Gas Fields and Large Shallow Seismogenic Reverse Faults Are Anticorrelated

Gianluca Valensise (gianluca.valensise@ingv.it)
Istituto Nazionale di Geofisica e Vulcanologia - INGV - Rome, Italy
Federica Donda
Istituto Nazionale di Oceanografia e Geofisica Sperimentale - OGS - Trieste, Italy
Alberto Tamaro
Istituto Nazionale di Oceanografia e Geofisica Sperimentale - OGS - Trieste, Italy
Giorgia Rosset
Istituto Nazionale di Oceanografia e Geofisica Sperimentale - OGS - Trieste, Italy
Stefano Parolai
Istituto Nazionale di Oceanografia e Geofisica Sperimentale - OGS - Trieste, Italy

Research Article

Keywords: Gas fields, seismogenic faults, anticorrelated, relationships, hydrocarbon, northern-central Italy, underground energy, carbon energy systems

Posted Date: September 8th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-860465/v1

Gas fields and large shallow seismogenic reverse faults are anticorrelated

Valensise G.[1*], Donda F.[2], Tamaro A.[2], Rosset G.[2], Parolai S.[2]

[1] Istituto Nazionale di Geofisica e Vulcanologia – INGV – Rome, Italy
[2] Istituto Nazionale di Oceanografia e di Geofisica Sperimentale – OGS – Trieste, Italy

Abstract (152 words)

We investigated the spatial relationships among 18 known seismogenic faults and 1,612 wells drilled for gas exploitation in the main hydrocarbon province of northern-central Italy, a unique dataset worldwide. We adopted a GIS approach and a robust statistical technique, and found a significant anticorrelation between the location of productive wells and of the considered seismogenic faults, which are generally overlain or encircled by sterile wells.

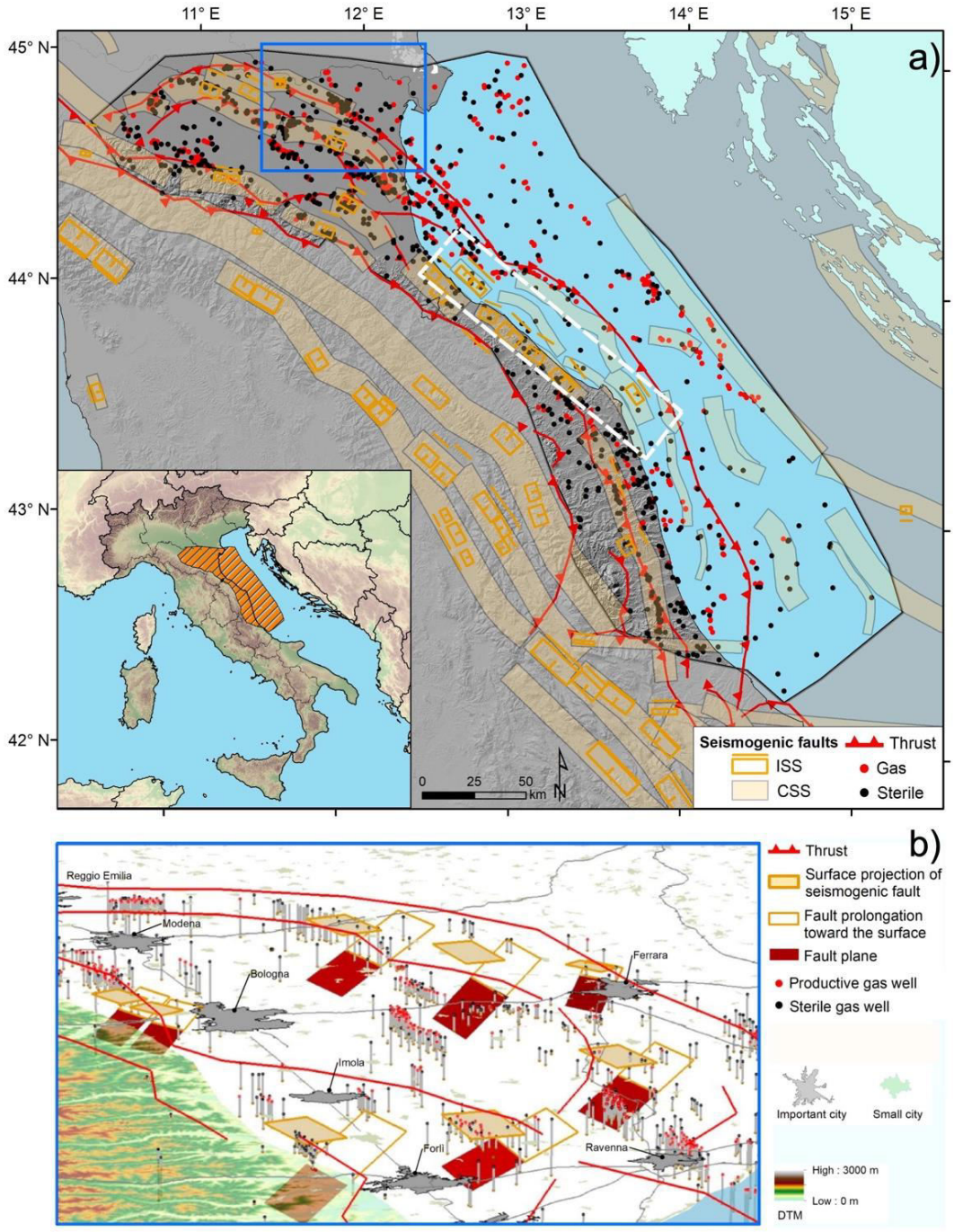
Our observations suggest that (a) earthquake ruptures encompassing much of the upper crust may cause gas to be lost to the atmosphere over geological time, and that (b) reservoirs underlain by smaller or aseismic faults are more likely to be intact.

These findings, which are of inherently global relevance, have crucial implications for future hydrocarbon exploitation, for assessing the seismic-aseismic behavior of large reverse faults, and for the public acceptance of underground energy storage facilities in tectonically active areas, a pillar of future low carbon energy systems.

Introduction (Total of Introduction, Results and Discussion: 4270 words)

Thrust faulting earthquakes are an inherent occurrence in any hydrocarbon-bearing active region: they may be events generated by regional-scale tectonic processes, but also events triggered at all stages of hydrocarbon exploitation activities[1-4]. Nevertheless, earthquakes are often irregularly scattered across relatively homogeneous tectonic trends in most hydrocarbon-bearing regions worldwide. For instance, a thorough investigation of the seismicity of the Zagros region of southern Iran[5], one of the largest oil and gas reserves worldwide, revealed not only that thrust faulting earthquakes are a minority, despite the characteristic fold and thrust structure of the region, but also that the observed seismicity accounts for a small fraction of the range shortening estimated from GPS, InSAR and other lines of evidence. In the frictional regime, the fraction of fault slip that is released in earthquakes is generally referred to as *seismic coupling*, or *c*, a dimensionless parameter originally introduced based on observations of great earthquakes in the circum-Pacific belt[6] . Based on an investigation of the compressional domains encircling the Italian peninsula, most of which are also important oil and gas reservoirs, *c* is about 50%, half than that estimated for the extensional domains straddling the Apennines chain[7]. Rather than being evidence for an episodically aseismic behavior of the large thrust faults occurring beneath Italy's hydrocarbon reservoirs, this suggests that about one out of two of such faults slips consistently in an aseismic fashion. Moreover, stick-slip behavior was suggested to occur only where previous tectonic histories caused the exhumation of competent rocks (e.g. limestones and dolostones) up to the characteristic 3-10 km depth of local upper crustal thrust faults[7].

Figure 1: Overview of the study area, showing the location of selected gas wells and seismogenic faults.



a) The study area, shown by a black polygon, area was precisely delineated following the structural scheme proposed by Mantovani et al.[8]. Gas reservoirs mostly lie within the external foothill belt between the leading thrusts of the internal zone and the outermost thrusts in the foreland basin.

The blue box shows the location of panel b). The white dashed rectangle indicates a 120 km-long coastal strip where seismic hazard is consistently higher than in structurally similar, adjacent areas of the Italian peninsula, but where there exist only fewer-than-average sterile wells and virtually no productive wells (see text).

b) Details of the geometrical calculations conducted over the northwesternmost portion of our study area. The solid orange boxes represent the surface projection of the faults, whose full 3D extent is shown by the red patches, which extend from the minimum to the maximum depth of seismogenic faulting reported in the DISS database (see Methods section). The empty orange boxes show the fault prolongation toward the surface. The vertical lines show the location and depth of the investigated gas wells, allowing the reader to appreciate their actual 3D spatial relationships with the seismogenic faults.

This study explores the spatial distribution of productive and sterile gas fields in relation with the occurrence of large active and potentially seismogenic faults, and hence of moderate to strong earthquakes (M_w 5.5+). To this end, we use well locations as a proxy for the location of econimically attractive gas fields and assess their spatial correlation with fault locations. The study area spreads over a width of about 100 km from the inner portions of the Italian Apennines fold and thrust belt to the associated foredeep, and over a distance of about 300 km from the western Po Plain to the central Adriatic Sea (Figure 1).

The Apennines are a northeastern-verging portion of the peri-Mediterranean orogenic belt, formed as a consequence of the convergence and collision between the European and African plates since the Late Cretaceous[9]. They developed in Neogene and Quaternary times and currently comprise both an extensional domain straddling the range axis, and a compressional domain running along its Adriatic margin[10-13]. More specifically, the northern and central Apennines hydrocarbon province is a result of the progressive thrusting and folding, hosting hydrocarbon reservoirs at the core of growing anticlines/antiforms. The whole Italian peninsula is one of the richest hydrocarbon-producing regions in southern Europe[14,15]. Major gas fields lie parallel to the structural trends, and structural traps usually occur all throughout the thrust belt, most commonly along its external boundary, in the adjacent foredeep basin, and in the Adriatic foreland (Figure 1). They concentrate in the Po Plain and in the northern and central Adriatic Basin, one of the largest hydrocarbon-bearing regions of the globe and also the target of our investigation. Oil fields occur mosty in the western Po Plain, in the southern Apennines and in Sicily[16, 17].

