

## Special Issue: V3-LAVA PROJECT

# Retrospective validation of a lava-flow hazard map for Mount Etna volcano

Annalisa Cappello<sup>\*,1,2</sup>, Annamaria Vicari<sup>1</sup>, Ciro Del Negro<sup>1</sup>

<sup>1</sup> *Istituto Nazionale di Geofisica e Vulcanologia, Sezione di Catania, Osservatorio Etneo, Catania, Italy*

<sup>2</sup> *Università di Catania, Dipartimento di Matematica e Informatica, Catania, Italy*

## Article history

Received November 29, 2010; accepted June 30, 2011.

## Subject classification:

*Lava-flow paths, MAGFLOW simulator, Fitting scores, Temporal validation.*

## ABSTRACT

*This report presents a retrospective methodology to validate a long-term hazard map related to lava-flow invasion at Mount Etna, the most active volcano in Europe. A lava-flow hazard map provides the probability that a specific point will be affected by potential destructive volcanic processes over the time period considered. We constructed this lava-flow hazard map for Mount Etna volcano through the identification of the emission regions with the highest probabilities of eruptive vents and through characterization of the event types for the numerical simulations and the computation of the eruptive probabilities. Numerical simulations of lava-flow paths were carried out using the MAGFLOW cellular automata model. To validate the methodology developed, a hazard map was built by considering only the eruptions that occurred at Mount Etna before 1981. On the basis of the probability of coverage by lava flows, the map was divided into ten classes, and two fitting scores were calculated to measure the overlap between the hazard classes and the actual shapes of the lava flows that occurred after 1981.*

## Introduction

Mount Etna in Sicily (Italy) is the largest active volcano in Europe. Over the last 400 years, it has erupted over sixty times from vents on its flanks, while the eruptive activity at its summit has been nearly continuous. Lava flows generated during flank eruptions can potentially damage towns located at medium-low elevations, and might reach the Ionian coast where the town of Catania is located.

Volcanic hazard maps are the most intuitive way to represent areas that can be damaged by dangerous volcanic processes, such as lava flows, ash falls, lahars, and pyroclastic explosions [Damiani et al. 2006]. They are usually very useful to identify regions that are differently ranked based on the probability that they will be affected by a specific volcanic event over a specific time period. For lava-flow invasion, several hazard maps have already been suggested for Mount Etna volcano, most of which derived from quantitative analysis of past eruptions [Andronico and Lodato 2005,

Behncke et al. 2005]. Wadge et al. [1994] proposed a probabilistic approach that was based on the historical record of past eruptions, to constrain a series of Monte Carlo simulations. Favalli et al. [2009] applied a similar approach to obtain a hazard map, with the simulation of the inundation areas for a large number of possible future eruptions using an empirical relationship for the maximum length of the lava flows. Recently, Crisci et al. [2010] elaborated a lava-invasion susceptibility map that was based on the results of numerical simulations of flows that erupted from a grid of 393 hypothetical vents that are located on the eastern sector of Mount Etna. To validate the results of the proposed methodology, Crisci et al. [2010] simulated a set of 13 documented actual eruptions on the present-day topography, and evaluated the overlap between the affected areas and the susceptibility classes.

In recent years, we have made significant progress in hazard assessment at Mount Etna through the development of accurate and robust physical–mathematical models that can forecast the spatial and temporal evolution of lava flows. In particular, during the 2001, 2004, 2006, 2008 and 2011 effusive eruptions at Mount Etna, we obtained encouraging results using the MAGFLOW cellular automata (CA) model [Vicari et al. 2007, Del Negro et al. 2008, H erault et al. 2009, Vicari et al. 2009, Vicari et al. 2011]. MAGFLOW was developed at INGV-Catania to simulate lava-flow paths and the temporal evolution of lava emplacement. More recently, it has also been applied to more sophisticated issues, including prospective hazard mitigation on Mount Etna [Scifoni et al. 2010], and for compiling a new lava-flow invasion hazard map for Mount Etna volcano [Cappello et al. 2011].

