

# Investigating correlations between earthquakes and extreme events in self-potential data recorded in a seismic area of Southeastern Appennine Chain (Italy)

Luciano Telesca <sup>(1)</sup>, Marianna Balasco <sup>(2)</sup>, Gerardo Colangelo <sup>(1)</sup> and Vincenzo Lapenna <sup>(1)</sup>

<sup>(1)</sup> *Istituto di Metodologie per l'Analisi Ambientale (IMAA), CNR, Tito Scalo (PZ), Italy*

<sup>(2)</sup> *Dipartimento di Geologia e Geofisica, Università degli Studi di Bari, Italy*

## Abstract

The Normalized Wavelet Cross-Correlation Function (NWCCF) was used to study correlations between the series of extreme events in self-potential data and earthquakes, both modelled as stochastic point processes. This method gives objective results, robust to the presence of nonstationarities that often affect observational time series. Furthermore, the NWCCF identifies the timescales involved in the cross-correlated behaviour between two point processes. In particular, we analyzed the cross-correlation between the sequence of extreme events in self-potential data measured at the monitoring station Tito, located in a seismic area of Southern Italy, and the series of earthquakes which occurred in the same area during 2001. To evaluate the influence of rain on the dynamics of geoelectrical variations, we applied the same approach between the selected extreme values and the rain data. We find that the anomalous geoelectrical values seem to cross-correlate with the rain at short and intermediate timescales ( $\tau < 500$  h), while they significantly cross-correlate only with earthquakes ( $M \geq 2.5$ ) at long timescales ( $\tau > 500$  h).

**Key words** *self-potentials – extreme events – earthquakes – cross-correlation*

## 1. Introduction

In recent years, seismic research has involved the scientific community in debates concerning the predictability/unpredictability of earthquakes (*i.e.* Geller, 1996; Evans, 1997). The most important key point is to perform in-depth experimental research aimed to develop i) well-based geophysical monitoring activities, ii)

quantitative methods for the analysis of precursory signals and iii) physical models able to explain the generation mechanism of anomalous geophysical signals in earthquake preparation zones (Wyss, 1997).

Geophysical anomalous temporal patterns, interpreted as seismic precursors, have been observed and attributed to stress and strain changes which were followed by earthquakes (*i.e.* Rikitake, 1988; Park, 1996), but the physics underlying the source mechanisms of geophysical precursory phenomena is still difficult to understand (Nur, 1972; Scholtz *et al.*, 1973; Scholtz, 1990). Therefore, it is necessary to objectively define criteria to evaluate the reliability of short-term predictions based on this type of precursory signals (*i.e.* Wyss, 1991; Geller, 1996).

In this research field, the main weak points are the sporadic and irregular monitoring activi-

---

*Mailing address:* Dr. Luciano Telesca, Istituto di Metodologie per l'Analisi Ambientale (IMAA), CNR Area della Ricerca di Potenza, Contrada S. Loja, 85050 Tito Scalo (PZ), Italy; e-mail: ltelesca@imaa.cnr.it

ties in active seismic areas and the absence of robust statistical methods able to discriminate extreme events from random noise in precursory time series. Currently, there is no scientific consensus for a significant correlation between the reported anomalies and the occurrence of earthquakes. Precursor research is required to demonstrate quantitatively that its data are anomalous and are causally related to earthquakes.

This work tackles the problem of the cross-correlation between the series of extreme events identified in geoelectrical time series and earthquake activity. After identifying anomalies in self-potential data by means of objective methods (run theory or crossing theory) (Cuomo *et al.*, 1996, 1997, 1998), the fundamental question to be solved is the correlation with the seismic events which occurred in the investigation area. Rather than simply observing the concomitantly occurring self-potential extreme events and earthquakes by visual inspection, the cross-correlation between the two series of events has been quantitatively stated by means of the so-called Normalized Wavelet Cross-Correlation Function (NWCCF). This function not only identifies the presence of significant cross-correlations between two sequences of events, but also detects the timescales mainly involved in such cross-correlation. This statistical measure was used in Lowen *et al.* (2000) concerning the investigation of cross-correlated behaviour in two spike trains of neural discharges in cat RGC and LGN cells.

