PyBet Tools User Guide

Table of contents

abilistic tools aiming to provide probabilistic estimation of volcanic unrest, of eruption, and of any to volcanic activity. For this reason, they can be used for probabilistic forecasting during crisis of unrest prehensive probabilistic volcanic hazard assessments (PVHA). The background theory of these models d-hoc Bayesian event trees, where each node is treated through a Bayesian inference approach. This ether different kind of available information, as theoretical/empirical models, a-priori beliefs, past data r for short-term analysis).

current software implementations of these models, with the specific goal of providing a user-friendly
e softwares have been developed by using the Python programming language, since the broad
to install them on all the most common operating systems (Linux, Mac OSX, Windows) and also on the
nere they can be freely run online or downloaded. In the next sub-sections they are briefly introduced.

dels background theory can be found in literature (among others see Marzocchi et al., 2004; Marzocchi
2010) and are out of the scope of this document. However, some basic principles will be briefly
2, in order to properly approach the use of the corresponding tools.

, free, open-source and cross-platform software implementation of the Bayesian Event Tree for Volcanic
se of the tool is to provide a graphical support to BET_VH model, which estimates the long-term
azardous phenomenon (i.e., lava flows, tephra fall, pyroclastic flows, lahars, etc.) occurring in a selected
the uncertainties. The BET_VH model represents a flexible tool to provide probabilities of any specific
ted in, by merging all the available information, such as theoretical models, a priori beliefs, and past
e Bayesian inference and it deals with long-term forecasting only, therefore it can be useful in land use
found in Marzocchi et al. (2010) and it has been used in several applications (Selva et al., 2010; Sandri
4; Thompson et al., 2015; Tonini et al., 2015b)

urves, which describe the distribution of the exceedance probability as a function of a user-selected
a load) on a grid of points covering the target area. The computed hazard curves are (i) absolute
of eruption in a given time frame, and for all the possible vent locations and eruptive sizes) and (ii)
ent percentiles, in order to quantify the epistemic uncertainty). Such curves allow representation of the
PVHA, and are well suited to become a main input to quantitative risk analyses. PyBetVH allows for
h the computed hazard curves, and the corresponding Bayesian hazard/probability maps. PyBetVH is
rts of end users, making PVHA results accessible to people who may be less experienced in
.g. decision makers.

can be found in Tonini et al. (2015a), where the use of the tool is illustrated through the example of
he Okataina Volcanic Centre (OVC), New Zealand, by highlight the range of outputs that the tool can