Most of the Italian gas originates in Northern Adriatic reservoirs and is associated with two main source rock types. The older one is thermogenic gas-prone, the main gas pools being found in turbiditic sands occurring in the Apennines foothills[17], whereas the younger one produces biogenic gas hosted in the outer Plio-Pleistocene foredeep domain and feeds the most important fields[15]. Biogenic gas is trapped in the highly efficient Plio-Pleistocene turbidite systems of the Apennines foredeep[18] and within synsedimentary traps on both the inner and outer flanks of the folds[19].

Table 1 Main parameters of all Individual Seismogenic Sources (ISSs) that occur in our study area.

DISS SourceID	DISS Source name	Min depth (km)	Max depth (km)	Dip (deg)	Associated earthquake	Productive gas wells	Sterile gas wells
ITIS024	Mondolfo	4.0	7.0	30	1 Feb 1924	-	4
ITIS029	Conero offshore	2.5	6.4	40	23 Dec 1690	-	1
ITIS030	Senigallia	4.0	7.5	30	30 Oct 1930	-	5
ITIS031	Fano Ardizio	3.0	7.0	30	unknown	-	-
ITIS032	Pesaro S. Bartolo	2.5	5.9	35	unknown	-	1
ITIS033	Rimini offshore south	3.0	5.5	30	16 Aug 1916	-	2
ITIS034	Rimini offshore	3.0	5.5	30	17 May 1916	-	1
ITIS035	Rimini	3	6.0	30	25 Dec 1786	-	1
ITIS036	Val Marecchia	3.0	6.0	30	17 Mar 1875	-	-
ITIS070	Offida	4.5	8.7	35	10 Mar 1943	1	3
ITIS090	Ferrara	1.4	4.5	50	17 Nov 1570	-	2
ITIS091	Casalecchio di Reno	2.0	4.2	35	03 Jan 1505	-	-

ITIS093	Faenza	4.5	7.8	35	04 Apr 1781	-	11
ITIS100	Bagnacavallo	2.5	5.0	25	11 Apr 1688	16	9
ITIS103	Crespellano	2.0	4.5	35	20 Apr 1929	-	-
ITIS107	Mirandola	4.0	7.0	30	29 May 2012	-	3
ITIS134	Finale Emilia	4.0	8.4	43	20 May 2012	-	-
ITIS141	Argenta	3.0	6.3	35	19 Mar 1624	1	6
					Wells falling within ISSs	**18**	**49**
					Total within study area	**793**	**819**
					Grand total	**1612**	

Summary of the seismogenic faults occurring in the study area, showing also the number of wells falling within the surface projection of each source (see the Methods section).

The occurrence of the 20 and 29 May 2012 thrust faulting earthquakes in the southern Po Plain (M_w 6.1 and 6.0, respectively) prompted a debate on the potential triggering role of hydrocarbon exploitation[20]. Subsequent investigations of the relationships between known seismogenic sources and the distribution of gas fields in that region[21] revealed an anticorrelation between productive reservoirs and the presence of relatively large seismogenic faults, i.e. faults capable of M_w 5.5+ earthquakes.

We aim at substantiating these hypotheses and results by investigating a four times larger – and unique, to our knowledge – sample of gas wells, and a twofold number of large seismogenic faults, using a more detailed metric and a more robust statistical approach.

Results

We explored in a GIS-environment the spatial relationships among 1,769 wells drilled for gas exploitation and 18 known seismogenic faults (Figure 1a and Table 1: see also the Methods section) occurring over an area of over 60,000 km^2.We first calculated both minimum *planar* (2D) distances, from the well head to the closest point of a 2 km buffer drawn around the surface projection of each seismogenic fault, and minimum *three-dimensional* (3D) distances, from the well bottom to the closest point on the actual fault plane. Then we used the Weight of Evidence method (WofE[22]) to detect any statistically significant spatial correlation/anticorrelation between faults and productive/sterile gas fields (Figure 1b: see the Methods section for further details).

2D distances. Table 2a and Figs 2a, 2b summarize the results obtained using the *training dataset* for the eight adopted distance bins in the 2D distance case. A positive C implies a good spatial correlation (i.e., the higher the value, the larger the correlation) between the well location and the fault plane; a C around 0 indicates a non-significant correlation; a negative C indicates absence of correlation. The Studentised contrast C/S(C), where S(C) is the standard deviation of C, provides insight into the statistical significance of the results.

For productive wells, C values are negative for distances <5 km (down to -2.025); they then increase, reaching a maximum in the 20-30 km bin, and become negative again for distances >50 km. These results indicate that productive wells generally occur rather far from active faults but not beyond 50 km, a distance that may exceed the half-width of the hydrocarbon province investigated in this work.

Table 2 Summary of the results obtained through the Weight of Evidence test.

(a)

2D distance - Productive wells						
Distance bin (km)	**Contrast weight (C)**	**C/S(C)**	Total wells (*training dataset*)	Total wells (*training dataset* %)	Total wells (*testing dataset*)	Total wells (*testing dataset* %)
<2	-0.325	-1.197	14	2.9	7	2.2
2-5	-2.025	-3.496	3	0.6	5	1.6
5-10	0.184	1.209	48	10.1	18	5.7
10-20	0.467	4.484	125	26.3	82	25.8
20-30	0.758	7.155	119	25.1	94	29.6
30-40	0.499	4.109	82	17.3	51	16.0
40-50	0.245	1.711	55	11.6	35	11.0
>50	-1.907	-9.951	29	6.1	26	8.2
2D distance - Sterile wells						
<2	1.235	9.093	62	12.6	30	9.1
2-5	0.336	1.791	31	6.3	22	6.7
5-10	0.236	1.612	52	10.6	39	11.9
10-20	0.676	6.895	150	30.6	96	29.3
20-30	0.373	3.210	91	18.5	59	18.0
30-40	-0.228	-1.456	45	9.2	34	10.4
40-50	-0.735	-3.440	23	4.7	16	4.9
>50	-1.681	-9.833	37	7.5	32	9.8

(b)

3D distance - Productive wells						
Distance bin (km)	**Contrast weight (C)**	**C/S(C)**	Total wells (*training dataset*)	Total wells (*training dataset* %)	Total wells (*testing dataset*)	Total wells (*testing dataset* %)
<4	0.108	0.339	10	2.1	8	2.5
4-6	-1.113	-2.923	7	1.5	4	1.3
6-10	-0.523	-2.544	25	5.3	8	2.5
10-20	0.541	5.402	142	29.9	84	26.4
20-30	0.699	6.620	120	25.3	101	31.8
30-40	0.520	4.361	86	18.1	51	16.0
40-50	0.266	1.883	57	12.0	35	11.0
>50	-1.880	-9.831	28	5.9	27	8.5
3D distance - Sterile wells						
<4	1.179	6.161	29	5.9	10	3.0
4-6	0.867	5.707	48	9.8	25	7.6
6-10	0.074	0.474	45	9.2	45	13.7
10-20	0.722	7.567	166	33.8	101	30.8
20-30	0.356	3.117	95	19.3	63	19.2
30-40	-0.241	-1.553	46	9.4	35	10.7
40-50	-0.667	-3.249	25	5.1	17	5.2
>50	-1.652	-9.663	37	7.5	32	9.8

Summary of the results we obtained for the 2D (a) and 3D distance scenario (b), respectively for productive and sterile wells.

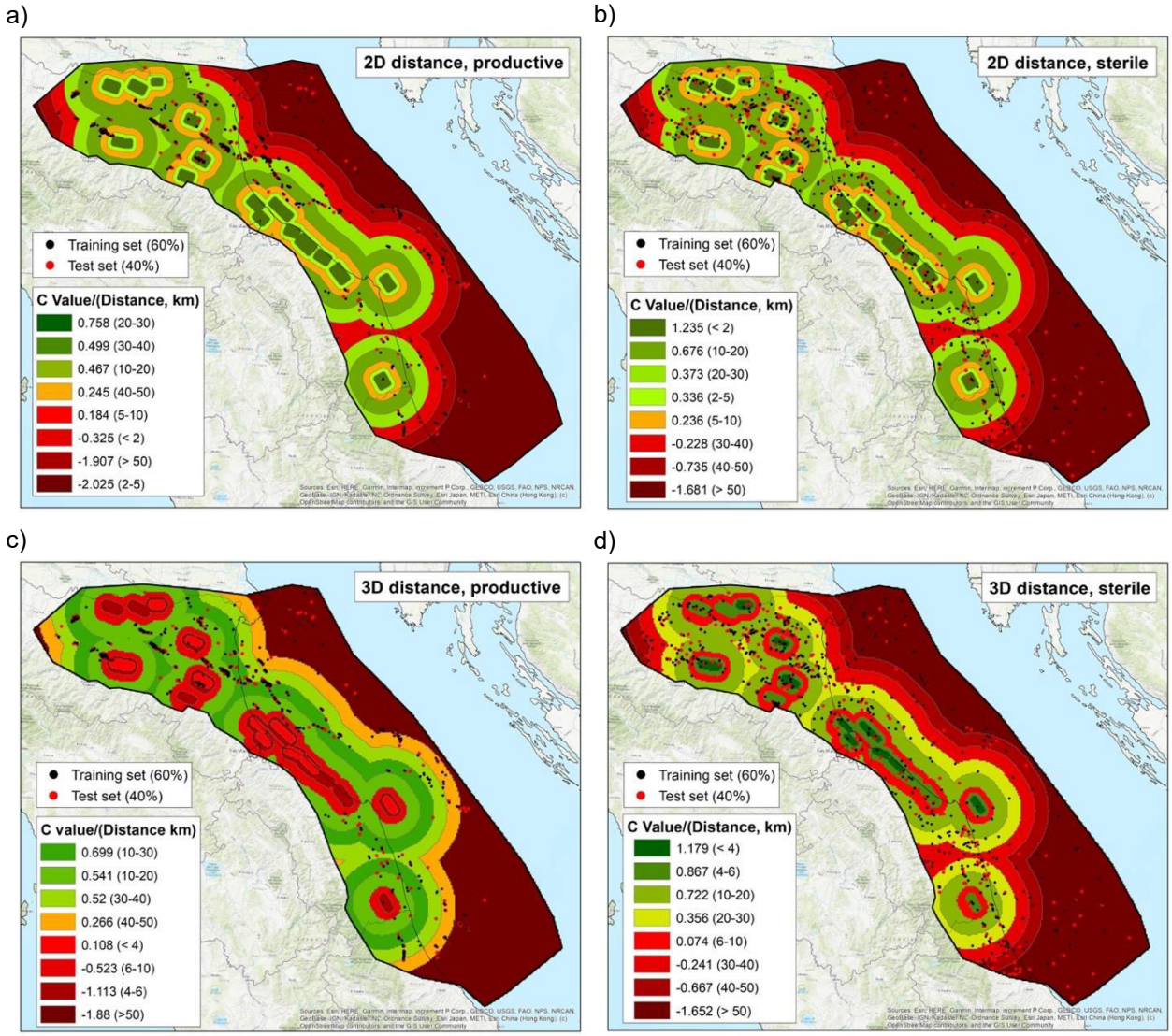
In marked contrast, for sterile wells the largest C values (1.235) and the largest Studentised contrast (9.093) are found for the first distance bin (0-2 km). C drops quickly for increasing

distance ranges, becoming negative (-1.681: C/S(C) -9.833) for a distance >50 km. These results indicate that gas wells falling very close to a seismogenic fault are likely to be sterile.

