Towards a Digital Twin of the Earth System: Geo-Soft-CoRe, a Geoscientific Software & Code Repository

Irene DeFelipe¹,², Juan Alcalde¹, Eldar Baykiev³, Isabel Bernal⁴, Kittiphon Boonma¹, Ramon Carbonell¹, Stephanie Flude⁵, Arnau Folch¹, Javier Fullea⁶,⁷, Daniel García-Castellanos¹, Adelina Geyer¹, Santiago Giralt¹, Armand Hernández⁸, Ivone Jiménez-Munt¹, Ajay Kumar⁹, Maria-Gema Llorens¹, Joan Martí¹, Cecilia Molina⁴, Andrés Olivar-Castaño¹⁰,¹¹, Andrew Parnell¹², Martin Schimmel¹,Montserrat Torné¹ and Sergi Ventosa¹

¹Geosciences Barcelona, GEO3BCN-CSIC, Barcelona, Spain, ²Departamento de Geología, Universidad de Salamanca, Salamanca, Spain, ³Institute of Geosciences, Christian-Albrechts-Universität Kiel, Kiel, Germany, ⁴Consejo Superior de Investigaciones Científicas, CSIC, Madrid, Spain, ⁵Earth Sciences, University of Oxford, Oxford, United Kingdom, ⁶Departamento de Física de la Tierra y Astrofísica, Universidad Complutense de Madrid, Madrid, Spain, ⁷Dublin Institute for Advanced Studies, Dublin, Ireland, ⁸GFZ Group, Centro de Investigaciones Científicas Avanzadas (CICA), Facultad de Ciencias, Universidad de Coruña, A Coruña, Spain, ⁹German Research Centre for Geosciences (GFZ), Potsdam, Germany, ¹⁰Department of Geology, University of Oviedo, Oviedo, Spain, ¹¹Universität Potsdam, Institut für Geowissenschaften, Potsdam, Germany, ¹²Insight Centre for Data Analytics, Hamilton Institute, National University of Ireland Maynooth, Maynooth, Ireland

The immense advances in computer power achieved in the last decades have had a significant impact in Earth science, providing valuable research outputs that allow the simulation of complex natural processes and systems, and generating improved forecasts. The development and implementation of innovative geoscientific software is currently evolving towards a sustainable and efficient development by integrating models of different aspects of the Earth system. This will set the foundation for a future digital twin of the Earth. The codification and update of this software require great effort from research groups and therefore, it needs to be preserved for its reuse by future generations of geoscientists. Here, we report on Geo-Soft-CoRe, a Geoscientific Software & Code Repository, hosted at the archive DIGITAL-CSIC. This is an open source, multidisciplinary and multiscale collection of software and code developed to analyze different aspects of the Earth system, encompassing tools to: 1) analyze climate variability; 2) assess hazards, and 3) characterize the structure and dynamics of the solid Earth. Due to the broad range of applications of these software packages, this collection is useful not only for basic research in Earth science, but also for applied research and educational purposes, reducing the gap between the geosciences and the society. By providing each software and code with a permanent identifier (DOI), we ensure its self-sustainability and accomplish the FAIR (Findable, Accessible, Interoperable and Reusable) principles. Therefore, we aim for a more transparent science, transferring knowledge in an easier way to the geoscience community, and encouraging an integrated use of computational infrastructure.

Systematic Review Registration: https://digital.csic.es/handle/10261/193580.

Keywords: digital twin, software, code, global change, hazards, solid earth
1 INTRODUCTION

The implementation of computational infrastructure (e.g., software) has represented a turning point in Earth science research. It has provided tools for large-scale digital data management, quality control during acquisition, digital signal analysis, data processing and inversion, and has fostered numerical modeling, advanced visualization, statistical and mapping tools, and data archiving (e.g., Kempler and Mathews, 2017). The development of different software has been grounded on the extraordinary scientific and technological advances achieved in the last half century, the progress of digitalization, and the generation of progressively larger and more precise datasets. Computer modeling, data-intensive analysis (big data), and artificial intelligence are now integral parts of many geoscience investigations that allow extracting more information from data than traditional assimilation approaches (Bergen et al., 2019; Reichstein et al., 2019). The integration of geological and geophysical data and generation of models that simulate different geological processes, may lead to digital models of certain aspects and/or processes of the Earth system, with the ultimate aim of building a digital twin of the Earth. Digital twins combine continuous observation, modeling and simulation of certain aspects of the Earth system, resulting in accurate predictions for possible scenarios (Bauer et al., 2021a; Bauer et al., 2021b). Among other applications, a digital twin of the Earth may generate the necessary knowledge to Nativi et al. (2021): tackle global change, ease the European Green transition into a more resource-efficient and competitive economy (European Commission, 2022), and take action for the Sustainable Development Goals (United Nations, 2022).

Earth science comprises a series of interconnected disciplines (e.g., geology, geophysics, meteorology, or oceanography) that target the nature, structure and evolution of the Earth. The global aim of Earth science is to characterize the processes happening in the Earth at different spatial and temporal scales; that is the Earth in 4D. Despite the huge advances achieved in those disciplines, geoscientists have yet to meet the digital transformation needs of society, a necessary adaptation for sustainable development that requires a deeper knowledge of the dynamics of the Earth system (Reid et al., 2010; Guo et al., 2020; Nativi et al., 2021). New paradigms in Earth science are accompanied by an exponential increase in the volume of data generated from a variety of sources. Thus, technologies capable of performing increasingly complex computing operations (e.g., Hashem et al., 2015) are needed to process large amounts of data in a relatively short time and to assess their precision. However, building a digital twin of such complexity as the entire Earth system is still a long-term and ambitious challenge, despite its critical relevance for our society.

High performance computing has provided researchers with the necessary capabilities to trace and follow the dynamic of the Earth system. With a profound understanding of physical phenomena (e.g., ongoing climate change, the natural variability and forced change in the cryosphere, the evolution of biodiversity, land use, and natural resources), modeling of predictive scenarios can be available. The ever-growing supercomputer capability has improved the assessment of hazards and risks associated with extreme events, allowing to anticipate them and to estimate their implications. As an example, the EU Center of Excellence for Exascale in solid Earth (ChEESE, 2022) is preparing open-source codes and workflows to enable services on hazard assessment, urgent computing, and early warning forecast. This type of initiative may ultimately allow designing and/or planning more effective measures in case of natural disasters, more efficient environmental policies, and relevant legislative measures. Therefore, high performance computing contributes to analyzing and characterizing events that involve a high socio-economic impact.

A huge international, interinstitutional, and interdisciplinary integrative effort in Earth science is needed to set the path towards more precise models of the Earth system. World-class, advanced, and multi-scale methods will have to be reconciled and integrated to model and predict future scenarios. In this sense, the implementation of computational infrastructure represents a step-forward in Earth science, as the monitoring and simulation of natural activity will provide a successful transition to a more sustainable development (Reid et al., 2010). Aware of the needs of the geoscientific community, the European Commission (2016) fosters global open science as a driver for enabling innovative research. For instance, the European Open Science Cloud (EOSC, 2022) is driving towards a virtual environment with open and seamless services for storage, management, analysis and reuse of data. Similarly, the European Plate Observation System (EPOS, 2022) was conceived in 2002 to ensure a long-term plan to facilitate integrated use of geoscientific data. More recently, Destination Earth (DestinE, 2022) was launched in 2021 to develop a highly precise digital model of the Earth. DestinE aims to integrate digital twins, giving users a customized access to high-quality information, services, models, scenarios, forecasts, and visualizations. A successful plan for the development of digital twins requires data and software tools. Thus, the preservation of Earth system related software schemes is mandatory. For example, Computational Infrastructure for Geodynamics (CIG, 2022) is a community-driven organization that promotes the advance of Earth science by developing and disseminating software for geodynamics and related fields. Other examples are the Society of Exploration Geophysicists (SEG Open Data, 2021), an international initiative that aims to share geophysical open data and software, or Vhub (2016) which provides tools to model volcanic processes. These initiatives set the basis for archiving all the source code generated by individuals or research groups.