## 2. Geological and seismological settings

The measuring station is located on the Southern Apennine Chain whose framework consists of a pile of thrust sheets forming a complex system orogenically transported over the flexured south-western margin of the Apulia foreland. It is the result of a complex sequence of tectonic events associated with the collision between Africa and Europe (Pantosti and Valensise, 1990).

Starting in the Middle Miocene up to the Upper Pliocene, five or more compressional tectonic phases (Patacca *et al.*, 1988) caused progressive thrusting and piling of different tec-

tonic units corresponding to different paleogeographic domains toward the stable external domains of the Apulo-Adriatic foreland. During the Quaternary, the Southern Apennines were affected by an important extensional tectonic phase, with NE-SW extensional trend, that caused further chain fragmentation into several isolated blocks (Doglioni *et al.*, 1996).

The Southern Apennine Chain is one of the most active areas of the Mediterranean region. In this area, on November 23, 1980 ( $M_S = 6.9$ ), a large normal-faulting earthquake occurred (Pantosti and Valensise, 1990). One of the most historically important events, the December 16, 1857 normal-faulting earthquake (Mallet, 1862), occurred in Val d'Agri. Seismic activity occurred after the 1980 event consisting of medium intensity events ( $M < 5.0$ ) located close to the border between Campania and Basilicata regions. The May 5, 1990 ( $M_D = 5.0$ , ING-Istituto Nazionale di Geofisica) and the May 26, 1991 ( $M_D = 4.7$ ) earthquakes may be considered the strongest events after the Irpinia 1980 earthquake (Tertulliani *et al.*, 1992) which occurred in this area. These events were followed by aftershock sequences that identify a fault structure located near Potenza town. Seismological analysis of the above events demonstrated that such earthquakes were generated by a strike-slip fault in WE direction, perpendicularly oriented toward the Apennine Chain (Ekström, 1994). This fault lies north of Potenza town and is located in such a way as to limit toward north and south two great seismogenetic faults that caused the 1857 Val d'Agri and 1980 Irpinia earthquakes respectively. The fault area outlined by the aftershocks extends approximately 20 km in length and 10 km in depth, making it significantly larger than expected for a  $M_L = 5.0$  earthquake. The aftershocks were concentrated between 15 and 25 km depth, which is deeper than over well determined focal depth in the Central and Southern Apennines (Ekström, 1994).

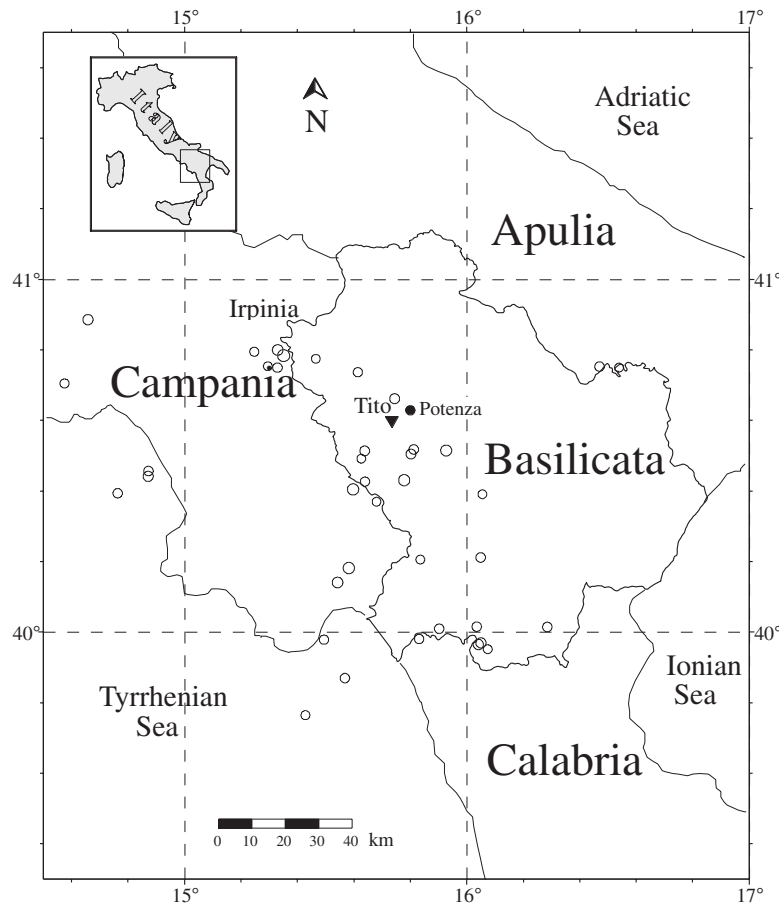
The geological and seismological settings allow us to consider the investigated area an ideal outdoor laboratory to study the possible correlations between tectonic activity and anomalous patterns in the geoelectrical signals.

