Mechanical Response of Shallow Crust to Groundwater Storage Variations: Inferences from Deformation and Seismic Observations in the Eastern Southern Alps, Italy

F. Pintori[1,2], E. Serpelloni[1], L. Longuevergne[3], A. Garcia[1], L. Faenza[1], L. D'Alberto[4], A. Gualandi[1,5], M. E. Belardinelli[2]

[1]Istituto Nazionale di Geofisica e Vulcanologia, Italy.

[2]Università di Bologna, Dipartimento di Fisica e Astronomia, Settore di Geofisica, Bologna, Italy.

[3]Univ Rennes, CNRS, Geosciences Rennes - UMR 6118, F-35000 Rennes, France

[4]ARPA Veneto, Inland Waters Office, Padova, Italy.

[5]Department of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA, USA.

Contents of this file

Introduction

In this supporting information we present details regarding the data processing and the modeling tools used for obtaining the results presented in the paper. The supplementary material is organized as follows: in Text S1 we present the procedure of GNSS data processing, including the linear trend remotion technique through the vbICA analysis. In Text S2 we present the details regarding the computation of the mean values of meteorological variables, which are then used as input of the hydrological model GR5J, whose ability to simulate the observed river flows is unraveled. In Text S3 we describe the models tested to reproduce the hydrologically-induced transient deformation, providing, during a specific time interval, the modeled displacements. We also discuss how the inclusion of the topography, rock heterogeneity, and fracture's initial opening affects the resulting displacements and stress field. Finally, Text S4 presents the analyses based on the seismic catalog, i.e., the spatio-temporal declustering process and the statistical technique used to evaluate the correlation between water storage level and seismicity rates.

Text S1. Geodetic analysis

S1.1. GNSS data set and data processing

The position time-series have been obtained adopting a three-step procedure approach, as in Serpelloni et al. (2006), that includes: 1) raw phase data reduction, 2) combination of loosely constrained network solutions and reference frame definition and 3) time-series analysis, including velocity estimates and spatial filtering of common mode errors.

The raw GPS observables have been analyzed using the 10.70 version of the GAMIT/GLOBK package (Herring et al., 2018) adopting standards defined in the framework of the IGS "Repro2 campaign" (http://acc.igs.org/reprocess2.html). The GAMIT software is used to estimate station positions, atmospheric delays, satellite orbits, and Earth orientation parameters from ionosphere-free linear combination GPS phase observables using double differencing techniques to eliminate phase biases caused by drifts in the satellite and receiver clock oscillators. GPS pseudo-range observables are used to constrain clock timing offsets and to improve automated editing of the phase data, assisting in the resolution of integer phase ambiguities. GPS phase data are weighted according to an elevation-angle-dependent error model (Herring et al., 2015) using an iterative analysis procedure whereby the elevation dependence is determined from the observed scatter of phase residuals. In this analysis the satellites orbit parameters are tightly constrained to the IGS final products. We use the IGS absolute antenna phase center model for both satellite and ground-based antennas, which improves the accuracy of estimates for the vertical components of site position by mitigating reference frame scale and atmospheric mapping function errors (Schmid et al., 2005; 2007). While the first-order ionospheric delay is eliminated by the ionosphere-free linear combination, the second-order ionospheric corrections are applied based on the formulation of (Petrie et al., 2010), using IONEX files from the Center for Orbit Determination in Europe (CODE). The tropospheric delay is modeled as piecewise linear model and estimated using the Vienna Mapping Function 1 (VMF1; Boehm et al., 2007) with a 10° cutoff. We use the Global Pressure and Temperature 2 (GPT2; Lagler et al., 2013) model to provide a priori hydrostatic delays. The pole tide was also corrected in GAMIT by IERS standards. The Earth Orientation Parameters (EOP) are tightly constrained to priori values obtained from IERS Bulletin B. Non-tidal atmospheric loading and ocean tidal loading are corrected using MIT filtered atmospheric displacements files (available at ftp://everest.mit.edu/pub/GRIDS) and the FES2004 (Lyard et al., 2006) model,

respectively. The International Earth Rotation Service (IERS) 2003 model for diurnal and semi-diurnal solid Earth tides was set. Because of the large number of stations included in our Euro-Mediterranean GPS processing (~3000), this step is performed for several sub-networks, each made by <50 stations, with each sub-network sharing a set of high-quality IGS stations, which are used as tie-stations in the combination step.

In the second step we use the ST_FILTER program of the QOCA software (http://qoca.jpl.nasa.gov), which adopts a Kalman filter estimation algorithm (Dong et al., 1998; 2002), to combine all the daily loosely constrained solutions with the global solution of the IGS network made available by MIT (http://sopac.ucsd.edu), and simultaneously realize a global reference frame by applying generalized constraints (Dong et al., 1998). Specifically, we define the reference frame by minimizing the velocities of the IGS core stations (http://igscb.jpl.nasa.gov), while estimating a seven- parameter transformation with respect to the GPS realization of the ITRF2008 frame (Altamimi et al., 2011), i.e., the IGb08 reference frame.

In the third step we analyze the position time series in order to estimate and correct offsets due to stations equipment changes, while simultaneously estimating annual and semi-annual periodic signals and a linear velocity term. The model derived from the combination of these signals is then subtracted from the position time series in order to get the residual positions. The residual time-series are then used to estimate the Common Mode Error (CME) performing a Principal Component Analysis (PCA), as described in Dong et al. (2006). The PCA is performed at a continental-scale, over the same area used by Serpelloni et al. (2013), and the first two PCs are here considered as CME. This prevents the removal of the eventual more localized signals of geophysical interests recorded by the GPS stations in the study region, since the PCA detects the signals common to a much larger region. As a result, after removing the CME, the typical repeatability in our analysis is ~1 mm for the horizontal components, and ~3 mm for the vertical component, with a 30% gain in the daily repeatability and a significant improvement of the signal to noise ratio. After the spatial filtering, the estimated seasonal motions are added back to the filtered time-series, obtaining position time series with a reduced scatter around the adopted model. The filtered displacement time-series are rotated in a Adria-fixed reference frame using the rotation pole parameter estimated from GPS velocities in Serpelloni et al. (2016).

S1.2. Description of the procedure used to remove the linear trend using vbICA

We applied a vbICA (Gualandi et al., 2016) to the GPS time series in the Adria-fixed reference frame. Through this process, the displacement time-series of the GPS stations are decomposed into a finite number of independent components (ICs), characterized by a spatial distribution (U), a temporal evolution (V) and a weight coefficient (S). As shown by Gualandi et al. (2016), the number of ICs used to decompose the observations must be chosen a-priori, and in order to decide on the number of components to retain several statistical tests can be applied. In this work we use the F-test.

The component that describes the largest variance of the data set is IC1, which we interpret as the linear tectonic trend (Fig. S2a), to which a sinusoidal annual signal is superimposed. The presence of an annual signal in the component representing the tectonic trend is due to the fact that ICA is not able to completely separate the tectonic signal from other processes.

Assuming that the tectonic trend is linear, we initially estimate it by fitting the temporal evolution of IC1 with the following function:

$$q + m \cdot x + A \cdot sin(2\pi \cdot x + \varphi) \tag{S1}$$

and considering only the terms

$$V_{lin} = q + m \cdot x \tag{S2}$$

Since the displacements reconstructed by the IC1 are $V_1 \cdot S_1 \cdot U_1$, where S_1 is the weight coefficient and U_1 the spatial distribution, the displacements associated only with the linear trend of IC1 are $V_{lin} \cdot S_1 \cdot U_1$.

It is then possible to remove this component of the displacement ($V_{lin} \cdot S_1 \cdot U_1$) from the displacement time series and then apply another vbICA to the new detrended time-series. This approach is effective in detrending short spanning time-series of GNSS stations affected by transient displacements, where a trajectory model (Bevis and Brown, 2014) would fail if the transient is not properly modeled with an ad hoc function.

