lubricants

MDPI

Article

Cleaning, Relubrication and Effectiveness Evaluation of Wind Turbine Main Bearings

Yuezhen Tu [1,*], Qingbiao Jian [2], Jun Liu [3], Tianhong Wang [2], Shuai Yang [4] and Weikang Kong [5]

[1] School of Renewable Energy, Shenyang Institute of Engineering, Shenyang 110136, China
[2] College of Electrical Engineering, Chongqing Electric Power College, Chongqing 400053, China
[3] School of Electric Power Engineering, Sichuan Water Conservancy Vocational College, Chengdu 611800, China
[4] School of Electrical Engineering and Electronic Information, Xihua University, Chengdu 610039, China
[5] Intelligent Transportation System Research Center, Southeast University, Nanjing 210096, China
* Correspondence: yuezhentu@163.com

Abstract

In-service grease cleaning and relubrication is a common maintenance practice intended to restore the condition of wind turbine main bearings. However, evaluating its field effectiveness remains challenging because lubricant displacement, debris transport, thermal dissipation, and structural degradation evolve across mismatched timescales. This study proposes a five-domain field-evaluation framework encompassing intervention fidelity, lubricant compatibility, debris response, operating-normalized thermal behaviour, and engineering outcomes. Sixteen commercial wind turbine main bearings were tracked in a non-randomized paired field setup (eight treatment units receiving 3.2 kg of fresh grease and eight comparators). Multi-source data comprising 537,306 quality-controlled 10 min supervisory control and data acquisition (SCADA) records, 64 condition-monitoring system (CMS) vibration waveforms, grease-iron content measured by inductively coupled plasma optical emission spectrometry (ICP-OES), and physical inspection records were analyzed. The replacement grease met all release criteria (base-oil viscosity 131.8 mm^2/s at 40 °C, four-ball weld load 4200 N) and formed homogeneous static blends across 25–75% mixing ratios with resident grease. In the only pair with complete 60-day grease sampling, iron concentration fell by 20.6% in the treated unit and rose by 1.8% in the comparator. The operating-normalized thermal residual increase in the treatment group was 12.36 residual units lower than that of the comparator group, treatment bearings recorded no CMS or temperature alarms compared with 25 alarm events in comparators, and seven of eight treatment versus two of eight comparator bearings were relatively normal at the post-intervention inspection. Given clipped CMS waveforms (61 of 64), single-pair debris evidence, a short post-intervention SCADA record and no pre-intervention inspection, these results indicate consistent favourable trends rather than a demonstrated causal effect. By integrating multi-stream evidence without relying on unverified single-metric health indices, this framework shows that the intervention was executed as intended and was followed by consistently favourable indicators, and it identifies the observations required before a state-restoring effect on wear or bearing life can be inferred.

Keywords: wind turbine bearing; grease-purge cleaning; relubrication; tribology; SCADA; condition monitoring; effectiveness evaluation

Check for updates

Received: 24 September 2026
Revised: 22 October 2026
Accepted: 28 October 2026
Published: 23 November 2026

Copyright: © 2026 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.

1. Introduction

Wind turbine main bearings operate under highly variable and coupled mechanical conditions. Rotor thrust, gravitational loading, torque, turbulence and transient control actions continuously change the bearing load distribution, while drivetrain deformation can further modify local roller–raceway contact conditions. The relationship between wind-field characteristics and main-bearing loading has been demonstrated through aeroelastic and multibody analyses, with different drivetrain configurations showing distinct loading behaviours [1]. A comprehensive review further indicates that main-bearing damage involves interacting mechanical, lubrication and operational factors rather than a single failure mechanism [2]. Field measurements and three-dimensional drivetrain modelling have also demonstrated the importance of realistic loading histories in evaluating main-bearing fatigue [3]. Therefore, the actual condition of a main bearing cannot be adequately represented by a single operating parameter or monitoring signal.

Lubrication provides a critical link between operating conditions and bearing degradation. Wind turbine main bearings can experience mixed or boundary lubrication under certain combinations of load, speed, temperature and lubricant starvation, while effective film formation is strongly influenced by grease properties and operating conditions [4,5]. Grease is not simply a lubricant reservoir; oil separation, thickener structure, mechanical working and contamination can alter its ability to supply lubricant to the loaded contact [6]. Experiments under wind turbine bearing conditions have shown that different greases can exhibit substantially different wear behaviour under oscillatory and highly loaded conditions [7]. Rating-life analyses have further shown that standard calculations predict main-bearing lives beyond the 20-year design life and therefore do not account for the failure rates reported in the field, indicating that damage drivers not represented in these calculations act on main bearings [8]. The life-modification analysis of grease-lubricated wind turbine bearings indicates that one such driver, the actual lubrication state, should be considered when evaluating bearing performance [9]. These findings suggest that lubricant replacement should be evaluated according to its influence on the actual lubrication condition rather than simply the quantity of fresh grease introduced.

Bearing degradation may also occur under operating conditions that appear moderate at the turbine level. Studies of wind turbine pitch bearings have shown that oscillatory motion, contact geometry and lubrication conditions strongly affect local wear and fretting behaviour [10,11]. Although pitch and main bearings have different operating modes, these studies demonstrate an important tribological principle: macroscopic operating variables do not necessarily reflect local contact severity. Experimental and scaled test-rig studies further show that the dynamic behaviour of wind turbine main bearings is strongly dependent on the simulated operating conditions and bearing configuration [12]. Consequently, degraded grease, wear debris and local surface damage may develop without an immediate or proportional change in conventional turbine-level measurements.

Supervisory control and data acquisition (SCADA)-based condition monitoring has become an important approach for detecting abnormal main-bearing behaviour because it can use existing operational measurements without additional sensors. Ensemble learning and SCADA-based prognosis have demonstrated the feasibility of detecting and predicting main-bearing faults from operational data [13,14]. One-class support-vector machines and multivariate time-series forecasting have further been used to identify deviations from normal bearing behaviour [15,16]. Principal-component analysis, recurrent neural networks and convolutional autoencoders have subsequently extended SCADA-based early fault detection and prognosis [17–19]. Other approaches, including physics-informed neural networks and acoustic-emission monitoring, have investigated main-bearing fatigue and condition characterization from complementary perspectives [20,21]. These studies

demonstrate the value of indirect monitoring, but their primary objectives are fault detection, prognosis or life estimation rather than determining whether a completed maintenance intervention has restored the bearing condition.

This distinction is particularly important for in-service grease-purge cleaning followed by relubrication. Unlike bearing disassembly or solvent cleaning, this procedure attempts to modify the internal lubrication state through the existing grease path. The intended process is to displace degraded grease and suspended contaminants, introduce compatible fresh grease and improve lubricant supply to the loaded contacts. However, the quantity of grease displaced does not necessarily represent the quantity of degraded material removed. Residual grease may remain in inaccessible regions, material compatibility may affect the resulting lubricant state, and operating conditions after maintenance may differ substantially from those before maintenance. Thus, a maintenance work order or grease addition record alone cannot demonstrate that the intended physical intervention was sufficiently achieved.

The existing literature therefore leaves a specific field-evaluation gap. Lubrication studies establish relationships between grease properties and tribological performance, whereas condition-monitoring studies establish relationships between operational signals and bearing faults. Even studies based on remaining-power estimation and field-oriented monitoring primarily focus on fault-related prediction rather than maintenance intervention effectiveness [22]. Similarly, SCADA-based unsupervised prognosis demonstrates that changes in a representative bearing temperature can be detected months before failure, but such models do not determine whether a subsequent maintenance action causes a change in bearing condition [23]. A field evaluation should instead connect the physical intervention with multiple independent forms of evidence, including lubricant displacement, material compatibility, debris response, operating-normalized thermal behaviour and engineering outcomes.

To bridge the gap between laboratory tribological characterization and operational condition monitoring, this study proposes a five-domain field-evaluation framework (Figure 1) to systematically evaluate the effectiveness of in-service grease-purge cleaning and relubrication across sixteen utility-scale wind turbine main bearings arranged in a non-randomized paired treatment-comparator design. The empirical investigation integrates physical mass displacement verification, static blend compatibility screening (25–75% mixing ratios), longitudinal grease-iron debris tracking via inductively coupled plasma optical emission spectrometry (ICP-OES), operational thermal residual normalization across 537,306 quality-controlled SCADA records, and engineering outcome assessment through condition monitoring system (CMS) vibration alarms, SCADA-temperature alarms and end-of-period physical inspections. The primary novelty and contribution of this work lie in establishing an end-to-end evidence chain that connects a physical tribological intervention with independent machine-level responses, rather than relying on isolated temperature channels or small-scale laboratory coupons. Furthermore, by introducing a power- and ambient-conditioned thermal residual formulation, the framework reduces the influence of operating exposure on the thermal comparison, while explicitly maintaining separate evidence boundaries across heterogeneous measurement streams. In this way, the framework shows which parts of a state-restoration claim are supported by field evidence and which remain unresolved, without making speculative, unverified claims regarding bearing service-life extension.

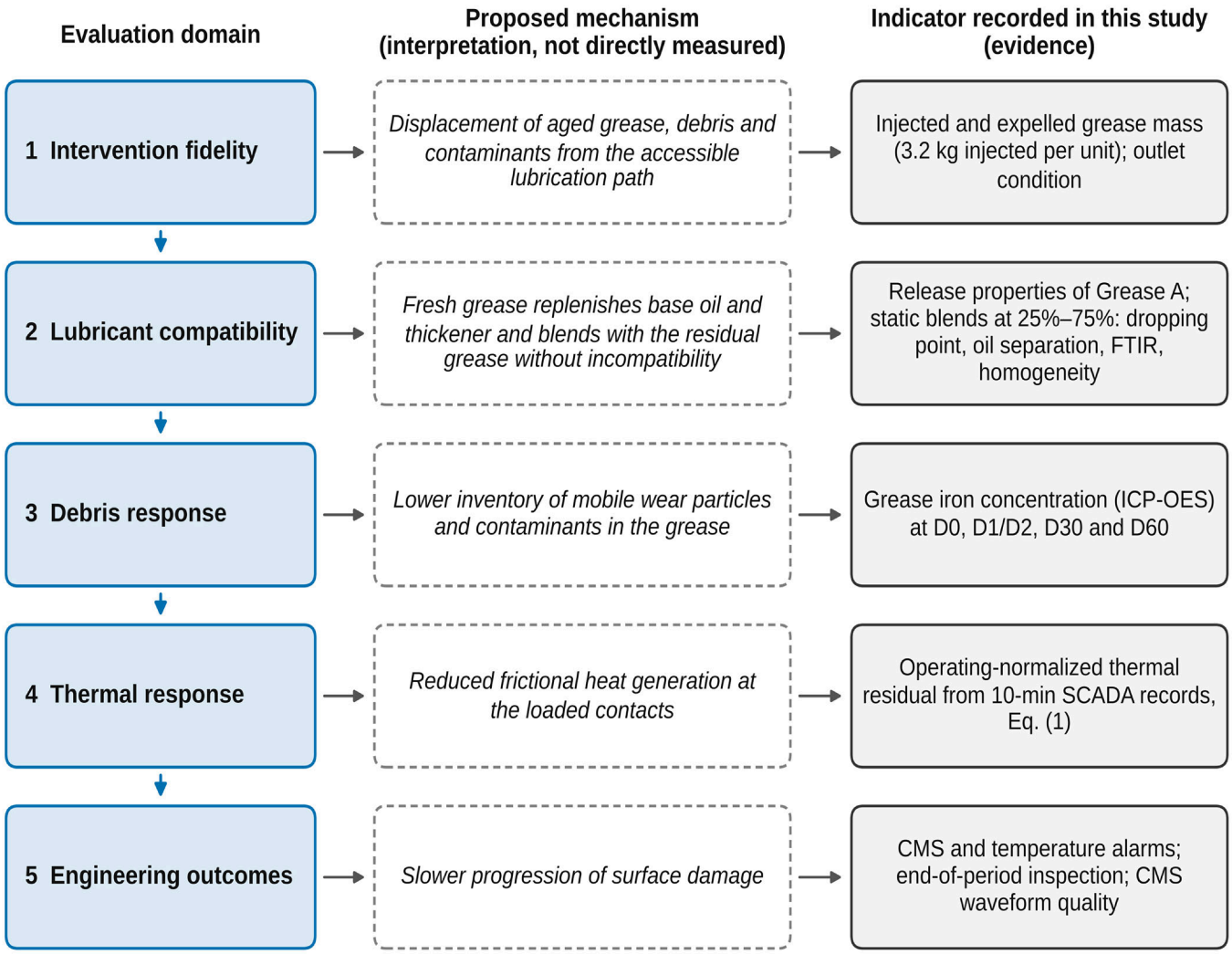

Figure 1. Mechanistic chain and five-domain evaluation framework, from cleaning and relubrication to condition response. Left: evaluation domains in their hypothesized causal sequence; middle (dashed outline): proposed mechanisms, which were not directly measured; right (solid outline): indicators recorded in this study.