### 3. Data analysis and discussion

In this work we analyze the time series of hourly means recorded at station Tito during 2001. Figure 1 shows the location of the station and the epicenters of the earthquakes which occurred in a circular area of 100 km radius from the measuring station. Technically, a geoelectrical or self-potential time series is a sequence of voltage differences measured with a selected sampling interval using a receiving electrode array. During the geoelectrical soundings, where a current is injected into the ground, the self-po-

tential represents the noise (Lapenna *et al.*, 1994). On the other hand, when we record using a passive measurement technique (*i.e.* without an energizing system), we measure the signal.

To avoid some self polarizing effects, we used ceramic electrodes made of ceramic vessels filled with a saturated solution of copper sulphate. The time series obtained with the different probes are constantly checked to remove anomalous patterns related to polarization effects (Di Bello *et al.*, 1994). The problem of the identification of extreme events in self-potential data, that are electrical values above or below a



**Fig. 1.** Location of the geoelectrical monitoring station Tito in Southern Apennine Chain. The epicentral distribution of the events occurring during the measurement period is also shown.

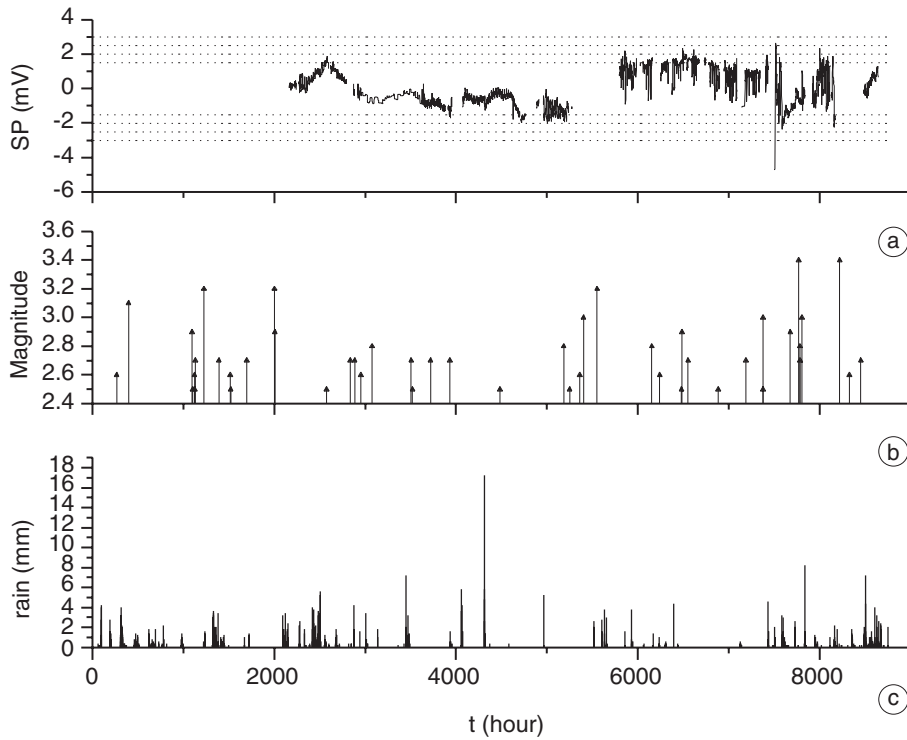
fixed threshold, has already been approached in previous works by Lapenna *et al.* (1994) and Cuomo *et al.* (1996, 1997, 1998), to which we refer the reader.