Regardless of these integrative initiatives, among the geoscience community there is still a lack of: 1) multidisciplinary integration of research products and software, and interoperability between different tools, code and data for analysis and processing, and 2) fully defined standardized formats understandable through the international community. Therefore, we present a comprehensive software collection that aims at: 1) integrating multidisciplinary software covering different aspects of the geosciences, 2) implementing interoperable software, as it was coded by standardized
programming languages, and applied to Earth science, and 3) setting the basis for research software management plans. This contribution presents a comprehensive overview of Geo-Soft-CoRe (Geoscientific Software & Code Repository), a geoscientific open software collection. We aim to narrow the gap between different disciplines in Earth science, enlarging the observation systems of natural processes to improve forecast, and encourage innovation in scientific and technological responses towards a more sustainable environment. The open source computational infrastructure presented here together with complementary initiatives of open data (e.g., García-Mayordomo et al., 2012; DeFelipe et al., 2021; Di Giacomo et al., 2021; Zahorec et al., 2021), will reinforce the understanding of the evolution of the Earth system to improve the predictability of future scenarios. Finally, we want to contribute to a harmonized strategy for software sharing in Earth science, making accessible tools that will help develop comprehensive digital twins of the Earth system in the future.

2 SOFTWARE AND CODE COLLECTION

STRUCTURE AND HANDLING

Geo-Soft-CoRe is stored in the hosting platform DIGITAL.CSIC, an institutional repository that comprises multidisciplinary scientific data and data products. This repository follows the institutional open access mandate, which requires deposit of peer reviewed publications and underlying data of CSIC’s authors, the international mandates of open access, and the FAIR principles: Findable, Accessible, Interoperable and Reusable (Wilkinson et al., 2016; Lamprecht et al., 2020). Figure 1 shows a workflow diagram of the functioning of the repository, including the actions and processes carried out by the three main stakeholders of the repository: the researchers or software developers, the hosting platform, and the end-users.

In a first stage, researchers/software developers are responsible for providing: 1) their software in an appropriate format (e.g., well documented and operative), 2) related documentation, and 3) quality control of the software. Researchers/software developers must provide the standardized metadata provided by DIGITAL.CSIC (2021b) which puts emphasis on contextualizing the objects being deposited. In a second stage, DIGITAL.CSIC will check the metadata completeness and will generate a DOI and a handle, making the item public, findable, citable, and interoperable by exposing it to external open databases, search engines and open science aggregators (OpenAire, DataCite Commons and BASE). Additionally, DIGITAL.CSIC enabled in 2020 the Scholix standard, a global initiative that relates DOI and metadata elements with other research outputs, aggregated by Scholexplorer (2022) and DataCite (2022). For that, DIGITAL.CSIC follows the DSpace workflow for software management (Rodríguez-Gairín and Sulé Duesa, 2008). Finally, the end-user can find and access the software via the repository and download it without restrictions, as both the metadata record and the readme file include a standard usage license that explicitly indicate the allowed uses. The information included in the repository can therefore be reused, giving to it an appropriate citation, and enabling transfer knowledge to the Earth science community.

Geo-Soft-CoRe has been developed by a wide-range of scientists from different institutions, and is multiscale, multidisciplinary, conciliatory, and interoperable. The software compiled here addresses three main research challenges: 1) global change (Section 3.1), 2) hazards (Section 3.2), and 3) structure and geodynamics of the solid Earth (Section 3.3 and Section 3.4). These challenges are in turn divided into nine targets: climate, volcanism, geological storage of CO2, seismic data processing, integrated geophysical and petrological methods, potential field, lithospheric and surface processes, lithospheric mantle buoyancy, and folding. All these targets are interconnected and pursue the

FIGURE 1 | Flow diagram showing the steps needed for researchers or software developers to upload their code in Geo-Soft-CoRe to achieve knowledge transfer following the FAIR principles.
grand challenges of the Earth sciences (Figure 2; Acocella, 2015). The observation of the Earth’s processes generates a wealth of data whose management requires the use of extensive computational resources. Geo-Soft-CoRe provides tools to manage these data, refining our understanding of the natural processes. In addition, computational infrastructure allows generation of simulations that improve the forecast of natural hazards to minimize their consequences for society and the environment. Figure 2 shows the existing feedback process, in which data collection and processing lead to the understanding of the involved processes, forecasts and resources, which in turn leads to data collection to improve the previous models. Finally, the dissemination and communication at all levels of society will encourage technological, policy, and social responses towards global sustainability (Reid et al., 2010) and will assure quality software sustainability.

Geo-Soft-CoRe comprises at present 24 software and code packages (Table 1) and can be accessed at DIGITAL.CSIC (2021a). Each software package is stored together with the most relevant information, including the title, list of authors, year, abstract, version, code language used, funding agencies, related literature, a Digital Object Identifier (DOI), and manuals, descriptions of use and/or examples of application. Complementarily, some of the software presented in this paper can be accessed through different platforms of code sharing (e.g., GitHub or researcher’s personal websites), which are also indicated in the repository. In the following sections, we provide a brief explanation of each piece of software and examples of their applicability. Geo-Soft-CoRe is the result of a fruitful national and international, inter-institutional collaboration. We are actively working on keeping this collection updated, trying to expand it with new software packages. Furthermore, we are open for international collaborations by offering a framework of data and software storage, or as an example of research software management.

3 GEOSCIENTIFIC SOFTWARE

3.1 Global Change

The North Atlantic Oscillation (NAO) is a large-scale atmospheric mode of variability that controls a large fraction of the European climate variability in winter, atmospheric circulations and weather patterns (Hurrell et al., 2003; Hernández et al., 2020a). Understanding the processes that govern the NAO variability is key to model global climate change, and has been used for important applications such as the development of better mitigation plans for agriculture or water management (Hernández et al., 2020b). To quantify its impact, the NAO index is defined as the surface sea-level pressure difference between the Subtropical High (Azores) and Subpolar Low (Iceland).

