Investigation of the Dynamic Relationship between Oil Temperature and Bearing Gearbox Condition Indicator Values for the Bell 407 Helicopter Based on Cointegration Analysis

Oumayma Babay[1], Lotfi Saidi1[1,2,4], Eric Bechhofer[3], Mohamed Benbouzid[4]

[1]*Department of Electronics and Computer Engineering, Higher School of Sciences and Technology of Hammam Sousse, University of Sousse, Sousse 4011, Tunisia*
Oumaymababay2000@gmail.com

[2]*ENSIT- Laboratory of Signal Image and Energy Mastery (SIME), University of Tunis, LR 13ES03, Tunis 1008, Tunisia; lotfi.saidi@ieee.org*

[3]*Green Power Monitoring Systems, LLC, VT 05753, USA*
eric@gpms-vt.com

[4]*University of Brest, UMR CNRS 6027 IRDL, 29238 Brest, France*

mohamed.benbouzid@univ-brest.fr

ABSTRACT

The primary objective of this study is to investigate the dynamic relationship between oil temperature and the Bearing Gearbox Condition Indicator (BGCI) values of the Bell 407 helicopter. To achieve this goal, we employ robust econometric tools, such as unit root tests, cointegration tests, and Autoregressive Distributed Lag (ARDL) models for both, long-run and short-run estimates. Our findings indicate that variable temperature tends to converge to its long-run equilibrium path in response to changes in other variables. The results of the ARDL analysis confirm that spectral kurtosis, inner race, cage, and ball energy significantly contribute to the increase in temperature. Furthermore, we use the impulse response function (IRF) to trace the dynamic response paths of shocks to the system. The identification of a cointegrating relationship between oil temperature and BGCI values suggests a practical and significant connection that can potentially be used to predict hazardous changes in oil temperature using BGCI values, which is an important implication for enhancing the safety and reliability of helicopter operations.

This study presents a promising direction for condition monitoring (CM) in rotating aircraft machinery, emphasizing the potential of integrating temperature data to simplify the diagnostic process while still achieving reliable results.

KEYWORDS: Bearing gearbox; Cointegration; condition monitoring (CM), Condition indicator (CI); Oil temperature

1. INTRODUCTION

The relationship between oil temperature and BGCI values is crucial for assessing the health of helicopter main gearboxes. Monitoring systems like Health Usage Monitoring Systems (HUMS), utilize vibration signatures to detect faults (Tabrizi et al., 2017, Zhang et al., 2012). Oil temperature is a key parameter affecting gearbox health (Huang et al., 2020). Data analysis and artificial intelligence techniques are employed to monitor the lubrication and cooling systems of modern helicopters (Li et al., 2019). Additionally, the vibration signatures of damaged components, like bearings, are used as condition indicators, which can vary based on the system design and operating conditions (Wei et al., 2022) and oil temperature. The use of advanced signal processing tools can help extract bearing fault signatures from vibration signals, thereby enhancing the fault detection capability. Overall, integrating oil temperature monitoring and bearing condition indicators can provide a comprehensive approach for ensuring the safety and reliability of helicopter gearboxes.

In Tabrizi et al. (2017), proposed a novel combined method for fault detection in rolling bearings based on cointegration for the development of fault features that are sensitive to the presence of defects but insensitive to changes in operational conditions.

According to Zhang et al. (2012), the root cause of high oil temperatures is improper lubricating oil selection and serious solid particle pollution. Oil analysis techniques are also applied to monitor the working conditions of oil, to prevent mechanical failure, and to extend machine life.

In (Huang et al., 2020), a new diagnostic method based on modulation signal bispectrum (MSB) analysis is proposed, which shows that the amplitudes coupled between the fault frequency and epicycle carrier frequency can have more deterministic information regarding bearing vibrations.
In (Li et al., 2019), the authors introduce a novel health indicator based on cointegration for rolling bearings' run-to-failure process, which can depict different run-to-failure data in a unified way. Through the cointegration test, the study found a certain degree of a cointegration relationship between energy and complexity features, leading to the development of a novel health indicator. The indicator exhibits "two-stage" characteristics, with a zero-line stage followed by a quickly raising stage resembling an exponential function, making it more suitable for an exponential degradation model compared to root mean square (RMS).
In (Wei et al., 2022), an improved deep deterministic policy gradient algorithm with a Convolutional long-short time memory basic learner is proposed, which can excavate the complex relationship between oil temperature and working conditions and a multi-critics network structure is adopted to solve the problem of inaccurate Q-value estimation.
In (Bayoumi et al., 2012), the effect of oil and grease on component performance and fault detection was examined in four different aircraft wetted-component case studies, which aimed to improve performance by examining the effect that oil and greases have on components.
Murrad et al. (2012) formed a study group and mandated it to improve the maintenance strategy of the S61-A4 helicopter fleet in the Malaysian Armed Forces (MAF). The strategy consisted of a structured approach to reassessment/ redefinition of suitable maintenance actions for the main rotor gearbox.
In (Xu et al., 2021), the influence of the dynamic wear model considering the tooth contact flash temperature on the dynamic characteristics of a gear-bearing system is studied, and the effects of initial wear, friction factors, and damping ratio on the system response are studied.
To overcome the high-temperature bearings problem, advanced materials, such as ceramic matrix composites, are being developed to withstand extreme temperatures without adverse effects, ensuring reliability in harsh conditions (Joseph & Loewe, 1990). While advancements in lubrication systems and materials are promising, the inherent risks of inadequate lubrication remain a critical concern for helicopter safety, necessitating ongoing research and development in this area.
Traditional vibration-based CM for rotating machines, especially those with multiple bearings like turbo-generator sets, is complex and requires extensive data collection. This complexity arises from needing multiple sensors at each bearing location, making fault diagnosis challenging even for experienced analysts. The study aims to simplify the fault diagnosis process by proposing a method that utilizes only one vibration sensor and one temperature sensor per bearing. This approach is intended to reduce the data collection burden while maintaining effective monitoring capabilities.
The present study investigates the dynamic relationship between oil temperature and BGCI in a Bell 407 helicopter. To achieve this goal, we employ robust econometric tools, such as unit root tests, cointegration tests, and ARDL models, for both long- and short-run estimates. We utilized the Dumitrescu Hurlin (2012) panel causality tests to confirm the causal relationships between the variables. In summary, the contributions of this paper lie in its innovative approach to simplifying fault diagnosis, integrating temperature data, applying cointegration tests, and ARDL models for enhanced analysis, providing experimental validation, and offering practical solutions for industrial applications. These contributions collectively advance the field of CM for rotating machinery.
This paper's remaining structure is outlined as follows: Section 2 describes the data collection. Section 3 presents the empirical model. Section 4 demonstrates the relationship results between oil temperature and BGCI Values. Last, Section 5 concludes this paper.