We remark that the values of both q and m are slightly affected by both the number of components chosen for the ICA, in particular when considering 4 ICs or more, and the initial decomposition parameters. Since once removed the linear trend, 3 components are necessary to reconstruct the observations according to the F-test, then we selected 4 ICs for the decomposition of the signal in the Adria-fixed reference frame (Fig. S2): one more than the detrended case, representing the tectonic motion.
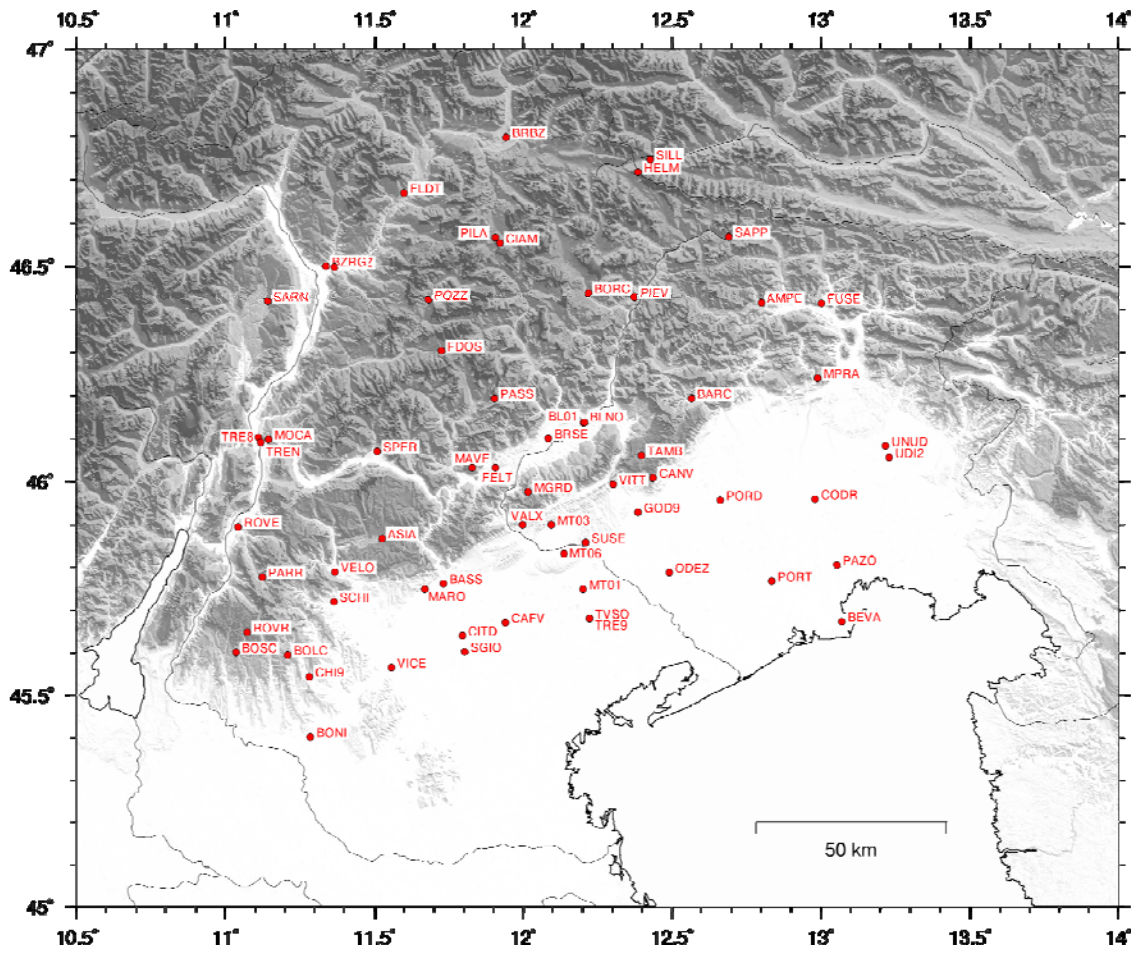
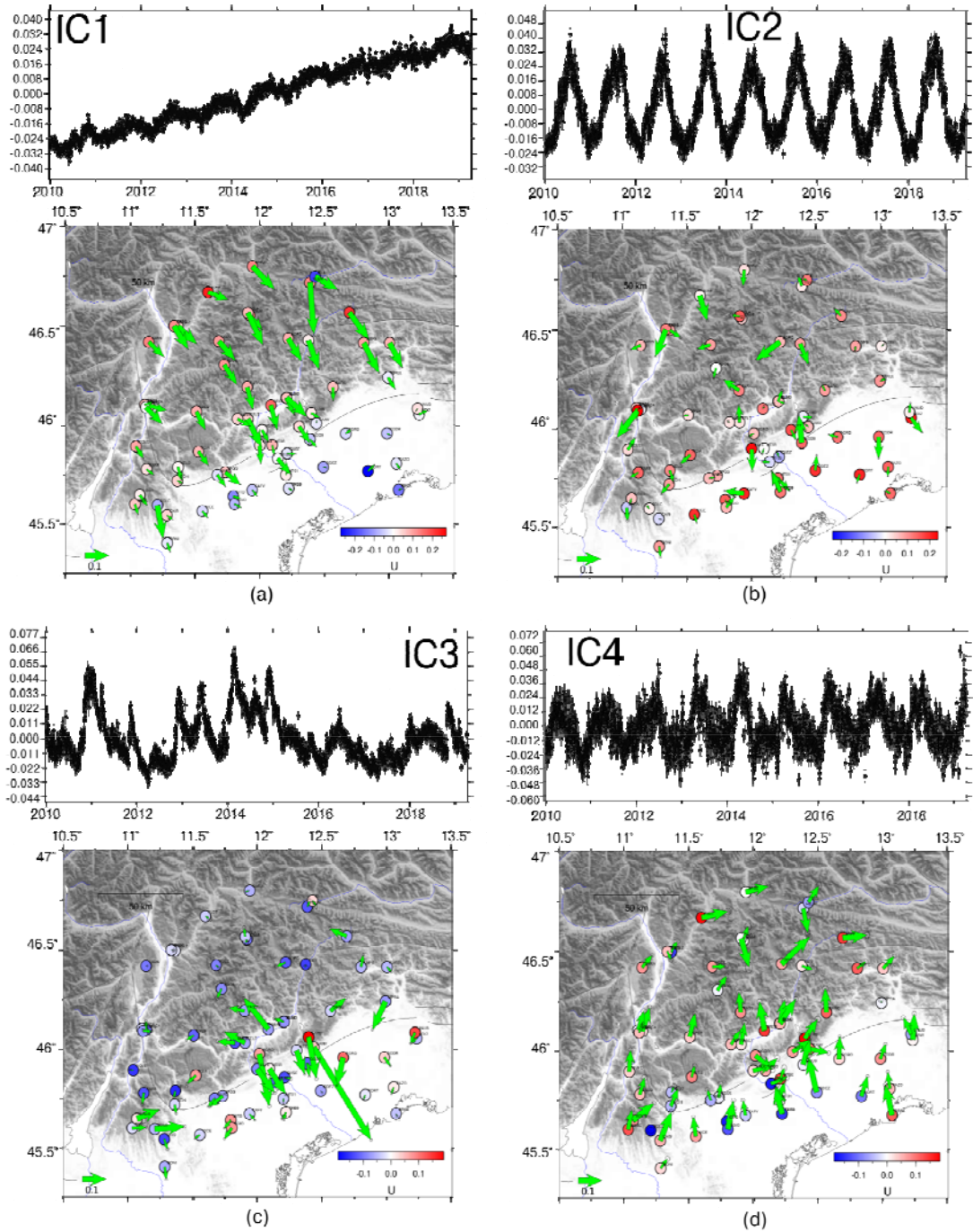**Figure S1.** Map of the GPS stations included in the vbICA analysis.

Figure S2. Result of the vbICA analysis using GNSS time series in the Adria-fixed reference frame. The (a), (b), (c), (d) panels represent IC1, IC2, IC3, IC4 respectively. The top of each panel represents the temporal evolution (V; in black), while in the map is plotted the corresponding spatial response in the horizontal (green arrows) and vertical (coloured circles) components. As regard IC1, the directions and amplitudes of the spatial response are in good agreement with horizontal and vertical velocities estimated with the classic trajectory model approach of Anderlini et al. (2020).
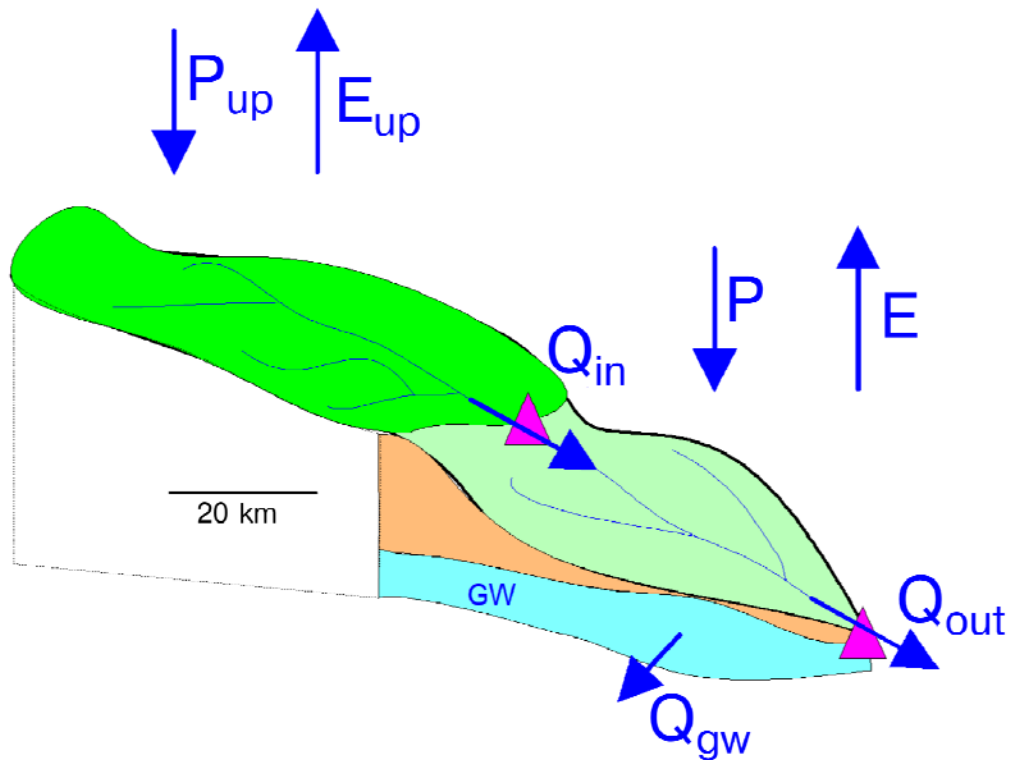
6

Text S2. Hydrological modeling

Figure S3. Schematic describing the modeling approach to estimate total water storage changes in a downstream sub-catchment (light green), based on precipitation (P), actual evapotranspiration (E), river discharge (Qin and Qout), and potential groundwater import/export from a surrounding basin Qgw (e.g. karstic system).