3D distances. Table 2b and Figs 2c, 2d summarize the results obtained using the *training dataset* for the eight adopted distance bins in the 3D case.

The results we obtained for the productive wells indicate a non-significant correlation (C=0.108) for the shortest distance bin (< 4 km). C is negative in the 4-6 km and 6-10 km bins (-1.113 and -0.441) and becomes positive C for distances >10 km, implying that productive wells generally "stay away" from seismogenic faults. For distances >50 km we found a negative C (-1.88), consistent with the results obtained for the 2D case. The results suggest no statistical correlation, or a negative correlation, between the location of the productive wells and that of the seismogenic faults.

Figure 2: Distribution of C values for the 2D and 3D distance cases.



Distribution of C values for the 2D (a, b) and 3D (c, d) distance cases, for productive (left) and sterile wells (right), respectively.

As detailed in the Methods section, we calculated the success-rate curve using our *training dataset*, then used the *testing dataset* for estimating the prediction-rate curve (Figure 3).

In the 2D case, 80.3 % and 78.3 % of the data in the *training dataset* are found in the 20% and 35% largest C bins, respectively for productive and sterile wells. The AUC estimated for the success-rate curve is very high for productive wells (90.7%), and good for sterile wells (77.0%).

In the 3D case, 85.5% and 68.8% of the data in the *training dataset* are found in the 20% and 30% of the largest C bins, respectively for productive and sterile wells (Figure 3). The AUC estimated for the success rate curve is very high for productive wells (90.7%) and good for sterile wells (80.0%).

In both the 2D and 3D cases the AUCs calculated from the prediction rate curves for the *testing dataset* are similar to those estimated based on the *training dataset*, suggesting that the model performs well in predicting the typology of the productive/sterile nature of a well depending on its 2D/3D distance from a seismogenic fault.

The results of our statistical analysis show a clear anticorrelation between the spatial distribution of productive wells and of seismogenic faults, whereas the correlation between sterile wells and faults is more tenuous, albeit still statistically significant. In our opinion this circumstance reflects a specific characteristic of the data that will be elucidated later on.

Overview of the results. Unlike other applications of the WofE method, in this work we had to deal with two widely independent sources of uncertainty: (a) the incompleteness of the dataset of gas reservoirs, mainly concerning their depth, and (b) the inevitably limited and fragmentary knowledge on the seismogenic faults; their location and geometry generally rely on good quality geological and exploration seismology data, but their activity is documented by earthquakes that are difficult to assign to a specific causative fault as they occurred up to four centuries ago.

Nevertheless, our analyses document a statistically significant relationship between the location of productive/sterile gas wells and the occurrence of seismogenic faults capable of M_w 5.5+ earthquakes (see the Results section, Table 2 and Figs 2, 3). More specifically, the WofE shows that the highest spatial correlation between the vast majority of productive wells and the selected seismogenic faults is obtained for distances in the range 10 to 50 km: in 381 out of 475 cases (80%) for the 2D distance scenario, and in 405 out of 475 (86%) for the 3D scenario, which is more sensitive to the actual distance of each well from a seismogenic fault. These figures imply that the existence of a productive well above an identified seismogenic source is a statistically rare occurrence throughout our study region. Conversely, sterile wells show a high spatial correlation with the location of seismogenic faults, although – somewhat unexpectedly – the results of the WofE method when using 2D distances are more robust than those obtained while considering 3D distances. These results indirectly support our working hypotheses, but their interpretation requires a closer inspection of the geographical distribution of the data.

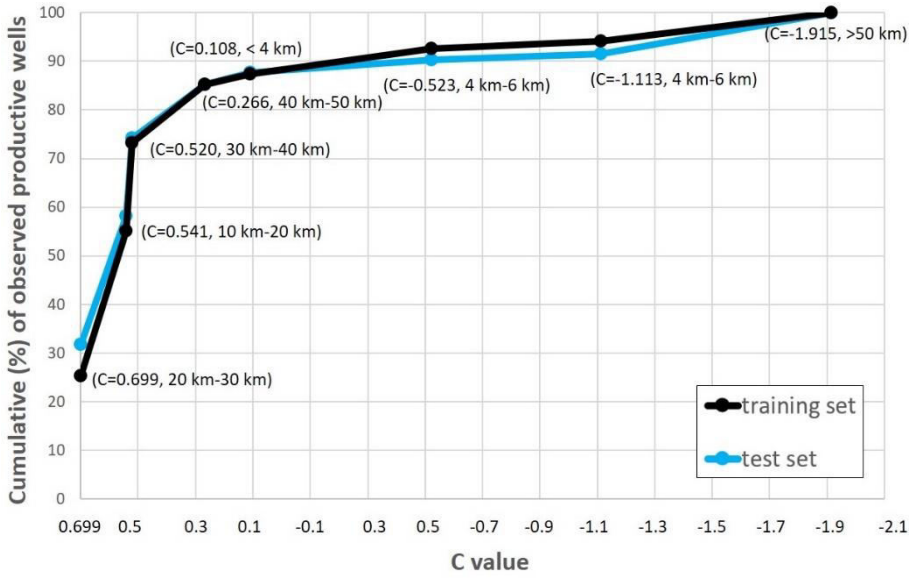
In summary, productive wells "stay away" from large seismogenic faults, whereas sterile wells are often associated with them. The only significant exceptions reported for the 2D scenario are 14 productive wells lying in the class 0-2 km (Table 2a), 11 of which are located within the surface projection of the individual source ITIS100 Bagnacavallo, two fall within the ITIS141 Argenta, and one falls within the ITIS070 Offida (Figure 1, Table 1). In the 3D scenario we find only 10 wells falling within 4 km of the fault.

The occasional presence of productive gas wells (and therefore of a major gas reservoir) above a seismogenic fault could be related to local heterogeneities in the lithology and/or stratigraphy with respect to the predominant Plio-Pleistocene geological and stratigraphic setting of the Apennines fold and thrust belt and foredeep. Although such variations could occur at a local

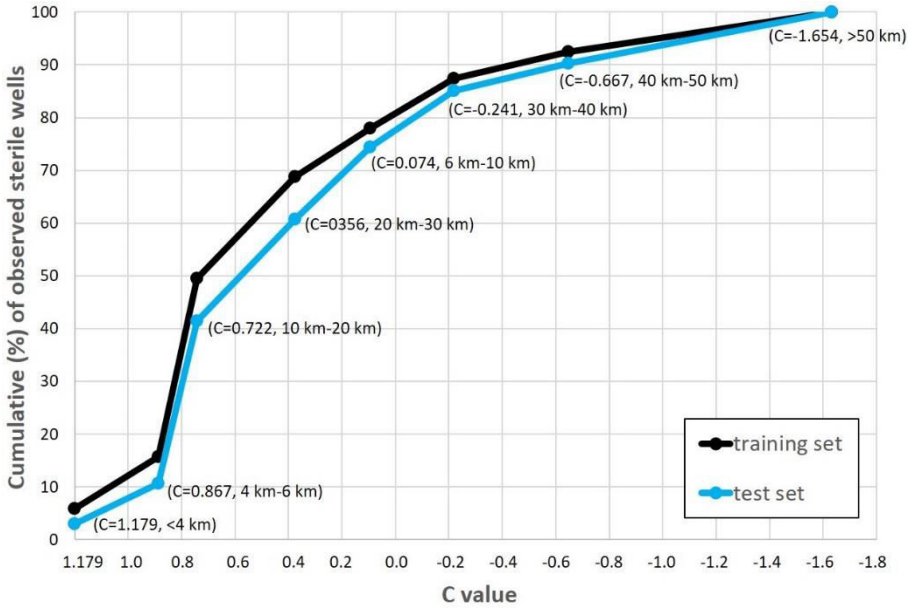
scale (well scale), the dataset we analyzed reveals the nearly exclusive occurrence of Plio-Pleistocene gas-charged or sterile sandy-silty sedimentary succession[16,18,19].

Figure 3: Success-rate curves and prediction-rate curves for the 3D scenario.

a) - Productive



b) - Sterile



Success-rate curves (in black) and prediction-rate curves (in blue) of C distributions calculated for a 3D distance, respectively for productive (a) and sterile wells (b).

Discussion

Do these unsuitably located productive wells challenge our findings? The occurrence of a gas reservoir above a seismogenic fault may indicate that the fault is poorly located, for example because it is deeper than assumed in the DISS database, or that there is a flaw in the physical behavior implied in the proposed anticorrelation between faults and productive wells. For

ITIS100 and ITIS141 the first scenario is more likely, as these two sources are held responsible for earthquakes that occurred over three centuries ago, whereas for the ITIS070 one might consider that the productive well overlying it reaches a depth of 2.5 km, whereas the fault is assumed to extend between 4.5 and 8.7 km depth. As all fault data are inherently uncertain within a few km, and in consideration of the large number of data analyzed in this work, we consider it legitimate to stick to the results of the statistical tests without chasing individual outliers.

