b lies mainly near the reference level for the pinion with temporary rises during intervals of high load, while H2 remains lower, consistent with its larger J, larger pitch diameter, and lower stage load. Contact stress magnitudes exceed bending stress (Hertzian pressure) but remain mostly below the chosen percentile cap.

The resulting bending and contact safety-factor histories for the two helical stages are shown in Figures 13–16. These curves combine the time-varying stress demand with the allowable strengths obtained from the upsized design. Three consistency checks are satisfied.

First, stage scaling is correct: H1 levels exceed H2 for W_t, σ_b, and σ_c, matching the drivetrain ratios and geometry.
Second, shape coherence holds: $W_t(t)$, $\sigma_b(t)$, and $\sigma_c(t)$ change together over time, with $\sigma_c(t)$ appearing as a compressed version of the $W_t(t)$ pattern and $\sigma_b(t)$ nearly proportional.

Third, factor influence behaves as expected: with time-invariant geometry and AGMA factors, only the absolute levels differ between stages, not the temporal pattern.

Pinions are the limiting elements because they have fewer teeth (lower J and thus higher σ_b for the same W_t), accumulate more load cycles for life factors,and in the baseline configuration the mating gears benefit from hardness ratio effects. Verifying the pinions is therefore conservative. The same computational pipeline could be applied to gears by substituting the gear geometry (J, diameters, and life factors); the $\sigma_c(t)$ timeline is unchanged, but it would be compared against the gear allowables instead. For the upsized pinions considered here, the time varying bending and contact safety factors $S_F(t)$ and $S_H(t)$ remain greater than unity over all observed operating points, supporting the conclusion that the gearbox design is sufficient for the intended service life under the reconstructed torque spectrum.

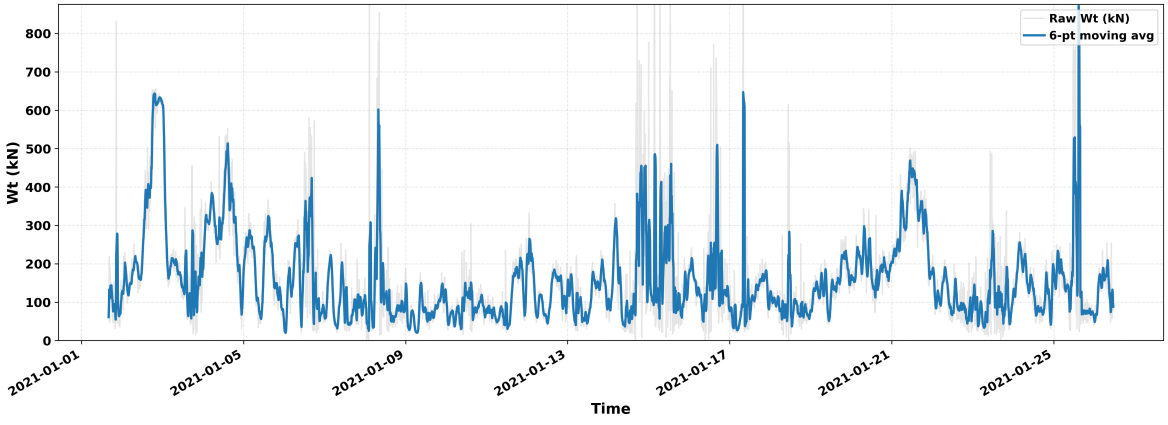


Figure 7: Tangential load W_t time series for Helical Stage 1 (H1) pinion on the test set, showing raw 10-minute values and the 6-point moving average.