In the present study, we present a new validation procedure for the evaluation of the accuracy of lava-flow invasion hazard maps. The validation approach consists of splitting the dataset that contains the information about eruptions at Mount Etna volcano into two parts: eruptions

that occurred before 1981, which was chosen as the reference time, and eruptions that happened after 1981. The first information was used as the learning dataset, to develop a lava-flow hazard map up to 1981, which was then discretized into ten classes. The second dataset represents the testing dataset: the actual shape of the lava flows that occurred after 1981 were superimposed on the hazard map, and two scores were calculated to measure their percentages of overlap.

**Lava-flow hazard map by a MAGFLOW model**

The lava-flow invasion hazard map at Mount Etna was developed by considering the main volcanological structures and the information on past eruptions. The methodology used is based on the following steps: (i) computation of the susceptibility map that provides the spatio-temporal probabilities of vent opening; (ii) characterization of the expected eruptions; (iii) simulations of lava-flow paths by the MAGFLOW model; (iv) computation of the lava-flow invasion hazard map; and (v) discretization of the resulting map into a number of fixed hazard levels.

The computation of the susceptibility map was based on the characterization of the zones with the highest probabilities of the opening of future eruptive vents. We defined a grid of potential vents that was spaced at regular intervals of 500 m, in both the vertical and horizontal directions, and that assigned a probability of activation ( $p_a$ ) to each potential vent, as follows:

$$p_a(x, y, \Delta t) = (1 - e^{-\lambda_t \Delta t})(1 - e^{-\lambda_{xy} \Delta x \Delta y}) \quad (1)$$

where  $\lambda_t$  and  $\lambda_{xy}$  are the temporal and spatial recurrence rates, respectively,  $\Delta t$  is the time interval of 50 years, and  $\Delta_x \Delta_y$  is an area of 0.25 km<sup>2</sup>.

For the temporal recurrence rate  $\lambda_t$ , we used the results obtained for the power intensity function that was applied to the Tanguy et al. [2007] catalog by Smethurst et al. [2009]. For the spatial recurrence rate  $\lambda_{xy}$ , we used the most common spatial point process model, which is based on computation of a kernel function. A kernel function is a probability density function that is symmetric about the origin and that spreads the probability away from the structure considered. We decided to use the Gaussian kernel, although other different kernel functions could have been used, including the Cauchy kernel [Martin et al. 2004] and the Epanechnikov kernel [Lutz and Gutmann 1995]. Connor and Hill [1995] and Lutz and Gutmann [1995] demonstrated that the shape of the kernel function chosen in this type of analysis generally has a trivial impact on the probability calculations, compared to other parameters. The formula of the Gaussian kernel is given by:

$$\lambda_{xy} = \frac{1}{2\pi N h^2} \sum_{i=1}^N e^{-\frac{d_i^2}{2h^2}} \quad (2)$$

where  $d_i$  is the distance from the point  $(x,y)$  to the  $i$ -th volcanic structure location,  $h$  is a smoothing parameter that controls the size of the zone to which each data point contributes an increased intensity, and  $N$  is the number of volcanic structures that are considered in the calculation. As  $N$  occurs in the denominator, the integral  $\lambda_{xy}$  across the whole area will be unity.

The smoothing parameter was calculated by comparing the observed nearest-neighbor distances of the main vents with the expected distribution of the nearest-neighbor distances [Weller et al. 2006]. As discussed in Cappello et al. [2011], plots for  $h = 750$  m and  $h = 2,200$  m gave the upper and lower bounds, respectively, to the cumulative nearest-neighbor distances of the volcanic events of Mount Etna. Hence, a medium value (1,500 m) between the lower and upper bounds was chosen and used for Equation (2).

For the characterization of the expected events, a statistical analysis of flank eruptions that have occurred on Mount Etna in the last 400 years was performed, and five eruption classes were obtained by combining three different values of the emitted lava volumes (30, 100 and  $>100 \times 10^6$  m<sup>3</sup>) with two times of eruptions (30, and  $>30$  days). A sixth type of event that was characterized by a duration of 30 days and a lava volume of more than  $100 \times 10^6$  m<sup>3</sup> was introduced as an extreme event, as the past cannot represent the future perfectly.

