Ⓔ

ELECTRONIC SEISMOLOGIST

SHERIFS: Open-Source Code for Computing Earthquake Rates in Fault Systems and Constructing Hazard Models

by Thomas Chartier, Oona Scotti, and Hélène Lyon-Caen

ABSTRACT

Modeling the seismic potential of active faults and their associated epistemic uncertainties is a fundamental step of probabilistic seismic-hazard assessment (PSHA). Seismic hazard and earthquake rate in fault systems (SHERIFS) is an open-source python code that builds hazard models including earthquake ruptures involving several fault sections or fault-to-fault (FtF) ruptures. It contains user-friendly tools to calculate the annual rate of FtF ruptures in a fault system based on the slip-rate estimates and accounting for associated background seismicity.

SHERIFS applies a forward incremental approach following three rules: (1) the FtF ruptures allowed in the fault system are defined as input by the user and explored randomly, (2) the magnitude–frequency distribution of the modeled seismicity in the fault system must follow an imposed shape, and (3) the slip-rate budget attributed to each fault section must be preserved in the calculation if the first two rules allow it. Indeed, in some cases, a fraction of the slip-rate budget must be considered as being spent in non-mainshock events such as creep or postseismic slip. Background seismicity rates are defined by the hazard modeler as the ratio of seismicity occurring on the modelled faults for different ranges of magnitude.

Given a coherent set of input hypotheses, SHERIFS allows end users to build the seismic-hazard fault model thanks to an interactive user-friendly interface. It aims to help interactions between field data collectors and hazard modelers to explore and weight epistemic uncertainties affecting the input hypotheses. To do so, SHERIFS includes tools to compare modeled earthquake rates with the available local data (earthquake catalog and paleoseismological data). This comparison can be used to weigh different hypotheses explored in a logic tree and discard the hypotheses that are not in agreement with the data. SHERIFS's outputs are in a format that can be used directly as inputs for PSHA in the OpenQuake engine (Pagani *et al.*, 2014).

Supplemental Content: Text and figure describing the calculation of the earthquake rates for a system of five faults using seismic hazard and earthquake rate in fault systems (SHERIFS).

INTRODUCTION

Developing a hazard model for a region in which active faults have been described requires to convert the geological and geophysical data into 3D fault geometries and slip rates and then into earthquake rates. Most commonly seismic-hazard modelers have to develop their own approach to manage the available data. Each modeler has a different approach for translating slip rates into earthquake rates on faults which sometimes does not allow a straightforward understanding of the resulting hazard model. Some assumptions have to be made when building a hazard model and, in the ideal case, the data collectors are involved in the modeling process to insure the working hypotheses are acceptable.

To help data collectors build hazard models, some tools are available (FiSH and OpenQuake toolkits). The FiSH (Pace *et al.*, 2016) modeling approach provides estimates of earthquake rates and time-dependent probability of earthquakes in the period of interest for the hazard assessment using the data available on a fault. This code can easily be operated by a user with limited coding skills. FiSH can be used to explore a range of epistemic uncertainties affecting the earthquake rates of the modeled fault. Because the approach in FiSH focuses on one fault only, multifault ruptures can only be modeled by considering average parameters for the whole structure. The OpenQuake toolkits allow building hazard models using as input shapefiles containing information on the faults but considering as well a fault-by-fault approach and not a fault system approach.

Thus, in most hazard models (e.g., Woessner *et al.*, 2015; Taiwan Earthquake Model [TEM], Wang *et al.*, 2016), small faults are usually grouped into large structures to allow for larger magnitude earthquakes. Moreover, background earthquakes are handled using a truncated approach in which earthquakes with a magnitude lower or equal to M_w 6.4 occur only in the background zone with a rate defined by the rate in the earthquake catalog, whereas magnitudes higher than M_w 6.4 are located on the faults with a rate defined using the average slip rate of the fault. This approach allows integrating faults within area zones while avoiding double counting. However, it can lead to

 doi: 10.1785/0220180332

discontinuities in the regional magnitude–frequency distribution (MFD) because the rate of low-magnitude earthquakes and the rate of high-magnitude earthquakes are calculated from different sources of information (Danciu *et al.*, 2017). Last but not least, magnitudes stronger than M_w 6.4 may potentially occur in the background as well, especially in the regions of slow tectonic deformation and in the presence of concealed faults (e.g., Darfield earthquake, Hornblow *et al.*, 2014).

It has been known for a long time several fault sections of a fault or of several faults can rupture during a single fault-to-fault (FtF) rupture event (e.g., Barka and Kadinsky-Cade, 1988; Wesnousky, 1988; Knuepfer, 1989). It is therefore paramount to have modeling procedures to allow a large set of possible FtF ruptures to occur in an aleatory fashion in the hazard model while reflecting the individual slip rate of each section.

The Uniform California Earthquake Rupture Forecast version 3 (UCERF3, Field *et al.*, 2014) tackles this issue and allows FtF ruptures to occur in the model by relaxing the segmentation assumption and treating the problem at the faults system level. The annual rate of FtF ruptures is inverted in a grand inversion process considering the large amount of data available in California. UCERF3 allows all possible ruptures in the fault system to follow a set of plausibility rules. The individual slip rates of each section, the earthquake rates at the level of California, and at the level of the paleoseismic sites are some of the constraints used to invert for the rate of each rupture. This model also allows small earthquakes to happen on faults and large earthquakes to happen in the area sources by defining off-fault seismicity without appealing to the artificially truncated distribution commonly used.

In this article, we propose an alternative approach to calculate the earthquake rates in a fault system, which preserves the main philosophy of UCERF3. The aim of the proposed approach, named SHERIFS (seismic hazard and earthquake rates in fault systems), is to provide a user-friendly tool with a graphical interface, which allows data collectors and fault modelers to build hazard models in a forward processing scheme with an exploration of the epistemic uncertainties. The approach consists in the conversion of the slip-rate budget of each fault of the system into rupture rates for all possible single fault and FtF ruptures. SHERIFS relies on a forward modeling incremental method with a simple layout of the assumptions about the input model which leads to an easy construction of the logic tree. The only required input data are an estimate of the geometry and slip rate of the modeled faults.

The theoretical background of the methodology is presented in Chartier *et al.* (2017).

In this article, we present how this methodology has been implemented in SHERIFS, an open-source tool coded in the Python language. The objective of SHERIFS is not only to provide the optimal hazard model but also to allow the exploration of a large range of epistemic uncertainties and we recommend using it as such. By providing an easy-to-use tool for modeling faults in hazard models, we wish to promote discussions and feedback between data providers and hazard modelers, one of the main objectives of the Fault2SHA European Seismological Commission (ESC) working group (see Data and Resources). SHERIFS includes a set of tools allowing to compare the modeled earthquake rates with the local earthquake catalog or paleoseismological data and to discuss the different input hypotheses. Thus, SHERIFS can be useful not only to hazard modelers but also to geologists as a means to test the coherency of the rupture hypotheses that best describe the seismic potential of a fault system.

This tool creates hazard models for fault systems that can be used as inputs for OpenQuake (Pagani *et al.*, 2014), but because most probabilistic seismic-hazard assessment (PSHA) computing codes use similar entries, the output of SHERIFS could easily be adapted to provide inputs for other codes as well (e.g., OpenSHA, Field *et al.*, 2003; CRISIS, Aguilar-Meléndez *et al.*, 2017).

