

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 Baruah¹, Chandan Dey^{1,2}, Sangeeta Sharma¹, Santanu Baruah¹

5 ¹Geoscience & Technology Division, CSIR-North East Institute of Science and Technology
6 (CSIR-NEIST), Jorhat-785006, Assam, India

7 ²Academy of Scientific and Innovative Research (CSIR-NEIST), Jorhat-785006, Assam, India

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

9

10 **Abstract.** 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 frequency quasi periodic component in the time series of soil radon emanation.
18 It was scrutinized that sum of just 9 eigenfunctions were sufficient enough for reproducing the
19 prominent characteristics of the overall variation. This perhaps also evinces that more
20 significantly produced fluctuations are mostly free from natural variations. The variations in
21 soil temperature was observed to be dominated by daily variations similar to radon variation
22 which account to 97.99 % whereas soil pressure accounts for 100 % of the total variance which
23 suggests that daily variations of soil radon (Rn-222) emanation are controlled by soil pressure
24 in MPGO, Tezpur during the investigation period followed by soil temperature. The study
25 concludes that SSA eliminates diurnal and semidiurnal components from time series of soil
26 radon emanation for better correlation of soil radon emanation with earthquakes.

27 1 **Introduction**

28 Radon (Rn-222) is a noble gas, a decay product of radium with atomic number, $Z=86$
29 and a half-life of nearly 3.8 days. Because of its short decay time, amount changes in its
30 production from the rock is quite evidenced. Ulomov and Mavashev in the year 1971 (Ulomov
31 and Mavashev, 1971) first suggested the correlation of radon concentration with earthquakes.
32 It has been scrutinized that the radon concentration has correlation to earthquakes and volcanic
33 eruptions (Walia et al. 2006, Singh et al. 2005). Significant radon concentration anomalies were
34 also observed prior to the Uttarkashi earthquake of 20th October, 1991; $m_b \sim 6.5$ (Virk and Singh
35 1994). Radon concentrations was monitored in the North West Himalaya for earthquake
36 prediction studies and empirical equation between earthquake magnitude, epicentral distance
37 and precursory time were examined (Kumar et al. 2009). Earthquake prediction depending
38 entirely on precursory phenomena is empirical and comprises many applied problems. Among
39 various geophysical parameters soil radon is preferred and used for earthquake precursory
40 studies because of its ease of detectability. Radon in nature has different isotopes: Rn-222 (half-
41 life~3.8 days), Rn-220 (Thoron, half-life~54.5 s) and Rn-219 (half-life~3.92 s). The utmost
42 prominent is Rn-222 because of its longer half-life which is a product of Ra-226 decay process.
43 The Rn-222 emanates from soil or crust either by diffusion or convection and reaches the
44 atmosphere. The soil radon emanation concentration is generally assigned to developments of
45 micro-cracks, fissure and fracture due to dilatancy prior to earthquake. This process enhances
46 the transportation of Radon from its original enclosure following the cracks into the
47 atmosphere. Significant radon concentration anomalies were also observed prior to the
48 Uttarkashi earthquake of 20th October, 1991; $m_b \sim 6.5$ (Virk and Singh, 1994).

49 North-East India (NE India) is highly vulnerable to earthquake and lies in seismic zone
50 V (BIS 2002) and frequent occurrence of earthquake facilitates the probability of finding
51 precursory phenomena which may lead to a successful prediction in near future. With this

52 objective a Multiparametric Geophysical Observatory (MPGO) in Ouguri Hills (Latitude
53 N26.61°; Longitude E92.78°, Elevation~82m), Tezpur, Assam, India with the installation of
54 several geophysical instruments collecting data simultaneously, portray an opportunity towards
55 identification of precursory signatures prior to earthquakes. Earthquake precursory and
56 prediction studies advanced in late 1970s and Heicheng earthquake of 4th February is a land-
57 mark which was in short-term successfully predicted in 1975 in China (Adams, 1976). The
58 accomplished medium term forecast of M~7.5 earthquake on 6th August, 1988 in northeast
59 Indian region (Gupta and Singh, 1988) encouraged to strengthen such studies in India. Another
60 successful short term prediction was done of ($M \geq 4$) in Koyna region of western India, famous
61 for Reservoir Triggered Seismicity (Verma and Bansal, 2012).

62 The study tries to correlate radon emanation in soil gas with earthquakes within the
63 epicentral distance of 100 km of $m_b > 3.1$ from MPGO, Tezpur which is situated in a highly
64 tectonically strained and seismically active region. The reason for considering the criteria of
65 100 km with $m_b > 3.1$ is the tectonic setting of the study area which is bounded by two major
66 fault namely Kopili and the Bomdila Faults. The region is highly stressed because of
67 geotectonic settings where earthquake do occur periodically within very short span of time.
68 Kopili fault zone is experiencing compressional stress due to the Indo-Burma arc and Himalyan
69 arc to the east and the north respectively which is characterized by transverse tectonics. The
70 Bomdila Fault inter-weaves across three major tectonic domains of the NE-India, namely
71 MCT, MBT and Naga-Disang thrust along NW-SE direction. (Kayal, 2010). The empirical
72 relation for the energy release E to the magnitude (Gutenberg, 1956) gives $E > 10^{9.2}$ for
73 earthquake of $m_b > 3.1$. Stress and strain are directly proportional and hence in order to obtain
74 adequate stress build up processes for observing precursory signals in temporal soil Radon
75 emanation the criteria of 100 Km with $m_b > 3.1$ is taken into consideration. The major problem
76 is the removal of quasi periodic, diurnal (mostly due to temperature) and semidiurnal (mostly

77 due to pressure) components (Kumar et al., 2015). Radon anomalies are governed by seismic
78 activity as well as by meteorological parameters (soil temperature, pressure, rainfall, moisture
79 and even wind for atmospheric radon (Stranden et al., 1984; Kumar et al., 2009; Walia et al.,
80 2005) which in turn makes it more complex for identifying the seismic induced anomalies.
81 Here in the investigation characteristics features of temporal soil radon concentrations
82 variability at MPMGO, Tezpur is scrutinized by applying singular spectrum analysis (SSA) in
83 concern to the objective of filtering the meteorological parameters on radon concentration. SSA
84 is a relatively innovative and powerful advanced method which has been used across many
85 applied problems for different scientific fields (e.g., Fraedrich, 1986, Serita et al., 2005).
86 Seismic induced emanation of soil gas radon is a complex phenomenon to be studies based on
87 mathematical, statistical analysis and modellings. Nonlinear methods such as power spectrum,
88 fractal-multifractal analysis and singular spectrum analysis etc. probably are commanding tools
89 which reveals the nonlinear features and mechanism of soil gas radon emanation processes.
90 The major advantages of Caterpillar-SSA method is that it does not require better
91 understanding of time series parametric model and extensively applicable to varied spectrum
92 of diverse temporal data. It can also be applied to match with non-stationary temporal data and
93 also permits to estimate structures even for short temporal data. Here Caterpillar-SSA method
94 is used to study the temporal soil gas radon and the foremost concept of SSA is applying PCA
95 on trajectory matrix acquired from the original time series following time series reconstruction.

96 **2 Seismotectonics of the region**

97 In middle of the active Kopili and Bomdila fault, the MPMGO is situated. The Kopili
98 and Bomdila faults comprise Neogene-Quaternary sediments, which directly were deposited
99 over the Archean basement. The Kopili fault zone in an approx is 100 km in width and 300 km
100 in length. It is a NW-SE trending strike-slip fault (Kayal et al., 2006, Bhattacharya et al., 2008,
101 2010). The two Precambrian massifs - the Shillong Plateau and the Mikir Hills is delineated by

102 the tectonic disposition of the Kopili fault. MPGO is bounded to the north by the Main
103 Boundary Thrust (MBT) and to the south by the NE-SW trending Belt of Schuppen (Figure 1).
104 The Bomdila fault is strike slip fault of about 400 km in length which trends along WNW-ESE
105 direction. The northern portion of the fault mostly lies in the Gondwana, Paleogene and
106 Neogene sediments. This fault is surrounded by the Belt of Schuppen to the east and south by
107 the Mikir massif and to the west. The fault cuts across the Himalayan fold belt towards the
108 north (Nandy and Dasgupta, 1991).

109 The Kopili Fault has produced two large earthquakes (Figure 1) $M_w \sim 7.7$, 1869 event
110 (Figure 1) in the south eastern edge of the fault contravening the Naga-Disang thrust and
111 $M_w \sim 7.2$, 1943 earthquake which occurred farther to the north of 1869 event within a period of
112 nearly 75 years (Kayal, 2008). It is highly active with strong seismicity discernible down up to
113 depth of about 50 km, and which extends to the Main Central Thrust (MCT) in the Bhutan
114 Himalaya. Even if MCT is dormant (Ni and Barazangi, 1984), intense activity is observed at
115 the region where Kopili Fault meets the MBT and MCT (Nandy, 2001, Kayal et al., 2010,
116 2012). This is demonstrated by the August 19, 2009 Earthquake ($M_w \sim 5.1$) in the Assam Valley
117 that occurred in the center of the Kopili fault zone and the September 21, 2009 strong Bhutan
118 Himalaya earthquake ($M_w \sim 6.3$) that occurred at the northern end of the Kopili fault where it
119 connects with the MCT (Kayal et al., 2012). Both the earthquakes are shallow focus (depth ~
120 10 km) showing right lateral strike-slip faulting (Kayal et al., 2010) which suggests that the
121 Kopili fault zone is experiencing compressional stress due to the Indo-Burma arc and Himalyan
122 arc to the east and the north respectively which is characterized by transverse tectonics. The
123 Bomdila Fault inter-weaves across three major tectonic domains of the NE-India, namely
124 MCT, MBT and Naga-Disang thrust along NW-SE direction. The earthquake events along
125 Bomdila fault occur in a diffused pattern having post-collisional intracratonic characteristics
126 (Nandy and Dasgupta, 1991). It is observed that, the Upper Brahmaputra Valley stretching

127 between the Bomdila Fault and almost near NW-trending Mishmi Thrust in the northeast is
128 seismically dormant, and is recognized as the Assam Gap (Khattri, 1983).