In this paper we deal with the cross-correlation between the set of seismic events and the series of geoelectrical extreme values. Both the sequences of earthquakes and extreme events were modelled by means of stochastic point processes

$$s(t) = \sum_n A_n \delta(t - t_n) \quad (3.1)$$

where  $A_n$  is a coefficient proportional to the magnitude of the earthquake or to the value of the geoelectrical extreme event. Fixing a threshold for the coefficient  $A_n$ , we obtain the spike

train representing the seismic sequence and that describing the sequence of the extreme events. Figure 2a-c shows the time variation of the «normalized» (obtained subtracting the mean and dividing by the standard deviation) geoelectrical signal; the horizontal dotted lines indicate the thresholds  $T$  ( $1.5 \sigma$ ,  $2.0 \sigma$ ,  $2.5 \sigma$ ,  $3.0 \sigma$ ). An extreme event is defined as a value  $s$  of the signal so that  $|s| \geq T$ . Furthermore, the earthquakes ( $M \geq 2.5$ ) and the rain, recorded during the observation period, are also shown in the figure. The earthquakes, extracted from the instrumental seismic catalogue of the Istituto Nazionale di Geofisica e Vulcanologia, were selected as those contained in the Irpinia-Basilicata seismogenetic zone, in which the geoelectrical station was installed.



**Fig. 2a-c.** a) Normalized self-potential data recorded at Tito station during the 2001. The horizontal dotted lines indicates the thresholds  $T = \pm 1.5 \sigma$ ,  $\pm 2.0 \sigma$ ,  $\pm 2.5 \sigma$ , and  $\pm 3.0 \sigma$ . The Extreme Event (EE) is defined if  $|s| \geq T$ . b) Earthquakes (EQ) with magnitude  $M \geq 2.5$  which occurred in the investigation area during the 2001. c) Rain data (RR) recorded for the Tito site.

One way by which the information in an experimental sequence of events can be made more digestible is to reduce the data into statistics that emphasize a particular aspect of the data. The NWCCF is a Haar-wavelet-based version of the correlation function and is therefore insensitive to linear trends and non-stationarities. To calculate the NWCCF at a particular timescale  $\tau$ , the two spike trains first are divided into contiguous windows of duration  $\tau$ . The number of spikes  $Z_{1,n}$  falling within the  $n$ -th window are recorded for all indices  $n$  corresponding to windows lying entirely within the first spike-train data set. This process is repeated for the second spike train, producing  $Z_{2,n}$ . The difference between the count numbers in a given window in the first spike train ( $Z_{1,n}$ ) and the one after it ( $Z_{1,n+1}$ ) is then computed for all  $n$ , with a similar procedure followed for the second spike train. The NWCCF is defined as

$$\text{NWCCF}(\tau) = \frac{E \left\{ [Z_{1,n}(\tau) - Z_{1,n+1}(\tau)] [Z_{2,n}(\tau) - Z_{2,n+1}(\tau)] \right\}}{2 \left\{ E [Z_{1,n}(\tau)] E [Z_{2,n}(\tau)] \right\}^{\frac{1}{2}}} \quad (3.2)$$

where  $E \{ \dots \}$  indicates the average value. The normalization has two salutary properties: 1) it is symmetric in the two spike trains, and 2) when the same homogeneous Poisson point process is used for both spike trains the normalized wavelet cross-correlation function assumes a value of unity for all counting times  $\tau$ . A NWCCF  $\sim 0$  indicates that the processes are independent of each other. To determine the significance of a particular value for the normalized wavelet cross-correlation function at a certain timescale  $\tau$ , we compared it with the value at the same timescale of NWCCF calculated for synthetic data sets, that are shuffled versions of the original data sets (same interevent intervals but in random order).

Figure 3 shows the NWCCF( $\tau$ ) $\sim\tau$  relation between geoelectrical Extreme Events (EE) ( $|s| \geq 1.5 \sigma$ ) and the earthquakes (EQ) ( $M \geq 2.5$ ), and the NWCCF( $\tau$ ) $\sim\tau$  relation between the same series of EE and the rain (RR). In order to evaluate the significance of the results, we also

