REPLY TO REVIEWER COMMENTS

COMMENT: General comments: This study is about anomalies in radon concentration related to earthquakes using singular spectrum analysis, model free method. More detail explanation is needed, since overall there is a lack of explanation of analysis. There are too many figures showing the results. Please provide as many figures as you need to support the results. It is necessary to explain the relevance and differences of previous studies using singular spectrum analysis for radon concentration data.

Reply

We convey our sincere thank you for your generous concern, comment and suggestions during review of the article. We have to best of our knowledge revised the manuscript with explanation of analysis (Has been included in revised version from Line No 64-75, 86-95, 178-187, 201-228, 247-249, 317-321). Yes, we do agree that there are too many figures showing the results but we wish it to be in the manuscript itself rather than as supporting file. To best of our knowledge some explanations on the relevance on studies using singular spectrum analysis for radon concentration data has been added in the Introduction section (Has been included in revised version from Line No. 86-95).

Specific comments

COMMENT 1. The way to calculate the covariance matrix showed in Fig. 3,7 and 11 should be described. Scientific implication and/or explanation should be included, since few descriptions regarding the figures are in the present paper.

Reply

Necessary description regarding calculation of co-variance matrix are given as follows:-

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

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

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

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

Here *k*=0,..., *M*-1 where *M* is the window size

The above are included in the revised version to calculate the covariance matrix showed in Fig. 3,7 and 11.

The above lines has been included in the revised version from Line No. 201-210

COMMENT 2. Eigenfunctions in Fig. 4, 8, 12 and principle components in Fig. 5, 9, 13 are almost the same considering the difference in sign. Detail explanations how to calculate them are needed in method section.

Reply

The Eigenfunctions in Fig. 4, 8, 12 and principle components in Fig. 5, 9, 13 are almost the same considering the difference in sign. We have to best of our knowledge tried to explain how to calculate them in method section in the revised version as follows:-

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

PC = *S* * matrix of eigenvectors

Where S is covariance matrix.

The above lines has been included in the revised version from Line No. 210-214.

COMMENT 3. Since the data used in the result section are filtered ones, diurnal and semidiurnal components seemed to be not included in the data. What is the basis of the description regarding lines 20-23 and lines 228-230?

Reply

The basis of the description is by observing the singular eigenvectors in matrix path SVD and principal components of soil radon, temperature and pressure (Figure 4, Figure 5, Figure 8, Figure 9, Figure 12 and Figure 13).

COMMENT 4. Regarding line 63, the reason of criteria, 100 km of mb > 3.1, for selecting seismic activity should be explained.

Reply

The reason for considering the criteria of 100 km with mb > 3.1 is the tectonic setting of the study area which is bounded by two major fault namely Kopili and the Bomdila Faults. The region is highly stressed because of geotectonic settings where earthquake do occur periodically within very short span of time. So we wish to wish correlate these earthquakes even though small, with geophysical observation made for yielding better reflection in data set due to stress release.

Kopili fault zone is experiencing compressional stress due to the Indo-Burma arc and Himalyan arc to the east and the north respectively which is characterized by transverse tectonics. The Bomdila Fault inter-weaves across three major tectonic domains of the NE-India, namely MCT, MBT and Naga-Disang thrust along NW-SE direction. (Kayal, 2010). The empirical relation for the energy release E to the magnitude (Gutenberg , 1956) gives $E > 10^{9.2}$ for earthquake of mb > 3.1. Stress and strain are directly proportional and hence in order to obtain adequate stress build up processes for observing precursory signals in temporal soil Radon emanation the criteria of 100 Km with mb > 3.1 is taken into consideration.

The red coloured lines above has been included in revised version from Line No. 64 to 75.

COMMENT 5. Regarding lines 158-163, lack of description of the way to group the elementary matrices.

Reply

We do agree there is a lack of description and hence we have added to best of our knowledge the way of grouping the elementary matrix:

The splitting of the elementary matrices Xi into several groups and summation of the matrices within individual group is grouping . If $I = (i_1, . .. ,i_p)$ is a group of indices $i_1, . . . , i_p$. Now for the group I of corresponding matrix X_I is defined as $X_I = X_{i1} + \ldots\ldots\ldots + X_{ip}$. The spilting of the of the set with indices $J = 1,\ldots, d$ on the disjoint subsets $I_1,..., I_m$ resembles to the illustration $X = X_{I1} + \cdot \cdot \cdot + X_{Im}$. This procedure of selecting the sets $I_1,\ldots, I_m$ is termed as eigentriple grouping. For given group I the contribution of the component X_I into $X_{I1}+\ldots + X_{Im}$ is estimated to the corresponding eigenvalues as: $\sum_{i\in I} \lambda i \sum_{i=1}^{d} \lambda i$.

