

Article

Cross-Correlation among Seismic Events, Rainfalls, and Carbon Dioxide Anomalies in Spring Water: Insights from Geochemical Monitoring in Northern Tuscany, Italy

Lisa Pierotti ¹, Cristiano Fidani ^{2,3,*} , Gianluca Facca ¹ and Fabrizio Gherardi ¹ 

¹ Istituto di Geoscienze e Georisorse CNR, 56124 Pisa, Italy; lisa.pierotti@igg.cnr.it (L.P.); gianluca.facca@igg.cnr.it (G.F.); fabrizio.gherardi@igg.cnr.it (F.G.)

² Istituto Nazionale di Geofisica e Vulcanologia, 00143 Roma, Italy

³ Central Italy Electromagnetic Network, 63020 Fermo, Italy

* Correspondence: cristiano.fidani@ingv.it

Abstract: Variations in the CO₂ dissolved in water springs have long been observed near the epicenters of moderate and strong earthquakes. In a recent work focused on data collected during the 2017–2021 period from a monitoring site in the Northern Apennines, Italy, we noticed a significant correlation between CO₂ anomalies and moderate-to-weak seismic activity. Here, we extended this analysis by focusing on data collected from the same site during a different period (2010–2013) and by integrating the CENSUS method with an artificial neural network (ANN) in the already-tested protocol. As in our previous work, a fit of the computed residual CO₂ distributions allowed us to evidence statistically relevant CO₂ anomalies. Thus, we extended a test of the linear dependence of these anomalies to seismic events over a longer period by means of binary correlations. This new analysis also included strong seismic events. Depending on the method applied, we observed different time lags. Specifically, using the CENSUS methodology, we detected a CO₂ anomaly one day ahead of the earthquake and another anomaly eleven days ahead. However, no anomaly was observed with the ANN methodology. We also investigated possible correlations between CO₂ concentrations and rain events and between rain events and earthquakes, highlighting the occurrence of a CO₂ anomaly one day after a rain event of at least 10 mm and no linear dependence of seismic and rain events. Similar to our previous work, we achieved a probability gain of around 4, which is the probably of earthquake increases after CO₂ anomaly observations.

Keywords: CO₂ anomalies; continuous monitoring; small earthquakes; statistical correlations; conditional probability



Citation: Pierotti, L.; Fidani, C.; Facca, G.; Gherardi, F. Cross-Correlation among Seismic Events, Rainfalls, and Carbon Dioxide Anomalies in Spring Water: Insights from Geochemical Monitoring in Northern Tuscany, Italy. *Water* **2024**, *16*, 739. <https://doi.org/10.3390/w16050739>

Academic Editors: Galina Kopylova and Svetlana Boldina

Received: 27 November 2023

Revised: 22 February 2024

Accepted: 24 February 2024

Published: 29 February 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The first observations of high concentrations of CO₂ in soil gas in correspondence with faults and major fractures were made in the Caucasus [1], and one of the earliest examples of the long-time monitoring of escaping gas during the occurrence of earthquakes was carried out in southern Dagestan [2]. The possible link between earthquake occurrence and gas release at the surface had been considered in relation to strong seismic events [3,4]. Water springs with elevated levels of dissolved CO₂ have been observed worldwide around seismic zones [5], and the presence of high concentrations of CO₂ of predominantly mantle and/or metamorphic origins has been considered a major factor controlling pore pressure in fractures. Moreover, carbon dioxide degassing has been a recurrent, observed phenomenon in many active tectonic regions [6,7]. Several observations of CO₂ and H₂ emanations nearby faults can prove to be a powerful means of identifying the state of the activity of those same faults [8]. Therefore, models based on the dilatancy phenomenon have been proposed to relate CO₂ and other gaseous emissions to earthquakes [9]. A close interplay between local magmatism and fluid release at the surface has been hypothesized

also in Iceland [10], and the link there was explained by advancing a “dual-mechanism model”. This model considers the reciprocal influence among fluids ascending from below the brittle crust and plate motion building strain [11]. During earthquakes, deep-seated fluids migrate upward through faults, resulting in micro-earthquakes.

CO₂ excesses have also been observed in extensional domains [12]. The coincidence of CO₂ emissions in highly seismic regions has been confirmed on occasion of major events [13,14], in many aftershock evolutions [15–20], as well as with small seismic events [21]. The release of deep-seated CO₂ may occur by diffuse venting/soil degassing or concentrated flow through lakes, vents, and hot and cold springs [22]. The amount of gaseous carbon escaping into the atmosphere is under quantification at a global scale [23]. The flux of endogenous CO₂ can affect aquifer dynamics at a regional scale and can be traced by variations in the total concentration of dissolved inorganic carbon and its isotope signature [24–27].

From a different perspective, high-pressure fluids can induce the movements of faults [28]. Specifically, the rise of fluids along fault planes can contribute to weakening processes in rocks [29]. Techniques resolving changes in V_p and V_p/V_s during strong seismic activity have recently suggested that a large volume of fluids pushed up, mobilizing groundwater [30], both in Central [31] and Northern Italy [32]. These have a potential impact on water precursory phenomena [33,34]. Simulation experiments have also been performed in underground laboratories, which confirmed the influence of fluids in the seismogenic process [35]. The occurrence of fluid-triggered seismicity has also been proven [36]. Finally, a relationship between earthquake dilatancy and resistivity has been proposed [37]. Moreover, pressure-impressed electric currents due to gas passage through fractured asperities have been observed in laboratory experiments [38]. These experiments were used to speculate a causal mechanism of the seismo-EM phenomena observed in recent earthquakes [39].

Since the early 1990s, the potential relationship between CO₂ degassing and earthquakes has been inferred by manual techniques in Italy [40]. The processing of continuous CO₂ time series began in the early 2000s with the operation of automatic monitoring stations [41–43]. This study aims to statistically assess the possible influence of low-magnitude earthquakes on the release of CO₂ at Galliciano, Northern Tuscany, where one of the automatic stations of the Geochemical Network of Tuscany (GNT) is currently monitoring selected geochemical parameters in the water issued by the local thermo-mineral spring. Following a satellite data processing procedure [44], we applied a new statistical interpretative approach to investigate CO₂ time series recorded by the Galliciano station during the period of 2010–2013. This procedure relies on the selection of relevant gas anomalies and on the identification of anomalies statistically correlated to small earthquakes, as has already been successfully carried out for the CO₂ time series collected during the period of 2017–2021 [45]. Compared to this previous study, in this paper we have considered the incorporation of artificial neural network (ANN) techniques in the identification of possible geochemical anomalies. Moreover, for the period of 2010–2013, we have analyzed not only weak main shocks, as had been performed in our past study [45], but also a second set of main shocks that comprises strong (M5+) events with major aftershocks. The earthquakes examined in this paper have been also characterized in terms of the ratio between the epicentral distance from the spring and the linear length of the seismic source. Finally, what is new in comparison with our previous work [45] is our verification of rain magnitude's influence on dissolved CO₂ variations, using two levels of rainfall in the procedure. We carried out this study with the multiple purpose of (i) validating the results of the 2017–2021 period [45], (ii) comparing the performance of different statistical approaches, and (iii) assessing in more detail any possible correlation between rainfall and seismic events. This ultimately helps in building confidence in assessing and quantifying mutual correlations among low-to-moderate seismic events, geochemical anomalies, and rain events by the integration of advanced statistical methods in the “conventional” analysis of hydrogeochemical data. The proposed statistical methods also allow for estimating

the conditional probabilities of earthquake occurrence around the Galliciano site, which represents a significant step forward in terms of forecasting improvement.

2. Seismicity

The Galliciano spring (Figure 1) is located in a sector of northwestern Apennines (Garfagnana Valley) where a NW–SE trending cluster of seismic activity is observed.

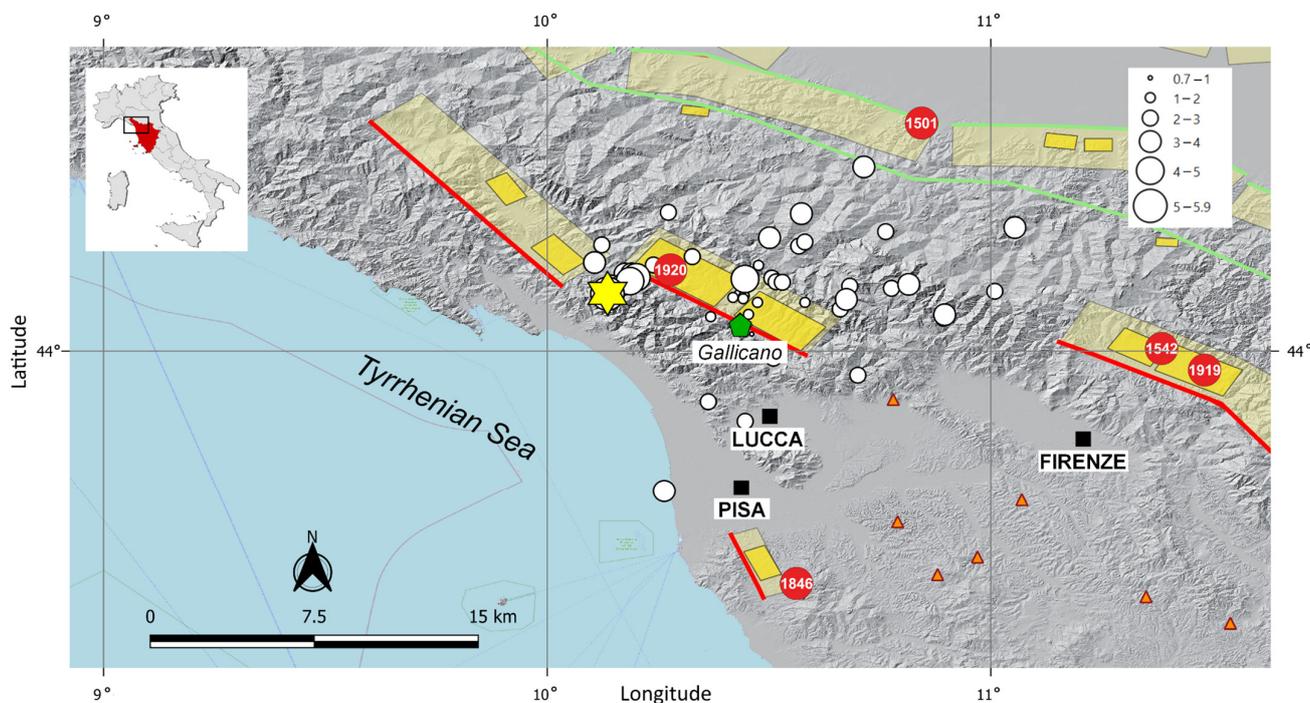


Figure 1. Location of the monitoring site at Galliciano (green pentagon). White dots = epicenters of the seismic events considered in this study; yellow star = epicenter of the main seismic event in Tuscany of 2010–2013; red dots = $M_W \geq 6$ earthquake after Rovida et al. [46] catalog (CPTI15 v.4.0); red lines = normal faults; green lines = inverse faults; yellow transparent boxes = seismogenic sources (after DISS Working group, 2021, version 3.3 [47]); and orange triangles = gaseous emissions points extracted from the Googas online catalog <http://hdl.handle.net/2122/3060> (accessed on 23 February 2024).

The area is characterized by an elevated seismic risk [48,49]. Several earthquakes with a magnitude (M_W) larger than 5.5 occurred in this area. Strong historical events are indicated by red circles in Figure 1. On the 14 August 1846, an earthquake with an estimated M_W of 6.0 occurred around 60 km south of the Galliciano spring, near the Orciano Pisano village (Colline Pisane earthquake). Two historical earthquakes took place in the Mugello Valley, to the NE of Firenze (13 June 1542: estimated magnitude of 6.0; 29 June 1919: estimated magnitude of 6.2). On the 5 June 1501, an earthquake with an estimated magnitude of 5.9 occurred near the Maranello village, in the Modena province, on the northwestern slopes of the Apennines. Lastly, the Lunigiana and Garfagnana Valleys experienced a catastrophic event of an estimated magnitude of 6.5 on the 7 September 1920 [46].

We retrieved information on regional earthquakes in the period of 2010–2013 from the ISIDE catalog <http://iside.rm.ingv.it> (accessed on 23 February 2024), and we used the International Catalog of Earthquakes to estimate the moment magnitude of the strongest earthquakes <https://earthquake.usgs.gov/earthquakes/search> (accessed on 23 February 2024). Our database considered 4046 seismic events with a magnitude greater than or equal to 0.7 that occurred within a radius of 50 km from the Galliciano site. They are reported by small squares whose dimensions are proportional to their magnitudes on the geographical

area on the left of Figure 2. The seismic events were declustered following the Reasenberg algorithm [50], using the publicly available routine of Z-map [51]; they are reported on the geographical area in Figure 2, in the middle. Figure 2, on the top right, describes the maximum curvature solution of the Gutenberg–Richter law, where gray events represent the non-cumulative distribution. The resulting magnitude of completeness is 1.5, and, in this region, b -value = 0.90 ± 0.05 and a -value = 3.909 ± 0.002 represent the coefficients of the red line. Figure 2, on the bottom right, highlights the depth distribution of the seismic events within the limit of 50 km chosen for this study.

