

# Citizens' smartphones unravel earthquake shaking in urban areas

Received: 3 February 2025

Accepted: 17 September 2025

Published online: 28 October 2025

 Check for updatesFrancesco Finazzi <sup>1</sup>✉, Fabrice Cotton <sup>2,3</sup> & Rémy Bossu <sup>4,5</sup>

Seismic shaking maps (known as ShakeMaps) have become an essential tool for assessing the impact of a recent earthquake or for scenario studies of future events. The spatial resolution of ShakeMaps remains low, particularly in urban areas where spatial variations can have a major impact on damage due to varying population density. Smartphones, with their internal accelerometers and widespread use, help to increase the density of seismic recording precisely where the potential impact of the earthquake may be greatest. In this work, we use statistical spatial modelling to show that smartphone measurements collected by the Earthquake Network citizen science initiative allow us to map site amplification at high spatial resolution, generate high-resolution ShakeMaps, and improve existing ground motion models. We apply the method to the red zone of Campi Flegrei, Italy, a high volcanic and seismic risk area with around 500,000 inhabitants and characterised by a complex spatial pattern of site amplification.

The assessment of future earthquake scenarios and the rapid estimation of earthquake effects depend on the generation of ground motion maps known as ShakeMaps<sup>1</sup>. In urban areas, where risk is increasingly concentrated because of global urbanisation<sup>2</sup>, and where small spatial variations in shaking intensity have a significant impact on damage and fatalities, the need for high-resolution ShakeMaps is paramount.

ShakeMaps rely on data from seismological networks that measure strong motions (accelerometric networks), complemented by ground motion models (GMMs). The sparser the accelerometric networks, the more ShakeMaps are dominated by GMMs, which usually cannot capture the particularities of the earthquake (e.g. rupture geometry, slip distribution, directivity effect, etc.) or the features of the region concerned (i.e. site conditions) which lead to spatial variation of seismic amplification/deamplification.

Site conditions are typically inferred from empirical relationships between relevant proxies available from regional or global maps<sup>3,4</sup>. The average shear wave velocity of the upper 30 m of the soil column (VS30) is used to characterize site amplification. For a single site, VS30 can be measured directly, but for larger areas and regions, it must be inferred from other parameters such as topographic slope<sup>5</sup>, geology, or from the depth of sediments<sup>6</sup>. The uncertainties of these amplification models as well as the uncertainties associated with GMMs

remain high<sup>6,7</sup>. In recent years, microzonation studies have helped to improve the prediction of site effects<sup>8,9</sup>. Such studies require considerable effort to map the local geological and geophysical properties of the subsurface layers and may have a limited predictive power for ground motion amplification when the effects are due to deep geological layers or complex 3D wave propagation effects.

To improve the spatial resolution of ShakeMaps and reduce their uncertainty, dense measurements of ground accelerations are necessary. Accelerometric networks, especially real-time ones, remain the exception, and the current alternative is to integrate macroseismic data (i.e. the felt experience of eyewitnesses), which can be rapidly crowdsourced after an earthquake<sup>10,11</sup>. However, macroseismic data collected from citizens are not direct physical measurements of ground shaking and may be affected by country-specific sociological and psychological biases<sup>12</sup>.

The enhancement of the resolution of peak ground acceleration (PGA) ShakeMaps in urban areas can be instead facilitated by the spatially dense accelerometer measurements collected by smartphones participating in the Earthquake Network (EQN) initiative<sup>13,14</sup>, which is part of an ongoing global effort to use smartphones for real-time seismic monitoring and early warning<sup>15,16</sup>. Unlike macroseismic data, which are subjective, smartphone data are physical measurements and are in real

<sup>1</sup>University of Bergamo, Bergamo, Italy. <sup>2</sup>GFZ Helmholtz Centre for Geosciences, Potsdam, Germany. <sup>3</sup>Institute of Geosciences, University of Potsdam, Potsdam, Germany. <sup>4</sup>European-Mediterranean Seismological Centre, Arpajon, France. <sup>5</sup>CEA, DAM, DIF, Arpajon, France. ✉ e-mail: [francesco.finazzi@unibg.it](mailto:francesco.finazzi@unibg.it)

time. Their integration in a ShakeMap, however, requires addressing two peculiarities of crowdsourced smartphone networks. First, smartphones are not anchored to the ground and the location of the device within the building (e.g. which floor and above which object) is unknown. As a result, smartphones do not directly measure free-field ground shaking and their records show amplitude shifts compared to measurements from seismological stations<sup>17,18</sup>. Secondly, crowdsourced smartphones and seismic stations are rarely co-located and measurements from citizens' smartphones cannot be easily calibrated against seismological station measurements.

The method we propose for generating PGA ShakeMaps is based on learning the site amplification at high spatial resolution from data fusion of station and smartphone measurements collected during multiple seismic events that have affected the same region. The data fusion is performed by a spatial statistical model that forces site amplification to emerge as a residual of an isotropic decay model centred on each earthquake epicentre. The resulting site amplification map is both time and event invariant, and is subsequently adopted in two relevant applications: (i) to generate the ShakeMap when station/smartphone data become rapidly available after an earthquake; (ii) to simulate the ShakeMap for any future event by assuming its epicentre and magnitude. The first problem is solved by a spatial statistical model which can incorporate both the high-resolution amplification map and the station and smartphone measurements collected for the specific event. The model achieves the incorporation of smartphone measurements without relying on conversion relations and allows for the quantification of the ShakeMap uncertainty. The second problem is solved by calibrating a regional GMM using the station measurements collected during past seismic events and using the amplification map as a site effect term.

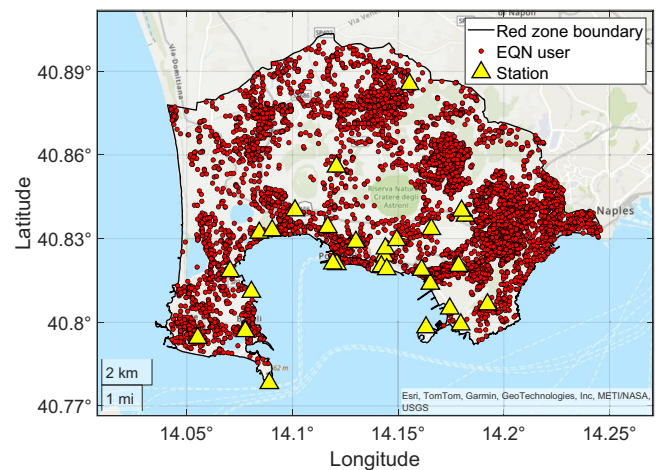
Here, we apply our method to the red zone<sup>19,20</sup> of Campi Flegrei, Italy, to produce a high-resolution amplification map of the region, to show the method's ability to generate high-resolution ShakeMaps after a seismic event, and to prove that our GMM outperforms existing regional GMMs based on VS30 information.

## Results

The red zone of Campi Flegrei is defined in the National Emergency Planning for Volcanic Risk in the Phlegraean area and was identified in the Decree of the President of the Council of Ministers of 24 June 2016<sup>1</sup>. The red zone, encompassing an area of ~130 km<sup>2</sup>, is characterised by elevated risk of pyroclastic flows and is inhabited by around 500,000 individuals. In the event of an emergency, the population is to be evacuated as a precautionary measure. The red zone is also the Campi Flegrei area most affected by local earthquakes, and the Italian area with the highest concentration of EQN participants in 2024. From April 2024 to June 2024, the number citizens in the red zone with the EQN app installed was between 7000 and 9000. The red zone is also monitored by the seismic stations of the Italian Seismic Network<sup>21</sup>, of the Italian Strong Motion Network<sup>22</sup> and of the Irpinia Seismic Network<sup>23</sup> (see Supplementary Fig. 1 and Supplementary Dataset 1). Figure 1 depicts the active seismic stations and active EQN smartphones on 8 June 2024. While the density of stations is relatively high with respect to the European density<sup>24</sup>, many highly populated areas of the red zone are only covered by smartphones. By integrating the accelerometric measurements from citizens' smartphones, we improve ShakeMaps right where people live.

### Seismic events for site amplification quantification

We considered four seismic events (detailed in Supplementary Dataset 2) that occurred between 27 April 2024 and 8 June 2024 during the Campi Flegrei seismic sequence started in 2023. These events were characterised by a similar coverage (in terms of network extension) of the red zone by stations and smartphones (see Fig. 2). Similar coverage is important to ensure that site amplification quantified at both the



**Fig. 1 | Station and smartphone networks.** Location of 29 seismological stations and 8151 citizens' smartphones participating in the Earthquake Network (EQN) citizen science initiative on 8 June 2024 in the red zone of Campi Flegrei. The EQN smartphone network is high density and covers areas of the red zone where no stations are installed. This allows the variability of the ground shaking to be captured at a higher spatial resolution over the entire red zone. Map was generated with MathWorks® MATLAB® Mapping Toolbox version 25.1 (R2025a) with map layer World Topographic Map<sup>41</sup> of Esri®.