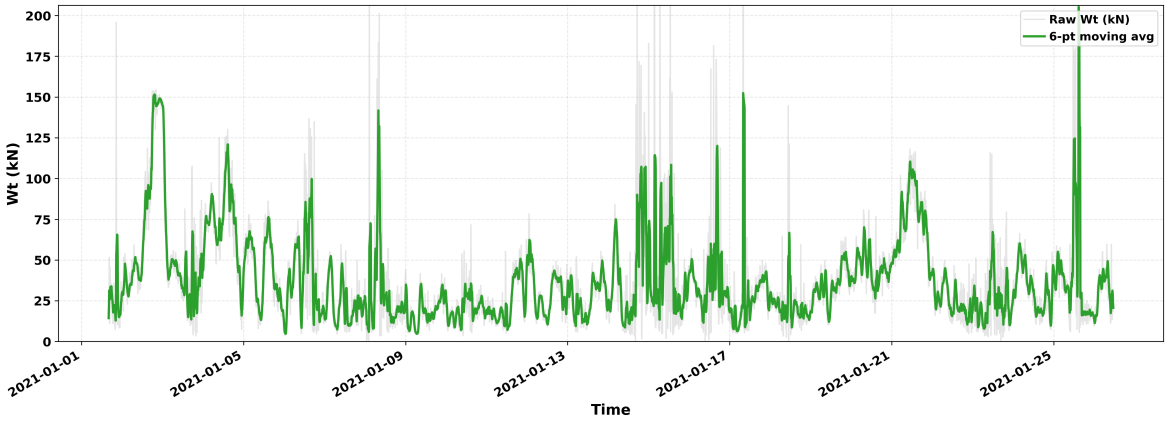


Figure 8: Tangential load W_t time series for Helical Stage 2 (H2) pinion on the test set, showing raw 10-minute values and the 6-point moving average.

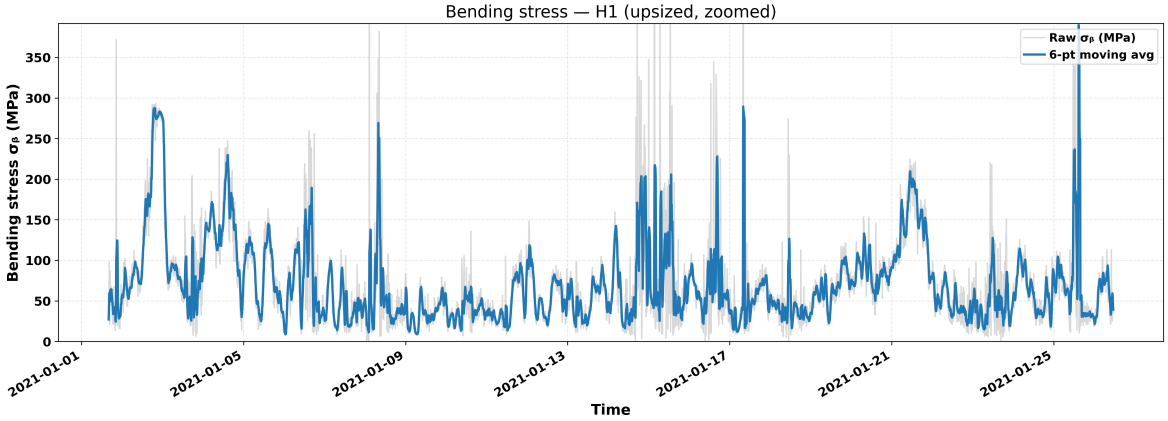


Figure 9: Bending stress time series for Helical Stage 1 (H1) pinion on the test set, showing raw values, 6-point moving average, and the 99.5th percentile cap.