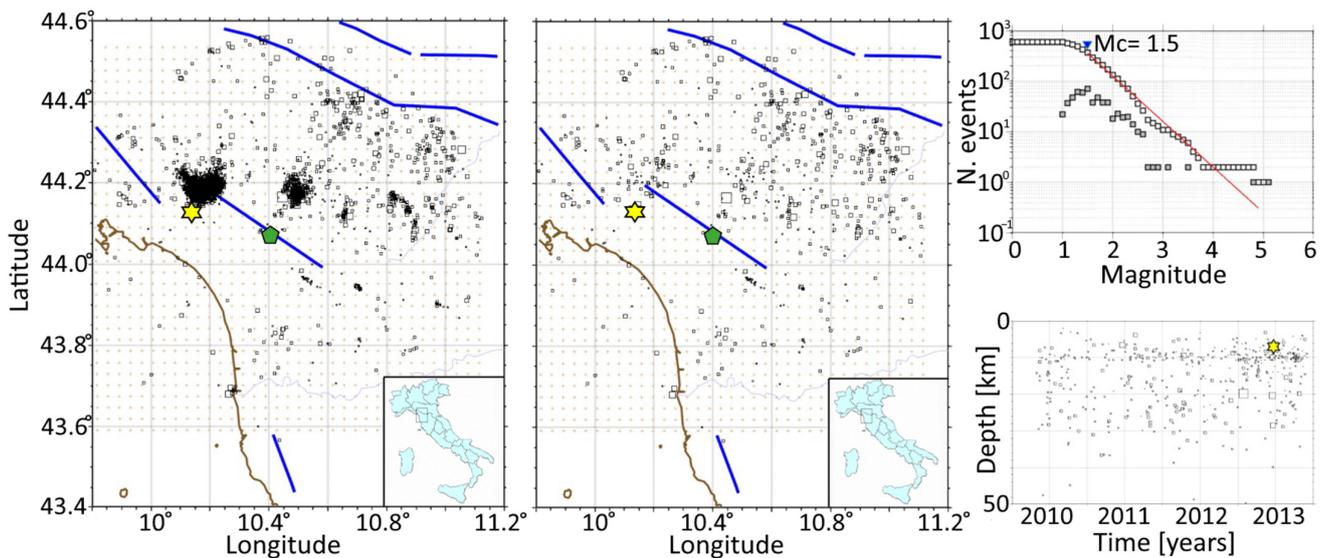


Figure 2. The distribution of the earthquakes from the dataset used in this analysis is on the left. Brown line: Tyrrhenian coastline; blue line: main local faults; green pentagon: Gallicano monitoring site; and yellow star: location of the most energetic seismic event occurred in the period of interest. The declustered earthquakes are distributed on the map in the middle. The frequency–magnitude distribution and the earthquakes’ depth distribution are on the right.

We identified two groups of declustered earthquakes consisting of the largest earthquakes of the day. A first set of shocks was obtained by collecting events occurring around the Gallicano spring (≤ 50 km) with a magnitude greater than or equal to 0.7, to which aftershocks were added again (46 events in total). A second set of main shocks ($M_l \geq 1.5$), which also includes two main events ($M_W = 5.9$ and $M_W = 5.8$), occurred in the Modena province, outside the Tuscany region (forty events in total). The identification of the first set of events represents a major advancement of our previous study [45] and allows for a deeper understanding of the possible influence of aftershocks on dissolved CO_2 . Moreover, the selection criterion for both sets was their distance from Gallicano being limited by their energy by three Dobrovolsky radii [52], as in [45]. The set of seismic events is reported in Table 1 with colors to highlight differences between the two groups. M_W was reported in column 7 for the major events so as to directly calculate the fault linear length from [53], as follows:

$$M_W = 4.24 + 1.67 \log(L_f) \quad (1)$$

where L_f is the fault length, considered to be adapted to the normal and inverse fault mechanisms commonly observed in Italy [54]. An M_l above 4 was also used with (1) when the M_W was not retrieved, as the difference between M_W and M_l can be ignored [54] for these magnitudes, while an approximate equation was used for the M_W of the smaller events [55]:

$$M_W = \frac{2}{3} M_l + 1.15 \quad (2)$$

where the fault area $A = L_f^2$ is related to [53].

$$M_w = 4 + \log A \tag{3}$$

Table 1. The entire set of seismic events used in this study is differentiated with a red color for the first group and with a green color for the second one. M_W is reported in column 7 when it is available from international databases or retrieved with Equation (2). Column 8 reports the distance between the epicenter and Galliciano. Column 9 reports the linear lengths of the seismic sources retrieved with Equation (1), when the M_W is available from international databases (*), or retrieved with Equation (3).

N.	Date and Time [YYYY-MM-DD hh:mm:ss]	Latitude [°]	Longitude [°]	Depth [km]	Magnitude M_I (INGV)	Magnitude M_W (2) (* USGS)	Distance from Galliciano [km]	L_f [km]	D/L_f
1	2010-07-14 09:10:42	10.3630	43.8850	7.4	2.4	2.7	20.3	0.25	81.2
2	2010-11-03 15:28:48	10.5810	44.1110	10.3	1.9	2.4	13	0.17	76.47
3	2010-12-29 02:15:35	10.5100	43.9840	5.3	2.1	2.5	10.3	0.2	51.5
4	2011-04-15 01:39:15	10.4330	44.1290	10.7	1.5	2.1	7.8	0.12	65
5	2011-05-22 12:26:48	10.4610	44.0390	4.9	0.7	1.6	3.2	0.07	45.71
6	2011-06-19 14:35:34	10.8150	44.1520	6.5	3.6	3.5	32.2	0.63	51.11
7	2011-07-12 04:11:39	10.4320	44.0110	1.3	0.8	1.7	5.5	0.07	78.57
8	2011-07-17 21:50:57	10.5670	44.2390	17.0	2.4	2.7	22.5	0.25	90
9	2011-08-20 20:02:48	10.8930	44.0750	3.9	2.9	3.1	36.7	0.37	99.19
10	2011-09-19 06:24:16	10.7630	44.2720	9.8	2.6	2.9	35.4	0.29	122.07
11	2011-10-26 16:26:28	10.7760	44.1430	12.4	2.4	2.7	29.2	0.25	116.8
12	2011-11-06 01:33:59	10.6820	44.1480	17.5	2.3	2.7	32.2	0.23	140
13	2012-01-22 05:31:15	10.4740	44.1110	10.1	1.9	2.4	6.6	0.17	38.82
14	2012-03-19 19:33:03	10.1230	44.2420	5.0	2.7	2.9	32.5	0.32	101.56
15	2012-03-24 19:18:04	10.2730	44.3160	19.0	2.8	3	31.5	0.34	92.65
16	2012-04-12 23:01:48	10.5730	44.3130	22.3	3.1	3.2	30.5	0.43	70.93
17	2012-04-13 22:13:56	10.8960	44.0830	9.5	3.3	3.3	37.1	0.5	74.2
18	2012-05-18 20:10:52	10.1752	44.1757	9.1	2.4	2.7	24.5	0.25	98
19	2012-05-20 02:03:53	11.2280	44.8890	6.3	6	5.9 *	113.7	11.29	9.55
20	2012-05-29 07:00:03	11.0860	44.8510	8.1	5.8	5.8 *	103.4	8.57	12.06
21	2012-08-28 01:56:14	10.6605	44.0958	16.2	2.4	2.7	18.7	0.25	74.8
22	2012-11-26 19:18:55	10.6742	44.1178	21.6	3.2	3.3	20.5	0.46	44.56
23	2012-12-04 16:38:27	11.0090	44.1363	8.8	2.9	3.1	46.8	0.37	126.49
24	2012-12-16 11:24:27	10.4437	44.1258	7.0	1.6	2.2	7.5	0.13	57.69
25	2013-01-25 14:48:18	10.4458	44.1643	19.8	4.8	4.9 *	11.6	2.48	4.68
26	2013-02-09 02:19:10	10.3685	44.0785	6.1	1.7	2.3	5.6	0.15	37.33
27	2013-02-20 11:09:59	10.5300	44.1568	13.5	3.0	3.1	13.4	0.4	33.5
28	2013-03-29 12:59:21	10.4187	44.1222	8.4	1.7	2.3	7.2	0.15	48
29	2013-04-01 23:35:11	10.4763	44.1953	14.8	2.0	2.5	15.5	0.18	86.11
30	2013-04-05 03:23:42	10.5065	44.1662	15.0	2.3	2.7	13.2	0.23	57.39
31	2013-04-08 11:38:57	10.5160	44.1578	13.5	2.6	2.9	12.9	0.29	44.48
32	2013-05-02 17:10:14	10.4537	44.0837	6.8	1.5	2.1	3.1	0.12	25.83
33	2013-05-13 18:35:47	10.4462	43.8403	8.3	2.5	2.8	24.3	0.27	90
34	2013-06-03 15:33:12	10.4418	44.1193	10.0	1.8	2.3	6.7	0.16	41.87
35	2013-06-14 18:22:23	11.0543	44.2812	20.3	3.7	3.6	55.4	0.68	81.47
36	2013-06-15 19:41:16	10.1462	44.1482	10.4	3.4	3.6 *	25.2	0.31	81.29
37	2013-06-19 10:51:07	10.7138	44.4193	28.5	3.5	3.5 *	45.8	0.36	127.22
38	2013-06-21 10:33:57	10.1357	44.1308	7.0	5.1	5.1*	25.1	3.26	7.7
39	2013-06-23 13:24:01	10.2008	44.1683	9.2	4.4	4.4 *	22.3	1.24	17.98
40	2013-06-29 05:21:34	10.1698	44.1513	5.9	3.5	3.5	23.5	0.58	40.52

Table 1. Cont.

N.	Date and Time [YYYY-MM-DD hh:mm:ss]	Latitude [°]	Longitude [°]	Depth [km]	Magnitude M_i (INGV)	Magnitude M_W (2) (* USGS)	Distance from Gallicano [km]	L_f [km]	D/L_f
41	2013-06-30 02:57:42	10.1867	44.1595	6.1	4.5	4.5 *	22.8	1.43	15.94
42	2013-07-09 09:12:57	10.1777	44.1625	6.2	3.1	3.2	23.7	0.43	55.12
43	2013-07-12 18:07:22	10.1067	44.2017	10.0	3.5	3.5	30.6	0.58	52.76
44	2013-08-10 13:11:09	10.2393	44.1962	5.8	2.7	2.9	21.9	0.32	68.44
45	2013-08-14 15:35:50	10.1903	44.1758	9.3	3.5	3.5 *	23.5	0.36	65.28
46	2013-08-21 05:32:13	10.5807	44.2485	15.4	2.8	3	24.1	0.34	70.88
47	2013-10-03 19:07:15	10.3270	44.2150	10.1	2.2	2.6	19.5	0.21	92.86
48	2013-10-19 12:29:35	10.2638	43.6807	10.5	3.5	3.5	43.7	0.58	75.34
49	2013-10-25 09:17:00	10.5012	44.2582	19.8	3.1	3.2	22.8	0.43	53.02
50	2013-11-19 21:05:48	10.7003	43.9453	5.0	2.2	2.6	25.2	0.21	120

The distance from the epicenters to the Gallicano spring was reported in column 8, and the linear length of the seismic sources was reported in column 9 of Table 1.

3. Hydrogeological and Hydrogeochemical Setting

The Gallicano spring emerges, with an average outflow of 1.5 L/s, on the Apuan side of the Serchio basin, in the bed of the “Turrte di Gallicano” river (right tributary of the Serchio river), in correspondence with a NW–SE-trending synthetic quaternary fault of the Serchio graben [49]. Regarding the low thermality of the waters ($T \sim 24$ °C), De Stefani [56] reports the following: “. . . .they are less saline and less hot waters due to the greater quantity of superficial and river waters that mix in them”. Chemical analyses available in the literature [57,58] indicate, for the Gallicano spring, a Cl-SO₄-Ca chemical composition and a salinity between 3 and 4 g/L. Near the spring, in the bed of the “Turrte di Gallicano” river, gaseous emissions have been reported for a long time [56], with an abundant presence of helium (7940 ppb dissolved in water and 2360 ppb directly from the gaseous vents), attributed to the presence of a fault crossing [57].