generated one hundred synthetic series of the three sequences. We generated the synthetic series by means of the random shuffling technique, well suited to test the significance of correlation in point processes (Viswanatham *et al.*, 1997; Duarte and Zatsiorsky, 2001; Pavlov *et al.*, 2001). The technique works as follows: first we generate a series of a normally distributed random numbers  $r(j)$ ,  $j = 1, \dots, N$ , where  $N$  is the number of the events (EE or EQ or RR). Next the series  $r(j)$  is rewritten in the order of increasing values, so we obtain the random sequence of indexes. The values in the original series are shuffled according to this sequence. The advantage of this method is that the synthetic series has the same mean, variance and probability density function of the inter-event times as the original one. Therefore, we calculated for each synthetic pair the NWCCF, then, we averaged all the synthetic NWCCFs. Figure 3 shows the mean synthetic NWCCF (thin line) and its  $1-\sigma_s$  range (dotted lines). We see that the mean surrogate NWCCF is approximately zero for almost all time-scales  $\tau$ , as expected for independent processes. The original NWCCF (bold line) can be considered at a particular timescale significantly different from zero if its value at that timescale is external to the surrogate  $1-\sigma_s$  range. In this case, we observe that the EE-EQ curve varies within the  $1-\sigma_s$  range of the shuffles, indicating that the EE with  $T = 1.5 \sigma$  are approximately independent of the EQ with magnitude  $M \geq 2.5$ . The correlations between EE and RR seem to be clearer at small time-scales (up to  $\sim 50$  h), indicating that the rain influences the generation of EEs, whose values exceed  $1.5 \sigma$ . At long timescales, the rain does not seem to be correlated with the EE series, since the NWCCF does not lie outside the shuffle standard deviation range. Figure 4 shows the NWCCF for EE-EQ and EE-RR pairs, with EE threshold  $T = 2.0 \sigma$ . We observe that the correlations between EE and EQ involve the intermediate timescales  $\sim 100$ - $300$  h, becoming rather evident at longer timescales. The EE series seems to cross-correlate with the RR series at small timescales, as observed in fig. 3; furthermore, intermediate timescales (up to approximately 500 h) also seem to be involved in the cross-correlated behaviour. Figure 5 shows the NWCCF

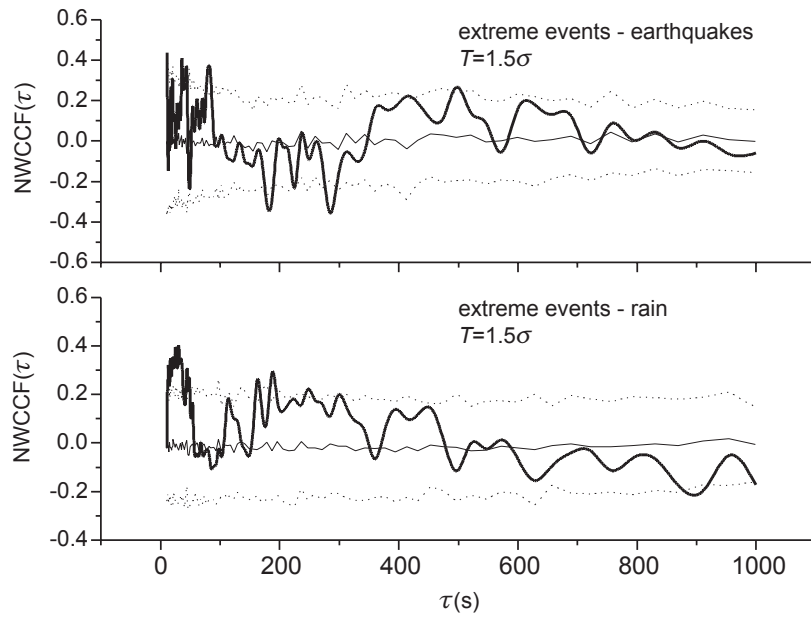


Fig. 3. NWCCF( $\tau$ )  $\sim$   $\tau$  relation for EE-EQ and EE-RR pairs, with  $T = 1.5 \sigma$ .

