IL NUOVO CIMENTO **41 C** (2018) 223
DOI 10.1393/ncc/i2018-18223-4

Colloquia: SIRR 2018

Innovative methodologies for the analysis of radon time series

F. Ambrosino([1])(*), M. Pugliese([2]), V. Roca([2]) and C. Sabbarese([1])

([1]) *Department of Mathematics and Physics, University of Campania "L. Vanvitelli" Caserta, Italy*
([2]) *Department of Physics "E. Pancini", University of Naples "Federico II" - Naples, Italy*

received 4 December 2018

Summary. — The present contribution is a part of a wide and detailed study on the long-term radon time series specific activity using novel methods of analysis with different aims: to extract the general trend; to evidence and distinguish endogenous and climatic dependences; to identify anomalies in order to show the correlation with geophysical phenomena; to forecast the original signal. Various examples of application of these methods are exhibited on radon time series collected in some Italian sites, in order to make a retrospective analysis to better study the existing correlation between anomalies and earthquakes. The results obtained from this study highlight the manifold capabilities of the proposed methods in the radon signal analysis, exploitable also for any kind of research framework that needs data analysis of time series.

1. – Introduction

In last years, the application of mathematical procedures for the analysis of the temporal signals has grown with the development of new and more sophisticated methodologies [1-4], as the hybrid methods. These consist in the sequential application of two or more methods with the aim of a better performance in terms of uncertainty and algorithm speed [5]. The application consists in using the output of the first method as the input of the second, and so on. Each component method has various potentials for the study of signal processing, which are mixed with different goals of analysis to better reach the features of the investigated time series. In this work, such procedure has been applied to the radon (^{222}Rn) signals recorded in Italian sites.

Radon signal monitored in soils contains both local and remote information [4,6-8]. It is well established that radon can provide useful information if it is continuously recorded in a site to characterize the local background signal and to record the complete trends

(*) Corresponding author. E-mail: fabrizio.ambrosino@unicampania.it

of all local and remote effects on the signal. The effects of sources, permeability and climate conditions of a site furnish the principal components of a radon time series. They are predominant and hide components due to remote geophysical effects and sources. Therefore, it is necessary to perform a specific analysis of the signal to distinguish and extract the desired information. Some of these analyses were carried out by using different individual methods on radon signal in different studies, and provided reasonable results [2, 4, 9]. In this study, more efficient procedures have been developed to obtain the complete information present in a radon time series of a site and also to forecast the signal on the basis of the known part. The proposed hybrid methods used for the analysis of radon time series are the appropriate combination and selection of several single-component methods in order to exploit the strengths of each one in obtaining information related to the specific characteristics of the signal.

2. – Materials and methods

The smoothing trend of the signal is made with the use of the STL (Seasonal and Trend decomposition using Loess) method [10] that decomposes the time series using a sequence of locally weighted regression models operations. The meteorological and seasonal effects are estimated with the MLR (Multiple Linear Regression) [4, 9]. A detailed decomposition of the signal in the largest possible number of components is performed by EMD (Empirical Mode Decomposition) through interpolation [11]. The calculation of a forecast signal based on the known one is made with SVR (Support Vector Regression) techniques [12]. The important part of the recognition of anomalous components in the series, cleaned from any kind of modulations, is made through the residues obtained essentially from the difference between the rough signal and the predicted one identified with the used hybrid methods.

The proposed hybrid methods are composed as follows:

- STL + MLR + EMD, used for a deeper decomposition of the signal, therefore, incorporating also the trend extrapolation;

- STL + MLR + EMD + SVR, used for the anomalies detection via the forecasting of the time series.

STL decomposes a time series into trend, seasonal and remainder components. On the remainder component the MLR model is applied, to consider environmental parameters that can influence the time series. On the residue between remainder and regression model, a further differentiation of seasonal-periodic modulations within the signal is made progressively in frequency through cubic-spline interpolation of the EMD method. After the ordered application of STL + MLR + EMD to decompose, distinguish and clean the time series from different principal component fraction of every kind, the trend analysis is obtained from the last component of the EMD decomposition that is the wanted representative trend of the signal. Instead, the forecasting part is obtained from the application of the SVR method on each component derived from the EMD method (that trains a regression model to predict them) and the final signal is made with the application of another SVR on the sum of the outputs from all predicted components. Finally, the anomalies detection is carried out forecasting the original time series, which is used as 95% confidence interval for the same signal in order to determine an anomalous peak when it is beyond that confidence level. In the case of anomalous growth, the anomaly is marked as positive, otherwise as negative.