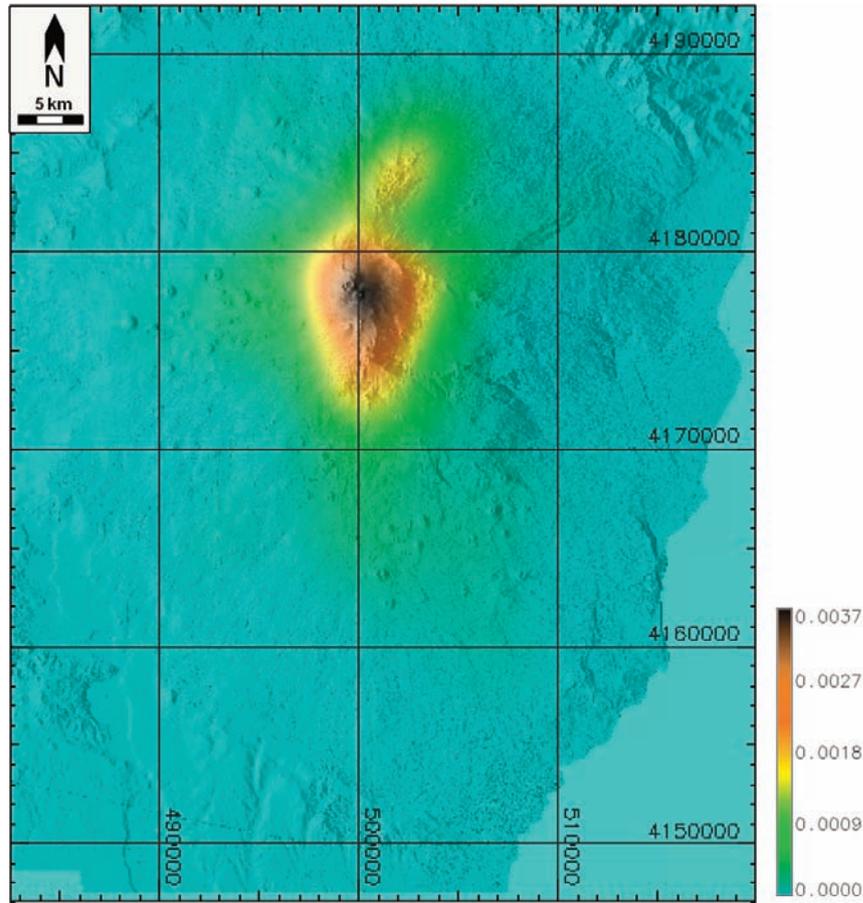
The event probability  $p_e$ , represents the probability that the  $e$ -th class of eruption occurs at the point  $(x,y)$ , and this was estimated as follows:

$$p_e(x, y) = \sum_{j=1}^{n_e} d_j^{-1} \quad (3)$$

where  $n_e$  is the total number of events in the  $e$ -th class, and  $d_j$  is the distance between point  $(x,y)$  and the  $j$ -th vent if  $d_j$  is less than or equal to  $h_e$ . Therefore, the value of  $p_e(x,y)$  depends on the number of main vents within a distance  $h_e$  of the point. The estimate  $p_e$  for each point  $(x,y)$  satisfies the following:

$$\sum_{e=1}^6 p_e(x, y) = 1 \quad (4)$$

Lava flow simulations were performed with the MAGFLOW code, which has been extensively used in lava-flow hazard applications for Mount Etna. The MAGFLOW model is based on the CA structure, in which the states of the cells are the thickness of the lava and the quantity of the heat. The states of the cells are synchronously updated according to local rules, which depend on the values of the cell and the values of the neighbors within a certain proximity. In this way, the CA can produce extremely complex structures from the evolution of rather simple and local rules. The evolution function of MAGFLOW is a steady-state solution of Navier-Stokes equations for the



**Figure 1.** Susceptibility map at Mount Etna, as updated using dikes, faults and eruptive fractures until 1981. The colors are associated with different levels of vent opening probability.

motion of a Bingham fluid on a plane that is subject to pressure force, in which the conservation of mass is guaranteed both locally and globally [Vicari et al. 2007]. The rheological properties were modeled using a variable viscosity relationship according to Giordano and Dingwell [2003], which was parameterized in terms of the temperature and water content.