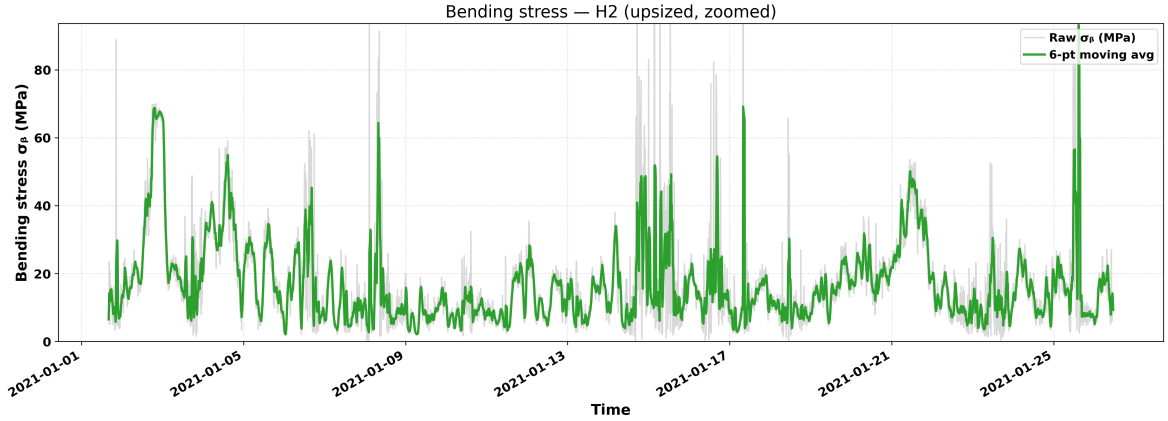


Figure 10: Bending stress time series for Helical Stage 2 (H2) pinion on the test set, showing raw values, 6-point moving average, and the 99.5th percentile cap.

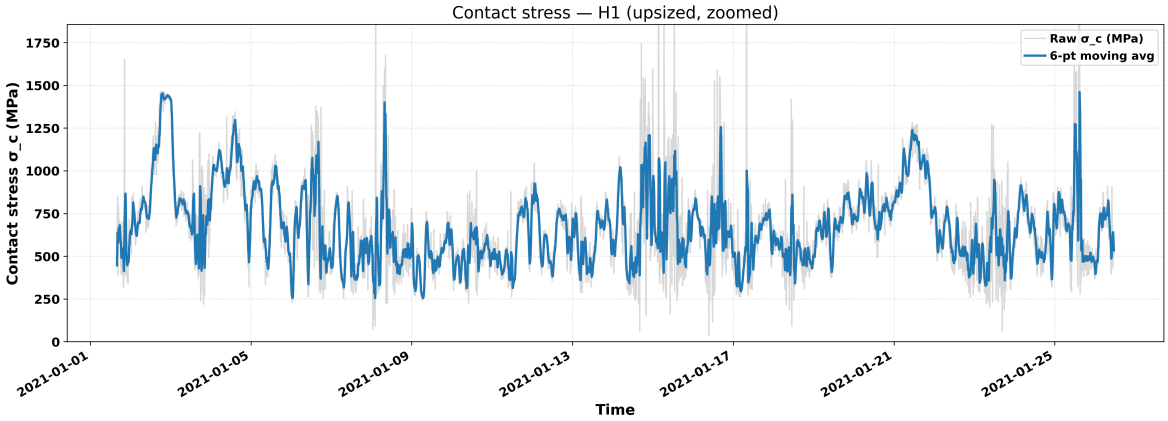


Figure 11: Contact stress σ_c time series for Helical Stage 1 (H1) pinion on the test set, showing raw values capped at the 99.5th percentile and the 6-point moving average.

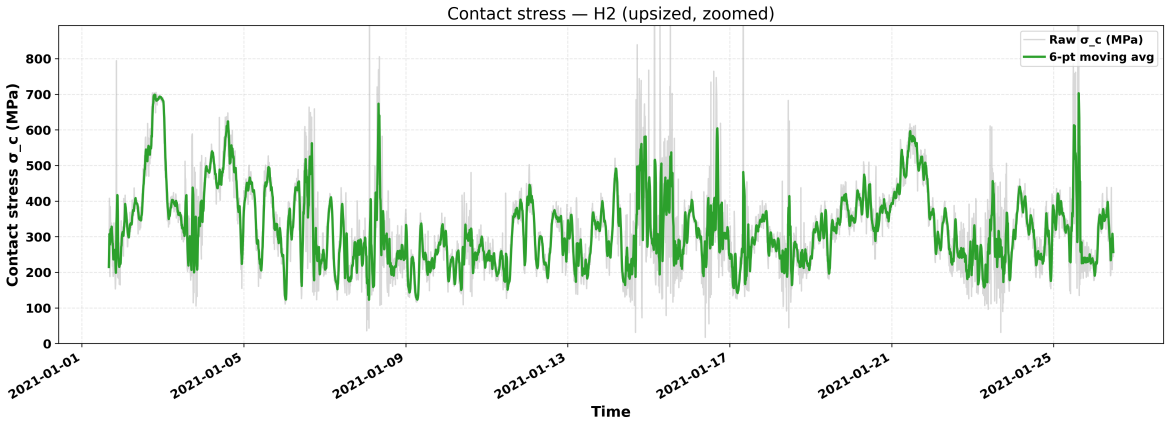


Figure 12: Contact stress σ_c time series for Helical Stage 2 (H2) pinion on the test set, showing raw values capped at the 99.5th percentile and the 6-point moving average.