A significant number of studies have been focused on the reconstruction of past NAO variability at different timescales (e.g., Glueck and Stockton, 2001; Luterbacher et al., 2001; Trouet et al., 2009; Olsen et al., 2012; Ortega et al., 2015). Nevertheless, the information related to decadal-scale NAO evolution beyond the last millennium is scarce and inconclusive. The NAO code presented here is a Bayesian modeling approach that models decadal variability of the NAO index in the last 2,000 years (Figure 3; Hernández et al., 2020b). This code also accounted for potential external factors such as volcanic eruptions and solar activity. Additionally, the interaction of the NAO with the second most important large-scale pattern of atmospheric circulation in the North Atlantic European sector, the Eastern Atlantic (EA) pattern, is addressed. A modulated impact is predicted when...
| Software/code name | Research challenge | Software target | Depth/time scale | Application | References | Section |
|--------------------|--------------------|-----------------|-----------------|-------------|------------|---------|
| NAO (North Atlantic Oscillation) | Global change | Climate | Atmosphere/present | Variation of the North Atlantic Oscillation | Parnell et al. (2018) | 3.1 |
| HAZMAP-2.4.4 | Hazards | Volcanism | Atmosphere and surface/present | Transport and deposition of particles (2D) | Macedonio et al. (2016) | 3.2.1.1 |
| FALL3D-8.0 | Hazards | Volcanism | Atmosphere and surface/present | Transport and deposition of particles (3D) | Folch (2021) | 3.2.1.1 |
| TWODEE-2.3 | Hazards | Volcanism | Atmosphere/present | Dispersion of gases | Folch et al. (2021) | 3.2.1.2 |
| SSC (Storage Security Calculator) | Hazards | Geological storage of CO₂ | Sedimentary basin/present | Quantification of CO₂ immobilization in a storage | Alcalde et al. (2018b) | 3.2.2 |
| Corr_stack_v04 (cross-correlation and stack, the phase coherence approach) | Crust and upper mantle characterization | Seismic data processing | Unrestricted | Cross-correlation and stack of event and/or ambient noise data | Schimmel (2020) | 3.3.1.1 |
| FastPCC (Fast Phase Cross-Correlation) | Crust and upper mantle characterization | Seismic data processing | Unrestricted | Interstation correlation for seismic ambient noise processing | Ventosa and Schimmel (2019) | 3.3.1.2 |
| ts-PWS (time-scale Phase-Weighted Stack) | Crust and upper mantle characterization | Seismic data processing | Unrestricted | Time-scale phase-weighted stack | Schimmel and Ventosa (2017) | 3.3.1.3 |
| Pofile | Crust and upper mantle characterization | Seismic data processing | Unrestricted | Time-frequency dependent polarization | Schimmel and Stutzmann (2019) | 3.3.1.4 |
| TS_PWS0_UG_1.5 | Crust and upper mantle characterization | Seismic data processing | Unrestricted | Group velocities of surface waves | Schimmel et al. (2018) | 3.3.1.5 |
| Rfun (Receiver functions) | Crust and upper mantle characterization | Seismic data processing | Lithosphere/present | Analysis of receiver functions | Olivar-Castaño (2021) | 3.3.1.6 |
| LitMod2D (Lithospheric Modeling) | Crust and upper mantle characterization | Integrated geophysical and petrological methods | Lithosphere/present | Physical properties of the lithosphere/upper mantle | Kumar et al. (2019) | 3.3.2 |
| LitMod3D (Lithospheric Modeling) | Crust and upper mantle characterization | Integrated geophysical and petrological methods | Lithosphere/present | Physical properties of the lithosphere/upper mantle | Fullea et al. (2019a,b) | 3.3.2 |
| FA2BOUG | Crust and upper mantle characterization | Potential field | Lithosphere/present | Bouguer gravity anomaly correction | Fullea (2008) | 3.3.3.1 |
| Geoid-elevation inversion | Crust and upper mantle characterization | Potential field | Lithosphere/present | Lithospheric density and thermal structure | Fullea et al. (2019c) | 3.3.3.2 |
| Magnetic tesseroids | Crust and upper mantle characterization | Potential field | Lithosphere/present | Induced and remanent magnetic field | Baykiev et al. (2020a) | 3.3.3.3 |
| Global magnetic inversion | Crust and upper mantle characterization | Potential field | Lithosphere/present | High-resolution magnetic inversion | Baykiev et al. (2020b) | 3.3.3.3 |
| Spilover | Geodynamics | Lithospheric and surface processes | Hydrosphere/Pleistocene to present | Landscape evolution of an overtopping lake | García-Castellanos (2018) | 3.4.1.1 |
| tAo | Geodynamics | Lithospheric and surface processes | Lithosphere to surface/Mesozoic to Pleistocene | 2D interplay between lithospheric flexure and surface transport | García-Castellanos (2015a) | 3.4.1.2 |
| TSC (Tectonics, Isostasy, Surface transport, and Climate) | Geodynamics | Lithospheric and surface processes | Lithosphere to surface/Mesozoic to Pleistocene | 3D interplay between lithospheric flexure, fault deformation and surface transport | García-Castellanos (2015b) | 3.4.1.3 |
| UhuruTSC | Geodynamics | Lithospheric and surface processes | Lithosphere to surface/Mesozoic to Cenozoic | Topography evolution integrating tectonics, river transport and climate Controls of buoyancy during convergence | Garcia-Castellanos and Jiménez-Munt (2015b) | 3.4.1.4 |
| LithBuoy (lithospheric mantle buoyancy) | Geodynamics | Lithospheric mantle buoyancy | Lithospheric mantle/Unrestricted | (Continued on following page) | | |
concentrations of CO₂ emitted from volcanic (but also non-atmospheric conditions and depending on the topography, active volcanic areas. For example, under favorable eruptive processes can also jeopardize the environment in areas (e.g., Bevilacqua et al., 2021). Added to these, non-syn-Aravena et al., 2020), and often affecting highly populated regional or even global (Bartolini et al., 2017; Martí, 2017; Aravena et al., 2020), and often affecting highly populated areas (e.g., Bevilacqua et al., 2021). Added to these, non-syn-eruptive processes can also jeopardize the environment in active volcanic areas. For example, under favorable atmospheric conditions and depending on the topography, concentrations of CO₂ emitted from volcanic (but also non-volcanic) diffuse sources can reach high local concentrations, putting humans and animals at risk (Costa et al., 2008; Folch et al., 2020). The understanding and modeling of the volcanic activity and the quantification of related volcanic hazards is thus pivotal in terms of prevention, mitigation and emergency preparedness before a hazard can occur (Vujicic-Lugassy and Frank, 2010). Geo-Soft-CoRe comprises three long-standing open-source community codes that aim to simulate different volcanic processes: HAZMAP-2.4.4, FALL3D-8.0 (Section 3.2.1.1) and TWODEE-2.3 (Section 3.2.1.2).

3.2.1 Transport and Deposition of Particles
HAZMAP-2.4.4 is a 2D analytical code to solve the equation of advection, diffusion and sedimentation (ADS) of small-size particles dispersed in the atmosphere from an explosive eruption column (Macedonio et al., 2005). The model simplifies the ADS equation for volcanic particles transport in the atmosphere using a semi-analytical approach that reduces the required computer time and memory. Complementary, FALL3D-8.0 is a 3-D time-dependent Eulerian model for atmospheric passive transport and deposition of particles, aerosols, and radionuclides (Folch et al., 2020). This code solves the full ADS equation and is designed to forecast particle mass concentration in the atmosphere, loading at ground, and transport of any kind of airborne solid particle (Figure 4A). The outputs of these software can be combined to generate probabilistic hazard maps of ash clouds, tephra dispersal, and ash fallout accumulation on the ground (Bonasia et al., 2018; Pore et al., 2018; Poulidis et al., 2019; Vázquez et al., 2019; Folch et al., 2020). The combination of HAZMAP 2.4.4 and FALL3D-8.0 is more useful to assess the hazard of ash fallout in a wider variety of scenarios than applying them individually, as FALL3D-8.0 is valid also within the atmospheric boundary layer (Macedonio et al., 2008; Selva et al., 2014; Folch et al., 2020).

3.2.1.2 Dispersion of Gases
Complementary to HAZMAP-2.4.4 and FALL3D-8.0 which simulate particle dispersion, TWODEE-2.3 is a code to compute dispersion of gases denser than air (e.g., CO₂) released from natural sources in complex terrains (Hankin and Britter, 1999a; Hankin and Britter, 1999b; Hankin and Britter, 1999c; Chiodini et al., 2010). This code solves the shallow water equations for fluid depth, depth-averaged horizontal velocities, and depth-averaged fluid density. TWODEE-2.3 can be used for forecasting gas dispersion near the ground and/or for hazard assessment over complex terrains. The outputs are maps of CO₂ flux (Figure 4B) and its evolution with time (Folch et al., 2009; Folch et al., 2017) that can be potentially applied for the CO₂ concentration distribution under different atmospheric conditions, and assessment of hazard related to CO₂ leakage in a geological storage site.

3.2.2 Geological Storage of CO₂
Carbon capture and storage (CCS) includes a wide range of technologies developed to safely dispose of the CO₂ produced.
during power generation and industrial processes, rather than being emitted to the atmosphere (Metz et al., 2005). CCS ensures that the CO\(_2\) generated during these operations is immobilized underground (Figure 5A) and does not contribute to anthropogenic climate change. CCS plays a significant role in the vast majority of emissions reduction scenarios that have the potential to meet net-zero by mid-century, because it is the only technology currently capable of decarbonizing many essential industries, and because of its potential for negative emissions through biomass energy CCS and direct air capture (Alcalde et al., 2018c; IEA, 2020). However, concerns about CO\(_2\) leaking from the storage site still exist, with negative consequences for CCS implementation (Scott et al., 2015). To address these concerns, Alcalde et al. (2018a) developed the Storage Security Calculator (SSC). The SSC is a numerical program to realistically evaluate how secure a geological CO\(_2\) storage is at the global industry scale, and over the tens of thousands of years. When CO\(_2\) is injected into the subsurface, it is subject to several physicochemical processes that immobilize a significant proportion of the CO\(_2\), enhancing the security of the storage. SSC quantifies this immobilization, along with likely CO\(_2\) leakage rates via wellbores and geological structures and calculates the proportion of CO\(_2\) remaining in the storage reservoir over 10,000 years. The model can be run as a “base case” scenario, where single values are used for each input parameter, or as a Monte-Carlo analysis, to incorporate uncertainty on the input parameters. Model outputs are the amount of CO\(_2\) leaked, immobilized, and retained in the storage reservoir over time (Figure 5B), and Monte-Carlo outputs report the 5th, 50th, and 95th percentile of the model results. This code has been applied to three industry-scale scenarios, that assess the differences in storage security between onshore and offshore CO\(_2\) storage, and between implementing CCS in the context of well-regulated and poorly-regulated hydrocarbon industries. The code can be adapted to different regional implementation scenarios, as a method to quantify storage security in a geological time scale.