The above lines has been included in the revised version from Line No. 178-187.

COMMENT 6. Regarding lines 186-190 and Table 1, Is which used to calculate correlation coefficient, original data or filtered data? Are the results just 0.5 and -0.5? If not, the smaller digit values should be indicated in Table 1 (e.g. 0.49)

Reply

The correlation coefficient was calculated for original temporal soil radon data. We have indicated the two digit values in Table 1

COMMENT 7. Explanation about the w-correlation matrix is needed at method section.

Reply

Explanation about w-correlation matrix has been included in the method section as follows:

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

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

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

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

The above lines has been included in the revised version from Line No. 215-228.

COMMENT 8. Regarding lines 240-242, It can be suggested that the pressure with larger change is dominant, only when the response of radon to temperature and pressure is equal. However, no evidence the response of radon to them are indicated in the paper. Therefore, there is lack of basis for this description.

Reply

This has been observed from re-constructed time series and residuals of reconstructed time series (Figure 15 and Figure 16 respectively). The very basis of this observation is principal component analysis which establish this fact (Figure 5, Figure 9 and Figure 13).

COMMENT 9. Regarding lines 253-255, detail descriptions indicating which earthquakes have positive anomaly and the others have negative anomaly should be added. Are these results corrected the effects of temperature and pressure? Otherwise, it cannot be distinguished whether it is a change due to an earthquake or a change due to temperature or pressure.

Reply

In Table 2 we have detailed the descriptions indicating which earthquakes indicate positive and negative anomaly so that these results can be distinguished whether it is a change due to an earthquake or a change due to temperature or pressure. It is observed that the earthquake occurs when the radon emanation reaches background level (base level) from the highest peak anomaly accomplished. Most of the events were preceded by positive anomaly but three events occurred during the decreasing trend of radon emanation which are considered as negative anomaly in the case.

COMMENT 10. Regarding lines 266-268, Raising the water level means that there is a pressure gradient, which means that the fluid flows from a place with high pressure to a place with low pressure. Radon can be thought of as moving through the ground as well, and rising water levels can also indicate an increase in radon concentration. Please indicate if there is any previous research that supports the argument, in lines 266-268, in the paper.

Reply:

Yes there are very few research that supports the argument “MEROLLA, P., MOSE, D., & KING, A. (2003). EFFECT OF WATER TABLE FLUCTUATIONS ON RADON EMANATION FROM SOIL stating that a falling water table, as might occur during a dry interval after a heavy rainfall interval, initially causes a decrease in soil radon concentration (probably because the increased depth having pore spaces not filled by water acts as a void into which soil radon diffuses, and/or the water table acts as a sinking source of radon enrichment to the overlying soil). When the lower water table stabilizes, radon emanation from the soil increases (probably because the increased volume of connected water-free pores)”. The reference has been added in the revised version.

The line No 266-268 has been re interpreted as “The fluid pressure increases may result in water level rise (pore spaces probably gets filled by water and diffusion rate of radon is more in void than in water filled pore space; Nielson et al., 1984; Merolla et al., 2003) and this probably does not allow the soil gas Rn-222 to escape from the surface which in turn reduces or stabilizes the emanation of Rn-222. (Has been included in revised version from Line No 317-321).” Further the diffusion coefficients of radon in soil increase with the median pore diameter, and decrease with increasing soil water contents and with increasing widths of the soil pore size distribution (Nielson et al., 1984).

Technical comments 1. Line 14, the complete name of MPGO also should be indicated here. **2.** Lines 52-53, “Latitude 26.61o; Longitude 92.78o” should be “Latitude N26.61o; Longitude E92.78o”. **3.** Line 64, “The major problem arises is the” should be “The major problem is the”. **4.** Line 68, add spaces like “Stranden et al., 1984; Kumar et al., 2009; Walia et al., 2005”. **5.** Line 72, “parameters on radon emanation” should be “parameters on radon concentration”. **6.** Line 150, add brackets like “i.e. (Ui, Uj)= 0 for”. **7.** Lines 182, and 183, “m3” should be “m3”. **8.** Line 197, does percentage correlation coefficient mean coefficient of variation? **9.** Fig. 4, 5, 8, 9, 12 and 13, need axis label. **10.** “0C” should be “oC”. **11.** Line 408, “(200-300 N and 860-980 E)” should be “(20o -30o N and 86o -98o E)”.

Reply

As per your generous technical comments and suggestion from **1 - 7** has been corrected in the revised version. Yes, the “percentage correlation coefficient” was a typos error and it means the “coefficient of variation”. The technical comment **8** has been corrected in the revised version. The technical comment **8** has been corrected and for comment **9** for Fig. 4, 5, 8, 9, 12 and 13 axis label are described and added in the figure caption of revised version. The technical comment **11** has also been corrected in the revised version.