Within a seismic prevention/prediction program, financially supported by the Regional Government of Tuscany, Italy, for studying possible geochemical precursors to seismic activity [59], the Gallicano spring was selected as a suitable site for the installation of a continuous automatic monitoring station of selected chemical–physical parameters of the water (for the selection criteria, see [41]). The automatic station was installed on the 15 April 2003, and it is still operative, with two considerable periods of inactivity (1 December 2008–30 June 2010 and 21 October 2013–12 December 2016). In addition to continuous monitoring, since 2003, the spring has been sampled every month for chemical and isotopic analyses [41].

To describe the deep hydrogeological structures which, in the Serchio river basin, constitute the locus of the deep hydrothermal circulation feeding also the Gallicano spring and to understand the mode of this circulation, the hydrostructural model of the northern Apennines in Figure 3A (modified after [60]) and the circulation model in Figure 3B were considered. In this model, the geological formations have been grouped into three hydrogeological units based on their permeability characteristics. The following hydrostructural elements are distinguished, from top to bottom: impermeable cover (IC), carbonate aquifer system (CAS), and impermeable substratum (IS).

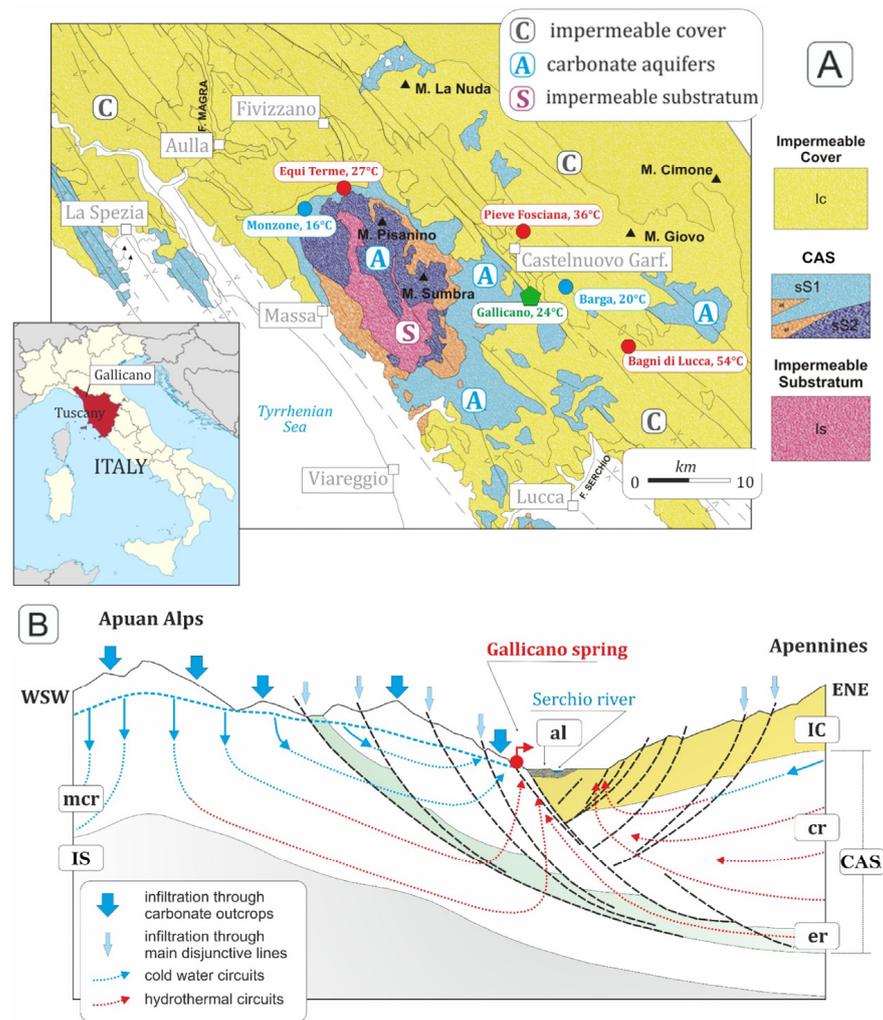


Figure 3. (A) Hydrostructural model of the Northern Apennines (modified after [60]). (B). Circulation model of the Gallicano hydrothermal system (al = alluvium; IC = impermeable cover; CAS = carbonate aquifer system; cr = non-metamorphic carbonate rocks; er = evaporites; mcr = permeable metamorphic carbonate rocks; and IS = impermeable substratum). Not to scale.

The impermeable cover (IC) consists of shales and sandstones of Tuscan nappe and by the clayey complex of the Liguride formations and Pliocene—Quaternary clay. It should be highlighted that the IC, thus defined with respect to the underlying CAS, hosts modest aquifer levels, which do not modify its hydrogeological role but are, in any case, sites of superficial circulation which feed springs characterized by near-ambient temperatures, very low salinity (<200 mg/L), and predominant Ca-HCO₃ facies (type-A water in Figure 4). The IC allows for connections between the underlying CAS and the outside, especially in correspondence with the cataclastic bands associated with NW–SE-trending antithetic quaternary faults; the latter can act as recharge zones as well as ascent paths for hydrothermal fluids.

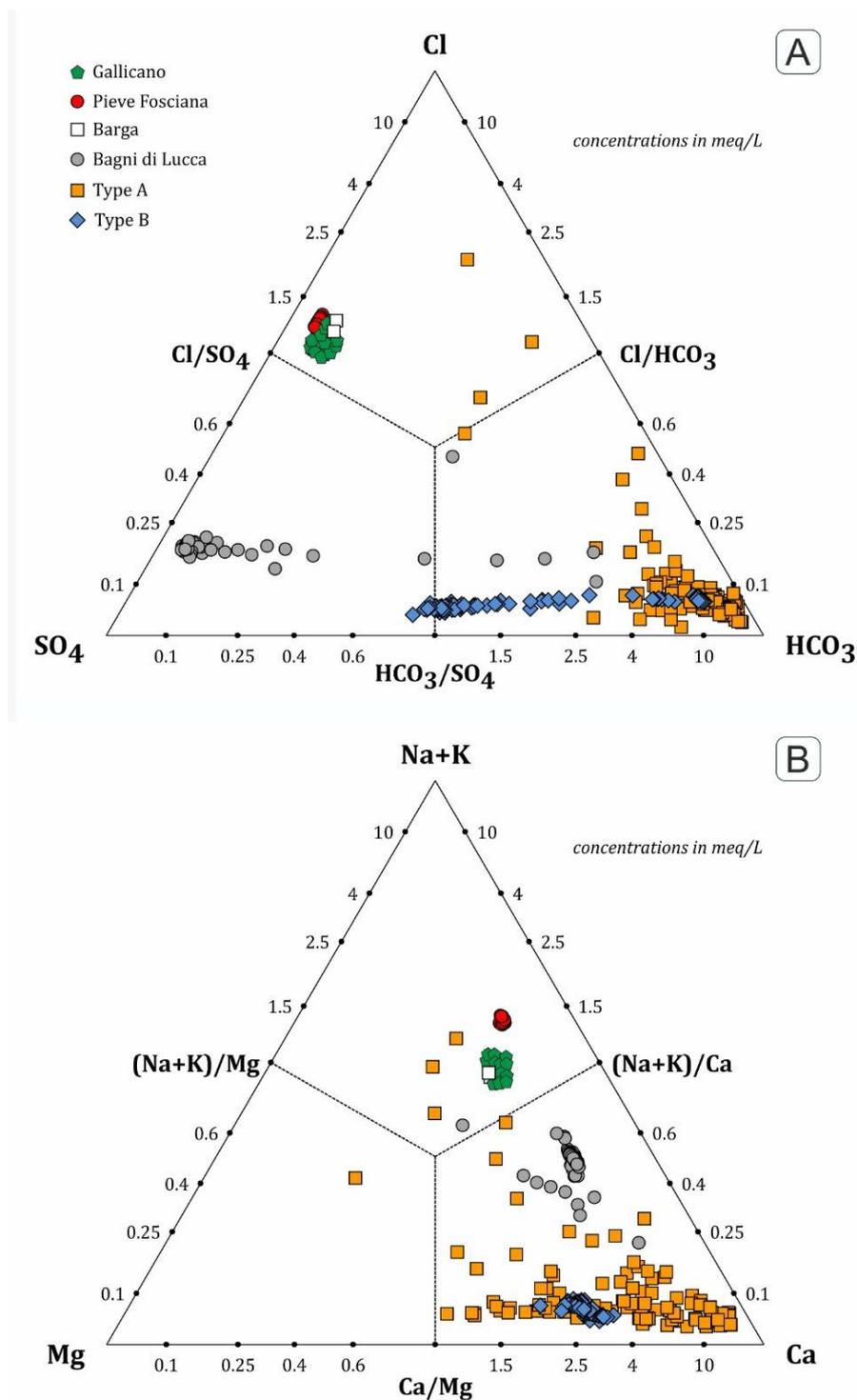


Figure 4. Triangular diagrams for the Serchio basin springs. Anions (A) and cations (B) relative to the total concentrations are computed on a meq/L basis. Green pentagon: Gallicano spring; red dots: Pieve Fosciana spring; white square: Barga spring; grey dots: Bagni di Lucca spring; yellow square: type-A water; and blue diamond: type-B water.

The CAS is composed of Mesozoic evaporitic and carbonate rocks belonging to non-metamorphic (subset Ss1) and metamorphic (subset sS2) Tuscan units, intercalated by low-permeability layers, which cause aquicludes and aquitards (at). If the low-permeability limit outcrops along their entire perimeter, the hydrostructures of the CAS give rise to

perched aquifers feeding depression, contact, and overflow springs, characterized by near-ambient water temperatures, low or high salinity (up to 1.6 g/L), and chemical composition varying from $\text{HCO}_3\text{-Ca}$ to $\text{SO}_4\text{-Ca}$ (type-B water in Figure 4). If the low-permeability limit deepens, the hydrostructures of the CAS feed hydrothermal circuits. It is noteworthy that the considerable depth of the low-permeability limit on the north-eastern side of the Serchio graben ensures the high piezometric loads and hydraulic gradients associated with these hydrostructures, allowing the large development of hydrothermal circuits in the Serchio Valley. In the eastern sector of the Serchio Valley (Apennines), the CAS is sandwiched between the impermeable cover (IC) and the low-permeability substratum (IS) and feeds thermo-artesian springs with AN ascent pattern through cataclastic belts associated with the NW–SE fault systems of the eastern side of the Serchio graben. The abundant presence of the IC ensures considerable protection to deep-circulating water, limiting possible mixing with the cold and hyposaline waters of more superficial circulation mechanisms. On the Apuan side, the CAS outcrops abundantly, while the IC is downthrown by the NW–SE normal faults that border the western side of the Serchio graben system. Therefore, the IC acts as a hydraulic barrier for water from the deep circulation system that rises to the surface, by overflow, driven by the synthetic neotectonic faults of the Serchio graben. Due to the lack of an impermeable cover, cold and dilute waters of superficial circulation mechanisms recharged by local rainwater are locally allowed to mix with deep waters. For this reason, the thermo-artesian springs of Pieve Fosciana ($T = 36\text{ }^\circ\text{C}$) and Bagni di Lucca ($T = 54\text{ }^\circ\text{C}$), which emerge on the eastern side of the basin, have higher temperatures than the thermo-mineral springs that emerge on the western side, such as the Gallicano spring. The ternary diagrams of Figure 4 allow one to identify the chemical facies of a thermo-mineral spring based on reciprocal relationships between the main anions and cations (Pieve Fosciana: Cl-Na facies; Bagni di Lucca: $\text{SO}_4\text{-Ca}$ facies; Gallicano and Barga: Cl-Na-Ca facies) acquired by interaction with the anhydrite and/or halite layers of a deep Triassic carbonate–evaporite reservoir [41,43].

The impermeable substratum, IS, is the low-permeability limit of the CAS and is represented by the Paleozoic basement. It mainly consists of quartzites, conglomerates, sandstones, and phyllites.

Continuous Geochemical Monitoring during the 2010–2013 Period

By elaborating the time series collected during the period of 2010–2013, we achieved an accurate definition of the geochemical background of the natural system. Table 2 and Figure 5 show the descriptive statistic and the continuous geochemical signals, respectively, of water temperature ($^\circ\text{C}$ —box A), electrical conductivity ($\mu\text{S}/\text{cm}$ —box B), pH (box C), redox potential (mV—box D), and the concentration of dissolved CO_2 (%—box E) acquired by the automatic station in the period of 2010–2013. The box F of Figure 5 shows the rainfall (mm) trend recorded at the Gallicano pluviometric station <https://www.sir.toscana.it/pluviometria-pub> (accessed on 23 February 2024).