2. HUMS DATA COLLECTION

The HUMS was developed to provide a holistic measure of aircraft health by providing automated flight data monitoring, rotor track and balance, engine performance monitoring, and drivetrain diagnostics/prognostics. The system incorporates ten smart accelerometers to collect and process vibration data into condition indicators. The accessory drive sensor monitors this study's duplex bearing. This bearing supports the accessory drive's hydraulic pump, which operates at 4450 RPM. The sensor's sample rate was 46875 sps for 2 seconds. The sensor performed an envelope (Abboud et al., 2017) using a window from 9 to 13 kHz. The system was designed to acquire data about every four minutes, depending on the aircraft regime. That is, after four minutes, if the aircraft is relatively straight and level, the data is captured. If the aircraft is manoeuvring, the acquisition is delayed until the aircraft is again straight and level.
Figures 1-2 show the oil temperature against BGCI of the Bell 407 helicopter (Fig. 1), and the regression of the CI data based on temperature against ball energy (Fig. 2). Here is important information. The gearbox oil was replaced in 2020, on December 22. It was hypothesized that the oil was contaminated with wear debris from the gearbox, which was causing the ball energy to increase (Fig.2). The CI energy then went down. However, we can see that from April 6, 2021, the fault started to propagate. The bearing was replaced on June 11, 2021.
The other way to look at this is the "peak" energy only seen at the start of a flight, so on the ground. That would be another avenue of research is to look at the regime in which the analysis was taken.

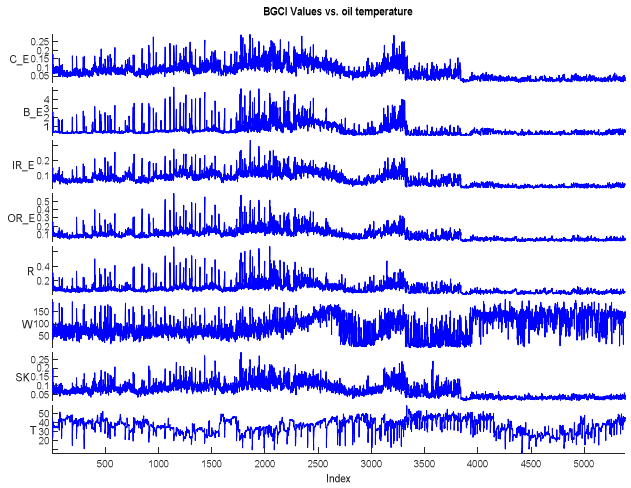

Figure 1. Oil temperature and BGCI values of the Bell 407 helicopter, from top to bottom: Cage Energy (C_E), Ball Energy (B_E), Inner Race Energy (IR_E), Outer Race Energy (OR_E), 1/*Rev* (R), Whip (W), Spectral Kurtosis (SK), and Temperature (T)

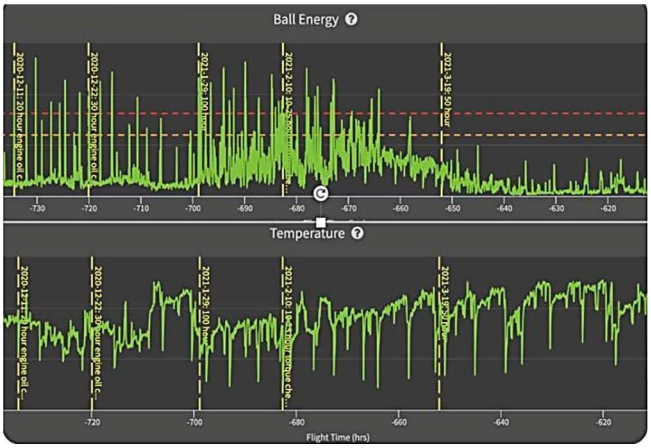

Figure 2. Regression of the CI data based on temperature against ball energy.

