1 **Singular spectrum and principal component analysis of soil radon**

2 **(Rn-222) emanation for better detection and correlation of seismic**

3 **induced anomalies**

4 Timangshu Chetia[1, 2], Saurabh Baruah[1], Chandan Dey[1, 2], Sangeeta Sharma[1], Santanu Baruah[1]

5 [1]Geoscience & Technology Division, CSIR-North East Institute of Science and Technology
6 (CSIR-NEIST), Jorhat-785006, Assam, India
7 [2]Academy of Scientific and Innovative Research (CSIR-NEIST), Jorhat-785006, Assam, India

8 *Correspondence to*: Santanu Baruah (santanub27@gmail.com)

9

10 **Abstact.** In the recent years there are several reporting's of anomalous seismic induced

11 temporal changes in soil radon emanation. It is however well known that radon anomalies apart

12 from seismic activity are also governed and controlled by meteorological parameters. This is

13 the major complication which arise for isolating the seismic induced precursory signals. Here

14 in the investigation the soil radon emanations temporal variability at Multiparametric

15 Geophysical Observatory (MPGO), Tezpur, is scrutinized in the lime light of singular spectrum

16 analysis (SSA). Further prior applying SSA Digital filter (Butterworth low pass) is applied to

17 remove the high frequen.o9cy quasi periodic component in the time series of soil radon

18 emanation. It was scrutinized that sum of just 9 eigenfunctions were sufficient enough for

19 reproducing the prominent characteristics of the overall variation. This perhaps also evinces

20 that more significantly produced fluctuations are mostly free from natural variations. The

21 variations in soil temperature was observed to be dominated by daily variations similar to radon

22 variation which account to 97.99 % whereas soil pressure accounts for 100 % of the total

23 variance which suggests that daily variations of soil radon (Rn-222) emanation are controlled

24 by soil pressure in MPGO, Tezpur during the investigation period followed by soil temperature.

25 The study concludes that SSA eliminates diurnal and semidiurnal components from time series

26 of soil radon emanation for better correlation of soil radon emanation with earthquakes.

1 Introduction

Radon (Rn-222) is a noble gas, a decay product of radium with atomic number, Z=86 and a half-life of nearly 3.8 days. Because of its short decay time, amount changes in its production from the rock is quite evidenced. Ulomov and Mavashev in the year 1971 (Ulomov and Mavashev, 1971) first suggested the correlation of radon concentration with earthquakes. It has been scrutinized that the radon concentration has correlation to earthquakes and volcanic eruptions (Walia et al., 2005; Singh et al., 2010; Chetia et al., 2019). Significant radon concentration anomalies were also observed prior to the Uttarkashi earthquake of 20th October, 1991; m_b~6.5 (Virk and Singh, 1994). Radon concentrations was monitored in the North West Himalaya for earthquake prediction studies and empirical equation between earthquake magnitude, epicentral distance and precursory time were examined (Choubey et al. 2009). Earthquake prediction depending entirely on precursory phenomena is empirical and comprises many applied problems. Among various geophysical parameters soil radon is preferred and used for earthquake precursory studies because of its ease of detectability. Radon in nature has different isotopes: Rn-222 (half-life~3.8 days), Rn-220 (Thoron, half-life~54.5 s) and Rn-219 (half-life~3.92 s). The utmost prominent is Rn-222 because of its longer half-life which is a product of Ra-226 decay process. The Rn-222 emanates from soil or crust either by diffusion or convection and reaches the atmosphere. The soil radon emanation concentration is generally assigned to developments of micro-cracks, fissure and fracture due to dilatancy prior to earthquake. This process enhances the transportation of Radon from its original enclosure following the cracks into the atmosphere. Variations of soil radon and thoron concentrations were also observed in fault zone prior to earthquakes SW Taiwan. Anomalous peaks were discernible usually few days or weeks prior to the earthquakes (Yang et al., 2005).

North-East India (NE India) is highly vulnerable to earthquake and lies in seismic zone V (BIS 2002) and frequent occurrence of earthquake facilitates the probability of finding

52   precursory phenomena which may lead to a successful prediction in near future. The estimated

53   cumulative probabilities show about a similar recurrence period of about 15-25 years from the

54   occurrence of last earthquake for a forecast of large earthquakes (Mw > 5) in the NE-India

55   region and its vicinity (Chetia et al., 2019). With this objective a Multiparametric Geophysical

56   Observatory (MPGO) in Ouguri Hills (Latitude N26.61^{o}; Longitude E92.78^{o}, Elevation~82m),

57   Tezpur, Assam, India with the installation of several geophysical instruments collecting data

58   simultaneously, portray an opportunity towards identification of precursory signatures prior to

59   earthquakes.   Earthquake precursory and prediction studies advanced in late 1970s and

60   Heicheng earthquake of 4th February is a land- mark which was in short-term successfully

61   predicted in 1975 in China (Adams, 1976). The accomplished medium term forecast of M~7.5

62   earthquake on 6th August, 1988 in northeast Indian region (Gupta, 1988) encouraged to

63   strengthen such studies in India. Another successful short term prediction was done of (M $\geqslant$ 4)

64   in Koyna region of western India, famous for Reservoir Triggered Seismicity (Verma and

65   Bansal, 2012).

66   The study tries to correlate radon emanation in soil gas with earthquakes within the

67   epicentral distance of 100 km of mb > 3.1 from MPGO, Tezpur which is situated in a highly

68   tectonically strained and seismically active region. The reason for considering the criteria of

69   100 km with mb > 3.1 is the tectonic setting of the study area which is bounded by two major

70   fault namely Kopili and the Bomdila Faults. The region is highly stressed because of

71   geotectonic settings where earthquake do occur periodically within very short span of time.

72   Kopili fault zone is experiencing compressional stress due to the Indo-Burma arc and Himalyan

73   arc to the east and the north respectively which is characterized by transverse tectonics. The

74   Bomdila Fault inter-weaves across three major tectonic domains of the NE-India, namely

75   MCT, MBT and Naga-Disang thrust along NW-SE direction. (Kayal, 2010). The empirical

76   relation for the energy release E to the magnitude (Gutenberg, 1956) gives E > $10^{9.2}$ for

77   earthquake of mb > 3.1. Stress and strain are directly proportional and hence in order to obtain

78   adequate stress build up processes for observing precursory signals in temporal soil Radon

79   emanation the criteria of 100 Km with mb>3.1 is taken into consideration. The major problem

80   is the removal of quasi periodic, diurnal (mostly due to temperature) and semidiurnal (mostly

81   due to pressure) components (Kumar et al., 2015). Radon anomalies are governed by seismic

82   activity as well as by meteorological parameters (soil temperature, pressure, rainfall, moisture

83   and even wind for atmospheric radon (Stranden et al., 1984; Choubey et al., 2009; Walia et al.,

84   2005) which in turn makes it more complex for identifying the seismic induced anomalies.

85   Here in the investigation characteristics features of temporal soil radon concentrations

86   variability at MPGO, Tezpur is scrutinized by applying singular spectrum analysis (SSA) in

87   concern to the objective of filtering the meteorological parameters on radon concentration. SSA

88   is a relatively innovative and powerful advanced method which has been used across many

89   applied problems for different scientific fields (e.g. Fraedrich, 1986; Serita et al., 2005).

90   Seismic induced emanation of soil gas radon is a complex phenomenon to be studies based on

91   mathematical, statistical analysis and modellings. Nonlinear methods such as power spectrum,

92   fractal-multifractal analysis and singular spectrum analysis etc. probably are commanding tools

93   which reveals the nonlinear features and mechanism of soil gas radon emanation processes.

94   The major advantages of Caterpillar-SSA method is that it does not require better

95   understanding of time series parametric model and extensively applicable to varied spectrum

96   of diverse temporal data. It can also be applied to match with non-stationary temporal data and

97   also permits to estimate structures even for short temporal data. Here Caterpillar-SSA method

98   is used to study the temporal soil gas radon and the foremost concept of SSA is applying PCA

99   on trajectory matrix acquired from the original time series following time series reconstruction.

100   **2   Seismotectonics of the region**

101   In middle of the active Kopili and Bomdila fault, the MPGO is situated. The Kopili

102   and Bomdila faults comprise Neogene-Quaternary sediments, which directly were deposited

103   over the Archean basement. The Kopili fault zone in an approx is 100 km in width and 300 km

104   in length. It is a NW-SE trending strike-slip fault (Kayal et al., 2006, Bhattacharya et al., 2008,

105   2010). The two Precambrian massifs - the Shillong Plateau and the Mikir Hills is delineated by

106   the tectonic disposition of the Kopili fault. MPGO is bounded to the north by the Main

107   Boundary Thrust (MBT) and to the south by the NE-SW trending Belt of Schuppen (Figure 1).

108   The Bomdila fault is strike slip fault of about 400 km in length which trends along WNW-ESE

109   direction. The northern portion of the fault mostly lies in the Gondwana, Paleogene and

110   Neogene sediments. This fault is surrounded by the Belt of Schuppen to the east and south by

111   the Mikir massif and to the west. The fault cuts across the Himalayan fold belt towards the

112   north (Nandy and Dasgupta, 1991).