2. Materials and Methods

2.1. Evaluated Bearings, Field Intervention and Mechanistic Hypotheses

The evaluated components were the grease-lubricated main bearings of 16 identical 2.0 MW doubly fed induction generator (DFIG) turbines at a single onshore wind farm; each turbine contributed one main bearing. Each bearing is a double-row spherical roller bearing with bore diameter × outside diameter × width of 280 × 580 × 175 mm and a machined brass cage, mounted with a three-point drivetrain suspension, fixed on the main shaft and located in the bearing housing. The grease is held in the free volume of the bearing and housing cavities (approximately 12 kg capacity); fresh grease is admitted through an annular groove and lubrication holes in the outer ring and spent grease leaves through a purge port at the bottom of the bearing housing. Before the study, the bearings were relubricated by a scheduled relubrication every six months with approximately 0.5 kg of grease per application. These design features are reported because they determine which internal regions fresh grease reaches during a purge, where wear particles accumulate and which wear modes are expected; main-bearing type and drivetrain suspension differ considerably between turbine designs [2].

The analysis unit was one wind turbine main-bearing assembly identified by an anonymous unit code. Sixteen units were arranged as eight treatment–comparator pairs using the available information on cumulative operating hours and operating context (Figure 2). Eight treatment units underwent the defined cleaning–relubrication procedure, and eight comparator units did not undergo the protocol purge. The assignment was retained as documented in the field records. It is described as paired rather than randomized because an independently verifiable allocation sequence was not available. Within each pair, the units were matched on cumulative operating hours and operating context with the same site and turbine model, and with a maximum difference in cumulative operating hours of approximately 5%; the unit that received the protocol purge was selected by maintenance schedule and access. Before the study, all units were maintained under the same scheduled relubrication regime and grease type, with no previous main-bearing replacement. Baseline comparability of the degradation state could be assessed only through the indicators available for all 16 units: the mean baseline grease-iron concentration was 132.2 mg/kg in the treatment group and 135.7 mg/kg in the comparator group (a difference of 2.6%; Section 3.4), the mean turbine-specific thermal baseline b_i of Equation (1) was 66.1 and 65.6 residual units, and the numbers of alarm events in a pre-intervention window of the same length as the post-intervention window were two and three, respectively. No pre-intervention inspection was available, so baseline surface damage could not be compared directly; this limitation is carried into the interpretation of the inspection outcome (Section 3.5).

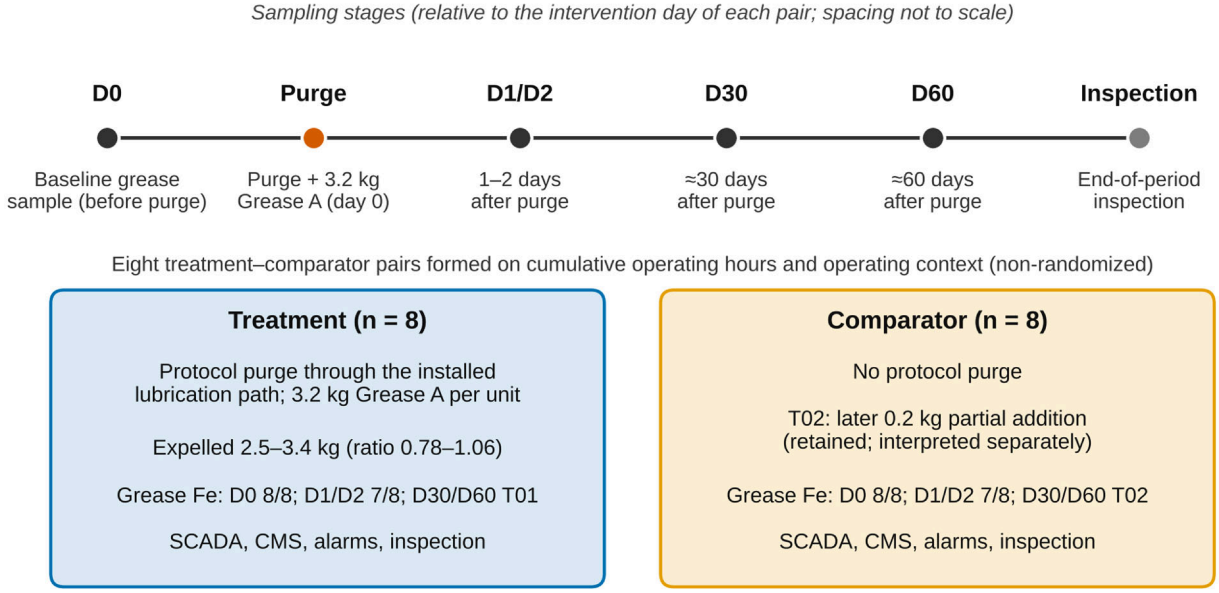

Figure 2. Paired field-evaluation design and evidence streams. Sampling stages are labelled by the nominal number of days after the intervention day of each pair: D0, baseline sample taken before the purge; D1/D2, 1–2 days after the purge; D30 and D60, approximately 30 and 60 days after the intervention. Spacing is not to scale.

Cleaning was operationalized as grease displacement through the installed lubrication path. No solvent wash, bearing opening or disassembly was recorded. Each treatment unit received a nominal 3.2 kg charge of Grease A. Expelled mass ranged from 2.5 to 3.4 kg, providing a directly observed measure of intervention exposure. A ratio close to or above unity does not imply that all grease inside the bearing was replaced; it indicates that a mass comparable with the injected charge emerged at the outlet and may include resident grease already present in the accessible flow path.

The comparator set was not completely untreated throughout the observation record. Unit T02 received a later 0.2 kg partial addition without the protocol purge. Rather than concealing this deviation, T02 was retained as a contaminated comparator and interpreted separately when its longitudinal iron trajectory was examined. This treatment avoids redefining the group after observing the result and acknowledges that a small addition may alter grease mobility without reproducing the cleaning exposure of the treatment units.

Three mechanistic hypotheses were specified (Figure 1). The material-displacement hypothesis predicted that a credible purge would be supported by expelled mass, outlet condition and a later reduction in the debris marker. The lubricant hypothesis required the replacement grease to satisfy recorded release criteria and to remain homogeneous when mixed with residual grease. The machine-response hypothesis predicted a smaller operating-normalized thermal increase, fewer condition-monitoring alarms and a higher proportion of relatively normal-inspection findings in the treatment group. These hypotheses were evaluated jointly, but no single endpoint was designated as a surrogate for bearing life.

2.2. Lubricant Characterization and Compatibility Experiment

The replacement lubricant was anonymized as Grease A. The available quality record reported base-oil viscosities of 131.8 mm^2/s at 40 °C and 12.6 mm^2/s at 100 °C, worked penetration of 280 in 0.1 mm units, a dropping point of 268 °C, a four-ball weld load of 4200 N, oxidation stability of 0.3 bar and water content of 0.05%. These quantities were used as release descriptors. They were not converted into a calculated film thickness because contact geometry, speed, inlet supply and pressure–viscosity response were not available at the resolution required for such a calculation.

Grease A was the replacement grease specified for the intervention of the turbine manufacturer's approved main-bearing grease, a lithium-complex thickener with a mineral base oil. Its suitability was judged against the duty of a main bearing, which rotates slowly under high and fluctuating loads and can therefore operate in mixed lubrication for a non-negligible part of its life, particularly when the inlet is starved [4,5]. Three requirements follow. First, the base oil must be viscous enough to form a separating film at low entrainment speed; the measured base-oil viscosities (131.8 mm^2/s at 40 °C and 12.6 mm^2/s at 100 °C) agreed with the reference values to within 0.2% and 0.8% (Table 1), and at low speed the thickener can raise the film thickness above that predicted from the base oil alone [24]. Second, because full film separation cannot be assured, the grease must protect the contacts under boundary conditions; the four-ball weld load of 4200 N exceeded the 4000 N acceptance value, indicating an extreme-pressure capability suited to heavily loaded, low-speed contacts. Third, the grease must remain mechanically and thermally stable and pumpable through the relubrication path: a worked penetration of 280 (0.1 mm) lies in the middle of the 265–295 range of National Lubricating Grease Institute (NLGI) consistency grade 2, the dropping point of 268 °C is far above normal main-bearing operating temperatures, and the recorded oxidation-stability value (0.3 bar) and water content (0.05%) are consistent with an unaged, uncontaminated product; consistency, oil bleed and oxidation stability govern how grease degrades in service [6,25]. The release properties therefore establish that Grease A met its specification and was suitable in kind for the application; they do not show that it was the optimal grease, which would require film-thickness and friction measurements under main-bearing conditions.

Static compatibility was assessed by blending Grease A with the residual grease at mass ratios of 100:0, 75:25, 50:50, 25:75 and 0:100. The endpoints were worked penetration, dropping point, oil separation after the recorded thermal hold, Fourier-transform infrared (FTIR) spectral consistency and visual homogeneity. The graduated mixture series was important because a purge does not create an instantaneous interface between pure old

and pure new grease. Intermediate compositions are expected in ports, channels and partially renewed bearing volumes. The design follows the logic of the ASTM D6185 practice for binary grease mixtures, which judges compatibility by comparing mixtures with the neat greases and notes that 25:75 and 75:25 mixtures are used by some laboratories in place of its prescribed 10:90, 50:50 and 90:10 ratios [26]; however, the prolonged-working (shear-stability) test and the post-storage consistency change in that practice were not part of the recorded series, so the present series is a static screen rather than a complete compatibility determination.

Table 1. Anonymized release properties and static compatibility endpoints.

Property	Fresh Grease	Acceptance/Reference	Interpretation
Base-oil viscosity at 40 °C	131.8 mm^2/s	132 mm^2/s	Supports film-forming comparison
Base-oil viscosity at 100 °C	12.6 mm^2/s	12.5 mm^2/s	Thermal viscosity reference
Worked penetration	280 (0.1 mm)	265–295	Within supplied range
Dropping point	268 °C	≥250 °C	Thermal release criterion
Four-ball weld load	4200 N	≥4000 N	Load-carrying indicator
Oil separation at 50:50 blend	2.1%	≤5%	No excessive separation
FTIR/visual compatibility	Good/homogeneous	No new peak or phase separation	Static material gate

The decision criterion was deliberately conservative. A mixture was considered to pass the static material gate when no phase separation or new FTIR inconsistency was recorded, consistency remained within the documented working range and oil separation remained below the recorded limit. The interpretation follows the grease literature, in which material properties constrain but do not uniquely determine contact performance [6,7,9]. A passed gate means that immediate incompatibility was not observed under the tested condition; it does not establish dynamic compatibility under high-shear, repeated load reversal, metal-catalyzed oxidation or long-duration churning.