### 3.3 Crust and Upper Mantle Characterization

#### 3.3.1 Seismic Data Processing

The different strategies to image and constrain the structure of the lithosphere require a deep understanding of wave propagation mechanisms and seismic sources. In principle, we can use the seismic wavefield generated by any source to study the structure of the lithosphere (Dias et al., 2015; Haned et al., 2016; Corela et al., 2017; Poveda et al., 2018; Romero and Schimmel, 2018; Andrés et al., 2020; Núñez et al., 2020; Olivar-Castaño et al., 2020; Ayarza et al., 2021) and to monitor natural and anthropogenic activity (Díaz et al., 2017; Lecocq et al., 2020; Díaz et al., 2022). In that respect, the use of seismic noise is progressively gaining more importance in Earth science (Campillo and Roux, 2015; Maciel et al., 2021) as it is ubiquitous in time and space (Stutzmann et al., 2009; Schimmel et al., 2011a). With seismic noise we usually refer to a diffuse and scattered wavefield which traditionally has been discarded as it interferes with ballistic signals from earthquakes and controlled sources. Seismic noise does not require a deterministic source of energy (e.g., earthquakes, explosions, etc) and represents a more cost-efficient and sustainable method in comparison with controlled source approaches. Nevertheless, noise studies cannot replace traditional active source and earthquake seismology in all type of applications but is very useful for imaging and monitoring shallow structures. The seismic processing tools which we describe below, have been designed to extract weak-amplitude signals hidden by other less coherent signals. The methods are independent, data-adaptive
and have been used for a broad range of data types and applications, mainly in exploration and earthquake seismology. Presently, the approaches are mostly employed to process seismic noise data for the different interferometry applications, and analysis of receiver functions from teleseismic sources.

To improve the quality of seismic signals and progress in our understanding of all the information contained in a seismogram, we develop tools to identify and extract more signals and efficiently process large volumes of seismic ambient noise data. A cross-correlation is theoretically sufficient to get the Empirical Green’s Function (EGF). In practice, we improve EGF convergence and signal-to-noise ratio using more elaborated processing flows (Bensen et al., 2007; Schimmel et al., 2011b; Moreau et al., 2017; Ventosa et al., 2017; Schimmel et al., 2018) organized in: 1) preprocessing, that may include instrument response correction, anomalous signal rejection, and spectral whitening, usually done with programs like ObsPy (Krischer et al., 2015), SAC (Goldstein and Snoke, 2005), or MSNoise (Lecocq et al., 2014); 2) correlation, a geometrically-normalized, 1-bit (Bensen et al., 2007) or phase (Schimmel, 1999; Schimmel et al., 2011b; Ventosa et al., 2019) correlation of many short data sequences; and 3) stacking, a sum of correlations that may include weights and denoising (Schimmel and Gallart, 2007; Ventosa et al., 2017). Aside, we include other tools, for instance to robustly measure group velocities of surface waves extracted from noise cross-correlations (Haned et al., 2016; Schimmel et al., 2017; Nuñez et al., 2020), and to characterize elliptically polarized signals within the noise wave field and corresponding sources (Schimmel et al., 2011a; Davy et al., 2015; Carvalho et al., 2019).

3.3.1.1 Cross-Correlation and Stack of Event and/or Ambient Noise Data
Ambient noise data are routinely used to extract EGFs for seismic noise monitoring and imaging studies with a wide range of applications, e.g., fault and volcano monitoring (Wegler and Sens-Schönfelder, 2007; Brenguier et al., 2008; D’Hour et al., 2016; Sánchez-Pastor et al., 2019) and imaging structures at different scales (mapping discontinuities, seismic ambient noise tomography, Shapiro and Campillo, 2004; Haned et al., 2016; Romero and Schimmel, 2018; Andrés et al., 2020, among others). The Corr_stack_v04 software package (Schimmel et al., 2011b) is coded to cross-correlate and stack event coda and/or ambient noise data. This software package contains additional programs to compute the phase cross-correlation, PCC (Schimmel, 1999), and the time-frequency domain phase weighted stacking, ts-PWS (Schimmel and Gallart, 2007), that uses the instantaneous phase coherence obtained through analytical signal theory. The software package also contains a code for data preprocessing (time and frequency domain normalization) and permits to extract the EGFs using the classical approach for comparison.

3.3.1.2 Interstation Correlation for Seismic Ambient Noise Processing
A basic operation in seismic ambient noise studies is the correlation between stations, which can be distributed locally, regionally or around the world. Cross-correlation methods measure the similarity between datasets from pairs of stations as a function of lag time. Instead of conventional correlation-based methods that favor the most energetic components, the phase cross-correlation (Schimmel, 1999) is amplitude unbiased, which helps us to detect weak signals, increase convergence, improve waveform coherence (Schimmel et al., 2011b), and entails a minor increase in computational cost (Ventosa et al., 2019). The FastPCC software package computes interstation correlations, including fast implementations of geometrically-normalized, 1-bit and phase correlations.

3.3.1.3 Time-Scale Phase-Weighted Stack
The stack of cross-correlations enhances coherent signals from surface and body waves and attenuates random noise. The time-scale phase-weighted stack (ts-PWS) is a software package based on a physically intuitive denoising method (Ventosa et al., 2017) that uses the phase coherence of the signals contained in the cross-correlations to filter the EGF and accelerates its convergence. In addition, this software package implements the two-stage stacking and unbiased phase coherence strategies to reduce signal attenuation and further increase noise attenuation. These methods can be used to obtain clean surface waveforms, key to extract accurate dispersion
observations (e.g., velocity variation with frequency) and constrain subsurface structure.

3.3.1.4 Time-Frequency Dependent Polarization
Seismic polarization studies exploit the vector character of the seismic wave field through analyzing the recorded ground motion in 3-component seismograms simultaneously. The polarization approaches measure the shape and orientation of particle motion to determine wave type and direction of propagation. They are usually employed to filter seismic records, to identify signals, and to characterize the wave field. The PolPre program determines the polarization attributes through a time-frequency dependent eigen-decomposition of spectral covariance matrices and is mainly used to characterize the wave field and to identify signals. The approach is based on the degree of polarization, which measures the stability, or robustness of any arbitrary polarization as presented in Schimmel and Gallart (2003), Schimmel and Gallart (2004). This program can be used to identify Rayleigh waves in seismic noise (e.g., Schimmel et al., 2011a; Davy et al., 2015; Carvalho et al., 2019) and to identify Rayleigh waves from seismic events. The data mining of the identified signals and polarization attributes as function of time and frequency, permits the detection of seasonal changes and to find the directions of the noise sources.

3.3.1.5 Group Velocities of Surface Waves
The arrival times of seismic surface waves at a sensor depend on their frequency. This frequency dispersion is caused by the frequency-dependent depth sensitivity of surface waves and the variation of the seismic velocities with depth. Dispersion measurements are used to constrain the underlying structure, for instance, through tomographic inversions. With this application in mind, Ts_PWS0_UG_1.5 was coded to provide a new strategy for the semi-automated measurement of seismic group arrival times or group velocities. The program includes a fast and efficient algorithm of the time-frequency phase weighted stack (Schimmel and Gallart, 2007) based on wavelet theory as presented in Ventosa et al. (2017). Thus, this software employs the phase coherence and uses resampling strategies to robustly measure group velocities as function of frequencies and instantaneous frequencies (Schimmel and Ventosa, 2017).