129 **3 Method and techniques**

130 **3.1 Soil radon (Rn-222) time series**

131 A BMC2 barasol manufactured by Algade is into operation for soil radon emanation
132 time series data in MPGO for earthquake prediction and precursory studies. Soil gas radon
133 emanation every 15 minutes is being continuously monitored. The barasol probe is kept fitted
134 inside a plastic tube (length 1.5m and diameter of 0.0635m) with open bottom dug inside the
135 ground in such a way that the detection unit (detector sensitivity-0.02 pulses/h for 1 Bq/m³ and
136 saturation volumetric activity-3MBq/ m³) which is at the bottom of the probe lies 1m from the
137 ground level. A silicon alpha detector detects the radon gas which enters the detection chamber
138 when it emanates from the soil. The radon pass in a detection volume over three cellulose filters
139 which allows to trap all the solid daughter products of radon. The sensor is a fixed silicon
140 detector with a depleted depth of 100 μ m and 400 mm² of sensitive area. It performs the
141 counting by atomic spectrometry of radon (Rn-222) and daughter products created in the
142 detection volume (with window customized between 1.5 MeV and 6 MeV). The probe system
143 is embedded with soil pressure and temperature sensor. The sensor calibration permits the
144 volumic activity of the radon (Rn-222) to be evaluated. In the present investigation the soil
145 radon emanation temporal variability at MPGO, Tezpur the radon data were prudently checked
146 for no gaps or discontinuous jump. Digital filter (Butterworth) is applied to eliminate the high
147 frequency quasi periodic components form the soil radon time series for better discernibility of
148 seismic induced anomalies.

149 **3.2 Singular Spectrum Analysis**

150 The SSA results and graphs in the investigation are acquired by using Caterpillar-SSA
 151 3.40 software (Alexandrov and Golyandina, 2004). Window selection rule applied to the time
 152 series is one half of the length of the time series to meet the theoretical requirements for the
 153 investigation (Golyandina, 2010, Khan, 2011, Hassani, 2007). The singular value
 154 decomposition (SVD) algorithm was applied as it is more accurate than QR iteration which
 155 are the common most algorithm for solving of eigenvalues and singular value problems
 156 (Demmel, 1992). The main objective of SSA is decomposing the original time series into sum
 157 of series such that each of the component in this sum can be known (either as a trend, periodic
 158 or quasi-periodic components) or noise. This is accomplished by decomposition and
 159 reconstruction. At the first the time series is decomposed following the reconstruction of the
 160 original time series (which is without noise). The methodology adopted here is first to
 161 embedding a 1-dimension time series say, $Y_T = (y_1, \dots, y_T)$ into a multi-dimensional time series
 162 X_1, \dots, X_K having vectors $X_i = (y_i, \dots, y_{i+L-1}) \in R^L$ (Golyandina et al., 2001, 2001). Here the value
 163 of $K = T - L + 1$. The X_i Vectors are called L -lagged vectors. The embedding depends on the
 164 window length L , such that $2 \leq L \leq T$ which results for trajectory matrix (Hankel matrix:
 165 diagonal elements $i + j = \text{const.}$ are equal) $X = [X_1, \dots, X_K] = (X_{ij})_{i,j=1}^{L,K}$. Secondly the Singular
 166 Value Decomposition (SVD) of the trajectory matrix is performed to represent it as an addition
 167 of bi-orthogonal elementary matrices having rank one. Represented by $\lambda_1, \dots, \lambda_L$ which are the
 168 Eigen-values of XX' in a descending order of magnitude ($\lambda_1 \geq \dots \lambda_L \geq 0$) and by U_1, \dots, U_L which
 169 are the orthonormal system (i.e. $(U_i, U_j) = 0$ for $i \neq j$) is the orthogonality property and $\|U_i\|=1$,
 170 of the eigenvectors of the matrix XX' corresponding to these eigenvalues. (U_i, U_j) is the inner
 171 product of the vectors U_i and U_j and $\|U_i\|$ is the norm of the vector U_i . The Set

$$172 \quad d = \max (i, \text{ such that } \lambda_i > 0) = \text{rank } X \quad (\text{I})$$

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

$$175 \quad X = X_1 + \dots + X_d \quad (II)$$

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

177 Thirdly the series reconstruction is accomplished by grouping, to split the elementary
 178 matrices (X_i) into various groups and addition of the matrices within each and every group. The
 179 splitting of the elementary matrices X_i into several groups and summation of the matrices
 180 within individual group is grouping. Say $I = (i_1, \dots, i_p)$ is a group of indices i_1, \dots, i_p . Now
 181 for the group I of corresponding matrix X_I is defined as $X_I = X_{i_1} + \dots + X_{i_p}$. The splitting
 182 of the of the set with indices $J = 1, \dots, d$ on the disjoint subsets I_1, \dots, I_m resembles to the
 183 illustration

$$184 \quad X = X_{I_1} + \dots + X_{I_m}. \quad (III)$$

185 This procedure of selecting the sets I_1, \dots, I_m is termed as eigentriple grouping. For given
 186 group I the contribution of the component X_I into $X_{I_1} + \dots + X_{I_m}$ is estimated to the corresponding
 187 eigenvalues as: $\sum_{i \in I} \lambda_i / \sum_{i=1}^d \lambda_i$.

188 Finally diagonally averaging transfers each I matrix into a time series, which is an
 189 additive component of the initial series Y_T . If z_{ij} is an element of the matrix Z . Then k-th term
 190 of the produced series is acquired by averaging z_{ij} for i, j such that $i + j = k + 2$. This procedure
 191 is known as diagonal averaging (Hankelization) of the matrix Z . The Hankelization of a matrix
 192 Z is the Hankel matrix HZ which is the trajectory matrix corresponding to the series obtained
 193 from diagonal averaging. Hankelization is a best logical technique that HZ matrix is the nearest
 194 to Z among all corresponding size of Hankel matrices (Golyandina et al., 2001). Now applying
 195 the Hankelization technique to each and every matrix components in equation III, we get

REVISED VERSION

196
$$X = \tilde{X}_{I_1} + \dots + \tilde{X}_{I_m} \quad (IV)$$

197 Here $X_{II} = HX$ corresponding to initial series $Y_T = (y_1, \dots, y_T)$ decomposition into a
 198 sum of m series as

199
$$y_t = \sum_{k=1}^m y_t^{(k)} \quad (V)$$

200 Here $y_t^{(k)} = y_1^{(k)} + \dots + y_T^{(k)}$ corresponds matrix X_{Ij} .

201 Covariance matrix is a matrix of covariances between the values $X(t)$ and $X(t+k)$., where
 202 k is a lag. When k is positive $X(t)$ and $X(t+k)$ tends to fluctuate together. The covariances among
 203 $X(t)$ and $X(t+k)$ can be estimated as

204
$$COV = \frac{\sum_{i=1}^n (X(t) - \overline{x(t)}) (X(t+k) - \overline{x(t+k)})}{n-1} \quad (VI)$$

205 Where $\overline{x(t)}$ and $\overline{x(t+k)}$ represents the mean of $x(t)$ and $x(t+k)$ correspondingly.
 206 Now covariance matrix S is estimated as

207
$$S = \frac{1}{n-1} \sum_{i=1}^n (X(t) - \bar{X}) (X(t+k) - \bar{X}) \quad (VII)$$

208 Here $k=0, \dots, M-1$ where M is the window size

209 From the eigenvectors of the covariance matrix measured at different lags the principal
 210 components of the time series is estimated. The principal components is also a time series
 211 having same length as the “embedded” time series. The computation of principal component is
 212 from simple matrix product as

213
$$PC = S * \text{matrix of eigenvectors} \quad (VIII)$$

214 Where S is covariance matrix.

215 The exemplification of how discretely different component can be separated from each
 216 other is based on studying the SSA properties. To successfully decompose (SSA) the series Y_T
 217 the subsequent additive components of the series should be almost separable from each other.
 218 The weighted correlation is also known as *w-correlation* which is the natural measurement of
 219 dependencies among two series Y_T^1 and Y_T^2 given as:

$$220 \quad \rho_{12}^\omega = \frac{(Y_T^1, Y_T^2)_\omega}{\|Y_T^1\|_\omega \|Y_T^2\|_\omega} \quad (\text{IX})$$

221 Here, $\|Y_T^1\|_\omega = \sqrt{(Y_T^1, Y_T^1)_\omega}$, $(Y_T^i, Y_T^j)_\omega = \sum_{k=1}^T \omega_k y_k^i y_k^j$, ($i, j = 1, 2$), $\omega_k = \min\{k, T -$
 222 $k\}$ while assuming $L \leq T/2$.

223 If the absolute value of *w*-correlations is small, the corresponding series are almost *w*-
 224 orthogonal. On the other hand if it is large, the series are not *w*-orthogonal and are considered
 225 probably un-separable. If the reconstructed components exhibit *w*-correlation zero, it signifies
 226 the two components to be separable. Large values of *w*-correlations between reconstructed
 227 components indicate that the components should possibly be gathered into one group and
 228 correspond to the same component in SSA decomposition (Hassani, 2007).

229 **4 Results**

230 The average value of radon for a period of six month from April 2017-September 2017
 231 was found to be in the range 55-117 kBq/m³. The average emanation of soil-gas radon at
 232 MPGO for April, May, June, July, August and September 2017 is reported to be 55.94, 93.11,
 233 109.12, 117.69, 101.45, 92.34 (kBq/m³) respectively with standard deviation (Std.) of 21.3,
 234 28.53, 19.07, 28.09, 25.86, 18.65 (kBq /m³) respectively. Simultaneously variation of soil
 235 temperature and pressure with radon emanation was observed. Usually, radon shows positive
 236 correlation with temperature i.e. the soil radon concentration increases with increase in

237 temperature and decreases as temperature decrease. The correlation coefficient (Pearson
238 correlation) between radon and temperature is found to be 0.5 signifying positive correlation,
239 while the correlation coefficient between radon and pressure is found to be -0.5 signifying
240 negative correlation with an average temperature and pressure of 28.60 °C and 991.03 mbar
241 respectively during afore mentioned period. The positive correlation of radon with temperature
242 might be due to the rise in diffusion rate with temperature (Sharma et al. 2000, Singh et al.
243 2008). The negative correlation coefficient was found for soil radon and pressure which signify,
244 with the increase in pressure, the radon emanation decreases while with the decrease in pressure
245 the radon emanation increases. In general, the negative correlation is due to the diffusion
246 process which slows down with increase in pressure, which in turn decreases the radon
247 concentration in the soil. The average value of pressure, temperature, standard deviation,
248 coefficient of variation and correlation coefficient for the observation period is detailed in
249 Table 1. The maximum and minimum temperature observed was 31.24 °C and 23.78 °C i.e. a
250 change of 7.46 °C during the period of observation. Simultaneously, the maximum and
251 minimum pressure during the period of observation was 999.32 mbar and 980.72 mbar i.e. a
252 change of 18.6 mbar. Digital filter (Butterworth) is applied to eliminate the high frequency
253 quasi periodic components from the soil radon time series for better discernibility of seismic
254 induced anomalies and is represented in Figure 2.