For each potential vent of the grid, we launched six simulations that correspond to the six expected eruptions. The effusion rate functions that are required to run MAGFLOW were established by following the information relating to historic eruptions. The short-medium time and the long time were set to 30 and 90 days of simulation respectively, and 30, 100 and 200 million cubic meters were fixed as the erupted lava volumes. From the combination of these values, we obtained six possible functions that represent the variations of the flux rates in relation to the times of the eruptions. The curves were considered to be a kind of bell shape, in which the eruption starts from a null value of flux rate, reaches its peak after a  $\frac{1}{4}$  of the entire time of simulation, and then gradually decreases until the end of the eruption.

The lava-flow invasion hazard map was computed for each point  $(x,y)$  of the area covered by the flow simulations, by taking into account the information on the lava-flow

overlap, the activation probability of vents  $v_i$ , and the event probability of eruptive classes  $c_j$ , as follows:

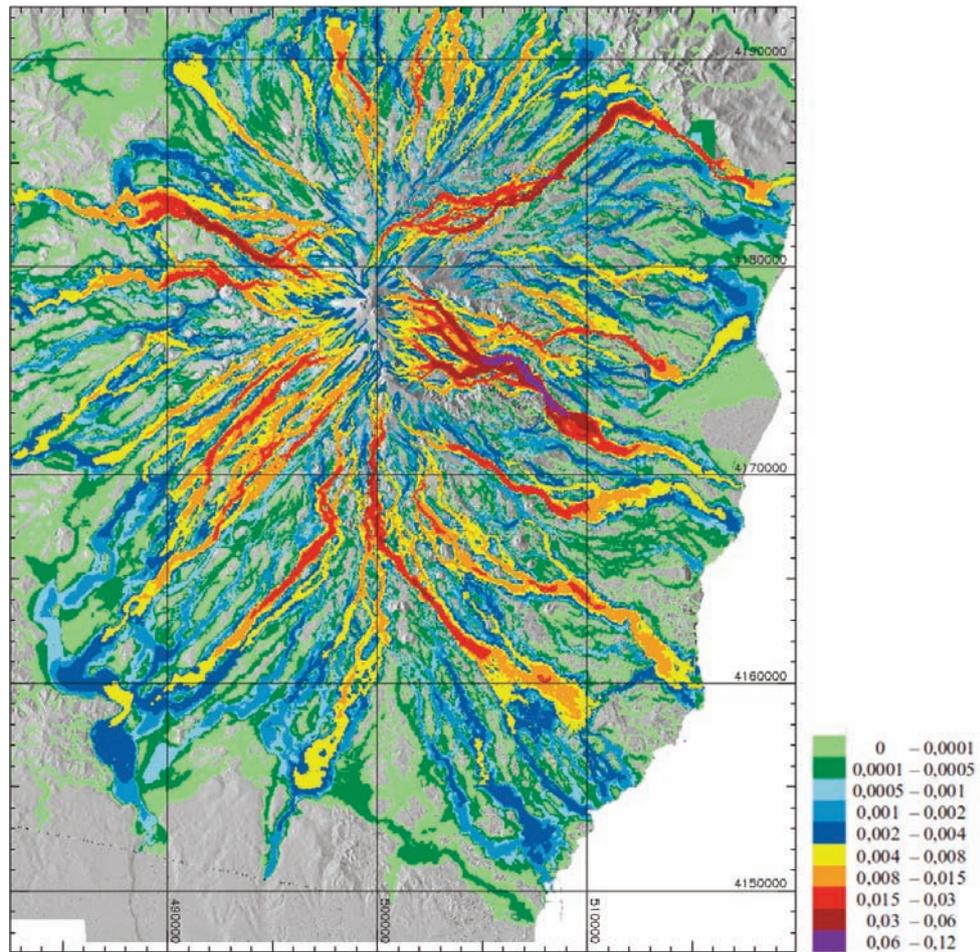
$$\text{Hazard}(x, y, \Delta t) = \sum_{i=1}^N \sum_{e=1}^6 [p_a(v_i) * p_e(v_i)] \quad (5)$$

if  $(x,y)$  is invaded by a simulated lava flow that erupted from  $v_i$ , or else.