3.3.1.6 Analysis of Receiver Functions
Seismic waves are not only reflected and refracted in the discontinuities of the subsurface, but P-waves are also converted to S-waves (Bath and Stefansson, 1966). Receiver functions are temporal series formed by P-to-S converted phases. The analysis of the time span between the direct P-wave arrivals and the P-to-S converted phases provides information on the depth of the discontinuity in which this conversion took place, e.g., the crust-mantle transition or upper mantle discontinuities (Neal and Pavlis, 1999). Receiver functions allow to constrain a detailed image of the crustal thickness and is a common methodology to unravel the structure of orogens (Zhu and Kanamori, 2000; Lombardi et al., 2008; Mancilla and Diaz, 2015; Diaz et al., 2018; Chiarabba et al., 2020). Here, we present the novel code Rfun, a GUI-based code aimed at providing an interactive tool to compute and analyze teleseismic P receiver functions. Rfun allows the user to select events automatically based on different parameters (e.g., magnitude, distance) and cut P-wave windows using locally stored data. The automatically cut P-wave windows can then be used to compute receiver functions using the waterlevel deconvolution method (Langston, 1979). Rfun offers two methods for analyzing receiver functions: 1) a semblance-weighted stacking method (Zhu and Kanamori, 2000; Eaton et al., 2006) aimed at estimating crustal thickness (H), and the P-to-S wave velocity ratio (k), and 2) a common conversion point (CCP) stacking algorithm, useful to image the geometry of subhorizontal seismic discontinuities below seismic stations. Rfun also provides images of the computed receiver functions, the H-k stack and the CCP cross-sections.

3.3.2 Integrated Geophysical and Petrological Methods
Understanding the present-day physical state (temperature and chemical composition) and architecture of the lithosphere and the upper mantle are grand challenges that require the integration of different geophysical, mineralogical, and petrological data (Deen et al., 2006; Griffin et al., 2009; Priestley and McKenzie, 2013; Afonso et al., 2016; Fullea et al., 2021). Nevertheless, deriving a coherent model of the physical state of the lithosphere and upper mantle is difficult due to the different sensitivities and resolutions of the various observables and methods at hand. Therefore, discrepancies in the model predictions based on different datasets still exist (Fullea et al., 2007; Afonso et al., 2008; Kumar et al., 2019).

LitMod2D (Afonso et al., 2008) is a finite-element forward modeling code developed to study the thermal, compositional, density, and seismological structure of the lithospheric and sublithospheric mantle. The code combines data from petrology, mineral physics, and geophysical observables within a self-consistent thermodynamic framework, thus reducing uncertainties arising from modeling individual observables and constraining the thermal and compositional structure of the lithospheric and sublithospheric mantle. LitMod2D has been applied to different geological settings, like the India-Eurasia collisional zone (Tunini et al., 2016), the Iberian Peninsula and its margins (Carballo et al., 2015; Pedreira et al., 2015), the Namibian volcanic margin (Fernández et al., 2010), the Zagros Mountains (Tunini et al., 2014), or the Western Carpathian (Tašárová et al., 2009). The latest update, LitMod2D_2.0 (Kumar et al., 2020) includes novel features: 1) a new graphic user interface; 2) the possibility of defining sublithospheric mantle anomalous bodies in terms of temperature, seismic velocities and their combination with the chemical composition; and 3) a post-processing tool-box that allows integration and calculation of additional observables (e.g., flexure, phase-diagrams, receiver functions, surface wave dispersion curves). Figure 6 shows an example of the thermal, density and P-wave seismic velocity distribution of the lithosphere and sublithospheric mantle. LitMod2D_2.0 has been used to study the opposite symmetry of the lithosphere
structure and the subducted paleo-Tethys in the Alboran and Algerian basins (Kumar et al., 2021), and the upper mantle anomalies below the Gibraltar arc and its topographic response (Jiménez-Munt et al., 2019).

3D versions of LitMod are also available for forward lithospheric modeling (LitMod3D, Fullea et al., 2009). The latest versions of LitMod3D incorporate an optimized thermal solver, the possibility to model crustal lithology, and a new module to integrate surface wave data. LitMod3D has been applied to unravel the structure and evolution of the lithosphere-asthenosphere boundary in the Atlantic-Mediterranean Transition Region (Fullea et al., 2010), the thermochemical structure of the mantle in Central Europe (Alasonati Tašárová et al., 2016), the Canary Islands (Fullea et al., 2015), the Scandinavian craton (Gradmann et al., 2013), or the Antarctic (Pappa et al., 2019).

3.3.3 Potential Field
3.3.3.1 Bouguer Gravity Anomaly
Density distribution in the Earth is one of the most important physical parameters to determine the structure and dynamics of the Earth’s interior. Lateral variations in density are responsible for gravity anomalies (e.g., Ayala et al., 2016), which are defined by the difference between the measured gravity at a particular location and the theoretical gravity given by a reference Earth model (e.g., Götze, 2011). Measured gravity data contain the effects of latitude, Earth tides, instrumental drift, distance from the reference ellipsoid, and masses between the actual topography and the reference ellipsoid. Bouguer gravity anomalies are obtained after applying several corrections to measured data, in particular the effects of the topographic masses (Bouguer slab plus slab curvature and terrain, e.g., Bullard A, B and C corrections respectively, Mallick et al., 2011). In this way the gravity effect of topography is reduced whereas deep crustal density variations are enhanced.

FA2BOUG (Fullea et al., 2008) is a code intended to calculate the complete Bouguer anomaly (Bullard A, B, and C) in continental and offshore environments from gridded free-air anomaly and elevation datasets (satellite or terrestrial). The code computes the complete Bouguer correction in several spatial domains according to the distance between the surrounding topography and the calculation point. FA2BOUG has provided a complete Bouguer anomaly map of the Atlantic-Mediterranean transition zone (Fullea et al., 2008), the Arabia-Eurasia collision zone in Iran (Jiménez-Munt et al., 2012), the Argentinian continental margin (Pedraza De Marchi et al., 2014), Central Tunisia for hydrogeological purposes (Azaiez et al., 2011), or the Tibetan Plateau with implications to large earthquakes hazard assessment (Wu and Gao, 2019).

3.3.3.2 Geoid-Elevation Inversion
Lateral and vertical subsurface density variations cause changes in the local gravity field. The geoid is the equipotential surface that coincides with the average sea level used as a reference to determine surface elevations (Fraczek, 2003). The lateral variations in the geoid (e.g., geoid anomalies with respect to the reference ellipsoid) are related to the mass distribution within the Earth. The pressure at the base of every vertical lithospheric column is determined by integrating the lithospheric density down to a certain compensation level. Local isostasy requires that the pressure at the base of the lithospheric column remains constant. For a wide range of relevant topographic wavelengths at lithospheric scale, local isostasy is a good approximation. Under certain general approximations, lateral geoid variations of lithospheric origin (typically wavelengths <4,000 km) can be described using a mass dipolar formula depending only on the vertical density distribution within the lithospheric column (e.g., Turcotte and Schubert, 2002). Therefore, it is possible to estimate the first order lithospheric density structure combining geoid anomaly and surface topography.

Geoid-elevation inversion is a code to simultaneously invert lithospheric geoid anomalies and surface elevation for the 1D crustal and lithospheric thickness coupled with temperature and density distributions (Fullea et al., 2006; Fullea et al., 2007). This software has been applied to map the crust-mantle boundary and the lithosphere-asthenosphere boundary in the Gibraltar Arc System, Atlas Mountains and adjacent areas (Fullea et al., 2007), the Iberian Peninsula (Torne et al., 2015), the Arabia-Eurasia collision zone in Iran (Jiménez-Munt et al., 2012), central
Eurasia (Robert et al., 2017), the African mainland (Globig et al., 2016), and in Alaska and surroundings shelves (Torne et al., 2020).

### 3.3.3.3 Magnetic Field
The lithospheric magnetic field reflects the tectonic setting and different geological domains of the Earth. One of the best-known examples are the magnetic anomalies in the ocean seafloor spreading centers, which have allowed dating the oceanic crust, setting the basis for global plate tectonic reconstructions (Müller et al., 2008; Granot, 2015). Magnetic anomalies are often constrained from satellite and airborne magnetic measurements. Despite the huge advances in determining the magnetization by those methods, they do not provide reliable magnetic anomaly predictions close to the Earth’s surface as the grid becomes finer than the dipole distance.

For easy testing geological lithospheric models against satellite and airborne magnetic measurements, the magnetic tesseroids code uses spherical prisms, the so-called tesseroids (Heck and Seitz, 2007) as a magnetic source. Therefore, this software calculates induced and remnant magnetic fields at global and regional scales and has been successfully applied in Fennoscandia and Central Africa (Baykiev et al., 2016).