255 The covariance matrix of the first 9 group of soil radon (Rn-222) time series is
256 represented in Figure 3. The singular value decomposition (SVD) to Rn-222 data evinced that
257 first 9 eigenfunctions (Figure 4) when grouped resulted for 99.90 % of the total variance in the
258 individual time series. The eigenfunction group 1 and 2 represents the aperiodic component and
259 group 3 to 9 represented periodic components. The periodic and aperiodic component mostly
260 corresponds to diurnal and semidiurnal variation (Kumar et al., 2015). The decomposed
261 eigenvectors in soil radon time series is grouped into two classes as diurnal and semidiurnal

262 variation. The sum of eigenfunction group 1 and 2 accounts for 98.62 % and group 3 to 9
263 accounts for 0.48 % of the variance. Radon variations is governed by daily variations, which
264 accounts to 99.90 % of the total variance in soil gas radon at the MPGO, Tezpur. The Principal
265 Component of soil radon (Rn-222) related to the first 9 grouping of eigentriples is represented
266 in Figure 5 and w-correlation matrix for the 9 reconstructed components is represented in
267 Figure 6.

268 The covariance matrix of the first 9 group of soil temperature and pressure time series
269 is represented in Figure 7 and Figure 11 respectively. The decomposed eigenfunctions for soil
270 temperature and pressure time series by applying SSA is represented in Figure 8 and Figure 12
271 respectively. The first 9 eigenfunction group from SVD to atmospheric temperature records
272 accounts nearly about 99.99 % and first 9 eigenfunction group of soil pressure 100 % of the
273 variation respectively. It is discernible that first eigenfunction 1 alone itself is capable of
274 producing 100 % variation but the time series is well modeled when first 9 eigenfunction is
275 grouped. This evinces SVD to radon, soil temperature and soil pressure data evince that first
276 9 eigenfunction accounts >98.00 % of the variance of the individual data series. This is
277 fascinating to observe that sum of first 9 eigenfunction is fairly sufficient to reproduce the
278 prominent features of the overall variation. This also suggests that the most significantly
279 produced variations are mostly free from naturally induced variations. Soil pressure variations
280 are dominated by semidiurnal variations by 100 % of the total variation in atmospheric pressure
281 at MPGO, Tezpur. On the other hand soil temperature variations is dominated by daily
282 variations like radon variation which account to 97.99 % whereas soil pressure accounts for
283 100 % of the total variance. This suggest that daily variations of soil radon (Rn-222) emanation
284 are controlled by soil pressure in MPGO, Tezpur during the investigation period followed by
285 soil temperature. The Principal Component of soil temperature related to the first 9 grouping
286 of eigentriples is represented in Figure 9 and w-correlation matrix for the 9 reconstructed

287 components is represented in Figure 10. The Principal Component of soil pressure related to
288 the first 9 grouping of eigentriples is represented in Figure 13 and w-correlation matrix for the
289 9 reconstructed components is represented in Figure 14. The reconstructed time series for soil
290 radon (Rn-222), temperature and pressure is represented in Figure 15 and its residual in Figure
291 16. It is also discernible that during the investigation time period the pressure change was
292 more than temperature change which also evinces the variation of soil radon at MPGO, Tezpur
293 was more governed by pressure followed by temperature change. The quasi periodic, diurnal
294 (periodic mostly due to temperature) and semidiurnal (aperiodic mostly due to pressure) were
295 eliminated by Singular Spectrum Analysis in the reconstructed time series by grouping and
296 analyzing the eigenfunctions and principle component of individual time series of Rn-222,
297 temperature and pressure respectively. The soil pressure and temperature were found to be
298 negatively correlated to each other (-0.62) which produces a pseudo effect in the soil radon
299 time series. The grouping and reconstruction of the time series also eliminates these pseudo
300 effect arising due soil pressure and temperature.

301 **5 Discussion**

302 The reconstructed soil radon time series along with the seismic activity during the
303 investigation period is shown in Figure 17. The hypocentral parameters of the earthquake
304 events found to have correlation with soil radon emanation is listed in Table 2. It was observed
305 that there were certain positive amplitude rise anomalies in radon emanation prior to six out of
306 9 earthquakes which occurred in the vicinity (100 km radially from MPGO, Tezpur). Increase
307 in the soil radon concentration is generally assigned to developments of microcracks, fractures
308 and fissure caused by dilatancy prior to earthquake. This process enhances the transportation
309 of radon from its original enclosure following the cracks. The rise in soil radon concentration
310 prior to an earthquake may be due to the strain buildup processes in the area. During this
311 process, very small fractures are developed in the rocks which enhances the contribution of

312 radon gas to the soil near the surface of earth (Miklavčić et al., 2008). Three earthquake events
313 were preceded by negative anomalies. The negative anomalies might be due to the
314 circumstance that during the final stage of dilatency model prior to an earthquake the Rn-222
315 emanation can be stable or it can decrease (Tomer, 2016). This is because, during the final
316 stage prior an earthquake, rupture occurs and fluid pressure and stress on rocks is released
317 (Bakhmutov and Groza, 2008). The fluid pressure increases may result in water level rise (pore
318 spaces probably gets filled by water and diffusion rate of radon is more in void than in water
319 filled pore space; Nielson et al., 1984; Merolla et al., 2003) and this probably does not allow
320 the soil gas Rn-222 to escape from the surface which in turn reduces or stabilizes the emanation
321 of Rn-222. Further a decreasing radon anomaly as observed in this study may be the result of
322 squeezing effect of compressional stress built up in the rock, which in turn result in soil porosity
323 changes at micro scale. There were certain events which occurred on the same day or just a
324 very short seismic gap of 1 or 2 days. Here in the case earthquake with higher magnitude might
325 also be the probable reason for the anomalous behavior of the soil gas radon emanation, as for
326 spatio-temporal clustered earthquakes, the largest magnitude earthquake is presumed to
327 precede the anomalies in radon emanation (Hartmann and Leavy, 2005). Positive as well as
328 negative anomalies were observed prior to 9 events which occurred in the vicinity (100 km
329 radially from MPGO) with in a short span of time.

330 **6 Conclusion**

331 The investigation concludes that digital filter assists in eliminating the high frequency
332 quasi periodic components from the time series. The SSA method helps eliminating the diurnal
333 and semidiurnal fluctuations from soil radon time series for improved correlating and detection
334 of the soil radon emanation with seismic activity. The investigation also evinced that radon is
335 dominated by daily variation at MPGO, Tezpur and is controlled by soil pressure followed by
336 temperature. It is also concluded that principle component analysis helps in removing the

337 pseudo effect pertaining to simultaneous soil pressure and temperature effect. It was observed
338 that there were certain positive amplitude rise anomalies in radon emanation prior to six events
339 out of 9 earthquakes which occurred in the vicinity (100 km radially from MPMO, Tezpur)
340 within a short span of time. The increase of radon emanation with temperature might be the
341 result of increasing diffusion rate with temperature. Three earthquake events were preceded by
342 negative anomalies. The negative anomalies might be due to the circumstance that during the
343 final stage of dilatancy model prior to an earthquake the soil gas radon emanation can be stable
344 or it can decrease. This is because, during the final stage prior an earthquake, rupture occurs
345 and fluid pressure and stress on rocks is released. Further a decreasing radon anomaly as
346 observed in this study may be the result of squeezing effect of compressional stress built up in
347 rock, which in turn changes porosity of soil at micro scale.

348 **Acknowledgements.** We acknowledge our sincere thanks to Ministry of Earth Sciences
349 (MoES) Government of India for providing funds vide project no. MoES/P.O.(Seismo)/NPEP-
350 16/2011. We thank Director, CSIR-NEIST Jorhat for giving necessary permission to publish
351 this paper.

352 **References**

- 353 Adams, R. D.: The Haicheng, China, earthquake of 4 February 1975; the first successfully
354 predicted major earthquake, *Earthquake Engineering & Structural Dynamics*, 4, 423-
355 437, 1976.
- 356 Alexandrov, T., and Golyandina, N.: The automatic extraction of time series trend and
357 periodical components with the help of the Caterpillar-SSA approach, *Exponenta Pro*,
358 3, 54-61, 2004.
- 359 Bakhmutov, V., and Groza, A.: The dilatancy-diffusion model: new prospects, *Proceedings*,
360 *International Conference: Problems of Geocosmos*, 7th, St. Petersburg, Russia, May,
361 2008, 26-30.
- 362 Bhattacharya, P. M., Kayal, J., Baruah, S., and Arefiev, S.: Earthquake source zones in
363 northeast India: seismic tomography, fractal dimension and b value mapping, in:
364 *Seismogenesis and Earthquake Forecasting: The Frank Evison Volume II*, Springer,
365 145-158, 2010.

- 366 Bhattacharya, P. M., Mukhopadhyay, S., Majumdar, R., and Kayal, J.: 3-D seismic structure
 367 of the northeast India region and its implications for local and regional tectonics,
 368 *Journal of Asian Earth Sciences*, 33, 25-41, 2008.
- 369 Demmel, J., and Veselić, K.: Jacobi's method is more accurate than QR, *SIAM Journal on*
 370 *Matrix Analysis and Applications*, 13, 1204-1245, 1992.
- 371 Fraedrich, K.: Estimating the dimensions of weather and climate attractors, *Journal of the*
 372 *atmospheric sciences*, 43, 419-432, 1986.
- 373 Golyandina, N., Nekrutkin, V., and Solntsev, V.: „Caterpillar“-SSA Technique for Analysis of
 374 Time Series in Economics, *New Models of Business: Managerial Aspects and Enabling*
 375 *Technology*, 198, 2001.
- 376 Golyandina, N., Nekrutkin, V., and Zhigljavsky, A. A.: *Analysis of time series structure: SSA*
 377 *and related techniques*, Chapman and Hall/CRC, 2001.
- 378 Golyandina, N.: On the choice of parameters in singular spectrum analysis and related
 379 subspace-based methods, arXiv preprint arXiv:1005.4374, 2010.
- 380 Gupta, H.: Medium-term earthquake prediction, *Eos, Transactions American Geophysical*
 381 *Union*, 69, 1620-1630, 1988.
- 382 Gutenberg, B.: The energy of earthquakes, *Quarterly Journal of the Geological Society*, 112,
 383 1-14, 1956.
- 384 Hartmann, J., and Levy, J. K.: Hydrogeological and gasgeochemical earthquake precursors—A
 385 review for application, *Natural Hazards*, 34, 279-304, 2005.
- 386 Hassani, H.: *Singular spectrum analysis: methodology and comparison*, 2007.
- 387 Kayal, J., Arefiev, S. S., Baruah, S., Tatevossian, R., Gogoi, N., Sanoujam, M., Gautam, J.,
 388 Hazarika, D., and Borah, D.: The 2009 Bhutan and Assam felt earthquakes (Mw 6.3
 389 and 5.1) at the Kopili fault in the northeast Himalaya region, *Geomatics, Natural*
 390 *Hazards and Risk*, 1, 273-281, 2010.
- 391 Kayal, J., Arefiev, S., Barua, S., Hazarika, D., Gogoi, N., Kumar, A., Chowdhury, S., and
 392 Kalita, S.: Shillong plateau earthquakes in northeast India region: complex tectonic
 393 model, *CURRENT SCIENCE-BANGALORE-*, 91, 109, 2006.
- 394 Kayal, J., Arefiev, S., Baruah, S., Hazarika, D., Gogoi, N., Gautam, J., Baruah, S., Dorbath, C.,
 395 and Tatevossian, R.: Large and great earthquakes in the Shillong plateau—Assam valley
 396 area of Northeast India Region: Pop-up and transverse tectonics, *Tectonophysics*, 532,
 397 186-192, 2012.
- 398 Kayal, J.: *Microearthquake seismology and seismotectonics of South Asia*, Springer Science
 399 & Business Media, 2008.
- 400 Khan, M. A. R., and Poskitt, D.: Window length selection and signal-noise separation and
 401 reconstruction in singular spectrum analysis, *Monash Econometrics and Business*
 402 *Statistics Working Papers*, 23, 2011-2023, 2011.
- 403 Khattri, K., and Tyagi, A.: Seismicity patterns in the Himalayan plate boundary and
 404 identification of the areas of high seismic potential, *Tectonophysics*, 96, 281-297, 1983.