The compatibility series was analyzed descriptively because the available record contained the mixture-level endpoints rather than replicate distributions. Monotonic changes across the blend ratio were examined for physical plausibility. In particular, the progression in dropping point and oil separation was evaluated to determine whether intermediate mixtures showed an abrupt discontinuity that would suggest interaction between the two grease systems.

2.3. Field Sampling, SCADA and CMS Processing

Grease samples were assigned to relative stages rather than calendar dates, each labelled Dn, where n is the nominal number of days after the intervention day of the corresponding treatment–comparator pair (Figure 2): D0 denoted the baseline sample of resident grease taken before the purge, D1/D2 the early post-intervention stage (1–2 days after the purge), and D30/D60 the longitudinal follow-up (approximately 30 and 60 days after the intervention); actual collection was within ±3 days of the nominal day. Iron concentration was measured by ICP-OES and expressed in mg/kg. Samples were drawn from the purge outlet with a clean sampling scoop (approximately 10 g per sample), acid-digested and measured by ICP-OES. Baseline samples were available for all 16 units and early samples for 14 units (seven per group). Complete D30/D60 trajectories were available only for T01 (treatment) and T02 (comparator). Missing stages were left missing; observations were not temporally reassigned to create a balanced laboratory panel.

Grease iron was treated as a condition marker rather than a direct wear-rate measurement. Its concentration can reflect newly generated wear particles, historical debris mobilized by the purge, non-uniform sampling from the lubrication path and dilution by fresh grease. Consequently, the group baseline and early-stage means were used to describe

comparability and short-term redistribution, while the complete T01-T02 trajectories were used only for a case-level longitudinal contrast. The analysis did not infer a fleet-level wear coefficient from two complete trajectories.

Four processes that can change the measured concentration without any change in wear were considered explicitly. (i) Displacement: The expelled-to-injected mass ratio was used only to document exposure; because residual grease remains in regions not swept by the purge, the post-purge concentration was not interpreted as the concentration of the whole bearing charge. (ii) Dilution: If the sampled grease after the purge contains a mass fraction f of fresh grease with little iron, its concentration falls to about $(1 - f)$ times the baseline value even when no wear occurs, so a relative decline x corresponds to an apparent fresh-grease fraction of at least x in the sampled grease; this reference was used to judge whether a decline exceeded what dilution alone could explain (Section 3.4). (iii) Replenishment: The later 0.2 kg addition to the comparator T02 could dilute or mobilize grease near the sampling point, which is why T02 was interpreted as a contaminated comparator. (iv) Sampling location: All stages of a given unit were sampled at the same purge-outlet location; because particle concentration differs between the outlet, the housing and the contact region, stage-to-stage changes were interpreted only within a unit. In addition, the response of ICP analysis to large wear particles depends on sample preparation: the standard determination of metals in greases acid-decomposes the sample and thereby dissolves particulate iron [27], whereas methods that dilute and aspirate the sample directly give low recoveries for particles larger than a few micrometres [28]. Large fatigue or adhesive-wear particles are therefore under-represented only if the samples were not fully digested (preparation used here: acid digestion (full decomposition)).

The SCADA archive contained 569,104 ten-minute records. Quality control excluded missing reports, zero-power observations, startup transients, maintenance windows, curtailment or icing states and remaining quality flags. The sequence retained 537,306 records, or 94.41% of the raw archive. This restriction was intended to reduce exposure heterogeneity before normalization; it was not used to select observations on the basis of the measured bearing temperature.

An operating-normalized thermal residual was then calculated from bearing temperature, ambient temperature and power fraction. The temperature difference was divided by the power fraction clipped to the interval [0.1, 1.0], and a turbine-specific pre-intervention baseline was subtracted. The lower clip prevents unstable amplification at very low load, whereas the upper clip preserves the rated-load scale. The resulting residual is expressed in descriptive residual units and should not be interpreted as an absolute heat-generation rate. This construction applies the central principle of SCADA normal-behaviour methods—conditioning temperature on operating exposure—without claiming the predictive complexity of the published machine-learning models [13–16,22,29].

$$\mathrm{ri,t} = (\mathrm{Tb,i,t} - \mathrm{Ta,i,t})/\mathrm{clip}(\mathrm{Pi,t/Pr},\, 0.1,\, 1.0) - \mathrm{bi} \quad (1)$$

where ri,t is the operating-condition-normalized thermal residual; Tb and Ta are bearing and ambient temperature; $Pi,t/Pr$ is the power fraction clipped to [0.1, 1.0]; and bi is the turbine-specific pre-intervention baseline.

Because the power fraction is dimensionless, r carries the dimension of a temperature difference; however, because it is scaled by the inverse power fraction (by up to a factor of 10 at the lower clip) and centred on a turbine-specific baseline, it is reported in residual units rather than as a physical temperature. For visualization, the quality-controlled archive was divided into nine consecutive time blocks in chronological order (B1–B9), with the purge boundary between B7 and B8; the group-mean residual of each block is reported in Section 3.3; each block was a consecutive one-ninth of the quality-controlled archive.

Because the blocks contain different numbers of valid records, the pre- and post-intervention period means reported in Section 3.3 are computed over all valid records and are therefore not simple averages of the block means confirmed.

The CMS inventory comprised 64 indexed WAV (Waveform Audio File Format) files, each containing a raw time-domain vibration waveform acquired by the CMS and stored as 16-bit signed integers (int16; full scale −32,768 to +32,767 counts). File metadata, coverage and amplitude-limit status were reviewed before any physical interpretation. A waveform was classified as clipped when its samples reached the int16 full-scale value. Sixty-one waveforms met this criterion and only three did not. Because saturation changes amplitude statistics and introduces spectral artefacts [30], the files were not used for an unrestricted comparison of root-mean-square (RMS) amplitude, kurtosis or band energy. CMS contributed evidence about alarm occurrence and signal quality, not a quantitative vibration effect size.

2.4. Effectiveness Indicators and Equations

Effectiveness was represented as a set of directionally interpretable indicators. Intervention fidelity was described by injected mass, expelled mass and their ratio. The material gate was described by release properties and mixture behaviour. Debris response was represented by the relative change in grease iron for the complete trajectories. Thermal response was represented by the difference between the comparator and treatment changes in the normalized residual. Engineering response was represented by alarm rates per unit and the proportion of relatively normal inspections. Table 2 summarizes, for each domain, the indicators, measurement procedures and data sources, evaluation criteria and evidence boundaries.

The indicators were kept separate because they arise from different physical and sampling processes. A temperature residual is based on hundreds of thousands of operational records, whereas an inspection classification is one unit-level observation and the complete iron contrast contains one treatment–comparator pair. Standardizing and averaging them would produce a numerically convenient index while assigning unjustified equal weight to measurements with unequal coverage. Directional convergence was therefore assessed qualitatively after each result had been reported in its original unit.

For lower-is-better outcomes, the contrast was oriented so that a positive value favoured treatment. The thermal contrast was calculated as the comparator within-group change minus the treatment within-group change. The iron contrast was calculated as the difference between the proportional D0-to-D60 reductions for T01 and T02. The inspection contrast was the difference in relatively normal proportions, and alarm rates were counts divided by the number of units in the corresponding group.

The directional effectiveness indicators were defined as

$$Eth = \Delta\mathrm{Y,C} - \Delta\mathrm{Y,T} \tag{2}$$

where $\Delta_{\mathrm{Y,C}}$ and $\Delta_{\mathrm{Y,T}}$ are the changes in the group-mean normalized residual from the pre- to the post-intervention period in the comparator and treatment groups, respectively, and *Eth* is the directional thermal contrast. A positive value means that the treatment group increased less than the comparator group; it is not a causal coefficient.

$$EFe = [(\mathrm{FeT,D0} - \mathrm{FeT,D60})/\mathrm{FeT,D0}] - [(\mathrm{FeC,D0} - \mathrm{FeC,D60})/\mathrm{FeC,D0}] \tag{3}$$

where E_{Fe} is calculated only for the complete T01-T02 pair and is reported as a case-level relative-change contrast; $\mathrm{Fe_T}$ and $\mathrm{Fe_C}$ denote the grease-iron concentrations of the treatment unit (T01) and the comparator unit (T02) at the stage indicated, and each bracketed term

is the proportional reduction from D0 to D60, so that a negative term denotes an increase in concentration.

$$Einsp = \mathrm{pT,normal} - \mathrm{pC,normal} \quad (4)$$

where *Einsp* is the difference in the proportion of units classified as relatively normal at inspection.

$$Ag = \Sigma\ \mathrm{alarmsg/Ng} \quad (5)$$

where *Ag* is the group alarm rate per unit. The effectiveness profile reports these domains separately rather than inventing a common unit.

Table 2. Five-domain evaluation framework: indicators, measurement procedures, evaluation criteria and evidence boundaries.

Domain	Indicator (s)	Measurement Procedure and Data Source	Evaluation Criterion	Evidence Boundary
Intervention fidelity	Injected mass; expelled mass; expelled-to-injected ratio; outlet condition	Injected charge and grease collected at the purge outlet weighed for each treatment unit; outlet grease described (colour, homogeneity, visible metallic debris)	Specified 3.2 kg charge injected; expelled mass comparable with the charge; clear outlet path	Document exposure; does not show complete renewal of resident grease
Lubricant compatibility	Release properties; static blend endpoints	Quality record of Grease A (Table 1); blends with residual grease at 100:0, 75:25, 50:50, 25:75 and 0:100: worked penetration, dropping point, oil separation after thermal hold, FTIR, visual homogeneity	Release properties within acceptance/reference values; no phase separation or new FTIR band; consistency within working range; oil separation below 5%; no abrupt change across the blend series	Static screen; no shear-stability, storage-stability or dynamic test
Debris response	Grease iron concentration; E_{Fe} (Equation (3))	ICP-OES iron (mg/kg) in grease sampled at D0, D1/D2, D30 and D60	Favourable direction if $E_{Fe} > 0$; reduced wear generation requires a decline beyond the dilution expected from the purge	One complete pair; dilution and sampling location not separable; large particles may be under-detected, depending on sample preparation
Thermal response	Normalized residual r (Equation (1)); Eth (Equation (2))	10 min SCADA bearing temperature, ambient temperature and active power after quality control; group means for the pre- and post-intervention periods	Favourable direction if Eth > 0 (smaller increase in the treatment group)	Descriptive; short post-intervention period; shared terminal rise; not a causal difference-in-differences estimate
Engineering outcomes	CMS and temperature alarm events per unit; units with ≥1 alarm; inspection class; CMS waveform quality	CMS and SCADA alarm logs for all 16 units; end-of-period inspection classified as relatively normal or abnormal; WAV files screened for int16 clipping	Favourable direction if alarm rates and the abnormal-inspection proportion are lower in the treatment group	Post-only inspection; threshold-dependent alarms; repeated alarms within units; 61 of 64 waveforms clipped

2.5. Statistical Interpretation and Evidence Boundary

The analysis was primarily descriptive because the evidence structure did not support a common inferential model. Baseline and early iron values were summarized by group means. The complete iron trajectories were reported at unit level. SCADA means were summarized for relative pre- and post-intervention periods, and the change difference was presented in a difference-in-differences (DID) form solely to clarify direction. It was not interpreted as a causal coefficient.