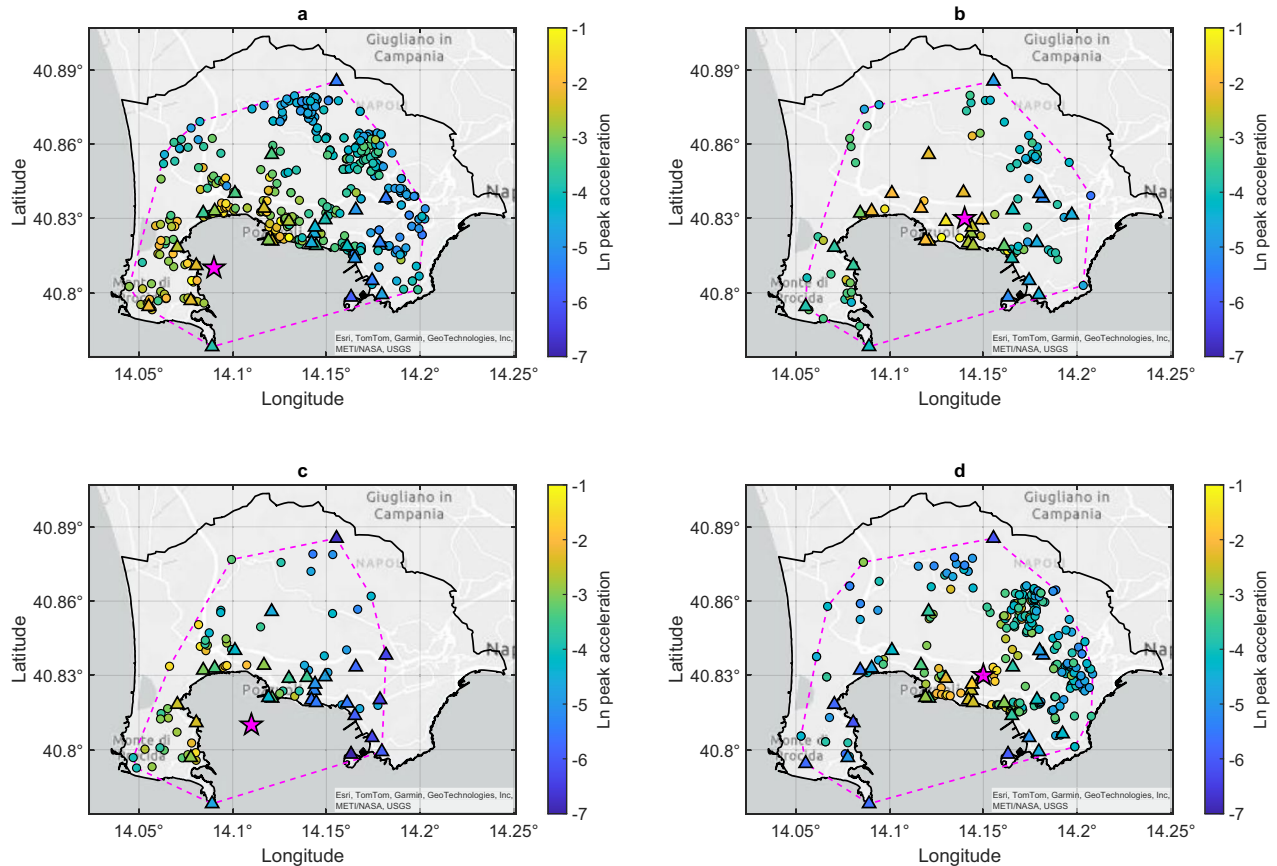
smartphones and the stations has the same reference (average amplification of the covered area). In addition, the events ranged in duration magnitude ( $M_d$ ) from 3.6 to 3.9. The relatively low magnitudes allow the assumption of radial decay without directivity effects. The number of stations in the red zone that measured each event ranged from 27 to 29, while the number of smartphones ranged from 56 to 441. The reason for this high variability is that only smartphones that are being charged perform seismic monitoring, and many more smartphones are charged at night. For stations, we considered PGA measurements, while for smartphones the peak smartphone acceleration (PSmA) as defined in the Methods section. Measurements collected during the four events are shown in Fig. 2.

### Bias between smartphone and seismological stations records

Our method is motivated by the fact that stations and smartphones measure different physical quantities. Stations directly measure the ground acceleration, while the acceleration measured by smartphones is linked to the ground acceleration through a transfer function that depends on the characteristics of the building in which the smartphone is located, the exact location of the smartphone within the building, and the characteristics of the object over which the smartphone is placed. Evaluating which factors are precisely controlling this bias between smartphone and seismological stations records is challenging. Moreover, when a citizen joins the EQN initiative, all of the above information is not available and does not have to be provided by the citizen to protect his/her privacy.

In the context of a citizen science initiative, the fact that we do not have access to citizens' smartphones, and that smartphones and stations are never co-located, makes it difficult to experimentally assess the bias of smartphone measurements. To overcome this problem, we adopted a statistical modelling approach. For each of the four seismic events, we fitted a spatial statistical model (Eq. 1 in the Methods section) on PSmA measurements. Then, we used the model to predict the PSmA at the station locations, where PGAs are available. The spatial model averages the building and smartphone-specific characteristics, allowing the average PGA-PSmA bias function to emerge.

Using linear regression, we assessed the relationship between the station PGAs and the predicted/filtered PSmA at the station locations. The regression lines depicted in Fig. 3, none of which are parallel to the



**Fig. 2 | Station and smartphone peak acceleration measurements.** The figure shows the station peak ground acceleration measurements (triangles) and the peak smartphone acceleration measurements (circles) for the four earthquakes used to derive the amplification map of Fig. 5 (accelerations are on natural logarithmic scale). **a** Md3.9, 27 April 2024 03:44:56 UTC (INGV ID 38381891). **b** Md3.9, 20 May 2024 19:46:14 UTC (INGV ID 38762031). **c** Md3.6, 22 May 2024 06:28:00 UTC (INGV

ID 38797691). **d** M3.7, 8 June 2023 02:09:03 UTC (INGV ID 39089101). The star markers are the earthquake epicentres, while the dotted lines are the convex hulls of the joint station and smartphone network. Maps were generated with MathWorks® MATLAB® Mapping Toolbox version 25.1 (R2025a) with map layer Light Gray Canvas Base<sup>42</sup> of Esri®.

bisector line, indicate that the average bias is amplitude dependent. On average, peak smartphone accelerations (PSmAs) are about 1.67 times higher than PGAs when PGA is 0.1g, and about 3 times higher when PGA is 0.01g. This bias prevents the direct integration of smartphone measurements into ShakeMaps, justifying the alternative strategy outlined in the following section.

### High resolution ground-motion amplification map derived from the data fusion of station and smartphone measurements

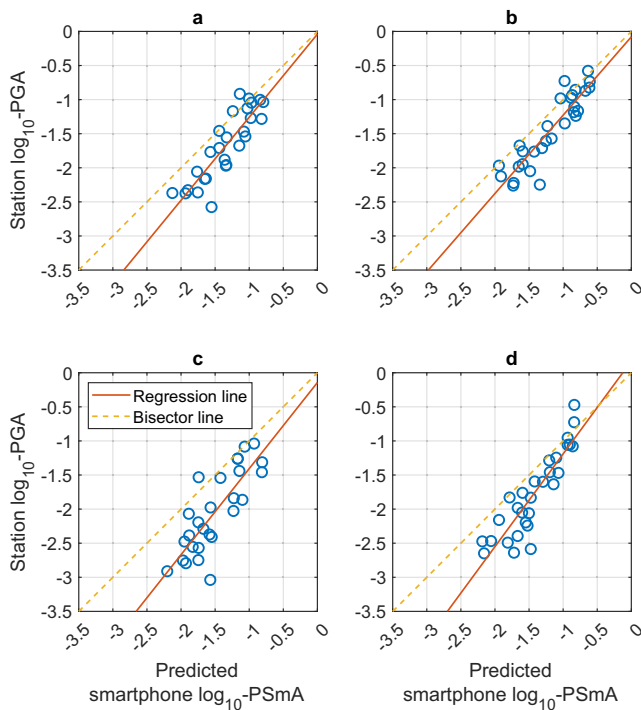
Although PSmAs exhibit a bias with respect to PGAs, we contend that the information from smartphones can be incorporated into ShakeMaps indirectly by modelling the spatial cross-correlation between the two types of measurements. Specifically, for each seismic event, we model the cross-correlation between PGAs and PSmAs residuals after an event-specific isotropic decay has been independently removed (see Supplementary Fig. 2). Station and smartphone residuals when considering the four events together are shown in Fig. 4. Smartphone residuals show higher variability at the small spatial scale because they are affected by both site amplification and smartphone- and building-specific characteristics. Station residuals depend only on the site amplification.