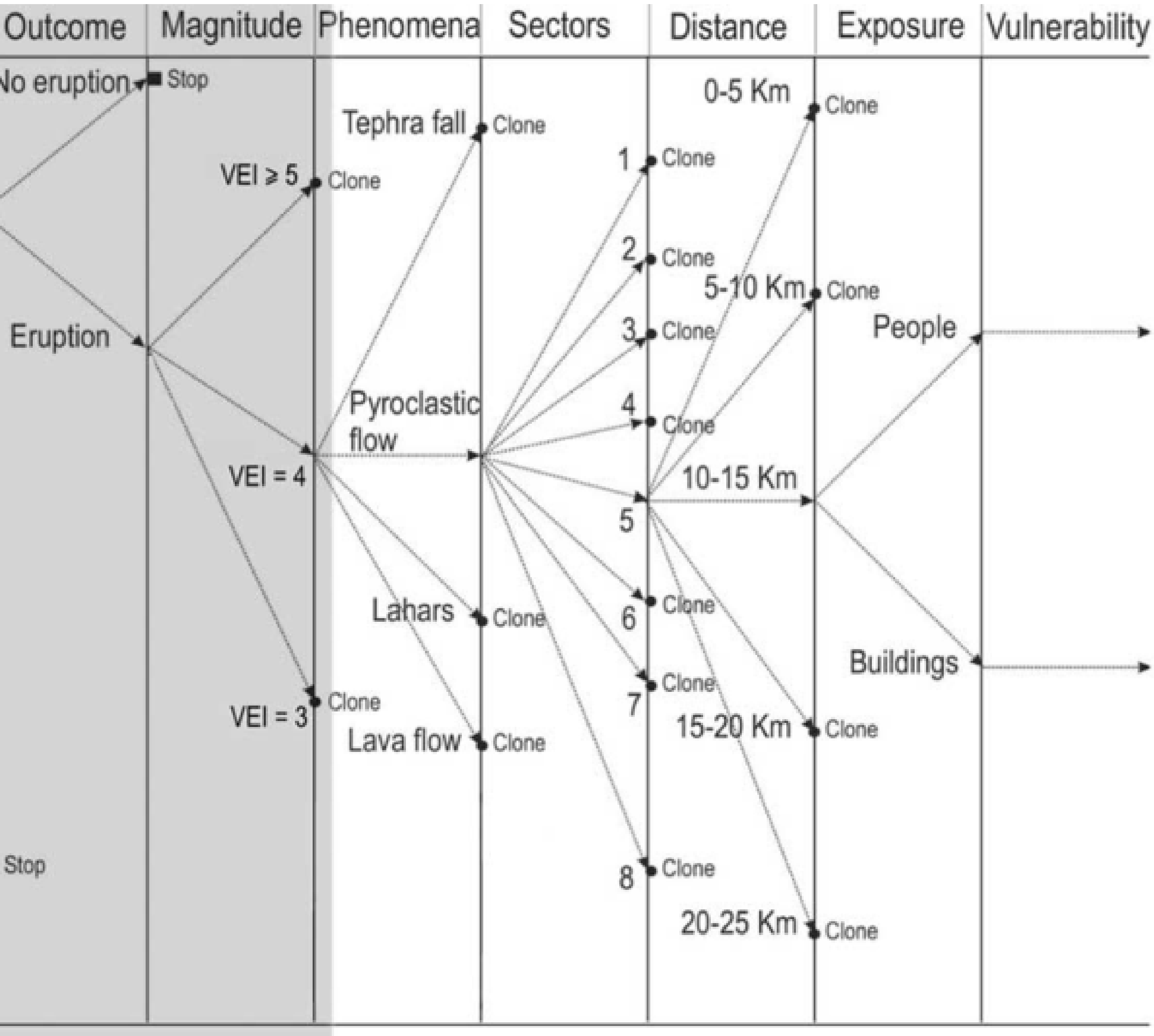
s-platform implementation of BET_EF (Bayesian Event Tree for Eruption Forecasting) equipped with a
_EF model aims to provide probabilities of unrest/eruption forecasting of magmatic events, by merging
mation such as theoretical models, a priori beliefs, monitoring measures, and any kind of past data.
ian procedure and it relies on the fuzzy approach to manage monitoring data. The method deals with
ng; therefore, it can be useful in many practical aspects such as land use planning and volcanic
., 2008). BET_EF model has been applied to several volcanic systems (Marzocchi et al., 2008; Sandri et
Selva et al., 2012b). **NOTE:** PyBetEF is not maintained anymore, since it has been included

cross-platform implementation of BET_UNREST model. BET_UNREST is an extension of BET_EF
anic unrest and its relating hazardous phenomena, by adding a specific branch to the event tree (Rouwet
gned and developed as one of the final products of the EU VUELCO project and tested during the last
anized in Dominica (West Indies) in the frame of VUELCO (Sandri et al., Under review).

tool can be found in Tonini et al. (2016), where the use of the tool is illustrated through an application to
sets of the Kawah Ijen stratovolcano, Indonesia. In particular, the tool is set on the basis of monitoring
000–2010, and is then blindly applied to the test period 2010–2012, during which significant unrest

eory

ET models are the definition of an event tree (ET, hereafter) and the use of the Bayesian inference. ET is
f events in which individual branches are alternative steps from a general prior event, state, or condition
subsequent events (intermediate outcomes) to final outcomes (see Figure 1). In this way, an ET shows
s of volcanic unrest at progressively higher degrees of detail. The branches at each node represent
ough these need not be mutually exclusive or exhaustive. The Bayesian approach is used at the
retical models of the eruptive process, past (historical and geological) data, and, for short-term
ring of the volcano. Thus it allows to formally account for both aleatoric (due to the intrinsic randomness
(due to limited data or knowledge) uncertainties. Through the ET, one can calculate both absolute and
orresponds to consider the selected full path or just a single node in the ET itself.



ee as designed after Marzocchiet al. (2004) for Mount Vesuvius.