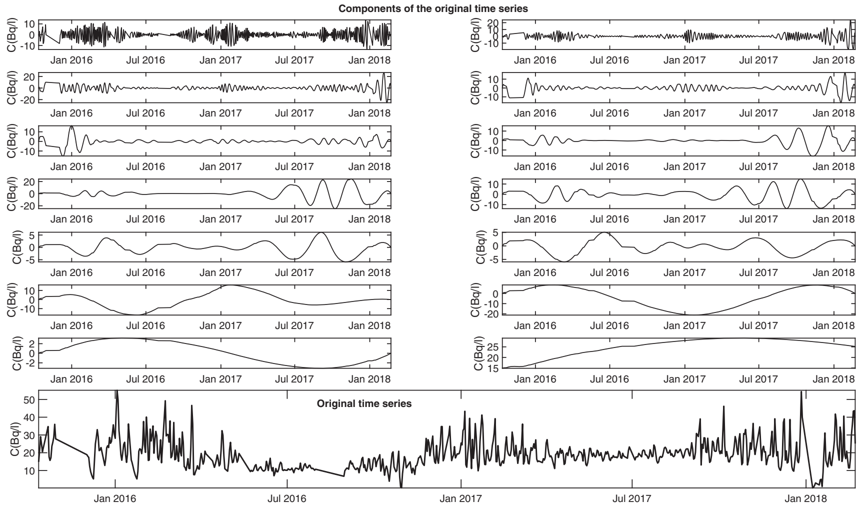

Fig. 1. – Application of STL + MLR + EMD hybrid method for the decomposition and the trend extrapolation to a radon time series (on the bottom) recorded at the S1 site from 01/10/2015 till 31/02/2018.

Examples of the application of the proposed hybrid methods are assessed on the analysed time series of radon specific activity C(Bq/l) data recorded in two sites of the Phlegrean Fields caldera located in the district of Naples (Italy) [4, 13]; the collected data also include meteorological parameters recorded at the nearby stations. The two selected sites are hereinafter referred to as S1 and S2.

3. – Results and discussion

Specific mathematical algorithms have been developed in Matlab® computing environment to implement the proposed hybrid method. In fig. 1 there is displayed an example of application of the STL + MLR + EMD method to highlight the decomposition of a radon time series recorded in S1. The trend signal is the last component of this analysis and is shown in the same figure. This analysis also performs the extraction of the periodic and seasonal components that can be indicative of the development of a deep process, the extrapolation of the noise from the signal that corresponds to the first

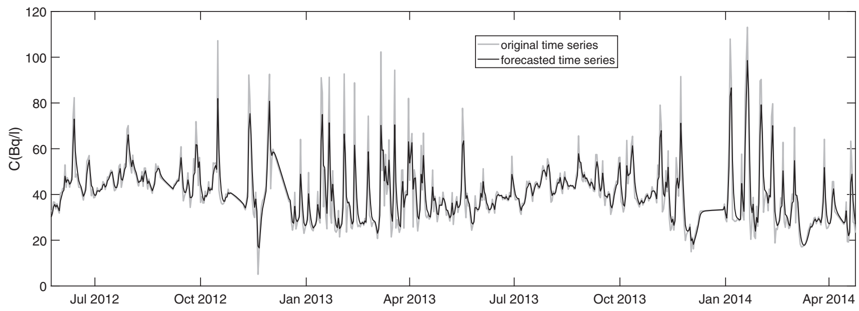

Fig. 2. – Application of the STL + MLR + EMD + SVR hybrid method for the forecasting (black line) to a radon time series (gray line) recorded at the S2 site from 31/05/2012 till 31/05/2014.

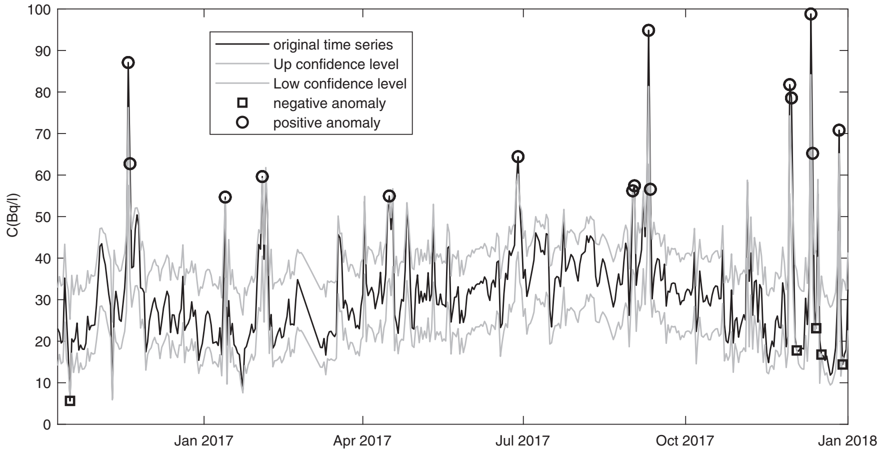

Fig. 3. – Application of the STL + MLR + EMD + SVR hybrid method for the anomalies identification to a radon time series (black line) recorded at the S2 site from 01/10/2016 till 01/01/2018. The gray lines are the two 95% confidence intervals beyond which a value is considered anomaly (positive as circle and negative as square).

component in fig. 1, and, of course, the determination of the residual deriving from the use of the algorithm code.