To demonstrate that station and smartphone residuals are spatially cross-correlated, we used the models of Eqs. (1) and (2) to derive a station-only-based log-amplification map and a smartphone-only-based map. Supplementary Fig. 3 depicts both maps, along with the corresponding standard deviation maps and the maps of the areas where the log-amplification is significantly positive or negative

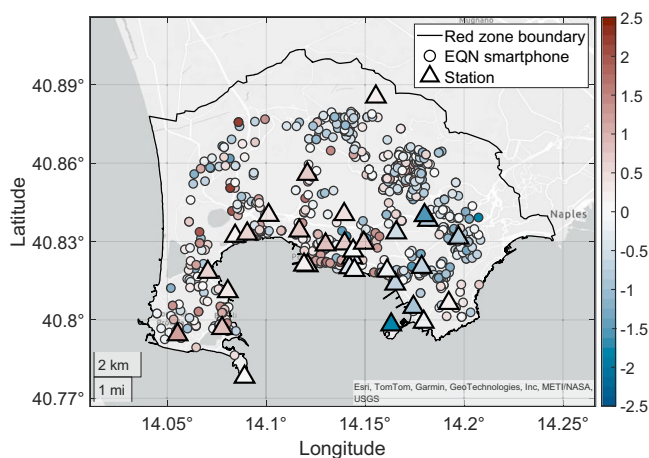
considering a 95% confidence interval (Eq. 10). The pixel-by-pixel spatial correlation between the two maps is 0.84. This result is significant as it demonstrates that station and smartphone residuals (when derived and modelled independently) capture similarly the local site amplification. The area where the station-only-based log-amplification map is significantly positive or negative corresponds to around 38.9 km<sup>2</sup> (29.8% of the red zone area) while the area for the smartphone-only-based map is around 64.5 km<sup>2</sup> (49.0%).

We then used the model in Eq. (11) to obtain the high-resolution log-amplification map based on the data fusion of station and smartphone measurements. The back-transformed (exponential of the log-amplification) map is depicted in Fig. 5, while the uncertainty map of the log-amplification is shown in Supplementary Fig. 4. We stress that the uncertainty map is not a secondary product as it enables the identification of areas where the amplification map can be relied upon, and of areas where the amplification significantly differs from the average amplification of the region. The level of uncertainty depends on the number of stations and smartphone measurements, their spatial distribution across the red zone, the number of seismic events used to derive the amplification map, and the errors and biases in the smartphone measurements.

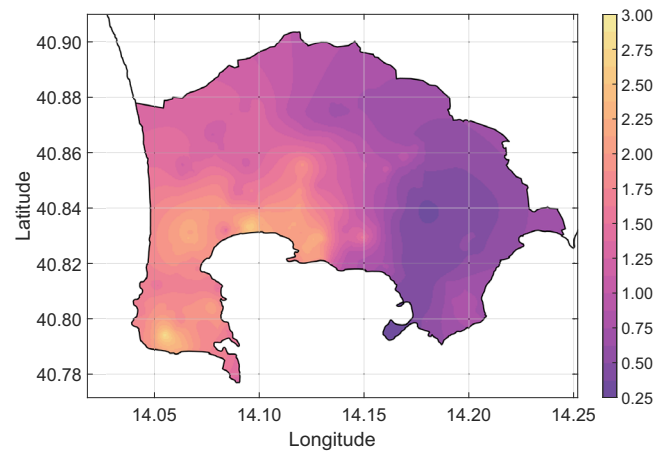
Using the data fusion model, the area where the log-amplification is significantly positive or negative is 79.0 km<sup>2</sup> (61.0% of the red zone area). Combining station and smartphone data increased the area by more than double compared to the map based on station data alone. Additionally, we assessed the power of the estimated log-amplification map in reducing the station residual variability using Eqs. (12) and (13).



**Fig. 3 | Amplitude-dependent bias between smartphone and station measurements.** Comparison between peak ground accelerations (PGAs) and peak smartphone accelerations (PSmAs) at the spatial locations of the stations. PGAs are those measured by the stations while the PSmAs are those observed by the smartphones and predicted at the station locations using the spatial statistical model in Eq. (1). The regression lines (detailed in Eq. 7), describe the relationship between predicted PSmAs and PGAs. **a** Md3.9, 27 April 2024 03:44:56 UTC (INGV ID 38381891). Regression parameters:  $\hat{\alpha}_0 = -0.04$ ,  $\hat{\alpha}_1 = 1.22$ . **b** Md3.9, 20 May 2024 19:46:14 UTC (INGV ID 38762031). Regression parameters:  $\hat{\alpha}_0 = -0.08$ ,  $\hat{\alpha}_1 = 1.15$ . **c** Md3.6, 22 May 2024 06:28:00 UTC (INGV ID 38797691). Regression parameters:  $\hat{\alpha}_0 = -0.14$ ,  $\hat{\alpha}_1 = 1.26$ . **d** M3.7, 8 June 2023 02:09:03 UTC (INGV ID 39089101). Regression parameters:  $\hat{\alpha}_0 = 0.17$ ,  $\hat{\alpha}_1 = 1.36$ .



**Fig. 4 | Residuals of station peak ground acceleration (PGA) and peak smartphone acceleration (PSmA), in natural logarithmic scale, obtained by fitting an event-specific isotropic decay model centred at each earthquake epicentre.** This figure shows the average station residuals and the individual smartphone residuals when considering the four events detailed in Fig. 2 together. In the eastern part of the region, negative residuals suggest comparatively lower amplification. Map was generated with MathWorks® MATLAB® Mapping Toolbox version 25.1 (R2025a) with map layer Light Gray Canvas Base<sup>42</sup> of Esri®.



**Fig. 5 | High-resolution site amplification map of the red zone of Campi Flegrei.** Each point of the map represents the local amplification factor needed to correct the peak ground acceleration estimated by an isotropic decay model centred on the earthquake epicentre. The map was generated with MathWorks® MATLAB® Mapping Toolbox version 25.1 (R2025a).

The standard deviation of the average station residual once the isotropic decay is removed is 0.6389 (0.2775 on the  $\log_{10}$  scale). Correcting the residual by the log-amplification map at the station decreased the standard deviation to 0.1827 (0.0793 on the  $\log_{10}$  scale), corresponding to a 71.4% reduction.

The log-amplification map and its uncertainty are also provided in Supplementary Dataset 3. The estimated model parameters for the four events are given in Supplementary Table 1.

### Exposure to site amplification

Considering the high-resolution amplification map of Fig. 5, and the population and building exact locations from the European buildings dataset<sup>25</sup>, we estimated the exposure distribution to site amplification for 498,817 people, for 8151 EQN users and for 334,837 buildings in the red zone. We found that all distributions (see Supplementary Fig. 5) are skewed toward amplifications lower than 1. This implies that people and buildings tend to be located in areas of the red zone where the PGA is deamplified rather than amplified (relative to the PGA estimated by an isotropic decay centred on the epicentre). About 19.4% of the population is exposed to an amplification factor of more than 1.5 while 5.8% to a factor of more than 2. About 24% of the buildings are exposed to an amplification factor of more than 1.5 and about 5.4% to a factor of more than 2. The EQN users and the population have a very similar exposure distribution, which means that the EQN users were a representative sample of the population (in terms of site amplification exposure).

### High-resolution PGA ShakeMap integrating smartphone measurements

The high-resolution amplification map issued from the fusion of station and smartphone data was used to produce a high-resolution PGA ShakeMap for the Md4.2 earthquake that occurred on 27 September 2023 with epicentre in the red zone (event detail in Supplementary Dataset 2). This was done using the model in Eq. (14). The event is the largest (in terms of magnitude) of the ongoing 2023–2025 Campi Flegrei sequence that was detected by at least 100 smartphones of the EQN initiative. Figure 6 depicts the PGA ShakeMap and its uncertainty derived using Eqs. (16) and (17), respectively. To better appreciate its high spatial resolution, the ShakeMap is restricted to an area of the red zone where the estimated PGA is significantly higher than zero. For a comparison, Supplementary Fig. 6 shows the ShakeMap provided by the National Institute of Geophysics and Volcanology (INGV). The PGA decay of the INGV map is nearly isotropic. On the contrary, Eq. (14)

produced a ShakeMap that exhibits small-scale variability thanks to the amplification map and thanks to the event-specific smartphone measurements.