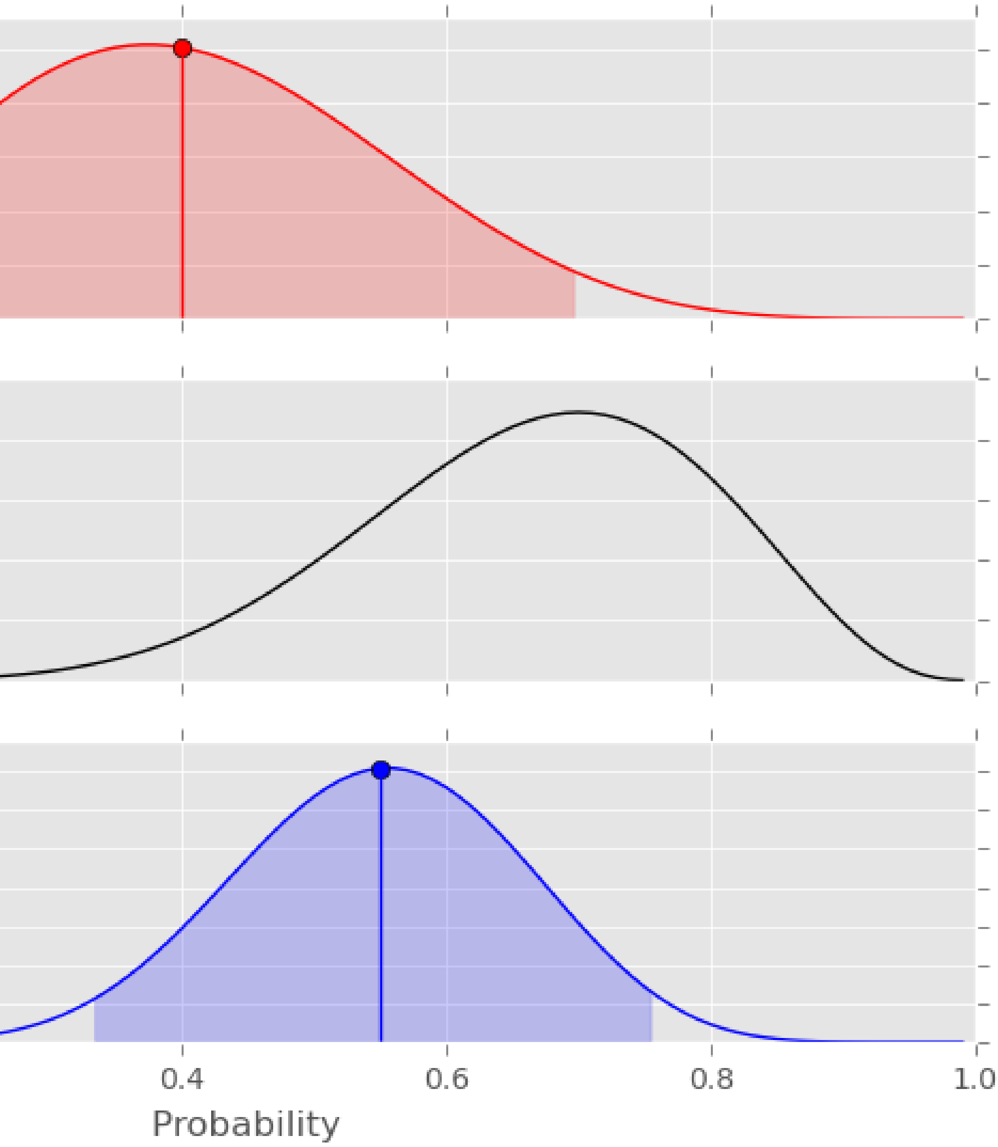
ce

lows to include a formal probabilistic treatment of the available data and their uncertainties and each esented by a probability density function (pdf) of the probability at the node. Each node is assigned a esian approach, meaning that a prior probability distribution (usually coming from theoretical models) a are statistically combined together to obtain a posterior probability. Quantities are estimated by nformation (analytical/empirical models, expert beliefs, data) as described in the following formulation