GENERAL METHODOLOGY APPLIED IN SHERIFS

The SHERIFS code allows end users to build the fault model thanks to an interactive user-friendly interface. The logic-tree structure and the input files can be easily modified for a rerun if different parameters need to be tested.

This section of the article synthesizes the main steps of the methodology implemented in SHERIFS as presented in Chartier *et al.* (2017). This methodology requires as input data fault section geometries and slip-rate estimates for all the faults of the fault system to calculate the earthquake rupture rates.

The SHERIFS iterative methodology requires first establishing possible FtF earthquake rupture scenarios in the fault system and fixing the shape of the target MFD for the entire fault system. At each iteration, a magnitude is randomly picked based on the target MFD and an earthquake rupture scenario able to host such magnitude is selected. For this scenario, an increment of the slip-rate budget of the involved faults is converted into an earthquake rupture rate for the considered magnitude. At the beginning of the iterative process, the absolute value of the target MFD is not known, only the shape is imposed. In SHERIFS, we explore different criteria to set the MFD target value which are detailed in the following sections. The iterative computation continues until the slip-rate budget of the faults is exhausted.

Required Information and Input Working Hypotheses

At first, the limits of the fault system have to be set. If the specific extent of the fault system is clearly limited in the field, we suggest using the enclosing boundary. In some applications, the fault system can be much larger than the area of interest for the hazard study. In such a case, the part of the system considered should be large enough to make sure an extension of the system would not affect the rupture rates inside the zone of interest. If the fault network of a region is composed of distinctive subregions in which it is believed no rupture can go through, we suggest treating them as different fault systems.

Within the defined fault system, the geometry (trace, dip, and upper and lower seismogenic depth), the slip rate (with uncertainties), and the kinematics (normal, strike slip, reverse, or the rake) of the fault sections must be provided. For partly creeping faults, we suggest the user to correct the geological or geodetic slip-rate input to reflect the seismic slip rate that should be considered in SHERIFS. Alternatively, the shear modulus or the geometry of the creeping sections can also be adapted to model only the locked part of the fault.

We suggest defining typical lengths of fault sections to be on the order of the seismogenic depth. Faults with a longer length should be cut in smaller sections (that can rupture together). The sections should be defined as small as necessary to account for local information. For example, a change in slip rate, rake, or dip should be considered for defining the sections.

The SHERIFS approach is based on the assumption the seismicity of the fault system taken as a whole follows the shape of a known MFD (i.e., Gutenberg and Richter, 1944; hereafter, GR, or Youngs and Coppersmith, 1985; hereafter, YC). If the local seismicity data do not show clearly a distinct shape, SHERIFS allows the user to explore several hypotheses in a logic tree.

Hypothesis on the FtF Ruptures

One of the most important working hypothesis of the SHERIFS methodology concerns the definition of possible FtF ruptures allowed in a given fault system. SHERIFS does not impose the way to choose the set of FtF ruptures, and only requires a list. The user is free to use rules (as done in UCERF3) or rely on physical simulations to define possible FtF ruptures. Because the rules to define FtF ruptures are subject to discussion (i.e., Schwartz, 2018), SHERIFS allows exploring different sets of possible FtF ruptures.

To implement in SHERIFS the single and FtF ruptures following the OpenQuake formalism, a single-fault rupture is defined in the input files with either the simple or complex fault typology, depending on the user's choice and can host magnitudes starting from the minimal magnitude up to the maximum magnitude the single fault can accommodate (scaling law-dependent value). The simple fault typology is defined by the fault trace, the dip, and the upper and lower seismogenic depth and the complex fault typology allows more details in the fault geometry such as listricity or seismogenic depth variable along strike. An FtF rupture, on the other hand, is defined with the characteristic fault source typology, hosting only the greater magnitudes activating the entire fault surface included in the FtF rupture.

Sampling the Model Space of Uncertainties

The slip rate of the faults, the magnitude scaling parameters, and the *b*-value are usually associated with important uncertainties. In SHERIFS, uncertainty bounds are considered and explored based on the user-defined number of random samples. For each branch of the logic tree, SHERIFS will create a number of models equal to the number of random samples. For each model, a slip-rate value is picked uniformly within the uncertainty bounds of the slip rate attributed to each fault, the parameters of the scaling law are picked independently following a Gaussian distribution within their error bounds; and a *b*-value is picked within the boxcar range of *b*-values considered by the user. The first sample provided in the output file (model number 1) is always the mean value of the slip rate, scaling law, and *b*-value parameters.

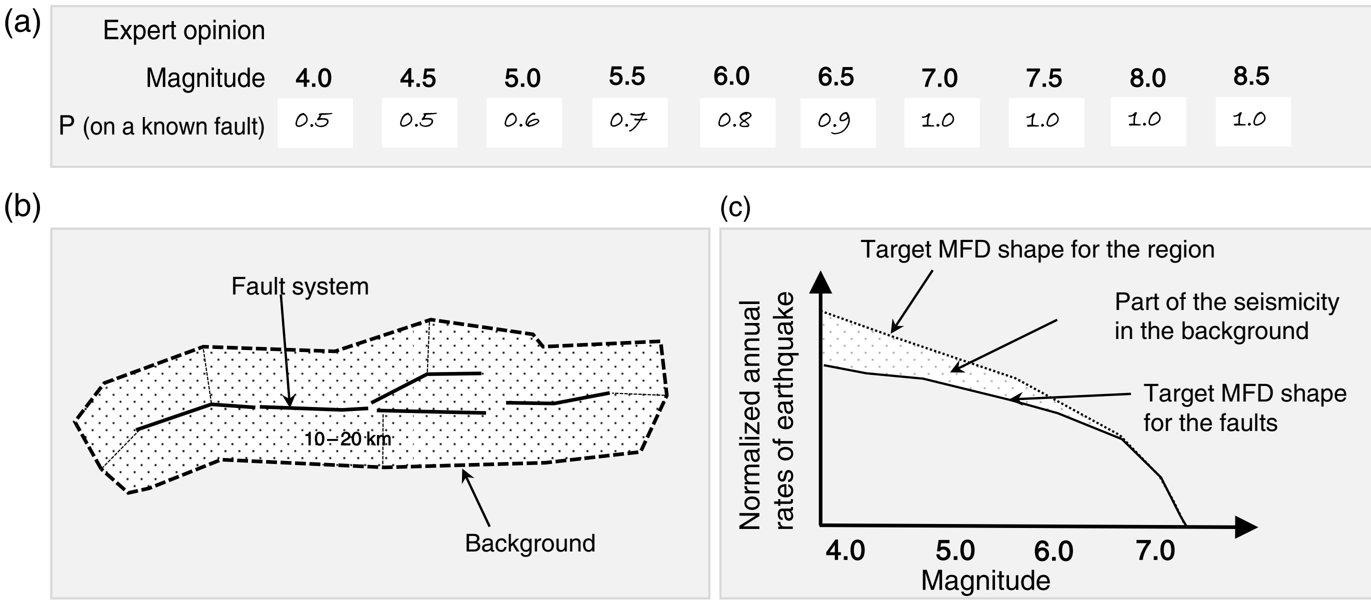
To explore slip-rate value uncertainties of neighboring sections, the user can choose between a correlated or random sampling method. If the random option is chosen, slip-rate values are selected in the distribution without consideration of the values sampled in neighboring fault sections. Therefore, a section of the fault can be sampled in the lower part of its distribution when its neighboring section of the same fault can be sampled in the upper part of this distribution. If the correlated option is chosen, sections, which can rupture together in FtF scenarios, will always be sampled in the same part of their distribution. In practice, the distribution of possible slip rate on each fault is divided in four quarters and if two faults can rupture together, their slip rates will be sampled in the same quarter of the distribution for this model. We suggest the correlated option be used when working with long-fault systems in which long faults have to be discretized in smaller sections. In the case in which two very different opinions concerning the slip rate of the faults are available, for example, slip rates based on geological or geodetic considerations, we suggest building two different models rather than exploring these options in one single distribution.