REVISED VERSION

- 405 Kumar, A., Walia, V., Arora, B. R., Yang, T. F., Lin, S.-J., Fu, C.-C., Chen, C.-H., and Wen,
406 K.-L.: Identifications and removal of diurnal and semidiurnal variations in radon time
407 series data of Hsinhua monitoring station in SW Taiwan using singular spectrum
408 analysis, *Natural Hazards*, 79, 317-330, 2015.
- 409 MEROLLA, P., MOSE, D., and KING, A.: EFFECT OF WATER TABLE FLUCTUATIONS
410 ON RADON EMANATION FROM SOIL, 2003.
- 411 Miklavčić, I., Radolić, V., Vuković, B., Poje, M., Varga, M., Stanić, D., and Planinić, J.: Radon
412 anomaly in soil gas as an earthquake precursor, *Applied radiation and isotopes*, 66,
413 1459-1466, 2008.
- 414 Nandy, D., and Dasgupta, S.: Seismotectonic domains of northeastern India and adjacent areas,
415 *Physics and Chemistry of the Earth*, 18, 371-384, 1991.
- 416 Nandy, D.: Geodynamics of Northeastern India and the adjoining region, ACB publications,
417 2001.
- 418 Ni, J., and Barazangi, M.: Seismotectonics of the Himalayan collision zone: Geometry of the
419 underthrusting Indian plate beneath the Himalaya, *Journal of Geophysical Research:*
420 *Solid Earth*, 89, 1147-1163, 1984.
- 421 Nielson, K., Rogers, V., and Gee, G.: Diffusion of Radon through Soils: A Pore Distribution
422 Model 1, *Soil Science Society of America Journal*, 48, 482-487, 1984.
- 423 Serita, A., Hattori, K., Yoshino, C., Hayakawa, M., and Isezaki, N.: Principal component
424 analysis and singular spectrum analysis of ULF geomagnetic data associated with
425 earthquakes, *Natural Hazards and Earth System Science*, 5, 685-689, 2005.
- 426 Sharma, A. K., Walia, V., and Virk, H.: Effect of meteorological parameters on radon
427 emanation at Palampur (HP), *Journal of Association of Exploration Geophysicists*, 21,
428 47-50, 2000.
- 429 Singh, H., Singh, J., Singh, S., and Bajwa, B.: Radon exhalation rate and uranium estimation
430 study of some soil and rock samples from Tusham ring complex, India using SSNTD
431 technique, *Radiation Measurements*, 43, S459-S462, 2008.
- 432 Singh, S., Kumar, A., Bajwa, B. S., Mahajan, S., Kumar, V., and Dhar, S.: Radon monitoring
433 in soil gas and ground water for earthquake prediction studies in North West Himalayas,
434 *India, Terr. Atmos. Ocean. Sci*, 21, 685-695, 2010.
- 435 Stranden, E., Kolstad, A., and Lind, B.: The influence of moisture and temperature on radon
436 exhalation, *Radiation Protection Dosimetry*, 7, 55-58, 1984.
- 437 Stranden, E., Kolstad, A., and Lind, B.: The influence of moisture and temperature on radon
438 exhalation, *Radiation Protection Dosimetry*, 7, 55-58, 1984.
- 439 Tomer, A.: Radon as a earthquake precursor: a review, *Int J Sci Eng Technol*, 4, 815-822, 2016.
- 440 Ulomov, V., and Mavashev, B.: The Tashkent Earthquake of 26 April, Tashkent: Acad. Nauk.
441 Uzbek, 1971.
- 442 Verma, M., and Bansal, B. K.: Earthquake precursory studies in India: Scenario and future
443 perspectives, *Journal of Asian Earth Sciences*, 54, 1-8, 2012.

REVISED VERSION

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

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

449

450

451

452

453

454

455

456

457

458

459

460

461

462

463

464

465

466

467

468

469

470

471

472

473

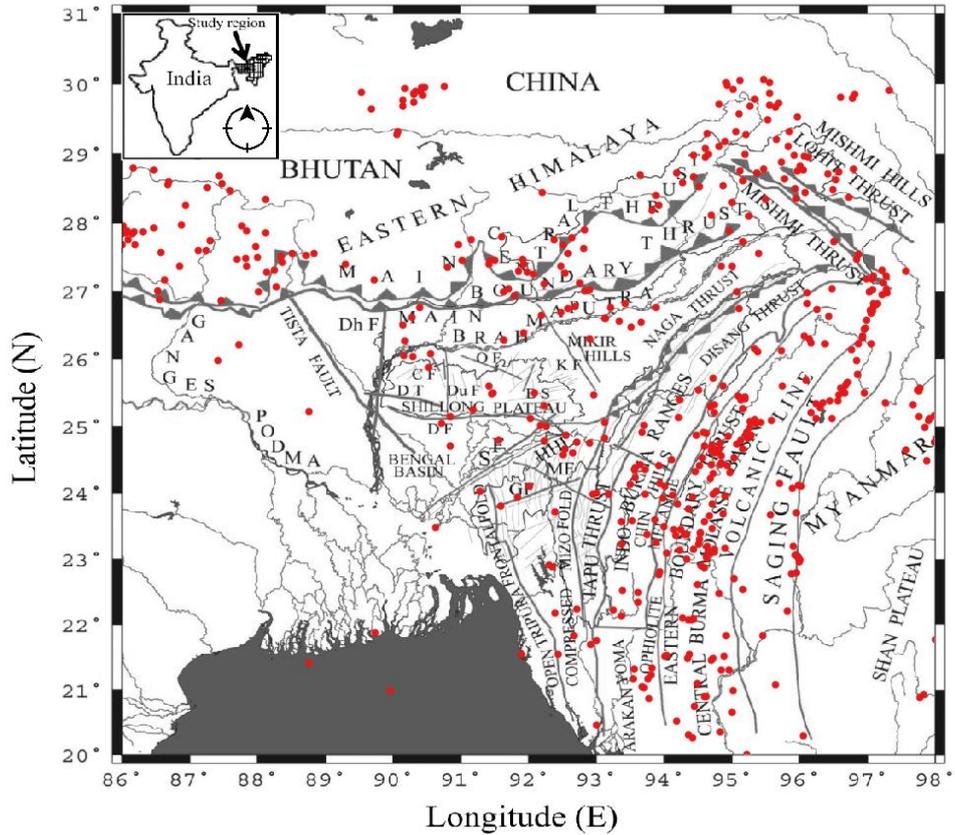
474

475 **FIGURES AND TABLES**

476

477

478



479

480 **Figure 1:** Map illustrates the earthquake events of $M_w \geq 5$ during 1918 to 2018, in NE-India
 481 and its border region (20^0 - 30^0 N and 86^0 - 98^0 E) along with the major tectonic features of the
 482 region.