][data | \theta]$$

or distribution (subjective beliefs about a parameter)
the likelihood function (available data)
the posterior distribution (updated beliefs after having observed data)

and posterior have a Beta distribution.



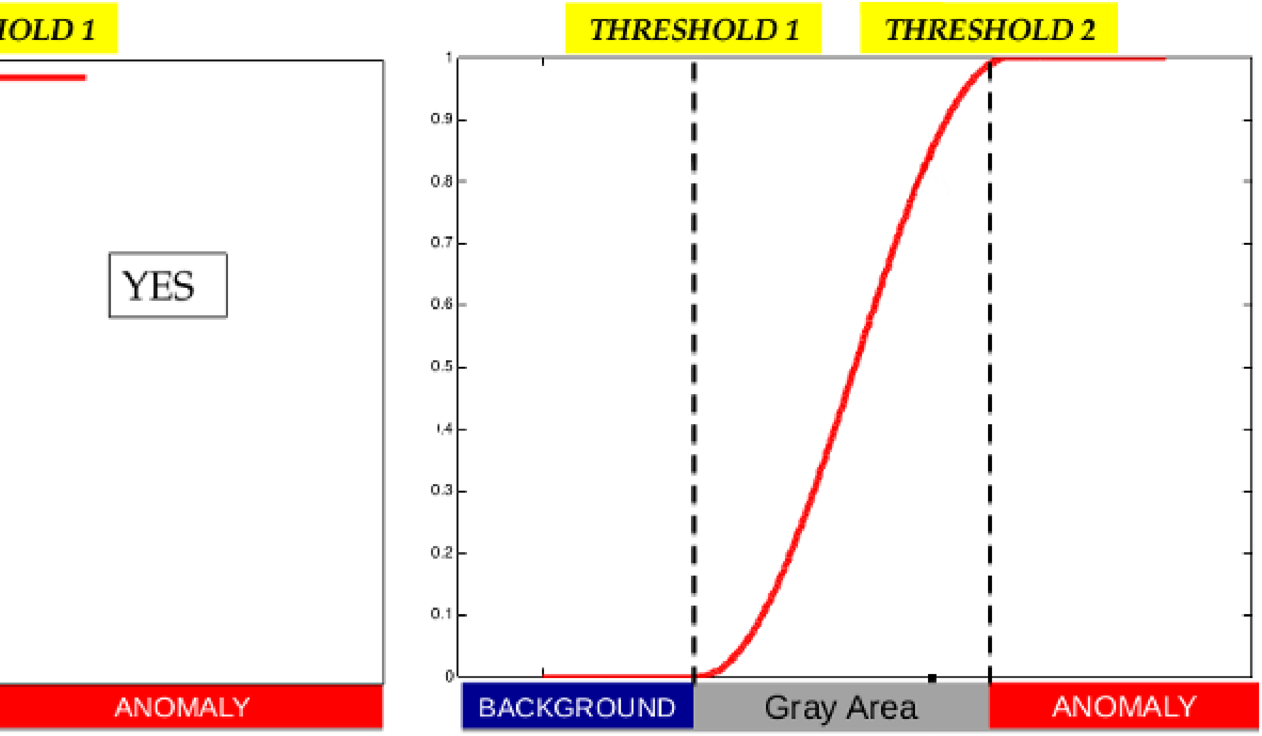
: posterior distribution as result of prior model updated through available observed data.