Concerning the magnitude scaling relationships, the user can choose a set of published options (Wells and Coppersmith, 1994; Leonard, 2010; Thingbaijam *et al.*, 2017) and explore them in a logic tree. For some options, it is possible to choose which dimension of the fault to use (length or area). A new branch of the logic tree will be created for each chosen option. Within this branch, the uncertainty affecting the estimate of the maximum magnitude is explored by random sampling of the parameters of each scaling relationship.

Setting the Background

Fault-based PSHA models usually take into account the possibility of an earthquake occurring on an unknown fault by defining a background zone around the faults. The seismicity rate of this background zone is usually deduced from the catalog and follows a GR distribution between the minimum magnitude M_{min} and a truncated magnitude M_{t} (usually $5.8 < M_{\mathrm{t}} < 6.5$). The earthquakes with a magnitude larger than M_{t} are confined to the modeled faults and their rates are defined by the slip rate and geometry of the faults.

In SHERIFS, earthquake rates are deduced from the geological information and are partitioned between background and faults (Fig. 1a). To define the background zone geometry, the user can rely on simple distance criteria (e.g., 10 or 20 km distance from the faults) to draw the background zone around the fault system (Fig. 1b). The background geometry and properties should be defined specifically for each model considered for the fault system. The seismicity rate of the background is defined by the user by setting the ratio R of the seismicity occurring on the modeled faults vs. in the background. In the graphical interface, this ratio has to be set for each magnitude $M_{\mathrm{w}} > M_{\mathrm{min}}$ of the distribution and is likely to increase with magnitude. To help set this ratio, the user may rely on analysis



▲ **Figure 1.** Modeling of the background seismicity in seismic hazard and earthquake rate in fault systems (SHERIFS). (a) Expert opinion of the probability of a future earthquake to occur on a modeled fault for magnitude bin. (b) Example of background geometry. (c) Target magnitude–frequency distribution (MFD) modified to take into account the background seismicity.

of the distance between earthquakes and modeled faults. In SHERIFS, the MFD for the whole system has been defined to follow a target shape. The ratio R defined by the user for each magnitude bin is used to deduce the target shape for the faults. At the end of the iterative process, when the absolute MFD value of the fault system is known, the MFD of the background is then calculated using this ratio (Fig. 1c).

Because R is very difficult to determine, several background hypotheses are meant to be explored. This straightforward definition of the background has been introduced in SHERIFS since the version used in Chartier *et al.* (2017). We hope to encourage the discussion about the expert opinion on the matter of the background seismicity and its effect on the seismic-hazard assessment.

Workflow of the SHERIFS Tool

SHERIFS workflow includes three tools that should be used sequentially (Fig. 2). The first tool sets up and runs the SHERIFS calculation, and the logic tree exploring the epistemic uncertainties with the help of a graphical interface. The second tool allows visualizing the hazard model and comparing the modeled earthquake rates with available data. After a critical analysis of this comparison, the third tool allows the end-user setting the weights of the logic-tree branches. The final outputs can then be directly used to perform PSHA.

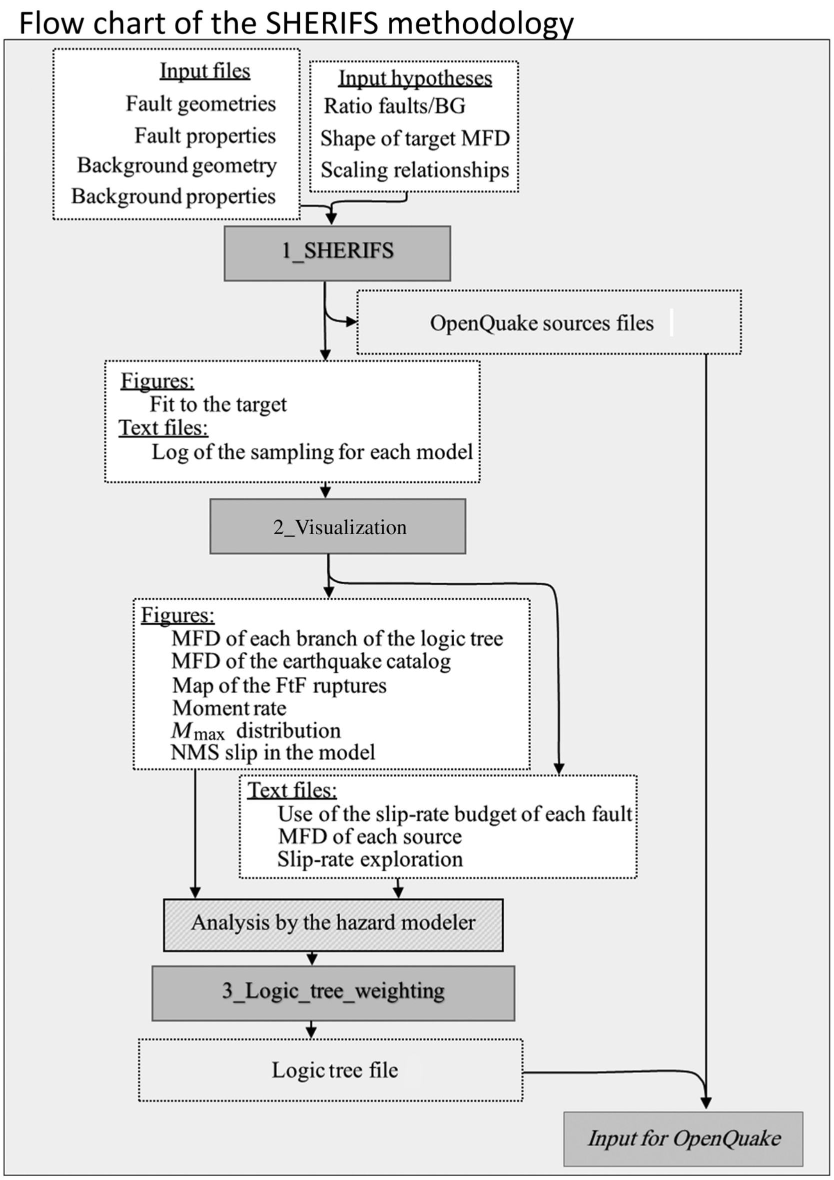
Once the necessary information are in the input files and the python file1_SHERIFS.py has been filled with information concerning the run, SHERIFS opens a series of windows using a graphical interface. The purpose of these windows is to help the user build the logic tree and the input files necessary for the PSHA calculation. The detail of each of the windows is explained in the user manual (see Data and Resources). It is worth noting the use of the graphical interface is not required by SHERIFS; an advanced user with a good knowledge of the required information and files structure can run SHERIFS independently of the graphical interface. It is also possible for the user to modify some hypotheses of the calculation or add or delete a logic-tree branch and rerun the calculation.

The workflow for the calculation of the earthquake rates is described in Figures 3 and 4.