Where the data structure allowed, uncertainty was quantified with methods suited to very small samples, and the unit of inference was the bearing or the matched pair rather than the individual record. (i) The 20.6% decline in T01 and the 22.4-percentage-point contrast E_{Fe} are single-unit and single-pair quantities; with one observation per group, neither a confidence interval nor a significance test can be computed, so these values are reported as case-level observations. (ii) For the thermal contrast, the 537,306 ten-minute records are strongly autocorrelated and nested within 16 units; treating them as independent observations would severely understate the standard error of a difference-in-differences contrast [31]. The thermal contrast was therefore evaluated at the unit level. For each unit, the change Δr_i (post-period mean minus pre-period mean of r) was computed, and for each pair j the within-pair difference $d_j = \Delta r_{C,j} - \Delta r_{T,j}$ was formed, so that a positive value favours treatment. The mean of d_j was tested with a two-sided exact sign-flip permutation test over all $2^8 = 256$ sign assignments, and a 95% percentile bootstrap confidence interval was obtained by resampling pairs (10,000 resamples); as a sensitivity analysis that ignores the pairing, the 16 units were also permuted over all 12,870 allocations to two groups of eight. Because a causal interpretation would additionally require parallel pre-intervention trends [32], these quantities describe uncertainty in the observed contrast, not in a causal effect. (iii) The proportions of relatively normal-inspection findings were compared with Fisher's exact test [33], and the difference in proportions was reported with a Newcombe hybrid-score 95% confidence interval [34]; because these statistics ignore the pairing, the exact McNemar test on pair-level outcomes was also computed. (iv) Alarm counts were summarized as events per unit with exact Poisson 95% confidence intervals, and the two groups were compared with an exact conditional test for two Poisson counts with equal exposure [35]. Because repeated alarms from the same bearing are not independent events [36], this event-level test is anti-conservative; the numbers of units with at least one alarm were therefore also compared with Fisher's exact test and, respecting the pairing, with the exact McNemar test. All tests were two-sided and computed in Python (version 3.9) with SciPy; *p*-values are reported as measures of compatibility with no between-group difference, not as confirmation of a treatment effect.

A causal difference-in-differences interpretation would require, among other conditions, a defensible parallel-trend period, an adequately aligned post-intervention interval and limited contamination between groups [32]. The intervention occurred close to the end of the available archive, so the post-intervention segment was short relative to the recurring operating pattern. T02 also received a partial addition. These features precluded the claim that the 12.36-unit contrast isolated the effect of cleaning–relubrication from all concurrent time-varying influences.

Physical inspection was reduced to relatively normal or abnormal using the recorded engineering classifications: a bearing was classified as abnormal when the inspection record reported wear, pitting, indentation, residue or fatigue spalling, and as relatively normal otherwise endoscopic inspection through the inspection port, classified by a certified maintenance engineer. Alarm events were counted within each group and divided by eight units to obtain rates per unit. No time-to-event analysis was attempted because there were no main-bearing shutdowns or replacements and exposure time was not observed with

the alignment required for a hazard model. The zero event count is therefore an outcome description, not evidence that future failure risk was zero.

Alarm events were taken from the alarm logs of the two monitoring systems over the observation window the 30 days before to the 30 days after the purge, with pre- and post-intervention counts reported separately. A CMS alarm event was defined as an alarm record generated by the CMS for a main-bearing measurement channel when a monitored vibration feature exceeded its configured alarm level overall RMS velocity and envelope acceleration, with the manufacturer's default alarm levels. A temperature alarm event was defined as a main-bearing temperature alarm record generated by the SCADA system when the bearing temperature exceeded its configured threshold 80 °C. The same alarm settings applied to treatment and comparator units were confirmed. Each logged alarm activation was counted as one event, so repeated alarms from the same bearing were counted separately activations within a short interval were not merged; each logged activation was counted separately; the event counts are therefore reported together with the number of bearings that recorded at least one alarm.

Evidence strength was judged from coverage, independence and consistency. Purge and release-property records were complete for the intervention; the SCADA panel was dense but temporally constrained; the longitudinal laboratory evidence was sparse; CMS amplitude evidence was quality-limited; and inspection was post-only. These distinctions were carried into the results and discussion so that a favourable direction was not confused with equal certainty across domains.

3. Results

3.1. Cleaning and Relubrication Fidelity

All eight treatment units received the same nominal fresh-grease mass (Table 3). Expelled mass ranged from 2.5 to 3.4 kg, yielding expelled-to-injected ratios of 0.78–1.06. The range shows that the intervention produced substantial material displacement in every treated unit while retaining realistic between-unit variation in the accessible lubrication volume and outlet response. A ratio above one is physically consistent with expulsion of resident grease in addition to the newly injected charge and should not be interpreted as a mass-balance error.

Table 3. Intervention fidelity and field evidence coverage.

Domain	Available Evidence	Coverage	Interpretation
Purge exposure	3.2 kg injected; 2.5–3.4 kg expelled	8/8 treatment units	Controlled exposure
Baseline Fe	ICP-OES grease iron	16/16 units	Group comparison available
Longitudinal Fe	D30/D60 trajectories	2/16 units	Complete-pair analysis only
SCADA	569,104 raw; 537,306 QC	16/16 units	5.59% excluded
CMS	64 indexed WAV; 61 clipped	64 files	Quality-limited vibration inference
Inspection/alarms	Inspection plus two alarm classes	16/16 units	Engineering cross-check

Field observations described light-coloured, homogeneous expelled grease and clear outlet paths in seven treatment units. T07 retained historical metallic debris in its D0 material. Taken together, the observations support procedural fidelity but also separate cleaning success from damage reversal. The purge can remove mobile contamination from

the accessible grease path; it cannot remove an indentation, repair a spall or guarantee that all material in a shielded or poorly supplied zone has been renewed.

3.2. Lubricant Gate: Release Properties and Static Compatibility

The measured release properties met the recorded acceptance criteria (Table 1; Figure 3a). Grease A had a worked penetration of 280, dropping point of 268 °C and four-ball weld load of 4200 N, while its base-oil viscosity closely matched the reference values at both recorded temperatures. These results establish that the grease used in the intervention corresponded to the intended material specification and provide a traceable basis for interpreting the subsequent mixture tests.

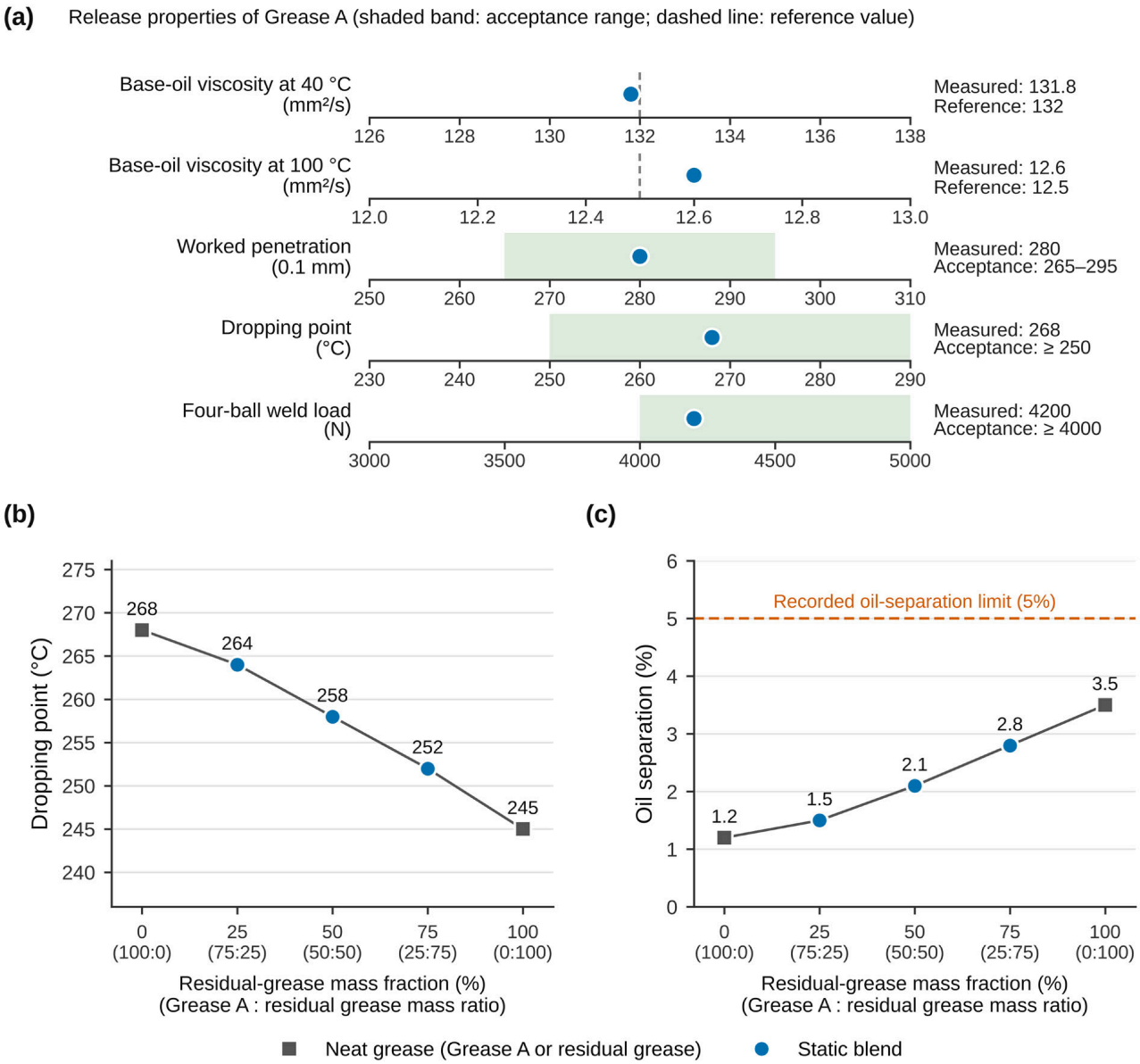

Figure 3. New-grease properties and static compatibility results. (**a**) Release properties of Grease A against the recorded acceptance range (shaded band) or reference value (dashed line). (**b**) Dropping point and (**c**) oil separation of Grease A, the residual grease and their static blends versus the residual-grease mass fraction, with the Grease A: residual grease mass ratio in parentheses; the dashed line in (**c**) is the recorded oil-separation limit of 5%.

All intermediate mixtures remained visually homogeneous and showed no recorded phase separation. Dropping point decreased progressively from 268 °C for Grease A to 264, 258 and 252 °C for the 75:25, 50:50 and 25:75 blends and 245 °C for the residual-grease reference (Figure 3b). Oil separation increased from 1.2% to 1.5%, 2.1% and 2.8% for the same blends and 3.5% for the residual grease (Figure 3c); every intermediate value remained below the recorded 5% limit. Each blend value lay between the values of the two neat greases. The absence of an abrupt discontinuity supports static compatibility across the tested compositions. The trend nevertheless shows that increasing residual-grease content changed the thermal and bleed characteristics, so the result should be read as a pass condition rather than equivalence between old and new grease.

3.3. Operating-Normalized Thermal Response

Quality control retained 537,306 of 569,104 SCADA rows. The treatment and comparator pre-period residual means were 1.21 and 0.79, respectively, indicating close group-level baselines on the normalized scale (Table 4); these means are close to zero by construction, because each turbine's baseline b_i is subtracted in Equation (1), so baseline comparability is assessed with the baseline levels themselves (Section 2.1). Post-period means increased to 46.05 for treatment and 58.00 for the comparator. The corresponding within-group changes were +44.84 and +57.21 residual units, giving a directional thermal contrast of +12.36 residual units in favour of treatment. At the unit level, the mean within-pair difference in change was 12.36 residual units in favour of treatment (95% bootstrap CI 7.2 to 17.3; exact sign-flip permutation $p = 0.016$; unpaired permutation $p = 0.036$).