nalysis procedure can be summarized by the following practical concepts:

native, the posterior is very much determined by the data (the posterior is data-driven)
ive, the posterior is a mixture of the prior and the data

the prior, the more data you need to change your beliefs

d by identifying all the monitored parameters that can provide information about a degree of anomaly at ntification of anomalies is where volcanologists introduce their conceptual models about the pre-eruptive naly of each measure is computed in two ways (see Figure 3 here below): the parameter's anomaly can ance of one threshold), as shown in the left panel of Figure 3, or through a fuzzy approach (by defining he right panel of Figure 3.



eters' thresholds

parameters have been identified and the corresponding thresholds defined, it is possible to compute for degree of anomaly any time new data become available. For the first node (unrest), the identification of r is enough to set the probability of unrest to one. For nodes 2 (magmatic unrest) and 3 (eruption) the ee of anomaly at the node into a conditional probability. The total degree of anomaly at the node is

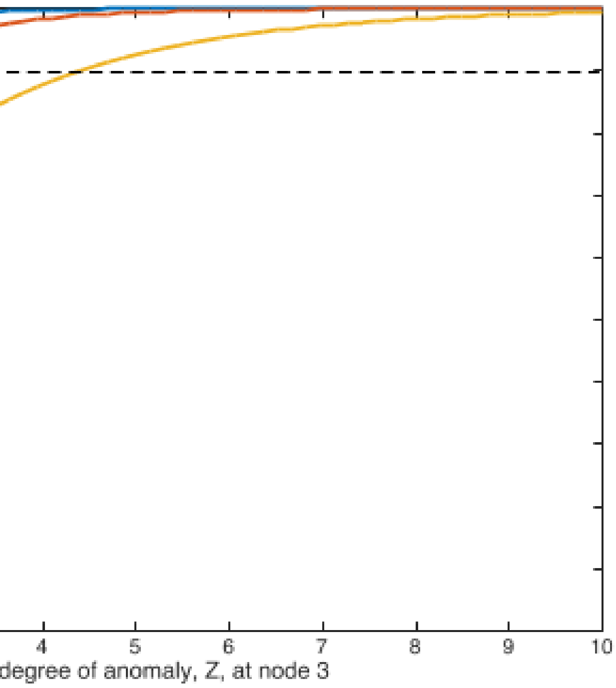
$_i\ w_i$

$_i$ is the

P_i the weight given at parameter i.

hip between te degree of anomaly Z and the entropy score H (defined by logarithm of (1-P)) which is a of the event of interest. In other words, when the degree of anomaly doubles, the predictability ree of anomaly is used to estimate the mean of a Beta distribution describing the conditional probability or eruption (node 3) through:

e defined for any volcanic systems, since their meaning is strictly connected to specific configuration of
nd the expert's beliefs on the selected volcano.



ith the y axis) is set from 1-a that is the probability that the event occurs when we observe Z=0 (no
for weakly monitored volcanoes, and smaller for well monitored volcanoes. The default for the parameter
. (at node 3, this means that there is 0.10 probability to have an eruption in the next time window without
node).

of the curve) is set defining what is the value of Z for which there is 0.90 probability that the event
esponds to about 0.90 probability of occurrence if 2 parameters are recognized as anomalous. The
so depends on the weighting scheme. We consider that the maximum weight for one parameter is 1.

equation (*) is the mean of a Beta distribution. The variance of the distribution is set by the parameter
ivalent number of data. The **equivalent number of data** mimics how much we believe to the
behind the choice of the anomalies. In other words, if we set $$\Lambda=5$$ we think that our model
ve can get from a probability calculated from 5 data. Of course the larger $$\Lambda$$, the higher the

on Requirements

n and some Python third party modules/libraries which need to be installed in order to be able to run the
k if the Python interpreter is installed. By typing “python” in a terminal window followed by the Enter key,
lar response:

```
 Apr 20 2012, 22:39:59)

t", "credits" or "license" for more information.
```

les/libraries are:

user interface library)
nputing library)
 library)
ormer Python Image library)

nd freely available and usable for all the most common operating systems (Linux, Mac OSX and
operating system and on the experience of the user in using its machine and/or Python, the procedure
e approached in different ways. **NOTE:** The current versions of all BET tools run on Python 2.7.x. The
he way.