3. EMPIRICAL MODEL

This study aims to analyze the dynamic effects of the cage, ball energy, inner race energy, outer race energy, 1/*Rev*, whip, and spectral kurtosis on temperature.

The general form of the empirical equation is indicated below:

$$T_t = f(\mathrm{C}_{Et}, \mathrm{B}_{Et}, \mathrm{IR}_{Et}, \mathrm{OR}_{Et}, \mathrm{R}_t, \mathrm{W}_t, \mathrm{SK}_t) \quad (1)$$

Where T is the temperature, C_E represents cage, B_E denotes ball energy, IR_E represents inner race energy, OR_E represents outer race energy, R represents 1/*Rev*, W represents whip, SK represents spectral kurtosis, and t denotes the acquisition time.

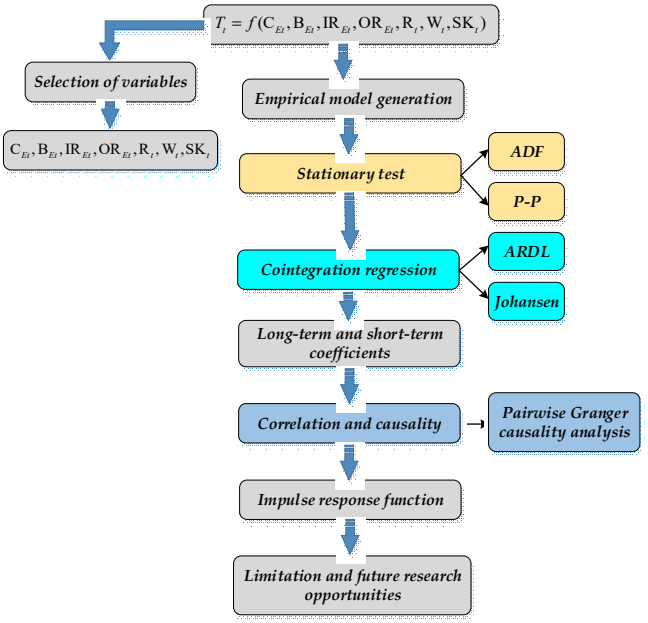

Figure 3. Flowchart of analytical techniques used in the study

To examine the long-run linkages between the variables, we employed the following equation derived from Eq. (1) :

$$\begin{aligned} T_t = {} & \alpha_0 + \alpha_1 \mathrm{C}_{Et} + \alpha_2 \mathrm{B}_{Et} + \alpha_3 \mathrm{IR}_{Et} + \alpha_4 \mathrm{OR}_{Et} \\ & + \alpha_5 \mathrm{R}_t + \alpha_6 \mathrm{W}_t + \alpha_7 \mathrm{SK}_t + \varepsilon_t \end{aligned} \quad (2)$$

The estimated econometric model above is not in linear form, which does not seem to present consistent results and is not helpful in the decision-making process. To address this issue, we transform all variables into natural logarithms to analyze the relationships between the dependent and independent variables. Utilizing a log-linear specification model offers several advantages, aiming to achieve consistent and robust empirical findings (Shafique et al., 2021; Ouni & Ben Abdallah, 2024). Therefore, the log-linear form is given in Eq. (3):

$$\begin{aligned} LnT_t = {} & \alpha_0 + \alpha_1 \mathrm{LnC}_{Et} + \alpha_2 \mathrm{LnB}_{Et} + \alpha_3 \mathrm{LnIR}_{Et} \\ & + \alpha_4 \mathrm{LnOR}_{Et} + \alpha_5 \mathrm{LnR}_t + \alpha_6 \mathrm{LnW}_t \\ & + \alpha_7 \mathrm{LnSK}_t + \varepsilon_t \end{aligned} \quad (3)$$

Where LnT is the natural logarithm of temperature, LnC_{Et} is the natural logarithm of the cage, $Ln\mathrm{B}_{Et}$ is the natural logarithm of ball energy, $LnIR_{Et}$ is the natural logarithm of inner race energy, $LnOR_{Et}$ is the natural logarithm of the outer race energy, LnR_t is the natural logarithm of rev, LnW_t is the natural logarithm of whip, and $LnSK_t$ is the natural logarithm of spectral kurtosis. α_0 represents the fixed effect. α_1, α_2, α_3, α_4, α_5, α_6, and α_7 are the slope coefficients. ε_t represents the white noise.

A cointegration-based computation procedure, consisting of three stages, was developed for this purpose. The

methodology presented in the paper is described in a general manner in Figure 3, with specific steps outlined for the application of econometric tools such as ARDL models, cointegration tests, and IRF. Indeed, in the first stage, a cointegration model of the monitored bearing gearbox is established using a set of condition indicator values. This model has the role of a bearing oil temperature monitoring model. In the second stage, the stationarity test was carried out. In the third stage, the cointegration procedure is deployed. The next stage used the ARDL model to examine the long and short-term relationships between the explanatory variables and the temperature. The correlation and causality tests were carried out in the fourth stage. In the last stage, we employed the IRF to measure the effects of shocks from independent variables on the dependent variable.