Table 4. Operating-normalized SCADA response.

Metric	Treatment	Comparator	Directional Contrast
Pre-period residual mean	1.21	0.79	Baseline levels close
Post-period residual mean	46.05	58.00	Both groups increased
Within-group change	+44.84	+57.21	Treatment increase lower
E_th = change_C − change_T	–	–	12.36 residual units favourable to treatment
Unit-level uncertainty of Eth	–	–	**7.2 to 17.3; sign-flip $p = 0.016$; unpaired $p = 0.036$**
Interpretive status	Descriptive	Descriptive	Not a causal DID estimate

Note: Eth was computed from unrounded group means; recomputing it from the rounded values shown gives 12.37.

The group curves also displayed a common recurring pattern and a marked increase in the terminal archive segment (Figure 4). Before the purge, the block-level difference between the treatment and comparator curves varied between approximately −7.5 and +5.9 residual units, and in the last pre-intervention block (B7) the treatment curve already lay about 7.5 residual units below the comparator curve. After the purge, the difference widened to approximately −12 and −17 residual units in B8 and B9. The post-intervention separation therefore exceeds the pre-intervention fluctuation only moderately, and part of the divergence began before the purge boundary. The smaller treatment increase is consistent with reduced thermal sensitivity after relubrication, but the shared movement indicates that group-independent operating or archive effects remained. The result therefore passes a directional screening criterion but does not establish a stable post-treatment temperature reduction. Its value lies in showing that the favourable difference persists after explicit adjustment for ambient temperature and load, while its limitation lies in the short temporal support beyond the purge boundary.

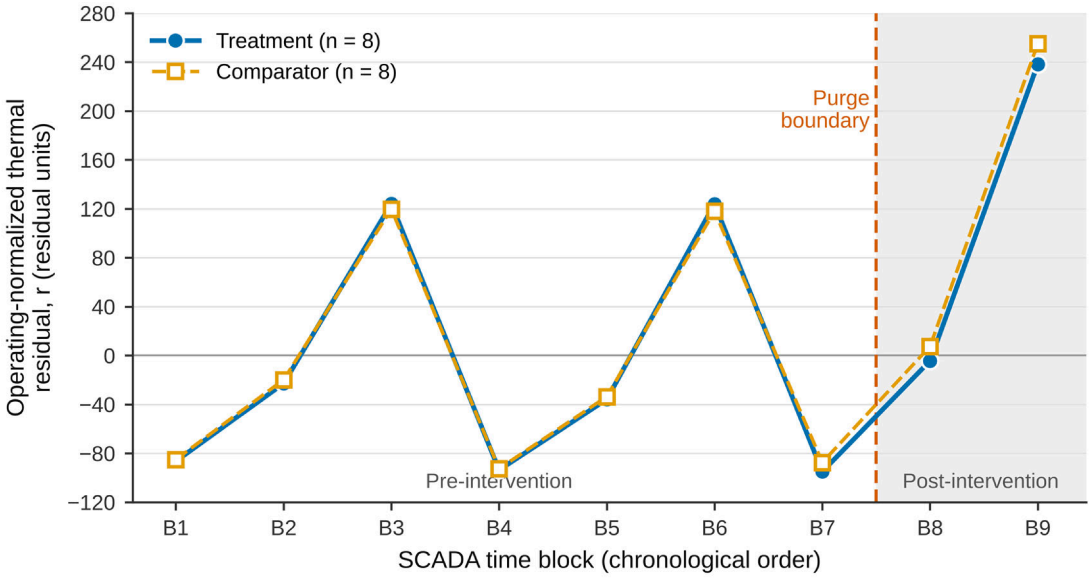

Figure 4. Operating-condition-normalized main-bearing thermal residual. Points are group means of the residual r (Equation (1); residual units) in nine consecutive time blocks (B1–B9) of the quality-controlled SCADA archive in chronological order; the dashed line marks the purge boundary between B7 and B8, and the shaded area the post-intervention period (treatment, n = 8; comparator, n = 8).

3.4. Debris Marker and Longitudinal Response

At D0, mean grease-iron concentrations were 132.2 mg/kg in treatment units and 135.7 mg/kg in comparator units (Table 5; Figure 5a). Early D1/D2 means were 146.4 mg/kg (n = 7) and 145.0 mg/kg (n = 7). Both group means therefore increased between D0 and D1/D2, by 14.2 mg/kg in the treatment group and 9.3 mg/kg in the comparator group (the D1/D2 means exclude one unit per group), so the early increase was not specific to the purged units. Because dilution by the fresh charge should be greatest immediately after the purge, this early rise indicates that any dilution was outweighed at that stage by mobilized particles or by a common sampling effect; T01 and T02 had no D1/D2 sample. The similar early group means do not indicate a treatment-related reduction at that stage. A short-term increase after purging is also compatible with the redistribution or mobilization of historical particles, which is why the early samples were not used as evidence of a new wear episode without supporting trajectories.

Table 5. Grease iron response and relative-change contrast.

Unit/Group	D0 (mg/kg)	D1/D2	D30	D60	Relative Change in Concentration (D60 vs. D0)
Treatment mean (n = 8)	132.2	146.4 (n = 7)	Not available	Not available	Not estimable
Comparator mean (n = 8)	135.7	145.0 (n = 7)	Not available	Not available	Not estimable
T01 treatment	124.7	N/A	96.2	99.0	−20.6%
T02 comparator	119.4	N/A	126.6	121.5	+1.8%
Complete-pair E_{Fe}	–	–	–	–	+22.4 percentage points

Note: Relative change = $(Fe_{D60} - Fe_{D0})/Fe_{D0}$; a negative value denotes a decrease in concentration. E_{Fe} (Equation (3)) is the difference between the proportional reductions in T01 (20.6%) and T02 (−1.8%), i.e., 22.4 percentage points.

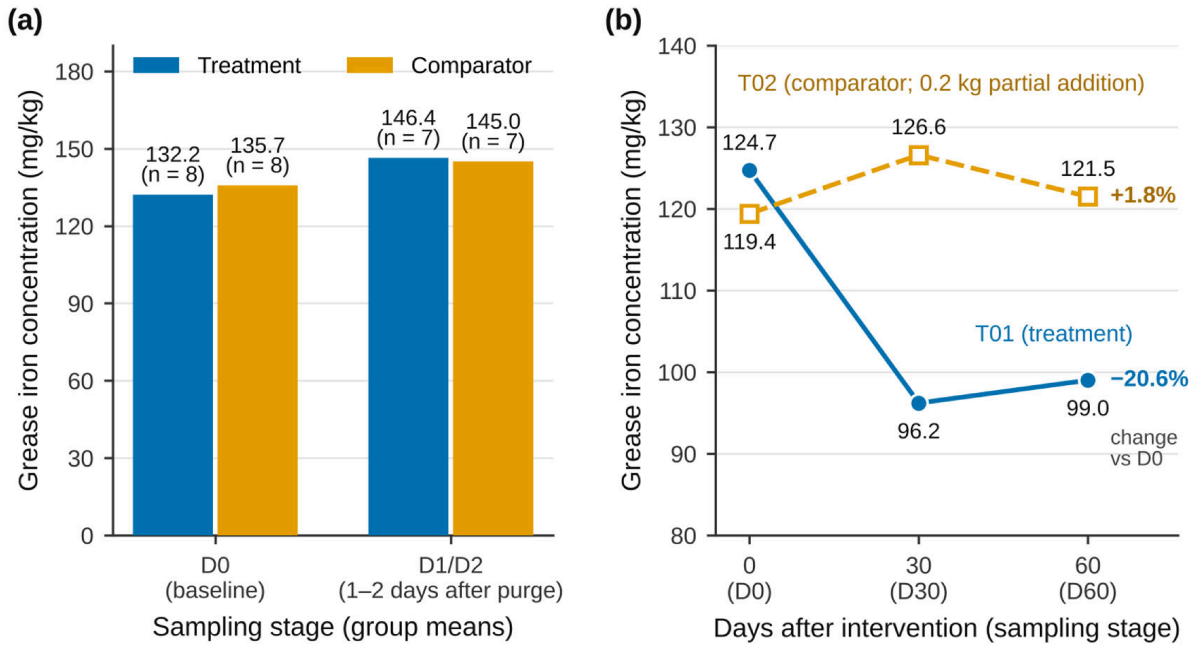

Figure 5. Group-level and unit-level responses of grease iron concentration. (**a**) Group means at D0 (n = 8 per group) and D1/D2 (n = 7 per group); the *y*-axis starts at zero. (**b**) Complete trajectories of T01 (treatment) and T02 (comparator, which received a later 0.2 kg partial addition) at nominal days 0, 30 and 60 after the intervention; percentages are changes in concentration relative to D0.

The complete treatment trajectory, T01, declined from 124.7 mg/kg at D0 to 96.2 mg/kg at D30 and 99.0 mg/kg at D60 (Figure 5b). The final concentration was 20.6% below baseline. T02 changed from 119.4 to 126.6 and 121.5 mg/kg, so that its D60 concentration was 1.8% above baseline (a proportional reduction of −1.8% in the convention of Equation (3)). The difference between proportional reductions was 22.4 percentage points in favour of T01. This contrast is directionally consistent with the removal of mobile debris and subsequent stabilization, but one complete pair cannot estimate a treatment-group distribution or distinguish dilution from a sustained change in particle generation. Under the dilution reference defined in Section 2.3, the 20.6% decline in T01 corresponds to an apparent fresh-grease fraction of about one-fifth in the sampled grease. Whether such a fraction is plausible depends on the 3.2 kg charge relative to the grease capacity of the bearing (Section 2.1) and on the sampling point, so dilution alone could account for the decline, but this can be neither confirmed nor excluded with the available data. The small change between D30 and D60 shows that the lower level persisted over the follow-up, not that it would persist beyond D60. Because a single trajectory has no within-group replication, no confidence interval or significance test can be attached to the 20.6% decline or to E_{Fe}; for scale, the comparator trajectory deviated from its own baseline by at most 6.0% over the same stages.

3.5. CMS, Inspection and Engineering Outcomes

3.5.1. CMS Waveform Quality: A Data Limitation

Of the 64 indexed CMS waveforms, 61 (95.3%) reached the int16 limit and three did not. The high saturation prevalence prevented a reliable comparison of amplitude-dependent vibration features, and three unclipped waveforms are too few to support a between-group comparison. This negative analytical result is important: treating clipped records as ordinary waveforms would have produced apparently precise but physically distorted RMS and spectral estimates, and bearing diagnostics depend on precisely these amplitude, envelope and spectral features [37]. The clipping is therefore a limitation of the

acquired data rather than a finding about either group: the waveforms provide no evidence for or against a vibration effect of the intervention. CMS was therefore retained as a record of coverage, quality limitation and alarm occurrence.

3.5.2. Alarm and Inspection Outcomes: Descriptive Results

The alarm and inspection outcomes below were obtained from the alarm logs and inspection records rather than from the stored waveform files. Because CMS alarms are generated within the same acquisition chain, saturation could also affect the alarm features if they are computed from the saturated signals as the CMS alarm features use the same channels and measurement range as the stored waveforms. Treatment units recorded no CMS alarms and no SCADA-temperature alarms. Comparator units recorded 10 CMS alarms and 15 temperature alarms, equivalent to 1.25 and 1.875 events per unit (exact Poisson 95% CI 0.60–2.30 and 1.05–3.09 events per unit, compared with 0–0.46 for the treatment group; Table 6; Figure 6b). Because 15 temperature alarms occurred in eight comparator units, at least some comparator bearings alarmed repeatedly. Five and six of the eight comparator units recorded at least one CMS alarm and one temperature alarm, respectively, against none of the eight treatment units; the corresponding unit-level exact tests are given in Table 6. The event-level exact conditional tests ($p = 0.002$ for CMS alarms and $p < 0.001$ for temperature alarms) treat repeated alarms as independent and are therefore anti-conservative; they are reported for completeness only. Neither group experienced a main-bearing shutdown or replacement.