Statistical significance of the data and of the results. We also noticed that for productive wells, large C values are obtained only for relatively large well-to-fault distances, whereas sterile wells correlate with seismogenic faults primarily for short distance ranges, but exhibit relatively high C values also for more distant ranges, up to 30 km. Also, productive gas wells tend to cluster in space, whereas sterile wells exhibit a more scattered distribution over the whole study area (this is clearly seen in Figs 1, 2). These two conditions are a clear indication that both productive and sterile wells may not provide a reliable portrait of their own distribution.

Why does this happen? Generally, once a newly drilled well is found productive, additional wells are drilled around or aligned with it; conversely, if that well is found sterile the exploration generally stops, as further activity nearby is unlikely to be economically attractive. The result of this quite understandable imbalance is that, at least in our study area, the exploration has carefully circumvented large areas that were suitable in principle, i.e. based on their geological setting, but had already shown to be unproductive. This is the case of the nearly 120 km-long coastal strip stretching from Rimini to Ancona, a region where seismic hazard is consistently higher than in structurally similar, adjacent areas (http://zonesismiche.mi.ingv.it/), but where there exist virtually no productive wells and only fewer-than-average sterile wells (Figs 1, 2).

The combination of these circumstances shows (a) that the wells are not randomly distributed in space, as an ideal statistical analysis would require, and (b) that certain "extreme" areas, where the absence of productive reservoirs has been known for decades, are severely undersampled. Nevertheless, we maintain that this inevitable imbalance of our sample (a) is at least partially compensated by its richness, and that (b), if anything, a more uniform sample would have returned an even clearer correlation/anticorrelation between gas wells and seismogenic faults.

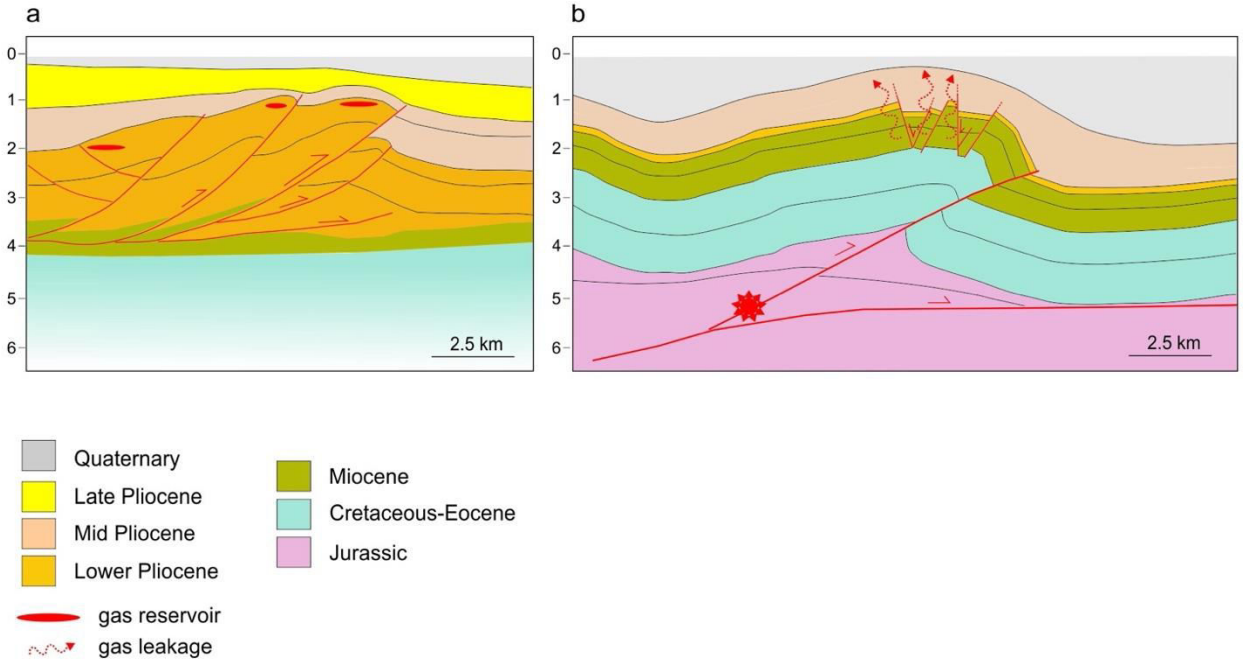
Structural interpretation of the results. Having established the statistical significance of the results, the next question is: what relates earthquakes and hydrocarbon reservoirs? Unquestionably, the success or failure of hydrocarbon exploration in a fold and thrust belt is primarily controlled by the relative timing of source rock maturation, hydrocarbon migration and trap development. The outcomes of our study, however, strongly suggest that an equally crucial pre-requisite for the formation of an efficient gas reservoir is that the deposits overlying the reservoir-bearing formations be unaffected by active fractures and faults which might allow fluids to escape to the atmosphere. This condition is not warranted in earthquake-prone areas[21].

A close inspection of Figure 1 – and of the ample literature that lies behind (see the Introduction) – shows that most productive reservoirs are hosted in small-scale anticlines (Figure 4a), generated by faults that are shorter and narrower with respect to the deep and large faults driving long-wavelength folds that may generate significant earthquakes and where gas is generally not found (Figure 4b). In the DISS database the larger and positively seismogenic faults responsible for creating potentially large folds are listed as Individual Seismogenic Sources (ISSs); smaller folds/reservoirs are either delineated as Composite Seismogenic Sources (CSSs), implying that they may generate M_w 5.5+ earthquakes but their ability to slip seismically vs. aseismically is

currently unknown, or are not mapped at all, implying that they are too small to generate M_w 5.5+ earthquakes.

Mucciarelli et al.[21] contended that if the fault that controls the evolution of a potential hydrocarbon trap generates a significant earthquake, the integrity of the sealing horizons will inevitably be jeopardized. They also stressed that any subsequent gas loss is not caused by the earthquake shaking *per se*, but rather by the actual fault slip and its consequences. A M_w 5.5 earthquake is necessarily the result of a few cm of coseismic slip of a fault that may extend over a considerable thickness of the upper crust – generally between 2-3 and 8-10 km in our study area – and potentially up to the surface; larger events will cause larger slip over a larger fault. Coseismic strain may propagate upward following different mechanisms, including the presence of overpressured fluids, leading to the creation or rejuvenation of fractures and faults that may immediately act as pathways for fluid escape. Such mechanisms are compatible with the limited distance we calculated between the location of sterile wells and the location of known seismogenic faults – less than 2 km for the 2D case, less than 6 for the 3D case (see Table 2 and S1) – even taking into account the uncertainties discussed earlier.

Figure 4: Schematic cross sections across the outermost fronts of the Apennines fold and thrust belt.



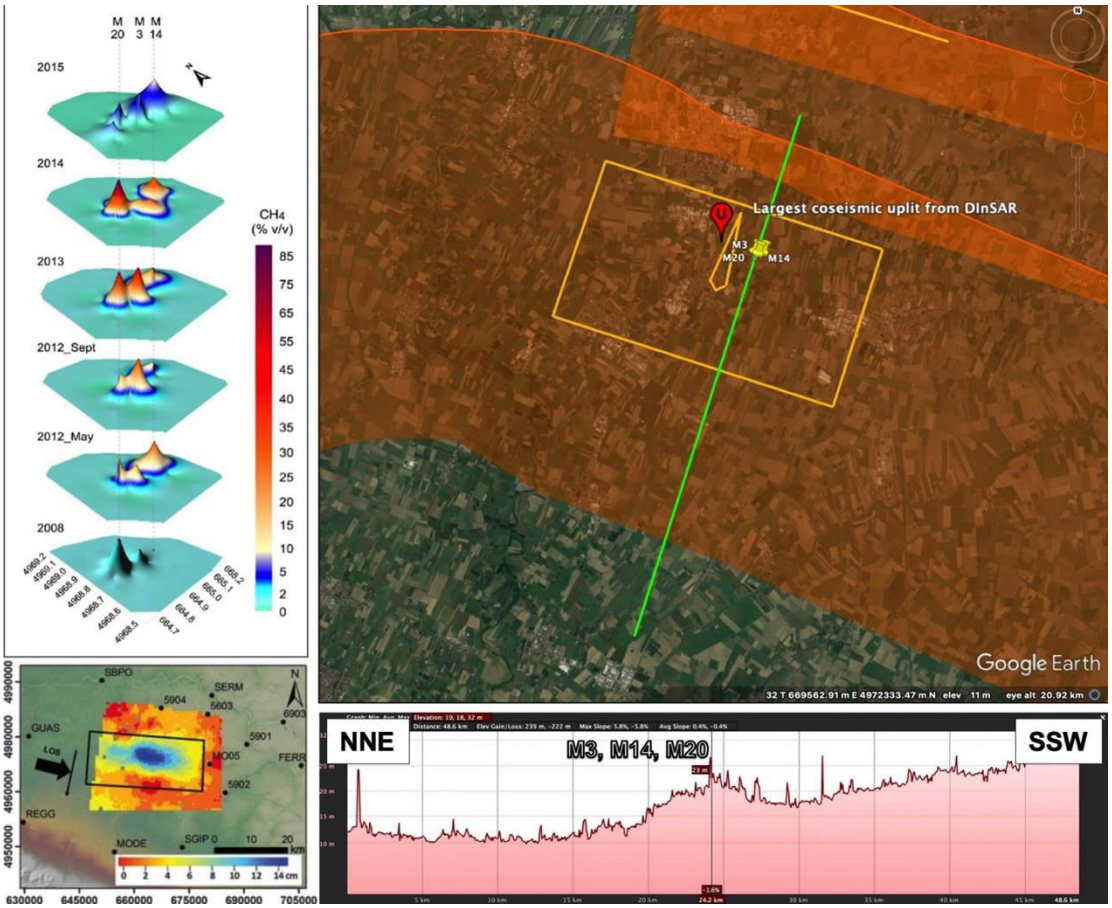
a) Small-scale anticlines hosting gas reservoir (modified from Casero & Bigi[17]). b) Larger-scale folds, driven by deep and large faults, which generate relatively large earthquakes (Mw 5.5+: red star; modified from Emami et al.[23]).