483

484

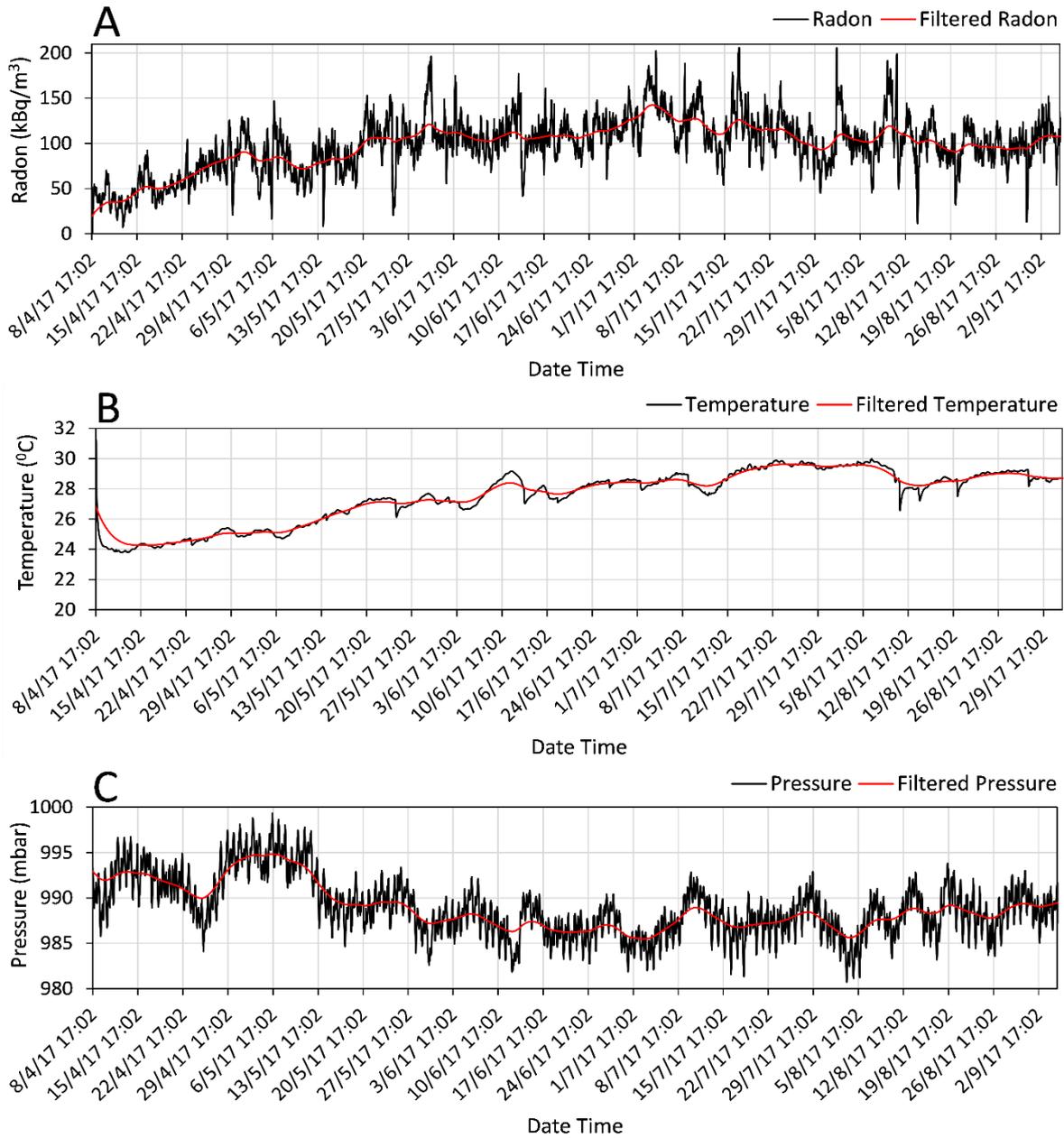
485

486

487

488

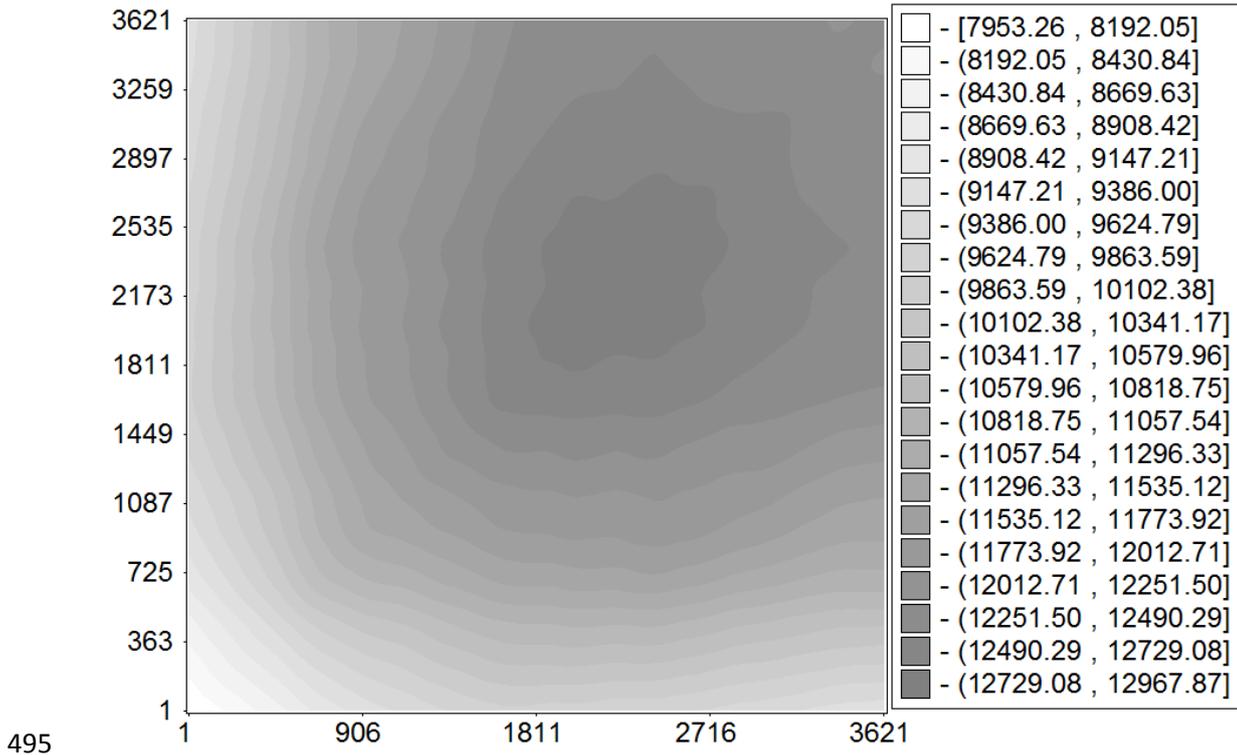
489



490

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

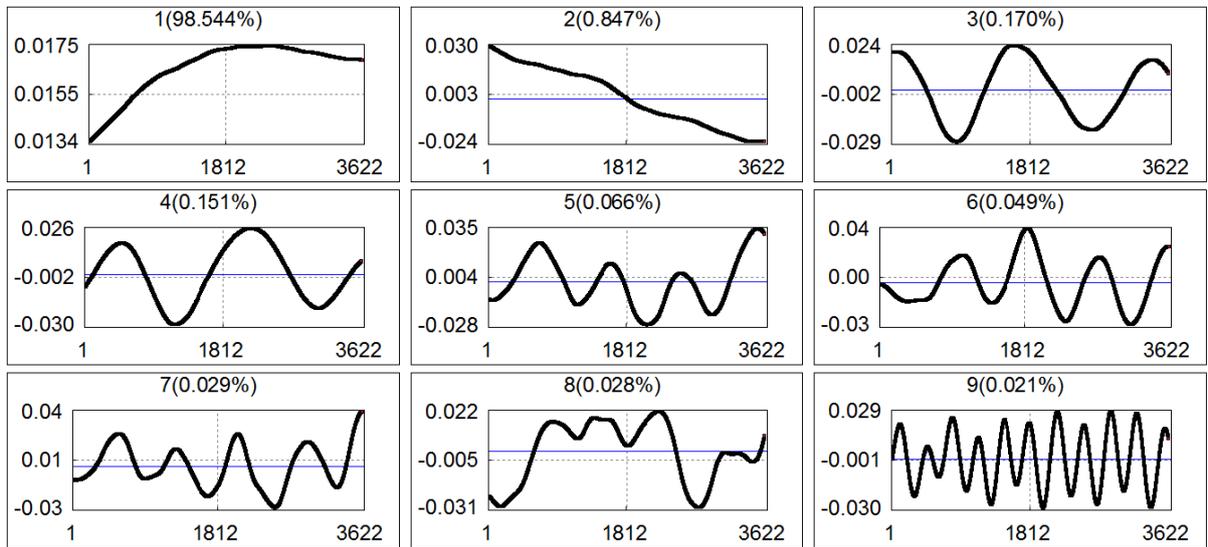
494



495

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

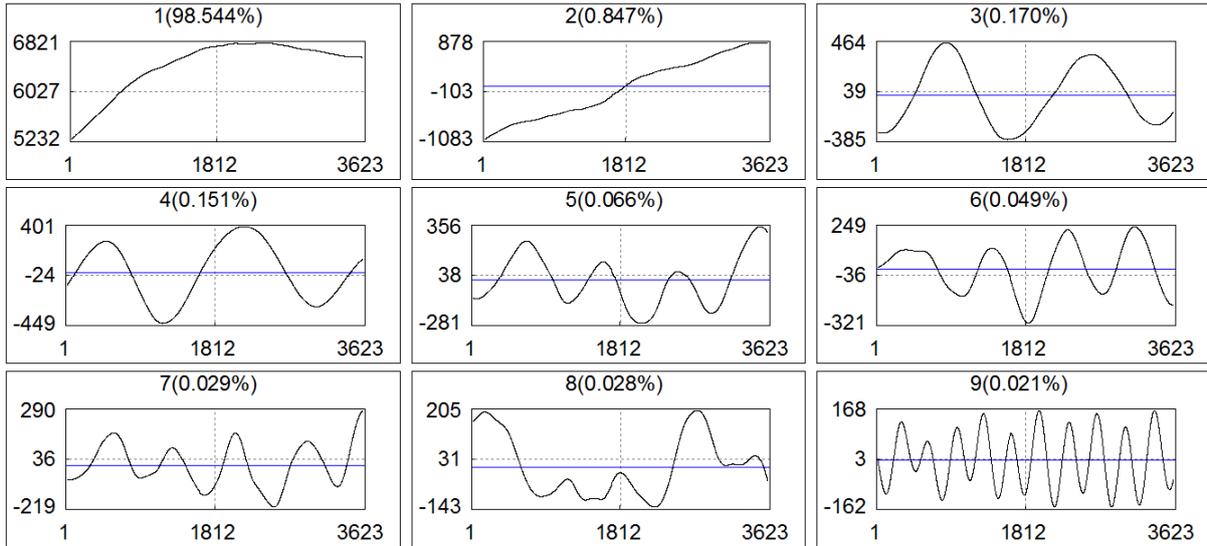
497



498

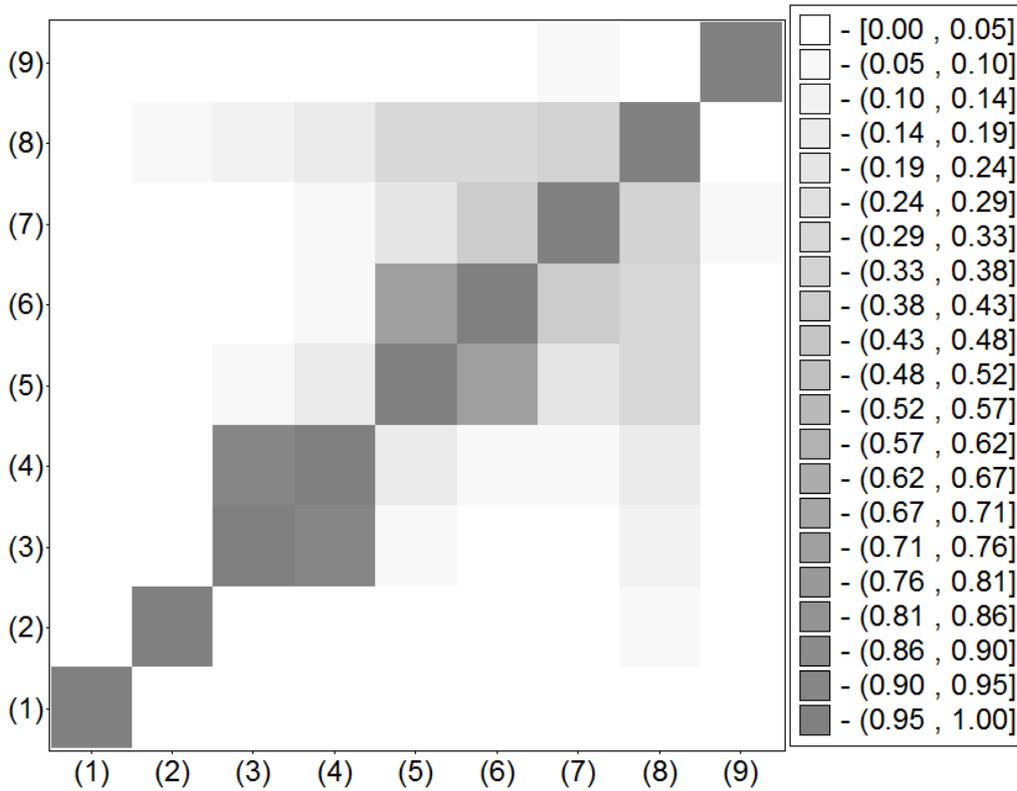
499 **Figure 4:** Eigenfunctions of soil radon (Rn-222) first 9 group. The scale in the horizontal x-
 500 axis represents the window length (L) and the scale in the vertical y-axis represents the singular
 501 eigenvectors (matrix path SVD).