S2.1. Spatially-distributed precipitation and temperature

As the single water flux increasing storage changes, an adequate estimation of precipitation is required to robustly quantify them. Since rainfall is highly heterogeneous over this mountainous region, we considered 54 pluviometers and 55 temperature observations distributed within or close to the Piave at Segusino catchment (Fig. S4). Data are available from 2010 to present.

We used the Thiessen polygons (i.e. nearest neighbour method) to compute mean rainfall over each of the three catchments. This method consists in dividing the entire basin in Thiessen polygons generated from a set of sample points, which in the case of precipitation estimation are the pluviometers. Each polygon has the property of the closest pluviometer. Note that the shape of the Thiessen might evolve over time to account for potential missing data.

Assuming that inside (or very close to) the hydrological basin we are considering there are n pluviometers, the weighted mean precipitation (P_m) at certain time is computed as follows:

$$P_m = \frac{\sum_{i=1}^n p_i A_i}{\sum_{i=1}^n A_i} \qquad (S3)$$

where p_i is the precipitation recorded by the i-th pluviometer and A_i the area of the polygon it represents. The time resolution of the data we are using is daily, then P_n is daily too.

This computation process can be complicated by the presence of missing data and by the fact that the calculus of polygon areas is not that easy. Then, in order to compute the areas of Thiessen polygons, we build a grid with squared cells over the basin with a resolution of 2 km both in the N-S and E-W direction, and count how many grid cells points (at the corners of each cell) lie inside each polygon: the larger the polygon is the more points are inside it (Fig. S4). In this way A_i in equation (S3) becomes the number of points that are inside the i-th polygon.

In order to compute the mean temperature of the area we follow the same process, using as sample points the position of the thermometers.
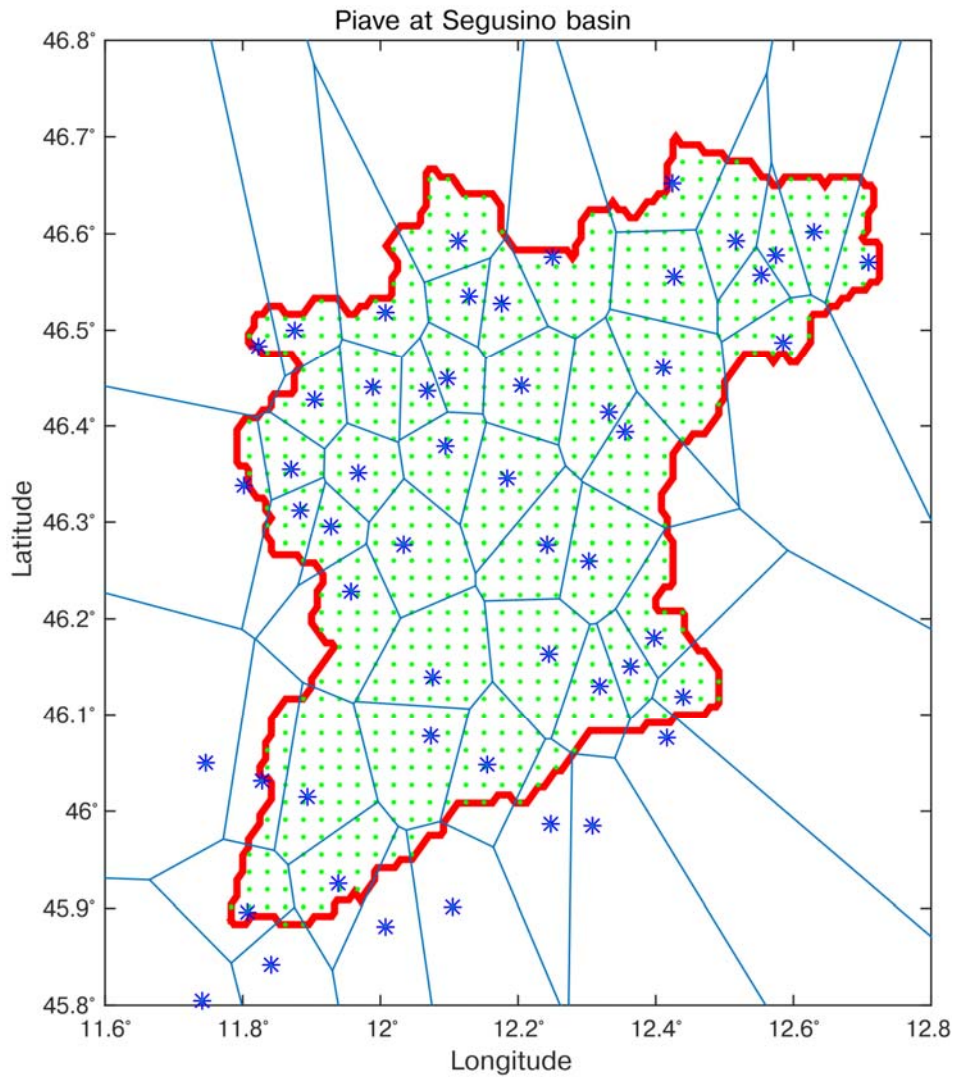

Figure S4. Example of division of the Piave at Segusino basin (red line) in Thiessen polygons

(blue lines) to compute mean precipitation. The blue stars represent the pluviometers that generate the Thiessen polygons, while the small green dots inside the basin are the grid points we used to compute A_i.

S2.2. Potential evapotranspiration

Actual evapotranspiration E is first estimated based on potential evapotranspiration (PET) - i.e. the atmospheric demand for moisture. We used the Jensen-Haise (Jensen et al. 1990) method to estimate PET:

$$PET = \frac{R_e(T+5)}{100\lambda\rho} \text{ if } T + 5 > 0; PET = 0 \text{ otherwise}$$ (S4)

This approach only requires a limited amount of information, temperature (T) and extraterrestrial radiation (R_e), depending only on latitude and julian day. $\lambda = 2.45\frac{MJ}{Kg}$ is the latent heat flux and $\rho = 1000\frac{Kg}{m^3}$ is the density of water. As underlined by Oudin et al. (2005), such a simplified model, which is based on local observations, does not affect the model performances.

S2.3. Modeling water storage changes with GR5J hydrological model (Belluno Valley, 883 km²)

The quality of the model is evaluated using the Nash-Sutcliffe efficiency (Nash and Sutcliffe, 1970). This normalized index can range from $-\infty$ to 1. An efficiency of 1 corresponds to a perfect description of observed discharge. When the efficiency falls to 0, model predictions are as accurate as the mean of the observed data. GR5J performs quite well on the three catchments (Table S1, Fig. S5).

Watershed	Catchment area [km²]	Nash-Sutcliffe efficiency	Correlation between modeled water storage changes and ICA V2 eigenvector
Cordevole@Ponte Mas	706	0.62	0.79
Piave@ Belluno	1907	0.72	0.84
Piave@ Segusino	3496	0.76	0.88
Belluno Valley	883	N/A	0.83

GRACE	~105	N/A	0.61

Table S1. Nash-Sutcliffe efficiency coefficient and correlation with V2 for the basins considered.
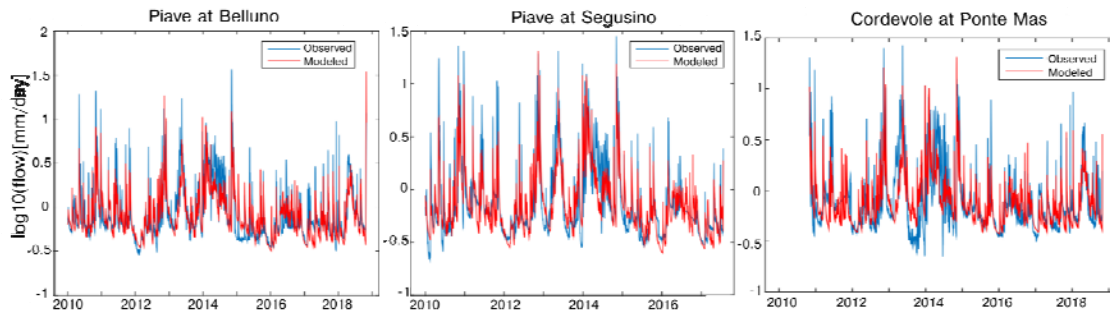

Figure S5. Comparison between observed (blue) and modeled (red) flows of the Piave river at Belluno (left), Segusino (center) and of the Cordevole river at Ponte Mas (right).