The first step consists in formatting the inputs provided by the user to run SHERIFS (Fig. 3). The slip-rate budget of each fault is subdivided in slip-rate (dsr) increments. The maximum magnitude is deduced using the geometry of each section of the maximum FtF rupture scenario and the scaling law. The shape of the target MFD is then corrected for the proportion of the seismicity assumed to occur in the background and then converted in terms of moment rate.

SHERIFS iterative methodology spends the slip-rate budget of each fault until the target MFD is reached and the slip-rate budgets of each fault are exhausted (Fig. 4). In some cases, the target is reached before the budget of all faults is exhausted; the remaining slip-rate budget is then considered as non-mainshock (NMS) slip and not converted into earthquake rates. The proportion of NMS slip on a specific fault of the system is an output of the methodology resulting from the combination of faults parameters, the set of FtF scenarios explored, and the imposed shape of the regional MFD.

More specifically, at each iterative loop, one magnitude bin of size 0.1 is randomly picked using the MFD defined in terms of moment rate and a source (single fault or FtF rupture) able to host such magnitude is selected (Fig. 4). If the target is not already set or not reached for this bin of magnitude, the rate of this magnitude for the picked source is increased and the



▲ **Figure 2.** Flowchart of the overall SHERIFS methodology. BG, Backgorund; FtF, fault-to-fault; NMS, non-mainshock.

slip-rate budget of the fault or faults involved in the source is reduced by one dsr increment; if the target is reached, the dsr increment is considered as NMS slip.

The selection of fault or FtF scenario, which can host the picked magnitude, is realized with faults preferentially picked according to the ratio of remaining slip-rate budget over the slip-rate initial budget. This *ad hoc* rule has been set up to help isolated faults and spend their slip-rate budget.

The absolute value of the target MFD is set when the system reaches one of the following:

1. The budget of at least one of the faults participating in the largest magnitude earthquakes is exhausted. The target is then set because it is not possible to change the rate of earthquakes of the three largest magnitude bins, because the budget of one of the faults participating in the largest magnitude FtF ruptures is exhausted. Using the shape of the regional target MFD, the absolute value of the target earthquake rate for all magnitude bins can then be set (Fig. 5a). The choice of the three largest bins was made to allow flexibility around the target shape for the largest bins and avoid a single odd rupture scenario in the last magnitude bin limits the rates of the whole system.
2. The rate of the two largest magnitude bins is limited by the budget of a fault being exhausted but the third largest one is not. The annual rate on the latter bin is then limited to twice the rate of the mean of the two largest bins to limit the differences of earthquake rates between the last magnitude bins and keep the consistency of the target MFD shape set (Fig. 5b).
3. The seismic moment required to fit the target shape and the rate of the largest magnitude earthquakes is equal to the remaining seismic moment in the system at a given moment of the iterative process. This *ad hoc* rule then sets the target to ensure the rate of the smaller magnitude earthquakes will also reach the set target (Fig. 5c).

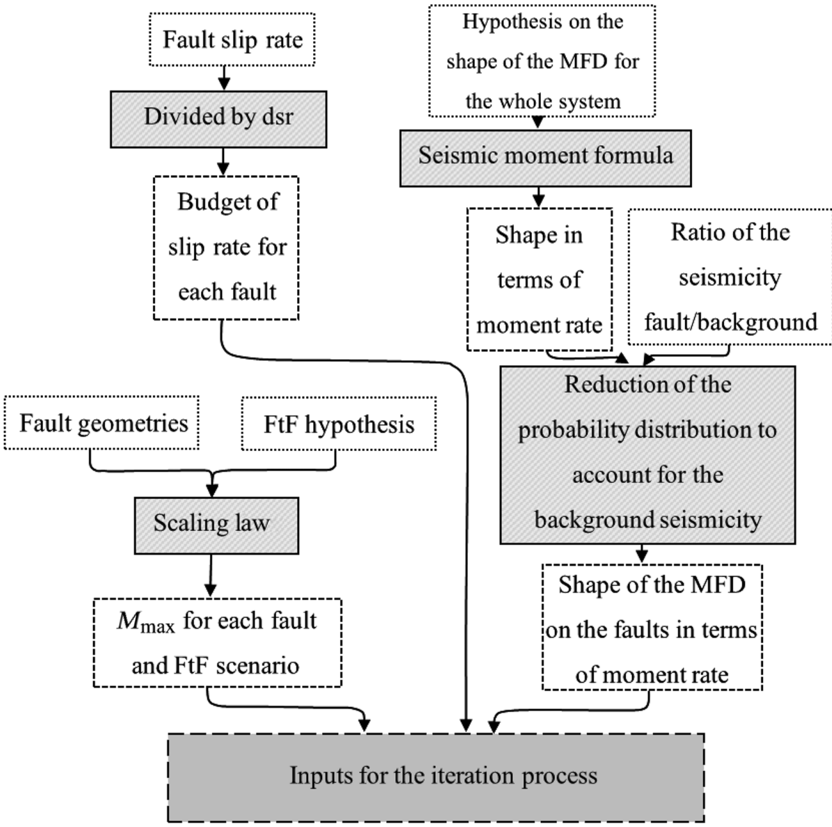
Rules 2 and 3 were newly introduced since the version of the code used in Chartier *et al.* (2017) to better model the target shape of certain fault systems proposed by some beta users. Additional novelties introduced in the latest version of SHERIFS are discussed in the Ⓔ supplemental content.

SHERIFS will check if the resulting MFD of the fault system matches the shape of the target imposed by the user. In the event the shape does not match, which occurs when the slip-rate increment is not small enough, then the model is rerun with a slip-rate increment divided by two. After three reruns, the model is accepted but a warning is displayed.

It is important to point out SHERIFS is not modeling earthquake ruptures as an earthquake simulator would do but only converting slip rate on faults into earthquake rates. Each spending of a slip-rate increment is not the modeling of a true earthquake but is converted into a rate of earthquakes, which can be allowed on a specified fault.

NMS Slip

NMS slip is the proportion of slip-rate budget of a fault section, which could not be transformed into earthquake rates within the SHERIFS procedure. A high-NMS proportion signifies the input slip rate for this fault section is not compatible with the set of scenarios, the target MFD and the input slip rate of the neighboring faults given the assumptions and rules of



▲ **Figure 3.** SHERIFS data processing workflow. Dsr, sub-divided in slip rate.

SHERIFS. In this case, NMS can be regarded as an error or misfit of the model. Models that predict a very high NMS, typically over 30%–40%, need to be reconsidered in detail to understand the origin of this NMS. Similarly, the high NMS possibly calculated for some fault sections needs to be understood. The origins for this NMS can be diverse; the set of scenario involving the fault section or the fault slip-rate set in input for this section is not realistic or the target MFD is not suitable for the fault system. Large NMS slip can also result from miscalculations in the code, when using a too large slip-rate increment for example. NMS slip differs from what could be strictly called a misfit or error, and it can be explained by physical phenomena and should therefore not necessarily be reduced to zero. NMS can reflect that part of the slip rate deduced from geology or geodetic inversions, which can be dissipated in nonseismic processes such as afterslip, creep, or slow-slip events but cannot be taken as a definitive proof of the existence of these phenomena in a specific fault system as it could reflect some limitations of the SHERIFS approach.

MODEL VISUALIZATION

This tool allows extracting and visualizing numerous results, which can be easily compared with the seismicity catalog of the region and plotted on a map. The user needs to specify in the Python file the name of the run and the data files to use as well as the spatial limits of the maps. The user can select the type of information to plot, very useful when running a large fault model. The user manual describes in more detail the use of this tool.