3.1. Estimation Procedures

The first step in our analysis is to employ unit root tests to verify the stationarity of all variables. We utilize the standard unit root tests of the Augmented Dickey-Fuller (ADF) test (Dickey and Fuller, 1979) and the Phillips-Perron (PP) test (Phillips and Perron, 1988). These tests indicate the null hypothesis that a series possesses a unit root (indicating non-stationarity), while the alternative hypothesis suggests stationarity. The ADF unit root test primarily assesses the presence of autocorrelation in the series, while the PP unit root test examines the presence of heteroskedasticity in the time series. A time series is considered non-stationary if one or more of its moments (mean, variance, or covariance) are not time-independent. A non-stationary series with a stochastic unit root needs to be differenced once to achieve stationarity. Before exploring cointegration analysis, it is imperative to establish empirically the integration process.

The empirical equation of the ADF unit root test can be represented as follows in Eq. (4):

$$\Delta Y_t = \beta_0 + Y_{t-1} + \sum_{i=1}^{m} d_i \Delta Y_{t-m} + \mu_t \tag{4}$$

Where Δ represents the first difference operator, μ_t denotes the error term, β_0 is the intercept term associated with the equations, m indicates the number of lags of the specific variable in the model, and t signifies the time measure.

The empirical equation of the PP unit root test can be expressed in Eq. (5):

$$\Delta Y_t = \beta + \theta \times Y_{t-1} + \mu_t \tag{5}$$

The long-term relationship between temperature, cage, bull energy, inner race energy, outer race energy, *Rev*, whip, and spectral kurtosis is investigated using the ARDL test and the Johansen and Juselius cointegration test.

The Johansen and Juselius (1990) cointegration approach is employed to examine the long-run relationship among the variables. The Johansen and Juselius cointegration technique is constructed on λ_{trace} and λ_{max} statistics. Trace statistics investigates the null hypothesis of r cointegrating relations against the alternative of N cointegrating relations and is computed as:

$$\lambda_{trace} = N \sum_{i=r+1}^{n} \log(1-\lambda_i) \tag{6}$$

Where N is the number of observations.

The maximum Eigen-value statistics tests the null hypothesis of r cointegrating relations against the :

$$\lambda_{\max} = N \log(1-\lambda_r + 1) \tag{7}$$

The ARDL model introduced by Pesaran et al. (2001) examines the existence of long-run and short-run relationships between the variables under examination. This method offers several advantages over traditional cointegration tests. Firstly, it addresses endogeneity issues by accommodating appropriate variable lag lengths for both independent and dependent (Narayan, 2005). Secondly, it can handle a mixture of stationary variables such as $I(0)$ and $I(1)$, but not $I(2)$ (Pesaran et al., 2001). Thirdly, the ARDL bounds testing approach demonstrates improved efficiency and robustness, effectively mitigating issues related to autocorrelation. Moreover, the ARDL model allows for varying lag lengths, enhancing its flexibility and enabling the estimation of both long-term and short-term dynamics through an error correction model (ECM). The ARDL model is specified as follows:

$$\begin{aligned}\Delta LnT_t = {} & \alpha_0 + \sum_{k=1}^{n} \alpha_{1k} \Delta LnT_{t-k} + \sum_{k=1}^{n} \alpha_{2k} \Delta Ln\,\mathrm{C}_{Et-k} + \sum_{k=1}^{n} \alpha_{3k} \Delta Ln\,\mathrm{B}_{Et-k} \\ & + \sum_{k=1}^{n} \alpha_{4k} \Delta Ln\,\mathrm{IR}_{Et-k} + \sum_{k=1}^{n} \alpha_{5k} \Delta Ln\,\mathrm{OR}_{Et-k} + \sum_{k=1}^{n} \alpha_{6k} \Delta Ln\,\mathrm{R}_{t-k} \\ & + \sum_{k=1}^{n} \alpha_{7k} \Delta Ln\,\mathrm{W}_{t-k} + \sum_{k=1}^{n} \alpha_{8k} \Delta Ln\,\mathrm{SK}_t + \delta_1 LnT_{t-1} + \delta_2 Ln\,\mathrm{C}_{Et-1} \\ & + \delta_3 Ln\,\mathrm{B}_{Et-1} + \delta_4 Ln\,\mathrm{IR}_{Et-1} + \delta_5 LnO\,\mathrm{R}_{Et-1} + \delta_6 Ln\,\mathrm{R}_{t-1} \\ & + \delta_7 LnW_{t-1} + \delta_8 LnSK_{t-1} + \varepsilon_t\end{aligned} \tag{6}$$

After obtaining the long-run coefficients from Eq. (6), we use the ECM described in Eq. (7) to analyze the short-term relationships.

$$\begin{aligned}\Delta LnT_t = {} & \alpha_0 + \sum_{k=1}^{n} \alpha_{1k} \Delta LnT_{t-k} + \sum_{k=1}^{n} \alpha_{2k} \Delta Ln\,\mathrm{C}_{Et-k} \\ & + \sum_{k=1}^{n} \alpha_{3k} \Delta Ln\,\mathrm{B}_{Et-k} + \sum_{k=1}^{n} \alpha_{4k} \Delta Ln\,\mathrm{IR}_{Et-k} \\ & + \sum_{k=1}^{n} \alpha_{5k} \Delta Ln\,\mathrm{OR}_{Et-k} + \sum_{k=1}^{n} \alpha_{6k} \Delta Ln\,\mathrm{R}_{t-k} \\ & + \sum_{k=1}^{n} \alpha_{7k} \Delta Ln\,\mathrm{W}_{t-k} + \sum_{k=1}^{n} \alpha_{8k} \Delta Ln\,\mathrm{SK}_t + \gamma ECM_{t-1} + \varepsilon_t\end{aligned} \tag{7}$$

Where the difference operator is defined by Δ, the optimal lag length of variables is denoted by n, the error correction term ECM_{t-1} coefficient is indicated by γ, and the residual error term is presented by ε_t. Under the ARDL framework, the null hypothesis of no cointegration; $H_0 = \delta_1 = \delta_2 = \delta_3 = \delta_4 = \delta_5 = \delta_6 = \delta_7 = \delta_8 = 0$ against the alternative hypothesis of cointegration $H_1 \neq \delta_1 \neq \delta_2 \neq \delta_3 \neq \delta_4 \neq \delta_5 \neq \delta_6 \neq \delta_7 \neq \delta_8 \neq 0$ is tested by taking the F-statistics used by Narayan (2005).