For all basins, water storage changes are highly correlated with the temporal variations in surface displacements, as described by the second eigenvector of ICA (Fig. S6), around or above 0.8. The correlation between V2 and GWS is high during all the time interval, implying that not only the large deformation events, but also the small ones are well modeled. One exception is the extreme weather event, named "tempesta Vaia", occurred at the end of October 2018, where the increase of the V2 is smaller than the GWS. The reason might be that, during this event, the precipitation in the northern sector of the basins was much higher than in the southern one, where the GPS stations responding to this hydrological signal are located. This is confirmed by Fig. 4 and Fig. S6, where the GWS peak during the Vaia storm computed in the Belluno Valley catchment basin is much smaller than in the river basins.
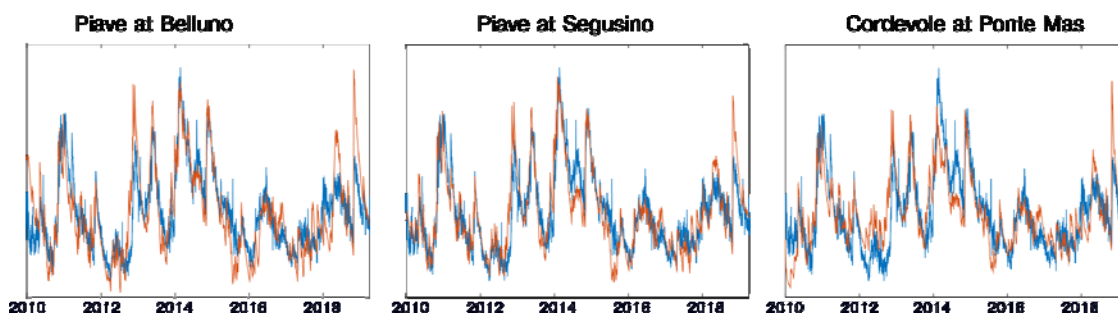

Figure S6. Comparison between V2 (blue) and GWS variations (red) computed in the Piave at Belluno (left), Piave at Segusino (center) and Cordevole at Ponte Mas (right).
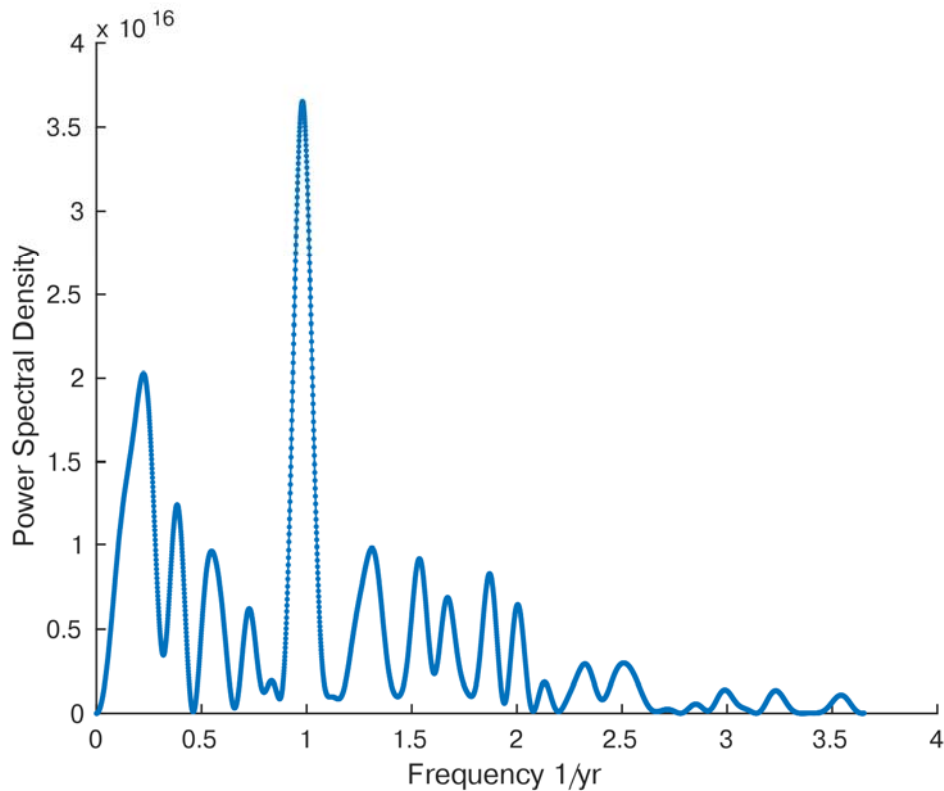
Figure S7. Power spectral density of GWS$_{res}$.

Text S3. Hydro-mechanical modeling

S3.1. Modeling displacements and stress changes

We used a finite element model (FEM) in order to explore physical linkages between water storage changes in the Belluno Valley and surface deformation recorded by GNSS stations, and compute subsurface stress variations. We define a geologically realistic model and test a set of hydrological structures that could explain the observed surface displacements.

In order to focus on physical processes linking the hydrological cycle and deformation, the problem was reduced to 2-D under the plane strain hypothesis, where linear elasticity is resolved considering small deformations. The modeled domain is 1000 x 500 km, i.e. 200 time wider than the area of interest to avoid boundary effects. The chosen boundary conditions are fixed constraints in the lateral and bottom edges, free surface on the top edge (Fig. S8). The mesh is triangular with 20 m size near the boundaries. The size of the domain and the associated mesh were determined as the best compromise between stability and accuracy of the deformed

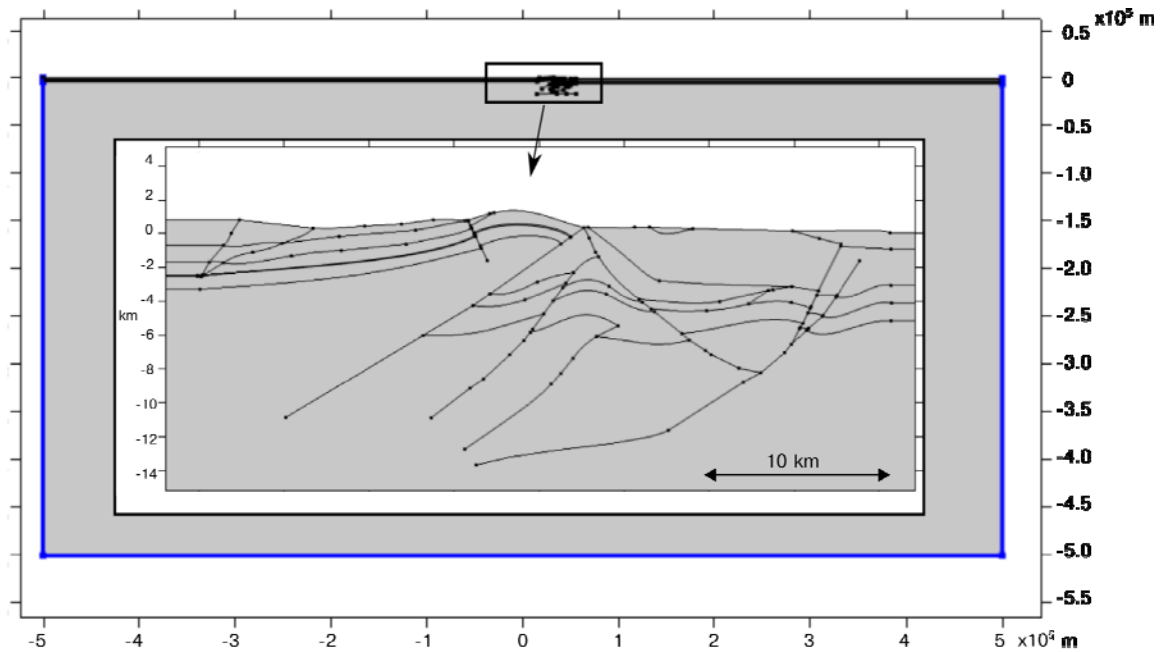domain and the calculation time. The FEM is performed with Comsol Multiphysics (https://www.comsol.com/).

Figure S8. Entire domain of the COMSOL model, in blue the fixed constraint. In the black rectangle a closer look to the modeled study area.

The definition of the model is based on the geological cross section shown in Fig. 5 of the main text (from Galadini et al., 2005), where the main geological units are explicitly described. Each layer is considered as elastic, homogeneous and isotropic. We considered perfect continuity between the different layers. No specific elastic behavior was considered for the faults following the small deformation hypothesis. Elastic parameters attributed to the different formations, according to Anselmi et al. (2011), are highlighted in Table S2.

Rock type	Young's modulus [GPa]	Poisson's ratio
Crystalline basement	92	0.25
Belluno Basin Units	59	0.26
Igne Formation	64	0.3
Flysch	26	0.2
Scaglia Rossa	38	0.3
Dolomites	95	0.32
Montello Conglomerates	26	0.2
Limestones	71	0.25
Sediments	15	0.35

Table S2. Rock parameters used in COMSOL.

S3.2. Testing hydrological pressure sources to explain surface displacements

In this section we provide additional information about how the different models we tested describe the relationship between GWS_{res} and the displacements associated with IC2 in a specific time-span. We assume that the pressure variations caused by the accumulation of water are directly proportional to the GWS_{res}.