The produced figures include the distribution of MFDs of the whole model and of each individual branch of the logic tree. These MFDs can be compared with the MFD of the local earthquake catalog. The local catalog is extracted using the geometry of the background defined by the user. The uncertainties of both the completeness period of the catalog and the magnitude of the earthquakes can be explored. A user who wishes to study the MFD of a region smaller than the whole system can define a subarea. The code will extract the MFD of the faults within this subarea and the MFD of the catalog for this subarea. This tool also plots the MFD of individual faults and fault sections and compares the results with the rate of historical or paleoearthquakes when this information is available for single faults.

The user can use this tool to extract statistics concerning the amount of NMS slip in the model, the maximum magnitude for each branch, the repartition of the moment budget between single-fault ruptures and complex FtF ruptures.

Maps of each FtF rupture scenario are generated allowing the user to check the input scenarios were correctly set up.

LOGIC-TREE WEIGHTING

Once the content of the logic tree has been analyzed and compared with the local data (catalog and paleoseismicity), the user needs to attribute weights to each branch of the logic tree. This tool opens a graphical interface to help building the logic-tree file used as an input for the PSHA calculation.

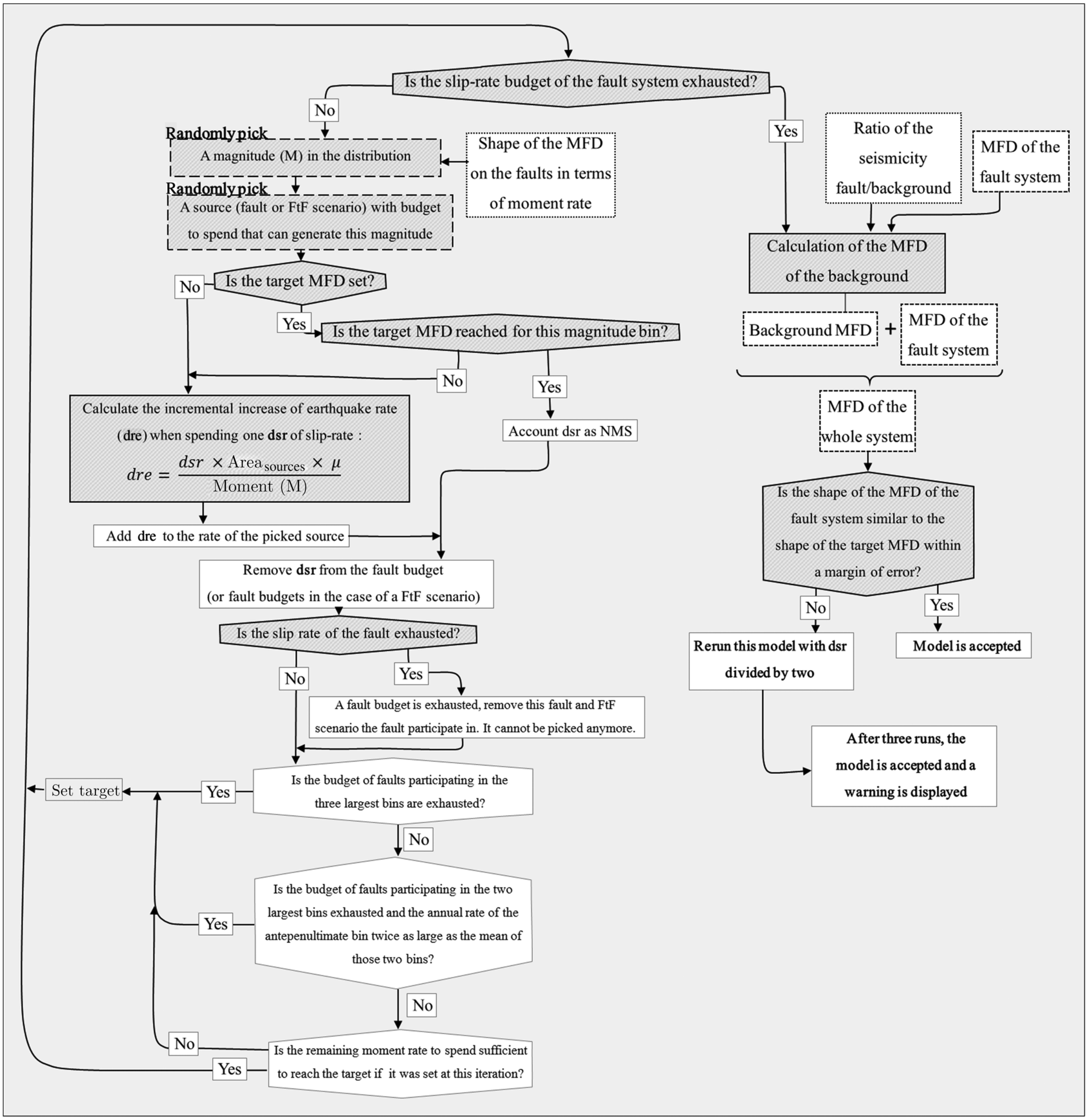
SHERIFS aims to promote the discussion of the uncertainties affecting the earthquake rates on faults. The logic-tree weighting tool should not be run before a proper discussion of each branch of the logic tree.

APPLICATIONS OF SHERIFS

Past Applications of SHERIFS

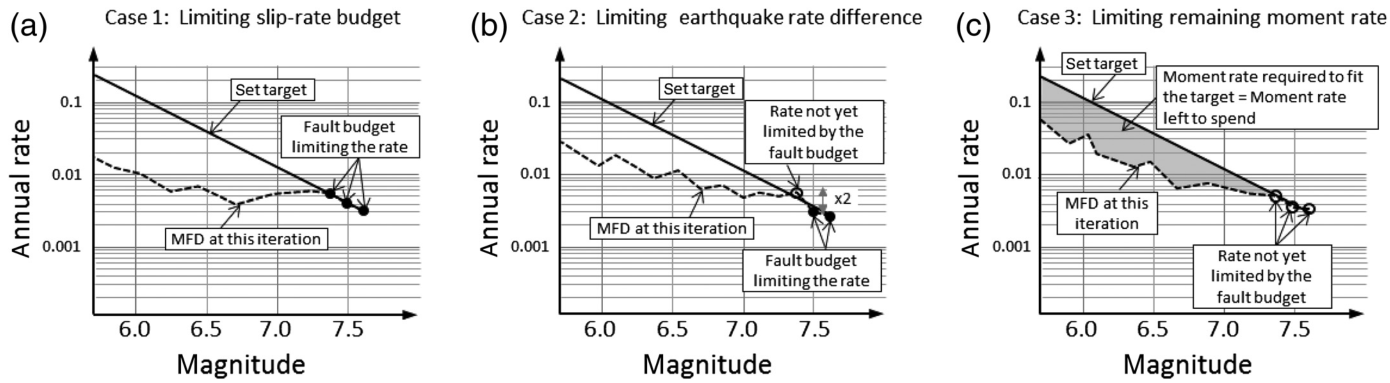
The proof of concept of the SHERIFS methodology was realized on the western Corinth rift (WCR) fault system (Chartier *et al.*, 2017). In this methodological exercise, a GR MFD target shape was considered, and three sets of FtF scenarios were tested: B14_s considered only single-fault scenarios, B14 considered FtF ruptures with a distance criteria of 3 km between two faults, and B14_hc considered a higher connectivity within the fault system and allowed FtF ruptures separated by 5 km or less.