When the computed F-value exceeds the upper bound, it indicates cointegration; conversely, when the computed F-value is below the lower bound, it suggests no cointegration. However, if the F-value falls between the upper and lower critical values, the decision of cointegration is inconclusive. Additionally, a significant error correction term suggests a long-term relationship between the variables.

Then, we examined the causal relationship between temperature and other variables using the Dumitrescu and Hurlin (2012) test, which is a simplified version of Granger's (1969) non-causality test. We chose this test because it considers two different types of heterogeneity: one in the regression model used to evaluate Granger causality and another in the causality relationship itself. In our analysis, we employed the linear model shown in Eq. (8):

$$Y_{it} = a_i + \sum_{i=1}^{k} \gamma_i^{(k)} Y_{it-k} + \sum_{i=1}^{k} \beta_i^{(k)} X_{it-k} + \varepsilon_{it} \quad (8)$$

Where represents the slope coefficients, ai represents the cross-sectional unit, and K shows the lag length. In this context, the null hypothesis suggests that there is no causal relationship in at least one cross-sectional unit. To test this null hypothesis, we used the Z-bar statistics ($\bar{Z}$) and W-bar statistic ($\bar{W}$) tests, which can be computed as follows:

$$\bar{W} = \frac{1}{N} \sum_{i=1}^{N} W_i \quad (9)$$

$$\bar{Z} = \sqrt{\frac{N}{2K}} (\bar{W} - K) \quad (10)$$

The innovation accounting technique is employed to analyze the forecast relationship between selected variables in a given dataset. This method utilizes the IRF to visually demonstrate how shocks in one variable affect others, altering their magnitudes over time. The IRF highlights the impacts of these shocks on both current and future values of the variables involved. Specifically, a standard error shock in one variable at time period '*t*+*s*' may positively, negatively, or bidirectionally influence another variable '*j*' at period '*t*'. This relationship can be expressed mathematically as given in Eq. (11) :

$$\psi_s = \frac{\varphi Y_{i,t+s}}{\varphi \mu_{j,t}} \quad (11)$$

Where Y represents dependent variables, and μ is the forecast error term.

4. RESULTS

4.1. Unit root analysis

To study the relationship between temperature, cage, ball energy, inner race energy, outer race energy, *Rev*, whip, and spectral kurtosis, we verify the presence of unit roots in the variables using ADF and PP tests. Table 1 shows the outcomes of both the ADF and PP tests. It indicates that all variables (*LnT*, LnC_E, LnB_E, $LnOR_E$, *LnR*, *LnW*, *LnSK*) are stationary at the level, except the variable $LnIR_E$, which is stationary at the first difference. Therefore, some variables are *I*(0), while others are *I*(1). It is concluded that the variables used in this study have a mixed order of integration, as evidenced by both ADF and PP unit root tests.

Table 1. Results of unit root tests

Variables	ADF		PP	
	Level	Δ	Level	Δ
LnT	-8.460***	-32.875***	-17.104***	-29.324***
LnC$_E$	-3.619***	-32.923***	-10.094***	-38.303***
LnB$_E$	-5.725***	-30.881***	-8.831***	-45.706***
LnIR$_E$	-0.703	-35.604***	-1.153	-37.367***
LnOR$_E$	-4.548***	-30.650***	-19.252***	-43.162***
LnR	-5.801***	-30.347***	-12.187***	-29.177***
LnW	-5.146***	-25.347***	-17.309***	-36.076***
LnSK	-3.817***	-34.203***	-12.572***	-31.278***

Note: *** indicates a 1% level of significance, Δ denotes the first difference

4.2. Cointegration tests

After confirming the integration order of the variables, this study employs the ARDL bounds model and the Johansen and Juselius cointegration test to investigate the cointegration among the variables under examination. The results of the ARDL bounds testing are shown in Table 2. When we analyzed the ARDL model with *LnT* as the dependent variable, we found evidence of cointegration in the series under consideration. This conclusion is drawn from observing that both upper and lower-bound critical values are well below the estimated F-statistic (5.534). Thus, we reject the null hypothesis of no cointegration, suggesting the presence of a long-term relationship between the variables. The results of the Johansen cointegration test, as shown in

Table 2: ARDL Bound tests for cointegration

Estimated model	F-statistics	
$LnT_t=f(LnC_{Et}, LnB_{Et}, LnIR_{Et}, LnOR_{Et}, LnR_t, LnW_t, LnSK_t)$	5.534*	
	Lower bound	*Upper bound*
Significance level		
1%	3.09	3.86
5%	2.93	3.83
10%	2.101	3.869

Note: * indicates a 10% level of significance.

Table 3. Johansen cointegration test.