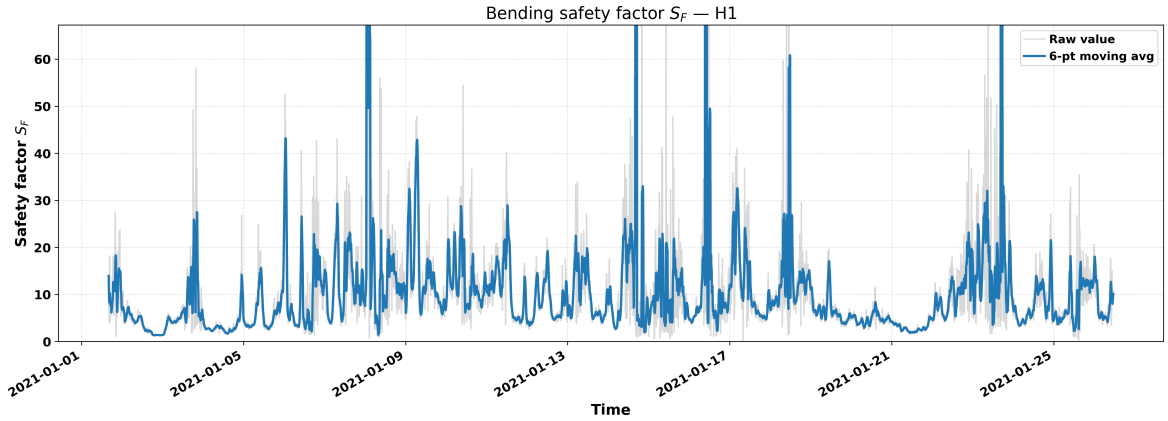


Figure 13: Bending safety factor S_F time series for Helical Stage 1 (H1) pinion on the test set, showing values capped at the 99.5th percentile and the 6-point moving average.

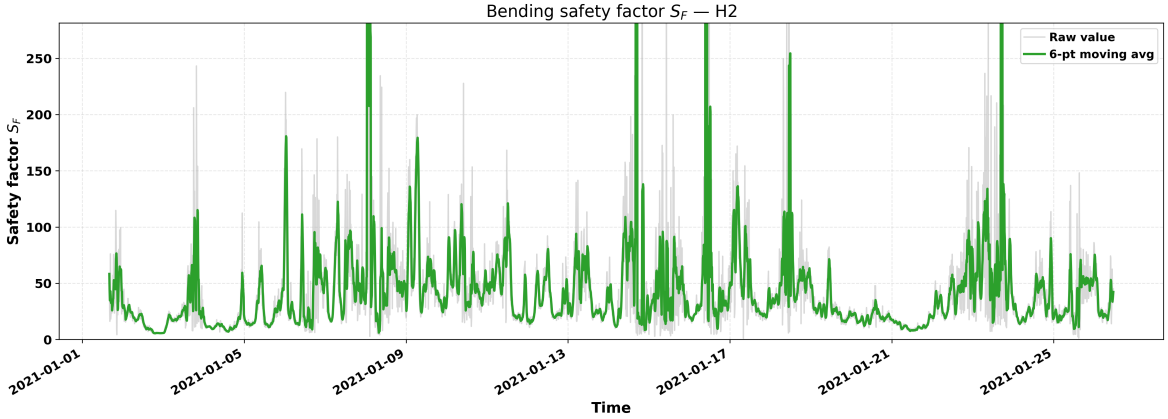


Figure 14: Bending safety factor S_F time series for Helical Stage 2 (H2) pinion on the test set, showing values capped at the 99.5th percentile and the 6-point moving average.