SHERIFS was run with these hypotheses and sanity checks were performed.

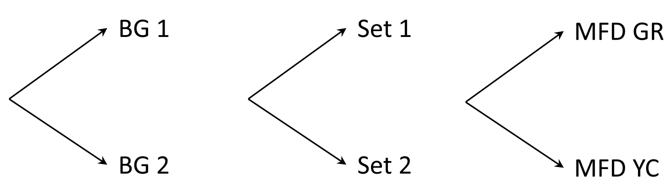


▲ **Figure 4.** SHERIFS iterative process workflow to compute earthquake rates on faults.

The comparison of the fault model results with the seismicity revealed the B14_s branch of the logic tree implied seismicity rates were incompatible with the rates computed from the catalog. Moreover, the B4_s branch did not produce the large earthquakes observed in the historical catalog. Finally, comparison of the modeled rates with paleoseismic data available for the Aigion fault, clearly indicated the B14_s branch cannot reproduce the rate of large earthquakes deduced from paleoseismic data. On the other hand, for the B14 and the B14_hc branches, in which FtF ruptures were allowed, the agreement between the modeled rates and the recorded seismicity and paleoseismic rate improved, with the 5 km criteria used in B14_hc explaining better the data than the 3 km distance criteria used in B14. The discussion of these results leads to the conclusion the model B14_s should be discarded from the logic tree, and the model B14_hc could have a stronger weight than B14.



▲ **Figure 5.** The three cases of setting the target in the iterative process. (a) Case 1: limiting slip-rate budget, (b) case 2: limiting earthquake rate difference, and (c) case 3: limiting remaining moment rate.



▲ **Figure 6.** Logic tree for the example calculation in which BG1 and BG2 are the two background hypotheses, set 1 and set 2 are the two rupture scenario set hypotheses, and MFD GR (Gutenberg–Richter) and MFD YC (Youngs and Coppersmith) are the two MFD shape hypotheses.

The SHERIFS methodology is presently being tested on different types of fault systems in France, Italy, Israel, Spain, Ecuador, India, and Turkey. A comparison exercise was also performed between SHERIFS and alternative methodologies that are currently being develop within the Fault2SHA ESC working group (see Data and Resources) using a fault system in northern Italy (Visini *et al.*, 2019).

Example of a SHERIFS Calculation

We present here a simplified example of application of SHERIFS for the five southern faults of the WCR (Greece) with a discussion of the resulting models. We hope this can be used as an illustration of the type of discussion that can be supported by SHERIFS. Two key features of SHERIFS have been used: a hazard model has been generated and a fundamental discussion on the input hypotheses has taken place. All the calculation files for this example can be found with the SHERIFS code (see Data and Resources). Using the SHERIFS graphical interface, the model was run (please refer to the user manual for a detailed description of the user interface). For this example, the logic tree presented in Figure 6 is used with 20 random samples on M_{max}, b-value (between 0.95 and 1.05), and the slip rates of the faults (Table 1). For this run, an increment of slip-rate dsr of 0.001 mm/yr was chosen.

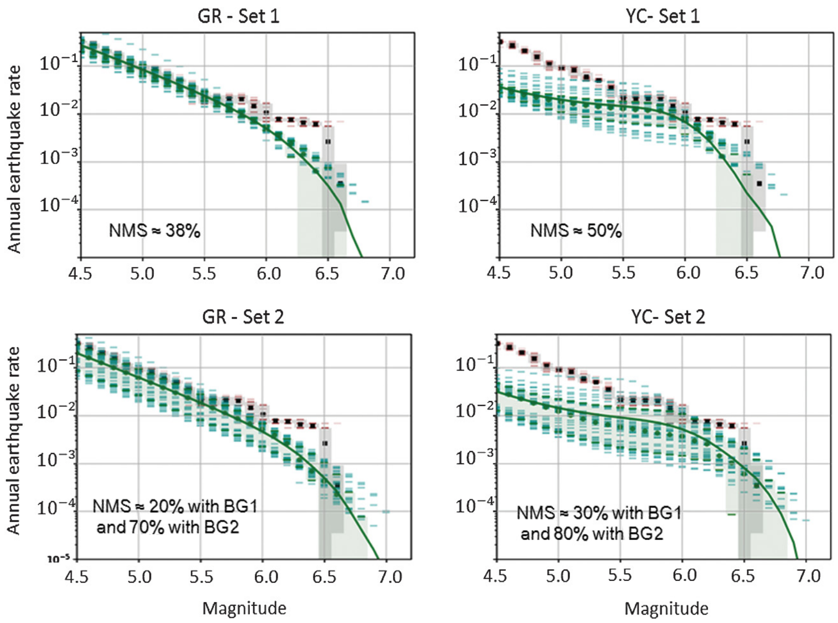
In the branch BG1, 100% of the seismicity of all magnitudes is considered to be modeled on the faults. In the hypothesis BG2, 60%, 70%, 80%, 90%, 95%, and 100% of the seismicity is considered to be on the faults for magnitude M_w 4.0, 4.5, 5.0, 5.5, 6.0, and 6.5 and more, respectively. In the set of rupture scenarios set 1, faults F1, F2, and F3 can rupture together in scenarios of two or three faults and fault F5 can only rupture alone. In the branch set 2, all four faults can rupture together. Finally, two hypotheses of target MFD are explored: GR and modified YC (see the Ⓔ supplemental content for the complete formulation of the modified YC equation).

After running the model visualization tool, it is possible to compare the modeled earthquake rates with those of the catalog for each hypothesis of the logic tree.

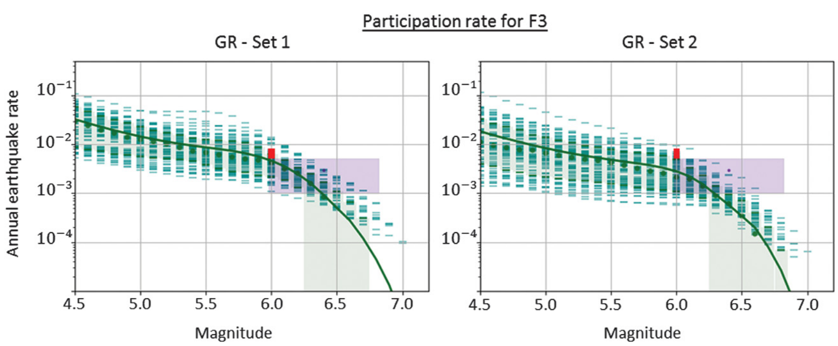
SHERIFS is a tool that allows discussing how each branch and each hypothesis of the logic tree performs against the

Table 1
Faults Parameters Used for the Example Calculation

Model_Name	Fault Name	Dip	Oriented	Mechanism	Seismogenic Depth (km) Upper	Seismogenic Depth (km) Lower	Slip Rate (mm/yr) Min	Slip Rate (mm/yr) Mean	Slip Rate (mm/yr) max	Domain	Shear Modulus (GPa)
Example_Model	F1	60	N	N	0	6	4.8	5.0	5.2	Active_Shallow_Crust	30
Example_Model	F2	55	N	N	0	7	3.0	3.2	3.4	Active_Shallow_Crust	30
Example_Model	F3	60	N	N	0	7	3.8	4.0	4.2	Active_Shallow_Crust	30
Example_Model	F5	60	N	N	0	7	3.3	3.5	3.7	Active_Shallow_Crust	30



▲ **Figure 7.** For each hypothesis of MFD shape and rupture scenario set, comparison between the modeled MFD of the whole logic tree and the earthquake rate calculated from the catalog. Dashed green lines are the MFD of each individual model, the solid green line is the mean MFD, and green patches represent the uncertainty (16–84 percentiles). The dotted black line is the rate from the catalog with uncertainties. The ratio of NMS is indicated for each background hypothesis.