	Eigenvalue	Trace Statistic	Critical value 0.05	Prob **
None *	0.149	3380.233	159.529	0.000
At most 1 *	0.135	2507.699	125.615	0.000
At most 2 *	0.110	1721.344	95.753	0.000
At most 3 *	0.087	1090.921	69.818	0.000
At most 4 *	0.056	596.760	47.856	0.000
At most 5 *	0.025	286.123	29.797	0.000
At most 6 *	0.022	147.548	15.494	0.000
At most 7 *	0.004	23.174	3.841	0.000

Note: Trace test indicates 8 cointegrating eqn(s) at the 0.05 level. Further, "*" denotes rejection of the hypothesis at the 0.05 level.

Table 3, indicates that both the maximum eigenvalue and trace tests reject the null hypothesis of no cointegration. Specifically, at a 5% significance level, the tests reveal the presence of eight cointegrating equations.

4.3. Long-run nd short-run analysis

Before applying the ARDL model, selecting the optimal lag length is crucial. Various criteria, such as the Akaike Information Criterion (AIC), Hannan-Quinn (HQ) Information Criterion, and Schwarz Bayesian Criterion (SBC), help determine the optimal lag length. In this study, the optimal lag length is based on the AIC criterion. Table 4 presents the results of the long-run and short-run analysis using the ARDL model.

The speed of ECM_{t-1} of the model satisfies the expected condition of negative and significant value that corrects the previous period's disequilibrium in the coming period (-0.811).

The results show that the variable whip significantly decreases with temperature both in the long term and short run. This implies that in the long term (short run), a 1% increase in whip results in a significant decrease in temperature of 0.215% (0.811%).

The variable *Rev* shows a negative and significant impact on temperature, both in the long term and short term. According to these results, a 1% increase in Rev leads to a decrease in temperature by 0.219% (0.277%) in the long run (short run). A positive and significant relationship has been established between ball energy and temperature. A 1% increase in ball energy results in a temperature increase of 0.112% and 0.155% in the long run and short run, respectively.

Table 4 presents the diagnostic test results for our model, Figure 4 gives the summary of empirical results. An R^2 value of 0.85 indicates a strong fit of the estimated model. Additionally, tests for serial correlation, heteroscedasticity, Ramsey tests, and normality confirm that our model is correctly specified and follows normality. Furthermore, these tests validate the absence of serial correlation and heteroscedasticity issues.

Table 4. Results of ARDL estimation (*LnT* is a dependent variable)

Regressors	Coefficient	t-statistics	Prob
Long-run analysis			
LnW	-0.215	-8.773	0.000***
LnR	-0.219	-6.621	0.000***
$LnOR_E$	-0.178	-4.398	0.000***
$LnSK$	0.025	0.636	0.524
$LnIR_E$	0.193	4.197	0.000***
LnC_E	0.035	0.823	0.410
LnB_E	0.112	3.557	0.000***
Short-run analysis			
ECM_{t-1}	-0.811	-1.470	0.000***
$\Delta(LnW)$	-0.248	-7.787	0.000***
$\Delta(LnR)$	-0.277	-5.673	0.000***
$\Delta(LnOR_E)$	-0.233	-3.477	0.000***
$\Delta(LnSK)$	0.030	0.433	0.664
$\Delta(LnIR_E)$	0.292	3.656	0.000***
$\Delta(LnC_E)$	-0.015	-0.202	0.839
$\Delta(LnB_E)$	0.155	3.772	0.000***
Constant	4.281	0.069	0.000***
Diagnostic tests			
R^2	0.85		
X^2 **ARCH**	0.388 (0.960)		
X^2 **Ramsey**	2.839 (0.213)		
X^2 **LM**	0.352 (0.399)		
X^2 **Normality**	0.144 (0.930)		

Note: *** indicates a 1% level of significance.

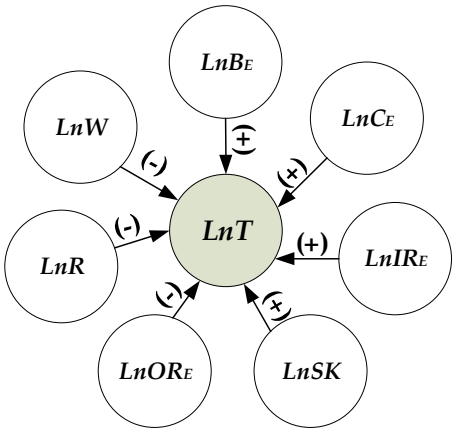

Figure 4. Summary of empirical results

4.4. Pairwise Granger causality analysis

The relationship between the variables suggests the presence of Granger causality, as determined by the F-statistic. The summary of pairwise Granger causality is provided in Table 5 and Figure 5, which includes the direction of causality between the variables. The results of the pairwise Granger causality tests indicate a unidirectional causality relationship from temperature to outer race energy, spectral kurtosis, inner race energy, and cage. Our findings suggest a bidirectional causal relationship between whip and temperature, Rev and temperature, as well as temperature and ball energy.