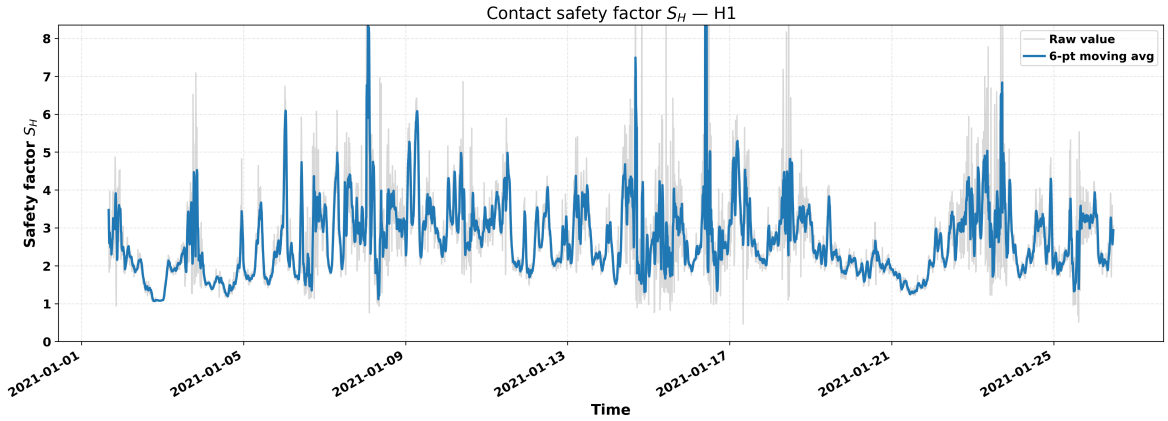


Figure 15: Contact safety factor S_H time series for Helical Stage 1 (H1) pinion on the test set, showing values capped at the 99.5th percentile and the 6-point moving average.

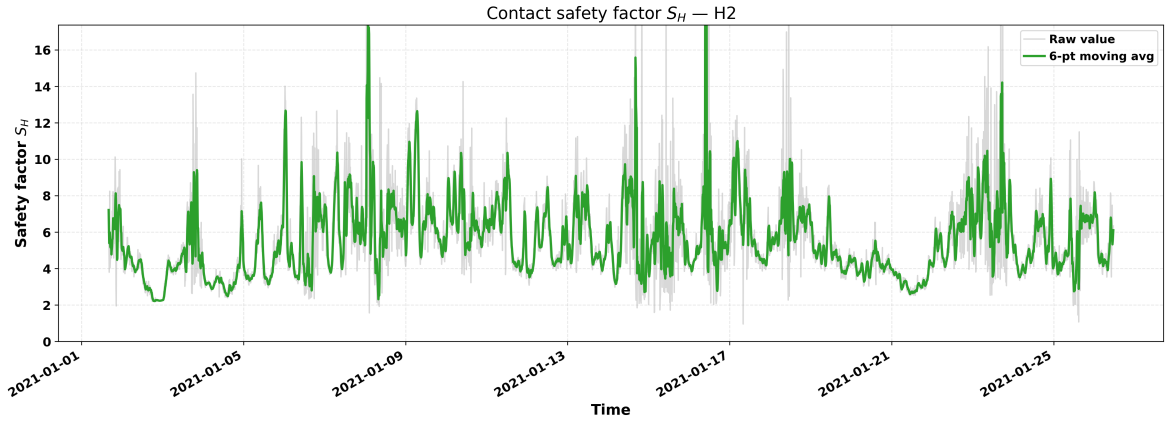


Figure 16: Contact safety factor S_H time series for Helical Stage 2 (H2) pinion on the test set, showing values capped at the 99.5th percentile and the 6-point moving average.