A further step in the characterization of the magnetic field encompasses a global high-resolution magnetic field inversion, providing a global heterogeneous coverage based on a spherical harmonic representation. Baykiev et al. (2020c) developed the global magnetic inversion code for global high-resolution inversion of satellite magnetic data using spherical harmonic models of tesseroids. Finally, the global magnetic inversion code can be easily linked to LitMod (see Section 3.3.2) to provide a more precise model of the lithosphere, proving its versatility.

### 3.4 Geodynamics

#### 3.4.1 Lithospheric and Surface Processes
The current topography of the Earth is the result of complex interactions between lithospheric-scale tectonics, sediment transport and climate. However, the feedback mechanisms and interplay among these natural processes is not yet fully understood. With the aim of reproducing the interactions between these processes in a multiscale approach (Figure 7A), we include here four codes: Spillover (Section 3.4.1.1), tAo (Section 3.4.1.2), TISC (Section 3.4.1.3) and UhuruTISC (Section 3.4.1.4).

#### 3.4.1.1 Landscape Evolution of an Overtopping Lake
Landscape evolution is highly influenced by river erosion over geological timescales, but rock erodibility is yet poorly understood. Specifically, megafloods or outburst floods, are catastrophic phenomena that can leave a significant erosion imprint in the landscape and represent a natural hazard (Burr et al., 2009). Spillover is a code that calculates the feedback between incision (spillway erosion) and water flow (discharge) at the outlet of an overtopping lake (Figure 7B). It is coded to evaluate the rock erodibility under well-constrained hydraulic conditions. Its use has allowed us to link quantitatively the erodibility derived from geologically recent settings with those obtained from long-term landscape evolution (Garcia-Castellanos and O’Connor, 2018). The program requires as an input the lake’s area and hypsometry (water depth distribution) and it yields a range of peak flooding discharge depending on rock erodibility at the outlet. Spillover has been applied to model the Lake Bonneville megafllood (Abril-Hernández et al., 2018), and 86 other floods caused by the overtopping of lakes (Garcia-Castellanos and O’Connor, 2018), as well as in catastrophic floods in the past (Garcia-Castellanos et al., 2009).

#### 3.4.1.2 2D Interplay Between Lithospheric Flexure and Surface Transport
The interplay between lithospheric flexure and surface transport, particularly during mountain building and foreland basin formation, the influence of the climate during the formation of high reliefs (Garcia-Castellanos, 2007), and the role of tectonic lakes (Garcia-Castellanos, 2006) are interconnected processes that shape the current topography. tAo is a software that provides 2D sections using 1D lithospheric flexure calculations with various rheological models, in combination with prescribed fault kinematics, other types of isostatic loads, and erosion/deposition/transport models. This code requires a choice of elastic thickness, fault geometry and velocity, rock erodability, climatic conditions and other parameters, and yields a section of topography, sedimentary thickness, moving tectonic blocks, and erosion-sedimentation rates. tAo has been applied to different foreland basins such as the Piedmont basin in Italy (Carrapa and Garcia-Castellanos, 2005), the Ebro and Guadalquivir Cenozoic basins in Spain (Gaspar-Escribano et al., 2001; Garcia-Castellanos et al., 2002), and the Zagros-Mesopotamian basin (Saura et al., 2015).

#### 3.4.1.3 3D Interplay Between Lithospheric Flexure and Surface Transport
Since 2D modeling approaches, such as tAo, fail to capture the importance of 3D sediment transport and hydrology during drainage and sedimentary basin formation, TISC (Tectonics, Isostasy, Surface transport, and Climate) was coded as the pseudo 3D version of tAo. TISC simulates the evolution of large-scale sediment transport together with prescribed fault-tectonic deformation and lithospheric isostatic movements (viscous or viscoelastic thin plate model) over geological time scales (Garcia-Castellanos et al., 2002; Garcia-Castellanos and Jiménez-Munt, 2015a). This code requires a choice of elastic thickness, fault geometries, rock erodability, climatic conditions and other parameters and yields maps of topography, sedimentary thickness, and erosion-sedimentation rates. TISC has been successfully applied to different geological settings such as: the evolution of the Paleotethys sedimentary basins during the Messinian Salinity Crisis (Bartol et al., 2012), the drainage evolution of the Ebro Cenozoic basin (Figure 7C; Garcia-Castellanos et al., 2003), the evolution drainage of the Rio Grande extensional Basin and the East African Rifts (Berry et al., 2019), and the evolution of foreland basins during
tectonic collision (García-Castellanos, 2002; Amadori et al., 2018).

3.4.1.4 Topography Evolution Integrating Tectonics, River Transport and Climate

To further investigate tectonic processes at crustal and lithospheric scales; the lithospheric deformation, the surface and climatic processes and the thermal and mechanical evolution during continental collision need to be integrated. 

*UhuruTISC* is a program resulting from two independent codes: *Uhuru* (Jiménez-Munt et al., 2001) and *TISC* (Section 3.4.1.3), coded to couple the thermo-mechanic lithospheric deformation with surface and climatic processes (Jiménez-Munt et al., 2005a; García-Castellanos and Jiménez-Munt, 2015b). *UhuruTISC* calculates 3D lithospheric deformation using the thin-sheet approach, which assumes a vertically averaged rheology and allows simulation of the lithosphere as a thin viscous layer subjected to plane stresses. This model accounts for isostatic and potential energy effects due to crustal and lithospheric thickness variations. *UhuruTISC* has been successfully applied to neotectonics processes in the Africa-Eurasia plate boundary, from Azores to Gibraltar (Jiménez-Munt et al., 2001), large-scale deformation of the Tibetan Plateau formation (Jiménez-Munt and Platt, 2006), and the postcollisional deformation and the present tectonic field in the Alps (Jiménez-Munt et al., 2005b). In addition, it has been used to study the feedbacks between orographic precipitation and inherited tectonic structures during the
3.4.2 Lithospheric Mantle Buoyancy

The driving mechanisms for plate tectonics are thought to be controlled by the relative buoyancy between the lithospheric mantle and the underlying asthenosphere. To investigate the influence of buoyancy during incipient plate convergence, LithBuoy is a 2D thermal-diffusive model of plate convergence that accounts for different chemical compositions and tectonothermal ages of the lithosphere. This code has been applied in three types of continental lithosphere and two types of oceanic lithosphere, and the theoretical models were also quantitatively compared to three geological scenarios: Alborán Sea, Zagros and Tibet (Boonma et al., 2019b). To illustrate the results of this code, Figure 7E shows the model output with the density difference relative to an initial density distribution during the subduction of a continental lithosphere.

3.4.3 Folding

Folding is a dynamic process that results from compressional or shear stress in a ductile domain. Fold patterns provide constraints on the deformation conditions, the stress field and the viscosity contrast between the folded layers and the surrounding rock matrix (Llorens et al., 2013b; Schmalholz and Mancktelow, 2016; Llorens, 2019b). To understand the development of folds, different field and analytical studies, analogue experiments, and numerical models have been carried out in the last 50 years (Ramsay and Huber, 1987; Tikoff and Peterson, 1998; Schmalholz and Podladchikov, 2000; Llorens et al., 2013a; Llorens et al., 2013b; Bulnes et al., 2019). However, most of these studies did not consider the oblique orientation of the natural beds with respect to the strain axes. To unravel the mechanical evolution of large-amplitude folding both in pure (e.g., co-axial) and simple (e.g., non-coaxial) shear and to reproduce the stress and strain evolution, we present FoldRock. It includes a set of numerical simulations run with the software package Elle (2022) (Jessell et al., 2001; Bons et al., 2007; Piazolo et al., 2019) including the finite element package BASIL (Housman et al., 2008). This simulation package allows: 1) observing the geometrical softening of rocks, 2) determining the efficiency of folding under different scenarios, and 3) new constraints during fold evolution that can be applied to strain analysis in the field. This code has been applied to study deformation of lithospheric layers around the crust-mantle boundary (Llorens, 2019b). Figure 8 shows the workflow of the simulations and an example of folding of a competent layer embedded in a softer matrix under simple shear boundary conditions.