113   The Kopili Fault has produced two large earthquakes (Figure 1) M_w~7.7, 1869 event

114   (Figure 1) in the south eastern edge of the fault contravening the Naga-Disang thrust and

115   M_w~7.2, 1943 earthquake which occurred farther to the north of 1869 event within a period of

116   nearly 75 years (Kayal, 2008). It is highly active with strong seismicity discernible down up to

117   depth of about 50 km, and which extends to the Main Central Thrust (MCT) in the Bhutan

118   Himalaya. Even if MCT is dormant (Ni and Barazangi, 1984), intense activity is observed at

119   the region where Kopili Fault meets the MBT and MCT (Nandy, 2001; Kayal et al., 2010,

120   2012). This is demonstrated by the August 19, 2009 Earthquake (M_w~5.1) in the Assam Valley

121   that occurred in the center of the Kopili fault zone and the September 21, 2009 strong Bhutan

122   Himalaya earthquake (M_w~6.3) that occurred at the northern end of the Kopili fault where it

123   connects with the MCT (Kayal et al., 2012). Both the earthquakes are shallow focus (depth ~

124   10 km) showing right lateral strike-slip faulting (Kayal et al., 2010) which suggests that the

125   Kopili fault zone is experiencing compressional stress due to the Indo-Burma arc and Himalyan

5

126   arc to the east and the north respectively which is characterized by transverse tectonics. The

127   Bomdila Fault inter-weaves across three major tectonic domains of the NE-India, namely

128   MCT, MBT and Naga-Disang thrust along NW-SE direction. The earthquake events along

129   Bomdila fault occur in a diffused pattern having post-collisional intracratonic characteristics

130   (Nandy and Dasgupta, 1991). It is observed that, the Upper Brahmaputra Valley stretching

131   between the Bomdila Fault and almost near NW-trending Mishmi Thrust in the northeast is

132   seismically dormant, and is recognized as the Assam Gap (Khattri, 1983).

133   **3   Method and techniques**

134   **3.1   Soil radon (Rn-222) time series**

135   A BMC2 barasol manufactured by Algade is into operation for soil radon emanation

136   time series data in MPGO for earthquake prediction and precursory studies. Soil gas radon

137   emanation every 15 minutes is being continuously monitored. The barasol probe is kept fitted

138   inside a plastic tube (length 1.5m and diameter of 0.0635m) with open bottom dug inside the

139   ground in such a way that the detection unit (detector sensitivity-0.02 pulses/h for 1 Bq/m^3 and

140   saturation volumetric activity-3MBq/ m^3) which is at the bottom of the probe lies 1m from the

141   ground level. A silicon alpha detector detects the radon gas which enters the detection chamber

142   when it emanates from the soil. The radon pass in a detection volume over three cellulose filters

143   which allows to trap all the solid daughter products of radon. The sensor is a fixed silicon

144   detector with a depleted depth of 100μm and 400 mm^2 of sensitive area. It performs the

145   counting by atomic spectrometry of radon (Rn-222) and daughter products created in the

146   detection volume (with window customized between 1.5 MeV and 6 MeV). The probe system

147   is embedded with soil pressure and temperature sensor. The sensor calibration permits the

148   volumic activity of the radon (Rn-222) to be evaluated. In the present investigation the soil

149   radon emanation temporal variability at MPGO, Tezpur the radon data were prudently checked

150   for no gaps or discontinuous jump. Digital filter (Butterworth) is applied to eliminate the high

151   frequency quasi periodic components form the soil radon time series for better discernibility of

152   seismic induced anomalies.

3.2   Singular Spectrum Analysis

154   The SSA results and graphs in the investigation are acquired by using Caterpillar-SSA

155   3.40 software (Alexandrov and Golyandina, 2004). Window selection rule applied to the time

156   series is one half of the length of the time series to meet the theoretical requirements for the

157   investigation (Golyandina, 2010, Khan, 2011, Hassani, 2007). The singular value

158   decomposition (SVD) alogorithm was applied as it is more accurate than QR iteration which

159   are the most common algorithm for solving of eigenvalues and singular value problems

160   (Demmel and Veselić, 1992).   The main objective of SSA is decomposing the original time

161   series into sum of series such that each of the component in this sum can be known (either as a

162   trend, periodic or quasi-periodic components) or noise. This is accomplished by decomposition

163   and reconstruction. At the first the time series is decomposed following the reconstruction of

164   the original time series (which is without noise). The methodology adopted here is first to

165   embedding a 1-dimension time series say, $Y_T = (y_1,...,y_T)$ into a multi-dimensional time series

166    $X_1,...,X_K$ having vectors $X_i = (y_i,...,y_{i+L-1}) \in R^L$ (Golyandina et al., 2001, 2001). Here the value

167   of $K = T - L + 1$. The X_i Vectors are called L-lagged vectors. The embedding depends on the

168   window length L, such that $2 \leq L \leq T$ which results for trajectory matrix (Hankel matrix:

169   diagonal elements $i + j$ = const. are equal) $X = [X_1,...,X_K]=(X_{ij})_{i,j=1}^{L,K}$. Secondly the Singular

170   Value Decomposition (SVD) of the trajectory matrix is performed to represent it as an addition

171   of bi-orthogonal elementary matrices having rank one. Represented by $\lambda_1,...,\lambda_L$ which are the

172   Eigen-values of XX' in a descending order of magnitude $(\lambda_1 \geq ...\lambda_L \geq 0)$ and by $U_1,...,U_L$ which

173   are the orthonormal system (i.e. $(U_i, U_j) = 0$ for $i \neq j$) is the orthogonality property and $||U_i||=1$,

174 of the eigenvectors of the matrix XX' corresponding to these eigenvalues. (U_i, U_j) is the inner

175 product of the vectors U_i and U_j and $\|U_i\|$ is the norm of the vector U_i. The Set

$$d = \max (i, \text{ such that } \lambda_i > 0) = rank\ X \tag{I}$$

177 If we represent $V_i = X'\ U_i / \sqrt{\lambda_i}$, then SVD of the trajectory matrix can be represented

178 as:

$$X = X_1 + \cdots + X_d \tag{II}$$

180 Here $X_i = \sqrt{\lambda_i} U_i V_i'$ $(i = 1,...,d)$.

181 Thirdly the series reconstruction is accomplished by grouping, to split the elementary

182 matrices (X_i) into various groups and addition of the matrices within each and every group. The

183 splitting of the elementary matrices Xi into several groups and summation of the matrices

184 within individual group is grouping. Say $I = (i_1, \dots, i_p)$ is a group of indices $i_1, \dots, i_p$. Now

185 for the group I of corresponding matrix X_I is defined as $X_I = X_{i1} + \dots\dots + X_{ip}$. The spilting

186 of the of the set with indices $J = 1,\dots, d$ on the disjoint subsets $I_1,..., I_m$ resembles to the

187 illustration

$$X = X_{I_1} + \cdots + X_{Im}. \tag{III}$$

189 This procedure of selecting the sets $I_1,\dots, I_m$ is termed as eigentriple grouping. For given

190 group I the contribution of the component X_I into $X_{I1} + \cdots + X_{Im}$ is estimated to the corresponding

191 eigenvalues as: $\sum_{i\in I} \lambda i / \sum_{i=1}^{d} \lambda i.$

192 Finally diagonally averaging transfers each I matrix into a time series, which is an

193 additive component of the initial series Y_T. If z_{ij} is an element of the matrix Z. Then k-th term

194 of the produced series is acquired by averaging z_{ij} for i, j such that $i + j = k + 2$. This procedure

195 is known as diagonal averaging (Hankelization) of the matrix Z. The Hankelization of a matrix

196   Z is the Hankel matrix HZ which is the trajectory matrix corresponding to the series obtained

197   from diagonal averaging. Hankelization is a best logical technique that HZ matrix is the nearest

198   to Z among all corresponding size of Hankel matrices (Golyandina et al., 2001). Now applying

199   the Hankelization technique to each and every matrix components in equation III, we get

200
$$X = \tilde{X}_{I_1} + \ldots + \tilde{X}_{I_m} \tag{IV}$$

201   Here $X_{I_1} = HX$ corresponding to initial series $Y_T = (y_1, \ldots, y_T)$ decomposition into a

202   sum of m series as

203
$$y_t = \sum_{k=1}^{m} \tilde{y}_t^{(k)} \tag{V}$$

204   Here $\tilde{y}_t^{(k)} = \tilde{y}_1^{(k)} + \ldots + \tilde{y}_T^{(k)}$ corresponds matrix X_{1j}.

205   Covariance matrix is a matrix of covariances between the values $X(t)$ and $X(t+k)$., where

206   k is a lag. When k is positive $X(t)$ and $X(t+k)$ tends to fluctuate together. The covariances among

207    $X(t)$ and $X(t+k)$ can be estimated as

208
$$\text{COV} = \frac{\sum_{i=1}^{n}(X(t) - \overline{x(t)})(X(t+k) - \overline{x(t+k)})}{n-1} \tag{VI}$$

209   Where $\overline{x(t)}$ and $\overline{x(t+k)}$ represents the mean of $x(t)$ and $x(t+k)$ correspondingly.

210   Now covariance matrix $(\tilde{C})$ can be obtained by putting equal values $\tilde{c}_{ij}$ at each matrix

211   antidiagona $|i - j| = k$ as

212
$$\tilde{c}_{ij} = \frac{1}{N-|i-j|} \sum_{m=1}^{N-|i-j|} x_m\, x_{m+|i-j|} \tag{VII}$$

213   Where $1 \le i, j \le L$, N is the length of the time series and $|i - j| = k$.

214   From the eigenvectors of the covariance matrix measured at different lags the principal

215   components of the time series is estimated. The principal components is also a time series

216   having same length as the "embedded" time series. The computation of principal component is

217   from simple matrix product as

218   PC = $\tilde{C}$ * matrix of eigenvectors   (VIII)

219   Where $\tilde{C}$ is the covariance matrix, whose elements are $\tilde{c}_{ij}$.

220   The exemplification of how discretely different component can be separated from each

221   other is based on studying the SSA properties. To successfully decompose (SSA) the series Y_T

222   the subsequent additive components of the series should be almost separable from each other.