The similarity in patterns for W_t, σ_b, and σ_c can be explained with the mesh load W_t is responsible for driving both stress responses. The bending relation is approximately linear in W_t,

$$\sigma_b \propto W_t,$$

The contact relation is monotone through a square root,

$$\sigma_c \propto \sqrt{W_t}.$$

Therefore the timevarying shape of $\sigma_b(t)$ and $\sigma_c(t)$ follows $W_t(t)$. The square root compresses peaks in σ_c relative to σ_b, but the temporal pattern remains aligned. Differences in level across stages arise from stage-specific geometry and factors ($d,\ F,\ J,\ I,\ K_v,\ K_m$).

The (H1) and (H2) W_t series exhibit the expected stage ratio (H1 larger than H2 due to drivetrain ratios and efficiencies). The moving average follows the operating envelope; transient spikes are visible in the raw series. The 99.5^{th} percentile cap is only a plotting aid and is not used in any stress computation.

For H1, the 6-point average of σ_b lies mainly near the reference level for the pinion with intermittent excursions during high-load intervals; H2 remains lower, consistent with its larger J, larger pitch diameter, and lower stage load. The H1-to-H2 level difference agrees with baseline geometry and factors, indicating correct implementation.

Contact stress magnitudes exceed bending (Hertz pressure) but remain mostly below the 99.5^{th} percentile cap. H1 exceeds H2 because of higher W_t, smaller d, and slightly different $K_v,\ K_m,\ I$. Peaks are less sharp than in bending due to the square-root dependence.

1. Stage scaling is correct: H1 levels exceed H2 for W_t, σ_b, and σ_c, matching train ratios and geometry.

2. Shape coherence holds: W_t, σ_b, and σ_c co-vary in time; σ_c is a compressed version of the W_t pattern; σ_b is nearly proportional.

3. Factor influence behaves as expected: with time-invariant geometry and AGMA factors, only levels shift between stages, not the temporal pattern.

Pinions are typically limiting because they have fewer teeth (lower J, higher σ_b for the same W_t), accumulate more load cycles for life factors, and, in the baseline, the mating gears benefit from hardness ratio effects. Verifying pinions is conservative. The same pipeline can be applied to gears by substituting the gear J, diameters, and life factors; $\sigma_c(t)$ is the same timeline but compared to gear allowables.

The figures provide stress histories consistent with the baseline geometry and factors. The moving average highlights sustained levels; the percentile cap prevents rare spikes from dominating axes. Absolute capacity and detailed life assessments require comparison with AGMA allowables and life factors; the present plots quantify the stress demand only. Within this scope, the results are internally consistent and technically plausible for the current baseline.

4. Conclusion

This project presented an integrated framework that links SCADA based wind power forecasting, drivetrain torque reconstruction, and AGMA gear stress analysis for a wind

turbine gearbox. A multivariate LSTM network was trained on 10-minute SCADA records, using wind speed, rotor speed, wind direction, yaw angle, and related operational signals to predict short-horizon electrical power output. The model achieved a test-set MAE of about 79.5 kW (≈3.8% of the 2.1 MW rated power), an RMSE of roughly 126.5 kW (≈6.0%), and a coefficient of determination R^2, indicating stable generalization across train, validation, and test splits for noisy field data. Directional error analysis showed that most wind sectors remain within 5–8% of rated power, with larger errors concentrated in a few inflow directions, that likely reflect wake or yaw effects rather than purely algorithmic limitations.

The predicted power time series was then converted to drivetrain torque using the measured shaft speeds, and the resulting torque histories were compared against available SCADA torque measurements. After smoothing, the curves of actual and reconstructed torque almost overlap, with peak loads on the order of 10^6 Nm that are physically consistent with a 2.1 MW turbine operating at the observed rotor speeds. This confirms that the LSTM based power forecasts, combined with power to torque conversion, can reproduce both the magnitude and temporal pattern of the mechanical loading that drives gearbox degradation.

Building on this torque history, a three stage planetary–helical gearbox was sized following AGMA 2001-D04 guidance. The design employs a planetary input stage followed by two parallel helical stages with practical tooth counts and carburized, hardened steel gears. Using the reconstructed torques, tangential tooth loads, bending stresses, and contact (pitting) stresses were computed for the helical pinions with all relevant AGMA modifying factors (overload, dynamics, size, load distribution, rim thickness, temperature, hardness ratio, and reliability). The time-varying bending and contact safety factors $S_F(t)$ and $S_H(t)$ remain greater than unity over all observed operating points, supporting the conclusion that the gearbox design is sufficient for the intended service life under the reconstructed torque spectrum.