▲ **Figure 8.** Each green dotted line is the sum of the rates of all different ruptures passing through this fault. The solid green line is the mean rate. The purple dot is the paleoearthquake rate and the purple box is the associated uncertainty. The red dot is the historical earthquake rate and the red box is the associated uncertainty.

earthquake catalog and paleoearthquake data. This analysis also allows us to understand which hypotheses on this simple fault system are able to reconcile the geologic data (slip rate, geometry of the faults) with the seismological and paleoseismic data.

Models using a target MFD shape GR match the rate of the catalog better than the YC MFD (Fig. 7). Given the large discrepancy between the modeled rates and the catalog rates for the YC hypothesis, this branch can have a weight of 0 in the logic tree. Between the sets 1 and 2, set 1 leads to slightly higher rates than set 2 which matches the rate of the catalog better. In this respect, the branch set 1 can be weighted stronger than the branch set 2 in the logic tree.

The modeled rate can be compared at the fault level with the rate calculated from paleoseismology data (Fig. 8). Here for fault 3, in which some trenches have been conducted, we compare the rate of all earthquake ruptures that involved fault 3 to the rate calculated from the paleoseismology study.

Although both rupture sets hypotheses agree with the rate of paleoearthquakes, within the range of uncertainty, the set 2 hypothesis leads to an average annual rate closer to the paleoseismic rate and therefore could be weighted stronger than set 1. Interestingly even though the regional target MFD is of the shape of a GR, the participation rates of the fault can, and often do, differ from the GR distribution. Similar observations have been made with different system level approaches such as integer programming (Geist and Parsons, 2018) and simulated annealing (Field *et al.*, 2014).

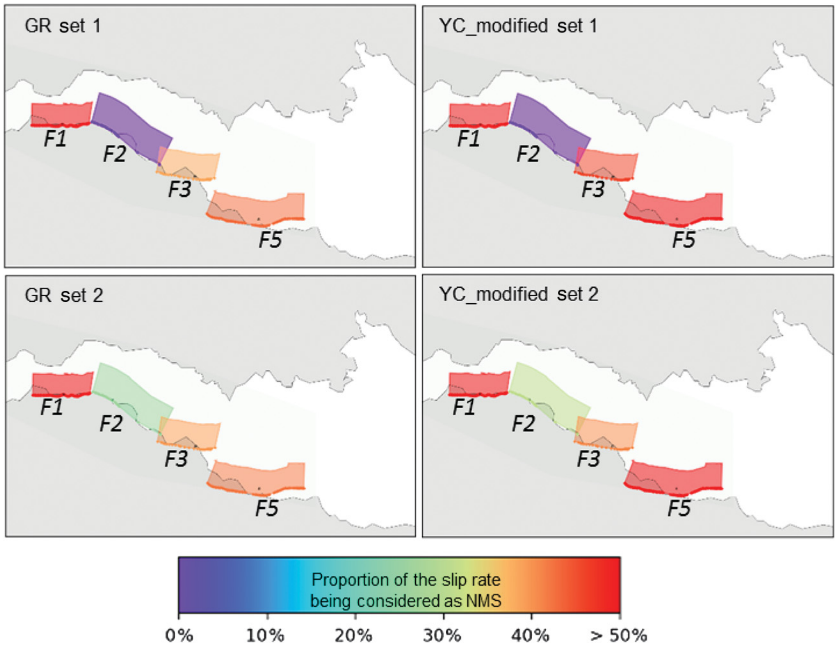
On the NMS slip map (Fig. 9), we can identify, which faults have trouble spending their slip-rate budget given the set of hypothesis. This can help identifying some possible issues with the model input hypotheses.

An NMS ratio of more than 30% for a fault might indicate some modeling issues for a particular model. In this example, all branches of the logic tree lead to large NMS slip on most of the faults of the system. This is due to this example system being a subset of the faults of the WCR fault system and not including the other faults of the system, and; therefore, the possible FtF ruptures with those faults lead to some difficulties to accommodate the input slip rate as seismic slip rate in the SHERIFS approach. In other words, identifying the faults belonging to a fault system is paramount.

LIMITATIONS OF SHERIFS

Although SHERIFS allows considering faults systems as a whole and modeling complex earthquake ruptures in seismic-hazard assessment models, it remains a simplification of the natural systems and several limitations have to be noted before any use of this methodology. Although, SHERIFS explores a wide range of epistemic uncertainties, a full exploration of the epistemic uncertainty would require comparison with other fault modeling approaches (Field *et al.*, 2014; Geist and Parsons, 2018; Visini *et al.*, 2019).

SHERIFS is not able to consider complex slip distributions along strike for a given rupture. In the present version, the methodology considers the spending of a slip-rate budget is uniform for all faults participating in a given rupture. SHERIFS being an open-source code, it can be modified at will, and we



▲ **Figure 9.** Map of the NMS slip ratio for each branch of the logic tree.

encourage any user willing to develop and test this feature to do so.

When considering only a part of a fault system, SHERIFS will not resolve the rates at the edge of the system correctly likely because ruptures with faults that lie outside of the modeled area were omitted. Users should verify the sites of interest are located in the center of the modeled area, where the hazard assessment is not sensitive to changes in the definition of the limits of the modeled area, ensuring all sources potentially contributing to the hazard for the site are correctly modeled.

Testing of SHERIFS on different fault systems has led to important improvements regarding the versatility of the code. Although many cases can now be supported by SHERIFS, it is not impossible some applications may require further adaptations of the code. Hence, we suggest SHERIFS be used with care, and the results be always checked by the end user.

To allow the code to be applicable to very diverse fault systems and hypotheses, a few controls on the input files have been implemented. It is the duty of the user to make sure the input files are consistent with the hypotheses wanted for the calculation. SHERIFS provides output files and figures to help judging the validity of the results. The output figures of SHERIFS are not in any case exhaustive, and we encourage any user who requires more specific outputs to take part in the development of Python routines producing these outputs and to share these developments with the community under the same GNU Lesser General Public License.

CONCLUSION AND PERSPECTIVES

SHERIFS is an open-source Python tool to calculate the rate of earthquakes in a fault system while relaxing segmentation hypotheses. It relies on three basic ingredients: an assumption on the shape of the regional MFD, hypotheses on the possible FtF ruptures, and preservation of slip-rate variability along faults, if needed. SHERIFS's objective is to provide data collectors and fault modelers with a friendly and easy-to-use interface to explore coherency of input hypotheses and the impact of uncertainties in data and input hypotheses on the resulting seismic hazard. Output files from SHERIFS are formatted to be directly usable as input files for the PSHA calculation. Future developments will include tools to analyze the hazard outputs with complex logic trees.

We hope the seismic-hazard models generated using SHERIFS can represent more accurately the vision of the data providers.