223   The weighted correlation is also known as w-*correlation* which is the natural measurement of

224   dependencies among two series Y_T^1 and Y_T^2 given as:

225                              $$\rho_{12}^{\omega} = \frac{(Y_T^1, Y_T^2)_{\omega}}{\left\|Y_T^1\right\|_{\omega} \left\|Y_T^2\right\|_{\omega}}$$ (IX)

226   Here, $\left\|Y_T^1\right\|_{\omega} = \sqrt{(Y_T^i, Y_T^i)_{\omega}}$, $(Y_T^i, Y_T^j)_{\omega} = \sum_{k=1}^{T} \omega_k y_k^i y_k^j$, (i, j = 1, 2), ω_k = min{k, L, T-

227   k} while assuming L ≤ = T/2.

228   If the absolute value of w-correlations is small, the corresponding series are almost w-

229   orthogonal. On the other hand if it is large, the series are not w-orthogonal and are considered

230   probably un-separable. If the reconstructed components exhibit w-correlation zero, it signifies

231   the two components to be separable. Large values of w-correlations between reconstructed

232   components indicate that the components should possibly be gathered into one group and

233   correspond to the same component in SSA decomposition (Hassani, 2007).

234   **4   Results**

235   The average value of radon for a period of six month from April 2017-September 2017

236   was found to be in the range 55-117 kBq/m^3. The average emanation of soil-gas radon and the

237   standard deviation at MPGO from April to September 2017 is illustrated in Table 1.

238   Simultaneously variation of soil temperature and pressure with radon emanation was observed.

239   Usually, radon shows positive correlation with temperature i.e. the soil radon concentration

240   increases with increase in temperature and decreases as temperature decrease. The correlation

241   coefficient (Pearson correlation) between radon and temperature is found to be 0.5 signifying

242   positive correlation, while the correlation coefficient between radon and pressure is found to

243   be -0.5 signifying negative correlation with an average temperature and pressure of 28.60 ^{0}C

244   and 991.03 mbar respectively during afore mentioned period. The positive correlation of radon

245   with temperature might be due to the rise in diffusion rate with temperature (Sharma et al.

246   2000; Singh et al. 2008). The negative correlation coefficient was found for soil radon and

247   pressure which signify, with the increase in pressure, the radon emanation decreases while with

248   the decrease in pressure the radon emanation increases. In general, the negative correlation is

249   due to the diffusion process which slows down with increase in pressure, which in turn

250   decreases the radon concentration in the soil. The average value of pressure, temperature,

251   standard deviation, coefficient of variation and correlation coefficient for the observation

252   period is detailed in Table 2. The maximum and minimum temperature observed was 31.24 ^{0}C

253   and 23.78 ^{0}C i.e. a change of 7.46 ^{0}C during the period of observation. Simultaneously, the

254   maximum and minimum pressure during the period of observation was 999.32 mbar and 980.72

255   mbar i.e. a change of 18.6 mbar. Digital filter (Butterworth) is applied to eliminate the high

256   frequency quasi periodic components from the soil radon time series for better discernibility of

257   seismic induced anomalies and is represented in Figure 2.

258   The covariance matrix of the first 9 group of soil radon (Rn-222) time series is

259   represented in Figure 3.The singular value decomposition (SVD) to Rn-222 data evinced that

260   first 9 eigenfunctions (Figure 4) when grouped resulted for 99.90 % of the total variance in the

261   individual time series. The eigenfuction group 1 and 2 represents the aperiodic component and

11

262 group 3 to 9 represented periodic components. It has been earlier studied that the periodic

263 components of radon data extracted by SSA consisted of diurnal and semidiurnal components

264 since the components included variation of 24-hour and 12- hour period (Kumar et al., 2015).

265 The decomposed eigenvectors in soil radon time series is grouped into two classes as diurnal

266 and semidiurnal variation. The sum of eigenfuction group 1 and 2 accounts for 98.62 % and

267 group 3 to 9 accounts for 0.48 % of the variance. Radon variations is governed by daily

268 variations, which accounts to 99.90 % of the total variance in soil gas radon at the MPGO,

269 Tezpur. The Principal Component of soil radon (Rn-222) related to the first 9 grouping of

270 eigentriples is represented in Figure 5 and w-correlation matrix for the 9 reconstructed

271 components is represented in Figure 6.

272 The covariance matrix of the first 9 group of soil temperature and pressure time series

273 is represented in Figure 7 and Figure 11 respectively. The decomposed eigenfunctions for soil

274 temperature and pressure time series by applying SSA is represented in Figure 8 and Figure 12

275 respectively. The first 9 eigenfunction group from SVD to atmospheric temperature records

276 accounts nearly about 99.99 % and first 9 eigenfunction group of soil pressure 100 % of the

277 variation respectively. It is discernible that first eigenfunction 1 alone itself is capable of

278 producing 100 % variation but the time series is well modeled when first 9 eigenfunction is

279 grouped. This evinces SVD to radon, soil temperature and soil pressure data evince that first

280 9 eigenfunction accounts >98.00 % of the variance of the individual data series. This is

281 fascinating to observe that sum of first 9 eigenfunction is fairly sufficient to reproduce the

282 prominent features of the overall variation. This also suggests that the most significantly

283 produced variations are mostly free from naturally induced variations. Soil pressure variations

284 are dominated by semidiurnal variations by 100 % of the total variation in atmospheric pressure

285 at MPGO, Tezpur. On the other hand soil temperature variations is dominated by daily

286 variations like radon variation which account to 97.99 % whereas soil pressure accounts for

287   100 % of the total variance. This suggest that daily variations of soil radon (Rn-222) emanation

288   are controlled by soil pressure in MPGO, Tezpur during the investigation period followed by

289   soil temperature. The Principal Component of soil temperature related to the first 9 grouping

290   of eigentriples is represented in Figure 9 and w-correlation matrix for the 9 reconstructed

291   components is represented in Figure 10. The Principal Component of soil pressure related to

292   the first 9 grouping of eigentriples is represented in Figure 13 and w-correlation matrix for the

293   9 reconstructed components is represented in Figure 14. The reconstructed time series for soil

294   radon (Rn-222), temperature and pressure is represented is Figure 15 and it's residual in Figure

295   16. It is also discernible that the during the investigation time period the pressure change was

296   more than temperature change which also evinces the variation of soil radon at MPGO, Tezpur

297   was more governed by pressure followed by temperature change. The quasi periodic, diurnal

298   (periodic mostly due to temperature) and semidiurnal (aperiodic mostly due to pressure) were

299   eliminated by Singular Spectrum Analysis in the reconstructed time series by grouping and

300   analyzing the eigenfunctions and principle component of individual time series of Rn-222,

301   temperature and pressure respectively. The soil pressure and temperature were found to be

302   negatively correlated to each other (-0.62) which produces a pseudo effect in the soil radon

303   time series. The grouping and reconstruction of the time series also eliminates these pseudo

304   effect arising due soil pressure and temperature.

305   **5   Discussion**

306   The reconstructed soil radon time series along with the seismic activity during the

307   investigation period is shown in Figure 17. The hypocentral parameters of the earthquake

308   events found to have correlation with soil radon emanation is listed in Table 3. The earthquakes

309   are selected from a catalogue which follow Stationary Poisson process where we have adopted

310   declustering method by Reasenberg (Reasenberg, 1985). After declustering of the database, we

311   have used the maximum likelihood algorithm provided by ZMAP (Wiemer, 2001) to perform

13

312 a Gutenberg-Richter regression (Gutenberg and Richter, 1944). Here, the magnitude range of

313 earthquakes catalog indicate relatively high magnitude of completeness (MC~4.7) for the

314 region. It was observed that there were certain positive amplitude rise anomalies in radon

315 emanation prior to six out of 9 earthquakes which occurred in the vicinity (100 km radially

316 from MPGO,Tezpur). Increase in the soil radon concentration is generally assigned to

317 developments of microcracks, fractures and fissure caused by dilatancy prior to earthquake.

318 This process enhances the transportation of radon from its original enclosure following the

319 cracks. The rise in soil radon concentration prior to an earthquake may be due to the strain

320 buildup processes in the area. During this process, very small fractures are developed in the

321 rocks which enhances the contribution of radon gas to the soil near the surface of earth

322 (Miklavčić et al., 2008). Three earthquake events were preceded by negative anomalies. The

323 negative anomalies might be due to the circumstance that during the final stage of dilatency

324 model prior to an earthquake the Rn-222 emanation can be stable or it can decrease (Tomer,

325 2016). This is because, during the final stage prior an earthquake, rupture occurs and fluid

326 pressure and stress on rocks is released (Bakhmutov and Groza, 2008). The fluid pressure

327 increases may result in water level rise (pore spaces probably gets filled by water and diffusion

328 rate of radon is more in void than in water filled pore space; Nielson et al., 1984; Merolla et

329 al., 2003) and this probably does not allow the soil gas Rn-222 to escape from the surface which

330 in turn reduces or stabilizes the emanation of Rn-222. Further a decreasing radon anomaly as

331 observed in this study may be the result of squeezing effect of compressional stress built up in

332 the rock, which in turn result in soil porosity changes at micro scale. There were certain events

333 which occurred on the same day or just a very short seismic gap of 1 or 2 days. Here in the case

334 earthquake with higher magnitude might also be the probable reason for the anomalous

335 behavior of the soil gas radon emanation, as for spatio-temporal clustered earthquakes, the

336 largest magnitude earthquake is presumed to precede the anomalies in radon emanation

337 (Hartmann and Leavy, 2005). Positive as well as negative anomalies were observed prior to 9

338 events which occurred in the vicinity (100 km radially from MPGO) with in a short span of

339 time.