References

[1] S. Hanifi, X. Liu, Z. Lin, S. Lotfian, A critical review of wind power forecasting methods—past, present and future, Energies 13 (15) (2020) 3764. doi:10.3390/en13153764.

[2] B. Liu, S. Zhao, X. Yu, L. Zhang, Q. Wang, A novel deep learning approach for wind power forecasting based on WD-LSTM model, Energies 13 (18) (2020) 4964. doi:10.3390/en13184964.

[3] Y. Wang, R. Zou, F. Liu, L. Zhang, Q. Liu, A review of wind speed and wind power forecasting with deep neural networks, Applied Energy 304 (2021) 117766. doi:10.1016/j.apenergy.2021.117766.

[4] K. Ömer Ali, Prediction of wind power with machine learning models, Applied Sciences 13 (20) (2023) 11455. doi:10.3390/app132011455.

[5] S. Kumar, K. K. Raj, M. Cirrincione, A. Buscarino, G. Capizzi, G. Limeo, V. Franzitta, R. R. Kumar, A comprehensive review of remaining useful life estimation approaches for rotating machinery, Energies 17 (22) (2024) 5538. doi:10.3390/en17225538.

[6] J. Carroll, S. Koukoura, A. McDonald, A. Charalambous, S. Weiss, S. McArthur, Wind turbine gearbox failure and remaining useful life prediction using machine learning techniques, Wind Energy 22 (3) (2019) 360–375. doi:10.1002/we.2290.

[7] D. Meng, P. Nie, S. Yang, X. Su, C. Liao, Reliability analysis of wind turbine gearboxes: past, progress and future prospects, International Journal of Structural Integrity 16 (1) (2025) 4–38. doi:10.1108/IJSI-08-2024-0129.

[8] L. Yang, Z. Zhang, Wind turbine gearbox failure detection based on SCADA data: A deep learning-based approach, IEEE Transactions on Instrumentation and Measurement 70 (1) (2021) 1–11. doi:10.1109/TIM.2020.3045800.

[9] N. Li, P. Xu, Y. Lei, X. Cai, D. Kong, A self-data-driven method for remaining useful life prediction of wind turbines considering continuously varying speeds, Mechanical Systems and Signal Processing 165 (2022) 108315. doi:10.1016/j.ymssp.2021.108315.

[10] S. Xiang, X. Zheng, J. Miao, Y. Qin, P. Li, J. Hou, M. Ilolov, Dynamic self-learning neural network and its application for rotating equipment RUL prediction, IEEE Internet of Things Journal 12 (9) (2025) 12257–12266. doi:10.1109/JIOT.2024.3520235.

[11] W. Dong, A. R. Nejad, T. Moan, Z. Gao, Structural reliability analysis of contact fatigue design of gears in wind turbine drivetrains, Journal of Loss Prevention in the Process Industries 65 (2020) 104115. doi:10.1016/j.jlp.2020.104115.

[12] A. K. Mahamad, S. Saon, T. Hiyama, Predicting remaining useful life of rotating machinery based on artificial neural network, Computers & Mathematics with Applications 60 (4) (2010) 1078–1087. doi:10.1016/j.camwa.2010.03.065.

[13] Y. Zhang, T. Zhou, X. Huang, L. Cao, Q. Zhou, Fault diagnosis of rotating machinery based on recurrent neural networks, Measurement 171 (2021) 108774. doi:10.1016/j.measurement.2020.108774.

[14] M. Xia, T. Li, L. Liu, L. Xu, S. Gao, C. W. de Silva, Remaining useful life prediction of rotating machinery using hierarchical deep neural network, in: 2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC), 2017, pp. 2778–2783. doi:10.1109/SMC.2017.8123047.

[15] Y. Xia, W. Li, W. Ren, A novel remaining useful life prediction based on transfer hybrid deep neural network under variable working conditions, Physica Scripta 99 (10) (2024) 106011. doi:10.1088/1402-4896/ad7ab9.