The value assigned to each point represents the probability of it being affected by a lava flow during the time interval considered. Finally, to make the map more readable, we discretized it into ten hazard intervals.

Parameter	Value	Unit
Density of lava	2600	$\text{kg m}^{-3}$
Specific heat	1150	$\text{J kg}^{-1} \text{K}^{-1}$
Emissivity of lava	0.9	—
Temperature of solidification	1143	K
Temperature of extrusion	1360	K

**Table 1.** Typical parameters for Mount Etna lava flows.



**Figure 2.** Lava-flow hazard map at Mount Etna, constructed using the learning dataset (reference time 1981). The colors represent different levels of lava-flow invasion hazards.

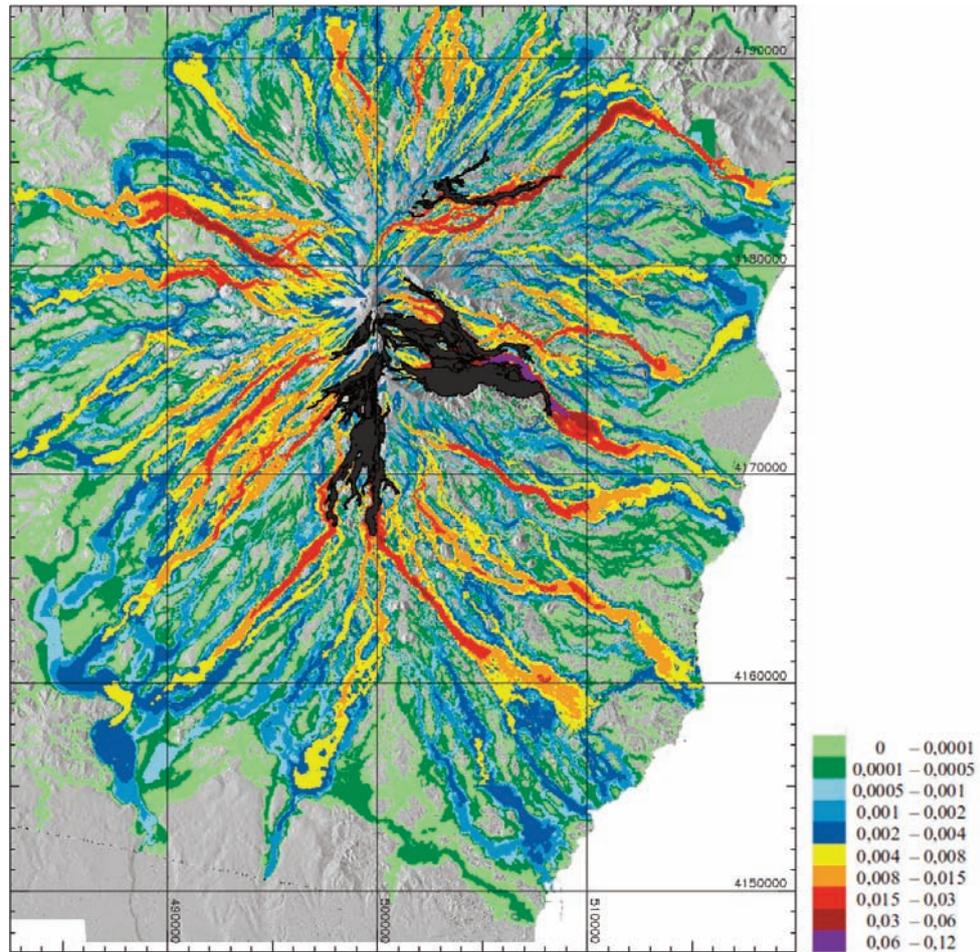
Hazard class	Area (km <sup>2</sup> )	Probability level ( $p_i$ )
[0 - 0.0001]	247.12	0.0001
[0.0001 - 0.0005]	203.54	0.0005
[0.0005 - 0.001]	109.27	0.001
[0.001 - 0.002]	129.20	0.002
[0.002 - 0.004]	139.43	0.004
[0.004 - 0.008]	119.64	0.008
[0.008 - 0.015]	81.38	0.015
[0.015 - 0.03]	38.41	0.03
[0.03 - 0.06]	11.15	0.06
[0.06 - 0.12]	1.67	0.12