340 **6   Conclusion**

341   The investigation concludes that digital filter assists in eliminating the high frequency

342 quasi periodic components from the time series. The SSA method helps eliminating the diurnal

343 and semidiurnal fluctuations from soil radon time series for improved correlating and detection

344 of the soil radon emanation with seismic activity. The investigation also evinced that radon is

345 dominated by daily variation at MPGO,Tezpur and is controlled by soil pressure followed by

346 temperature. It is also concluded that principle component analysis helps in removing the

347 pseudo effect pertaining to simultaneous soil pressure and temperature effect. It was observed

348 that there were certain positive amplitude rise anomalies in radon emanation prior to six events

349 out of 9 earthquakes which occurred in the vicinity (100 km radially from MPGO,Tezpur)

350 within a short span of time. The increase of radon emanation with temperature might be the

351 result of increasing diffusion rate with temperature. Three earthquake events were preceded by

352 negative anomalies. The negative anomalies might be due to the circumstance that during the

353 final stage of dilatency model prior to an earthquake the soil gas radon emanation can be stable

354 or it can decrease. This is because, during the final stage prior an earthquake, rupture occurs

355 and fluid pressure and stress on rocks is released. Further a decreasing radon anomaly as

356 observed in this study may be the result of squeezing effect of compressional stress built up in

357 rock, which in turn changes porosity of soil at micro scale.
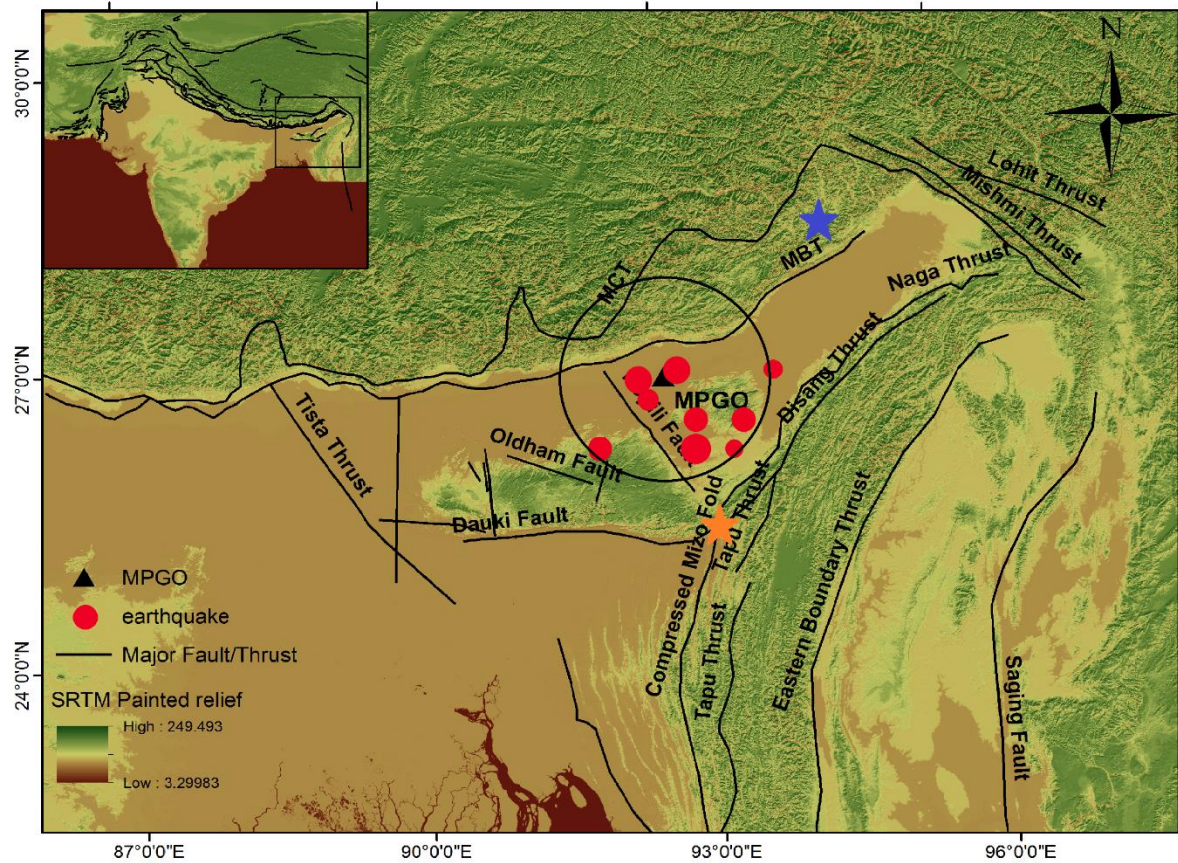
References

363 Adams, R. D.: The Haicheng, China, earthquake of 4 February 1975; the first successfully
364   predicted major earthquake, Earthquake Engineering & Structural Dynamics, 4, 423-
365   437, 1976.

366 Alexandrov, T., and Golyandina, N.: The automatic extraction of time series trend and
367   periodical components with the help of the Caterpillar-SSA approach, Exponenta Pro,
368   3, 54-61, 2004.

369 Bakhmutov, V., and Groza, A.: The dilatancy-diffusion model: new prospects, Proceedings,
370   International Conference: Problems of Geocosmos, 7th, St. Petersburg, Russia, May,
371   2008, 26-30.

372 Bhattacharya, P. M., Mukhopadhyay, S., Majumdar, R., and Kayal, J.: 3-D seismic structure
373   of the northeast India region and its implications for local and regional tectonics,
374   Journal of Asian Earth Sciences, 33, 25-41, 2008.

375 Bhattacharya, P. M., Kayal, J., Baruah, S., and Arefiev, S.: Earthquake source zones in
376   northeast India: seismic tomography, fractal dimension and b value mapping, in:
377   Seismogenesis and Earthquake Forecasting: The Frank Evison Volume II, Springer,
378   145-158, 2010.

379 Chetia, T., Baruah, S., Dey, C., Sharma, S., and Baruah, S.: Multi-parametric approach for
380   Earthquake Precursor Detection in Assam Valley (Eastern Himalaya, India) using
381   Satellite and Ground Observation Data, Geotectonics, In Press, 2019.

382 Chetia, T., Baruah, S., Dey, C., Sharma, S., and Baruah, S.: Probabilistic analysis of seismic
383   data for earthquake forecast in North-East India and its vicinity, Current Science, In
384   Press, 2019.

385 Choubey, V., Kumar, N., and Arora, B.: Precursory signatures in the radon and geohydrological
386   borehole data for M4. 9 Kharsali earthquake of Garhwal Himalaya, Science of the total
387   environment, 407, 5877-5883, 2009.

388 Demmel, J., and Veselić, K.: Jacobi's method is more accurate than QR, SIAM Journal on
389   Matrix Analysis and Applications, 13, 1204-1245, 1992.

390 Fraedrich, K.: Estimating the dimensions of weather and climate attractors, Journal of the
391   atmospheric sciences, 43, 419-432, 1986.

392 Golyandina, N., Nekrutkin, V., and Solntsev, V.: „Caterpillar"-SSA Technique for Analysis of
393   Time Series in Economics, New Models of Business: Managerial Aspects and Enabling
394   Technology, 198, 2001.

395 Golyandina, N., Nekrutkin, V., and Zhigljavsky, A. A.: Analysis of time series structure: SSA
396   and related techniques, Chapman and Hall/CRC, 2001.

397 Golyandina, N.: On the choice of parameters in singular spectrum analysis and related
398   subspace-based methods, arXiv preprint arXiv:1005.4374, 2010.