Table 2. Descriptive statistic for the parameters monitored continuously at the Gallicano station.

	Mean	Median	Minimum	Maximum	Std. Dev
Temperature ($^\circ\text{C}$)	22.5	22.6	20.4	23.7	0.76
Electrical Conductivity ($\mu\text{S}/\text{cm}$)	5444	5571	4007	6088	461
pH	6.97	6.97	6.89	7.07	0.03
ORP (mV)	387	393	171	461	38.8
CO_2 (%)	2.37	2.19	1.46	5.06	0.60

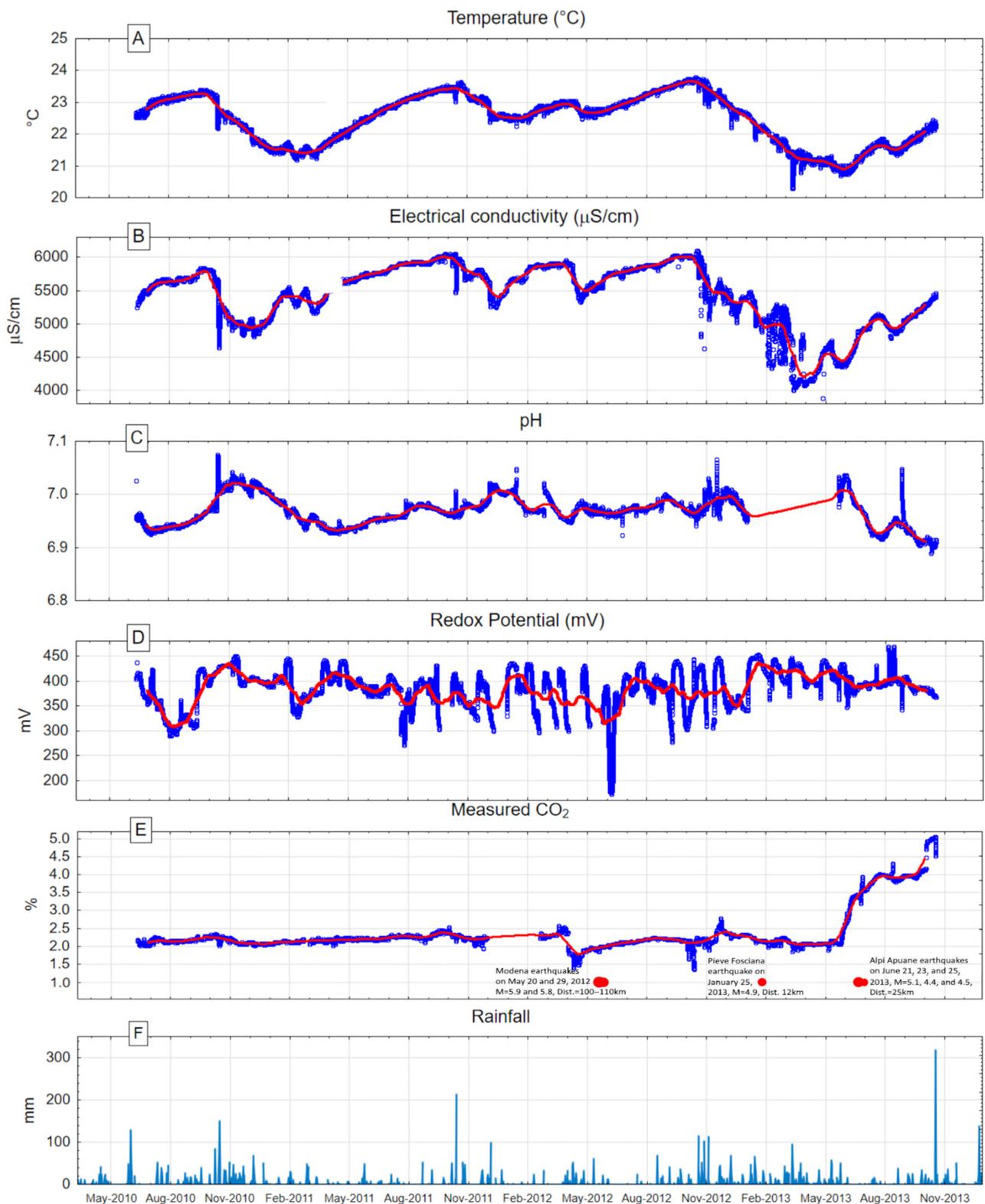


Figure 5. Geochemical parameters measured continuously in the Galliciano station (blue lines) compared with seismic events: (A) temperature (°C); (B) electrical conductivity (µS/cm); (C), pH; (D) redox potential (mV); (E) CO₂ (vol.% of the headspace of the measurement cell). Red lines: temporal trend of the mobile average of each parameter (average on a 30-day window). (F) Rainfall (mm) recorded at the Galliciano pluviometric station data from <https://www.sir.toscana.it/pluviometria-pub> (accessed on 23 February 2024).

The signals of temperature (Figure 5A) and electrical conductivity (Figure 5B) are consistent over time, with the highest values recorded in the months of October, before the autumn rains. The effects of rainfall on these signals appear with seasonal decreases but also with an abrupt fall (up to 1 °C) after intense rainfall events. This agrees with the circulation model (Figure 2B), according to which rainfall directly feeds the cold hyposaline superficial circulation which, in turn, due to the lack of CI, dilutes the deep component before it rises to the surface. This mixing also involves an increase in pH (Figure 5C), which shows an anti-correlated trend to the two previous signals. The redox potential (Figure 5D) has an average value of 387 mV, indicative of predominantly oxidized conditions in the aquifer. The discontinuous trend of this parameter is not real but is due to the long stabilization times needed for the electrode after each monthly cleaning operation.

Regarding the %CO₂ signal, four “macroscopic” CO₂ anomalies can be recognized in Figure 5E. A slight increase in dissolved CO₂ was observed at the beginning of April 2012, followed by a sharp decrease in the middle of the month. Months later, sharp drops in concentration were observed in October 2012 and a significant peak towards the end of November. All three variations almost reached a 1% CO₂ concentration. Finally, from the end of May 2013, a sharp rise in dissolved CO₂ was observed, which increased by approximately 1.5%. Three of the four macroscopic anomalies preceded the most significant seismic events around the Gallicano spring. The variation in April 2012 anticipated the strong events in Modena on the 20 and 29 May 2012 ($M_W = 5.9$ and 5.8); the variation in November 2012 anticipated the event that occurred in Pieve Fosciana ($M_W = 4.9$); and the variation in May 2013 anticipated the events that occurred under the Apuan Alps chain ($M_W = 5.1, 4.4,$ and 4.5). The three abrupt variations anticipated all seismic events with a magnitude $M > 4$, reported in Figure 5E using red circles, taking place within one or two months from the variations.

The significant variations in dissolved CO₂ observed in May 2013, approximately one month before the Apuan Alps’ seismic event ($M_W = 5.1$), can be interpreted using the described hydrogeological model. A possible decrease in permeability during the preparatory phase of an earthquake has been suggested [41]. In fact, the phases of crustal deformation that precede earthquakes can induce permeability variations in the deeper layers of the CAS and contribute to a greater release of gases towards the surface, including CO₂. Therefore, water from the deep circulation system may reach the surface with a higher concentration of CO₂ before a seismic event. Moreover, minor variations in dissolved CO₂ concentrations are generally observed during extended rainfall, as observed in our research for the months of October 2010, 2011, and 2012, in which they were all negative and 0.2–0.5% in amplitude. This depends on the fact that hydrothermal circuits, richer in dissolved CO₂, are diluted by cold-water circuits, with less concentration of dissolved CO₂, fed by rainfall infiltrated through carbonate outcrops and main disjunctive lines (Figure 3B). However, the other negative and positive variations in CO₂ concentrations in Figure 5E lack of a visual association with seismic or rain events. Finally, there are no settlements near the spring nor are there any known anthropogenic influences on the concentration of CO₂. Thus, several sudden minor variations remain unexplained at this level of the analysis.

4. Statistical Methods

A first attempt to assess the statistical correlation between mild-to-moderate seismic events and small-amplitude CO₂ anomalies in Gallicano was recently carried out based on data collected during the period of April 2017–March 2021. The CO₂ anomalies, highlighted by fitting the residual component of the signals with Pearson type-VII functions, were tested in our previous work for their linear dependence to seismic events by binary correlations [45]. In that framework, correlations were used also to estimate the conditional probabilities of seismic events, allowing for a sort of “earthquake forecasting experiment” based on continuous geochemical signals. Continuing with the same approach, here, we focus on the possible binary correlation between geochemical anoma-

lies and mild-to-moderate seismic events that occurred around Gallicano in the period of 2010–2013.

4.1. Processing of CO₂ Time Series to Highlight Anomalies

Similar to [45], we fitted the CO₂ residual signal using a Pearson type-VII function by defining a threshold over which CO₂ fluctuations had a 99% probability of being anomalies. We followed and compared two distinct approaches. Firstly, we decomposed the time series according to the Census I method [61], and then we trained an ANN [62] to reproduce a synthetic signal of the CO₂. The residual signal was obtained by a comparison of real (i.e., CO₂ time series acquired by the station) and synthetic data (i.e., CO₂ time series built by the ANN).

Before applying these two procedures, the raw data were filtered with appropriate moving median-smoothing methods [63,64] to filter out the outliers (Table 2). The CO₂ residual signal obtained from the comparison between the CO₂ measured signal and a CO₂ synthetic signal was analyzed to detect anomalous variations. Two kinds of synthetic signals were generated using CENSUS and ANN. The CO₂ residuals were compared with several distribution functions to obtain the best fit. Then, the anomalies were defined by the cumulative probability function as having a 95–99% probability of not belonging to the distribution.

4.1.1. Decomposition of CO₂ Time Series by Census I Method

We decomposed the CO₂ time series with the Census I method [61], as implemented in the statistical software STATISTICA 12.0 [65]. This procedure allowed us to detrend the CO₂ signal for external influences based on a reference analysis spanning 3 months. More details on the procedure can be found in [41]. Data series of CO₂ irregular components and residuals are plotted in Figure 6 together with selected earthquakes for comparison.

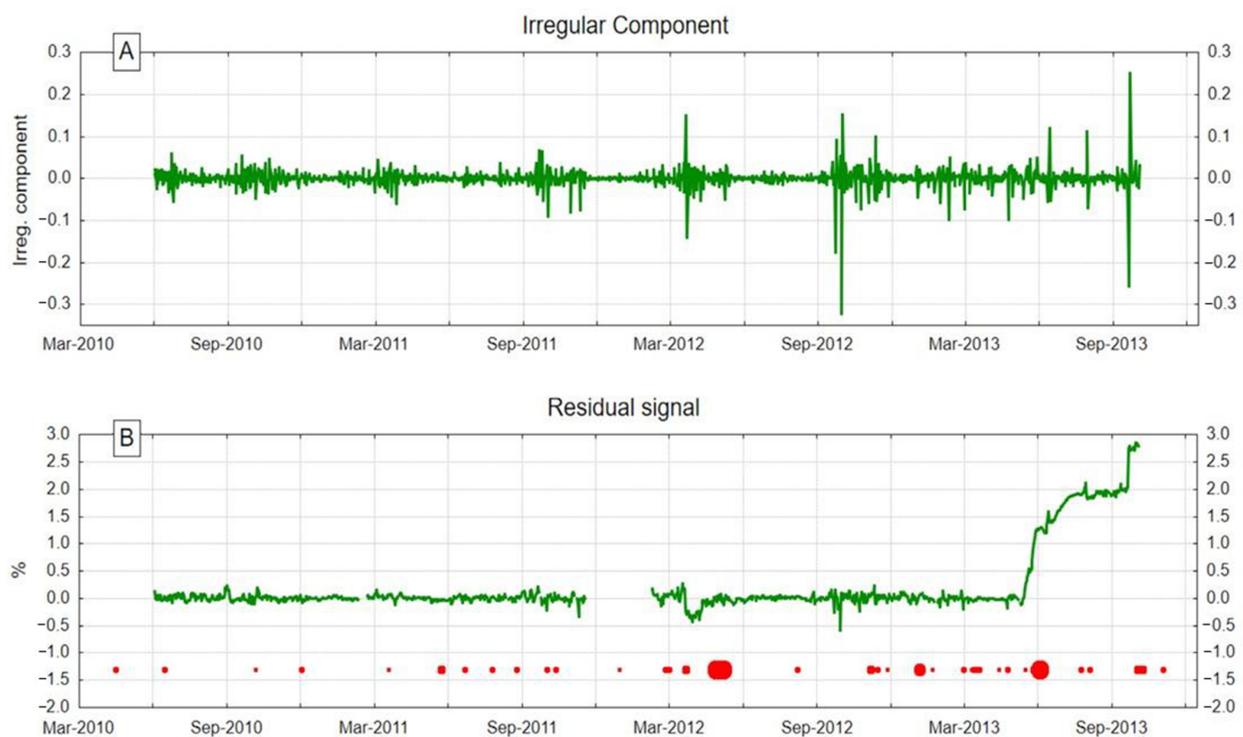


Figure 6. (A): Irregular component of CO₂ time series. (B): Residual signal obtained by the difference between the CO₂ measured signal and the CO₂ simulated signal via ANN. The red dots mark the selected seismic events in Section 3. The dots' dimensions progressively increase with the local magnitude (M_l).