REVISED VERSION



502

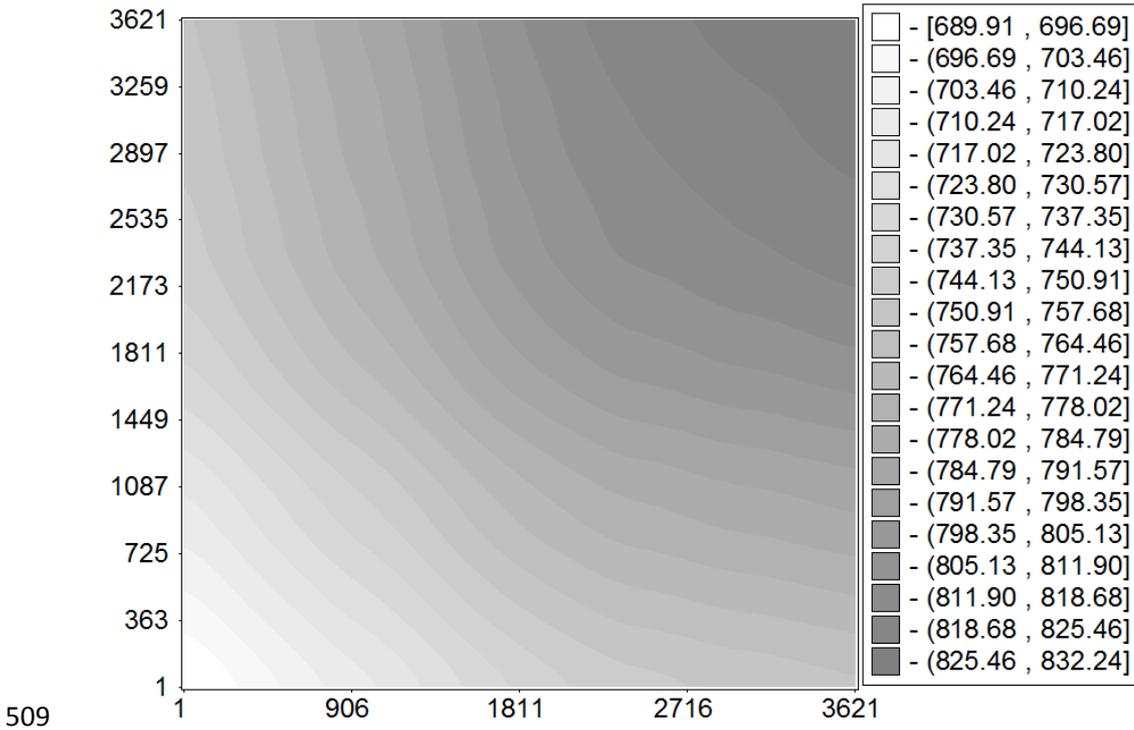
503 **Figure 5:** Principal Component of soil radon (Rn-222) related to the first 9 eigentriples. The
 504 scale in the horizontal x-axis represents the window length (L) and the scale in vertical y-axis
 505 represents the Principal Component (PC).



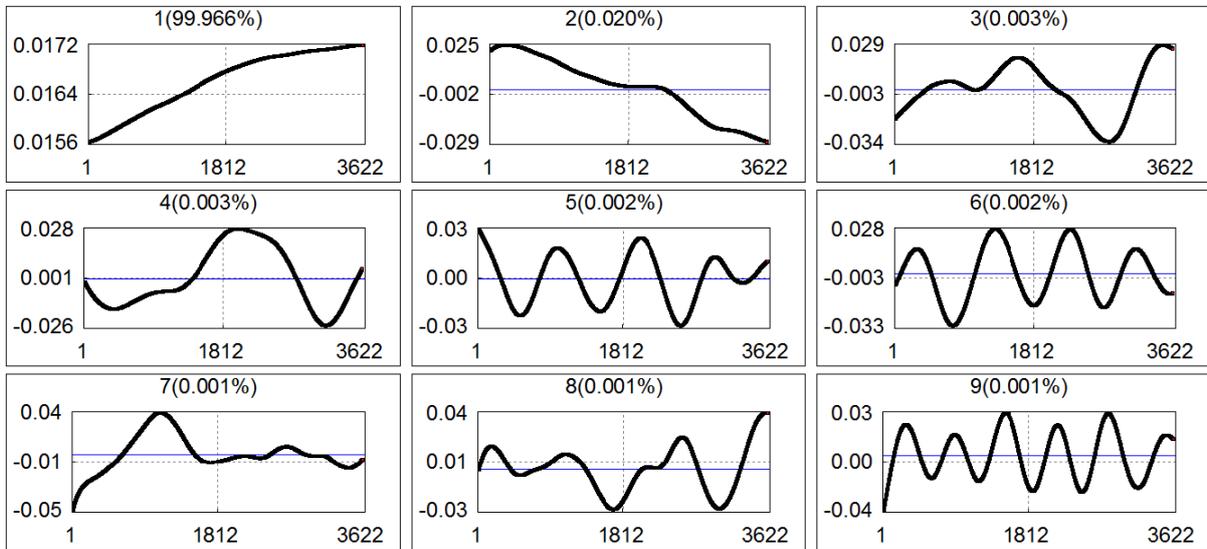
506

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

508



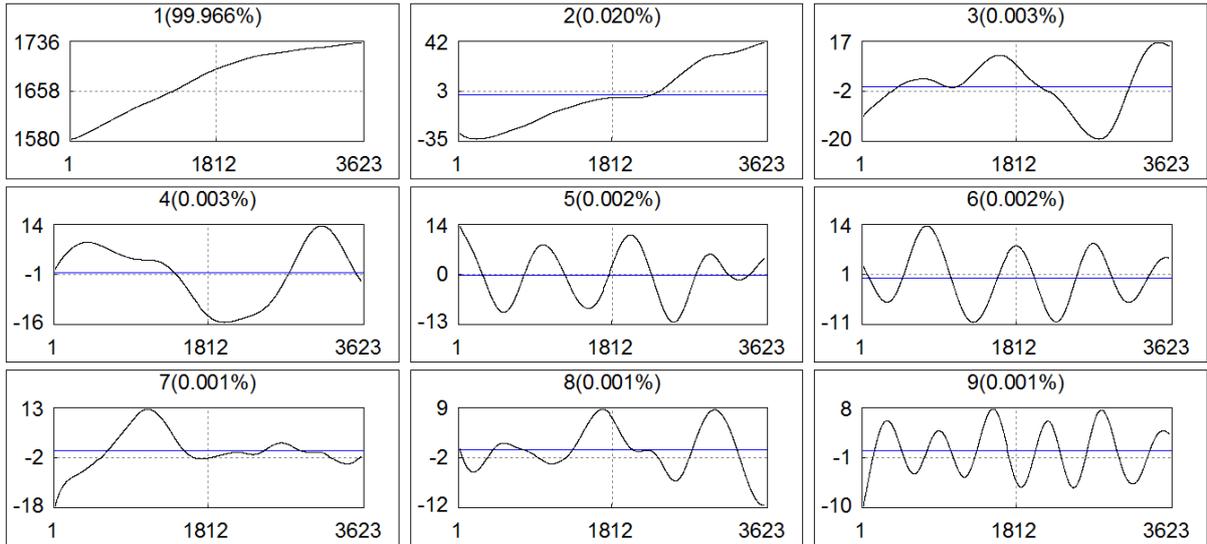
509
510 **Figure 7:** Covariance matrix of the first 9 group of soil temperature time series.



511
512 **Figure 8:** Eigenfunctions of soil temperature first 9 group. The scale in the horizontal x-axis
513 represents the window length (L) and the scale in the vertical y-axis represents the singular
514 eigenvectors (matrix path SVD).