As deformation is linear with pressure, the different mechanical models are evaluated on the ability of reproducing the spatial distribution of relative surface displacements extracted from IC2. In this way, the mechanical model remains independent from any hydrological model. Noise levels are also considered in the evaluation. It is worth noting that the vertical displacements associated with IC2 are noisier than the horizontal ones, as shown in Fig. 7, so that the analysis of the model performance is mainly driven by the horizontal displacements. In particular, when considering the vertical displacements associated with IC2 during T1, (10 October 2013 - 22 February 2014), 11 of the 12 stations considered have a value smaller than the mean noise level, which is, for each GPS station, the mean error associated with the daily measurements. The model displacements, for each configuration, are shown in Fig. S9.
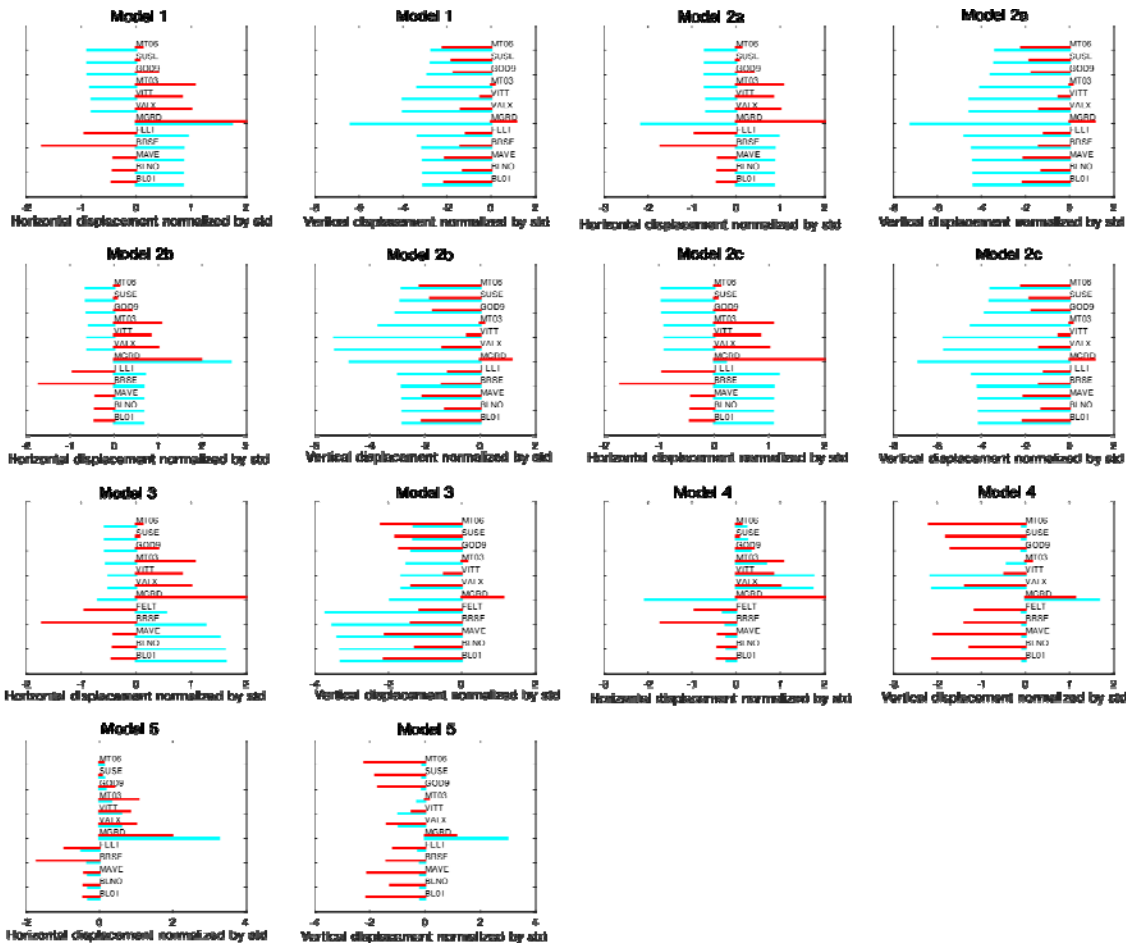
Figure S9. Comparison between displacements associated with IC2 during T1 (red) and modeled using COMSOL using 7 different sources of deformation (blue). In order to better compare the displacement patterns, both horizontal and vertical displacements have been normalized with their standard deviation.
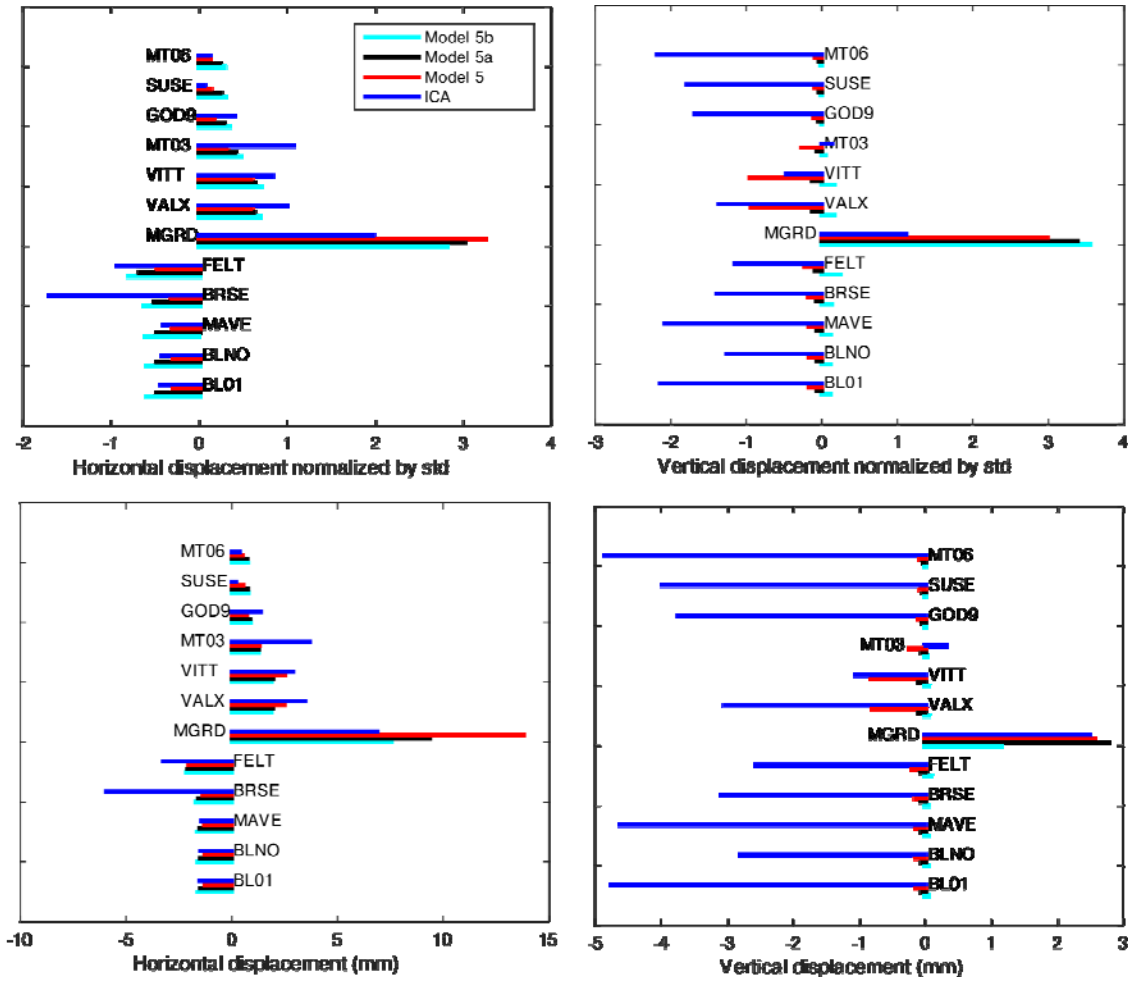
Figure S10. Top: comparison among displacements generated by using Model 5 (red); Model 5a (black); Model 5b (blue). To better compare the displacement patterns, both horizontal and vertical displacements have been normalized with their standard deviation. Bottom: same as top, but without normalization.

S3.3. Computation of the Coulomb failure function

We consider Coulomb stress change (CFF) in order to characterize the conditions for rock failure. It is defined on a fault plane as

$$CFF = \Delta\tau_r + \mu'\Delta\sigma_n \tag{S5}$$

where $\Delta\tau_r$ is the change in the shear stress on the plane in the expected slip direction on the target fault, $\Delta\sigma_n$ is the change on normal stress (positive for extension) and μ' is the apparent coefficient of friction (Harris and Simpson, 1992), which we assumed = 0.5.

We compute $\Delta\tau_r$ and $\Delta\sigma_n$ starting from the three components of the stress tensor ($\Delta\sigma_{xx}$, $\Delta\sigma_{yy}$, $\Delta\sigma_{xy}$) obtained by the COMSOL model in a x, y reference frame, where y is vertical positive with elevation and x horizontal, positive toward SW.

Since the receiving faults we are considering are inverse:

$$\Delta\sigma_n = \Delta\sigma_{xx}sin^2(\delta) - 2\Delta\sigma_{xy}\big(sin(\delta)cos(\delta)\big) + \Delta\sigma_{yy}cos^2(\delta) \tag{S6}$$

$$\Delta\tau_r = cos(\delta)sin(\delta)\big(\Delta\sigma_{yy} - \Delta\sigma_{xx}\big) + \Delta\sigma_{xy}\big(cos^2(\delta) - sin^2(\delta)\big) \tag{S7}$$

Where δ is the dip angle.