4 DISCUSSION AND GENERAL REMARKS

New paradigms in Earth science are driving the research community towards a more sustainable development through monitoring, modeling and forecasting natural processes. Current international initiatives aim for digital twins of different aspects of the Earth system, in which computational infrastructure is essential. Aware of this, geoscientists have been developing a wide variety of software and code to process and manage large amounts of data, generate models and simulations of natural processes, and improve the predictability of future scenarios. Its success is thus reliant on making this software easily accessible for the research community and assuring its sustainability. The integration of data and computing infrastructure facilitates modeling the Earth system, and thus, addressing global challenges of environmental sustainability and the digital transformation of society (Nativi et al., 2021).

We presented here Geo-Soft-CoRe, a multidisciplinary collection of research software intended to facilitate accessibility and knowledge exchange in Earth sciences. Geo-Soft-CoRe comprises 24 pieces of software and code that were developed to target different aspects of the Earth system in a multiscale and multidisciplinary approach. The research challenges addressed by these software focus on: 1) global change, 2) hazards, and 3) crust-upper mantle characterization and geodynamics of the solid Earth. In turn, these challenges are classified into nine specific targets: climate, volcanism, geological storage of CO₂, seismic data processing, integrated geophysical and petrological methods, potential field, lithospheric and surface processes, lithospheric mantle buoyancy, and folding. Through DIGITAL.CSIC, we provide the software and code with a DOI,
achieving the FAIR principles for research software and promoting its self-sustainability. Geo-Soft-CoRe complements previous initiatives of open data to promote a more transparent science (e.g., DeFelipe et al., 2021), and provides tools to implement best practices for scientific publications (Gil et al., 2016). The aim behind Geo-Soft-CoRe is not to substitute any current repository (e.g., GitHub) but to gather, curate and display geoscientific software to make it more visible and accessible.

5 OPEN QUESTIONS AND FURTHER STEPS

Making research products (data and software) accessible is key to ensure the advancement of Earth science. Data and metadata granularity, usage of community vocabularies, and rich supporting documentation are essential components in the FAIRification process, and software developers need to be aware of their importance. The FAIR principles also refer to the software hosting infrastructures that need to be supported and maintained. Therefore, both researchers/software developers and repositories need to enhance the visibility to promote reuse of the research software, and to enable computational reproducibility and transparency (Ramachandran et al., 2021) to speed up scientific advance and innovation. Within the framework of the EOSC Synergy Project (2022), DIGITAL.CSIC is developing an evaluation tool that will automatically assess the FAIRness of their digital objects to help: 1) researchers to better describe and manage their research products and, 2) repository administrators to identify technological issues for further development. Certainly, the monitoring of how research products are reused is a challenge to establish new innovative initiatives. To ease this monitoring, DIGITAL.CSIC will set up an API (Application Programming Interfaces) REST (REpresentational State Transfer) for a programed access to the data and software that hosts.

Geo-Soft-CoRe intends to be a dynamic collection, open to grow by the addition of new geoscientific software. For that reason, we aim at enlarging this collection by including software addressing other pressing research challenges in Earth science, for instance, planetary science (Azari et al., 2016), paleoclimate and climate science for a sustainable development (Asrar et al., 2012; Ludwig et al., 2019), and/or environmental quality and its impact on the economy and society (e.g., Ike et al., 2020; Usman, 2022), as well as engage citizens within Earth sciences (Lee et al., 2020). In addition, with this initiative we want to meet the Agenda 2030 for the Sustainable Development Goals by ensuring public access to information (e.g., Bernal, 2021).

Open access data and software will help to produce a more transparent science, ease the knowledge transfer for the next generations of geoscientists, and set the ground for developing improved models of the Earth system. We encourage an integrated use of geoscientific data and computational infrastructure fostering research innovation, and we are open to new collaborations, representing an example of scientific software management. In the longer term, advances in the understanding of the different processes in the Earth system, coupled with the improvement in computing power will enable the integration of some of the tools hosted in Geo-Soft-CoRe. This in turn will bring the geoscience community one step closer to produce a digital twin of the Earth system, which will be decisive to address the grand challenges of our time.

CODE AVAILABILITY STATEMENT

Geo-Soft-CoRe can be found in the Spanish National Research Council repository (DIGITAL.CSIC, 2021a).

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding authors.

AUTHOR CONTRIBUTIONS

IDF and JA wrote the draft and all the authors reviewed the manuscript. RC conceptualized the study. CM and IB provided the facilities to upload the software into DIGITAL.CSIC. IDF managed the software and code collection. JA, EB, KB, SF, AF, JF, DG-C, SG, AH, JJ-M, AK, M-GL, AO-C, AP, MS, MT, and SV developed the software. IDF, JA, KB, SF, AF, DG-C, JJ-M, AK, and M-GL designed the figures.

FUNDING

This research has been funded by the Projects EPOS IP 676564, EPOS SP 871121, SERA 730900, GeoCAM (PGC2018-095154-B-I00, Spanish Government) and the Center of Excellence for Exascale in Solid Earth (ChEESE) under the Grant Agreement 823844. IDF was funded by a FEDER-Junta de Castilla y León Postdoctoral contract (SA0084P20). JA and M-GL are funded by the Spanish Ministry of Science and Innovation through the Juan de la Cierva fellowship (IJIC 2018-036074-I and IJC 2018-036826-I, respectively), funded by MCIN/AEI/10.13039/501100011033. AH is grateful for his Ramón y Cajal contract (RYC 2020-029253-I). Additional funding was provided by the Spanish Ministry of Science and Innovation (RTI 2018-095594-B-I00, PGC 2018-095154-B-I00) and the Generalitat de Catalunya (AGAUR, 2017SGR1022). AP’s work was supported by: a Science Foundation Ireland Career Development Award (17/CDA/4695); an investigator award (16/IA/4520); a Marine Research Programme funded by the Irish Government, co-financed by the European Regional Development Fund (Grant-Aid Agreement No. PBA/CC/18/01); European Union’s Horizon 2020 research and innovation programme InnoVar under grant agreement No 818144; SFI Centre for Research Training in Foundations of Data Science 18/CRT/6049, and SFI Research Centre awards I-Form 16/RC/3872 and Insight 12/RC/2289_P2. AH and SG thank the
Spanish research project PaleoModes (CGL2016-75281-C2-1-R) which provided some of their financial support. JF is supported by an Atracción de Talento senior fellowship (2018-T1/AMB/11493) funded by Comunidad Autonoma de Madrid (Spain), and a project funded by the Spanish Ministry of Science and Innovation (PID2020-114854GB-C22).

REFERENCES

Abriol-Hernández, J. M., Pérez-Ráez, R., O’Connor, J. E., and García-Castellanos, D. (2018). Computational Fluid Dynamics Simulations of the Late Pleistocene Lake Bonneville Flood. J. Hydrol. 561, 1–15. doi:10.1016/j.jhydrol.2018.03.065

Accella, V. (2015). Grand Challenges in Earth Science: Research Toward a Sustainable Environment. Front. Earth Sci. 3, 68. doi:10.3389/feart.2015.00068

Afonso, J. C., Fernàndez, M., Ranalli, G., Grifo, and Science and Innovation (PID2020-114854GB-C22). (2018-T1/AMB/11493) funded by Comunidad Autonoma de Madrid (Spain), and a project funded by the Spanish Ministry of Science and Innovation (PID2020-114854GB-C22).

ACKNOWLEDGMENTS

Part of this work has been done using the facilities of the Laboratory of GeoDynamic Modeling from GEO3BCN-CSIC. We thank Alvaro Aravena, Carmen Gaima and Abani Kumar Patra for their valuable comments that improved the manuscript.