Examples of the application of the STL + MLR + EMD + SVR hybrid method are carried out on two radon time series recorded in the S2 site for the forecasting of the time series and for the detection of anomalies; the results are displayed in fig. 2 and fig. 3, respectively. The comparison between the original signal and the forecast signal is displayed in fig. 2; the average and the standard deviation of the uncertainty are 5.2% and 1.2% calculated as the average of the difference between the original and the forecast signal divided by the original one. The procedure of extracting the anomalous changes that do not match the overall pattern bounded by the two statistical confidence levels is shown in fig. 3. The main purpose of the latter analysis is the possible selection of the

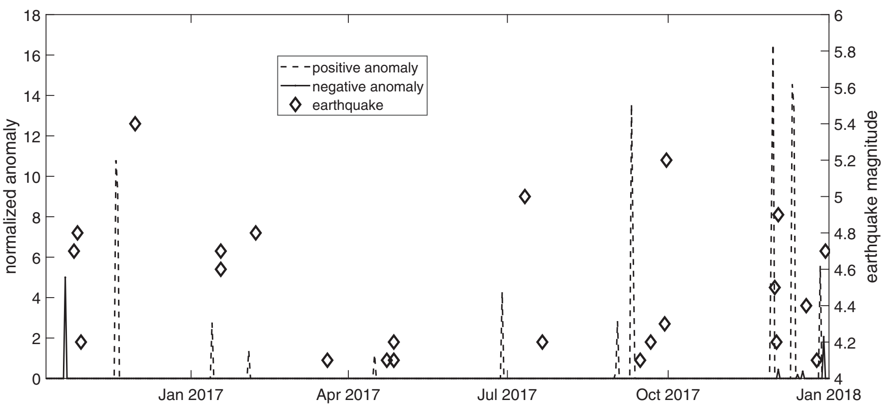

Fig. 4. – Normalized anomalies from fig. 3 compared with selected earthquakes influencing the radon monitoring site S1 in the Phlegrean Fields area. The dashed lines and black lines are positive and negative anomalies, respectively, while the diamonds are the selected earthquakes.

detected radon anomalies that can be caused by endogenous phenomena in the Earth crust. The normalized anomalies (as the distance of a data point from the average in terms of the standard deviation) found from the analysis in fig. 3 and compared with the earthquakes that influence the Phlegrean Fields radon monitored area (selected from the Earthquake National Center database of the National Institute of Geophysics and Volcanology of Italy -`cnt.rm.ingv.it`) are shown in fig. 4. These anomalies indicate that the gas could be a geochemical forerunner of selected earthquakes, as already consolidated by many studies in the search for clear indicators of tectonic phenomena that could contribute to reducing the effects of these natural risks [4, 6-9, 14-16].

4. – Conclusion

In the analysis of radon time series, the hybrid methods provide a potentially powerful tool for capturing and interpreting the signal due to both exogenous and endogenous factors. In this study, two hybrid methodologies are proposed through the grafted use of the different statistical-mathematical methods: STL, MLR, EMD and SVR. Both have different characteristics and are used for different purposes in signal analysis: decomposition with trend identification, seasonal-periodic component extraction with noise and residual estimation; forecasting of the original series; anomalous component identification as the difference between the original signal and the forecast one. The study demonstrates the effectiveness of signal analysis methods applied to some radon time series, and shows a retrospective analysis of the existing correlation between anomalies and earthquakes influencing the radon monitored area. This confirms the origin of the gas from the subsoil and shows how the gas could be a geochemical precursor of earthquakes.

REFERENCES

[1] Breiman L., *Mach. Learn.*, **24** (1996) 49.
[2] Pasini A. *et al.*, *Nuovo Cimento C*, **24** (2001) 331.
[3] Alessio S. *et al.*, *Nuovo Cimento C*, **27** (2004) 73.
[4] Sabbarese C. *et al.*, *Radiat. Prot. Dosim.*, **177** (2017) 202.
[5] Wang J. *et al.*, *Energy J.*, **148** (2018) 59.
[6] Stranden E. *et al.*, *Health Phys.*, **47** (1984) 480.
[7] Segovia N. and Mena M., *Nuovo Cimento C*, **22** (1999) 275.
[8] Friedmann H., *Radiat. Prot. Dosim.*, **149** (2012) 177.
[9] Singh S. *et al.*, *Pure Appl. Geophys.*, **174** (2017) 2793.
[10] Cleveland R. B. *et al.*, *J. Off. Stat.*, **6** (1990) 3.
[11] Yu H. *et al.*, *Pure Appl. Geophys.*, **175** (2018) 1683.
[12] Drucker H. *et al.*, *Adv. Neural Inf. Process. Syst.*, **9** (1997) 155.
[13] Sabbarese C. *et al.*, *Appl. Radiat. Isot.*, **122** (2017) 180.
[14] Inan S. and Seyis C., *Acta Geophys.*, **58** (2010) 828.
[15] Ghosh D. *et al.*, *J. Appl. Geophys.*, **69** (2009) 67.
[16] Virk H. S., *Nuovo Cimento C*, **22** (1999) 423.