Considering the high correlation between water storage changes, deformation and seismicity, with no significant time delay, pore pressure changes is not considered in the definition of CFF, and pore pressure changes won't be modeled.
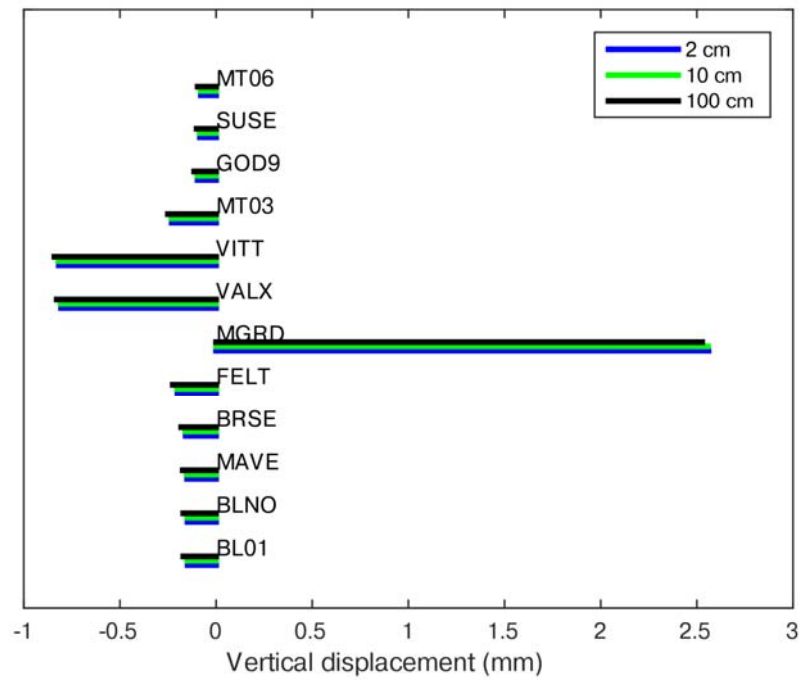
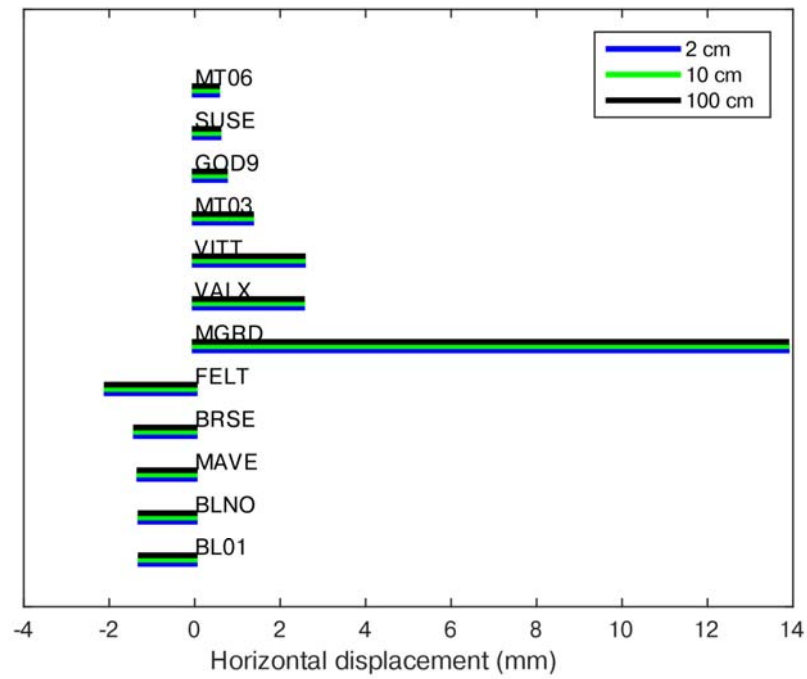**Figure S11.** Horizontal (top) and vertical (bottom) displacements generated by the fracture described as Model 5, considering different values of initial opening: 2 cm (blue), 10 cm (green), 100 cm (black).
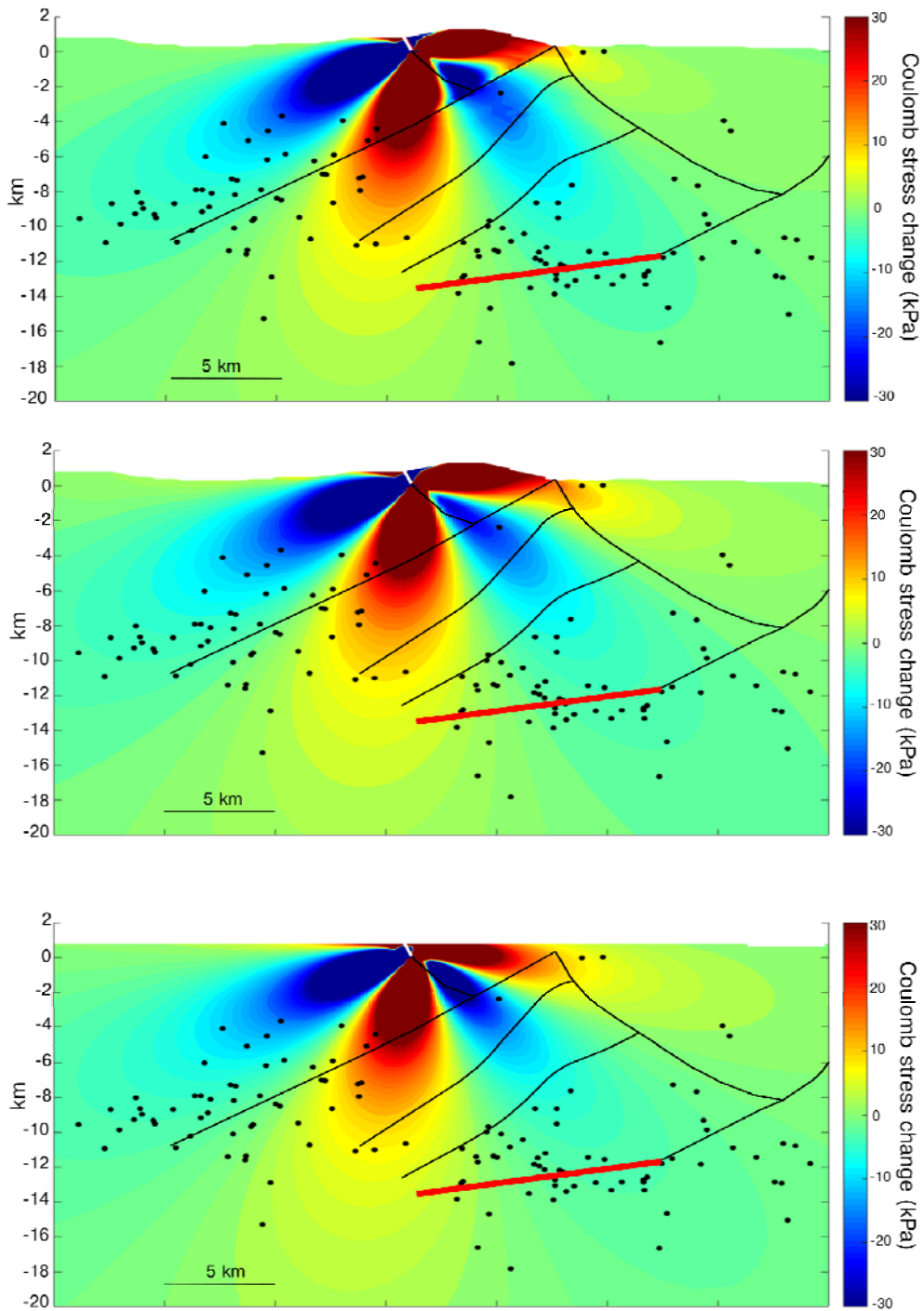
Figure S12. Same as Fig. 9, considering a domain with: both topography and rock stratification (Model 5, top); only topography (Model 5a, central); a flat free surface (Model 5b, bottom).

Text S4. Correlation between GWS$_{res}$ and seismicity rates

S4.1 ETAS modelling and temporal declustering

Completeness magnitude determination

The first step to investigate the possible correlation between the GWS$_{res}$ and seismicity rate is to conduct statistical analysis on the seismic catalogue. Assessing the magnitude of completeness (Mc), defined as the minimum magnitude above which we have reliably recorded all the earthquakes in the time and the region under investigation is the first, essential step to analyse seismicity rates. Romano et al elaborated a Mc map for this catalog, identifying higher Mc values for the seismicity located in the NE of the domain (Mc~0.7-1.0), respect to the Mc for the areas with better network coverage (where Mc reach values below 0.0).

In this work we calculate a Mc for the full data set. Numerous algorithms for data-driven Mc selection have been proposed; Mignan and Woessner (2012) provide a quite exhaustive overview. To estimate Mc we use the "Completeness Magnitude Estimation" tool available on the IS-EPOS platform (IS-EPOS, 2016), which provides four different Mc estimations using different approaches (the goodness of fit at 90% and 95% confidence bounds, the Maximum Curvature Method (Wiemer, 2000) and the Modified Goodness of fit (Leptokaropoulos et al., 2013). The obtained Mc solutions are summarized in Table S3; since the value of 0.5 leads to unstable results, we adopt the more conservative value of Mc = 0.7, which is close to the higher values for the seismicity located in the NE part of the domain in the Mc map presented by Romano et al. 2019. The resulting complete catalogue is composed by 731 events.

Method	Mc
Goodness of fit test at 90% confidence bounds	0.5
Goodness of fit test at 95% confidence bounds	**0.7**
Maximum Curvature Method	0.5
Modified Goodness of fit test	**0.7**

Table S3. Summary of the Mc obtained using different methods (tools from IS-EPOS, 2016).

Temporal declustering using the ETAS model

The seismic declustering consists in the identification and separation of the contribution of foreshocks, mainshocks and aftershocks, and to construct a catalogue composed only by independent events, i.e., those considered as not triggered by any preceding event. There are several approaches for solving the declustering problem, with a not unique solution. Van Stiphout et al. (2012) provides an overview of this issue, describing the pros and cons of the most popular algorithms. In this application, we adopt the strategy of Zhuang et al. (2002) which is based on the Epidemic Type Aftershock Sequence (ETAS) model. This model, firstly developed by Ogata (Ogata, 1988; 1998), considers the seismicity as the sum of background earthquakes thought to be caused by (stationary) tectonic loading and triggered earthquakes, thought to be caused by stress transfer. The stochastic declustering gives to each earthquake in the catalogue

the probability of being an aftershock of all past events. Starting from the ETAS model, the conditional intensity, in the temporal domain, can be defined as:

$$\lambda_0(t) = \mu + \sum_{\{i:t_i<t\}} \frac{Ke^{\alpha(m-m_c)}}{(t-c)^p} \tag{S8}$$

where μ represents the background events, that we assume time-independent; the cascade of aftershocks (in the summation) is described by the empirical Omori–Utsu law (Utsu et al., 1995), where k is the productivity factor of the sequence that depends on the magnitude (m) and m_c is the completeness magnitude. For the joint inversion of the five parameters for our complete catalogue, the maximum-likelihood method (Ogata et al., 1993) is used; we obtain: μ = 0.102 day−1, K = 0.025, α = 0.940, p = 0.907 and c = 0.226−2 yr.