Notice that the crustal volume located at the tip of a thrust fault has long been known to be characterized by extensive fracturing, allowing for an efficient migration of fluids toward the surface[24-28] . The density of fractures and the wide range of their orientation is typical of the fault damage zone, a volume of intense rock deformation resulting from the initiation, propagation and build-up of slip along the main underlying fault[29-32].

All of these processes are well described in the literature and have been extensively investigated with analogue and analytical models, but they have seldom been observed in the field. A notable exception is discussed by Sciarra et al.[33], who investigated in detail the release of underground fluids, including CH_4, following the 29 May 2012, M_w 6.0 earthquake near Mirandola (southern

Po Plain). The event was generated by a thrust fault lying beneath the Mirandola anticline and was assumed to be potentially seismogenic as early as 2000 in a prototype of the DISS database. At least six sterile wells (Bignardi, Camurana, Cavone, Concordia, Medolla, San Biagio) lie on the surface projection of the fault, supporting the anticorrelation advocated in this work. The area located just northeast of the village of Medolla has been known at least since 1838 for the existence of gas macroseeps located exactly above the culmination of the buried Mirandola anticline (Figure 5), as suggested by subsurface geology and topographic evidence and by the pattern of coseismic elevation changes following the 2012 earthquake[34]. The combination of this evidence with the data presented by Sciarra et al.[33] suggests that these macroseeps are the surface evidence of extrados normal faults associated with the buried Mirandola anticline (Figure 5). Their CH_4 flux measurements show that the sealing horizon of the local reservoir is partially efficient during the interseismic period but fails catastrophically following an earthquake and the subsequent remobilization of the extrados faults, causing accelerated degassing for at least five years.

Figure 5: Location of the historically documented Medolla macroseeps.



Summary of occurrrences in the region of *Terre calde* (Warm grounds), a set permanent microseeps documented historically near Medolla. Their name relates to the continuous degassing, which eventually heated the ground causing anticipated melting of the snow cover in winter.

Upper left panel): changes in the natural release of CH_4 between 2008 and 2015 for the M3, M14 and M20 emissions[33]. Upper right panel): surface projection of the Individual Seismogenic Source responsible for the 29 May 2012 earthquake, shown in yellow, overlying Compisite Seismogenic Sources, in orange (all from the DISS database). Lower left panel): vertical dislocation detected by DInSAR interferometry following the 29 May 2012 Mirandola (Emilia) earthquake[34]. Lower left panel): topographic profile of the area along a NNE-SSW trend.

Such mechanism of continuous degassing due to extrados faulting has been shown to dominate in the northern Apennines, at the western end of our study area[35, 36], and may be common elsewhere. We contend that a similar mechanism might explain many of our observations over the entire study area, and over any similar areas around the globe. See text for further discussion.

The case of Medolla illustrates a mechanism of progressive loss of underground fluid pressure, ultimately motivated by earthquake activity but not restricted to the coseismic phase. Based on the number of wells that have been unsuccessfully drilled in the area one should conclude that the geological setting of the area is fully compatible with the existence of an effective hydrocarbon reservoir associated with the Mirandola anticline, but also that the integrity of this reservoir is long gone.

Seismic hazard implications. Our dataset shows that in a fold and thrust hydrocarbon province the lack of productive gas reservoirs is likely to be controlled by seismogenic faulting, Conversely, the presence of significant reservoirs is in itself an indication of a predominantly aseismic behavior of the underlying faults; in other words, the ability of a thrust fault to generate significant earthquakes depends primarily on the rheology of the basin deposits. Their behaviour ultimately controls the fate of the reservoirs formed at the core of the largest anticlines; favouring their integrity by promoting aseismic slip over the underlying fault, or causing gas to be lost to the atmosphere by promoting stick-slip.

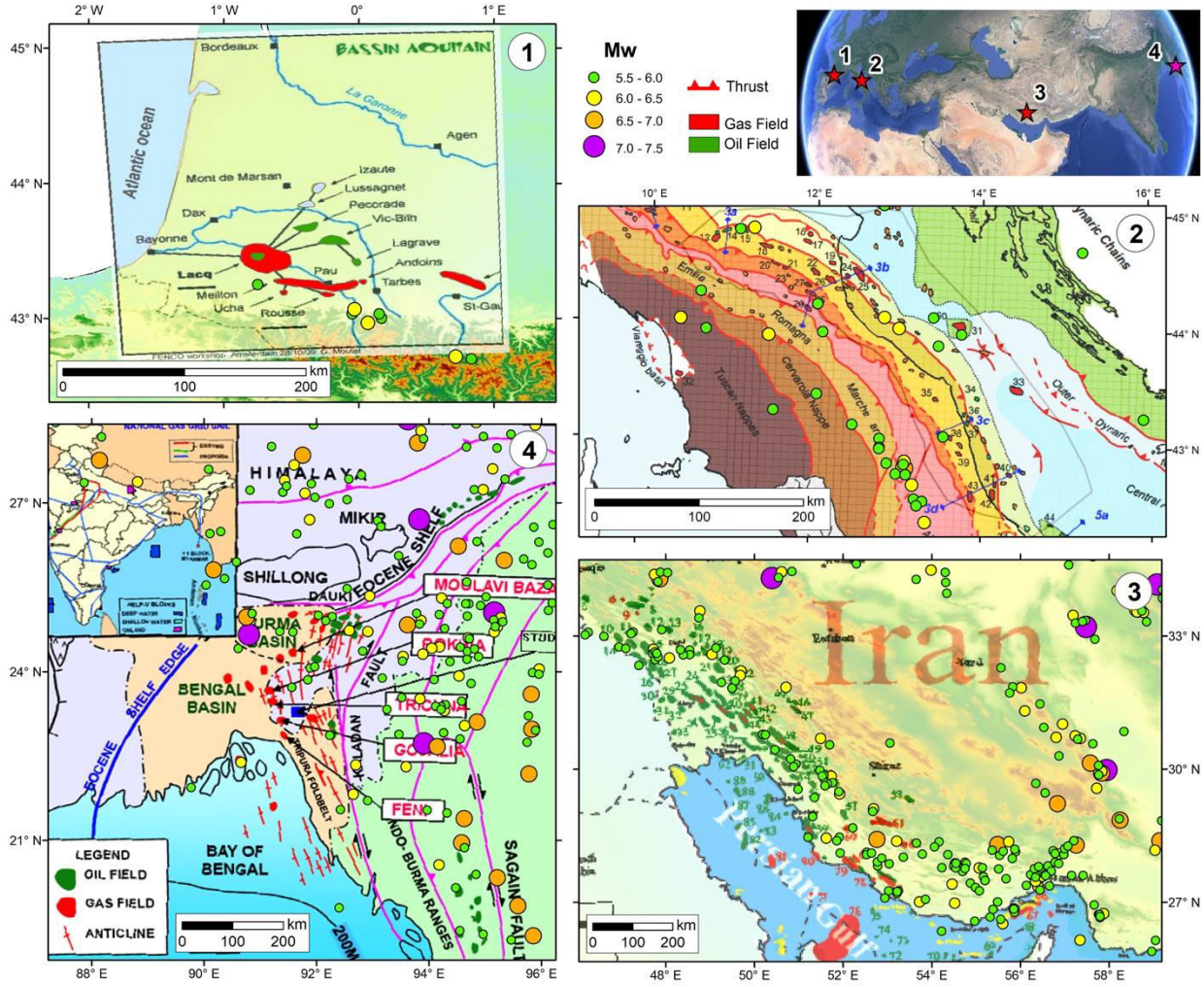
These results are consistent with the observations of Carafa et al.[7], who showed that the seismic moment released by Italian contractional domains is about 50% of the total moment expected based on fault geometries and slip rates. This may imply that Italian thrust faults alternate seismic and aseismic phases in equal proportions, but our findings rather suggest that the existence of a productive gas field proves that the underlying thrust fault slips always aseismically, and viceversa. If no other indication of aseismic behaviour is available, for instance from GPS, this evidence alone could justify a significant seismic hazard reduction over thrust faulting systems that exhibit little or no long-term seismicity.

Worldwide analogs. Our hypotheses are qualitatively supported by other hydrocarbon-bearing fold and thrust belts developed at convergent plate boundaries worldwide (see Figure 2 of Cooper[37], and Figure 1 of Goffey et al.[38]), which contain 14% of the known reserves worldwide. About half (49%) of these reserves are found in the Zagros fold belt of Iran, Iraq, and Turkey[37].

We selected three specific areas that share a comparable geological setting with our study area and for which both reliable seismological data and qualitative data on the location of productive gas fields are available: the Zagros-Persian Gulf area (southern Iran), the Aquitaine Basin (southwestern France) and the Bengal Basin (Bangladesh-India boundary; Figure 6). Unfortunately, none of these regions features rich, reliable and openly accessible databases such as ViDEPI and DISS. For this reason, and to avoid any bias in the interpretation of the scant available evidence, we deliberately adopted original maps from published literature, after carefully georeferencing them, and plotted the seismicity from publicly accessible databases.

Aquitaine Basin. This historical hydrocarbon province is located in the southwest of France, covering an area of about 35,000 km^2 (Figure 6, panel 1). It comprises several geological domains, including the Pyrenean fold and thrust belt and its foreland, formed as a consequence of the subduction of the Iberian plate beneath the European plate[42]. To the north, the leading edge of the fold and thrust belt hosts the most important hydrocarbon field of metropolitan France, the Lacq gas field[43]. Gas was confined in a structural trap within the pre-Aptian carbonate sequence located in an anticline at a depth of 3,500 to 4,050 m, below the oil field, located at 650-700 m depth. Gas was exploited until 2013, when the field became one of the most important pilot sites for Carbon Capture and Storage (CCS[44]).

Figure 6: Earthquake-prone, hydrocarbon-bearing fold and thrust belts that share similarities with our study area.



Panel 1: Aquitaine Basin (from Moutet[39]). Panel 2: our study area (from Casero[19]). Panel 3: Zagros fold and thrust belt (from Esrafili-Dizaji & Rabbani[440]). Panel 4: Bengal Basin (from Brahma et al.[41]). Following a widespread practice in the hydrocarbon industry, gas fields are outlined in red, whereas oil fields are shown in green.