399 Gupta, H.: Medium-term earthquake prediction, Eos, Transactions American Geophysical
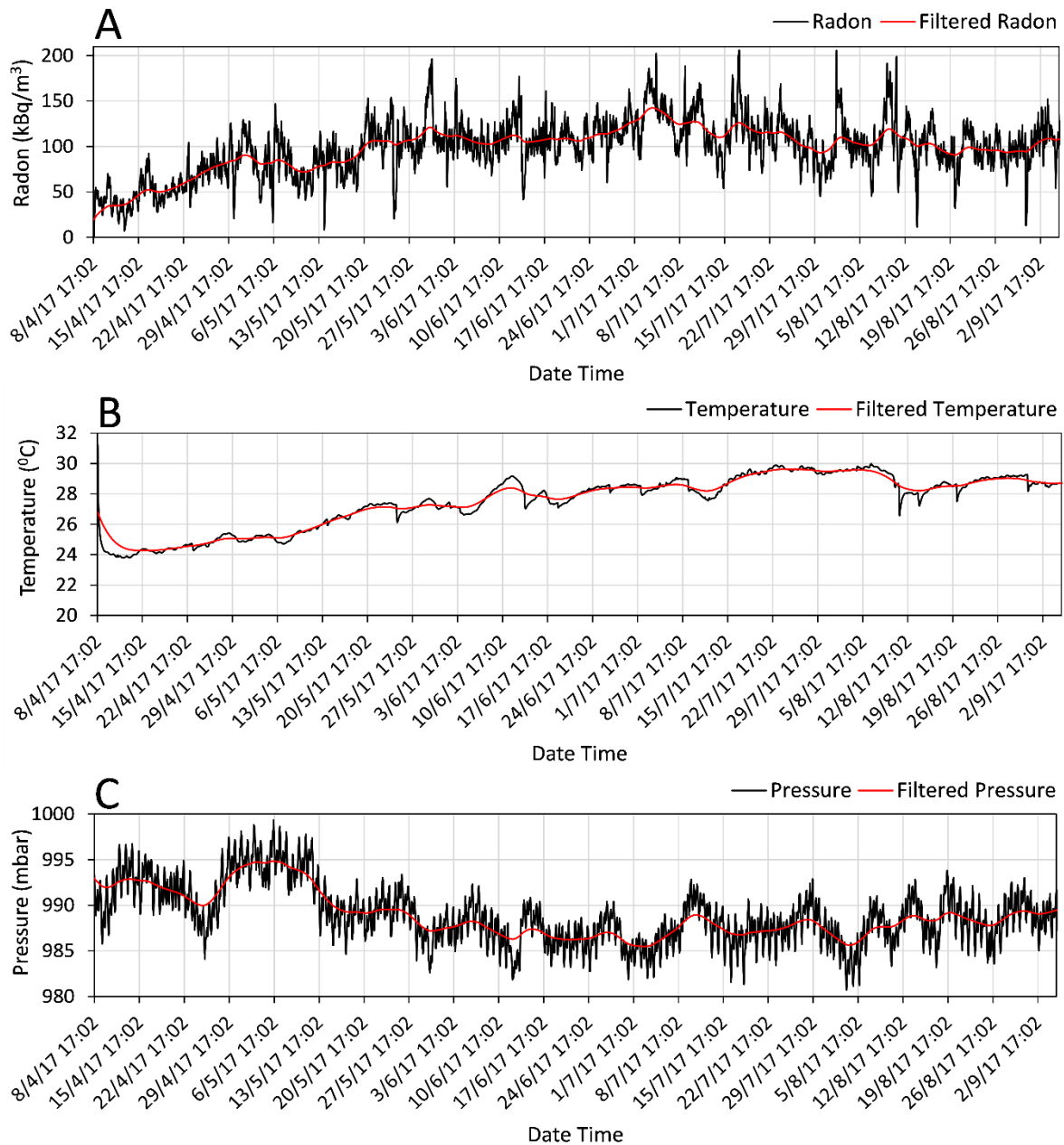400   Union, 69, 1620-1630, 1988.

401 Gutenberg, B., and Richter, C. F.: Frequency of earthquakes in California, Bulletin of the
402   Seismological society of America, 34, 185-188, 1944.

403 Gutenberg, B.: The energy of earthquakes, Quarterly Journal of the Geological Society, 112,
404   1-14, 1956.

405 Hartmann, J., and Levy, J. K.: Hydrogeological and gasgeochemical earthquake precursors–A
406   review for application, Natural Hazards, 34, 279-304, 2005.

407 Hassani, H.: Singular spectrum analysis: methodology and comparison, 2007.

408 Kayal, J., Arefiev, S., Barua, S., Hazarika, D., Gogoi, N., Kumar, A., Chowdhury, S., and
409   Kalita, S.: Shillong plateau earthquakes in northeast India region: complex tectonic
410   model, CURRENT SCIENCE-BANGALORE-, 91, 109, 2006.

411 Kayal, J.: Microearthquake seismology and seismotectonics of South Asia, Springer Science
412   & Business Media, 2008.

413 Kayal, J., Arefiev, S. S., Baruah, S., Tatevossian, R., Gogoi, N., Sanoujam, M., Gautam, J.,
414   Hazarika, D., and Borah, D.: The 2009 Bhutan and Assam felt earthquakes (Mw 6.3
415   and 5.1) at the Kopili fault in the northeast Himalaya region, Geomatics, Natural
416   Hazards and Risk, 1, 273-281, 2010.

417 Kayal, J., Arefiev, S., Baruah, S., Hazarika, D., Gogoi, N., Gautam, J., Baruah, S., Dorbath, C.,
418   and Tatevossian, R.: Large and great earthquakes in the Shillong plateau–Assam valley
419   area of Northeast India Region: Pop-up and transverse tectonics, Tectonophysics, 532,
420   186-192, 2012.

421 Khan, M. A. R., and Poskitt, D.: Window length selection and signal-noise separation and
422   reconstruction in singular spectrum analysis, Monash Econometrics and Business
423   Statistics Working Papers, 23, 2011-2023, 2011.

424 Khattri, K., and Tyagi, A.: Seismicity patterns in the Himalayan plate boundary and
425   identification of the areas of high seismic potential, Tectonophysics, 96, 281-297, 1983.

426 Kumar, A., Walia, V., Arora, B. R., Yang, T. F., Lin, S.-J., Fu, C.-C., Chen, C.-H., and Wen,
427   K.-L.: Identifications and removal of diurnal and semidiurnal variations in radon time
428   series data of Hsinhua monitoring station in SW Taiwan using singular spectrum
429   analysis, Natural Hazards, 79, 317-330, 2015.

430 Merolla, P., Mose, D., and King, A.: Effect of water table fluctuations on radon emanation
431   from soil, 2003.

432 Miklavčić, I., Radolić, V., Vuković, B., Poje, M., Varga, M., Stanić, D., and Planinić, J.: Radon
433   anomaly in soil gas as an earthquake precursor, Applied radiation and isotopes, 66,
434   1459-1466, 2008.

435 Nandy, D., and Dasgupta, S.: Seismotectonic domains of northeastern India and adjacent areas,
436   Physics and Chemistry of the Earth, 18, 371-384, 1991.

437 Nandy, D.: Geodynamics of Northeastern India and the adjoining region, ACB publications,
438   2001.

17

439 Ni, J., and Barazangi, M.: Seismotectonics of the Himalayan collision zone: Geometry of the
440   underthrusting Indian plate beneath the Himalaya, Journal of Geophysical Research:
441   Solid Earth, 89, 1147-1163, 1984.

442 Nielson, K., Rogers, V., and Gee, G.: Diffusion of Radon through Soils: A Pore Distribution
443   Model 1, Soil Science Society of America Journal, 48, 482-487, 1984.

444 Reasenberg, P.: Second-order moment of central California seismicity, 1969–1982, Journal of
445   Geophysical Research: Solid Earth, 90, 5479-5495, 1985.

446 Serita, A., Hattori, K., Yoshino, C., Hayakawa, M., and Isezaki, N.: Principal component
447   analysis and singular spectrum analysis of ULF geomagnetic data associated with
448   earthquakes, Natural Hazards and Earth System Science, 5, 685-689, 2005.

449 Sharma, A. K., Walia, V., and Virk, H.: Effect of meteorological parameters on radon
450   emanation at Palampur (HP), Journal of Association of Exploration Geophysicists, 21,
451   47-50, 2000.

452 Singh, S., Kumar, A., Bajwa, B. S., Mahajan, S., Kumar, V., and Dhar, S.: Radon monitoring
453   in soil gas and ground water for earthquake prediction studies in North West Himalayas,
454   India, Terr. Atmos. Ocean. Sci, 21, 685-695, 2010.

455 Stranden, E., Kolstad, A., and Lind, B.: The influence of moisture and temperature on radon
456   exhalation, Radiation Protection Dosimetry, 7, 55-58, 1984.

457 Tomer, A.: Radon as a earthquake precursor: a review, Int J Sci Eng Technol, 4, 815-822, 2016.

458 Ulomov, V., and Mavashev, B.: The Tashkent Earthquake of 26 April, Tashkent: Acad. Nauk.
459   Uzbeck, 1971.

460 Verma, M., and Bansal, B. K.: Earthquake precursory studies in India: Scenario and future
461   perspectives, Journal of Asian Earth Sciences, 54, 1-8, 2012.

462 Virk, H., and Singh, B.: Radon recording of Uttarkashi earthquake, Geophysical Research
463   Letters, 21, 737-740, 1994.

464 Walia, V., Su, T., Fu, C., and Yang, T.: Spatial variations of radon and helium concentrations
465   in soil-gas across the Shan-Chiao fault, Northern Taiwan, Radiation Measurements, 40,
466   513-516, 2005.