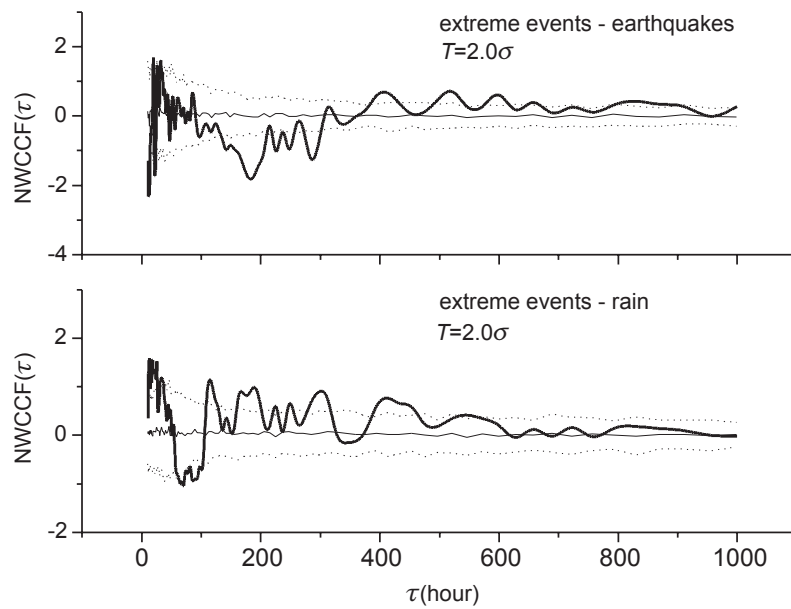
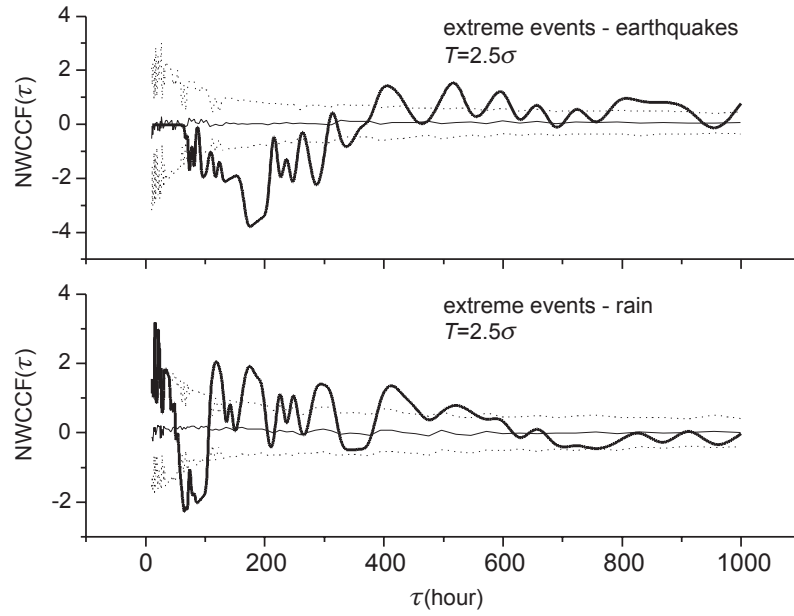
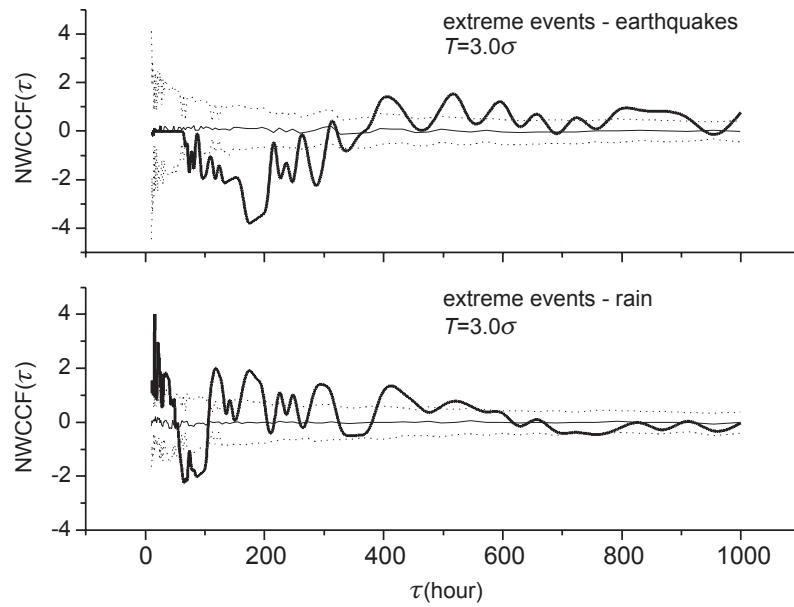


Fig. 4. NWCCF( $\tau$ )  $\sim$   $\tau$  relation for EE-EQ and EE-RR pairs, with  $T = 2.0 \sigma$ .