utions, Python libraries can be easily installed by means of their own package managers, which install,
he system in a consistent manner, by checking for all the required dependencies and shared libraries.
you can use the synaptic tool or simply the following command from the terminal:

```
ython-numpy python-matplotlib python-wxgtk3.0 python-pil
```

butions is very similar.

ws

do not have a native package manager to handle libraries and softwares in a smart way as Linux
of Python libraries could not be so straightforward. One can download and install the single libraries
being very careful to the Python Standard Library version installed on his/her own computer as
ct files. Moreover, Mac OSX users can try to use MacPorts (https://www.macports.org/) or Homebrew
n their own expertise in using these tools, which should play as package managers similarly to the ones
lternatively, as easiest way, one can install the Enthought Canopy free version
Canopy is a very rich Python package distribution, which will install a scientific Python environment in
but more solid and it takes care of the compatibility among all the libraries. Once you have installed
e in your own operating system), one should set the Canopy Python environment as the default Python.
nd for Windows and Mac OSX.

e above procedure does not set the Python Canopy environment as your default Python, you can
le as follows:

```
URUSERNAME/Library/Enthought/Canopy_64bit/User/bin:$PATH
```

default Python environment, by running the python command from your terminal, you should get

```
on
n 2.7.9 | 64-bit | (default, Jun 30 2015, 19:41:21)
pple Inc. build 5658) (LLVM build 2335.6)] on darwin
t", "credits" or "license" for more information.
```

erification

g system you are using and on the method used to install these Python libraries, their correct
entering in the Python interactive shell and manually importing the libraries:

```
 Apr 20 2012, 22:39:59)

t", "credits" or "license" for more information.


age
```

are satisfied, you can proceed with the installation of the tool(s) on your system.

ols

om the PyBetVH main page on the VHub web site (Download does not require to be registered users). on the right, by clicking on download, it will be asked to save a compressed file (betvh-rxxx.tar.gz). The older where the user has full permissions (reading, writing and executing) on files. This will be the and should not be moved anymore. After having extracted the files from the tar.gz file, the following opear:

```
 ls -hot betvh-rxxx

Jan  8 22:04 bin
Jan  8 22:02 src
Nov  9 19:01 data
Nov  9 19:01 doc
Nov  9 19:01 examples
Nov  9 19:01 middleware
Nov  9 19:01 rappture
Nov  9 19:01 LICENSE.txt
Nov  8 16:56 COPYING
```

oy VHub repository. After having opened a terminal window and moved into src/ folder, it must be and:

```
$ make install all
```

all the executable files in the bin/ directory. Open a terminal and move inside the bin/ folder:

```
rpath/betvh-rxxx/bin
```

with:

```
betvh.py
```

system, you can now set a link to the main executable, which is in /yourpath/betvh-rxxx/bin/betvh.py.
c OSX platforms:

```
+x /yourpath/betvh-rxxx/bin/betvh.py
yourpath/betvh-rxxx/bin/betvh.py /somewhereinyourPATH/PyBetVH
```

ove have been successfully concluded, the tool can be launched by simply doing, from any position in
al window:

n VHub. You can download it from here together with a working example. Uncompress the .zip file and
you prefer. The folder contains the PyBetEF main program and all the modules needed to be properly
t modify, remove or do any change to this folder, since the risk is to compromise the functionality of the
unched for a first trial. Open a terminal and move inside the folder containing the PyBetEF source files:

```
r/path/to/PyBetEF/folder/
```

```
PyBetEF.py
```

ke this operation faster. In Linux and Mac OSX systems, as illustrative example, one can make the e executable and then create a (symbolic) link to it in a folder belonging to its PATH environment

```
+x /yourpath/PyBetEF/PyBetEF.py
yourpath/PyBetEF/PyBetEF.py /somewhereinyourPATH/PyBetEF
```

e to run the tool from any place of her/his filesystem by opening a terminal and simply doing:

link (i.e., on the Desktop) by clicking with the mouse's right button ans selecting create a link. Then l directly launch the tool.

