A statistical analysis of the relationship between Pc4 and Pc5 ULF waves, solar winds and geomagnetic storms for predicting earthquake precursor signatures in low latitude regions

S N A Syed Zafar¹, Roslan Umar¹*, N H Sabri², M H Jusoh³, A Yoshikawa⁴, S Abe⁴, T Uozumi⁴

¹ East Coast Environmental Research Institute, Universiti Sultan Zainal Abidin, 21300 Kuala Nerus, Terengganu, Malaysia
² Advanced Nano Materials (ANoMa) Research Group, Faculty of Science and Marine Environment, Universiti Malaysia Terengganu, 21030 Kuala Nerus, Terengganu
³ Applied Electromagnetic Research Group (AERG), Faculty of Electrical Engineering, Universiti Teknologi Mara, 40450 Shah Alam, Selangor, Malaysia
⁴ International Center for Space Weather Science and Education (ICSWSE), Kyushu University, Fukuoka, Japan

* Corresponding author: roslan@unisza.edu.my

Abstract. Short-term earthquake forecasting is impossible due to the seismometer's limited sensitivity in detecting the generation of micro-fractures prior to an earthquake. Therefore, there is a strong desire for a non-seismological approach, and one of the most established methods is geomagnetic disturbance observation. Previous research shows that disturbances in the ground geomagnetic field serves as a potential precursor for earthquake studies. It was discovered that electromagnetic waves (EM) in the Ultra-Low Frequency (ULF) range are a promising tool for studying the seismomagnetic effect of earthquake precursors. This study used a multiple regression approach to analyse the preliminary study on the relationship between Pc4 (6.7-22 mHz) and Pc5 (1.7-6.7 mHz) ULF magnetic pulsations, solar wind parameters and geomagnetic indices for predicting earthquake precursor signatures in low latitude regions. The ground geomagnetic field was collected from Davao station (7.00° N, 125.40° E), in the Philippines, which experiences nearby earthquake events (Magnitude <5.0, depth <100 km and epicentre distance from magnetometer station <100 km). The Pc5 ULF waves show the highest variance with four solar wind parameters, namely SWS, SWP, IMF-Bz, SIE and geomagnetic indices (SYM/H) prior to an earthquake event based on the regression model value of $R^2 = 0.1510$. Furthermore, the IMF-Bz, SWS, SWP, SWE, and SYM/H were found to be significantly correlated with Pc5 ULF geomagnetic pulsation. This Pc5 ULF magnetic pulsation behaviour in solar winds and geomagnetic storms establishes the possibility of using Pc5 to predict earthquakes.
1. Introduction

The ULF wave anomaly prior to large earthquakes has been intensively investigated during the past decade. Experimental results show that electromagnetic emissions have a frequency range when rocks are under stress, which means the pressure is triggered by a large current [1]. The earthquake precursor is most commonly analysed in a frequency range of $f < 10$ Hz with skin depth less than 100 km and the epicentre distance from magnetometer being within 100 km [2]. Studies have found that ULF waves are generated due to the enhanced Sun-Earth coupling mechanism rather than some geological phenomena. The study shows that the $Pc3$-$Pc5$ wave range is a range that witnesses magnetized plasma wave propagation (Alfvén waves) in the magnetosphere and ionosphere [3][4]. The $Pc3$ magnetic pulsation appears to be very closely correlated to magnetic storms with seismic events [5]. Takla et al. (2018) discovered abnormal geomagnetic variations in a $Pc3$ range near the epicentre of some seismic events at the KTB station in Sumatra, Indonesia [6]. The significant change in polarization ratio analysis of the Pc4-Pc5 range before an earthquake event was revealed by several studies [7][8]. The $Pc3$-$Pc5$ range and occurrence with storms have shown a clear dependence associated with IMF and solar wind parameters [9][10]. Meanwhile, the relationship between $Pc3$-$Pc5$ magnetic pulsations and changes in the solar wind have been extensively studied in recent decades. The existence of $Pc5$ pulsations was triggered by fluctuations in solar wind dynamic pressure [11]. The $Pc5$ was found to be statistically correlated with solar wind speed in high latitude regions (Ji Kwon et al. 2019). The highest ULF power in the $Pc5$ range with high solar wind changes on the ground was observed in a declining solar cycle phase of 23 [12]. To determine and validate the ULF seismomagnetic phenomenon in solar winds and geomagnetic storms, this paper examined the possible ULF earthquake precursor signature with solar wind parameters and geomagnetic storms in low latitude regions. The ULF range ($Pc4$-$Pc5$) in solar winds and geomagnetic storms prior to an earthquake precursor was also examined.