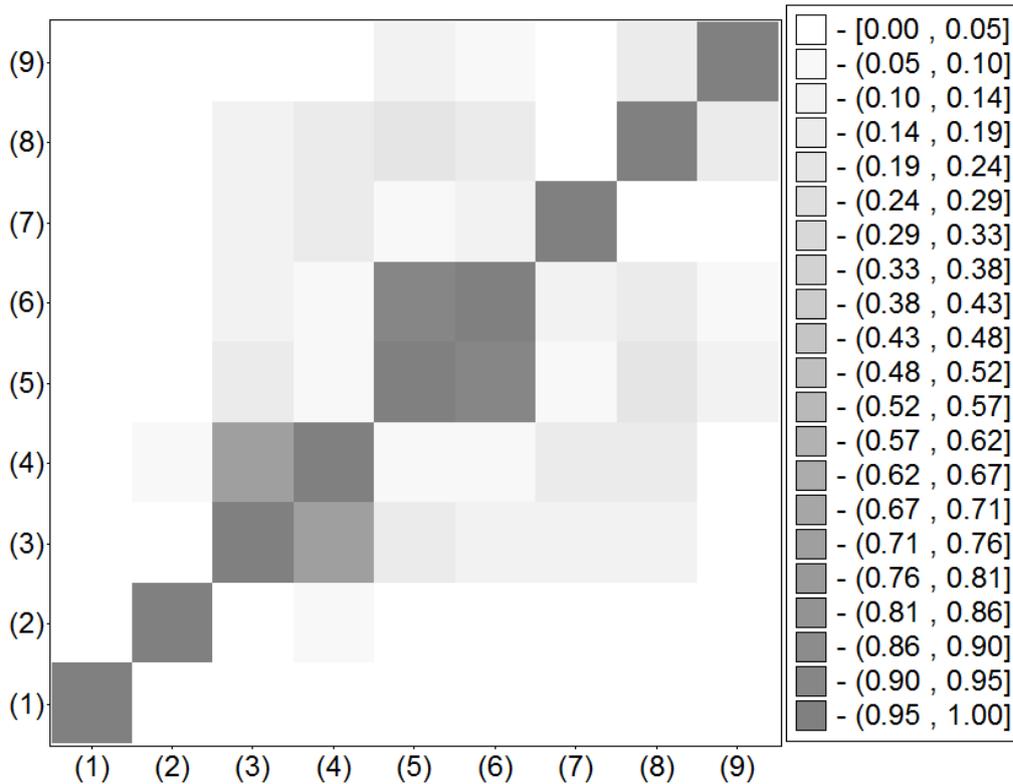
515

REVISED VERSION



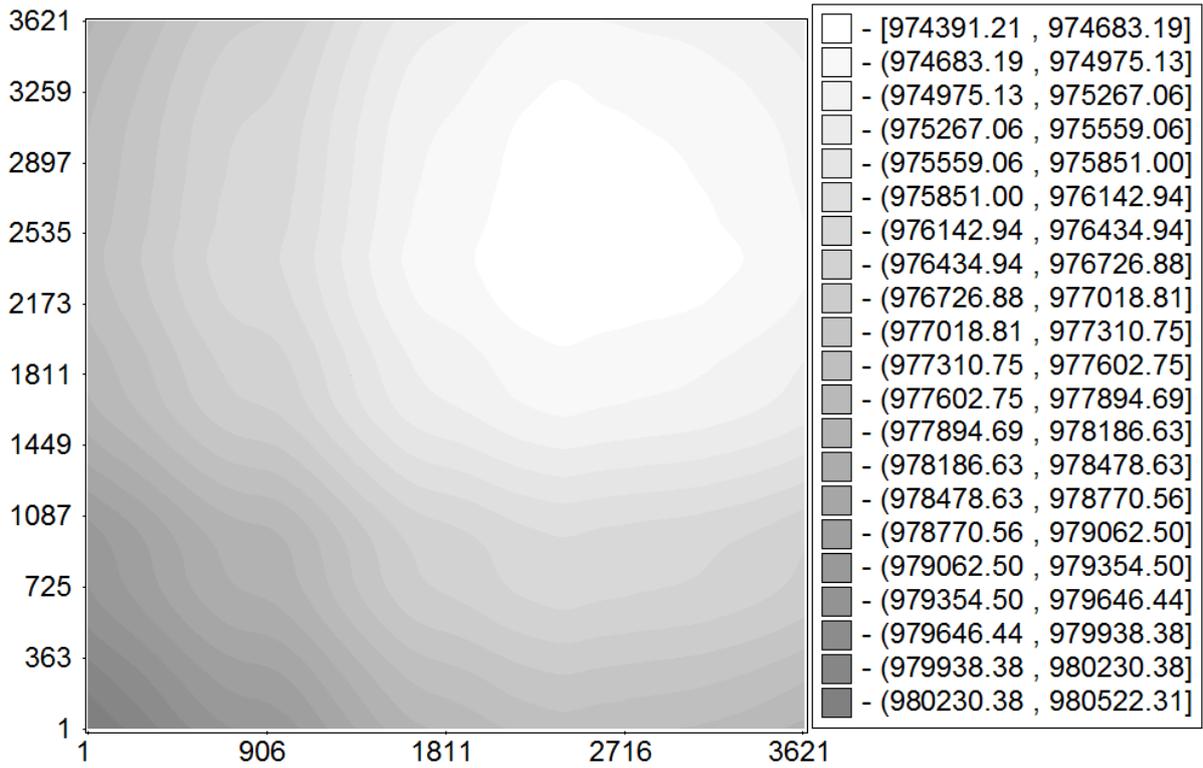
516

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



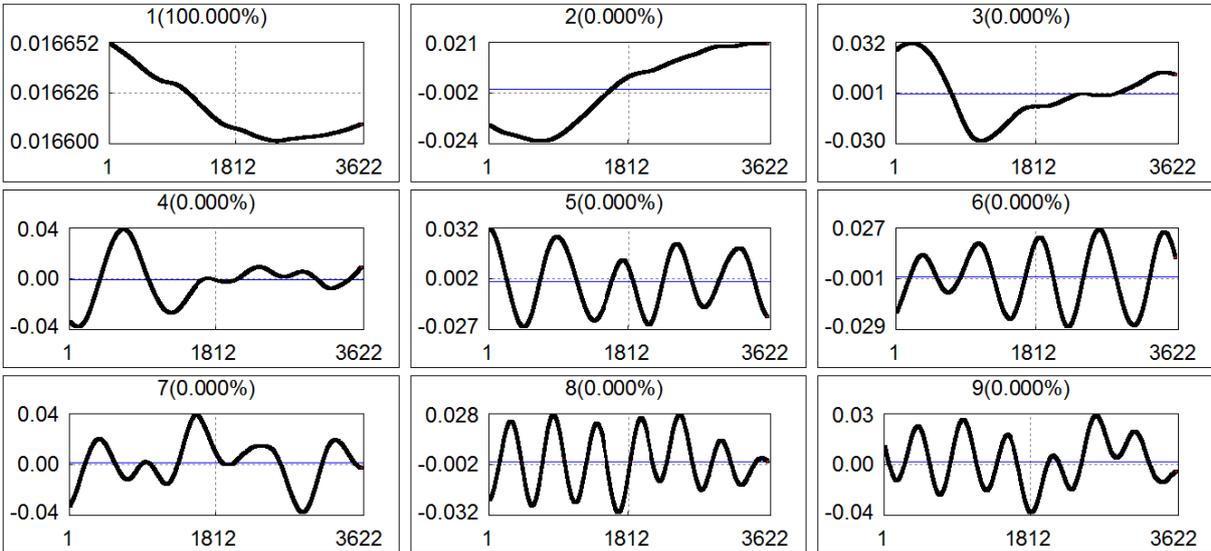
520

521 **Figure 10:** w-Correlation matrix for the 9 reconstructed components of soil temperature time
 522 series.



523

524 **Figure 11:** Covariance matrix of the first 9 group of soil pressure time series