**Fig. 5.** NWCCF( $\tau$ )  $\sim$   $\tau$  relation for EE-EQ and EE-RR pairs, with  $T = 2.5 \sigma$ .



**Fig. 6.** NWCCF( $\tau$ )  $\sim$   $\tau$  relation for EE-EQ and EE-RR pairs, with  $T = 3.0 \sigma$ .



variations with a higher value for the threshold  $T = 2.5 \sigma$ ; both in the EE-EQ and EE-RR curves the cross-correlation is very clear. We can see that the original EE-EQ NWCCF is significantly negative at small timescales (from 100 to almost 300 h), and positive for approximately  $\tau > 400$  h. Furthermore, negative EE-RR NWCCF characterizes the cross-correlation from approximately 50 to 100 h. Positive NWCCF is displayed at intermediate timescales, up to  $\sim 500$  h, although in the range 300-400 h the rain and the geoelectrical data seem to be independent of each other. For timescales  $\tau > 500$  h the rain does not seem to cross-correlate with the series of self-potential anomalous values, which, instead, show a significant correlation with the earthquakes. Figure 6 shows the NWCCF for the pairs EE-EQ and EE-RR, for the threshold  $T = 3.0 \sigma$ : the results are approximately similar to those obtained for  $2.5 \sigma$  threshold.

The results obtained by the NWCCF analysis suggest that:

i) The rain could influence the generation of extreme events in geoelectrical signals at small and intermediate timescales ( $< 500$  h), while for longer timescales the processes seem to be independent of each other.

ii) The earthquakes seem to be correlated with the geoelectrical anomalous values for several timescale ranges.

iii) The cross-correlated behaviours seem to be independent of the threshold of selection of the geoelectrical anomalous values, especially when the threshold is high.

iii) Excluding the timescale ranges, characterized by the correlation between rain and self-potential data (that cannot allow us to draw firm conclusions on the correlation between self-potentials and earthquakes), our findings can allow us to objectively establish that earthquakes could be responsible for the generation of geoelectrical extreme values at long timescale ranges, suggesting a transfer of information between the geophysical processes governing them.

## 5. Conclusions

In this work we performed a correlation analysis by means of the NWCCF between the

series of extreme events in the time series of hourly self-potential means, recorded at monitoring geoelectrical station Tito (Southern Italy) and earthquakes which occurred in the investigation area during 2001. We found significant correlations between geoelectrical anomalous values and seismic events in long timescale ranges ( $> 500$  h), indicating correlated mechanisms of generation of both anomalous self-potential data and earthquakes. Longer data sets can improve the analysis, especially in investigating the presence of a relationship between the values of the extreme events and the magnitudes of the earthquakes, or between the intensity of the self-potential anomaly and the distance of the earthquake from the station location. The results presented in this paper, although still preliminary disclose the potential of the NWCCF, and similar investigations performed over other geophysical data sets measured in seismic areas can give a contribution to the earthquake prediction problem.

## REFERENCES

- CUOMO, V., G. DI BELLO, V. LAPENNA, M. MACCHIATO and C. SERIO (1996): Parametric time series analysis of extreme events in electrical earthquake precursors, *Tectonophysics*, **262** (1-4), 159-172.
- CUOMO, V., V. LAPENNA, M. MACCHIATO and C. SERIO (1997): Autoregressive models as a tool to discriminate chaos from randomness in geoelectrical time series: an application to earthquake prediction, *Ann. Geofis.*, **XI** (2), 385-400.
- CUOMO, V., G. DI BELLO, V. LAPENNA, M. MACCHIATO, C. SERIO and L. TELESCA (1998): Linear and nonlinear dynamics in electrical precursory time series: implications for earthquake prediction, *Tectonophysics*, **287** (1-4), 279-298.
- DI BELLO, G., V. LAPENNA, C. SATTRIANO and V. TRAMUTOLI (1994): Self-potential time series analysis in a seismic area of the Southern Apennines: preliminary results, *Ann. Geofis.*, **XXXVII** (suppl. to n. 5), 1137-1148.
- DOGLIONI, C., P. HARABAGLIA, G. MARTINELLI, F. MONGELLI and G. ZITO (1996): A geodynamic model of the Southern Apennines accretionary prism, *Terra Nova*, **8**, 540-547.
- DUARTE, M. and V.M. ZATSIORSKY (2001): Long-range correlations in human standing, *Phys. Lett. A*, **283**, 124-128.
- EKSTRÖM, G. (1994): Teleseismic analysis of the 1990 and 1991 earthquakes near Potenza, *Ann. Geofis.*, **XXXVII** (6), 1591-1599.
- EVANS, J.R. (Editor) (1997): Assessment of schemes for earthquake prediction: Editor's introduction, *Geophys. J. Int.*, **131** (3), 413-420.