2. Methodology

In order to investigate the relationship between geomagnetic pulsation and solar wind parameters associated with geomagnetic storms near equatorial (SYM/H) regions for earthquake precursor research, geomagnetic pulsation data near seismic events with a depth of less than 100 km, magnitude less than 5, and an epicentre distance less than 100 km from the magnetometer station were collected. The International Association of Geomagnetism and Aeronomy (IAGA) classified two magnetic pulsations in the ULF band, which were investigated based on solar wind parameters and geomagnetic storms by using $Pc4$ (6.7-22 mHz) and $Pc5$ (1.7-6.7 mHz) for earthquake precursor signature research. These geomagnetic pulsations were obtained and collected from the SuperMag website (http://supermag.jhuapl.edu/) that services the Davao magnetometer station (7.00° N, 125.40° E) located in the Philippines [13][14]. Figure 1 depicts the location of the Davao magnetometer station. The sample interval for these data was 1 s. The observation was analysed within two months of the data being collected, prior to the occurrence of earthquake events.
This study selected four solar wind parameters, namely solar wind speed (SWS), solar wind pressure (SWP), IMF-Bz component (Bz), solar wind input energy (SWE) and symmetric part of the magnetospheric ring’s current activity at the equator, which is (SYM/H). The sample interval for these data was 1-minute. Data were acquired from (https://omniweb.gsfc.nasa.gov/form/omni_min.html), which were provided by Space Physics Data Facility (SPDF) based in NASA's Goddard Space Flight Centre in Greenbelt, MD U.S.A. Two months of data prior to the earthquake event were analysed with Pc4 and Pc5 ULF pulsation. The association of these parameters with Pc4 and Pc5 were determined for low latitude regions.

Multiple regression was used to examine the relationship of the strength between Pc4 and Pc5 magnetic pulsations, solar wind parameters and geomagnetic storms in low latitude regions for earthquake precursor signatures. The multivariate regression analysis was formulated according to the model equation below;

\[ y = \beta_0 + \beta_1x_1 + \ldots \beta_nx_n + \varepsilon \]

Where, \( y \) is the dependent variable, \( x_1 \) is the independent variable, \( \beta_1 \) through \( \beta_n \) are the estimated regression coefficients, \( \beta_0 \) is the value of \( Y \) when all independent variables (\( x_1 \) through \( x_n \)) are equal to zero and \( \varepsilon \) is the error. This equation model will be used for predicting the \( y \) variable as a linear function of the \( x \) variables [15]. In this study, there were two dependent variables, namely Pc4 and Pc5 ULF pulsations, which were analysed based on five independent variables, such as solar wind speed (SWS), solar wind dynamic pressure (SWP), IMF-Bz component (Bz), solar wind input energy (SIE) and magnetospheric ring current activity at the equator (SYM/H).

The multiple regression came out with the \( R^2 \) value, which could be used to estimate how a variation of the dependent variable changes as each independent variable change, as well as the strength of the variation of the dependent variable compared with selected independent variables. The \( R^2 \) coefficient (beta value) in the regression analysis determines the correlation between the independent variables, multicollinearity of independent variables (VIF-value) and significance (p-value), as discussed in the regression analysis [16]. Finally, performance of the regression model was determined. However, before performing the multivariate regression, the normality of assumptions for each variable was examined using the Shapiro-Wilk test. As the p-value >0.05 (alpha-value) for all variables, it was found to be normally distributed, thus, the null hypothesis was rejected because the null hypothesis of the observation was not normally distributed [17][18].
3. Result