ou have to unzip the file and place the FakeVolcano folder wherever in your file system. The folder PyBetEF. Instructions on how to prepare the input files are given in Input preparation section. The next k on Load Volcano button in the upper left panel and select the FakeVolcano folder. If everything went r this example is loaded in the lower left panel and the user can explore it by clicking on the nodes. Once the user can press the COMPUTE button on the lower right panel. A new window frame will be opened, output.

hown in 3.1 for PyBetVH

on

sists in a set of text files organized in a main folder representing the volcano in which the user is nformation is driven by a configuration file called pybet.cfg, which must always exist in order to run the xt file in which the user can input the general information about the volcano, and must be formatted as The user should not change the names of the various blocks (identified by square brackets) and neither ft of the sign '=', but only fill in the spaces on its right (for instance, in the block "Main Settings", the able name and cannot be changed, while on the right the user can set the values for the desired case

tion to PyBetVH can be found in Tonini et al. (2015) main paper and in the relating supplementary

```
lcano
4, 5115607

 45

ename/None
55000, 570000, 5108000, 5123000
None
eters = 0.1, 1.0, 1.0

5
mbda) = 1
= 0

5
mbda) = 1
= 0

n]
5
mbda) = 1
= 0

cation]
txt
mbda) = 1
 node_vent_monitoring.txt/None

False
```

.txt

s"

e =": the user should add volcano name

er =": the user should add longitude and latitude in UTM coordinates (meters),

s allows to select the geometry of the possible vent positions. Please add the keyword "Field" for a
r a caldera, corresponding to a rectangular grid of possible vent locations, or "Cone" for a central
acterized by a central vent position over the summit crater area, and 4 lateral sectors

geometry "Field" is selected in the previous line, the user should input 5 values (comma separated): the
ongitude length (in meters) of the field, the along-latitude length (in meters) of the field, the number of
ns (Nc) to divide the field along the longitude, the number of rows (Nr) to divide the field along the
, and the strike (in decimal degrees) from North;

geometry "Cone" is selected in the previous line, the user should input 3 values (comma separated): the
of the central vent position above the summit crater area (in meters), the outer radius (to the apron) of
cano edifice (in meters), the strike (in decimal degrees) from North to rotate clockwise the four
rees outer sectors;

the user should add the UTM zone together with the hemisphere, N=North or S=South (i.e, 33N or

=": the user should add a number that will determine the exposure time. For example, if the user wants
assesment for a 1 year exposure time, he must input the value "1" here, and keep in mind that time unit

the user should add the number of samples that PyBetEF will draw at each node of the event tree to
mic uncertainty. Acceptable values (compromise between accuracy and runtime) are between 400 and
user could insert any value (integer).

map =": the user should enter a name of file (in png format) containing the background image to be
ntour hazard maps or probability map. If keyword "None" is provided, the tool will try to connect to the
load a map from Google Maps based on the limits provided in the next item. A good practice could be to
wnloading a map from Google Maps the first time, then inserting here the filename of the downloaded
ther runs of the tool (the current PyBetEF version can load .png image format only).

: the user should provide SW and NE corners of the background map in UTM coordinates (m).

: the user should provide file name where monitoring parameters are set.

ction parameters =": a, b, lambda

e 1)

he user should input the prior probability value (a number between 0 and 1) for the volcano entering
et al 2008; 2010 for more details)

mbda) =": the user should input a number expressing how much he is confident on the prior probability.
ts a fictitious number of data he would like to collect before changing significantly his mind on that prior
s range from 1 (very low confidence on the prior probability value) to (theoretically) infinite; however, a
eady a very high confidence on the prior probability value. Please note that the combination "prior
"equivalent n. data (lambda) = 1" expresses the maximum ignorance probability distribution (uniform
nain !0,1!)