For each event, it is possible to calculate the contribution to the background and the whole conditional intensity from equation (S8) and, from here, to select the most likely background events. We have found that 51% of the events are classified as background (372 earthquakes). The spatial distribution of these events is shown in Fig. 7 in the main text.

S4.2 Covariate model

The covariate model from Garcia-Aristizabal (2018) provides a way for studying correlations between seismicity rates and proxies of possible forcing processes of interest. The exponential distribution is defined as the basic template function for modelling the t_{IET} as in equation (3) in the main text. The possible dependencies on hydrological data are modelled writing the μ parameter of the exponential distribution in terms of deterministic functions of the explanatory covariate according to equation (4) in the main text. The inference of model parameter values is performed using a Bayesian approach based on a Markov chain Monte Carlo (MCMC) method, whereas the model selection is based on Bayes factor (B_{KL}) calculations (Garcia-Aristizabal et al., 2015; 2016; 2018). B_{KL} provides a possible solution for hypothesis testing and model selection in Bayesian inference problems (e.g. Kass and Raftery, 1995; Raftery, 1995; Lewis and Raftery, 1997) and is calculated as the ratio of the posterior odds for M_K against M_L to the prior odds. When the models M_K and M_L are a priori equally probable, B_{KL} reduces to the ratio of the marginal likelihoods of the two competing models (Lewis and Raftery, 1997) :

$$B_{KL} = \frac{f(x|M_K)}{f(x|B_L)} \tag{S9}$$

The Bayes factor summarizes the evidence provided by the data x in favour of one specific model as opposed to another. Reference values for interpreting B_{KL} have been provided by Jeffreys (1961) and Raftery (1995) and are summarized in Table S4 (from Garcia-Aristizabal, 2018).

$2log_{10}(B_{KL})$	B_{KL}	**Evidence for M_K**
< 0	< 1	Negative (supports M_L)
0 - 2	1 - 3	Barely worth mentioning
2 - 5	3 - 12	Positive
5 - 10	12 - 150	Strong
> 10	> 150	Very strong

Table S4. Categories used for the interpretation of the Bayes factors for the model selection (from Raftery, 1995; Jeffreys, 1961; Garcia-Aristizabal, 2018).

S4.3 Spatial cluster analysis

For spatial cluster analysis we use the k-means partitioning approach (MacQueen, 1967). This algorithm partitions a data set into k clusters so that the resulting intra-cluster similarity is high and the inter-cluster similarity is low (Han et al., 2011). We select the "optimal number of clusters" using the silhouette method (Rousseeuw, 1987), which is one of the most popular approaches used in literature (Kaufman and Rousseeuw, 1990; Al-Zoubi and Rawi, 2008; Garcia-Aristizabal et al., 2017). Fig. S13 shows the plot of the silhouette coefficient; the maximum of this plot, which is reached for k=2, indicates the preferred number of clusters to partition the data. In this way, the partition in two clusters (A and B in Fig. 7) was obtained. It is worth noting that the classification of events located in the boundary between the two selected clusters may be discutible. To test the robustness of the obtained results we performed different tests in order to perform correlation analyses using resizing clusters A and B after manually changing the boundary between them (as e.g., following topographic features). These changes didn't modified the resulting correlations between seismicity in the (modified) clusters A and B and GWS_{res}.
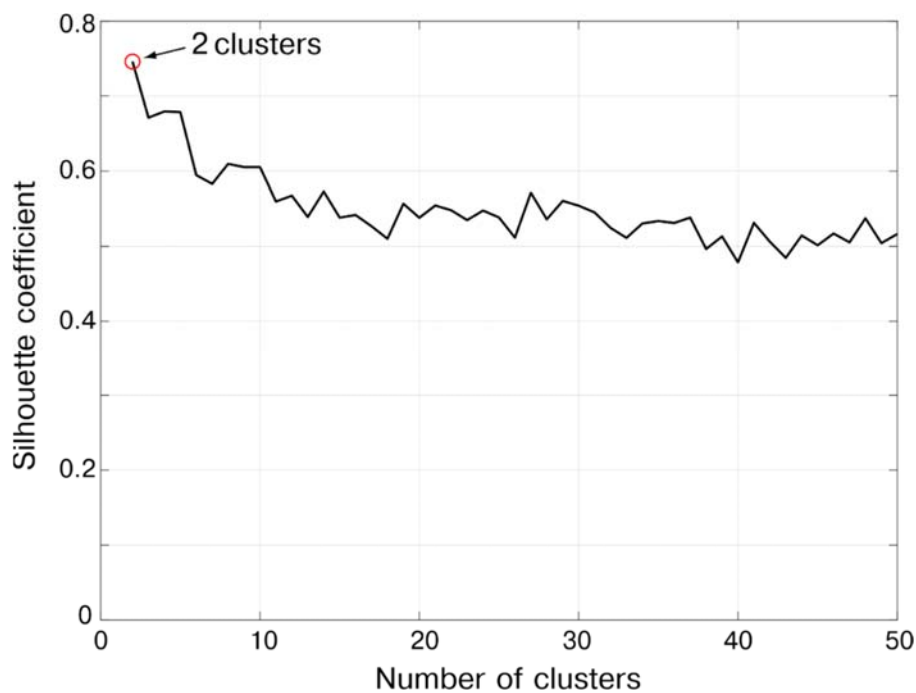

Figure S13. Plot of the mean Silhouette coefficient. The maximum value is found for a partition in two clusters (identified as clusters A and B in the paper).

Movie S1.

In the top panel we show the temporal evolution of GWS$_{res}$ (same as Fig. 4 in the maintext). In the bottom panel, the red segments connecting black dots and red dots represent the horizontal displacements associated with IC2 of each GNSS station. The displacements are calculated, as in section S1.2, with respect to the position at the instant corresponding to the absolute minimum of V2 (Fig. 4).

References

Al-Zoubi, M.B., & Rawi, M. (2008). An efficient approach for computing silhouette coefficients. *Journal of computer science*, 4, 252-255. https://doi.org/10.3844/jcssp.2008.252.255

Altamimi, Z., Collilieux, X., & Métivier, L. (2011). ITRF2008: an improved solution of the international terrestrial reference frame. *Journal of Geodesy*, 85, 457–473. https://doi.org/10.1007/s00190-011-0444-4

Anderlini, L., Serpelloni, E., Tolomei, C., De Martini, P.M., Pezzo, G., Gualandi, A., & Spada, G. (2020). New insights into active tectonics and seismogenic potential of the Italian Southern Alps from vertical geodetic velocities. https://doi.org/10.5194/se-2020-10

Andermann, C., Longuevergne, L., Bonnet, S., Crave, A., Davy, P., & Gloaguen, R. (2012). Impact of transient groundwater storage on the discharge of Himalayan rivers. *Nature Geoscience,* 5, 127–132. https://doi.org/10.1038/ngeo1356

Anselmi, M., Govoni, A., De Gori, P. & Chiarabba, C. (2011). Seismicity and velocity structures along the south-Alpine thrust front of the Venetian Alps (NE-Italy). *Tectonophysics*, 513, 37–48. https://doi.org/10.1016/j.tecto.2011.09.023

Bevis, M., & Brown, A. (2014). Trajectory models and reference frames for crustal motion geodesy. *Journal of Geodesy*, 88, 283–311. https://doi.org/10.1007/s00190-013-0685-5

Boehm, J., Heinkelmann, R., & Schuh, H. (2007). Short Note: A global model of pressure and temperature for geodetic applications. *Journal of Geodesy*, 81, 679–683. https://doi.org/10.1007/s00190-007-0135-3

de Lavenne, A., Thirel, G., Andréassian, V., Perrin, C., & Ramos, M.-H. (2016). Spatial variability of the parameters of a semi-distributed hydrological model. *Proceedings of the International Association of Hydrological Sciences*, 373, 87–94. https://doi.org/10.5194/piahs-373-87-2016

Dong, D., Fang, P., Bock, Y., Cheng, M., & Miyazaki, S. (2002). Anatomy of apparent seasonal variations from GPS-derived site position time series. *Journal of Geophysical Research*, 107(B4), 2075. https://doi.org/10.1029/2001JB000573

Dong, D., Fang, P., Bock, Y., Webb, F., Prawirodirdjo, L., Kedar, S., & Jamason, P. (2006). Spatiotemporal filtering using principal component analysis and Karhunen-Loeve expansion approaches for

regional GPS network analysis. *Journal of Geophysical Research*, 111, B03405. https://doi.org/10.1029/2005JB003806

Dong, D., Herring, T., & King, R. (1998). Estimating regional deformation from a combination of space and terrestrial geodetic data. *Journal of Geodesy*, 72(4), 200–214. https://doi.org/10.1007/s001900050161

Galadini, F., Poli, M.E., & Zanferrari, A. (2005). Seismogenic sources potentially responsible for earthquakes with M ≥ 6 in the eastern Southern Alps (Thiene-Udine sector, NE Italy). *Geophysical Journal International*, 161, 739–762. https://doi.org/10.1111/j.1365-246X.2005.02571.x