Table 5. Pairwise Granger causality analysis

Null hypothesis	F-statistic	Prob
LnW **does not Granger cause** LnT	8.935	0.0000***
LnT **does not Granger cause** LnW	9.945	5.E-05***
LnR **does not Granger cause** LnT	3.223	0.039**
LnT **does not Granger cause** LnR	5.953	0.002***
$LnOR_E$ **does not Granger cause** LnT	0.841	0.431
LnT **does not Granger cause** $LnOR_E$	5.985	0.002***
$LnSK$ **does not Granger cause** LnT	0.317	0.727
LnT **does not Granger cause** $LnSK$	6.083	0.002***
$LnIR_E$ **does not Granger cause** LnT	1.154	0.315
LnT **does not Granger cause** $LnIR_E$	5.189	0.005***
LnC_E**does not Granger cause** LnT	0.672	0.510
LnT **does not Granger cause** LnC_E	3.791	0.022**
LnB_E **does not Granger cause** LnT	4.412	0.012**
LnT **does not Granger cause** LnB_E	4.853	0.007***

Note: *, **, and *** indicate 1%, 5%, and 10% levels of significance, respectively.

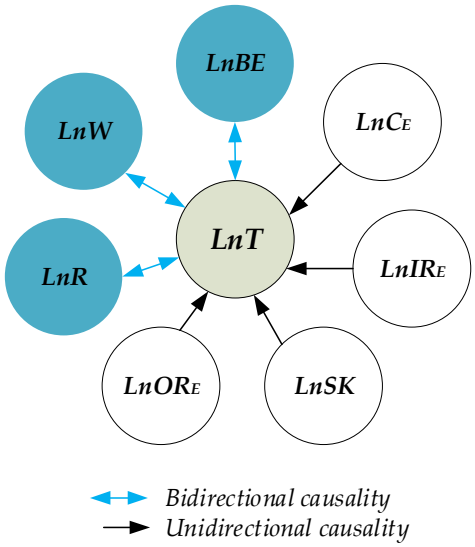

Figure 5. Granger causality results

4.5. Impulse response function

The IRF plays a crucial role in measuring the impacts of shocks from independent variables on the dependent variable. This method enables us to analyze the dynamic interactions among variables in our model, helping us to quantify the effects of these shocks. Figure 3 illustrates the impulse response functions of the temperature model.

1. CONCLUSION

The paper introduces a novel health indicator based on cointegration for BGCI. Through the cointegration test, the study found a certain degree of cointegration relationship between oil temperature and BGCI values for the Bell 407 helicopter leading to the development of the novel health indicator.

The paper stands out by applying econometric models, which are not commonly used in the CM community, to analyze rotating machinery data. This approach offers a fresh perspective and could inspire new avenues of research within the field. The findings have potential practical applications, particularly in improving predictive maintenance and monitoring of helicopter gearboxes. The identification of a cointegrated relationship between oil temperature and BGCI values could lead to more effective health monitoring systems.

While the results are promising, one of the challenges noted in the study is the potential overlap between different fault conditions. The study focused on commonly encountered gearbox bearing faults, but it may not encompass all possible fault scenarios that could occur in rotating machinery in rotorcraft. This limitation could affect the robustness of the proposed diagnostic methods when applied to less common or more complex fault conditions. These factors highlight areas for future research and improvement in the proposed diagnostic methods.

Moreover, the authors suggest that further experimentation on different rigs and with various fault types is necessary to fully validate the method's effectiveness and explore its potential for broader industrial applications.

References

1. Tabrizi, A.A., Al-Bugharbee, H., Trendafilova, I. and Garibaldi, L., (2017). A cointegration-based monitoring method for rolling bearings working in time-varying operational conditions. Meccanica, 52, pp.1201-1217.
2. Zhang, Y.G., Jiang, X.F., Wu, X.W. and Ying, Z., (2012). Diagnosis of High Oil Temperature of the Bearings Based on Oil Analytical Techniques. Applied Mechanics and Materials, 224, pp.119-122.
3. Huang, B., Xu, Y., Mba, D., Li, X., Hu, N., Gu, F. and Ball, A., (2020), April. Helicopter Planet Bearing Fault Diagnosis Based on Modulation Signal Bispectrum Analysis. In International Conference on Maintenance Engineering (pp.784-797). Cham: Springer International Publishing.
4. Abboud, D., Antoni, J., Sig-Zieba, S., Eltabach, M. (2017), Evneope analysis of rotating machine in variable speed conditions: A comprehensive treatment, Mechanical Systems and Signal Processing, Volume 84, Part I.
5. Li, H., Li, Y. and Yu, H., (2019). A novel health indicator based on cointegration for rolling bearings' run-to-failure process. Sensors, 19(9), p.2151.
6. Wei, L., Cheng, Z., Cheng, J., Hu, N. and Yang, Y., (2022). A fault detection method based on an oil temperature forecasting model using an improved deep deterministic policy gradient algorithm in the helicopter gearbox. Entropy, 24(10), p.1394.
7. Bayoumi, A., McKenzie, A., Gouda, K., McVay, J. and Carr, D., (2012), May. Impact of Lubrication Analysis on the Improvement of AH-64D Helicopter Component Performance. In Proceedings of the AHS 68th Annual Forum, Ft. Worth, TX.
8. Murrad, M. and Leong, M.S., (2008). Application of Vibration and Oil Analysis for Reliability Information on Helicopter Main Rotor Gearbox. Journal of System Design and Dynamics, 2(1), pp.370-381.
9. Xu, J., Li, X., Chen, R., Wang, L., Yang, Z. and Yang, H., (2021). Dynamic characteristics analysis of gear-bearing system considering dynamic wear with flash temperature. Mathematics, 9(21), p.2739.
10. Shafique, M., Azam, A., Rafiq, M. and Luo, X., (2021). Investigating the nexus among transport, economic growth and environmental degradation: Evidence from panel ARDL approach. Transport Policy, 109, pp.61-71.