Zagros fold and thrust belt**.** This belt forms the external portion of the Zagros active orogenic wedge (Figure 6, panel 3). It is the most recent result of the convergence and closure of the Neo-Tethys oceanic domain between Arabia and Eurasia[45]. In the Zagros fold and thrust belt the Arabian passive margin sequence has been decoupled from its basement and deformed by large-scale folding and thrusting. A rich seismicity record indicates that within the underlying Panafrican basement, shortening is presently accommodated by reverse faulting[5,46,47].

The Zagros fold and thrust belt comprises one of the most prolific hydrocarbon provinces worldwide. Most of the gas fields are concentrated in the Fars region and in the contiguous offshore area and accumulated in the Permo-Triassic carbonates, within the Kangan and Dalan formations, which are the most important carbonate reservoir rocks in the southwestern part of Iran[48,49]. Gas reservoirs lie at depths ranging from 2,600 to 3,600 m (e.g. Salman gas field[49]).

Bengal Basin**.** This large, highly productive and still largely unexplored hydrocarbon province lies on the eastern side of the Indian subcontinent, between the Shillong Plateau to the north, and the Indo-Burman Ranges to the east, in the Tripura district (Figure 6, panel 4). This basin originated during the collision of India with Eurasia and Burma, building the extensive

Himalayan and Indo-Burman mountain ranges, and thereby loading the lithosphere to form flanking sedimentary basins[50]. The Bengal Basin is a prolific petroleum-bearing basin where 25 economically-viable fields provide about 18% of India's known reserves. It contains up to 22,000 m of Cretaceous to Holocene sedimentary fill[51], which includes 4,000-5,000 m of hydrocarbon-bearing Surma Group sandstones[52], deposited as a consequence of major Miocene uplift in the Himalayas which funneled large volumes of sediments into the basin.

In spite of all these limitations, all three cases exhibit a pattern that is broadly consistent with the evidence from our study region. Clearly, further work is needed to explore these cases as we did for Italy, also in view of the limited resolution of the data available for them. Italian hydrocarbon data have become available after a 60 years-phase of intense exploitation that is inevitably coming to an end, and solid earthquake and seismogenic source data are regularly made available by INGV through an agreement with the Italian Civil Protection Department.

Yet, the lesson learned from Italian data may be useful for all scientists and hydrocarbon industry practitioners who are puzzled by the relationships between the exploitation of gas reservoirs and the occurrence of significant earthquakes.

Understanding these relationships has crucial implications for three independent issues of great societal relevance: (a) supporting the decision-making in the exploitation of as yet unexplored gas pools; (b) offering crucial insight into the seismic hazard of areas undergoing active crustal shortening, and specifically on the seismic/aseismic ratio of the local strain release; and (c) providing guidelines for the identification of suitable underground energy storage sites, which represent one of the viable options for a gradual transition from fossil fuels to a carbon-neutral economy[53]. Our findings indicate that the best option for planning such facilities is to stay away from large seismogenic faults and opt for a depleted reservoir, as choosing a reservoir whose past performance is unknown would greatly increase the risk of dispersion of CH_4 to the atmosphere. In fact, all existing Italian gas storage sites run by Stogit-SNAM and by Edison Stoccaggio SpA totally leave out the Rimini-Ancona stretch, the highest seismic hazard portion of our study area (Figure 1).

Methods **(1589 words)**

To investigate any spatial correlation between the location of productive gas fields and the occurrence of strong earthquakes (M_w 5.5+) in our study area (Figure 1) we adopted the Weight of Evidence (WofE) method, a bivariate statistical analysis based on the Bayesian probability framework[22]. The data considered in the analysis were collected from two public-domain datasets: the Database of Individual Seismogenic Sources, built and maintained by INGV (http://diss.rm.ingv.it/diss/), and the "Visibility of Petroleum Exploration Data in Italy" (ViDEPI) database of borehole data, assembled by the Italian Ministry of Economic Development (https://www.videpi.com/videpi/pozzi/pozzi.asp).

The Database of Individual Seismogenic Sources (DISS). DISS is an original georeferenced repository of tectonic, fault, and paleo-seismological information for the whole of Italy and its surrounding countries and seas. It was conceived in the late 1990s as a tool devoted to supporting seismic hazard assessment in Italy[54]. After its first publication in 2001[55], it has been revised and updated several times. The most recent version was published in 2018 (v. 3.2.1[56]).

405 The DISS contains information on seismogenic faults located in the Italian peninsula and the
406 surrounding regions. The information is supplied through two categories of seismogenic
407 sources[54]: Individual (ISSs), and Composite (CSSs). The ISSs are simplified (rectangular) 3D
408 representations of seismogenic fault planes, generally associated with significant known
409 earthquakes. For each ISS the database provides a full set of geometric (strike, dip, length, width,
410 and depth), kinematic (rake) and seismological parameters (single event displacement, slip rate,
411 recurrence interval), including the expected maximum earthquake size. In contrast, the CSSs are
412 simplified and 3D representations of crustal fault systems that are not segmented nor associated
413 with one or more specific earthquakes: each CSS hence may contain an unknown number of
414 ISSs. For this study we considered the eighteen ISSs falling within the study area (Figure 1a),
415 whose parameters (including the most recent earthquakes that can be associated with each of
416 them) are summarized in Table 1. All ISS are shown using a 3D perspective in Figure 1b, which
417 allows one to appreciate the relationships between their position and the location of the
418 productive and sterile gas wells, considered as proxies for the relevant reservoirs. The CSSs are
419 not used in this work, but are shown in Figure 1 to highlight the continuity of the main
420 seismogenic trends.

421 **The ViDEPI database of Italian hydrocarbon data**. We carried out a comprehensive analysis
422 of data from 1,769 wells, a subset of a large body of information acquired since 1957 by several
423 oil companies and made available by the Italian Ministry of Economic Development within the
424 ViDEPI project (http://www.videpi.com/). Current Italian regulations establish that the oil
425 companies shall provide the Ministry with technical reports on the activities that have been
426 carried out. The data, which were originally available only as hard copies, have been scanned,
427 geo-referenced and distributed free of charge through the project website. The available data
428 consist of composite logs containing the following information: (1) lithology, (2) geological
429 formation name(s), (3) formation(s) age(s), (4) depths, (5) litho-stratigraphy, (6) fluid
430 occurrence, (7) depositional environment of each formation, (8) biostratigraphy, and (9)
431 geophysical logs. Pressure and temperature values are sometimes reported.

432 We analysed in detail each well falling within the study area, gathering the parameters needed
433 for our investigation and for assessing their reliability based on the quality of accompanying
434 information. We first removed 49 poorly documented wells, 59 shallow gas wells (<500 m) and
435 49 dominantly oil wells, as they do not provide relevant information for our scopes, thereby
436 focusing exclusively on deeper gas wells. We then subdivided the remaining 1,612 wells into
437 two categories: 1) positively sterile, when they encountered no exploitable hydrocarbons, and 2)
438 positively productive, if they have been or are currently being exploited.

439 **Building a model for the GIS analysis**. Information regarding the selected wells was loaded to
440 ArcGis along with the 18 selected Individual Seismogenic Sources from the DISS database
441 (Figure 1). Thanks to the information provided by the ViDEPI, all wells were first projected
442 down to their actual maximum depth, then classified based on the criteria delineated above and
443 on the depth of the exploited hydrocarbon pool (Figure 1b).

444 **Weight of Evidence (WofE) method.** This method consists of a bivariate statistical analysis
445 based on the Bayesian probability framework[22]. It allows the association of potential evidence in
446 support of a given hypothesis and is easy to run within common GIS software packages. We
447 adopted it to evaluate the spatial relationships between the distribution of productive or sterile
448 gas wells and the location of major seismogenic faults.

449 The WofE method is commonly used in the geostatistic field, for example for identifying
450 potentially productive mining areas, locating flowing wells, or mapping cliff instabilities
451 associated with mine subsidence[57]. Recently it was used also for assessing landslide
452 susceptibility[58-60]. The computational framework of the method[22,57] is briefly summarized below.

453 The method estimates the weight for evidence (W), which can be positive or negative:

$$W^{+} = ln \frac{P\{F|E\}}{P\{F|\overline{E}\}} \quad (1)$$

454

$$W^{-} = ln \frac{P\{\overline{F}|E\}}{P\{\overline{F}|\overline{E}\}} \quad (2)$$

455

456 where $P\{F|E\}$ and $P\{F|\overline{E}\}$ are the conditional probabilities of being and not being within the
457 factor F, given the presence or the absence of the event E. $P\{\overline{F}|E\}$ and $P\{\overline{F}|\overline{E}\}$ indicate the
458 conditional probability of not being within the factor F, given the presence or the absence of the
459 event E.

460 The difference between the positive and negative weights is defined as the weight contrast C (=
461 W+ - W-), which provides a measure of the spatial correlation between a certain factor and the
462 occurrence of an event (see Table 2, S1, and Figs 2, S1). A negative C indicates the absence of a
463 factor where an event occurs, whereas a C around zero implies no significant relationship with
464 the occurrence of the event.

465 The statistical significance of the weights and of C can be estimated by considering their
466 standard deviation and variances as:

$$S(C) = \sqrt{S^2_{W+} + S^2_{W-}} \quad (3)$$

467

$$S^2_{W^+} = \frac{1}{P\{F|E\}} + \frac{1}{P\{F|\overline{E}\}} \quad (4)$$

468

$$S^2_{W^-} = \frac{1}{P\{\overline{F}|E\}} + \frac{1}{P\{\overline{F}|\overline{E}\}} \quad (5)$$

469

470 where S^2_{W+} and S^2_{W-} are the variance of W^+ and W^-, respectively[61]. Finally, the Studentised
471 Contrast C/S(C) provides an assessment of the quality of the results.

472 To estimate the conditional probabilities of equation (1) and (2) we first discretised the entire
473 study area with a regular 10x10 m grid comprising 31,541 x 39,330 pixels (lines x columns). We
474 then calculated the distance of each pixel to the nearest seismogenic fault, following two
475 strategies: in 2D (planar), or in full 3D, which we consider more appropriate for investigating the
476 motivations of the statistical evidence. These distances correspond to the factor F in equations
477 (1) and (2).