### Enhancement of regional ground motion models

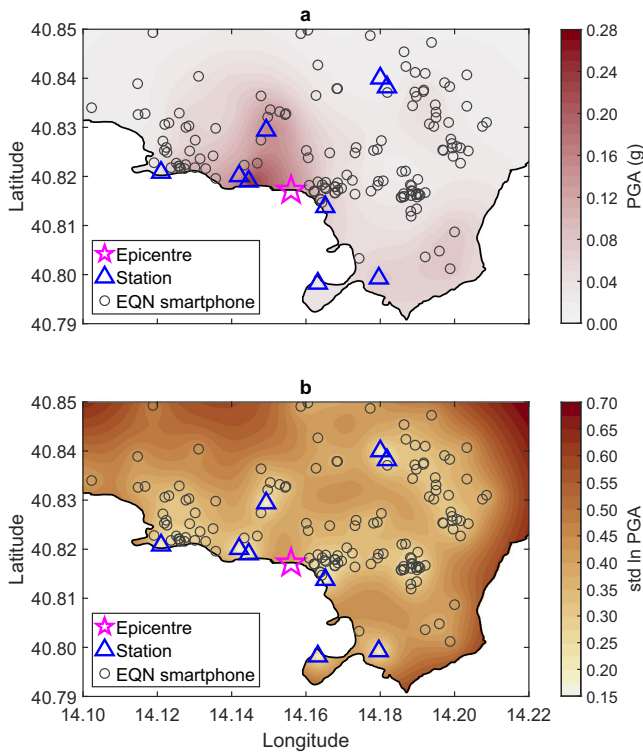
As a further relevant application, we used the high-resolution amplification map to improve the recently published regional GMM<sup>26</sup> with site classes based on VS30<sup>27</sup>. We considered a total of 661 station PGA measurements from 34 seismic events that occurred between 5 February 2023 and 18 June 2024, with an epicentre in the red zone and a magnitude between Md2.9 and Md4.4 (details in Supplementary Dataset 4). We then compared three GMMs: the regional GMM<sup>26</sup>, the same GMM recalibrated on the 34 events, and our GMM of Eq. (19) based on the amplification map in Fig. 5 and calibrated on the same events. The standard deviation of  $\log_{10}$ -PGA residuals is around 0.410 for the original GMM, around 0.397 for the recalibrated GMM and

around 0.357 for our GMM. This is a 13% and 10% reduction, respectively, which is significant given the high impact of random variability in GMMs on probabilistic seismic hazard assessments<sup>28</sup>. Additionally, the reduction in the standard deviation of the residuals is a measure of the GMM enhancement and an implicit validation of the amplification map. The estimated GMM coefficients are given in Table 1.

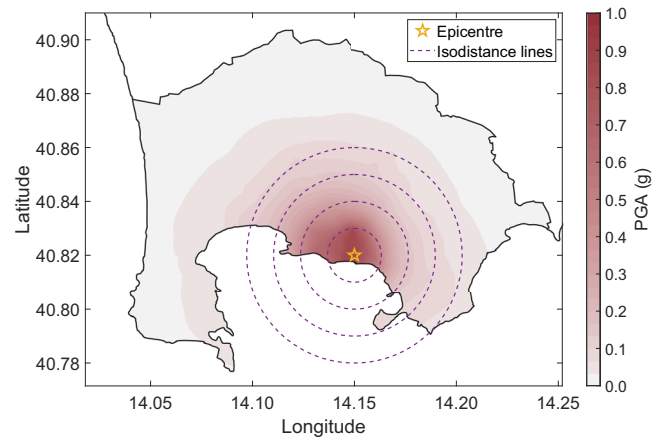
We then used our calibrated GMM to simulate the PGA ShakeMap assuming a Md5.0 event with epicentre in 40.82° latitude and 14.15° longitude. This is the same epicentre of the Md4.6 event of 13 March 2025. The ShakeMap, which is depicted in Fig. 7, retains the high spatial resolution and the spatial patterns of the amplification map. The PGA ShakeMaps obtained from the other GMMs are depicted in Supplementary Figs. 7 and 8, and they exhibit a radial decay with interaction of the VS30 site classes.

### Discussion

In the last decade, since the proliferation of smartphone technology and the ambition of adopting it for seismic monitoring, scientists have focused on comparing the monitoring capabilities of smartphones with scientific-grade accelerometers. This has been done through laboratory experiments, with smartphones and accelerometers installed on the same shake table<sup>29</sup> or, in the context of building structural health monitoring, with devices installed in the same building<sup>30,31</sup>. While these results have demonstrated that smartphones are suitable for seismic monitoring, they have very limited applicability when it comes to modelling data from crowdsourcing initiatives, with smartphones that are not fixed to the ground, that can be placed



**Fig. 6 | High-resolution peak ground acceleration (PGA) ShakeMap and its uncertainty for the Md4.2 event of 27 September 2023 (INGV ID 36299321).** **a** When both the smartphone measurements collected during the event and the site amplification information (Fig. 5) are considered, the spatial distribution of the PGA shows spatial variations not predicted by a conventional isotropic model (compare with Supplementary Fig. 6). Such ShakeMap is generated without the need to determine the earthquake’s magnitude or location and can be produced in real time, provided the data are accessible. **b** The ShakeMap uncertainty is described by the standard deviation of the predicted PGA on natural logarithmic scale. Maps were generated with MathWorks® MATLAB® Mapping Toolbox version 25.1 (R2025a).



**Fig. 7 | Simulated peak ground acceleration ShakeMap for a potential Md5.0 event with epicentre in the red zone of Campi Flegrei.** The ShakeMap is generated using the ground motion model in Eq. (19), with model parameters given in the last row of Table 1. Unlike the ShakeMap in Fig. 6, which is obtained for an event that has occurred and relies on station and smartphone measurements, this ShakeMap is derived for a potential future event, with the earthquake magnitude incorporated in Eq. (19). The ShakeMap is therefore suitable for earthquake scenario modelling and seismic hazard assessment. Isodistance lines with respect to the epicentre help to appreciate the ShakeMap anisotropy, which reflects higher amplification in the western part of the region.

**Table 1 | Comparison of ground motion models in terms of parameters and standard deviation of the residual PGA**

GMM type	GMM parameters							
	a	b	c	c2	c3	h	e	$\sigma$
lervolino et al. (2024)	-0.3306	0.6515	-2.8249	0.3362	-0.0128	0.7862	0.3492	0.4107
lervolino et al. (2024) recalibrated	-1.6075	0.7429	-4.8769	0.0853	0.1600	2.2873	0.0881	0.3968
Amplification-based	-2.8051	0.8479	-1.4008	-0.1332	0.0378	1.5625	0.6457	0.3569

The lervolino et al. (2024)<sup>26</sup> model and its recalibrated version include a site effect term based on VS30, while the amplification-based model includes a site effect term based on the site amplification map of Fig. 5. To be consistent with lervolino et al. (2024)<sup>26</sup>,  $\sigma$  is evaluated considering peak ground accelerations in the  $\log_{10}$  scale (Eqs. 18 and 19).

anywhere and on any floor within the building, and that provide measurements that cannot be calibrated against some reference measurements. As a result, no solutions have been available to exploit the high spatial density of crowdsourced smartphone networks in the generation of ShakeMaps.

The method we have developed provides a solution that does not require to directly model the bias in smartphone measurements relative to seismological station measurements, and that relies on the consistent spatial cross-correlation between smartphone and station measurements rather than on the absolute accelerations recorded by the smartphones. The method thus achieves data fusion without requiring any calibration between smartphones and stations. Another key feature of the method is its ability to iteratively learn the site amplification from subsequent seismic events that have affected the same region. Since the EQN smartphone network is dynamic in the location of the smartphones, the information from multiple seismic events allows us to derive a high spatial resolution amplification map across the whole region. This high-resolution information is then used to derive high-resolution PGA ShakeMaps for any new seismic event, supplemented by the information coming from station and smartphone measurements collected during that event. The amplification map issued from the fusion of seismological stations and smartphone records is also used to recalibrate conventional GMMs and reduce the random variability of seismic motion prediction.

The ability to obtain high-resolution, high-quality amplification and ShakeMap maps depends on the number of seismic stations and smartphones, as well as their spatial coverage of the region of interest. The results obtained in this paper for the red zone of Campi Flegrei can only be extrapolated to other regions if they are covered by a similar density of stations and smartphones. One advantage of our method, however, is its ability to assess the uncertainty and reliability of all model outputs. This enables us, on the one hand, to avoid overselling the model outputs and, on the other, to understand the impact of the uneven spatial distribution of stations and smartphones.