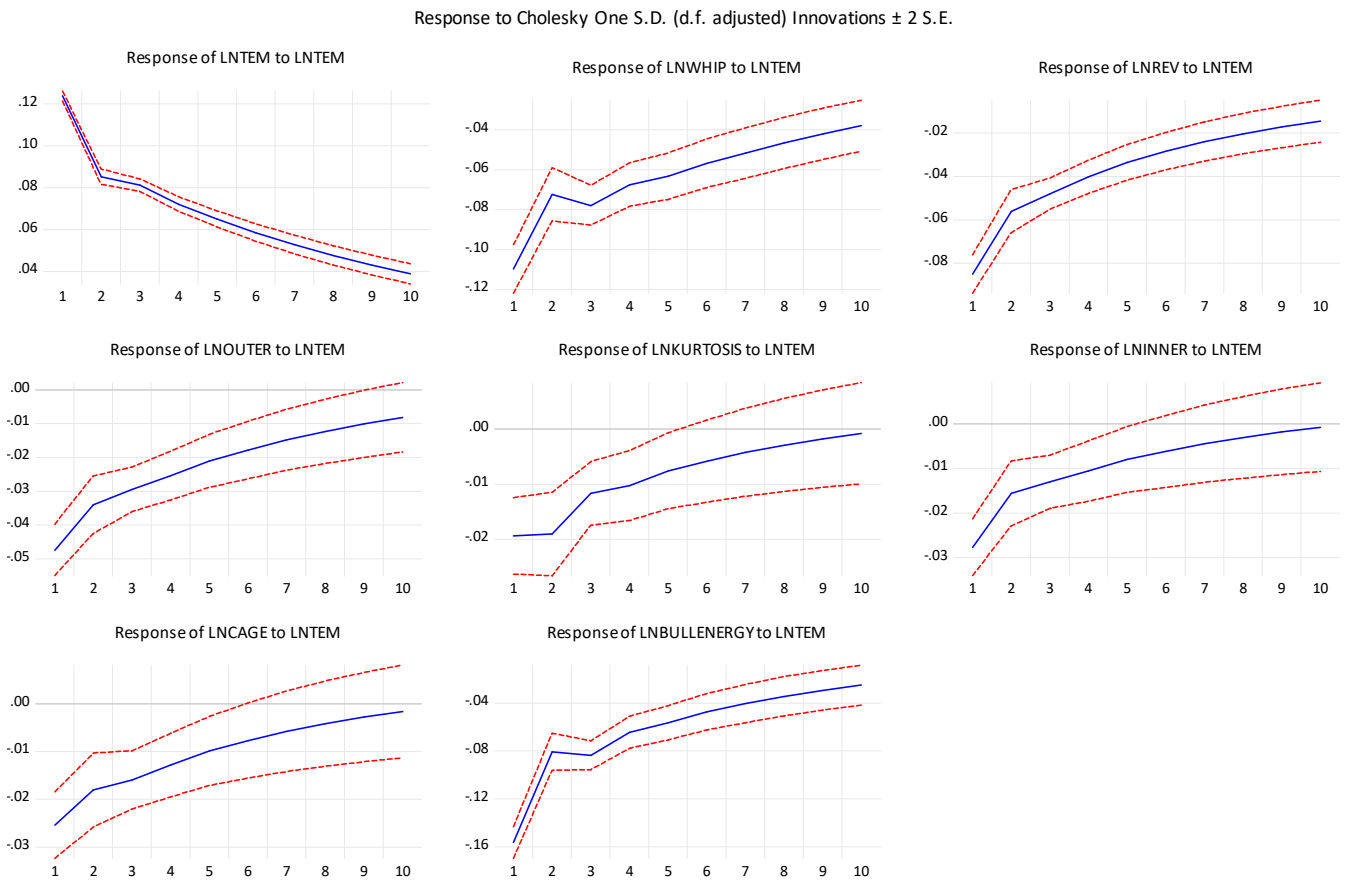

Figure 6. The impulse response functions of the temperature model

11. Ouni, M., & Ben Abdallah, K. (2024). Environmental sustainability and green logistics: Evidence from BRICS and Gulf countries by cross-sectionally augmented
12. autoregressive distributed lag (CS-ARDL) approach. Sustainable Development.
13. Phillips, P. C., & Perron, P. (1988). Testing for a unit root in time series regression. biometrika, 75(2), 335-346.
14. Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. Journal of the American Statistical Association, 74(366a), 427-431.
15. Narayan, P. K. (2005). The saving and investment nexus for China: evidence from cointegration tests. Applied Economics, 37(17), 1979-1990.
16. Pesaran, M. H., Shin, Y., & Smith, R. J. (2001). Bounds testing approaches to the analysis of level relationships. Journal of Applied Econometrics, 16(3), 289-326.
17. Dumitrescu, E. I., & Hurlin, C. (2012). Testing for Granger non-causality in heterogeneous panels. Economic modelling, 29(4), 1450-1460.
18. Granger, C. W. (1969). Investigating causal relations by econometric models and cross-spectral methods. Econometrica: journal of the Econometric Society, 424-438.
19. Ciokajlo, J. J., Jeffre, G. L. (1990). 5. Bearing assembly for use in high temperature operating environment.