Structural Framework Inference and Hydrogeological Implication. Jgg 02, 373–383. doi:10.4236/jgg.2011.23040

Azar, A., Biersteker, J. B., Dewey, R. M., Doran, G., Forsberg, E. J., Harris, C. D. K., et al. (2021). Integrating Machine Learning for Planetary Science: Perspectives for the Next Decade. Bull. AAS 53. doi:10.3847/25c2efeb.aa328727

Bartol, J., Matencio, L., García-Castellanos, D., and Leever, K. (2012). Modelling Depositional Shifts between Sedimentary Basins: Sediment Pathways in Parathys Basins during the Messinian Salinity Crisis. Tectonophysics 536–537, 110–121. doi:10.1016/j.tecto.2012.03.007

Bartolini, S., Martí, J., Sobradelo, R., and Becerril, L. (2017). Probabilistic E-Tools for hazard Assessment and Risk Management. Adv. Volcanol., 47–61. doi:10.1007/11157_2017_14

Bauer, P., Dueben, P. D., Hoefler, T., Quintino, T., Schulthess, T. C., and Widi, N. P. (2021a). The Digital Revolution of Earth-System Science. Nat. Comput. Sci. 1, 104–113. doi:10.1038/s43588-021-00023-0

Bauer, P., Stevens, B., and Hazeleger, W. (2021b). A Digital Twin of Earth for the green Transition. Nat. Clim. Chang. 11, 80–83. doi:10.1038/s41558-021-0096-y

Baykiew, E., Brönnr, M., Ebbing, J., Fabian, K., Holzrichter, N. and Yixiad, D. (2020b). Magnetic Tesseroid Inversion: A Program for Global Magnetic Tesseroid Inversion. DIGITAL.CSIC. doi:10.20350/digitalCSIC/13838

Baykiew, E., Brönnr, M., Ebbing, J., and Fabian, K. (2020a). Magnetic Tesseroids: A Collection of Command-Line Tools for Modelling of the Magnetic Field with Spherical Prisms (Tesseroids) Used as Magnetic Sources. DIGITAL.CSIC. doi:10.20350/digitalCSIC/13839

Baykiew, E., Ebbing, J., Brönnr, M., and Fabian, K. (2016). Forward Modeling Magnetic fields of Induced and Remanent Magnetization in the Lithosphere Using Tesseroids. Comput. Geosci. 96, 124–135. doi:10.1016/j.cageo.2016.08.004

Baykiew, E., Yixiad, D., and Ebbing, J. (2020c). Global High-Resolution Magnetic Field Inversion Using Spherical Harmonic Representation of Tesseroids as Individual Sources. Geosciences 10, 147–215. doi:10.3390/geosciences10040147

Bath, M., and R. Stefansson, E. (1966). S-P Conversion at the Base of the Crust. Ann. Geophys. 19, 119–130. doi:10.4011/ag-5041

Bensen, G. D., Ritzwoller, M. H., Barmin, M. P., Levshin, A. L., Lin, F., Moschetti, M. P., et al. (2007). Processing Seismic Ambient Noise Data to Obtain Reliable Broad-band Surface Wave Dispersion Measurements. Geophys. J. Int. 169, 1239–1260. doi:10.1111/j.1365-246X.2007.03374.x

Bergen, K. J., Johnson, P. A., De Hoop, M. V., and Beroza, G. C. (2019). Machine Learning for Data-Driven Discovery in Solid Earth Geoscience. Science 363, 365. doi:10.1126/science.aau0323

Bernal, I. (2021). El acceso a la información en la Agenda 2030: ¿Cómo contribuyen los repositorios de acceso abierto? Enredadera 36, 19–22. doi:10.20350/digitalCSIC/13939

Berry, M., Wijk, J., Cadol, D., Emry, A., and García-Castellanos, D. (2019). Endorheic-Exorheic Transitions of the Rio Grande and East African Rifts. Geochim. Geophys. Geosyst. 20, 3705–3729. doi:10.1029/2018gc008176

Baykiew, A., Aravena, A., Neri, A., Gutiérrez, E., Escobar, D., Schliz, M., et al. (2021). Thematic Vent Opening Probability Maps and hazard Assessment of Small-Scale Pyroclastic Density Currents in the San Salvador Volcanic Complex (El Salvador) and Nejapa-Chiltepe Volcanic Complex (Nicaragua). Nat. Hazards Earth Syst. Sci. 21, 1639–1665. doi:10.5194/nhess-21-1639-2021

Bonnais, R., Costa, A., Folch, A., Macedonio, G., and Capra, L. (2012). Numerical Simulation of Tephras Transport and Deposition of the 1982 El Chichón Eruption and Implications for hazard Assessment. J. Volcanol. Geotherm. Res. 231–232, 39–49. doi:10.1016/j.jvolgeores.2012.04.006

DeFelipe et al. 

Geoscientific Software & Code Repository

Frontiers in Earth Science | www.frontiersin.org 15 April 2022 | Volume 10 | Article 828005
the Lithosphere and Sublithospheric Upper Mantle. Geochem. Geophys. Geosyst. 10. doi:10.1029/2009GC002391
Fullea, J., Baykier, E., Afonso, J. C., Connolly, J. A. D., Mather, B., and Guerri, M. (2019b). LitMod3D_v4.0: a Software for Integrated Geophysical-Petrological Thermochemical Modelling of Lithosphere. DIGITAL.CSIC. doi:10.20350/digitalCSIC/13876
Fullea, J., Camacho, A. G., Negredo, A. M., and Fernández, J. (2015). The Canary Islands Hot Spot: New Insights from 3D Coupled Geophysical-Petrological Modelling of the Lithosphere and Uppermost Mantle. Earth Planet. Sci. Lett. 409, 71–88. doi:10.1016/j.epsl.2014.10.038
Fullea, J. (2008). FA2BOUG: a Software to Compute Complete Bouguer Gravity Anomalies. DIGITAL.CSIC. doi:10.20350/digitalCSIC/13871
Fullea, J., Fernández, M., and Jiménez-Munt, I. (2019c). Geoid_Elevation: Software to Compute Bouguer Gravity Anomalies from Gridded Free-Air Anomalies: Application to the Atlantic-Mediterranean Transition Region. Lithos 120, 74–95. doi:10.1016/j.lithos.2010.03.003
Fullea, J., Fernández, M., and Jiménez-Munt, I. (2019c). Geoid_Elevation: Software to Compute Bouguer Gravity Anomalies from Gridded Free-Air Anomalies: Application to the Atlantic-Mediterranean Transition Region. Lithos 120, 74–95. doi:10.1016/j.lithos.2010.03.003
Fullea, J., Fernández, M., and Jiménez-Munt, I. (2019c). Geoid_Elevation: Software to Compute Bouguer Gravity Anomalies from Gridded Free-Air Anomalies: Application to the Atlantic-Mediterranean Transition Region. Lithos 120, 74–95. doi:10.1016/j.lithos.2010.03.003
Fullea, J., Fernández, M., and Jiménez-Munt, I. (2019c). Geoid_Elevation: Software to Compute Bouguer Gravity Anomalies from Gridded Free-Air Anomalies: Application to the Atlantic-Mediterranean Transition Region. Lithos 120, 74–95. doi:10.1016/j.lithos.2010.03.003
Fullea, J., Fernández, M., and Jiménez-Munt, I. (2019c). Geoid_Elevation: Software to Compute Bouguer Gravity Anomalies from Gridded Free-Air Anomalies: Application to the Atlantic-Mediterranean Transition Region. Lithos 120, 74–95. doi:10.1016/j.lithos.2010.03.003
Fullea, J., Fernández, M., and Jiménez-Munt, I. (2019c). Geoid_Elevation: Software to Compute Bouguer Gravity Anomalies from Gridded Free-Air Anomalies: Application to the Atlantic-Mediterranean Transition Region. Lithos 120, 74–95. doi:10.1016/j.lithos.2010.03.003
Fullea, J., Fernández, M., and Jiménez-Munt, I. (2019c). Geoid_Elevation: Software to Compute Bouguer Gravity Anomalies from Gridded Free-Air Anomalies: Application to the Atlantic-Mediterranean Transition Region. Lithos 120, 74–95. doi:10.1016/j.lithos.2010.03.003
Fullea, J., Fernández, M., and Jiménez-Munt, I. (2019c). Geoid_Elevation: Software to Compute Bouguer Gravity Anomalies from Gridded Free-Air Anomalies: Application to the Atlantic-Mediterranean Transition Region. Lithos 120, 74–95. doi:10.1016/j.lithos.2010.03.003
Fullea, J., Fernández, M., and Jiménez-Munt, I. (2019c). Geoid_Elevation: Software to Compute Bouguer Gravity Anomalies from Gridded Free-Air Anomalies: Application to the Atlantic-Mediterranean Transition Region. Lithos 120, 74–95. doi:10.1016/j.lithos.2010.03.003