4.1.2. Creation of Synthetic CO₂ Time Series by ANN

The ANN has already been used to detect CO₂ anomalies in another spring of GNT [42]. In the current study, an ANN was used in parallel with Census I to compare their respective performances. The ANN method relies on a progressive learning process that emulates the training process of the human brain [66]. For this study, we used the multi-layer perceptrons' network (MLP) method [67], as implemented in the STATISTICA 12 software, also suitable for non-Gaussian data distributions [68].

CO₂ anomalies were identified on the residual signal after a comparison of the measured signal and the synthetic signal processed by ANN analysis. To reproduce the synthetic CO₂ signal, we trained the ANN with the signals recorded by the station for water temperature, electrical conductivity, and redox potential. We provided air temperature and rainfall data downloaded from <http://www.sir.toscana.it/> (accessed on 23 February 2024) as the input, using the measured CO₂ signal as the optimization target. We used 70% of the total dataset for training, 15% for validation (15%), and the remaining 15% for the test sets. One hundred synthetic models were retrieved by ANN, and the best reconstruction of the CO₂ signal was determined based on the best correlation between the synthetic and the original signals. Figure 6B shows that the best reconstruction by ANN failed to describe the CO₂ recordings in the last 4 months of the time series. Accordingly, we excluded this observation period from further investigations.

4.2. Meteorological Data

Operative since 2000, the automatic meteorological station of Gallicano (179 m asl) is located in the municipality of Gallicano, E-NE of the spring, at a distance of about 800 m. This station collects cumulative rain data every 5 min. In this work, we used meteorological data from the web archive of the Gallicano pluviometric station <https://www.sir.toscana.it/pluviometria-pub> (accessed on 23 February 2024). The data from June 2010 to October 2013 agree with the estimated value of about 1800 mm/yr for the previous 20 years. The Gallicano station was flooded after the exceptional rain event of the 21 October 2013 (318.8 mm of rain recorded in a single day; Figure 5F).

Previous studies ([41]) have demonstrated that local precipitations may dilute the CO₂ concentration of the deep-water component feeding the Gallicano spring. Because of this mixing relationship, we tentatively identified a couple of rain time series in correspondence with daily rain amounts of 10 and 20 mm, under the hypothesis that these thresholds could pinpoint the onset of meteoric-dominated conditions.

5. Results

The identification of CO₂ anomalies was not simple due to the complex variability in the observable phenomena and their dependence on several environmental parameters, such as mixing with deep and shallow water, and, therefore, on meteorological parameters. The main starting approach was to attempt to apply some methods to subtract from the signal the possible seasonal and external dependence. The irregular component obtained from Census I and the residuals obtained from ANN were two attempts to reduce the CO₂ input dataset. Then, the next step was to statistically determine what were the possible anomalies in these residuals' time series. Moreover, the final step was to correlate them with the seismic events selected around the Gallicano spring.

5.1. CO₂ Anomalies

Both time series distributions were modeled by several functions, and the best fits were obtained, in both cases, using the Pearson type-VII function [69], which means that such distributions exhibited peaked and symmetric distributions which were preliminary modeled using Gaussian and Lorentzian functions. However, these widely recognized functions failed to accurately depict the tails of the data. The Pearson type-VII function [69], which is essentially a Lorentzian function raised to a power, proved more adept at realistically handling tails than both Gaussian and Lorentzian functions, thanks to its adjustable

power parameter. The outcome was an improved confidence level in the fitting, also with a straightforward cumulative integral. The parametric distribution function was as follows:

$$P(x) = \frac{A\Gamma[m]}{\Gamma\left[m - \frac{1}{2}\right]\sqrt{\pi(2m - 1)}} \left[1 + \frac{(x - \mu)^2}{\sigma^2(2m - 1)}\right]^{-m} \tag{4}$$

where A is a multiplicative factor, Γ is the function gamma of Euler, μ the average, and σ^2 its variance. The best fit for the distribution by Census I was obtained using the parameters $m = 1.51$, $\mu = 3.5 \times 10^{-4}$, $\sigma^2 = 5.76 \times 10^{-3}$, and $A = 13.699$. The distribution was described in 56 intervals minus four parameters so as to have 52 degrees of freedom. χ^2 was 6.9, which ensured an approximation with a goodness > 99%. The best fit for the distribution by ANN, excluding the last 140 samples (days) in which the time series was not well described, was obtained using the parameters $m = 1.99$, $\mu = 2.0 \times 10^{-4}$, $\sigma^2 = 4.55 \times 10^{-3}$, and $A = 28.611$. The distribution was described in 51 intervals minus four parameters to have 47 degrees of freedom. χ^2 was 9.4, which ensured an approximation with a goodness > 99%.

The anomalies were identified by setting thresholds that correspond to well-defined probabilities of occurrence. The cumulative probability corresponding to a threshold value was calculated to evaluate the probability of each test input overcoming such a value. The cumulative of the probability distribution (4) was summed through Gauss’s hypergeometric function [70]:

$$Pr(x) = \frac{1}{2} + \frac{(x - \mu)}{\sigma[\pi(2m - 1)]^{1/2}} \frac{\Gamma[m]}{\Gamma[m - 1/2]} {}_2F_1 \left[m; \frac{1}{2}; \frac{3}{2}; -\frac{(x - \mu)^2}{\sigma^2(2m - 1)} \right] \tag{5}$$

The probability for the CO₂ residual time series’ fluctuations to be considered anomalies was chosen to equal 99%. Thus, Equation (5) produced a $\Delta x = 0.07409$ for the Census I residuals and a $\Delta x = 0.2278$ for the ANN residuals. The symmetry of the distribution required positive and negative thresholds around the averages μ , which were defined as $x^+ = 0.07444$ and $x^- = -0.07374$ and $x^+ = 0.2280$ and $x^- = -0.2276$ for Census I and ANN cases, which produced 23 and 21 CO₂ anomalies, respectively. The residual time series were transformed into binary series with a time step of one day, where “1” appeared when the datum went beyond these thresholds and “0” when it appeared somewhere else. The probability of an anomalous measurement was defined by its frequency:

$$P = \frac{N(1)}{N_{tot}} \tag{6}$$

with N_{tot} being the number of days, equal to 1208 for the measurement retrieved by Census I and 1068 for that retrieved by ANN. Thus, probabilities of 0.019 and 0.020 occurred for the two sets of anomalies, respectively.

5.2. Correlations

Binary time series of seismic and rainfall events were defined during the same period as the identified anomalies. The respective symbol “1” appeared when earthquakes were selected and when rainfalls went beyond the thresholds of 10 mm and 20 mm, and “0” was used in other cases. Then, Pearson correlation coefficients R were calculated, which assumed values between -1 and 1 , with a negative R meaning an anti-correlation, $R = 0$ no correlation, and a positive R denoting a correlation. A total of six time series was considered in this study, comparing them two-by-two, for a total of twelve correlation coefficient histograms: both Census I and ANN CO₂ anomalies with the first set of earthquakes and, after, with the second ones, as defined in Section 2; following this, both Census I and ANN CO₂ anomalies were compared with rainfall amounts of 10 mm and 20 mm; and, finally, both series of earthquakes were compared with rainfall amounts of 10 mm and 20 mm.

Formally defining CO₂ anomalous events with (EC), earthquake events with (EQ), and rainfall events with (ER), the Pearson correlation coefficients are as follows:

$$R(E_i ; E_j) = \frac{\frac{(\sum E_i \times E_j)}{N_{tot}} - P(E_i)P(E_j)}{\sqrt{P(E_i)[1 - P(E_i)]P(E_j)[1 - P(E_j)]}} \quad (7)$$

where $E = \{EQ, EC, ER\}$, i and j run over the number of considered days, and $\sum E_i \times E_j$ runs over each couple of different events, summing their coincidences. Both $P(E_i)$ and $P(E_j)$ were defined by (6). When, as in this case, a correlation is calculated between binary series, it is called a Matthews correlation [71], and Equation (7) holds [72]. R -histograms were obtained by introducing a time shift Δt , which lasted from 1 to 20 days earlier and later than the anomalous event, thus

$$R(E_i; E_j(\Delta t)) = \frac{\frac{(\sum E_i \times E_j(\Delta t))}{N_{tot}} - P(E_i)P(E_j)}{\sqrt{P(E_i)[1 - P(E_i)]P(E_j)[1 - P(E_j)]}} \quad (8)$$

Δt in (8) indicates that the time of the event E_j is given by $i \times t + \Delta t$, where $i \times t$ is the time step of the event E_i . $P(E_j) = N(E_j = 1)/N_{tot}$ does not depend on Δt . A conditional probability of the event E_i , given the observation of E_j , can be demonstrated to be related to the R coefficient [73]. Thus, when $\Delta t < 0$, the event E_j precedes the event E_i , and the conditional probability becomes a forecasting probability written as

$$P(E_i | E_j(\Delta t)) = P(E_i) + R(E_i; E_j(\Delta t)) \sqrt{P(E_i)[1 - P(E_i)] \left[\frac{1}{P(E_j)} - 1 \right]} \quad (9)$$

Equation (9) retrieves the probability of E_i following the observation of the E_j anomaly by Δt . The ratio $P(E_i | E_j(\Delta t))/P(E_i)$ defines the event E_i 's probability increase that is observed after the E_j anomaly measurement; this is, for definition, the probability gain $G_{ij}(\Delta t)$.

The 12 cross-correlation histograms were plotted with a time step Δt of ± 1 to a Δt of ± 20 days. The correlations between rainfalls and earthquakes are reported in Figure 7. The four combinations of two couples of time series show no significant correlation between the rainfalls and the considered seismic phenomena, as the p -value associated with each peak is greater than 0.05. The plot of Figure 8 reports four new correlation histograms between seismic and CO₂ events, where a significant peak is observed having time differences of 0, -1 , -10 , and -11 days in Figure 8A, and -1 and -11 days in Figure 8B. Being Δt the time difference $T_{CO_2} - T_{EQ}$, the CO₂ anomalies anticipated the earthquakes by one day and by eleven days. The cross-correlation peaks are around 0.08. We evaluated the significance of the 0.08 correlation peak by the p -value, resulting in $p < 0.05$, which tells us that the alternative hypothesis cannot be discarded. Based on Equation (9), such peaks are equivalent to increased probabilities that a seismic event may occur the same day, one day, 10 days, and 11 days after a CO₂ anomaly, i.e., $P(EQ | EC)$. Compared with the frequency of the earthquakes, in our case, $P(EQ) = 0.0323$, with the conditional probability $P(EQ | EC) = 4.14 P(EQ)$ being a probability gain of $G \approx 4$.

The histograms in Figure 9 describe the four correlations of combined time series between rainfalls and CO₂ anomalies. Here, another correlation peak at $\Delta t = T_{CO_2} - T_{Rain} = +1$ days descends when CO₂ anomalies are identified by Census I, indicating CO₂ anomalies following rainfalls by one day, with a high significance (p -value < 0.01). More weakly and even less defined in time are the peaks occurred when CO₂ anomalies are identified by ANN. In these cases, less significant (p -values even < 0.05) correlation peaks are observed for events on the same day, $+1$ day, and $+4$ days from rainfall, indicating that rainfalls tend to anticipate CO₂ anomalies by up to 4 days. This last, less clear result is probably due to

the way in which CO₂ anomalies are recovered by ANN, a method which takes into account rain time series themselves, therefore already partially subtracting rainfall contributions.