Garcia-Aristizabal, A. (2018). Modelling fluid-induced seismicity rates associated with fluid injections: examples related to fracture stimulations in geothermal areas. *Geophysical Journal International*, 215, 471–493. https://doi.org/10.1093/gji/ggy284

Garcia-Aristizabal, A., Bucchignani, E., & Manzi, M.P. (2017). Patterns in Climate-Related Parameters as Proxy for Rainfall Deficiency and Aridity: Application to Burkina Faso. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering, ISSN (online)*: 2376-7642, https://doi.org/10.1061/AJRUA6.0000860

Garcia-Aristizabal, A., Bucchignani, E., Palazzi, E., D'Onofrio, D., Gasparini, P., & Marzocchi, W. (2015). Analysis of non-stationary climate-related extreme events considering climate change scenarios: an application for multi-hazard assessment in the Dar es Salaam region, Tanzania. *Natural Hazards*, 75, 289–320. https://doi.org/10.1007/s11069-014-1324-z

Garcia-Aristizabal, A., Caciagli, M., & Selva, J. (2016). Considering uncertainties in the determination of earthquake source parameters from seismic spectra. *Geophysical Journal International*, 207, 691–701. doi:10.1093/gji/ggw303

Gualandi, A., Serpelloni, E., & Belardinelli, M.E. (2016). Blind source separation problem in GPS time series. *Journal of Geodesy*, 90, 323–341. https://doi.org/10.1007/s00190-015-0875-4

Han, J., Kamber, M., & Pei, J. (2011). *Data mining concepts and techniques third edition*. Morgan Kaufmann.

Harris, R.A., & Simpson, R.W. (1992). Changes in static stress on southern California faults after the 1992 Landers earthquake. *Nature*, 360, 251–254. https://doi.org/10.1038/360251a0

IS-EPOS. (2016). Completeness Magnitude Estimation [Web application/Source code]. Retrieved from https://tcs.ah-epos.eu/

Jeffreys, H., (1961). *Theory of probability, 3rd edn oxford*: Oxford university press.

Jensen, M.E., Burman, R.D., & Allen, R.G., (1990). Evapotranspiration and irrigation water requirements. ASCE Manuals and Reports on Engineering Practices No. 70., *American Society of Civil Engineers*, New York, NY, 360 p.

Kass, R.E., & Raftery, A.E. (1995). Bayes Factors. *Journal of the American Statistical Association*, 90, 773–795. https://doi.org/10.1080/01621459.1995.10476572

Kaufman, L., & Rousseeuw, P.J. (Eds.) (1990). *Partitioning around medoids (program PAM),* in: Finding Groups in Data, Wiley Series in Probability and Statistics. John Wiley & Sons, Inc., Hoboken, NJ, USA, pp. 68–125. https://doi.org/10.1002/9780470316801.ch2

Lagler, K., Schindelegger, M., Böhm, J., Krásná, H. & Nilsson, T. (2013). GPT2: Empirical slant delay model for radio space geodetic techniques. *Geophysical Research Letters*, 40, 1069–1073. https://doi.org/10.1002/grl.50288

Leptokaropoulos, K.M., Karakostas, V.G., Papadimitriou, E.E., Adamaki, A.K., Tan, O. & Inan, S. (2013). A homogeneous earthquake catalog for western turkey and magnitude of completeness determination. *Bulletin of the Seismological Society of America*, 103, 2739–2751. https://doi.org/10.1785/0120120174

Lewis, S.M., & Raftery, A.E. (1997). Estimating Bayes Factors via Posterior Simulation with the Laplace—Metropolis Estimator. *Journal of the American Statistical Association*, 92, 648–655. doi:10.1080/01621459.1997.10474016

Longuevergne, L., Florsch, N., Boudin, F., Oudin, L., & Camerlynck, C. (2009). Tilt and strain deformation induced by hydrologically active natural fractures: application to the tiltmeters installed in Sainte-Croix-aux-Mines observatory (France). *Geophysical Journal International*, 178, 667–677. https://doi.org/10.1111/j.1365-246X.2009.04197.x

Lyard, F., Lefevre, F., Letellier, T., & Francis, O. (2006). Modelling the global ocean tides: modern insights from FES2004. *Ocean Dynamics*, 56, 394–415. https://doi.org/10.1007/s10236-006-0086-x

MacQueen, J. (1967). *Some methods for classification and analysis of multivariate observations. Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, 1, 281–297. University of California Press, Berkeley, Calif., 1967. https://projecteuclid.org/euclid.bsmsp/1200512992

Mignan, A., & Woessner, J. (2012). Estimating the magnitude of completeness for earthquake catalogs. *Community Online Resource for Statistical Seismicity Analysis*. https://doi.org/10.5078/corssa-00180805

Nash, J.E., & Sutcliffe, J.V. (1970). River flow forecasting through conceptual models part I — A discussion of principles. *Journal of Hydrology*, 10, 282–290. https://doi.org/10.1016/0022-1694(70)90255-6

Ogata, Y. (1988). Statistical Models for Earthquake Occurrences and Residual Analysis for Point Processes. *Journal of the American Statistical Association*, 83, 9–27. https://doi.org/10.1080/01621459.1988.10478560

Ogata, Y. (1998). Space-Time Point-Process Models for Earthquake Occurrences. *Annals of the Institute of Statistical Mathematics*, 50, 379–402. https://doi.org/10.1023/A:1003403601725

Ogata, Y., Matsu'ura, R.S., & Katsura, K. (1993). Fast likelihood computation of epidemic type aftershock-sequence model. *Geophysical Research Letters*, 20, 2143–2146. https://doi.org/10.1029/93GL02142

Oudin, L., Hervieu, F., Michel, C., Perrin, C., Andréassian, V., Anctil, F., & Loumagne, C. (2005). Which potential evapotranspiration input for a lumped rainfall–runoff model? *Journal of Hydrology*, 303, 290–306. https://doi.org/10.1016/j.jhydrol.2004.08.026

Petrie, E.J., King, M.A., Moore, P., & Lavallée, D.A. (2010). Higher-order ionospheric effects on the GPS reference frame and velocities. *Journal of Geophysical Research: Solid Earth*, 115. https://doi.org/10.1029/2009JB006677

Pushpalatha, R., Perrin, C., Le Moine, N., Mathevet, T., & Andréassian, V. (2011). A downward structural sensitivity analysis of hydrological models to improve low-flow simulation. *Journal of hydrology*, 411, 66–76. https://doi.org/10.1016/j.jhydrol.2011.09.034

Raftery, A.E. (1995). Hypothesis testing and model. Markov chain Monte Carlo in practice 165–187.

Rodell, M. (2016). GLDAS Noah Land Surface Model L4 monthly 0.25 x 0.25 degree, Version 2.1. *NASA Goddard Earth Sciences Data and Information Services Center*. https://doi.org/10.5067/sxavczfaqlno

Rousseeuw, P.J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics* 20, 53–65. https://doi.org/10.1016/0377-0427(87)90125-7

Schmid, R., Rothacher, M., Thaller, D., & Steigenberger, P. (2005). Absolute phase center corrections of satellite and receiver antennas. *GPS Solutions*, 9, 283–293. https://doi.org/10.1007/s10291-005-0134-x

Schmid, R., Steigenberger, P., Gendt, G., Ge, M., & Rothacher, M. (2007). Generation of a consistent absolute phase-center correction model for GPS receiver and satellite antennas. *Journal of Geodesy*, 81, 781–798. https://doi.org/10.1007/s00190-007-0148-y

Serpelloni, E., Casula, G., Galvani, A., Anzidei, M., & Baldi, P. (2006). Data analysis of Permanent GPS networks in Italy and surrounding region: application of a distributed processing approach. *Annales Geophysique*, 49, 897–928. https://doi.org/10.4401/ag-4410

Serpelloni, E., Faccenna, C., Spada, G., Dong, D., & Williams, S. D. P. (2013). Vertical GPS ground motion rates in the Euro-Mediterranean region: New evidence of velocity gradients at different spatial scales along the Nubia-Eurasia plate boundary. *Journal of Geophysical Research: Solid Earth*, 118, 6003–6024. https://doi.org/10.1002/2013JB010102

Serpelloni, E., Vannucci, G., Anderlini, L., & Bennett, R. A. (2016). Kinematics, seismotectonics and seismic potential of the eastern sector of the European Alps from GPS and seismic deformation data. *Tectonophysics*, 688, 157–181. https://doi.org/10.1016/j.tecto.2016.09.026

Utsu, T., Ogata, Y., S, R., & Matsu'ura (1995). The Centenary of the Omori Formula for a Decay Law of Aftershock Activity. *Journal of Physics of the Earth*, 43, 1–33. https://doi.org/10.4294/jpe1952.43.1

van Stiphout, T., Marsan, D., & Zhuang, J. (2012). Seismicity declustering. *Community Online Resource for Statistical Seismicity Analysis*. https://doi.org/10.5078/corssa-52382934

Wiemer, S. (2000). Minimum Magnitude of Completeness in Earthquake Catalogs: Examples from Alaska, the Western United States, and Japan. *Bulletin of the Seismological Society of America*, 90, 859–869. https://doi.org/10.1785/0119990114

Zhuang, J., Ogata, Y., & Vere-Jones, D. (2002). Stochastic Declustering of Space-Time Earthquake Occurrences. *Journal of the American Statistical Association*, 97(458), 369–380. https://doi.org/10.1198/016214502760046925