A seismic amplification map depends on the reference chosen. In our approach, site amplification and deamplification are relative to the average site amplification in the region. This differs from some conventional amplification factors used in engineering seismology, where the site amplification is relative to a rock reference, but is very similar to amplification maps recently derived at the regional scale for European risk models<sup>7</sup>. A fine-scale interpretation of the amplification map may be unreliable due to local uncertainty and the map can only be compared with those of other regions if it has been verified that the average amplification effects due to subsurface layers are similar in both regions. Furthermore, the amplification values in our map can be rescaled relative to a specific station or subset of stations (provided these are included in the calibration dataset), thereby enabling more tailored interpretations. This amplification map is derived from moderate-magnitude earthquake recordings and therefore represents the linear response of the ground.

The analysis performed in this paper uses accelerations from stations and smartphones. To compute the PGA ShakeMaps, we assume that the spatial correlation of the station PGA is the same as the spatial correlation of the peak smartphone acceleration, which is only true if the two types of sensors measure the acceleration in the same ground motion frequency band. Tests on smartphone accelerometers<sup>29</sup> have shown that smartphone records match short-period spectral accelerations and that high-end accelerometers are capable of detecting motion in the 2–30 Hz frequency range, which is also the common frequency band that controls the PGA of seismological network accelerometers. However, these tests indicate that smartphone data cannot be used in combination with seismological stations to predict ground motion parameters that depend on frequencies lower than 2 Hz (e.g. peak ground velocity).

By improving the spatial resolution of PGA ShakeMaps in urban areas, where, by definition, exposure is greatest, the operational application of our approach has the potential to improve hazard assessment and reduce uncertainties in rapid impact estimates. It also confirms that crowdsourced data are a valuable and affordable complement to real-time accelerometer networks for operational seismology<sup>32,33</sup>.

## Methods

### Station and smartphone measurements

Amplification map and ShakeMaps are based on the joint analysis of station PGA measurements and PSmA measurements. PGA measurements are routinely provided by seismological agencies. For all the seismic events considered in this work, we used the ShakeMap published by INGV on their ShakeMap website ([shakemap.ingv.it](http://shakemap.ingv.it)). PSmA measurements are instead collected and provided by the EQN initiative. Citizens join the initiative by installing the EQN application on their smartphones. Seismic monitoring starts when the smartphone is charging and not in use. The application reads the smartphone's accelerometer output in real time and calculates the resulting acceleration from the usual three acceleration components. As EQN's main service is earthquake early warning, the app calculates the maximum resultant acceleration (MRA) over windows of 3 s. If the MRA is above a certain threshold (which depends on the sensor noise at rest), the information is sent to the EQN server. For a given earthquake, each smartphone collects one or more MRAs, depending on the duration of the shaking. The PSmA associated with a smartphone is the maximum of the MRAs collected during the seismic event.

### Smartphone and station data modelling

The method for the analysis of smartphone and station data is based on spatial statistical models with Gaussian latent variables. For the generic earthquake  $E$  and site  $S$  we model PSmA measurements as

$$\ln[Y_{\text{PSmA},ES}] = f_{\text{PSmA},E}(R_{ES}^h) + \delta W_{\text{PSmA},ES} + \delta \text{Inst}_{\text{PSmA},ES}, \quad (1)$$

while the PGA measurements as

$$\ln[Y_{\text{PGA},ES}] = f_{\text{PGA},E}(R_{ES}^h) + \delta W_{\text{PGA},ES} + \delta \text{Inst}_{\text{PGA},ES}. \quad (2)$$

The functions

$$\begin{aligned} f_{\text{PSmA},E} &= \beta_{0,\text{PSmA},E} + \beta_{1,\text{PSmA},E} R_{ES}^h, \\ f_{\text{PGA},E} &= \beta_{0,\text{PGA},E} + \beta_{1,\text{PGA},E} R_{ES}^h, \end{aligned} \quad (3)$$

describe a deterministic event-specific isotropic decays with respect to the hypocentral distance  $R_{ES}^h$ . Terms  $\delta W_{\text{PSmA},ES}$  and  $\delta W_{\text{PGA},ES}$  represent on the other hand the site and event specific variability of ground motion that is not explained by the isotropic decay. Assuming that site amplification varies smoothly across space, they are modelled as univariate zero-mean spatial Gaussian processes. A spatial Gaussian processes is continuous in space (i.e. the process is defined at any point in space and not just at the locations of the smartphone or station) and is also spatially correlated. This means that the process measured at two different locations tends to assume similar value if the distance between the locations is small, and different values if the distance is large. The spatial correlation of  $\delta W_{\text{PSmA},ES}$  and  $\delta W_{\text{PGA},ES}$  is determined by the spatial covariance functions  $v_{\text{PSmA},E} \exp(-\frac{D_{S,S'}}{\theta_{\text{PSmA},E}})$  and  $v_{\text{PGA},E} \exp(-\frac{D_{S,S'}}{\theta_{\text{PGA},E}})$ , respectively, where  $D_{S,S'}$  is the geodetic distance between any two sites  $S$  and  $S'$ ,  $v_{\text{PSmA},E}$  and  $v_{\text{PGA},E}$  are scale parameters while  $\theta_{\text{PSmA},E}$  and  $\theta_{\text{PGA},E}$  define the strength of the spatial correlation (the higher  $\theta$  the stronger the spatial correlation).

Model terms  $\delta W_{\text{PSmA},ES}$  and  $\delta W_{\text{PGA},ES}$  thus describe the event-dependent site amplification which emerges as a residual from the isotropic decay, and which is consistently correlated in space. Without these terms, site amplification would only be quantifiable at the station and smartphone locations, and the residual variability captured by  $\delta W_{\text{PSmA},ES}$  and  $\delta W_{\text{PGA},ES}$  would instead contribute to increase the uncertainty modelled by the random errors  $\delta \text{Inst}_{\text{PSmA},ES} \sim N(0, \phi_{\text{PSmA},E}^2)$  and  $\delta \text{Inst}_{\text{PGA},ES} \sim N(0, \phi_{\text{PGA},E}^2)$ . These errors are spatially uncorrelated, capturing the extra variability not explained by either the isotropic decay or the site amplification.

In general,  $\phi_{\text{PSmA},E}^2 \gg \phi_{\text{PGA},E}^2$  since stations are expected to measure the true ground shaking while the smartphone error is affected by many unobserved factors that contribute to a noisy and biased smartphone measurement (e.g. the location of the smartphone within the building, the object above which the smartphone is located and the smartphone sensor type).

### Model estimation

For each event  $E$ , the parameter sets of the two models to be estimated are  $\Psi_{\text{PSmA},E} = \{\beta_{0,\text{PSmA},E}, \beta_{1,\text{PSmA},E}, \phi_{\text{PSmA},E}^2, \nu_{\text{PSmA},E}, \theta_{\text{PSmA},E}\}$  and  $\Psi_{\text{PGA},E} = \{\beta_{0,\text{PGA},E}, \beta_{1,\text{PGA},E}, \phi_{\text{PGA},E}^2, \nu_{\text{PGA},E}, \theta_{\text{PGA},E}\}$ . Let  $\mathcal{D}_{\text{PSmA},E}$  and  $\mathcal{D}_{\text{PGA},E}$  be the data set of all the PSmA and PGA measurements for event  $E$ , respectively. Model estimation is based on the maximum-likelihood approach and is carried out adopting the expectation-maximization (EM) algorithm<sup>34</sup> implemented by the D-STEM software<sup>35</sup>.

Outputs of the EM algorithm are the estimated parameter sets  $\hat{\Psi}_{\text{PSmA},E}$  and  $\hat{\Psi}_{\text{PGA},E}$ , as well as

$$\begin{aligned} \widehat{\delta W}_{\text{PSmA},ES} &= \mathbb{E}_{\hat{\Psi}_{\text{PSmA},E}}(\delta W_{\text{PSmA},ES} | \mathcal{D}_{\text{PSmA},E}), \\ \widehat{\delta W}_{\text{PGA},ES} &= \mathbb{E}_{\hat{\Psi}_{\text{PGA},E}}(\delta W_{\text{PGA},ES} | \mathcal{D}_{\text{PGA},E}), \end{aligned} \tag{4}$$

namely the conditional expectations of  $\delta W_{\text{PSmA},ES}$  and  $\delta W_{\text{PGA},ES}$  given the measurements in  $\mathcal{D}_{\text{PSmA},E}$  and  $\mathcal{D}_{\text{PGA},E}$ . Uncertainties of  $\widehat{\delta W}_{\text{PSmA},ES}$  and  $\widehat{\delta W}_{\text{PGA},ES}$  are given by the conditional variances

$$\begin{aligned} \text{var}(\widehat{\delta W}_{\text{PSmA},ES}) &= \mathbb{V}_{\hat{\Psi}_{\text{PSmA},E}}(\delta W_{\text{PSmA},ES} | \mathcal{D}_{\text{PSmA},E}), \\ \text{var}(\widehat{\delta W}_{\text{PGA},ES}) &= \mathbb{V}_{\hat{\Psi}_{\text{PGA},E}}(\delta W_{\text{PGA},ES} | \mathcal{D}_{\text{PGA},E}). \end{aligned} \tag{5}$$

The computing of  $\mathbb{E}_{\hat{\Psi}}$  and  $\mathbb{V}_{\hat{\Psi}}$  is based on standard conditional multivariate normal distribution formulas<sup>34</sup>.