Table 6. Confidence intervals and exact tests.

Outcome	Treatment (n = 8)	Comparator (n = 8)	Directional Difference; 95% CIs and Exact Tests
CMS alarms	0	10	1.25 fewer per unit; rates 0 (95% CI 0–0.46) vs. 1.25 (0.60–2.30); $p = 0.002$ [a]
SCADA-temperature alarms	0	15	1.875 fewer per unit; rates 0 (0–0.46) vs. 1.875 (1.05–3.09); $p < 0.001$ [a]
Units with ≥1 CMS alarm	0/8	**5**/8	Fisher's exact $p =$ **0.026**; paired McNemar $p =$ **0.0625** [b]
Units with ≥1 temperature alarm	0/8	**6**/8	Fisher's exact $p =$ **0.007**; paired McNemar $p =$ **0.0313** [b]
Main-bearing shutdowns	0	0	No between-group difference
Relatively normal inspection	7/8 (87.5%)	2/8 (25.0%)	+62.5 percentage points (95% CI 13.9 to 83.1); Fisher's exact $p = 0.041$; paired McNemar $p = 0.0625$ or 0.125 [c]
Abnormal inspection	1/8 (12.5%)	6/8 (75.0%)	−62.5 percentage points

[a] Exact conditional test for two Poisson counts with equal exposure; repeated alarms from the same bearing are not independent, so this event-level test is anti-conservative. [b] Unit-level comparison, robust to repeated alarms within a bearing; Fisher's exact test ignores and the exact McNemar test respects the pairing. [c] Post-only inspection; the difference describes end-of-period condition and cannot be attributed to the intervention. Fisher's exact test and the Newcombe interval ignore the pairing; the exact McNemar p is 0.0625 or 0.125, depending on whether the abnormal treatment unit was paired with an abnormal or a relatively normal comparator. Rate intervals are exact Poisson 95% CIs; the interval for the difference in proportions is the Newcombe hybrid-score 95% CI.

Inspection classified seven of eight treatment units as relatively normal, compared with two of eight comparator units; the remaining comparator findings included wear, pitting, indentation, residue or early fatigue spalling. The difference in proportions was 62.5 percentage points (95% CI 13.9 to 83.1; Fisher's exact $p = 0.041$; Figure 6a), but these unpaired statistics ignore the matched design. The exact McNemar test, which respects the pairing, gives $p = 0.0625$ if the abnormal treatment unit was paired with an abnormal comparator and $p = 0.125$ if it was paired with a relatively normal one; the abnormal treatment unit was paired with an abnormal comparator, giving $p = 0.0625$; the paired comparison therefore does not reach $p < 0.05$. Because inspection was performed only after the intervention, this difference describes the end-of-period condition of the two groups and cannot be attributed to the purge; the groups may already have differed in surface condition before the intervention. The alarm and inspection directions agree, although the post-only inspection design cannot establish whether all between-group physical differences arose after the intervention.

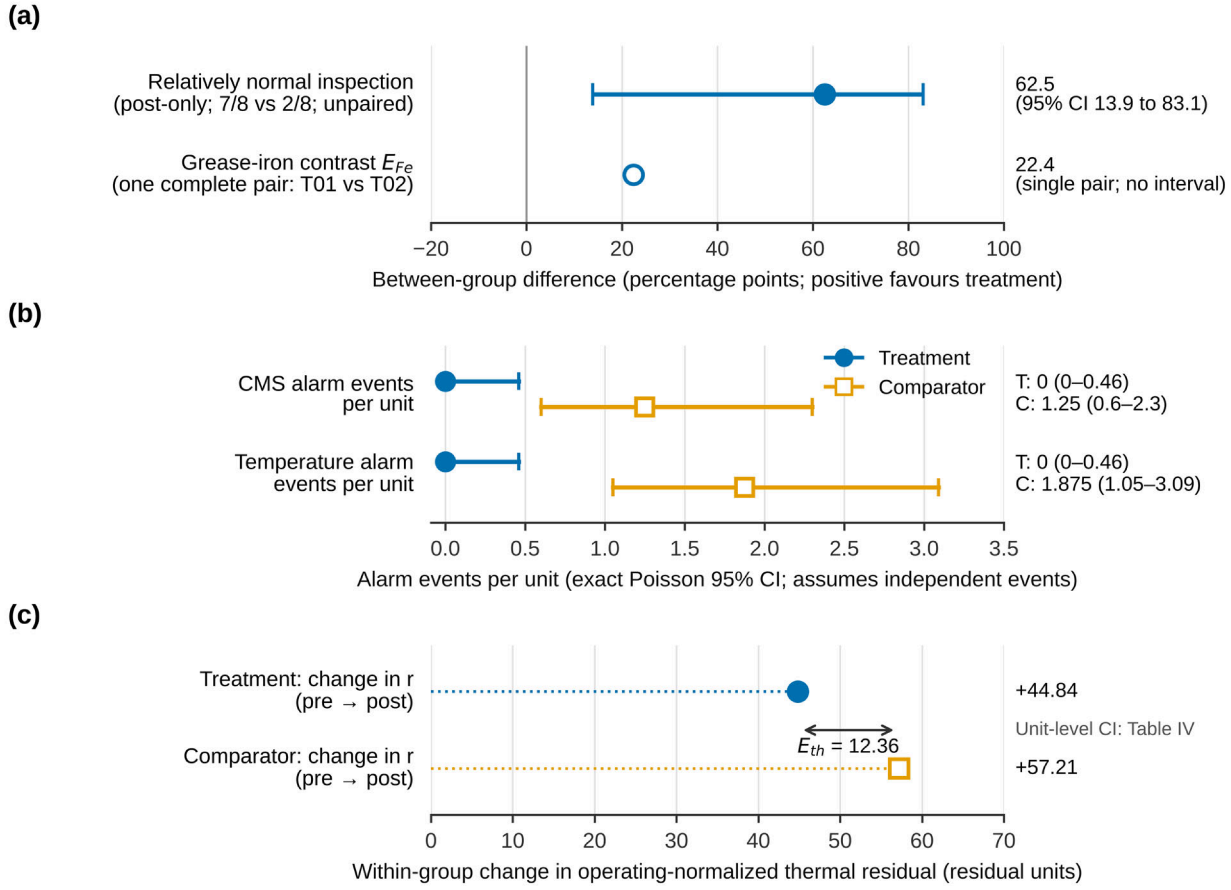

Figure 6. Directional effectiveness profile with the uncertainty that can be quantified from the available data. (**a**) Between-group differences in proportions: the inspection difference with its Newcombe hybrid-score 95% CI (unpaired analysis; the paired exact McNemar test gives $p \geq 0.0625$), and the iron contrast E_{Fe}, which is based on one complete pair and has no interval (open symbol). (**b**) Alarm events per unit with exact Poisson 95% CIs, which assume independent events and are therefore optimistic when repeated alarms occur in the same bearing. (**c**) Within-group change in the operating-normalized thermal residual from the pre- to the post-intervention period; the contrast *Eth* is descriptive, and its unit-level uncertainty is reported in Section 3.3 and Table 7.

Table 7. Integrated effectiveness profile.

Domain	Indicator	Observed Result	Evidence Status and Limits
Cleaning	Purge mass/injected mass	0.78–1.06 across treatment units	Complete for intervention; documents exposure, not complete grease renewal
Lubrication	Static mixture compatibility	No phase separation in 25:75–75:25 mixtures	Material gate; static screen only
Debris	E_Fe	+22.4 percentage points in one complete pair	Case-level (one pair); compatible with dilution; no interval
Thermal	E_th	+12.36 residual units favourable to treatment	Descriptive; short post-period and shared terminal rise; unit-level CI: **7.2 to 17.3**
Monitoring	CMS and temperature alarms	0 vs. 10; 0 vs. 15 events	Engineering signal; threshold-dependent; repeated alarms within units
Inspection	Normal-inspection fraction	87.5% vs. 25.0% (unpaired Fisher $p = 0.041$; paired McNemar $p \geq 0.0625$)	Post-only; baseline condition not inspected
Vibration	CMS waveform features	61 of 64 waveforms clipped	Not interpretable; no vibration effect size

3.6. Directional Effectiveness Profile

Table 7 and Figure 6 summarize the direction, size and evidential status of each indicator. Two properties must be distinguished: directional consistency, meaning that each indicator moves in the direction expected if the intervention were effective, and demonstrated effectiveness, meaning that a between-group difference has been shown to be caused by the intervention. The present data establish the first but not the second. The integrated profile was favourable to treatment in every interpretable outcome domain, but evidence completeness varied materially. Purge exposure was complete and directly observed. The lubricant gate was complete for the static measurements. The iron contrast was favourable but limited to one complete pair and is compatible with dilution alone. The thermal contrast was based on a dense panel but had a short aligned post-intervention period, a shared terminal rise and pre-intervention block differences in comparable order. Alarm counts and inspection proportions favoured treatment; the inspection difference had a 95% CI excluding zero in the unpaired analysis, but the paired McNemar test did not reach $p < 0.05$ and the inspection is post-only, and the event-level alarm comparison is anti-conservative because repeated alarms are clustered within bearings, whereas CMS waveform saturation prevented any vibration inference.

This pattern is consistent with a state-restoration interpretation because the material, thermal and engineering observations do not contradict one another. However, agreement in direction among indicators that are individually descriptive, post-only, quality-limited or based on a single pair does not by itself demonstrate treatment effectiveness, because a shared bias could move several indicators in the same direction. For example, if the comparator bearings had more pre-existing surface damage than the treatment bearings, they would be expected to show more alarms, more abnormal-inspection findings and a larger thermal increase even without any effect of the purge. The profile also does not justify conversion into a life-extension percentage. A composite value would conceal the fact that the strongest evidence concerns intervention execution and static compatibility, while the evidence most directly related to long-term wear and failure remains temporally or analytically limited.

4. Discussion

4.1. Observed Findings and Their Evidential Weight

The central result is convergence in direction rather than a single optimum metric. The purge records establish that the treatment produced substantial material displacement. The compatibility series excludes an observed static incompatibility across the expected transition compositions. The complete treatment trajectory shows a lower final iron concentration, the operating-normalized thermal contrast favours treatment, and the alarm and inspection outcomes point in the same direction. Because these observations arise from different measurement systems, their agreement is more informative than repetition of one correlated signal, but the strength of the evidence differs sharply between domains (Table 7). The evidence chain is nevertheless only as strong as its least complete link, which in this study is the longitudinal and vibration coverage.

To keep observations, mechanisms and conclusions separate, the findings can be grouped into three tiers. Directly observed and complete: every treated unit received the specified charge and expelled a comparable mass of grease, and Grease A met its release criteria and blended with the residual grease without static incompatibility. Observed but descriptive or limited: the iron decline in T01 (one pair, compatible with dilution), the smaller thermal increase in the treatment group (short post-intervention period and shared terminal rise), the absence of alarms in the treatment group (event counts clustered within bearings) and the higher proportion of relatively normal inspections (post-only; not significant in the paired analysis). Not observed and therefore not concluded: a reduced wear rate, reversal of surface damage, a lower failure probability and an extended bearing life. The mechanisms discussed in Section 4.2 are proposed explanations for the second tier, not additional evidence.