DATA AND RESOURCES

Seismic hazard and earthquake rate in fault systems (SHERIFS) can be accessed free of charge at https://github.com/tomchartier/SHERIFS and can be used under the license (GNU Lesser General Public License). This open-source code can be freely modified by the user as long as the modified code is also available with the same license. Questions about SHERIFS can be asked on the Google group (https://groups.google.com/forum/#!forum/sherifs). This Google group aims to be a platform to share the questions related to the use of SHERIFS for the modeling of fault systems. We hope to be able to answer the questions and to improve the points in which users identified shortcomings of SHERIFS. The information about Fault2SHA European Seismological Commission (ESC) working group is available at https://Fault2SHA.net/. The detail of each of the windows is explained in the user manual (provided with the code: www.github.com/tomchartier/SHERIFS). All websites were last accessed on January 2019.

ACKNOWLEDGMENTS

The authors would like to thank Ned Field and Bruno Pace for their valuable comments and suggestions that helped reshape and improve both the article and the user manual to a great extent.

Development of the seismic hazard and earthquake rate in fault systems (SHERIFS) code was supported financially by Institut de Radioprotection et de Sûreté Nucléaire, Ecole Normale Supérieure, and the Axa Research Fund Joint Research Initiative "Earthquake geology and seismic hazard assessment: tools for risk-informed decision-making."

The authors would like to warmly thank the beta testers of SHERIFS, in particular Octavi Gomez Novell, Etienne Marti, Stephane Baize, Ludmila Provost, Hervé Jomard, and the participants at the training of the December 2017 Fault2SHA Paris, France.

REFERENCES

Aguilar-Meléndez, A., M. G. Ordaz Schroeder, J. De la Puente, S. N. González-Rocha, H. E. Rodriguez-Lozoya, A. Córdova-Ceballos, A. García-Elías, C. Calderón-Ramón, J. E. Escalante-Martínez, J. R. Laguna-Camacho, and A. Campos-Rios (2017). Development and validation of software CRISIS to perform probabilistic seismic hazard assessment with emphasis on the recent CRISIS2015, *Computación y Sistemas* **21,** no. 1, 67–90.

Barka, A. A., and K. Kadinsky-Cade (1988). Strike-slip fault geometry in Turkey and its influence on earthquake activity, *Tectonics* **7,** 663–684.

Chartier, T., O. Scotti, H. Lyon-Caen, and A. Boiselet (2017). Methodology for earthquake rupture rate estimates of fault networks: Example for the western Corinth rift, Greece, *Nat. Hazards Earth Syst. Sci.* **17,** no. 10, 1857–1869.

Danciu, L., K. Şeşetyan, M. Demircioglu, L. Gülen, M. Zare, R. Basili, A. Elias, S. Adamia, N. Tsereteli, H. Yalçın, and M. Utkucu (2018). The 2014 earthquake model of the middle east: Seismogenic sources, *Bull. Earthq. Eng.* **16,** 3465–3496.

Field, E. H., R. J. Arrowsmith, G. P. Biasi, P. Bird, T. E. Dawson, K. R. Felzer, D. D. Jackson, K. M. Johnson, T. H. Jordan, C. Madden, and A. J. Michael (2014). Uniform California earthquake rupture forecast, version 3 (UCERF3)—The time-independent model, *Bull. Seismol. Soc. Am.* **104,** no. 3, 1122–1180.

Field, E. H., T. H. Jordan, and C. A. Cornell (2003). OpenSHA: A developing community-modeling environment for seismic hazard analysis, *Seismol. Res. Lett.* **74,** no. 4, 406–419.

Geist, E. L., and T. Parsons (2018). Determining on-fault earthquake magnitude distributions from integer programming, *Comput. Geosci.* **111,** 244–259.

Gutenberg, B., and C. F. Richter (1944). Frequency of earthquakes in California, *Bull. Seismol. Soc. Am.* **34,** no. 4, 185–188.

Hornblow, S., M. Quigley, A. Nicol, R. Van Dissen, and N. Wang (2014). Paleoseismology of the 2010 Mw 7.1 Darfield (Canterbury) earthquake source, Greendale fault, New Zealand, *Tectonophysics* **637,** 178–190.

Knuepfer, P. L. K. (1989). Implications of the characteristics of endpoints of historical surface fault ruptures for the nature of fault segmentation, *U.S. Geol. Surv. Open-File Rept. 89-315*, 193–228.

Leonard, M. (2010). Earthquake fault scaling: Self-consistent relating of rupture length, width, average displacement, and moment release, *Bull. Seismol. Soc. Am.* **100,** no. 5A, 1971–1988.

Pace, B., F. Visini, and L. Peruzza (2016). FiSH: MATLAB tools to turn fault data into seismic-hazard models, *Seismol. Res. Lett.* **87,** no. 2A, 374–386.

Pagani, M., D. Monelli, G. Weatherill, L. Danciu, H. Crowley, V. Silva, P. Henshaw, L. Butler, M. Nastasi, L. Panzeri, and M. Simionato (2014). OpenQuake engine: An open hazard (and risk) software for the global earthquake model, *Seismol. Res. Lett.* **85,** no. 3, 692–702.

Schwartz, D. P. (2018). Review: Past and future fault rupture lengths in seismic source characterization—The long and short of it, *Bull. Seismol. Soc. Am.* **108,** no. 5A, 2493–2520, doi: 10.1785/0120160110.

Thingbaijam, K. K. S., P. Martin Mai, and K. Goda (2017). New empirical earthquake source-scaling laws, *Bull. Seismol. Soc. Am.* **107,** no. 5, 2225–2246.

Visini, F., A. Valentini, T. Chartier, O. Scotti, and B. Pace (2019). Computational tools for relaxing the fault segmentation in probabilistic seismic hazard modelling in complex fault systems, *Pure Appl. Geophys.* 1–23.

Wang, Y. J., C. H. Chan, Y. T. Lee, K. F. Ma, J. B. H. Shyu, R. J. Rau, and C. T. Cheng (2016). Probabilistic seismic hazard assessments for Taiwan, *Terr. Atmos. Ocean. Sci.* **27,** no. 3, 325–340.

Wells, D. L., and K. J. Coppersmith (1994). New empirical relationships among magnitude, rupture length, rupture width, rupture area, and surface displacement, *Bull. Seismol. Soc. Am.* **84,** no. 4, 974–1002.

Wesnousky, S. G. (1988). Seismological and structural evolution of strike-slip faults, *Nature* **335,** 340–343.

Woessner, J., D. Laurentiu, D. Giardini, H. Crowley, F. Cotton, G. Grünthal, G. Valensise, R. Arvidsson, R. Basili, M. B. Demircioglu, and S. Hiemer (2015). The 2013 European seismic hazard model: Key components and results, *Bull. Earthq. Eng.* **13,** no. 12, 3553–3596.

Youngs, R. R., and K. J. Coppersmith (1985). Implications of fault slip rates and earthquake recurrence models to probabilistic seismic hazard estimates, *Bull. Seismol. Soc. Am.* **75,** no. 4, 939–964.

Thomas Chartier[1]
Hélène Lyon-Caen
Laboratoire de Géologie
CNRS UMR8538, Ecole Normale Supérieure
PSL University
24 rue Lhomond
75005 Paris, France
chartier@geologie.ens.fr
helene.lyon-caen@ens.fr

Oona Scotti
Institut de Radioprotection et de Sûreté Nucléaire
31 avenue de la Division Leclerc
92260 Fontenay aux Roses, France
oona.scotti@irsn.fr

Published Online 15 May 2019

[1] Also at Institut de Radioprotection et de Sûreté Nucléaire, 31 avenue de la Division Leclerc, 92260 Fontenay aux Roses, France.