Due to the spatial correlation induced by the Gaussian process,  $\widehat{\delta W}_{\text{PSmA},ES}$  is a smoothed/filtered version of the raw event-dependent site amplification (residual from the isotropic decay) measured by the smartphone. In particular, the estimated  $\widehat{\delta W}_{\text{PSmA},ES}$  at site  $S$  depends on all the smartphone measurement (and not only on the smartphone measurement at  $S$ ). When the variance  $\phi_{\text{PSmA},E}^2$  of  $\delta \text{Inst}_{\text{PSmA},ES}$  is high,  $\widehat{\delta W}_{\text{PSmA},ES}$  tends to be close to zero (which is the marginal mean of  $\delta W_{\text{PSmA},ES}$ ) and the uncertainty  $\text{var}(\widehat{\delta W}_{\text{PSmA},ES})$  tends to be large. On the other hand, when  $\phi_{\text{PSmA},E}^2$  is close to zero,  $\widehat{\delta W}_{\text{PSmA},ES}$  tends to be equal to the raw site amplification and  $\text{var}(\widehat{\delta W}_{\text{PSmA},ES})$  tends to be close to zero. The same considerations apply to  $\widehat{\delta W}_{\text{PGA},ES}$ .

Additionally, because  $\delta W_{\text{PSmA},ES}$  and  $\delta W_{\text{PGA},ES}$  are modelled as spatially continuous Gaussian processes,  $\widehat{\delta W}_{\text{PSmA},ES}$  and  $\widehat{\delta W}_{\text{PGA},ES}$  are available at any location in space once the models in Eqs. (1) and (2) have been estimated.

### Smartphone measurement bias assessment

For a given earthquake, the model in Eq. (1) is estimated using the PSmA measurements in  $\mathcal{D}_{\text{PSmA},E}$ . The estimated model is then used to

predict the PSmA at the spatial locations of the stations

$$\ln[\widehat{Y}_{\text{PSmA},ES_i}] = \widehat{\beta}_{0,\text{PSmA},E} + \widehat{\beta}_{1,\text{PSmA},E} R_{ES_i}^h + \widehat{\delta W}_{\text{PSmA},ES_i}, \tag{6}$$

where  $S_i$  is the location of the  $i$ th station. The event-specific bias function is assessed using the regression model

$$\log_{10}[Y_{\text{PGA},ES_i}] = \alpha_0 + \alpha_1 \log_{10}[\widehat{Y}_{\text{PSmA},ES_i}] + \varepsilon_i, \tag{7}$$

where  $Y_{\text{PGA},ES_i}$  is the PGA measured at the  $i$ th station,  $\alpha_0$  and  $\alpha_1$  are the model parameters while  $\varepsilon_i \sim N(0, \sigma_\varepsilon^2)$ . When  $\widehat{\alpha}_1 = 1$  and  $\widehat{\alpha}_0 \neq 0$ , the bias is purely multiplicative and  $\frac{Y_{\text{PGA},ES_i}}{\widehat{Y}_{\text{PSmA},ES_i}} = 10^{\widehat{\alpha}_0}$ . On the other hand, if  $\widehat{\alpha}_1 \neq 1$  then the bias is amplitude dependent.

### Site amplification and cross-correlation

The  $\delta W_{\text{PSmA},ES}$  and  $\delta W_{\text{PGA},ES}$  site amplifications obtained in Eq. (4) and their uncertainties in Eq. (5) are event-specific and are affected by the event-specific station and smartphone network geometry. By cumulating the information from multiple events, we aim to obtain the event-independent site amplification. To do this, we consider  $N$  seismic events that have affected the red zone of Campi Flegrei and for which  $\delta W_{\text{PSmA},ES}$  and  $\delta W_{\text{PGA},ES}$  have been estimated using Eq. (4). The estimated site-term event-independent log-amplification is given by

$$\begin{aligned} \widehat{\delta S2S}_{\text{PSmA},S} &= \frac{\sum_{E=1}^N \widehat{\delta W}_{\text{PSmA},ES} \omega_{\text{PSmA},ES}}{\sum_{E=1}^N \omega_{\text{PSmA},ES}}, \\ \widehat{\delta S2S}_{\text{PGA},S} &= \frac{\sum_{E=1}^N \widehat{\delta W}_{\text{PGA},ES} \omega_{\text{PGA},ES}}{\sum_{E=1}^N \omega_{\text{PGA},ES}}, \end{aligned} \tag{8}$$

where  $\omega_{\text{PSmA},ES} = \frac{1}{\text{var}(\widehat{\delta W}_{\text{PSmA},ES})}$  and  $\omega_{\text{PGA},ES} = \frac{1}{\text{var}(\widehat{\delta W}_{\text{PGA},ES})}$  are weights based on the variance. The standard deviations of  $\widehat{\delta S2S}_{\text{PSmA},S}$  and  $\widehat{\delta S2S}_{\text{PGA},S}$  are given by

$$\begin{aligned} \text{std}(\widehat{\delta S2S}_{\text{PSmA},S}) &= \frac{1}{\sum_{E=1}^N \omega_{\text{PSmA},ES}} \sqrt{\sum_{E=1}^N (\omega_{\text{PSmA},ES} \cdot \text{std}(\widehat{\delta W}_{\text{PSmA},ES}))^2}, \\ \text{std}(\widehat{\delta S2S}_{\text{PGA},S}) &= \frac{1}{\sum_{E=1}^N \omega_{\text{PGA},ES}} \sqrt{\sum_{E=1}^N (\omega_{\text{PGA},ES} \cdot \text{std}(\widehat{\delta W}_{\text{PGA},ES}))^2}. \end{aligned} \tag{9}$$

Standard deviations in Eq. (9) represent the uncertainties of  $\widehat{\delta S2S}_{\text{PSmA},S}$  and  $\widehat{\delta S2S}_{\text{PGA},S}$ , respectively. As customary in statistical inference, standard deviations related to estimates are used to compute confidence intervals. Specifically, the 95% confidence interval for  $\widehat{\delta S2S}_{\text{PSmA},S}$  and  $\widehat{\delta S2S}_{\text{PGA},S}$  are given by

$$\begin{aligned} & \left[ \widehat{\delta S2S}_{\text{PSmA},S} - 1.96 \cdot \text{std}(\widehat{\delta S2S}_{\text{PSmA},S}), \widehat{\delta S2S}_{\text{PSmA},S} + 1.96 \cdot \text{std}(\widehat{\delta S2S}_{\text{PSmA},S}) \right] \\ & \left[ \widehat{\delta S2S}_{\text{PGA},S} - 1.96 \cdot \text{std}(\widehat{\delta S2S}_{\text{PGA},S}), \widehat{\delta S2S}_{\text{PGA},S} + 1.96 \cdot \text{std}(\widehat{\delta S2S}_{\text{PGA},S}) \right] \end{aligned} \tag{10}$$

As depicted in Supplementary Figs. 3 and 4, the above confidence intervals are for example used to determine whether  $\widehat{\delta S2S}_{\text{PSmA},S}$  and  $\widehat{\delta S2S}_{\text{PGA},S}$  are significantly different from zero at the generic location  $S$ . If the zero value is included in the confidence interval, then it cannot be claimed that the event-independent log-amplification at  $S$  is either higher or lower than the average log-amplification of the red zone.

As a measure of statistical association between the log-amplifications given by Eq. (8), both quantities are evaluated over a regular grid covering the red zone, and the spatial correlation  $\text{corr}(\widehat{\delta S2S}_{\text{PSmA},S}, \widehat{\delta S2S}_{\text{PGA},S})$  is then computed. High correlation implies that stations and smartphones measurements are capturing

the same underlying site amplification spatial pattern, and that the joint modelling of both measurement types may be beneficial.