478 For the 3D case, we calculated the shortest distance from the bottom of each well to the surface
479 of the nearest seismogenic fault. In contrast, for the 2D case we calculated the shortest distance
480 between the head of each well and a point along the external boundary of a 2 km-wide buffer
481 area encircling the surface projection of the nearest seismogenic fault. All calculated distances
482 were then grouped into eight bins: 0-2, 2-5, 5-10, 10-20, 20-30, 30-40, 40-50, 10-50 and >50 km

for the 2D case; and 0-4, 4-6, 6-10, 10-20, 20-30, 30-40, 40-50, 10-50, >50 km for the 3D case, respectively. Notice that the width of the distance bins was decided based on a series of trial-and-error tests, aimed at exploring any possible bias in the results due to the choice of the bin size. We maintain that the selected bin widths allow a fine sampling of well-fault distances for the shorter distance values, which is crucial for any geological and geophysical interpretation of the results, without generating any undesired bias. Finally, each pixel was assigned to a group, depending on the above classification. Table S1 and Figure S1 show the distribution of gas wells and the spatial distribution of the different classes of pixels with respect to the binning scheme (for the 3D case only).

Based on equations (1) and (2), the weights of evidence can be modified (when analysing one group at a time) in numbers of pixels, as in the following:

$$W_i^+ = \ln \frac{\frac{N\text{pix}_1}{N\text{pix}_1+N\text{pix}_2}}{\frac{N\text{pix}_3}{N\text{pix}_3+N\text{pix}_4}} \qquad W_i^- = \ln \frac{\frac{N\text{pix}_2}{N\text{pix}_1+N\text{pix}_2}}{\frac{N\text{pix}_4}{N\text{pix}_3+N\text{pix}_4}} \tag{6}$$

where $N\text{pix}_1$ is the number of pixels – that is to say, wells – falling within the considered distance bin (e.g., the 0-2 km interval); $N\text{pix}_2$ is the number of pixels that correspond to a gas well (either productive or sterile) but fall outside the considered distance bin; $N\text{pix}_3$ is the number of pixels related to a potential event predictive factor (selected group of distance) but without any wells; and $N\text{pix}_4$ is the number of pixels where neither the considered potential event predictive factor nor a well are observed.

The predictive capabilities of the proposed model were checked through a cross-validation procedure which consists in subdividing a data sample into two subsets[62] The analysis is then performed on one subset, referred to as *training dataset*, which includes 60% of the wells, randomly selected from the whole available sample, while the validation is carried out on the remaining 40% of the data, referred to as *test dataset*.

The validity and the accuracy of the C distributions were tested using the *success-rate* and *prediction-rate* curves in combination with the Area Under the Curve (AUC): a method that provides a measure of the total accuracy based on the rate curves, where a total area equal to one indicates perfect accuracy[63] (Figure 3). The success-rate curve, which is based on the *training dataset*, shows how good the selected parameter curves are in fitting the known wells (productive or sterile). The prediction rate curve, which is based on the *testing dataset*, provides quantitative information for forecasting the position of a productive or sterile well. The most common procedure involves sorting in descending order the calculated index values that refer to the total number of cells in the study area.

Data availability

All hydrocarbon data used in this work come from the ViDEPI project database (https://www.videpi.com/videpi/videpi.asp).

All seismogenic source data were obtained from DISS, the Database of Individual Seismogenic Sources, v. 3.2.1 (http://diss.rm.ingv.it/diss/).

The seismological data used to construct panel 1 of Figure 6 were taken from 'The French seismic CATalogue (FCAT-17)'[64], whereas for panels 2, 3 and 4 they are from the ISC database ISC (1904-2017): http://www.isc.ac.uk/iscgem/download.php.

The underground storage facilities operated by Stogit-SNAM SpA are available at https://www.snam.it/en/about-us/geographical-presence/index.html.

The underground storage facilities operated by Edison Stoccaggio SpA are available at https://www.edisonstoccaggio.it/en/activities-and-facilities/our-plants/.

All of these sites were last accessed on 30 August 2021.

Acknowledgments

This paper is gratefully dedicated to the memory of Marco Mucciarelli, who passed away prematurely in 2016. Marco was a curious investigator and an outstanding scientist who loved 'wandering off the beaten path'. In 2014 he started research on the object of this paper with two of us (GV and FD), following the rather unexpected May 2012 earthquakes in the southern Po Plain. We are deeply indebted with him for his intuitions, his downright rigorous approach and his continuous encouragement.

Author contributions

G.V., F.D. and S.P. laid out the project. F.D. and G.R. searched the ViDEPI database extensively, did a far-reaching literature search and selected the data that were suitable for the analysis, based on extensive discussions with all coauthors. F.D. also compiled the "Worldwide analogs" section of the paper. G.V. provided seismogenic fault data, managed the manuscript organization and did most of the writing. S.P. laid out the statistical tests and participated in the writing. A.T. did the GIS processing of the data and the statistical calculations, drafted most of the figures. In cooperation with S.P. he also wrote the Methods section. All authors discussed the results and laid out the section dealing with the main findings of the work.

References

1. Ellsworth, W.L. Injection-induced earthquakes, *Science* **341**, 6142, doi: 10.1126/science.1225942 (2013).
2. Keranen, K.M., Savage, H.M., Abers, G.A. & Cochran, E.S. Potentially induced earthquakes in Oklahoma, USA: Links between wastewater injection and the 2011 Mw 5.7 earthquake sequence. *Geology* **41** (2013).
3. Rubinstein, J.L. & Mahani A.B. (2015). Myths and facts on wastewater injection, hydraulic fracturing, enhanced oil recovery, and induced seismicity. *Seismol. Res. Lett*. **86**, 1060–1067 (2015).
4. Yeck, W.L., Hayes, G.P., McNamara, D.E., Rubinstein, J.L., Barnhart, W.D., Earle, P.S., & Benz, H.M. Oklahoma experiences largest earthquake during ongoing regional wastewater injection hazard mitigation efforts. *Geophys. Res. Lett*. **44** (2017).
5. Nissen, E., Tatar, M., Jackson, J. A. & Allen, M. B. New views on earthquake faulting in the Zagros fold-and-thrust belt of Iran. *Geophysical Journal International* **186**, 928-944 (2011).
6. Kanamori, H. Great earthquakes at island arcs and the lithosphere. *Tectonophysics*, **12**, 187-198 (1971).
7. Carafa, M.M.C., Valensise, G. & Bird, P. Assessing the seismic coupling of shallow continental faults and its impact on seismic hazard estimates: a case-study from Italy. *Geophysical Journal International* **209/1**, 32–47 (2017).
8. Mantovani, E., Viti, M., Babbucci, D., Tamburelli, C. & Cenni, N. How and why the present tectonic setting in the Apennine belt has developed. *J Geol Soc* **176** 1291-1302 (2019).
9. Kruse, S. & Royden, L. Bending and unbending of an elastic lithosphere - The Cenozoic history of the Apennine and Dinaride foredeep basins. *Tectonics* **13**, 278-302 (1994).
10. Malinverno, A. & Ryan, W.B.F. Extension in the Tyrrhenian Sea and shortening in the Apennines as a result of arc migration driven by sinking of the lithosphere. Tectonics, 5, 227-245 (1986).
11. Royden, L., Patacca, E. & Scandone, P. Segmentation and configuration of subducted lithosphere in Italy: an important control on thrust-belt and foredeep-basin evolution. *Geology* **15**, 714-717 (1987).
12. Patacca E. & Scandone P. Geology of the Southern Apennines, *Boll. Soc. Geol. It*. (*Ital. J. Geosci*.) **7**, 75-119 (2007).
13. Barchi, M. The Neogene-Quaternary evolution of the Northern Apennines: crustal structure, style of deformation and seismicity. In Journal of the Virtual Explorer (eds Beltrando, M. et al.), 36, paper 11, doi: 10.3809/jvirtex.2010.00220 (2010).
14. Bertello, F., Fantoni, R. & Franciosi, R. Overview on Italy's petroleum systems and related oil and gas occurrences. CD Extended Abstract & Exhibitor's Catalogue, 70th EAGE Conference & Exhibition, A018, 4 pp (2008).
15. Bertello, F., Fantoni, R., Franciosi, R., Gatti, V., Ghielmi, M. & Pugliese, A. From thrust-and fold belt to foreland: hydrocarbon occurrences in Italy. *Petroleum Geology Conference series*, **7**, 113-126 (2010).

16. Cazzini, F., Dal Zotto, O., Fantoni, R., Ghielmi, M., Ronchi, P. & Scotti, P. Oil and gas in the Adriatic foreland, Italy. *Journal of Petroleum Geology* **38(3)**, 255-279 (2015).

17. Casero, P. & Bigi, S. Structural setting of the Adriatic basin and the main related petroleum plays. *Mar Petr Geol* **42**, 135-147 (2013).

18. Ghielmi M., Minervini M., Nini C., Rogledi S. & Rossi M. Late Miocene –Middle Pleistocene sequences in the Po Plain-Northern Adriatic Sea (Italy): the stratigraphic record of modification phases affecting a complex foreland basin. *Mar Petr Geol* **42**, 50-81 (2013).

19. Casero, P. Structural setting of petroleum exploration plays in Italy. Italian Geological Society for the IGC 32, Florence-2004, Special Volume, 189-199 (2004).

20. Cartlidge, E. Human activity may have triggered fatal Italian earthquakes, panel says. *Science* **344**, 141 (2014).

21. Mucciarelli, M., Donda, F., Valensise, G. Earthquakes and depleted gas reservoirs: Which comes first? *Natural Hazards and Earth System Sciences* **15**, 2201-2208 (2015).

22. Bonham-Carter, G.F. (Merriam D.F. ed.) *Geographic Information Systems for Geoscientists: modelling with GIS Computer methods in the geosciences* **13**, Pergamon Press, Oxford (1994).

23. Emami, H., Vergés, J., Naipas, T., Gillespie, P., Sharp, I., Karpuz, R., Blanc, E.P. & Goodarzi, M.G.H. Structure of the Mountain Front Flexure along the Anaran anticline in the Pusht-e Kuh Arc (NW Zagros, Iran): insights from sand box models. *Geol Soc Spec. Pub*. **330**, 155-178, (2010).