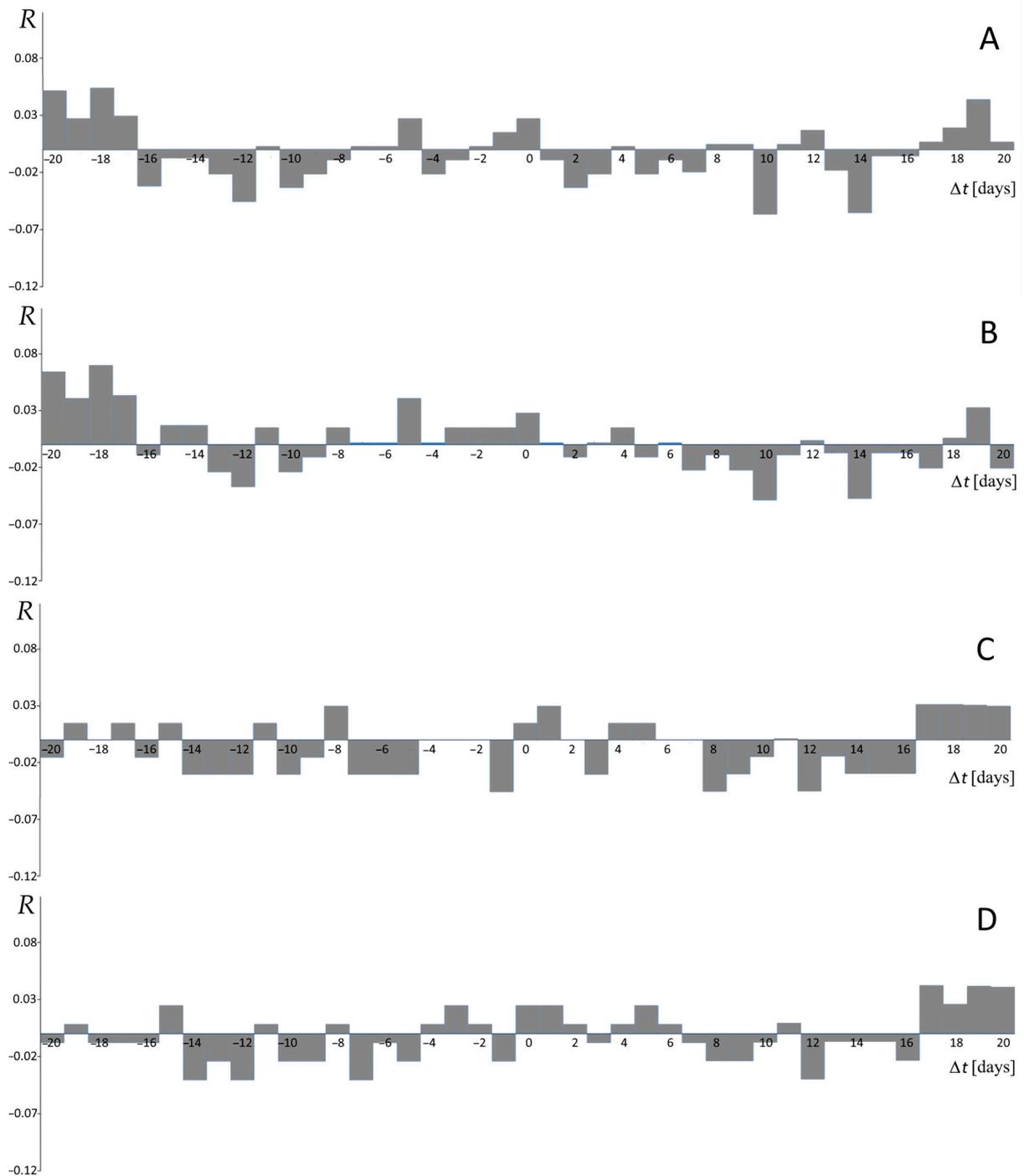


Figure 7. The four correlation histograms with Δt from -20 to 20 days, obtained between (A): first group of earthquakes and rainfalls > 10 mm; (B): main shocks and rainfalls > 10 mm; (C): first group of earthquakes and rainfalls > 20 mm; and (D): main shocks and rainfalls > 20 mm. No significant correlation peaks appear.

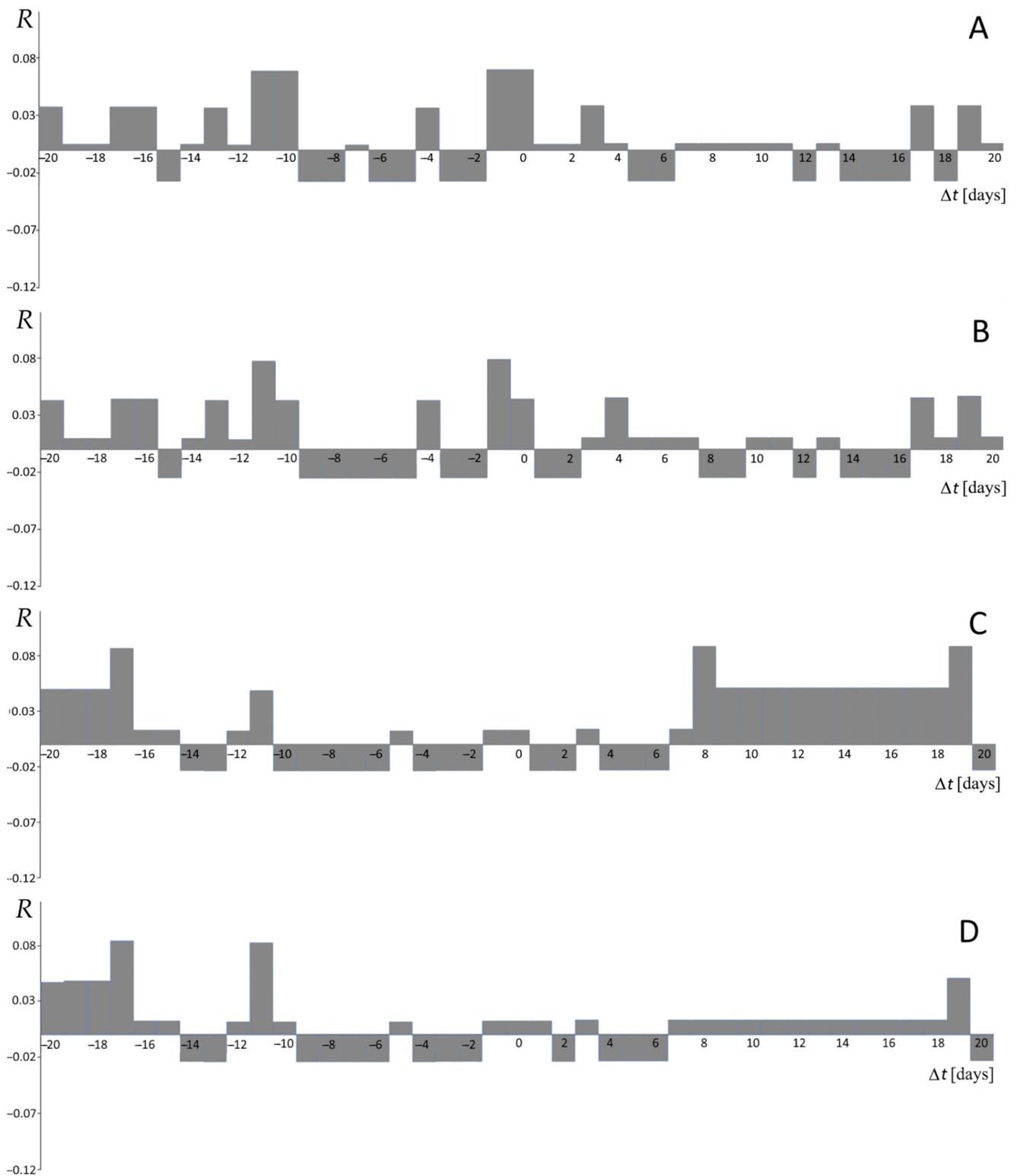


Figure 8. The four correlation histograms with Δt from -20 to 20 days, obtained between (A): CO_2 Census I and first group of earthquakes; (B): CO_2 Census I and main shocks; (C): CO_2 ANN and first group of earthquakes; and (D): CO_2 ANN and main shocks. Significant correlation peaks (p -value < 0.05) appear in (A,B) at $\Delta t = -1$ and -11 days, which means that the CO_2 anomalies anticipate the earthquakes.

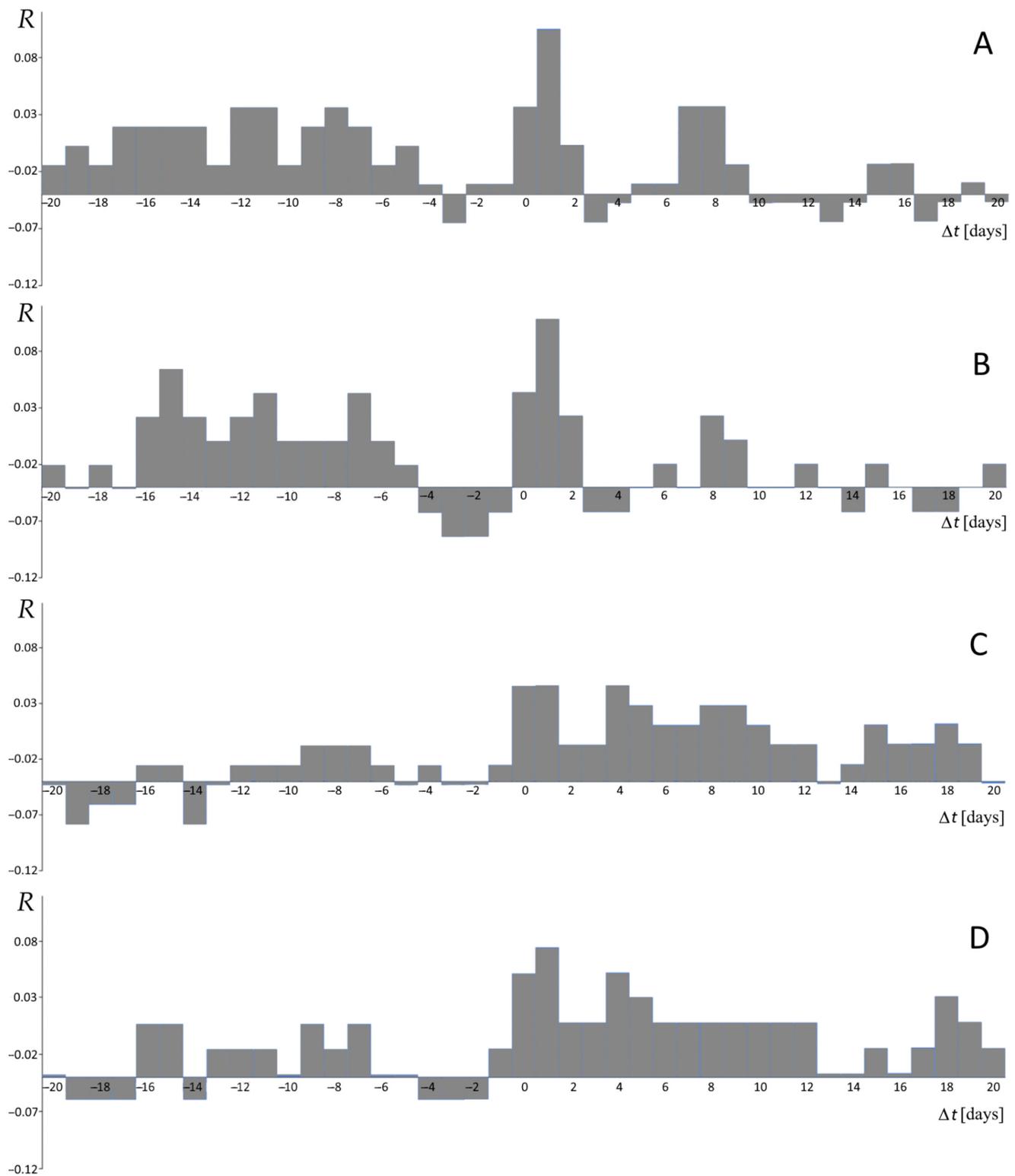


Figure 9. The four correlation histograms with Δt from -20 to 20 days, obtained between (A): CO₂ Census I and rainfalls > 10 mm; (B): CO₂ Census I and rainfalls > 20 mm; (C): CO₂ ANN and rainfalls > 10 mm; and (D): CO₂ ANN and rainfalls > 20 mm. Significant correlation peaks (p -value < 0.05) appear more markedly in (A,B) at $\Delta t = +1$ and less markedly in (C,D), which means that more intense rainfalls are followed by CO₂ anomalies.

6. Discussion and Conclusions

A first statistical processing of water-dissolved CO₂ data acquired continuously at the thermo-mineral spring of Gallicano, Italy, disclosed their positive correlation with earthquakes, where significant dissolved gas variations anticipated small seismic events by two days [45]. The extended procedure to obtain CO₂ vs. rainfalls and earthquakes vs. rainfalls time series showed only a slight correlation between the CO₂ anomalies and precipitation (in particular, rainfall occurred 1 day before the CO₂ anomalies), while it highlighted the complete absence of correlation between seismic and meteoric events. The evidence of a lack of any relationship between earthquakes and rainfall is reinforced by the results shown in Figure 7. It reports no significant correlation peaks at different time lags from −20 to 20 days for all the combinations of seismic main shocks or strongest shocks with 10 mm or 20 mm rainfall events.