The regression prediction model for Pc4 and Pc5 ULF geomagnetic pulsations with four solar wind parameters and geomagnetic indices was generated, as shown in equations (1) and (2). According to the results of the multiple regression analysis, the three solar wind parameters and geomagnetic indices discovered were significantly predictive using Pc4 ULF geomagnetic pulsation with $p<0.05$. The strength effect of each solar wind parameter and geomagnetic index on Pc4 variance is shown in Table 1 as a coefficient value. The coefficient values of five independent variables are as follows: $(Bz = -0.0164, T = -10.28, p < 0.05)$, $(SWS = 0.0007, T = -21.65, p < 0.05)$, $(SWP = 0.0740, T = 31.65, p<0.05)$, $(SIE = -0.0001, T = -1.16, p > 0.05)$ and $(SYM/H = -0.0038, T = -16.25, p < 0.05)$. Based on the coefficient results, the variable with the most contribution was solar wind pressure, which had the highest coefficient value compared to other variables. This was followed by the IMF-$Bz$ component, solar wind speed, geomagnetic indices (SYM/H) and solar wind input energy. In this analysis, the contribution of solar wind input energy to Pc4 was found to be insignificant, with a $p$-value greater than the significant value ($\alpha = 0.05$) and a very small coefficient contributing towards a small variance in Pc4. The VIF value in table 1 indicates how much of the variance is inflated in each coefficient. As a rule of thumb, a VIF value of 1 means it is not correlated, between 1 and 5 means it is moderately correlated, and more than 5 means highly correlated [19]. The result shows that the VIF value for solar wind parameters is more than 1, which means the association of solar wind parameters and geomagnetic indices is moderately colinear with each other in the model. The model degree explaining the variance of Pc4 ULF waves in $R^2$ value indicates how well Pc4 varies with respect to all solar wind parameters. Thus, the variance in Pc4 based on the four solar wind parameters and geomagnetic indices is described as $R^2 = 0.0482$, which corresponds to a 4.82 percent variance explained by four wind parameters and geomagnetic indices in low latitude regions. The $R^2$ for Pc4 in Table 2 shows that $R^2$-predicted is greater than $R^2$, indicating that the Pc4 prediction model is overfit [20][21]. In future studies with extended observation time, the unimportant independent variable should be excluded in order to fit the model and initiate the prediction model. Simultaneously, in this multiple regression analysis, the residual plot is used to verify whether the model is adequate to meet the assumptions of the analysis. The residual plot vs fitted value is used in this analysis. When the residual centre is zero, it indicates that the prediction model is correct on average. The residual plot of the Pc4 model shows a central tendency (Figure 2), despite the presence of a few outliers, indicating that this model cannot predict well.

$\text{Pc4}= 1.1683 - 0.0164 \text{Bz} + 0.0007 \text{SWS} + 0.0740 \text{SWP} - 0.0001 \text{SIE} - 0.0038 \text{SYM/H}$  \hspace{1cm} (1)

| Variables | Coefficients (Coef) | Standard error coef (SE coef) | T-value | p-value | Variation Inflation factor (VIF) |
|-----------|---------------------|------------------------------|---------|---------|---------------------------------|
| Constant  | 1.1683              | 0.0128                       | 91.20   | 0.000   |                                 |
| Bz        | -0.0164             | 0.0015                       | -10.98  | 0.000   | 2.05                            |
| SWS       | 0.0007              | 0.0000                       | 21.65   | 0.000   | 1.27                            |
| SWP       | 0.0740              | 0.0023                       | 31.65   | 0.000   | 1.36                            |
| SIE       | -0.0001             | 0.0000                       | -1.16   | 0.244   | 1.72                            |
| SYM/H     | -0.0038             | 0.0002                       | -16.25  | 0.000   | 1.32                            |

$F(5, 48001) = 485.76$
Table 2. Multiple regression result of Magnetic pulsation of Pc4.

| R² | R²-adjusted (R²-adj) | R²-predicted (R²-pred) |
|----|----------------------|------------------------|
| 4.82 % | 4.81 % | 4.79 % |

![Versus Fits (response is Pc4)](image)

Figure 2. Residual versus fitted value plot of Pc4 with four solar wind parameters and geomagnetic storm (SYM/H).

The Pc5 prediction model with four solar wind parameters and geomagnetic indices was created based on Equation (2). Meanwhile, Table 3 shows the coefficient values for four solar wind parameters and geomagnetic indices with a Pc5 magnetic pulsation. These five independent variables’ coefficient robustness (Bz = 0.0048, T = 2.82, p <0.05), (SWS = 0.0017, T = -43.80, p <0.05), (SWP = 0.1481, T =54.89, p <0.05), (SIE = -0.0001, T =9.79, p <0.05) and (SYM/H = -0.0046, T = -16.79, p <0.05) and it was found to be significantly predictive of Pc5 ULF magnetic pulsation. Solar wind pressure contributes the most to Pc5 variance, with a higher coefficient value than others. Sequence of the coefficient values is as follows: IMF-Bz component, solar wind speed, geomagnetic indices (SYM/H), and solar wind input energy. The solar wind input energy contributes very little to the variance of Pc5. The association between solar wind parameters and geomagnetic indices is moderately colinear, as indicated by the VIF values. Following that, a prediction model was created with R² = 0.150, which explains the variations of Pc5 with SWS, SWP, IMF-Bz, SIE, and SYM/H, implying that Pc5 has a 15.10 percent variance with selected solar wind parameters. As shown in Table 4, the R²-predicted is greater than R², indicating a poor predictive ability. The residual plot of the Pc5 model reveals a general trend with a few outliers, similar to that of the Pc4 model (see Figure 3).