467 Wiemer, S.: A software package to analyze seismicity: ZMAP, Seismological Research Letters,
468   72, 373-382, 2001.

469 Yang, T., Walia, V., Chyi, L., Fu, C., Chen, C.-H., Liu, T., Song, S., Lee, C., and Lee, M.:
470   Variations of soil radon and thoron concentrations in a fault zone and prospective
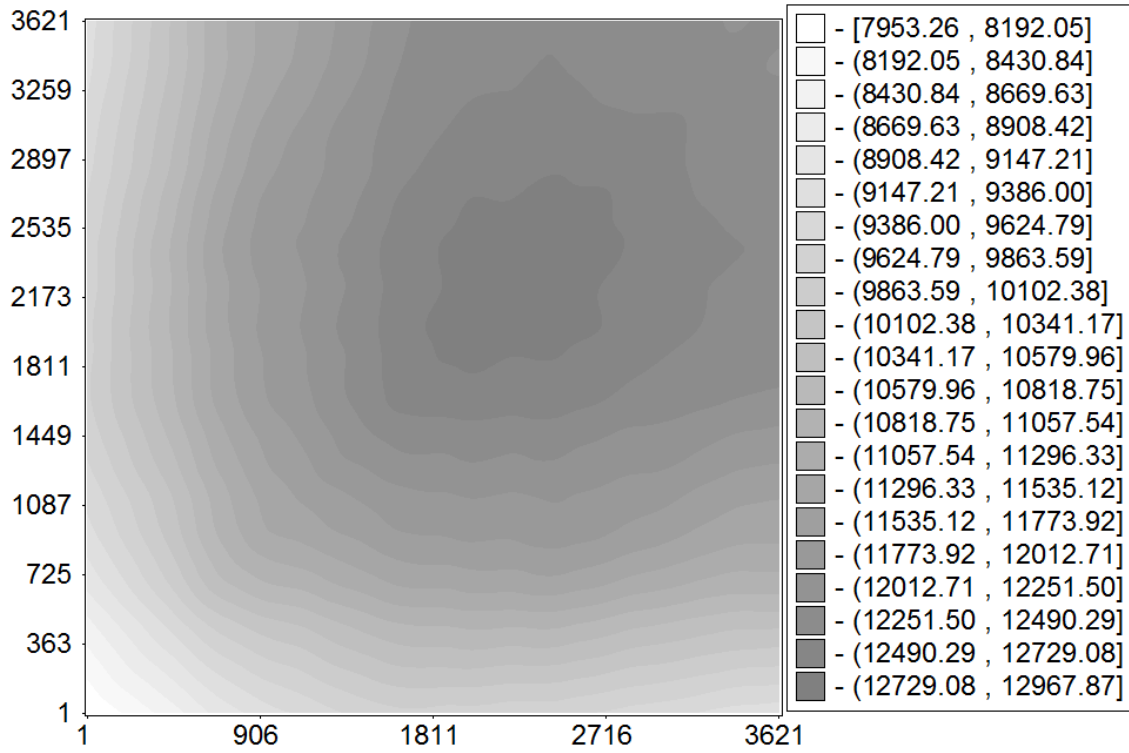471   earthquakes in SW Taiwan, Radiation Measurements, 40, 496-502, 2005.

472

473

474

475

476   **FIGURES AND TABLES**
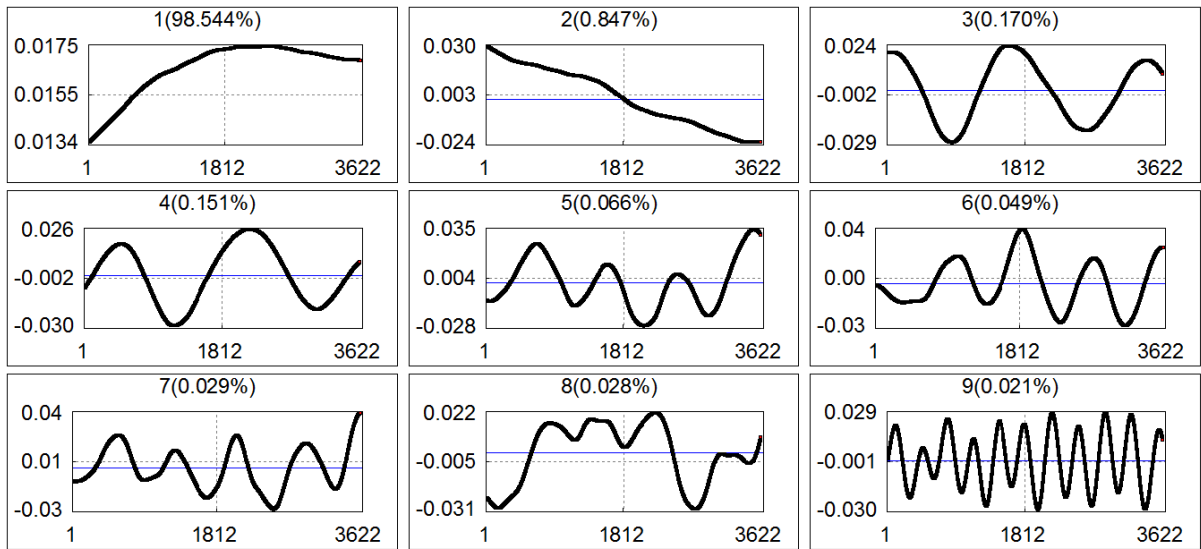
477

478

479

480

481   **Figure 1:** Map illustrates the main tectonic features of North-East (after Kayal, 2006), India

482   along with the study area represented by black color circle and the earthquake events (3.1 ≥ mb

483   ≤ 3.7) which were found to have correlation with soil radon emanation anomalies are

484   represented by the red color dots. The black color triangle represents the Multiparametric

485   Geophysical Observatory. The blue color star represents the earthquake of Mw~7.2, 1947 and

486   orange color star represents the earthquake of Mw~7.7, 1869.

487

488

489

490

Figure 2: The plot represents the removal of high frequency quasi periodic component for A) filtered time series of soil radon, B) filtered time series of soil temperature and C) filtered time series of soil pressure.

491

492

493

494

Figure 3: Covariance matrix of the first 9 group of soil radon (Rn-222) time series.

Figure 4: Eigenfunctions of soil radon (Rn-222) first 9 group. The scale in the horizontal x-axis represents the window length ($L*15$ minutes) and the scale in the vertical y-axis represents the singular eigenvectors (matrix path SVD).
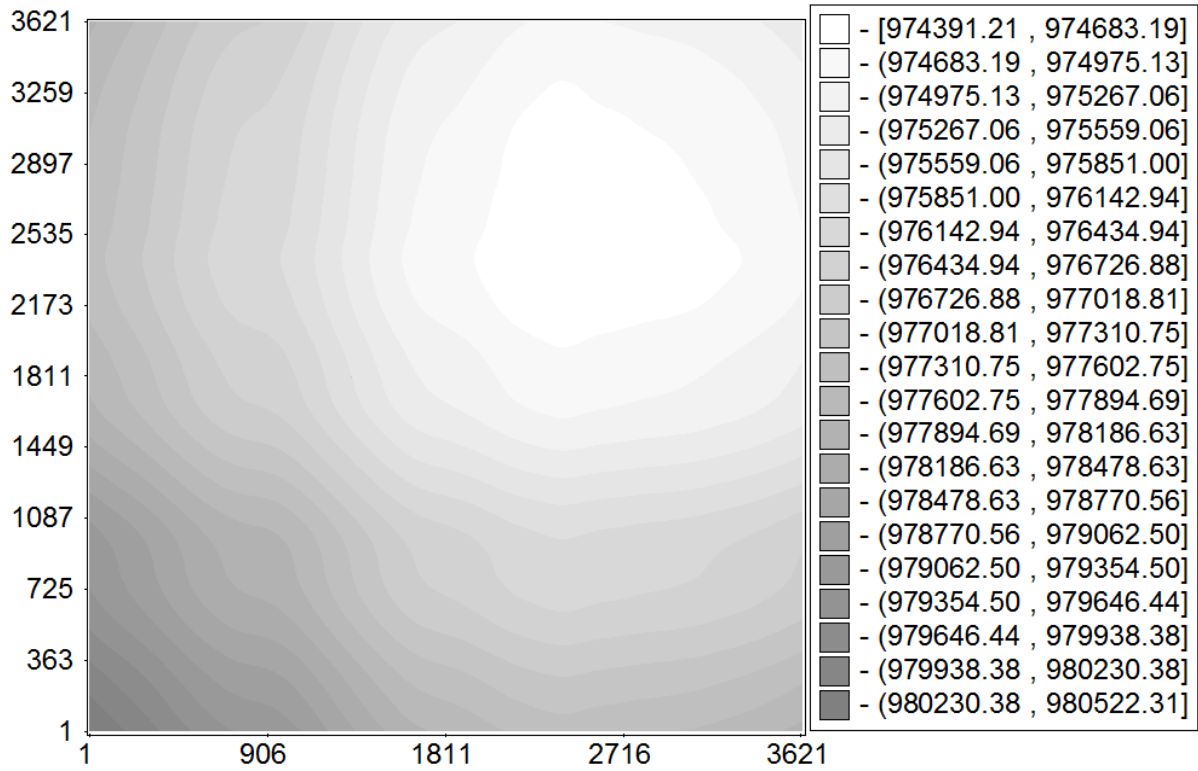
502

Figure 5: Principal Component of soil radon (Rn-222) related to the first 9 eigentriples. The scale in the horizontal x-axis represents the window length ($L*15$ minutes) and the scale in vertical y-axis represents the Principal Component (PC).
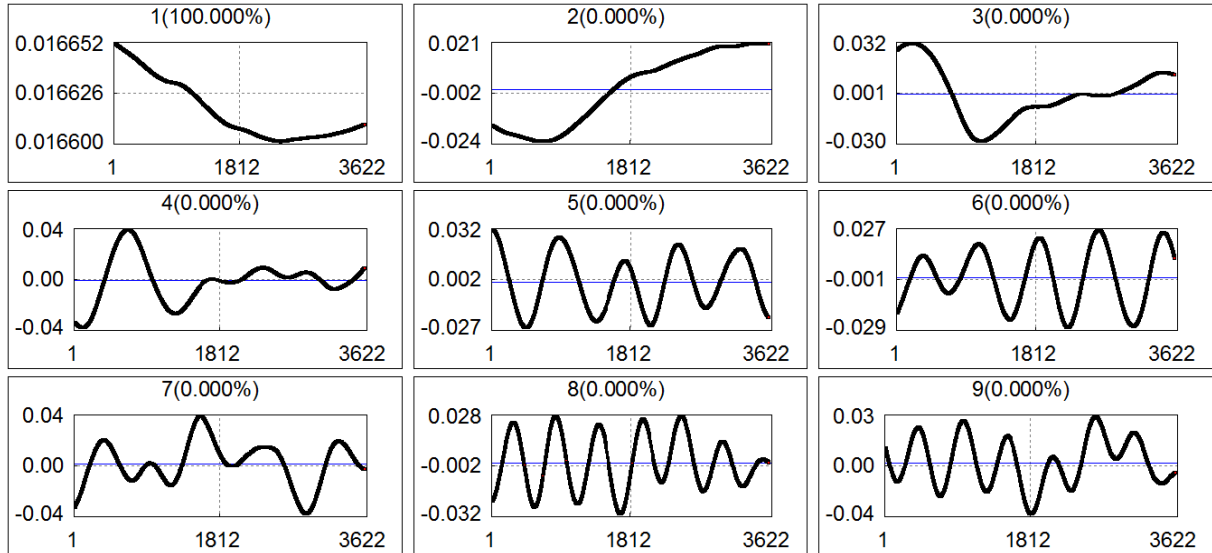

506

Figure 6: w-correlation matrix for the 9 reconstructed components of soil radon time series.