s) =”: the user should input the number of observed unrest episodes in the available record (that must
that it is reasonable to assume that no unrest has been lost over that time period covered by the

he user should input the number of observed time windows, in the time unit defined in the block “Main
e example, the number of time windows of 1 year), beginning in a state of NO UNREST. The user
e windows where there was already unrest from the beginning

should input the conditional prior probability value (a number between 0 and 1) of the unrest being due
t (see Marzocchi et al 2008; 2010 for more details)
=”: as above but related to the prior probability of magmatic unrest
e user should input the number of unrest episodes of known magmatic origin in the complete record
should input the number of unrest episodes of known origin (magmatic, hydrothermal, tectonic) in the

3)
nput the conditional prior probability value (a number between 0 and 1) of the magmatic unrest leading
ee Marzocchi et al 2008; 2010 for more details)
bove but related to the prior probability of magmatic unrest leading to eruption
hould input the number of magmatic unrest episodes having led to an eruption in the complete record
input the number of magmatic unrest episodes in the complete record

e of the text file in which the prior spatial probability of vent opening, and past occurrences at each vent
low for its structure).
ut related to the prior spatial probability of vent opening

lse” (if there is no reason to postulate a different probability distribution of the sizes depending on the
ample for volcanic structures that are partially submerged by the sea, implying a different explosivity
der water)
erent size classes (Ns) characterising magmatic eruptions
text file in which the prior probability of the different size classes, and the past occurrences of
node_style.txt”, see below for its structure)

esponding to one possible vent position. The number Nv of possible vent positions is implicitly declared
geometry. If geometry is “Field”, Nv is equal to Nr*Nc (see above in the block “Main Settings”); if it is
each line of the file, there must be 2 numbers: the first one is the prior probability of that vent position,
ataset) occurred from that vent position. The sum of the prior probabilities over all the Nv vent positions
represent a complete and mutually exclusive set).

n node 4 (“node 4-5 dependence = False” in the “Node Style” block of pybet.cfg), in this file there must

es of the Ns size classes (and the sum must be 1)
f data of such prior probability distribution
he number of past eruptions (in a complete dataset) of each size class
(“node 4-5 dependence = True” in the “Node Style” block of pybet.cfg), in this file there must be Nv

tus of the selected monitored quantities. The path to the this file is given by setting the parameter
te the status of monitoring parameters, he/she can create a new file by simply coping the old one and

er of the parameter (i.e, Parameter 1, Parameter 2, …) for each node. Then, it follows a series of rows that

sholds;

played by hazard assessments in civil protection issues prevents us from leaving the user free to use all
rovide by himself/herself all of the information of the volcano he/she is interested in.

ations. By using or downloading one of the BET tools, the user accepts that INGV and authors are not

ties of volcanic events: the example of volcanic hazard at Mt. Vesuvius. J. Geophys. Res., 109, B11201,

term eruption forecasting , Bull. Volcanol., 70, 623-632, doi:10.1007/s00445-007-0157-y
anic hazard assessment, Bulletin of Volcanology, Volume 72, Number 6, 705-716, doi:

014) Recognizing and tracking volcanic hazards related to non-magmatic unrest: a review. J. Appl.

ort-term probabilistic volcanic hazard assessment with cost-benefit analysis to support decision making
–723. http://dx.doi.org/10.1007/s00445-011-0556-y.
i-hazard assessment for El Misti volcano (Peru). Bull. Volcanol. 76 (2), 1–26. doi:

ernard B., VUELCO Book chapter, Under Review.
natural uncertainties on long-term hazard from tephra fallout at Campi Flegrei (Italy), Bulletin of

ng at high-risk volcanoes: The case of Campi Flegrei, Naples, J. Appl. Volcanol., 1, 5.
p for future vent opening at the Campi Flegrei caldera, Italy, Bull. Volcanol., 74, 497–510,

(2015) Exploring the influence of vent location and eruption style on tephra fall hazard from Okataina
926-y
ic hazard assessment and for generation of Bayesian hazard curves and maps. Comput Geosci

nerged vents on probabilistic hazard assessment for tephra fallout. Nat. Hazards Earth Syst. Sci. 15,

sian Event Tree tool to track and quantify volcanic unrest and its application to Kawah Ijen volcano,