24. Berkowitz, B. Analysis of fracture network connectivity using percolation theory. Mathematical Geology 27, 467-483 (1995).

25. Davatzes, N.C. & Aydin, A. Overprinting faulting mechanisms in high porosity sandstones of SE Utah. *J. Struct Geol* **25**, 1795-1813 (2003).

26. Gartrell, A., Zhang, Y., Lisk, M. & Dewhurst, D. Fault intersections as critical hydrocarbon leakage zones: integrated field study and numerical modelling of an example from the Timor Sea, Australia. *Mar Petr Geol* **21**, 1165-1179 (2004).

27. Kostakioti, A., Xypolias, P., Kokkalas, S. & Doutsos, T. Quantitative analysis of deformation along the fault Damage zone of the Klimatia thrust (NW Greece, Ionian Zone). *Bulletin of the Geological Society of Greece* XXXVI, 2004 Proceedings of the 10th International Congress, Thessaloniki (2004).

28. Dimmen, V., Rotevatn, A., Peacock, D.C.P., Nixon C.W., Naerland, K. Quantifying structural controls on fluid flow: Insights from carbonate-hosted fault damage zones on the Maltese Islands. *J Struct Geol* **101**, 43-57 (2017).

29. Cowie, P.A. & Scholz, C.H. Physical explanation for the displacement–length relation-ship of faults, using a post-yield fracture mechanics model. *J. Struct. Geol.* **14**, 1133–1148 (1992).

30. McGrath, A.G. & Davison, I. Damage zone geometry around fault tips. *J. Struct. Geol*. **17**, 1011–1024 (1995).

31. Kim, Y-S., Peacock, D.C. & Sanderson, D.J. Fault damage zones. *J Structr Geol* **26**, 3, 503-517 (2004).

32. Choi J-H., Edwards, P., Ko, K. & Kim, Y-S. Definition and classification of fault damage zones: a review and a new methodological approach. *Earth-Science Reviews* **152**, 70-87 (2016).

33. Sciarra, A., Cantucci, B. & Coltorti, M. Learning from soil gas change and isotopic signatures during 2012 Emilia seismic sequence. *Sci. Rep.* **7(1),** 14187 (2017).

34. Pezzo, G., Merryman Boncori, J.P., Tolomei, C., Salvi, S., Atzori, S., Antonioli, A., Trasatti, E., Novali, F., Serpelloni, E., Candela, L. & Giuliani, R. Coseismic deformation and source modeling of the May 2012 Emilia (Northern Italy), earthquakes. *Seism. Res. Lett*. **84**, 645–655 (2013).

35. Bonini, M. Fluid seepage variability across the external Northern Apennines (Italy): structural controls with seismotectonic and geodynamic implications. *Tectonophysics* **590**, 151–174. doi:10.1016/j. tecto.2013.01.020 (2013).

36. Vannoli, P., Martinelli, G. & Valensise, G. The seismotectonic significance of geofluids in Italy. *Front. Earth Sci.* **9**, doi: doi.org/10.3389/feart.2021.579390 (2021).

37. Cooper, M. Structural style and hydrocarbon prospectivity in fold and thrust belts: a global review. *Geol Soc Spec Pub* **272**, 1, 447-472 (2007).

38. Goffey, G.P., Craig, J., Needham, T. & Scott, R. (eds). Hydrocarbons in contractional belts. Geological Society, London, Special Publications, 348, https:// doi.org/10.1144/SP348.0 (2010).

39. Moutet, G., The Lacq integrated CCS project. Presentation to FENCO workshop, Amsterdam, 28 October 2009. Available at http://www.fenco-era.net/lw_resource/datapool/__pages/pdp_166/20091028_gerard_moutet_new.pdf (2009).

40. Esrafili-Dizaji, B., & Rabbani, J. Geology of Iran's Hydrocarbon reservoirs. Avaie zamin, Geological Association of students - University of Tehran, 3(7), 6-11, https://idoc.pub/documents/iran-oil-and-gas-fields-map-old-reservoirs-8jlkqk6om7l5 (in Persian) (2007).

41. Brahma, J., Sircar, A. & Karmakar, G. P..Hydrocarbon prospectivity in central part of Tripura, India using an integrated approach. *Journal of Geography and Geology* **5(3)**, 116-134 (2013).

42. Biteau, J.-J., Le Marrec, A., Le Vot M. & Masset J-M. The Aquitaine Basin. *Petroleum Geoscience* **12**, 247-273 (2006).

43. Winnock, E. & Pontalier, Y. Lacq Gas Field, France. In: Geology of Giant Petroleum Fields (Halbouty M.T. ed.), Vol. 14, American Association of Petroleum Geologists (1970).

44. Gal, F., Pokryszka, Z., Labat, N., Michel, K., Lafortune, S. & Marblè, S. Soil-Gas Concentrations and Flux Monitoring at the Lacq-Rousse CO2-Geological Storage Pilot Site (French Pyrenean Foreland): From Pre-Injection to Post-Injection. *Applied Sciences* **9(4)** (2019).

45. Molinaro, M., Leturmy, P., Guezou, J.-C., Frizon de Lamotte, D. & Eshraghi, A. The structure and kinematics of the southeastern Zagros fold-thrust belt, Iran: From thin-skinned to thick-skinned tectonics. *Tectonics*, **24** (2005).

46. Jackson, J.A. & Fitch T. Basement faulting and the focal depths of the larger earthquakes in the Zagros mountains (Iran). *Geophys. J. R. Astron. Soc*. **64**, 561– 586 (1981).

47. Berberian, M. Master “blind” thrust faults hidden under the Zagros folds: Active basement tectonics and surface morphotectonics, Tectonophysics 241, 193–224 (1995).

48. Bordenave, M.L. Gas Prospective Areas in the Zagros Domain of Iran and in the Gulf Iranian Waters. AAPG Annual Meeting, March 10-12, 2002, Houston, Texas (2003).

49. Nosrati, A., Kadkhodaie, A., Amini, A., Chehrazi, A., Mehdipour, V., Moslemnezhad, T. Reservoir properties distribution in the framework of sequence stratigraphic units: A case study from the Kangan Formation, Iranian offshore gas field, the Persian Gulf basin. *Journal of Natural Gas Science and Engineering*, **65**, 1-15, doi: 10.1016/j.jngse.2019.02.004 (2019).

50. Uddin, A. & Lundberg, N. Cenozoic history of the Himalayan-Bengal system: sand composition in the Bengal basin. Bangladesh. *Geol. Soc. Am. Bull*. **110**, 497-511 (1998).

51. Alam, M., Alam, M.M., Curray, J.R., Chowdhury, A.L.R. & Gani, M.R. An overview of the sedimentary geology of the Bengal Basin in relation to the regional tectonic framework and basin-fill history. *Sediment. Geol*. **155**, 179-208 (2003).

52. Rahman, M.J.J & Worden, R.H. Diagenesis and its impact on the reservoir quality of Miocene sandstones (Surma Group) from the Bengal Basin, Bangladesh. *Mar Petr Geol***. 77**, 898-915 (2016).

53. European Commission Document MJ-03-19-092-EN-N. Clean energy for all Europeans. Publications Office of the European Union, ISBN 978-92-79-99843-0, doi: 10.2833/21366 (2019).

54. Basili, R., Valensise, G., Vannoli, P., Burrato, P., Fracassi, U., Mariano, S., Tiberti, M.M. & Boschi, E. The Database of Individual Seismogenic Sources (DISS), version 3: Summarising 20 years of research on Italy's earthquake geology. *Tectonophysics* **453**, 20–43 (2008).

55. Valensise, G. & Pantosti, D. (eds.) Database of potential sources for earthquakes larger than M 5.5 in Italy. *Ann. Geophys.* **44,** 797–807 (2001).

56. DISS Working Group. Database of Individual Seismogenic Sources (DISS), Version 3.2.1: a Compilation of Potential Sources for Earthquakes Larger Than M 5.5 in Italy and Surrounding Areas. Istituto Nazionale di Geofisica e Vulcanologia, https://doi. org/10.6092/INGV.IT-DISS3.2.1. http://diss.rm.ingv.it/diss/ (2018).

57. Bonham-Carter, G.F., Agterberg, F.P. & Wright, D.F. Weights of evidence modelling: a new approach to mapping mineral potential. *Stat Appl Earth Sci Geol Surv Canada* **89**, 171–183 (1989).

58. Dahal, K.R., Hasegawa, S., Nonomura, A., Yamanaka, M., Takuro, M. & Nishino, K. GIS-based weights-of-evidence modelling of rainfall-induced landslides in small catchments for landslide susceptibility mapping. *Environmental Geology* **54**, 314-324 (2008).

59. Saponaro, A., Pilz, M., Wieland, M., Bindi, D., Moldobekov, B. & Parolai, S. Landslide susceptibility analysis in data-scarce regions: the case of Kyrgyzstan. *Bull. Eng. Geol. Environ.* **74**, 1117–1136 (2015).

60. Vakhshoori,V. & Zare, M. Landslide susceptibility mapping by comparing weight of evidence, fuzzy logic, and frequency ratio methods. *Geomatics, Natural Hazards and Risk* **7**, 1731-1752 (2016).

61. Agterberg, F.P., Bonham-Carter, G.F. & Wright, D.F. Statistical pattern integration for mineral exploration. In Computer Applications in Resource Estimation Prediction and

712 Assessment for Metals and Petroleum (Gabor, G. & Merriam, D.F. eds.), 1-21, Pergamon
713 Press, Oxford, (1990).

714 62. Chung C.-J.F. & Fabbri A.G. Validation of spatial prediction models for landslide hazard
715 mapping. *Nat Hazards* **30**, 451–472 (2003).

716 63. Lee, S. & Sambath, T. Landslide susceptibility mapping in the Damrei Romel area, Cambodia
717 using frequency ratio and logistic regression models. *Environmental Geology* **50**, 847-855
718 (2006).

719 64. Manchuel, K., Traversa, P., Baumont, D. et al. The French seismic CATalogue (FCAT-
720 17). *Bull Earthquake Eng* **16**, 2,227–2,251, doi: 10.1007/s10518-017-0236-1 (2018).

721

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementaryinformation.docx