4.2. Proposed Mechanisms and Their Limits

The proposed purge mechanism is consistent with the tribological literature on oscillatory and heavily loaded rolling contacts [7,10,11]. Resident grease can carry oxidized material and hard particles through the contact, while starvation and repeated reversal restrict lubricant replenishment. Introducing fresh grease changes both the contamination inventory and the oil-thickener structure available to the loaded zone. This mechanism explains why a purge may improve the observable state without reversing existing surface damage. T07 illustrates the distinction: historical metallic debris can remain detectable even when outlet condition and purge exposure indicate that the maintenance procedure itself was executed correctly. A recent laboratory and field study of a rescue grease introduced into micropitted wind turbine main bearings likewise reported suppressed damage progression together with lower bearing temperature and vibration [38], supporting the

plausibility of a lubricant-mediated change in bearing state; that study, however, concerned a different grease and intervention, and its findings cannot substitute for direct evidence in the present units.

The static compatibility result should be interpreted within the same mechanistic boundary. The monotonic decline in dropping point and increase in oil separation as residual-grease content increased are physically coherent with ageing of the residual grease, which alter consistency, increase oil separation [25,39], and contain no abrupt mixture anomaly. They support the absence of immediate incompatibility under the recorded test. However, grease behaviour in an operating bearing also depends on shear degradation, churning, reflow, contact temperature and oxidation [6,7,9]. The material gate therefore establishes plausibility for relubrication; it does not substitute for the shear- and storage-stability tests of a complete compatibility evaluation [26], an in-contact friction measurement or a long-duration wear test.

4.3. Interpretation of the Debris, Thermal, Vibration and Inspection Evidence

Grease iron provides a useful but non-unique marker. A decline after cleaning can result from removal of contaminated grease, dilution by the new charge, reduced generation of particles, or a combination of these mechanisms. The early group means in this study did not decrease and were consistent with short-term redistribution. The later T01 decline is more informative than the early group means because it persisted through D60, yet its magnitude corresponds to an apparent fresh-grease fraction of about one-fifth in the sampled grease, and the absence of complete trajectories for the remaining units prevents the separation of purge dilution from treatment-wide changes in wear generation. Reporting the 22.4-percentage-point contrast as a case result preserves the observation without inflating its inferential scope. The comparator of this pair, T02, also received a 0.2 kg partial addition during follow-up, so the contrast compares a purged unit with a partially relubricated comparator rather than an untreated one. The direction of the resulting bias on the iron contrast is uncertain, because dilution would lower and mobilization would raise the T02 concentration, whereas any lubrication benefit of the addition would bias the thermal and alarm contrasts towards zero.

The damage modes that a purge can influence depend on the bearing type and internal geometry. Main bearings are commonly spherical, tapered or cylindrical roller bearings in three-point or four-point drivetrain suspensions, and their failure modes and monitoring techniques have been reviewed [2,40]. In addition to subsurface-initiated contact fatigue, roller bearings suffer surface-initiated damage from debris indentation [41], and wind turbine bearings also show adhesive wear, which appears as smearing where rollers slide under light load or rapid acceleration and abrasive wear by hard particles and contaminants carried in the lubricant [42]. In tapered and cylindrical roller bearings, the roller-end/flange contacts operate with substantial sliding and are particularly sensitive to lubricant supply, and the roller–cage contacts generate wear particles whose composition depends on the cage material. These modes produce particles of different size and composition: fine particles from abrasive and adhesive wear are captured by ICP-based iron measurement, whereas large fatigue flakes may be under-represented unless the sample is fully digested [27,28], and wear of brass or polymer cages does not appear in iron at all but could be tracked through copper and zinc or filler elements. The abnormal-inspection findings (wear, pitting, indentation, residue and early fatigue spalling) span these modes. This is consistent with the intended action of the purge, which is to remove the hard contaminants and degraded grease that drive abrasive wear and debris denting, but a purge cannot repair an existing indentation or spall. A more complete debris assessment in future studies should therefore

combine iron with cage-related elements and particle-size information, for example, from analytical ferrography.

The thermal analysis addresses a different part of the mechanism. Raw bearing temperature would have been an unsuitable endpoint because power and ambient conditions varied throughout the archive. Conditioning the temperature difference on power fraction and subtracting a turbine-specific baseline follows the logic of normal-behaviour modelling in the SCADA literature [13–16,29]. The treatment group showed a smaller increase, but both groups moved strongly in the same direction near the archive boundary, and the treatment curve already lay below the comparator curve in the last pre-intervention block, so part of the contrast may predate the purge. That shared movement indicates remaining temporal structure and makes a stable causal slope impossible to identify. The normalized result is therefore useful as supporting evidence, not as an isolated estimate of frictional heat reduction.

The CMS limitation is substantive rather than procedural. Saturation at the int16 boundary changes the waveform shape, clips peaks and creates harmonics that are not generated by the bearing. A comparison of RMS, crest factor or spectral bands from those files would confound machine condition with acquisition limits. Published SCADA and machine-learning studies demonstrate that sophisticated algorithms can exploit multivariate structure [13–16,22,29], but no model can recover the original amplitude distribution from pervasive unverified clipping without additional information. Excluding an amplitude effect size was therefore necessary to maintain physical interpretability.

Alarm and inspection outcomes provide an engineering cross-check, but they also require restraint. Zero treatment alarms versus 10 CMS and 15 temperature alarms in the comparator set is operationally relevant, and the inspection difference of 62.5 percentage points is large. Nevertheless, alarms are affected by thresholds and observation exposure, while post-only inspection cannot establish pre-intervention equivalence of physical damage. If the comparator bearings had entered the study with more surface damage, they would be expected to alarm more often and to be classified as abnormal more often regardless of the purge. The absence of shutdown or replacement in both groups further means that the study evaluates state indicators rather than hard failure events. These outcomes reinforce the favourable direction but do not independently prove reduced hazard.

4.4. Sample Size and Generalizability

The sample of 16 bearings in eight non-randomized pairs limits both precision and generalizability. With eight units per group, only large differences can be distinguished from chance and confidence intervals are wide; for example, the 95% CI for the inspection difference extends from 13.9 to 83.1 percentage points, and with eight pairs, an exact paired test can reach $p < 0.05$ only if at least six pairs are discordant, all in the same direction. All units came from a single onshore site in a cold, dry climate with flat terrain, so the numerical results describe this configuration rather than wind turbine main bearings in general. Transfer to other turbines requires attention to bearing type and size, grease product and thickener, relubrication path and outlet design, operating temperature and load spectrum, and the condition of the bearings at the time of intervention, because each of these changes how much resident grease is displaced and which damage modes are present. The evaluation framework is transferable because its indicators are available on most utility-scale turbines; the effect sizes are not, and they should be re-established in larger studies with randomized or stepped allocation, complete repeated sampling and pre- and post-intervention inspection.

4.5. Implications for Field Evaluation and Maintenance Practice

The multi-source framework resolves a common tension between laboratory tribology and field monitoring. Controlled grease and contact studies offer clear mechanisms but limited machine-level external validity. SCADA studies offer long operational records but can detach a residual from the material condition that produced it. Here, the lubricant gate, debris marker, normalized thermal response and engineering outcomes are reported as connected but non-commensurate observations. This structure allows the field intervention to be discussed scientifically without treating every available signal as if it measured the same latent health quantity.

Relative to fault detection studies, the contribution lies in evaluating an intervention rather than identifying a fault. Relative to grease and bearing-contact studies, it lies in tracing a material action into a field response. The resulting argument is deliberately conditional: the executed purge is credible, the replacement grease is statically compatible, and the observed machine state is favourable to treatment. The study does not infer a life-modification factor because the duration, event count and longitudinal laboratory coverage do not support that calculation. This distinction is essential when evidence is intended to guide maintenance rather than merely rank models.

For practice, the results support the verification of bearing cleaning through multiple independent observations. Mass displacement should establish exposure; mixture behaviour should exclude immediate incompatibility; debris markers should be interpreted longitudinally; thermal channels should be conditioned on operating state; and alarm or inspection outcomes should be used as external checks. In the present dataset, this approach shows that the intervention was executed as intended and was followed by favourable indicators; whether it acts as a state-restoration measure remains to be confirmed by longer and complete observations. Decisions about service-life extension, optimized grease interval or failure-rate reduction remain outside the identifiable scope of the evidence and should not be derived from the favourable profile alone.

5. Limitations

The intervention was located close to the end of the available SCADA archive. The aligned post-intervention interval was therefore short relative to the recurring operating pattern, preventing a stable trend estimate and a causal difference-in-differences interpretation. Likewise, the 60-day grease follow-up is too short to establish a long-term reduction in wear or an extension of bearing life; such claims would require observation over at least one complete relubrication interval.

Complete D30/D60 grease-iron trajectories were available only for T01 and T02. The 22.4-percentage-point contrast is a case-level observation and cannot characterize between-unit variability or a treatment-group mean effect, and the decline in T01 is compatible with dilution of the sampled grease by the fresh charge.

Sixty-one of 64 CMS waveforms reached the int16 limit. Without validation of acquisition range and unclipped source signals, amplitude- and spectrum-dependent vibration features could not be used as quantitative endpoints; future acquisitions should verify sensor range and gain settings, and screen for saturation, before waveform analysis [30].

The comparator allocation was not independently documented as randomized, T02 received a partial addition, and physical inspection was available only after the intervention. Residual baseline imbalance and time-varying confounding therefore remain possible; in particular, a between-group difference in pre-existing surface damage could account for part of the alarm and inspection contrasts.

The compatibility experiment was static and the available record did not contain replicate distributions for inferential testing. The results exclude an observed static in-

compatibility but do not reproduce high-shear bearing operation or establish a lubricant life-modification factor, and the shear- and storage-stability tests of a complete compatibility evaluation were not performed [26].

The study comprised only 16 bearings in eight non-randomized pairs from a single configuration confirmed. The small sample yields wide confidence intervals and low statistical power, and the numerical results cannot be generalized to other bearing types, greases, relubrication systems or operating environments without further evidence (Section 4.4).

Alarm outcomes depend on the configured thresholds and on how repeated activations are logged. Event counts from one bearing are not independent, so event-level tests overstate statistical certainty, and the unit-level comparison has low power with eight units per group.

6. Conclusions

In-service grease-purge cleaning followed by controlled relubrication was associated with a consistently favourable but not conclusive evidence profile in the observed main-bearing units. Every treatment unit received the specified grease mass and expelled a substantial mass of resident grease. Grease A met the recorded release criteria and remained homogeneous in the tested transition mixtures. The only complete treatment iron trajectory ended 20.6% below baseline, whereas its paired comparator ended 1.8% above baseline, the normalized thermal increase was 12.36 residual units smaller than in the comparator group, and treatment units had fewer alarms and a higher proportion of relatively normal inspections.

The findings are consistent with a mechanistic interpretation in which removal of accessible contaminated grease and restoration of the lubricant supply were followed by favourable material and machine-state observations, but they do not establish a causal treatment effect. The favourable direction is consistent across independently measured domains; this consistency strengthens the case for further evaluation but does not substitute for a controlled comparison, because the debris evidence rests on one complete pair whose decline is compatible with dilution, the post-intervention SCADA period is short and both groups rose sharply at its end, 61 of 64 CMS waveforms were clipped, and no pre-intervention inspection was available. The evidence is strongest for intervention fidelity and static compatibility, and weaker for longitudinal debris response, vibration amplitude and causal thermal estimation.