Moreover, the results about the significance of the peaks in Figure 8 still underline the trend of observing significant dissolved CO₂ variation events before small earthquakes around the Gallicano site. However, some differences in the result of −2 days observed for the years from 2017 to 2021 exist with the time lags being −1 and −11 days in these cases, considering both main shocks and strongest shocks. Such differences were probably due to the different geological settings of the seismic events, which occurred at different hypocenters with respect to the hypocenters of the previously considered period. Specifically, time lags of −10 and −11 days can be linked to seismic events of magnitudes generally higher than the events' magnitude in the previously considered period. Noteworthy is the fact that an increase in the correlation coefficient was also observed for the time lags of −10 and −11 days in Figure 4 (middle) for the 2017–2021 period, which resulted not significant in our previous analysis [45]. It was also observed, for the extreme magnitudes in this study, that an M5.1 event was anticipated by an anomaly observed 20 days before, and, in our previous study, an M3.5 event was anticipated by an anomaly observed 3 days before. The ratio between the epicenter of an earthquake and the spring's distance and the linear length of the seismic source apparently decreases with increased earthquake magnitudes (Table 1), a phenomenon which agrees with the evidence that CO₂ anomalies related to closer seismic sources and/or larger seismic events can be detected earlier, at the point of monitoring. A time lag of 0 days was also observed only when strongest shocks were considered, meaning that the aftershocks were, in this case, contributing to such a peak. Correlation peaks were not observed when ANN was used to select dissolved CO₂ concentration anomalies, which should mean that this method was unable to select significant dissolved gas variations possibly related to earthquakes.

Finally, the results of Figure 9 strongly confirm the observation of dissolved CO₂ concentration anomalies one day after rainfall events. The weak correlation peak (+1 day) observed in our previous study became very significant when dissolved CO₂ concentration anomalies were selected by Census I for the period of 2010–2013, independently from whether the rainfall consisted of 10 or 20 mm cumulative rain events. Significant results, even if a little less so, were observed when selecting CO₂ anomalies by ANN, this time only for 20 mm cumulative rain events, which further underlined the poor effectiveness of this method. CO₂ variations subsequent to rainfall events agree with the groundwater circulation model developed by [41] for the Gallicano spring. According to this model, rainfall directly feeds a shallow circulation of water that is cold and poor in CO₂, which mixes with hydrothermal circuits, richer in CO₂, developed in a deep evaporitic-carbonatic reservoir. The mixing between the two components occurs before the outflow of water to the surface (Figure 3B).

In conclusion, we proposed a cross-correlation analysis between seismic events and rainfalls, seismic events and CO₂ anomalies, and rainfalls and CO₂ anomalies, highlighted in the thermo-mineral waters of Gallicano spring, a site in which automatic and continuous monitoring has been carried out since 2003. We focused our analysis on the time series collected during the period of June 2010–October 2013.

We processed CO₂ time series with both the Census and ANN methods to reconstruct two synthetic trends. Then, we fit residual distributions using Pearson type-VII functions to evidence anomalous fluctuations in dissolved CO₂ contents. These anomalies were defined based on a threshold according to which the values had a 99% probability of not happening by chance. By cross-correlating the binary series of CO₂ anomalies and earthquakes with a low-to-moderate magnitude, positive correlations were highlighted for CO₂ anomalies that occurred 1 and 11 days before the earthquakes, unlike our previous analysis, of the period of 2017–2021, which provided an anticipation of the anomalies with respect to the earthquakes 2 days earlier. Nevertheless, the probability gain resulted again to be around 4. We also expanded the same statistical approach to double the amount of cumulative precipitation and to an earthquake set also containing aftershocks. We estimated negligible correlations among any seismic and rain event sets compared to the CO₂ vs. earthquakes case, confirming the results from our previous study. Instead, a strong correlation appeared between rainfall and CO₂, occurring mainly one day after intense rainfalls, which reinforces the slight evidence which emerged in our previous work. The time lag of 1 day between CO₂ anomalies and earthquake observations in this work is similar to the time lag of 2 days obtained when studying the 2017–2021 period [45]. Moreover, a new time lag of 10–11 days was observed here with another significant correlation, which fits two different earthquake sets around Gallicano, which could be related to the generally higher magnitude of the earthquakes considered in this research. For what concerns the correlation between CO₂ anomalies and rainfalls, which were weakly observed in our past study [45], here, they were confirmed with a strong significance.

Figure 3B reports how mixing thought between hydrothermal circuits, richer in dissolved CO₂, and cold-water circuits, with less concentration of dissolved CO₂, is conceived. Since the deep-water contribution is richer in CO₂ (we obtain this information via chemical analyses), an increase in CO₂ may arise from deeper levels of the reservoir, as a result of changes in permeability associated with the preparatory phases of earthquakes. Moreover, diffused water from significant rainfalls can quickly intercept the rising spring water in the more superficial layers, so as to produce the observed negative anomalies. Overall, this study suggests that statistical methods are an efficient tool for supporting the “standard” analysis of hydrogeochemical data in the search of correlations between seismicity and variations in observational data. Furthermore, it also emphasizes that the availability of continuous time series is an essential prerequisite to advance the mere observation of “anomalous” data towards the identification of reliable precursors of seismic events.

Author Contributions: Conceptualization, L.P., C.F. and F.G.; methodology, L.P. and C.F.; software, C.F.; validation, L.P., C.F., F.G. and G.F.; resources, L.P. and G.F.; data curation, L.P. and G.F.; writing—original draft preparation, C.F.; writing—review and editing, L.P., C.F. and F.G.; supervision, F.G.; project administration, L.P., G.F. and F.G.; funding acquisition, L.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partially funded by Seismic Prevention Service of the Tuscany Region grant number [Decreto Dirigenziale Regione Toscana n.1666—22/03/2002].

Data Availability Statement: The raw data supporting the conclusion of this article will be made available by the authors after request.

Acknowledgments: The authors are grateful to the Seismic Sector of the Tuscany region, Italy, which has financed this research. We acknowledge many people at INGV for many constructive discussions and their insightful comments on a previous draft, which allowed us to improve the manuscript. We are also indebted to three anonymous reviewers and the Editor for their careful reading and constructive comments.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Netroba, A.V.; Fridman, A.I.; Plotnikov, I.A.; Khurin, M.I. On the large-scale 504 mapping of closed ore-bearing areas in the north Caucasus with the uses of gas surveying as a geochemical method. *Geokhimiya* **1971**, *8*, 1016–1021. (In Russian)

2. Kravtsov, A.I.; Voitov, G.I. Evaluation of the role of faults in the gas exchange between the lithosphere and the atmosphere illustrated in southern Daghestan. *Proc. High. Educ. Inst. Geol. Explor.* **1976**, *19*, 18–26. (In Russian)
3. Gold, T. Terrestrial sources of carbon and earthquake outgassing. *J. Petroleum Geol.* **1979**, *1*, 3–19. [[CrossRef](#)]
4. Gold, T.; Soter, S. Natural Explosive Noises. *Science* **1981**, *212*, 1297–1298. [[CrossRef](#)]
5. Irwin, W.P.; Barnes, I. Tectonic relations of carbon dioxide discharges and earthquake. *J. Geophys. Res.* **1980**, *85*, 3115–3121. [[CrossRef](#)]
6. Barnes, I.; Irwin, P.W.; White, D.E. *Global Distribution of Carbon Dioxide Discharges, and Major Zones of Seismicity*; Barnes, I., Irwin, P.W., White, D.E., Eds.; Water Resources Investigation WRI 78-39; U.S. Geological Survey: Washington, DC, USA, 1978; pp. 1–17. [[CrossRef](#)]
7. Gold, T.; Soter, S. Fluid ascent through the solid lithosphere and its relation to earthquakes. *Pure Appl. Geophys.* **1984**, *122*, 492–530. [[CrossRef](#)]
8. Sugisaki, R.; Ido, M.; Takeda, H.; Isobe, Y.; Hayashi, Y.; Nakamura, N.; Satake, H.; Mizutani, Y. Origin of hydrogen and carbon dioxide in fault gases and its relation to fault activity. *J. Geol.* **1983**, *91*, 239–258. [[CrossRef](#)]
9. King, C.Y. Gas geochemistry applied to earthquake prediction: An overview. *J. Geophys. Res.* **1986**, *91*, 12269–12281. [[CrossRef](#)]
10. Stefánsson, R. Advances in Earthquake Prediction, Research and Risk Mitigation. In *Springer and Praxis Book*; Springer: Berlin/Heidelberg, Germany, 2011; p. 274.
11. Stefánsson, R.; Halldórsson, P. Strain release and strain build-up in the South Iceland seismic zone. *Tectonophysics* **1988**, *159*, 267–276. [[CrossRef](#)]
12. Tamburello, G.; Pondrelli, S.; Chiodini, G.; Rouwet, D. Global-scale control of extensional tectonics on CO₂ Earth degassing. *Nat. Commun.* **2018**, *9*, 4608. [[CrossRef](#)]
13. Kingsley, S.P.; Biagi, P.E.; Piccolo, R.; Capozzi, V.; Ermini, A.; Khatkevichs, Y.M.; Gordeevs, E.I. Hydrogeochemical precursors of strong earthquakes: A realistic possibility in Kamchatka. *Phys. Chem. Earth C* **2001**, *26*, 769–774. [[CrossRef](#)]
14. Bräuer, K.; Kämpf, H.; Strauch, G.; Weise, S.M. Isotopic evidence (3He/4He, 13C-CO₂) of fluid-triggered intraplate seismicity. *J. Geophys. Res.* **2003**, *108*, 2070. [[CrossRef](#)]
15. Ventura, G.; Cinti, F.R.; Di Luccio, F.; Pino, N.A. Mantle wedge dynamics versus crustal seismicity in the Apennines (Italy). *Geochem. Geophys. Geosys.* **2007**, *8*, Q02013. [[CrossRef](#)]
16. Massin, F.; Farrell, J.; Smith, R.B. Repeating Earthquakes in the Yellowstone Volcanic Field: Implications for Rupture Dynamics, Ground Deformation, and Migration in Earthquake Swarms. *J. Volcanol. Geotherm. Res.* **2013**, *257*, 159–173. [[CrossRef](#)]
17. Miller, S.A. *Advances in Geophysics*; Dmowska, R., Ed.; Elsevier: Amsterdam, The Netherlands, 2013; Volume 54, pp. 1–46.
18. Fischer, T.; Matyska, C.; Heinicke, J. Earthquake-enhanced permeability—evidence from carbon dioxide release following the ML 3.5 earthquake in West Bohemia. *Earth Planet. Sci. Lett.* **2017**, *460*, 60–67. [[CrossRef](#)]
19. Yoshida, K.; Hasegawa, A. Hypocenter Migration and Seismicity Pattern Change in the Yamagata-Fukushima Border, NE Japan, Caused by Fluid Movement and Pore Pressure Variation. *J. Geophys. Res. Solid Earth* **2018**, *123*, 5000–5017. [[CrossRef](#)]
20. Chiodini, G.; Cardellini, C.; Di Luccio, F.; Selva, J.; Frondini, F.; Caliro, S.; Rosiello, A.; Beddini, G.; Ventura, G. Correlation between tectonic CO₂ Earth degassing and seismicity is revealed by a 10-year record in the Apennines, Italy. *Sci. Adv.* **2020**, *6*, eabc2938. [[CrossRef](#)] [[PubMed](#)]
21. Heinicke, J.; Koch, U.; Martinelli, G. CO₂ and Radon measurements in the Vogtland area (Germany)—A contribute to earthquake prediction research. *Geophys. Res. Lett.* **1995**, *22*, 771–774. [[CrossRef](#)]
22. Heinicke, J.; Braun, T.; Burgassi, P.; Italiano, F.; Martinelli, G. Gas flow anomalies in seismogenic zones in the Upper Tiber Valley, Central Italy. *Geophys. J. Int.* **2006**, *167*, 794–806. [[CrossRef](#)]
23. Hunt, J.A.; Zafu, A.; Mather, T.A.; Pyle, D.M.; Barry, P.H. Spatially variable CO₂ degassing in the main Ethiopian rift: Implications for magma storage, volatile transport, and rift-related emissions. *Geochem. Geophys. Geosyst.* **2017**, *18*, 3714–3737. [[CrossRef](#)]
24. Rose, T.P.; Davison, M.L. Radiocarbon in hydrologic systems containing dissolved magmatic carbon dioxide. *Science* **1996**, *273*, 1367–1370. [[CrossRef](#)]
25. Sorey, M.L.; Evans, W.C.; Kennedy, B.M.; Farrar, C.D.; Hainsworth, L.J.; Hausback, B. Carbon dioxide and helium emissions from a reservoir of magmatic gas beneath Mammoth Mountain, California. *J. Geophys. Res. Solid Earth* **1998**, *103*, 15303–15323. [[CrossRef](#)]
26. Chiodini, G.; Frondini, F.; Kerrick, D.M.; Rogie, J.; Parello, F.; Peruzzi, L.; Zanzari, A.R. Quantification of deep CO₂ fluxes from Central Italy. Examples of carbon balance for regional aquifers and of soil diffuse degassing. *Chem. Geol.* **1999**, *159*, 205–222. [[CrossRef](#)]
27. Caliro, S.; Chiodini, G.; Avino, R.; Cardellini, C.; Frondini, F. Volcanic degassing at Somma–Vesuvio (Italy) inferred by chemical and isotopic signatures of groundwater. *Appl. Geochem.* **2005**, *20*, 1060–1076. [[CrossRef](#)]
28. Cox, S.F. Faulting processes at high fluid pressures: An example of fault valve behavior from the Wattle Gully Fault, Victoria, Australia. *J. Geophys. Res. Solid Earth* **1995**, *100*, 12841–12859. [[CrossRef](#)]
29. Collettini, C.; Cardellini, C.; Chiodini, G.; De Paola, N.; Holdsworth, R.E.; Smith, S.A.F. Fault weakening due to CO₂ degassing in the Northern Apennines: Short- and long-term processes. *Geol. Soc. Lond. Spec. Publ.* **2008**, *299*, 175–194. [[CrossRef](#)]
30. Chiarabba, C.; De Gori, P.; Valoroso, L.; Petitta, M.; Carminati, E. Large extensional earthquakes push-up terrific amount of fluids. *Sci. Rep.* **2022**, *12*, 14597. [[CrossRef](#)]