According to the results of the analysis concerning the Pc4 and Pc5 prediction model, Pc5’s magnetic pulsation has a higher variance in the IMF-Bz, solar wind speed, solar wind dynamic pressure, solar wind input energy components and geomagnetic indices (SYM/H) compared to Pc4. Pc5 has a higher R² value than Pc4, and its magnetic pulsation is better suited for prediction models with four solar wind parameters and geomagnetic indices as well as an earthquake precursor signature for earthquake depths of 100 km and an epicentre distance of 100 km. In addition, the regression model results show that Bz, SWS, SWP, SWE, and SYM/H have a significant influence on the ULF of geomagnetic pulsations (Pc5), which is useful for investigating earthquake precursor signatures in low latitude regions. According to Currie et al. [22], Sharma et al. (2012) [23] and Marchitelli et al. [24],
there is a statistically significant correlation between solar wind changes and geomagnetic storms prior to the occurrence of an earthquake event with different correlation strengths at different latitudes [22][23][24]. Several previous studies have looked into the possibility of earthquake precursor signatures based on Pc4-Pc5 magnetic pulsations [25][26]. The solar wind-magnetosphere coupling mechanism results in the occurrence of geomagnetic storms, implying that charge carriers from plasma densities, driven by solar winds, are capable of flowing out into the lithosphere and emitting ULF waves. This may cause stress in the rock to increase, resulting in an earthquake [27]. Consequently, a preliminary study of the relationship strength of Pc4 and Pc5 geomagnetic pulsations with solar wind parameters and geomagnetic storm prior to predicting the earthquake precursor was carried out using a multiple regression approach. Moreover, the study also intended to determine the significance of solar parameters and geomagnetic storms used in the prediction of the Pc4 and Pc5 ULF magnetic pulsation and their impact on earthquake event preparation in low latitude regions.

\[ \text{Pc5} = 1.9250 + 0.0048 \text{Bz} + 0.0017 \text{SWS} + 0.1481 \text{SWP} + 0.0001 \text{IE} - 0.0046 \text{SYM/H} \]  

(2)

**Table 3.** Multiple Regression of Coefficient of four solar wind parameters and geomagnetic index for magnetic pulsation of Pc5.

| Variables | Coefficients (Coef) | Standard error coef (SE coef) | T-value | p-value | Variation Inflation factor (VIF) |
|-----------|---------------------|-------------------------------|---------|---------|-------------------------------|
| Constant  | 1.9250              | 0.0148                        | 130.39  | 0.000   |                               |
| Bz        | 0.0048              | 0.0017                        | 2.82    | 0.005   | 2.05                          |
| SWS       | 0.0017              | 0.0003                        | 43.80   | 0.000   | 1.27                          |
| SWP       | 0.1481              | 0.0027                        | 54.89   | 0.000   | 1.36                          |
| SIE       | 0.0001              | 0.0000                        | 9.79    | 0.000   | 1.72                          |
| SYM/H     | -0.0046             | 0.0003                        | -16.79  | 0.000   | 1.32                          |

\[ F (5, 48001) = 1707.37 \]

**Table 4.** Multiple regression result of Magnetic pulsation of Pc5.

|        | R²   | R²- adjusted (R²-adj) | R²-predicted (R²-pred) |
|--------|------|-----------------------|------------------------|
|        | 15.10% | 15.09%                | 15.07%                 |

**Figure 3.** Residual versus fitted value plot of Pc5 with four solar wind parameters and geomagnetic indices (SYM/H).
4. Conclusion

This study concluded that the Pc5 ULF magnetic pulsation based on IMF-Bz, solar wind speed, solar wind dynamic pressure, solar wind input energy components and geomagnetic indices (SYM/H) has a greater effect than Pc4. Pc5 geomagnetic pulsation was found to be significantly correlated with the IMF-Bz, SWS, SWP, SWE, and SYM/H based on the generation of the regression model. This indicates that Pc5 is an appropriate wave band to study the possible earthquake precursor signature in ULF wave records by combining selected solar parameters and the SYM/H index. However, it appears that the Pc5 prediction model did not perform well and did not match model assumptions because $R^2$-predicted was greater than $R^2$ and a large residual with a few outliers was obtained. Future studies should fit and improve the model in order to realise the Pc5 prediction model for earthquake precursors at low latitude regions by eliminating irrelevant variables and using extended observation period data. In addition, the residual plot should be adjusted so that it corresponds with the prediction model’s assumptions.