### Statistical data fusion and calibration model

Models in Eqs. (1) and (2) separately describe smartphone and station measurements. Log-amplifications  $\delta S2S_{PSmA,S}$  and  $\delta S2S_{PGA,S}$  are thus estimated independently, they are characterized by a different uncertainty and cannot be easily and naturally combined into a single log-amplification value. To overcome this limit, the following bi-variate model is considered

$$\begin{aligned} \ln[Y_{PSmA,ES}] &= f_{PSmA,E}(R_{ES}^h) + \delta W_{PSmA,ES} + \delta Inst_{PSmA,ES} \\ \ln[Y_{PGA,ES}] &= f_{PGA,E}(R_{ES}^h) + \delta W_{PGA,ES} + \delta Inst_{PGA,ES} \end{aligned} \quad (11)$$

Model in Eq. (11) differs from models in Eqs. (1) and (2) because  $\delta W_{PSmA,ES}$  and  $\delta W_{PGA,ES}$  are jointly modelled in the vector  $\delta W_{ES} = (\delta W_{PSmA,ES}, \delta W_{PGA,ES})$  as a bi-variate zero-mean spatial Gaussian process with matrix spatial covariance function given by  $V_E \exp(-\frac{D_{S,S}}{\theta})$ , with  $V_E$  a  $2 \times 2$  variance-covariance matrix. This implies that, net of the isotropic decay, there is an exchange of information between model terms  $\delta W_{PSmA,ES}$  and  $\delta W_{PGA,ES}$ . The higher the covariance in  $V_E$ , the more  $\delta W_{PGA,ES}$  is affected by the residual smartphone measurements  $\ln[Y_{PSmA,ES}] - f_{PSmA,E}(R_{ES}^h)$ .

The model parameter set becomes  $\Psi_E = \{\beta_{0,PSmA,E}, \beta_{1,PSmA,E}, \beta_{0,PGA,E}, \beta_{1,PGA,E}, \phi_{PSmA,E}^2, \phi_{PGA,E}^2, V_E, \theta\}$  and it is estimated using the EM algorithm.

Model in Eq. (11) is then adopted to produce  $\delta S2S_{PGA,S}$  defined in Eq. (8). The advantage of using the model in Eq. (11) lies in the fact that  $\delta S2S_{PGA,S}$  is now obtained using the information of both station and smartphone measurements. In addition, the model does not require the smartphone measurements to be calibrated against the station measurements. This aspect is crucial in the crowdsourced setting, where people's smartphones are not co-located with seismic stations.

To measure the efficacy of  $\delta S2S_{PGA,S}$  to reduce the residual variability at the stations, the following measures of site-to-site variability<sup>36</sup> are considered

$$std_{S_i} \left( \sum_{E=1}^N \ln[Y_{PGA,ES_i}] - \hat{f}_{PGA,E}(R_{ES_i}^h) \right), \quad (12)$$

$$std_{S_i} \left( \sum_{E=1}^N \ln[Y_{PGA,ES_i}] - \hat{f}_{PGA,E}(R_{ES_i}^h) - \delta S2S_{PGA,S_i} \right), \quad (13)$$

where  $std_{S_i}$  is the standard deviation computed across the stations. Equation (12) is based on the raw average station residual, where the residual is obtained by simply removing the fitted isotropic decay. The residual in Eq. (13) is instead a corrected version since it is obtained by also removing the event-independent site amplification. The decrease from the raw value to the corrected value is a measure of the quality of  $\delta S2S_{PGA,S}$ <sup>37</sup>.

### Enhanced statistical data fusion and calibration model

Once  $\delta S2S_{PGA,S}$  is available, an improved version of the model in Eq. (11) is

$$\begin{aligned} \ln[Y_{PSmA,ES}] &= f_{PSmA,E}(R_{ES}^h, \delta S2S_{PGA,S}) + \delta W_{PSmA,ES} + \delta Inst_{PSmA,ES} \\ \ln[Y_{PGA,ES}] &= f_{PGA,E}(R_{ES}^h, \delta S2S_{PGA,S}) + \delta W_{PGA,ES} + \delta Inst_{PGA,ES} \end{aligned} \quad (14)$$

where

$$\begin{aligned} f_{PSmA,E}(R_{ES}^h, \delta S2S_{PGA,S}) &= \beta_{0,PSmA,E} + \beta_{1,PSmA,E} R_{ES}^h + \beta_{2,PSmA,E} \delta S2S_{PGA,S} \\ f_{PGA,E}(R_{ES}^h, \delta S2S_{PGA,S}) &= \beta_{0,PGA,E} + \beta_{1,PGA,E} R_{ES}^h + \beta_{2,PGA,E} \delta S2S_{PGA,S} \end{aligned} \quad (15)$$

Terms  $\delta W_{PSmA,ES}$  and  $\delta W_{PGA,ES}$  describe the variability of station and smartphone measurements which is not explained neither by the isotropic decay nor by the site-specific amplification (e.g. directivity and non-linear effects).

Starting from event  $N+1$ , model in Eq. (14) exploits  $\delta S2S_{PGA,S}$  which is based on the previous  $N$  events. Such aggregation, earthquake by earthquake, of event and decay-corrected residuals to produce high-resolution amplification maps, is similar to the approach with classical seismological data<sup>38</sup>.

### Resulting ground shaking map and uncertainty

Once estimated, model in Eq. (14) is used to produce the spatial map of the PGA and its uncertainty for the event  $E$ . The estimated In-PGA at  $S$  is given by

$$\ln[\hat{Y}_{PGA,ES}] = \hat{f}_{PGA,E}(R_{ES}^h, \delta S2S_{PGA,S}) + \delta W_{PGA,ES}, \quad (16)$$

where  $\delta W_{PGA,ES} = \mathbb{E}_{\Psi_E}(\delta W_{PGA,ES} | \mathcal{D}_E)$ , with  $\mathcal{D}_E$  the data set of all PSmA and PGA measurements for event  $E$ . The uncertainty of  $\ln[\hat{Y}_{PGA,ES}]$  is given by

$$\text{var}(\ln[\hat{Y}_{PGA,ES}]) = \mathbb{V}_{\Psi_k}(\delta W_{PGA,ES} | \mathcal{D}_E). \quad (17)$$

Interestingly,  $\ln[\hat{Y}_{PGA,ES}]$  in Eq. (16) is the In-PGA estimated at site  $S$  borrowing information from the log-amplification map  $\delta S2S_{PGA,S}$ , from all the event-specific PGAs measurements and from all the event-specific PSMA measurements.

### Enhancement of the regional GMM

The recently proposed GMM<sup>26</sup> for Campi Flegrei area has functional form

$$\log_{10}[\hat{Y}_{PGA,ES}] = a + bM_{w,E} + (c + c_2 M_{w,E}) \log_{10}(\sqrt{R_{ES}^2 + h^2}) + c_3 \sqrt{R_{ES}^2 + h^2} + e\tilde{S} \pm \gamma, \quad (18)$$

where  $R_{ES}$  is the epicentral distance,  $M_{w,E}$  is the moment magnitude,  $\{a, b, c, c_2, c_3, h, e\}$  is the set of parameters to be calibrated and  $\gamma$  is the residual modelled as a zero-mean Gaussian random variable with standard deviation  $\sigma$ . The term  $e\tilde{S}$  in Eq. (18) describes the potential amplification of the ground-motion due to the site effect. Site conditions are classified according to the European Seismic Building Code EC8<sup>39</sup> in two classes: stiff soils (EC8 site class B) and soft soils (EC8 site class C).  $\tilde{S}$  is then equal to 0 and 1 for B and C soil classes, and the modelled amplification  $e\tilde{S}$  can only take two values depending on the site class.

The equation of our enhanced GMM is

$$\begin{aligned} \log_{10}[\hat{Y}_{PGA,ES}] &= a + bM_{d,E} + (c + c_2 M_{d,E}) \log_{10}(\sqrt{R_{ES}^2 + h^2}) \\ &+ c_3 \sqrt{R_{ES}^2 + h^2} + e \cdot \delta S2S_{PGA,S} \pm \gamma, \end{aligned} \quad (19)$$

where  $M_{d,E}$  is the duration magnitude and where  $\tilde{S}$  has been replaced by the estimated log-amplification  $\delta S2S_{PGA,S}$  depicted in Fig. 5.

## Data availability

The data used in this study are available on the Zenodo repository <https://zenodo.org/records/16420181><sup>40</sup>. Correspondence and request for materials should be addressed to Francesco Finazzi.

## Code availability

The MATLAB code used in this study is available on the Zenodo repository <https://zenodo.org/records/16420181><sup>40</sup>.