31. Gabrielli, S.; Gabrielli, S.; Akinci, A.; Ventura, G.; Napolitano, F.; Del Pezzo, E.; De Siena, L. Fast Changes in Seismic Attenuation of the Upper Crust due to Fracturing and Fluid Migration: The 2016–2017 Central Italy Seismic Sequence. *Front. Earth Sci.* **2022**, *10*, 909698. [[CrossRef](#)]
32. Fonzetti, R.; Valoroso, L.; De Gori, P.; Chiarabba, C. Fault and Fluid Interaction during the 2012 Emilia (Northern Italy) Seismic Sequence. *Seismol. Res. Lett.* **2023**, *94*, 671–684. [[CrossRef](#)]
33. Manga, M.; Wang, C.Y. Earthquake Hydrology. In *Treatise on Geophysics*, 2nd ed.; Elsevier: Oxford, UK, 2015; Volume 4, pp. 305–328.
34. Martinelli, G.; Facca, G.; Genzano, N.; Gherardi, F.; Lisi, M.; Pierotti, L.; Tramutoli, V. Earthquake-related signals in Central Italy detected by hydrogeochemical and satellite techniques. *Front. Earth Sci.* **2020**, *8*, 584716. [[CrossRef](#)]
35. Gischig, V.S.; Giardini, D.; Amann, F.; Hertrich, M.; Krietsch, H.; Loew, S.; Maurer, H.; Villiger, L.; Wiemer, S.; Bethmann, F.; et al. Hydraulic stimulation and fluid circulation experiments in underground laboratories: Stepping up the scale towards engineered geothermal systems. *Geomech. Energy Environ.* **2020**, *24*, 100175. [[CrossRef](#)]
36. Guglielmi, Y.; Cappa, F.; Henry, P.; Elsworth, E. Seismicity triggered by fluid injection-induced aseismic slip. *Science* **2015**, *348*, 1224–1226. [[CrossRef](#)]
37. Wang, W.; Shan, J.; Ni, Z.; Cai, J. Relationship between earthquake dilatancy and electric precursor phenomena. *Nat. Hazards* **2015**, *79*, 249–262. [[CrossRef](#)]
38. Enomoto, Y. Coupled interaction of earthquake nucleation with deep Earth gases: A possible mechanism for seismo-electromagnetic phenomena. *Geophys. J. Int.* **2012**, *191*, 1210–1214. [[CrossRef](#)]
39. Enomoto, Y.; Yamabe, T.; Okumura, N. Causal mechanisms of seismo-EM phenomena during the 1965–1967 Matsushiro earthquake swarm. *Sci. Rep.* **2017**, *7*, 44774. [[CrossRef](#)] [[PubMed](#)]
40. Martinelli, G.; Tamburello, G. Geological and geophysical factors constraining the occurrence of earthquake precursors in geofluids: A review and reinterpretation. *Front. Earth Sci.* **2020**, *8*, 596050. [[CrossRef](#)]
41. Pierotti, L.; Botti, F.; D’Intinosante, V.; Facca, G.; Gherardi, F. Anomalous CO₂ content in the Gallicano thermo-mineral spring (Serchio Valley, Italy) before the 21 June 2013, Alpi Apuane earthquake (M = 5.2). *Phys. Chem. Earth Parts A/B/C* **2015**, *85*, 131–140. [[CrossRef](#)]
42. Pierotti, L.; Gherardi, F.; Facca, G.; Piccardi, L.; Moratti, G. Detecting CO₂ anomalies in a spring on Mt. Amiata volcano (Italy). *Phys. Chem. Earth Parts A/B/C* **2017**, *98*, 161–172. [[CrossRef](#)]
43. Gherardi, F.; Pierotti, L. The suitability of the Pieve Fosciana hydrothermal system (Italy) as a detection site for geochemical seismic precursors. *Appl. Geochem.* **2018**, *92*, 166–179. [[CrossRef](#)]
44. Fidani, C. West Pacific Earthquake Forecasting Using NOAA Electron Bursts with Independent L-Shells and Ground-Based Magnetic Correlations. *Front. Earth Sci.* **2021**, *9*, 673105. [[CrossRef](#)]
45. Fidani, C.; Gherardi, F.; Facca, G.; Pierotti, L. Correlation between small earthquakes and anomalies in spring waters: A statistical experiment on the probability of seismic occurrence. *Front. Earth Sci.* **2023**, *11*, 1128949. [[CrossRef](#)]
46. Rovida, A.; Locati, M.; Camassi, R.; Lolli, B.; Gasperini, P.; Antonucci, A. *Catalogo Parametrico dei Terremoti Italiani (CPTI15), Versione 4.0*; Istituto Nazionale di Geofisica e Vulcanologia: Italy, 2022.
47. DISS Working Group. *Database of Individual Seismogenic Sources (DISS), Version 3.3.0: A Compilation of Potential Sources for Earthquakes Larger than M 5.5 in Italy and Surrounding Areas*; Istituto Nazionale di Geofisica e Vulcanologia: Italy, 2021. [[CrossRef](#)]
48. Stucchi, M.; Meletti, C.; Montaldo, V.; Crowley, H.; Calvi, G.M.; Boschi, E. Seismic Hazard Assessment (2003–2009) for the 686 Italian Building Code. *Bull. Seismol. Soc. Am.* **2011**, *101*, 1885–1911. [[CrossRef](#)]
49. Molli, G.; Manighetti, I.; Bennett, R.; Malavieille, J.; Serpelloni, E.; Storti, F.; Giampietro, T.; Bigot, A.; Pinelli, G.; Giacomelli, S.; et al. Active fault systems in the Inner Northwest Apennines, Italy: A reappraisal one century after the 1920 Mw~6.5 Fivizzano earthquake. *Geosciences* **2021**, *11*, 139. [[CrossRef](#)]
50. Reasenber, P. Second-Order Moment of Central California Seismicity, 1969–1982. *J. Geophys. Res.* **1985**, *90*, 5479–5495. [[CrossRef](#)]
51. Wiemer, S. A Software Package to Analyze Seismicity: ZMAP. *Seismol. Res. Lett.* **2001**, *72*, 373–382. [[CrossRef](#)]
52. Dobrovolsky, I.P.; Zubkov, S.I.; Miachkin, V.I. Estimation of the size of earthquake preparation zones. *Pure Appl. Geophys.* **1979**, *117*, 1025–1044. [[CrossRef](#)]
53. Leonard, M. Earthquake Fault Scaling: Self-Consistent Relating of Rupture Length, Width, Average Displacement, and Moment Release. *Bull. Seismol. Soc. Am.* **2010**, *100*, 1971–1988. [[CrossRef](#)]
54. Trippetta, F.; Petricca, P.; Billi, A.; Collettini, C.; Cuffaro, M.; Lombardi, A.M.; Scrocca, D.; Ventura, G.; Morgante, A.; Doglioni, C. From mapped faults to fault-length earthquake magnitude (FLEM): A test on Italy with methodological implications. *Solid Earth* **2019**, *10*, 1555–1579. [[CrossRef](#)]
55. Munafò, I.; Malagnini, L.; Chiaraluce, L. On the Relationship between Mw and ML for Small Earthquakes. *Bull. Seismol. Soc. Am.* **2016**, *106*, 2402–2408. [[CrossRef](#)]
56. De Stefani, C. Le acque termali di Torrite in Garfagnana. *Boll. Soc. Geol. It.* **1904**, *XXIII*, 117–148. (In Italian)
57. Bencini, A.; Duchi, V.; Martini, M. Geochemistry of thermal springs of Tuscany, (Italy). *Chem. Geol.* **1977**, *19*, 229–252. [[CrossRef](#)]
58. Fanelli, M.; Bellucci, L.; Nachira, F. Manifestazioni idrotermali aliene. CNR–PFE. In *Geothermal Energy Subproject RF 13*; Appendix; 1982. CNR Edition. (In Italian)
59. Cioni, R.; Guidi, M.; Pierotti, L.; Scozzari, A. An automatic monitoring network installed in Tuscany (Italy) for studying possible geochemical precursory phenomena. *Nat. Hazards Earth Syst. Sci.* **2007**, *7*, 405–416. [[CrossRef](#)]

60. Baldacci, F.; Raggi, G. Lineamenti geologici e idrogeologici della regione dell'Appennino Ligure Toscano a nord dell'Arno. Contributo alle conoscenze delle risorse geotermiche del territorio italiano, CNR–PFE. In *Geothermal Energy Subproject RF 13*; 1982; pp. 68–85. CNR Edition. (In Italian)
61. Makridakis, S.; Wheelwright, S.C.; Hyndman, R.J. *Forecasting Methods and Applications*, 3rd ed.; John Wiley and Sons: New York, NY, USA, 1998; p. 642.
62. Muller, B.; Reinhardt, J. *Neural Networks: An Introduction*; Springer: Berlin/Heidelberg, Germany, 1990; p. 280.
63. Box, G.E.P.; Jenkins, G.M. *Time Series Analysis: Forecasting and Control*; Holden-Day: San Francisco, CA, USA, 1976.
64. Velleman, P.F.; Hoaglin, D.C. *Applications, Basics, and Computing of Exploratory Data Analysis*; Duxbury Press: Boston, MA, USA, 1981.
65. Statsoft Inc. *STATISTICA (Data Analysis Software System) Version 12 for Windows: Statistics*; STATSOFT Inc.: Tulsa, OK, USA, 2013.
66. Muller, B.; Reinhardt, J.; Strickland, M.T. *Physics of Neural Networks, Neural Networks: An Introduction*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 1995.
67. Rumelhart, D.E.; McClelland, J.L.; PDP Research Group. *Parallel Distributed Processing—Explorations in the Microstructure of Cognition, Volume 1: Foundations*; MIT Press: Cambridge, MA, USA, 1986; p. 547.
68. Solomatine, D.P.; Ostfeld, A. Data-driven modelling: Some past experiences and new approaches. *J. Hydroinformatics* **2008**, *10*, 3–22. [[CrossRef](#)]
69. Pearson, K. Mathematical contributions to the theory of evolution, XIX: Second supplement to a memoir on skew variation. *Philos. Trans. R. Soc. A* **1916**, *216*, 429–457. [[CrossRef](#)]
70. Johnson, N.L.; Samuel Kotz, S.; Balakrishnan, N. *Continuous Univariate Distributions*, 2nd ed.; Wiley: New York, NY, USA, 1995; Volume 2.
71. Matthews, B.W. Comparison of the predicted and observed secondary structure of T4 phage lysozyme. *Biochim. Biophys. Acta (BBA)-Protein Struct.* **1975**, *405*, 442–451. [[CrossRef](#)]
72. Fidani, C. Probability, Causality and False Alarms Using Correlations between Strong Earthquakes and NOAA High Energy Electron Bursts. *Ann. Geophys.* **2020**, *63*, 543. [[CrossRef](#)]
73. Fidani, C. Improving Earthquake Forecasting by Correlations between Strong Earthquakes and NOAA Electron Bursts. *Terr. Atmos. Ocean. Sci.* **2018**, *29*, 117–130. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.