508

Figure 7: Covariance matrix of the first 9 group of soil temperature time series.

Figure 8: Eigenfunctions of soil temperature first 9 group. The scale in the horizontal x-axis represents the window length ($L*15$ minutes) and the scale in the vertical y-axis represents the singular eigenvectors (matrix path SVD).

Figure 9: Principal Component of soil temperature related to the first 9 eigentriples. The scale in the horizontal x-axis represents the window length (*L**15 minutes) and the scale in the vertical y-axis represents the Principal Component (PC).

Figure 10: w-Correlation matrix for the 9 reconstructed components of soil temperature time series.
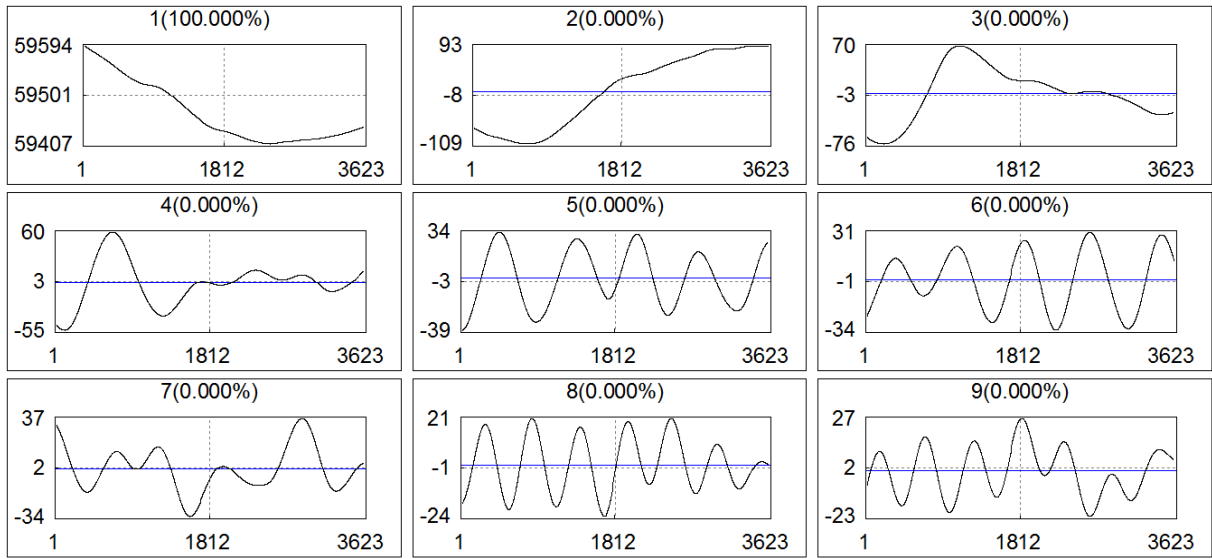
Figure 11: Covariance matrix of the first 9 group of soil pressure time series
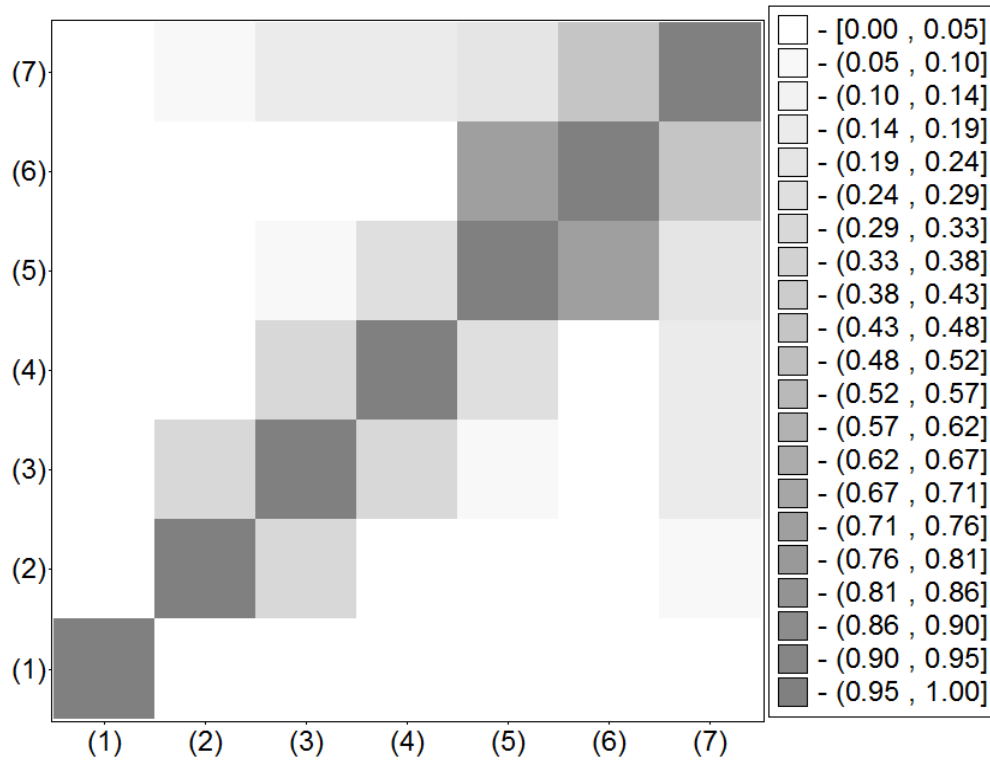

Figure 12: Eigenfunctions of soil pressure first 9 group. The scale in the horizontal x-axis represents the window length ($L*15$ minutes) and the scale in the vertical y-axis represents the singular eigenvectors (matrix path SVD). .
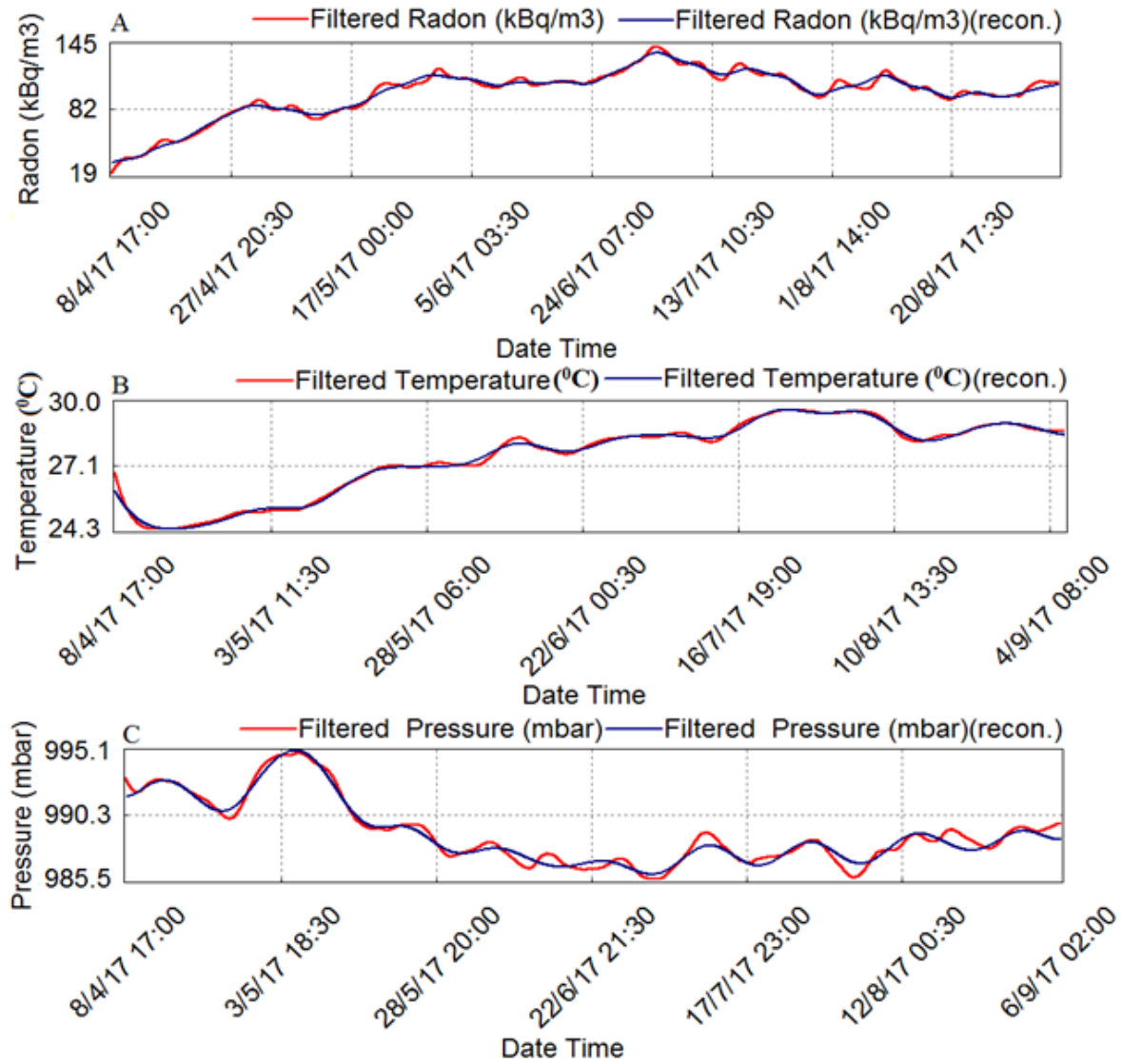
530