## References

- Wald, D. J. et al. TriNet “ShakeMaps”: rapid generation of peak ground motion and intensity maps for earthquakes in southern California. *Earthq. Spectra* **15**, 537–555 (1999).
- Bilham, R. The seismic future of cities. *Bull. Earthq. Eng.* **7**, 839–887 (2009).
- Thompson, E. M., Baise, L. G., Kayen, R. E., Tanaka, Y. & Tanaka, H. A geostatistical approach to mapping site response spectral amplifications. *Eng. Geol.* **114**, 330–342 (2010).
- Bergamo, P., Hammer, C. & Fäh, D. Correspondence between site amplification and topographical, geological parameters: collation of data from Swiss and Japanese stations, and neural networks-based prediction of local response. *Bull. Seismol. Soc. Am.* **112**, 1008–1030 (2022).
- Wald, D. J. & Allen, T. I. Topographic slope as a proxy for seismic site conditions and amplification. *Bull. Seismol. Soc. Am.* **97**, 1379–1395 (2007).
- Weatherill, G. et al. Modelling site response at regional scale for the 2020 European Seismic Risk Model (ESRM20). *Bull. Earthq. Eng.* **21**, 665–714 (2020).
- Loviknes, K., Cotton, F. & Weatherill, G. Exploring inferred geomorphological sediment thickness as a new site proxy to predict ground-shaking amplification at regional scale: application to Europe and eastern Türkiye. *Nat. Hazards Earth Syst. Sci.* **24**, 1223–1247 (2024).
- Peruzzi, G. et al. Assessing 1D litho-stratigraphical amplification factor for microzoning studies in Italy. *Bull. Earthq. Eng.* **14**, 373–389 (2016).
- Moscatelli, M., Albarello, D., Scarascia Mugnozza, G. & Dolce, M. The Italian approach to seismic microzonation. *Bull. Earthq. Eng.* **18**, 5425–5440 (2020).
- Bossu, R. et al. Rapid public information and situational awareness after the November 26, 2019, Albania earthquake: lessons learned from the LastQuake system. *Front. Earth Sci.* **8**, 235 (2020).
- Quitoriano, V. & Wald, D. J. USGS “Did You Feel It?”—Science and lessons from 20 years of citizen science-based macroseismology. *Front. Earth Sci.* **8**, 120 (2020).
- Finazzi, F. Fulfilling the information need after an earthquake: statistical modelling of citizen science seismic reports for predicting earthquake parameters in near realtime. *J. R. Stat. Soc. A Stat. Soc.* **183**, 857–882 (2020).
- Finazzi, F. The Earthquake Network project: toward a crowdsourced smartphone-based earthquake early warning system. *Bull. Seismol. Soc. Am.* **106**, 1088–1099 (2016).
- Bossu, R., Finazzi, F., Steed, R., Fallou, L. & Bondár, I. “Shaking in 5 Seconds!”—Performance and user appreciation assessment of the earthquake network smartphone-based public earthquake early warning system. *Bull. Seismol. Soc. Am.* **93**, 137–148 (2022).
- Kong, Q., Allen, R. M., Schreier, L. & Kwon, Y. W. MyShake: a smartphone seismic network for earthquake early warning and beyond. *Sci. Adv.* **2**, 1501055 (2016).
- Allen, R. ichardM. et al. Global earthquake detection and warning using Android phones. *Science* **389**, 254–259 (2025).
- Marcou, S., Allen, R. M., Abrahamson, N. A. & Sung, C. H. Ground-motion modeling using MyShake smartphone peak acceleration data. *Bull. Seismol. Soc. Am.* **115**, 86–105 (2024).
- Patel, S. C. & Allen, R. M. The MyShake App: user experience of early warning delivery and earthquake shaking. *Seismol. Res. Lett.* **93**, 3324–3336 (2022).
- Ricci, T., Barberi, F., Davis, M. S., Isaia, R. & Nave, R. Volcanic risk perception in the Campi Flegrei area. *J. Volcanol. Geoth. Res.* **254**, 118–130 (2013).
- Charlton, D., Kilburn, C. & Edwards, S. Volcanic unrest scenarios and impact assessment at Campi Flegrei caldera, Southern Italy. *J. Appl. Volcanol.* **9**, 1–26 (2020).
- Istituto Nazionale di Geofisica e Vulcanologia (INGV). Rete Sismica Nazionale (RSN) [Data set]. Istituto Nazionale di Geofisica e Vulcanologia (INGV). <https://doi.org/10.13127/SD/XOFXNH7QFY> (2005).
- Presidency of Council of Ministers - Civil Protection Department. Italian Strong Motion Network [Data set]. International Federation of Digital Seismograph Networks. <https://doi.org/10.7914/SN/IT> (1972).
- Weber, E. et al. An advanced seismic network in the southern Apennines (Italy) for seismicity investigations and experimentation with earthquake early warning. *Seismol. Res. Lett.* **78**, 622–634 (2007).
- Massin, F., Clinton, J. & Böse, M. Status of earthquake early warning in Switzerland. *Front. Earth Sci.* **9**, 707654 (2021).
- Schorlemmer, D., Oostwegel, L. J. N., Lars, L. & Evaz Zadeh, T. A model of European buildings. <https://doi.org/10.5880/GFZ.2.6.2023.011> (2024).
- Iervolino, I. et al. Seismic risk mitigation at Campi Flegrei in volcanic unrest. *Nat. Commun.* **15**, 1–14 (2024).
- Forte, G. et al. Seismic soil classification of Italy based on surface geology and shear-wave velocity measurements. *Soil Dyn. Earthq. Eng.* **122**, 79–93 (2019).
- Bommer, J. J. & Abrahamson, N. A. Why do modern probabilistic seismic-hazard analyses often lead to increased hazard estimates?. *Bull. Seismol. Soc. Am.* **96**, 1967–1977 (2006).
- D’Alessandro, A. & D’Anna, G. Suitability of low-cost three-axis MEMS accelerometers in strong-motion seismology: tests on the LIS331DLH (iPhone) accelerometer. *Bull. Seismol. Soc. Am.* **103**, 2906–2913 (2013).
- Kong, Q., Allen, R. M., Kohler, M. D., Heaton, T. H. & Bunn, J. Structural health monitoring of buildings using smartphone sensors. *Seismol. Res. Lett.* **89**, 594–602 (2018).
- Vezio, P., Mingozzi, E., Passarella, A. & Poli, N. Characterizing smartphone capabilities for seismic and structural monitoring. *Sci. Rep.* **14**, 23017 (2024).
- Böse, M., Julien-Laferrrière, S., Bossu, R. & Massin, F. Near real-time earthquake line-source models derived from felt reports. *Seismol. Res. Lett.* **92**, 1961–1978 (2021).
- Bossu, R., Böse, M., Steed, R. & Wald, D. J. The potential of crowd-sourced data for the rapid impact assessment of large earthquakes: the 2023 M 7.8 Kahramanmaraş-Pazarçık, Türkiye, Earthquake. *Seismol. Res. Lett.* **95**, 2058–2070 (2024).
- Fassò, A. & Finazzi, F. Maximum likelihood estimation of the dynamic coregionalization model with heterotopic data. *Environmetrics* **22**, 735–748 (2011).
- Wang, Y., Finazzi, F. & Fassò, A. D-STEM v2: a software for modeling functional spatio-temporal data. *J. Stat. Softw.* **99**, 1–29 (2021).
- Atik, L. A. et al. The variability of ground-motion prediction models and its components. *Seismol. Res. Lett.* **81**, 794–801 (2010).
- Zhu, C. et al. How well can we predict earthquake site response so far? Site-specific approaches. *Earthq. Spectra* **38**, 1047–1075 (2022).
- Parker, G. A. & Baltay, A. S. Empirical map-based nonergodic models of site response in the Greater Los Angeles Area. *Bull. Seismol. Soc. Am.* **112**, 1607–1629 (2022).
- CEN. EN 1998–1—Eurocode 8: Design of Structures for Earthquake Resistance—Part 1: General Rules, Seismic Actions and Rules for Buildings (Committee European de Normalisation, 2004).

40. Finazzi, F. Supporting code and data for the article “Citizens’ smartphones unravel earthquake shaking in urban areas”. Zenodo. <https://doi.org/10.5281/zenodo.16420181> (2025).
41. Esri. “Topographic” [basemap]. Scale Not Given. “World Topographic Map”. June 14, 2013. <http://www.arcgis.com/home/item.html?id=30e5fe3149c34df1ba922e6f5bbf808f> (2025).
42. Esri. “Light Gray Canvas” [basemap]. Scale Not Given. “Light Gray Canvas Base”. October 26, 2017. <https://www.arcgis.com/apps/mapviewer/index.html?layers=291da5eab3a0412593b66d384379f89f> (2025).

## Acknowledgements

This work has been partly funded by the European Commission under the Geo-INQUIRE project (number 101058518) within the HORIZON-INFRA-2021-SERV-01 call (FC). Opinions expressed in this article solely reflect the authors’ views; the EU is not responsible for any use that may be made of the information it contains. RB thanks the Fondation MAIF for supporting this research. The authors thank Iunio Iervolino and Pasquale Cito of the University of Naples Federico II for providing the soil classification of the Campi Flegrei area.

## Author contributions

F.F., F.C., and R.B. jointly conceived the study, discussed the results, and contributed to and commented on the manuscript at all stages. F.F. designed the statistical method and models, performed the data analysis and prepared the figures.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-025-64543-3>.

**Correspondence** and requests for materials should be addressed to Francesco Finazzi.

**Peer review information** *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work. A peer review file is available.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>

**Publisher’s note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025