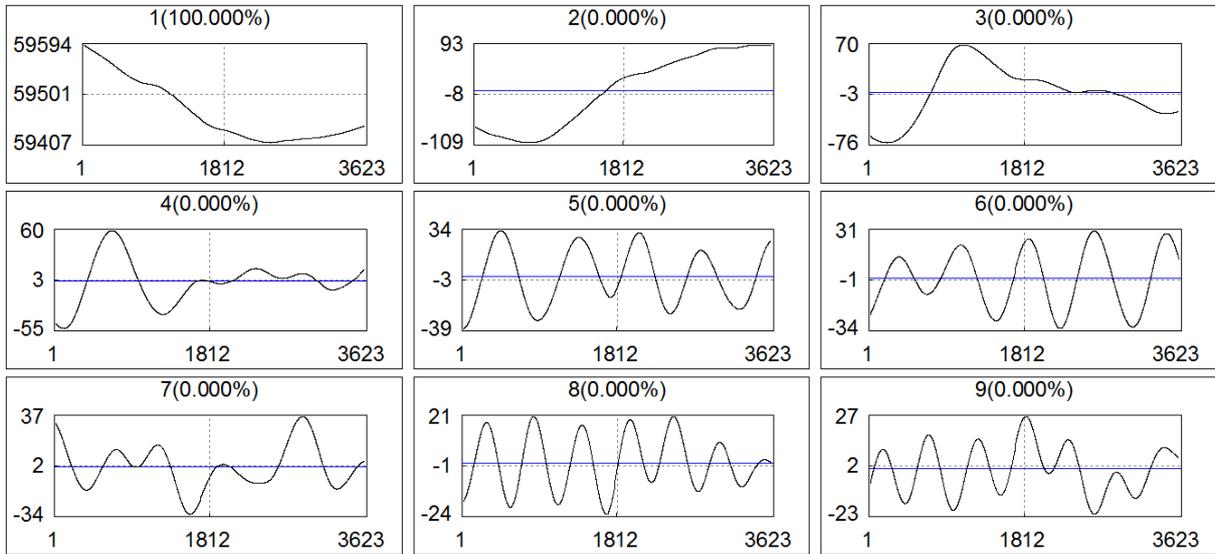


525

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

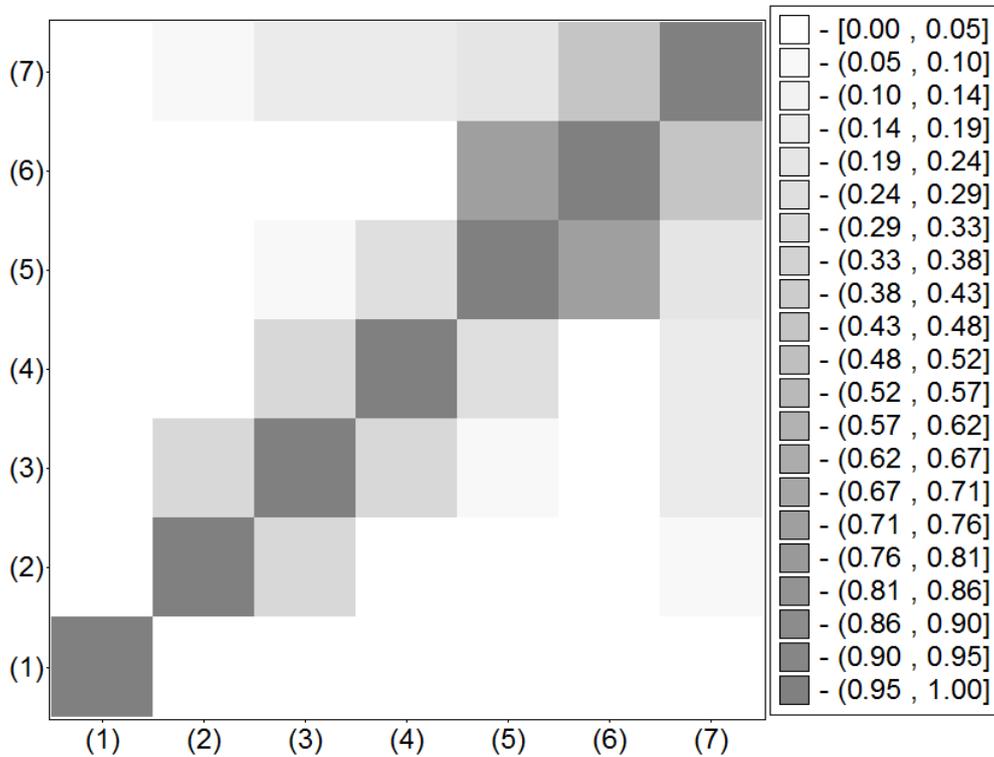
529

530



531

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

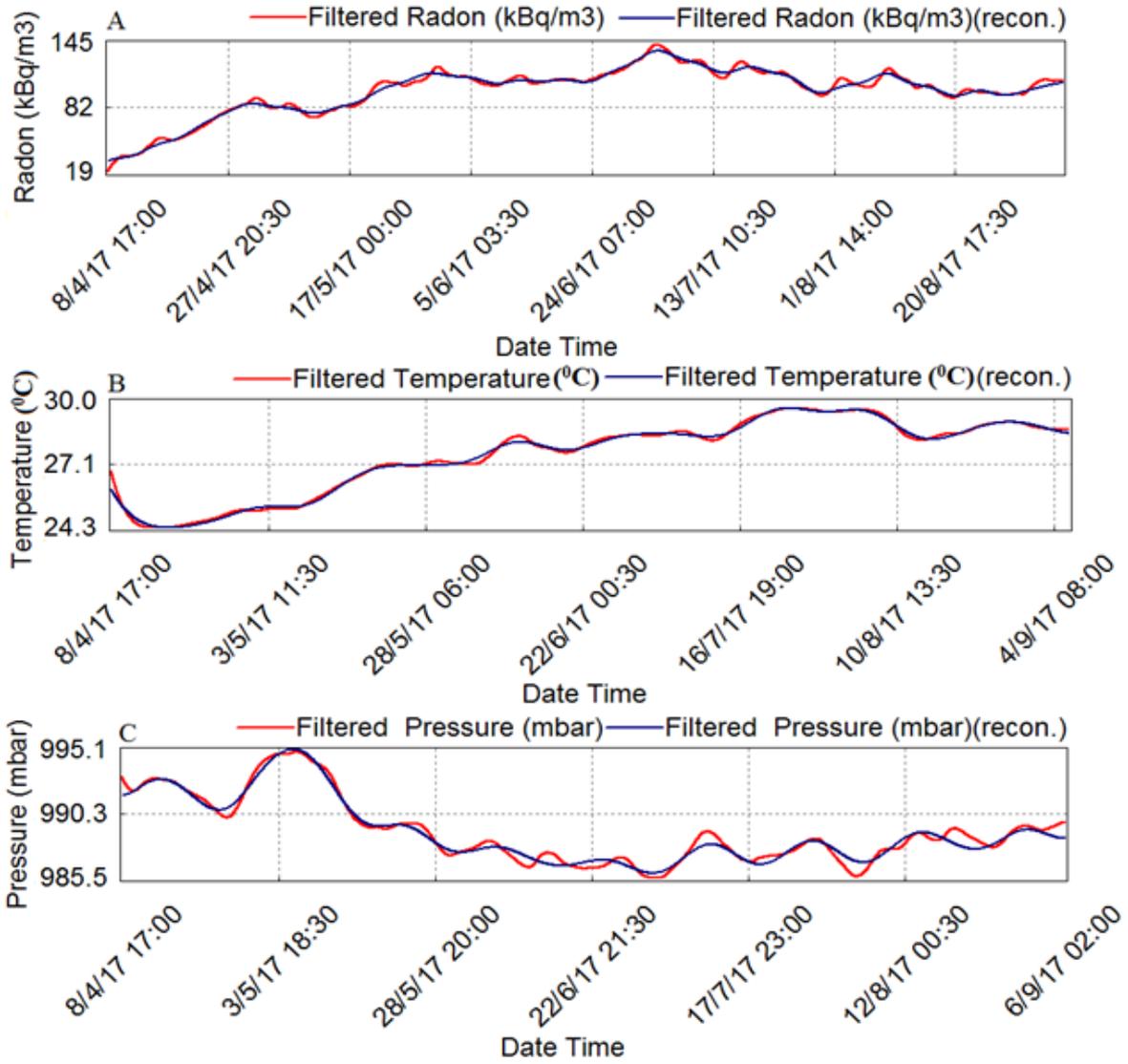


535

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

537

538



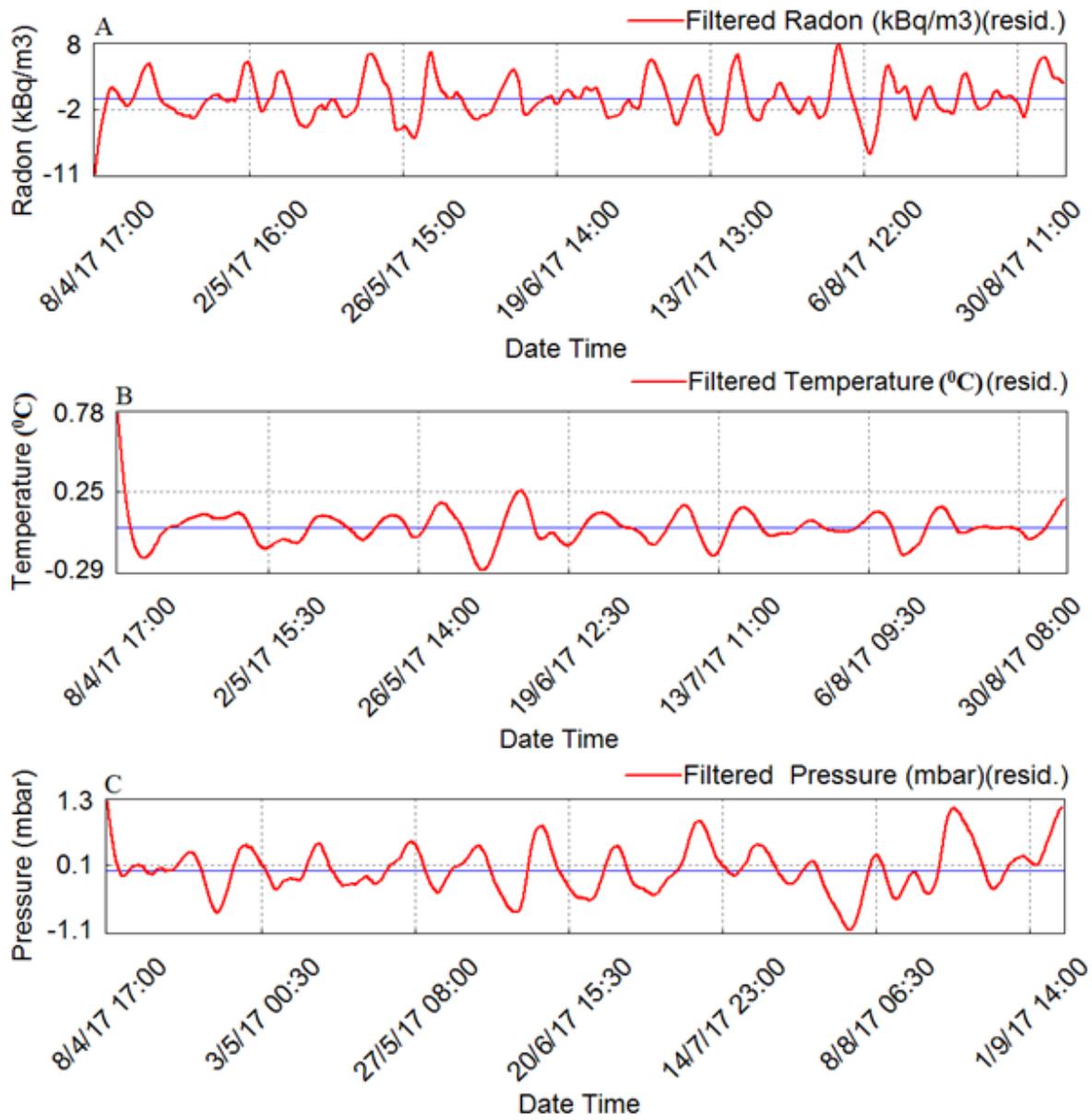
539

540 **Figure 15:** Plot showing the reconstructed time series of A) filtered soil radon, B) filtered
541 temperature and C) filtered pressure.

542

543

544



545

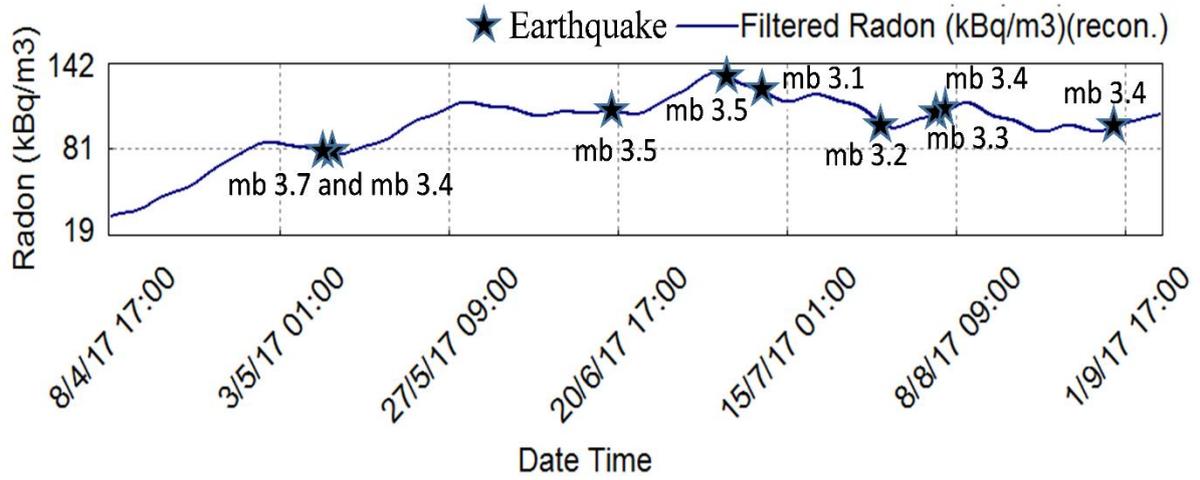
546 **Figure 16:** Residual of reconstructed time series of A) Filtered soil radon, B) temperature and

547 C) pressure respectively.

548

549

550



551

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

555

556

557 **Table 1:** The correlation co-efficient of soil radon gas concentration with soil pressure and
 558 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 (°C)	28.60	0.62	2.19	0.51
Pressure (mbar)	991.03	2.48	0.25	-0.52

559

560

561

REVISED VERSION

562 **Table 2:** Hypocentral parameters of the earthquake events found to have correlation with radon
 563 anomaly.

Date of Event	UTC TIME	Lat (°N)	Long (°E)	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	+

564

565

566

567

568

569