531

Figure 13: Principal Component of soil temperature related to the first 9 Eigentriples. The scale in the horizontal x-axis represents the window length (*L**15 minutes) and scale in the vertical y-axis represents the Principal Component (PC).
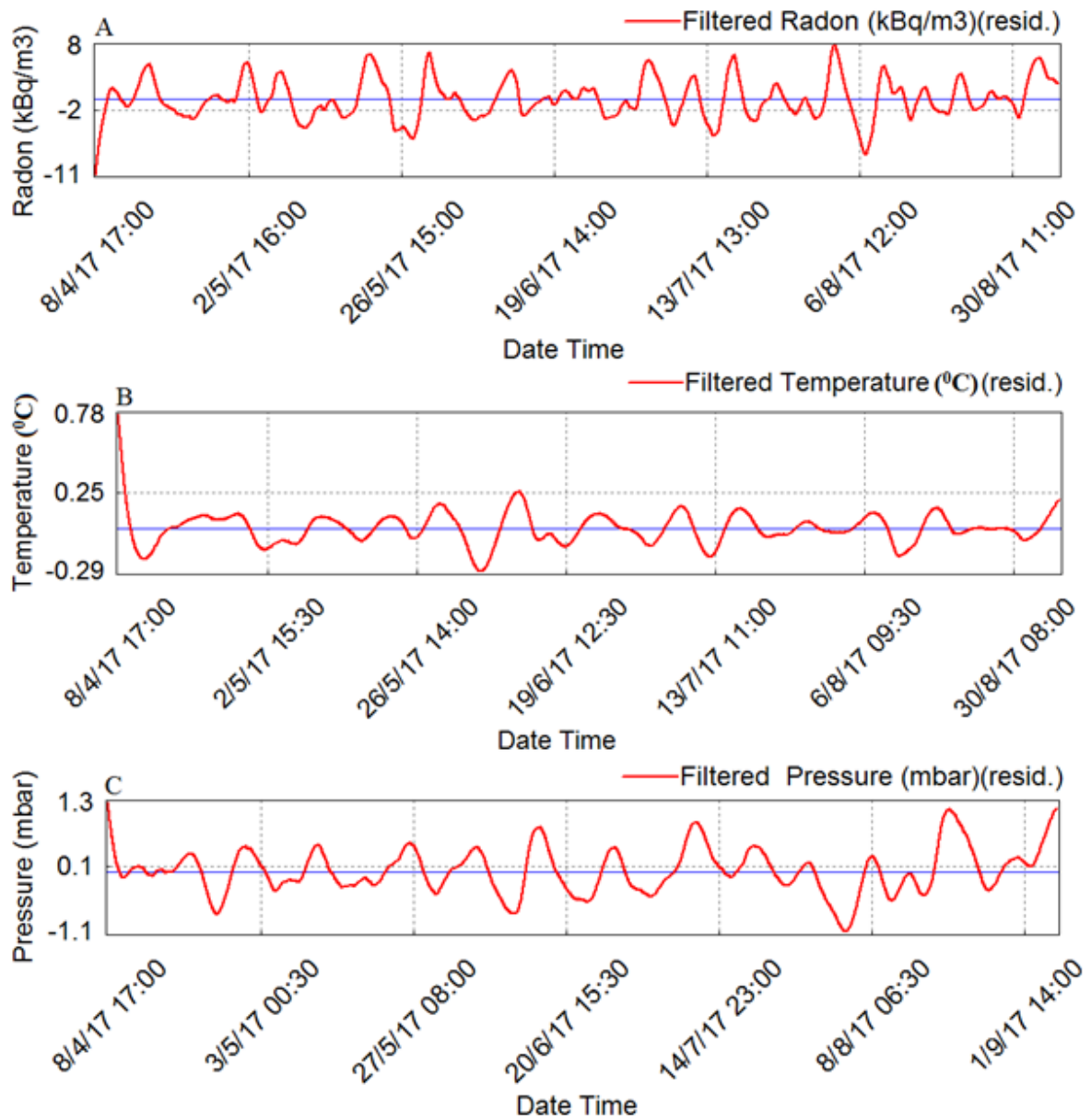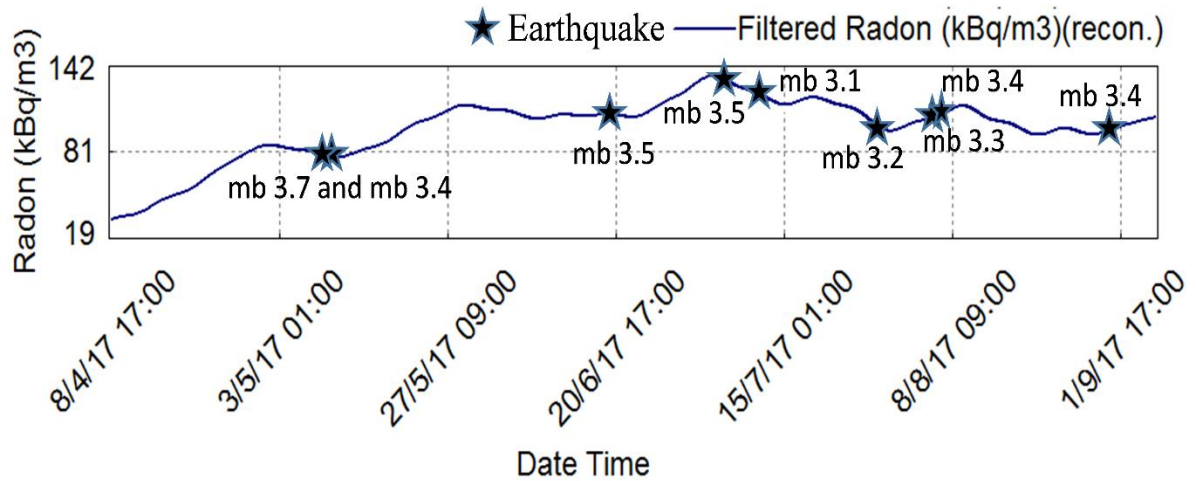

535

Figure 14: w correlation matrix for the 9 reconstructed components of soil pressure time series.

537

538

539

Figure 15: Plot showing the reconstructed time series of A) filtered soil radon, B) filtered

temperature and C) filtered pressure.

542

543

544

Figure 16: Residual of reconstructed time series of A) Filtered soil radon, B) temperature and C) pressure respectively.

Figure 17: Plot showing the reconstructed filtered time series of soil radon emanation along with earthquake during the investigation period in the vicinity of MPGO, Tezpur (~100 km radially from MPGO) which occurred in a very short span of time.

565 **Table 1:** The average emanation of soil-gas radon and the standard deviation at MPGO from

566 April to September, 2017.

Month (2017)	Average Rn-222 emanation (kBq /m3)	Standard Deviation (kBq /m3)
April	55.94	21.3
May	93.11	28.53
June	109.12	19.07
July	117.69	28.09
August	101.45	25.86
September	92.34	18.65

567

568

569 **Table 2:** The correlation co-efficient of soil radon gas concentration with soil pressure and

570 temperature at OH-MPGO during year 2017.

Parameters	Average (Avg.)	Standard Deviation (Std.)	Coefficient of Variation (Std./Avg. %)	Correlation Coefficient
Radon (KBq/m³)	94.94	23.58	24.84	----
Temperature (^{0}C)	28.60	0.62	2.19	0.51
Pressure (mbar)	991.03	2.48	0.25	-0.52

571

572

573

574 **Table 3:** Hypocentral parameters of the earthquake events found to have correlation with radon

575 anomaly.

Date of Event	UTC TIME	Lat (ºN)	Long (ºE)	Place	Depth (km)	Mag (m$_b$)	Distance from MPGO (km)	Type of Anomaly (+ or -)
9/5/17	01:53:55	26.3	92.7	Assam	25	3.7	44	+
9/5/17	03:26:54	26.6	93.2	Assam	28	3.4	67	+
20/6/17	04:31:58	27.1	92.5	West Kameng, Arunachal Pradesh	10	3.5	67	+
4/7/17	10:05:47	27.0	92.1	West Kameng, Arunachal Pradesh	10	3.5	78	-
10/7/17	23:28:30	27.1	93.8	Papumpare, Arunachal Pradesh	10	3.1	78	-
25/7/17	18:28:00	26.3	93.1	Karbi Anglong, Assam	28	3.2	67	-
5/8/17	12:24:56	26.8	92.2	Darrang, Assam	10	3.3	44	+
7/8/17	11:25:07	26.3	91.7	Kamrup, Assam	10	3.4	100	+
31/8/17	17:57:26	26.6	92.7	Sonitpur Assam	10	3.4	67	+

576

577

578

579

580

581