References

[1] Hattori K et al 2013 Investigation of ULF seismo-magnetic phenomena in Kanto, Japan during 2000–2010: case studies and statistical studies. Surveys in Geophysics 34 (3) 293-316
[2] Han P et al 2014 Statistical analysis of ULF seismomagnetic phenomena at Kakioka, Japan, during 2001–2010 Journal of Geophysical Research: Space Physics 119 (6) 4998-5011
[3] Silva G B, Padilha A L and Alves L R 2020 Latitudinal variation of Pc3–Pc5 geomagnetic pulsation amplitude across the dip equator in central South America. In Annales Geophysicae 38 (1) 35-49
[4] McPherron R L 2005 Magnetic pulsations: their sources and relation to solar wind and geomagnetic activity Surveys in Geophysics 26 (5) 545-592
[5] Masruri M F I et al 2018 Analysis of Preseismic Event Using Seismo-electromagnetics Data. In 2018 2nd International Conference on Applied Electromagnetic Technology 1-7 IEEE
[6] Takla E M et al 2018 Anomalous ultra-low frequency signals possibly linked with seismic activities in Sumatra, Indonesia NRIAG Journal of Astronomy and Geophysics 7 (2) 247-252
[7] Yusofa K A et al 2020 Normalized Polarization Ratio Analysis for ULF Precursor Detection of the 2009 M7, 6 Sumatra and 2015 M6 8 Honshu Earthquakes Engineering Journal SI 3 (1) 35-41
[8] Hayakawa M et al 2000 ULF electromagnetic precursors for an earthquake at Biak, Indonesia on February 17, 1996, Geophysical Research Letters 27 (10) 1531-1534
[9] Hynönén R Tanskanen E I and Francia P 2020 Solar cycle evolution of ULF wave power in solar wind and on ground Journal of Space Weather and Space Climate 10 43
[10] Filawati S 2021 Analysis of IMF Orientation and Pc5 Power During St. Patrick Geomagnetic Storm on 17 March 2015 In Journal of Physics: Conference Series 1805 (1) 012026
[11] Pranoto S C et al 2020 Ap c5 ULF waves during the geomagnetic storms on 7–8 September 2017 observed in the Indonesian region In Journal of Physics: Conference Series 1523 (1) 012025
[12] Kwon H J et al 2019 Characteristics of Pc5 activity at high latitudes stations in Antarctica Journal of Atmospheric and Solar-Terrestrial Physics 193 105087
[13] Gjerlov J W 2012 The SuperMAG data processing technique Journal of Geophysical Research: Space Physics 117 (A9)
[14] Syed Zafar S N A et al 2021 Behaviour of Geomagnetic storm, horizontal geomagnetic field and solar wind parameters during solar flare and CMEs event Journal of Physics: Conference Series, IOP Publishing 1768 (1) 012003
[15] Uyanik G K and Güler N 2013 A study on multiple linear regression analysis Procedia-Social and Behavioral Sciences 106 234-240
[16] Syed Zafar S N A et al 2019 Radio Frequency Interference on Nearby Radio Astronomical Lines: Relationship between Wind Speed and Radio Signal Strength Measured at East Coast of Peninsular Malaysia. Sains Malaysia E 48 (1) 183-189
[17] Sabri N H et al 2017 Correlation analysis of tropical rainfall climate effect on radio signal strength at KUSZA observatory, Terengganu. Advanced Science Letters 23 (2) 1268-1271
[18] Shapiro S S and Wilk M B 1965 An analysis of variance test for normality (complete samples). Biometrika, 52 (3/4) 591-611.
[19] Tranmer M and Elliot M 2008 Multiple linear regression. The Cathie Marsh Centre for Census and Survey Research (CCSR) 5 (5) 1-5
[20] Mat R et al 2018 The modelling of tropical weather effects on ultra-high frequency (UHF) radio signals using SmartPLS. In IOP Conference Series: Materials Science and Engineering 440
[21] Olive D J 2017 Multiple linear regression. In Linear regression 17-83
[22] Currie J L and Waters C L 2014 On the use of geomagnetic indices and ULF waves for earthquake precursor signatures Journal of Geophysical Research: Space Physics 119 (2) 992-1003
[23] Sharma A K, Patil A V and Haridas R N 2012 Investigation of ULF magnetic anomaly before moderate earthquakes. Exploration Geophysics 43 (1) 36-46
Acknowledgement

This study was supported by the grant FRGS/1/2019/STG02/UNISZA/02/3. Special thanks are dedicated to Electromagnetic Research Group (EMRG) for their assistance in this work. Last but no least we want to acknowledge use of the data from SuperMAG and OMNIWeb.