Cleaning–relubrication can therefore be regarded as a technically credible maintenance measure for the evaluated units: it was executed as intended and was followed by favourable indicators, but its state-restoring benefit is plausible rather than demonstrated. Claims of restored bearing state, reduced wear or improved bearing life should be treated as tentative interpretations rather than demonstrated outcomes. The present evidence does not identify a numerical extension of bearing life, an optimized relubrication interval or a reduction in failure probability. Those claims would require longer aligned observation, complete repeated grease sampling in all units, pre- and post-intervention inspection, validated CMS acquisition and event-based outcome coverage, preferably in a larger randomized or stepped design.

Author Contributions: Conceptualization, Y.T., Q.J., J.L., T.W., S.Y. and W.K.; Methodology, Y.T., Q.J., J.L., T.W. and S.Y.; Investigation, Y.T., Q.J., J.L., T.W., S.Y. and W.K.; Data curation, Y.T., Q.J., J.L., T.W. and S.Y.; Writing—original draft, Y.T., Q.J., J.L., T.W., S.Y. and W.K.; Writing—review & editing, Y.T., Q.J., J.L., T.W., S.Y. and W.K.; Visualization, Y.T., Q.J., J.L., T.W., S.Y. and W.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to confidentiality agreements with the wind farm operator.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Hart, E.; Turnbull, A.; Feuchtwang, J.; McMillan, D.; Golysheva, E.; Elliott, R. Wind turbine main-bearing loading and wind field characteristics. *Wind Energy* **2019**, *22*, 1534–1547. [CrossRef]
2. Hart, E.; Clarke, B.; Nicholas, G.; Kazemi Amiri, A.; Stirling, J.; Carroll, J.; Dwyer-Joyce, R.; McDonald, A.; Long, H. A review of wind turbine main bearings: Design, operation, modelling, damage mechanisms and fault detection. *Wind Energy Sci.* **2020**, *5*, 105–124. [CrossRef]
3. Ishihara, T.; Wang, S.; Kikuchi, Y. Fatigue prediction of wind turbine main bearing based on field measurement and three-dimensional elastic drivetrain model. *Eng. Fail. Anal.* **2025**, *167*, 108985. [CrossRef]
4. Hart, E.; de Mello, E.; Dwyer-Joyce, R. Wind turbine main-bearing lubrication—Part 1: An introductory review of elastohydrodynamic lubrication theory. *Wind Energy Sci.* **2022**, *7*, 1021–1042. [CrossRef]
5. Hart, E.; de Mello, E.; Dwyer-Joyce, R. Wind turbine main-bearing lubrication—Part 2: Simulation-based results for a double-row spherical roller main bearing in a 1.5 MW wind turbine. *Wind Energy Sci.* **2022**, *7*, 1533–1550. [CrossRef]
6. Lugt, P.M. Modern advancements in lubricating grease technology. *Tribol. Int.* **2016**, *97*, 467–477. [CrossRef]
7. Schwack, F.; Bader, N.; Leckner, J.; Demaille, C.; Poll, G. A study of grease lubricants under wind turbine pitch bearing conditions. *Wear* **2020**, *454*, 203335. [CrossRef]
8. Kenworthy, J.; Hart, E.; Stirling, J.; Stock, A.; Keller, J.; Guo, Y.; Brasseur, J.; Evans, R. Wind turbine main bearing rating lives as determined by IEC 61400-1 and ISO 281: A critical review and exploratory case study. *Wind Energy* **2024**, *27*, 179–197. [CrossRef]
9. Wu, Q.; Wu, P.; Zhang, L.; Zhao, S.; Wu, M. Research on the Analysis Method of the Life Modification Factor for Grease-Lubricated Wind Turbine Pitch Bearings. *Lubricants* **2026**, *14*, 108. [CrossRef]
10. Schwack, F.; Halmos, F.; Stammler, M.; Poll, G.; Glavatskih, S. Wear in wind turbine pitch bearings—A comparative design study. *Wind Energy* **2022**, *25*, 700–718. [CrossRef]
11. Hurtado Molina, E.; Long, H. Fretting wear in raceway-ball contact under wind turbine pitch bearing operating conditions. *Wear* **2025**, *564–565*, 205696. [CrossRef]
12. Gbashi, S.M.; Olatunji, O.O.; Adedeji, P.A.; Madushele, N. Development and evaluation of a laboratory-scale test rig prototype for full characterization of wind turbine main bearing dynamics and condition monitoring. *Mech. Syst. Signal Process.* **2025**, *238*, 113205. [CrossRef]
13. Beretta, M.; Julian, A.; Sepulveda, J.; Cusidó, J.; Porro, O. An ensemble learning solution for predictive maintenance of wind turbines main bearing. *Sensors* **2021**, *21*, 1512. [CrossRef] [PubMed]
14. Encalada-Dávila, Á.; Puruncajas, B.; Tutivén, C.; Vidal, Y. Wind turbine main bearing fault prognosis based solely on SCADA data. *Sensors* **2021**, *21*, 2228. [CrossRef] [PubMed]
15. Tutivén, C.; Vidal, Y.; Insuasty, A.; Campoverde-Vilela, L.; Achicanoy, W. Early fault diagnosis strategy for WT main bearings based on SCADA data and one-class SVM. *Energies* **2022**, *15*, 4381. [CrossRef]
16. Xiao, X.; Liu, J.; Liu, D.; Tang, Y.; Zhang, F. Condition monitoring of wind turbine main bearing based on multivariate time series forecasting. *Energies* **2022**, *15*, 1951. [CrossRef]
17. Campoverde, L.; Tutivén, C.; Vidal, Y.; Benalcázar-Parra, C. SCADA data-driven wind turbine main bearing fault prognosis based on principal component analysis. *J. Phys. Conf. Ser.* **2022**, *2265*, 032107. [CrossRef]
18. Encalada-Dávila, Á.; Moyón, L.; Tutivén, C.; Puruncajas, B.; Vidal, Y. Early fault detection in the main bearing of wind turbines based on gated recurrent unit (GRU) neural networks and SCADA data. *IEEE/ASME Trans. Mechatron.* **2022**, *27*, 5583–5593. [CrossRef]
19. Tutivén, C.; Encalada-Dávila, Á.; Vidal, Y.; Benalcázar-Parra, C. Detecting bearing failures in wind energy parks: A main bearing early damage detection method using SCADA data and a convolutional autoencoder. *Energy Sci. Eng.* **2023**, *11*, 1395–1411. [CrossRef]
20. Yucesan, Y.A.; Viana, F.A.C. Wind turbine main bearing fatigue life estimation with physics-informed neural networks. *Proc. Annu. Conf. PHM Soc.* **2019**, *11*, 1–14. [CrossRef]
21. Mohammad, H.; Vlašic, F.; Žáček, J.; Maya, B.; Mazal, P. Using acoustic emission for condition monitoring of the main shaft bearings in 4-point suspension wind turbine drivetrains. *Nondestruct. Test. Eval.* **2024**, *39*, 2108–2131. [CrossRef]
22. Wiese, B.; Pedersen, N.L.; Nadimi, E.S.; Herp, J. Estimating the remaining power generation of wind turbines—An exploratory study for main bearing failures. *Energies* **2020**, *13*, 3406. [CrossRef]

23. Insuasty, A.; Tutivén, C.; Vidal, Y. SCADA data-driven wind turbine main bearing fault prognosis based on one-class support vector machines. *Renew. Energy Power Qual. J.* **2021**, *19*, 338–343. [CrossRef]
24. Cen, H.; Lugt, P.M.; Morales-Espejel, G.E. On the film thickness of grease-lubricated contacts at low speed. *Tribol. Trans.* **2014**, *57*, 668–678. [CrossRef]
25. Rezasoltani, A.; Khonsari, M.M. On monitoring physical and chemical degradation and life estimation models for lubricating greases. *Lubricants* **2016**, *4*, 34. [CrossRef]
26. ASTM International. *ASTM D6185-24*; Standard Practice for Evaluating Compatibility of Binary Mixtures of Lubricating Greases. ASTM International: West Conshohocken, PA, USA, 2024. [CrossRef]
27. ASTM International. *ASTM D7303-23*; Standard Test Method for Determination of Metals in Lubricating Greases by Inductively Coupled Plasma Atomic Emission Spectrometry. ASTM International: West Conshohocken, PA, USA, 2023. [CrossRef]
28. ASTM International. *ASTM D5185-26*; Standard Test Method for Multielement Determination of Used and Unused Lubricating Oils and Base Oils by Inductively Coupled Plasma Atomic Emission Spectrometry (ICP-AES). ASTM International: West Conshohocken, PA, USA, 2026. [CrossRef]
29. Xiao, X.; Liu, J.; Liu, D.; Tang, Y.; Qin, S.; Zhang, F. A normal behavior-based condition monitoring method for wind turbine main bearing using dual attention mechanism and Bi-LSTM. *Energies* **2022**, *15*, 8462. [CrossRef]
30. Kannan, V.; Dao, D.V.; Li, H. Detection of signal integrity issues in vibration monitoring using one-class support vector machine. *J. Vib. Eng. Technol.* **2024**, *12*, 601–611. [CrossRef]
31. Bertrand, M.; Duflo, E.; Mullainathan, S. How much should we trust differences-in-differences estimates? *Q. J. Econ.* **2004**, *119*, 249–275. [CrossRef]
32. Roth, J.; Sant'Anna, P.H.C.; Bilinski, A.; Poe, J. What's trending in difference-in-differences? A synthesis of the recent econometrics literature. *J. Econom.* **2023**, *235*, 2218–2244. [CrossRef]
33. Agresti, A. A survey of exact inference for contingency tables. *Stat. Sci.* **1992**, *7*, 131–153. [CrossRef]
34. Newcombe, R.G. Interval estimation for the difference between independent proportions: Comparison of eleven methods. *Stat. Med.* **1998**, *17*, 873–890. [CrossRef]
35. Ng, H.K.T.; Gu, K.; Tang, M.L. A comparative study of tests for the difference of two Poisson means. *Comput. Stat. Data Anal.* **2007**, *51*, 3085–3099. [CrossRef]
36. Glynn, R.J.; Buring, J.E. Ways of measuring rates of recurrent events. *BMJ* **1996**, *312*, 364–367. [CrossRef] [PubMed]
37. Randall, R.B.; Antoni, J. Rolling element bearing diagnostics—A tutorial. *Mech. Syst. Signal Process.* **2011**, *25*, 485–520. [CrossRef]
38. Vengudusamy, B.; Thimmidi, P.; Krämer, J.; Püler, T.; Püler, M.; Lupovic, R.; Wunder, T.; Jörgensen, T.; Bogen, O.; Narnhammer, D.; et al. Micropitting mitigation in wind turbine main bearings through laboratory and field testing of an advanced rescue grease. *J. Tribol.* **2026**, *148*, 112601. [CrossRef]
39. Tian, H.; Liu, Y.; Fan, Y.; Wang, G.; Wang, Z. Research on the static thermal degradation law of lubricating grease for wind power bearings. *Lubricants* **2025**, *13*, 134. [CrossRef]
40. Gbashi, S.M.; Madushele, N.; Olatunji, O.O.; Adedeji, P.A.; Jen, T.C. 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; IEEE: New York, NY, USA, 2022; pp. 127–134. [CrossRef]
41. Ai, X. Effect of debris contamination on the fatigue life of roller bearings. *Proc. Inst. Mech. Eng. Part J J. Eng. Tribol.* **2001**, *215*, 563–575. [CrossRef]
42. Peng, H.; Zhang, H.; Shangguan, L.; Fan, Y. Review of tribological failure analysis and lubrication technology research of wind power bearings. *Polymers* **2022**, *14*, 3041. [CrossRef] [PubMed]

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.