- GELLER, R.J. (Editor) (1996): Debate on VAN method, *Geophys. Res. Lett.*, **23** (11), 1291-1452.
- LAPENNA, V., M. MACCHIATO, D. PATELLA, C. SATRIANO, C. SERIO and V. TRAMUTOLI (1994): Statistical analysis of non-stationary voltage recordings in geoelectrical prospecting, *Geophys. Prospect.*, **42**, 917-952.
- LOWEN, S.B., T. OZAKI, E. KAPLAN, B.E.A. SALEH and M.C. TEICH (2000): Fractal features of dark, maintained, and driven neural discharges in the cat visual system, in *Methods: a Companion to Methods in Enzymology*, edited by T. SMITH and D. LANGE (Academic Press, San Diego), **27**, 377-394.
- MALLET, R. (1862), *The first principle of observational seismology as developed in the report to the Royal Society of London of the expedition made by command of the Society into the interior of the Kingdom of Naples to investigate the circumstances of the great earthquake of December 1857, London* (reprint by ING, Roma, 1987).
- NUR, A. (1972): Dilatancy pore fluids, and premonitory variations of  $t_p/t_s$  travel times, *Bull. Seismol. Soc. Am.*, **62**, 1217-1222.
- PANTOSTI, D. and G. VALENSISE (1990): Faulting mechanism and complexity of the November 23, 1980, Campania Lucania earthquake, inferred from surface observations, *J. Geophys. Res.*, **95** (B10), 15329-15341.
- PARK, S.K (1996): Precursors to earthquakes: seismo-electromagnetic signals, *Surv. Geophys.*, **17**, 493-516.
- PATACCA, E., P. SCANDONE, M. BELLATALLA, N. PERILLI and U. SANTINI (1988): L'Appennino Meridionale: modello strutturale e palinspastica dei domini esterni, in *74° Congresso della Società Geologica Italiana: «L'Appennino Campano-Lucano nel Quadro Geologico dell'Italia Meridionale»*, Sorrento 17 Settembre 1988, 67-69 (relazioni).
- PAVLOV, A.N., W. EBELING, L. MOLGEDEY, A.R. ZIGANSHINA and V.S. ANISHCHENKO (2001): Scaling features of texts, images and time series, *Physica A*, **300**, 310-324.
- RIKITAKE, T. (1988): Earthquake prediction: an empirical approach, *Tectonophysics*, **148**, 195-210.
- SCHOLZ, C.H. (1990): *The Mechanics of Earthquakes and Faulting* (Cambridge University Press), pp. 439.
- SCHOLZ, C.H., L.R. SYCHES and Y.P. AGGARWAL (1973): Earthquake prediction: a physical basis, *Science*, **181**, 803-810.
- TERTULLIANI, A., M. ANZIDEI, A. MARAMAI, M. MURRU and F. RIGUZZI (1992): Macroseismic study of the Potenza (Southern Italy) Earthquake of 5 May 1990, *Natural Hazards*, **6**, 25-38.
- VISWANATHAN, G.M., C.-K. PENG, H.E. STANLEY and A.L. GOLDBERGER (1997): Deviations from uniform power law scaling in nonstationary time series, *Phys. Rev. E*, **55**, 845-849.
- WYSS, M. (Editor) (1991): *Evaluation of Proposed Earthquake Precursors* (American Geophysical Union, Washington), pp. 94.
- WYSS, M. (1997): Cannot earthquakes be predicted?, *Science*, **278**, 487-488.