[16] H. Zhang, Q. Zhang, S. Shao, T. Niu, X. Yang, Attention-based LSTM network for rotatory machine remaining useful life prediction, IEEE Access 8 (2020) 132188–132199. doi:10.1109/ACCESS.2020.3010066.

[17] S. Siami-Namini, N. Tavakoli, A. S. Namin, The performance of LSTM and BiLSTM in forecasting time series, in: 2019 IEEE International Conference on Big Data (Big Data), 2019, pp. 3285–3292. doi:10.1109/BigData47090.2019.9005997.

[18] S. Siami-Namin, N. Tavakoli, A. Siami Namin, A comparison of ARIMA and LSTM in forecasting time series, in: 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA), 2018, pp. 1394–1401. doi:10.1109/ICMLA.2018.00227.

[19] J. Deutsch, D. He, Using deep learning-based approach to predict remaining useful life of rotating components, IEEE Transactions on Systems, Man, and Cybernetics: Systems 48 (1) (2018) 11–20. doi:10.1109/TSMC.2017.2697842.

[20] M. Pagitsch, G. Jacobs, D. Bosse, Remaining useful life determination for wind turbines, Journal of Physics: Conference Series 1452 (1) (2020) 012052. doi:10.1088/1742-6596/1452/1/012052.

[21] I. Delgado, M. Fahim, Wind turbine data analysis and LSTM-based prediction in SCADA system, Energies 14 (1) (2021) 125. doi:10.3390/en14010125.

[22] Y. Aafif, A. Chelbi, L. Mifdal, S. Dellagi, I. Majdouline, Optimal preventive maintenance strategies for a wind turbine gearbox, Energy Reports 8 (2022) 803–814. doi:10.1016/j.egyr.2022.07.084.

[23] F. Elasha, S. Shanbr, X. Li, D. Mba, Prognosis of a wind turbine gearbox bearing using supervised machine learning, Sensors 19 (14) (2019) 3092. doi:10.3390/s19143092.

[24] W. Zou, Z. Lu, Z. Hu, L. Mao, Remaining useful life estimation of bearing using deep multiscale window-based transformer, IEEE Transactions on Instrumentation and Measurement 72 (2023) 1–11. doi:10.1109/TIM.2023.3268453.

[25] Y. Wang, Y. Zhao, S. Addepalli, Remaining useful life prediction using deep learning approaches: A review, Procedia Manufacturing 49 (2020) 81–88. doi:10.1016/j.promfg.2020.06.015.

[26] G. Li, L. Yu, J. He, W. Zhang, S. Liu, J. Wei, A remaining useful life prediction method for rotating machinery based on interactive BiLSTM, Measurement Science and Technology 36 (1) (2025) 016128. doi:10.1088/1361-6501/ad89ee.

[27] G. Van Houdt, C. Mosquera, G. Nápoles, A review on the long short-term memory model, Artificial Intelligence Review 53 (8) (2020) 5929–5955. doi:10.1007/s10462-020-09838-1.

[28] J. Wang, X. Han, Y. Jia, T. Song, Reliability assessment of wind farm with wind turbine gearbox fatigue, in: Proceedings of the 2021 International Conference on Energy Internet (EI), Vol. 38, 2021, pp. 2706–2711. doi:10.1109/EI252483.2021.9713075.

[29] Y. Lei, N. Li, S. Gontarz, J. Lin, S. Radkowski, J. Dybala, A model-based method for remaining useful life prediction of machinery, IEEE Transactions on Reliability 65 (3) (2016) 1314–1326. doi:10.1109/TR.2016.2570568.

[30] T. Lin, H. Wang, X. Guo, P. Wang, L. Song, A novel prediction network for remaining useful life of rotating machinery, The International Journal of Advanced Manufacturing Technology 124 (11) (2023) 4009–4018. doi:10.1007/s00170-021-08351-1.

[31] R. G. Budynas, J. K. Nisbett, Shigley's Mechanical Engineering Design, 8th Edition, McGraw-Hill, 2008.