**Table 2.** Probabilities and areas of the ten hazard intervals into which the lava-flow invasion hazard map was discretized.

### Validation of the hazard map

The validation of the lava-flow hazard map is the last crucial step of the whole process, as this provides a quantitative evaluation of its reliability. As our knowledge only includes the historical information, we propose a retrospective approach that uses the more recent past eruptions as the potential future events. Our strategy is based on the following concept: if the invasion probabilities reflect the past lava flows, then it is reasonable that they will also predict the areas that will be most likely to be invaded in the future. Hence, the comparison can evaluate the probability distribution of lava-flow hazards at the reference time against the historical lava-flow paths that occurred successively.

The first step of the validation consists of dividing the available dataset into two parts that are based on the ages of the past eruptions: eruptions that occurred before 1981, and eruptions that happened after 1981. This time limit was chosen for two reasons. From a mathematical point of view, as the reference time, the year 1981 determines two statistically meaningful sets, which divides the dataset into 49 events used for learning and 14 used for testing. From a volcanological point of view, starting from 1981, the activity of Mount Etna volcano showed a temporal trend (an increase



**Figure 3.** Lava-flow hazard map for Mount Etna constructed using the learning dataset (reference time 1981) with superimposed actual lava-flow paths from eruptions that occurred after 1981 (in black). The colors represent different levels of lava-flow invasion hazards.

in the number of eruptions), which is seen in particular by the activity over the last 20 years or so [Salvi et al. 2006].

The learning dataset is used to calculate for each point of the study area the probability of a vent opening  $p_a$  for a time period of 25 years (Equation 1), and the event probability  $p_e$  (Equation 3). The susceptibility map obtained is shown in Figure 1. Then we ran the lava-flow simulations through the MAGFLOW model, using the typical material properties of Etnean basaltic lava (Table 1). The basis for the numerical simulations was a digital elevation model of the volcano area updated to 2005, which was represented as a 10 m cell grid and a regular vent grid (of 36 km × 32.5 km) that was spaced in regular intervals of 500 m, in both the vertical and horizontal directions. Using Equation (6), we developed the lava-flow hazard map for Mount Etna volcano up to 1981. A logarithm scale was then applied (Figure 2) to discretized it into ten levels of hazard (Table 2).

The testing dataset includes 14 eruptions that occurred after 1981 (Table 3), and this was superimposed on the hazard map (Figure 3). The two scores were then calculated to measure the overlap.

The first validation index  $VI_1$  represents the percentage that the observed values hit the highest hazard class, and therefore the best fit corresponds to 100. This was calculated as:

Eruption year	Onset	End
1983	March 28, 1983	August 6, 1983
1985	March 12, 1985	July 13, 1985
1986-1987	October 30, 1986	March 1, 1987
1989	September 27, 1989	October 9, 1989
1991-93	December 14, 1991	March 31, 1993
2001 (2700 m)	July 18, 2001	August 9, 2001
2001 (2100 m)	July 18, 2001	August 9, 2001
2001	July 26, 2001	August 2, 2001
2002	October 27, 2002	November 4, 2002
2002-2003	October 27, 2002	January 28, 2003
2004-2005	September 7, 2004	March 8, 2005
2006	October 27, 2006	November 27, 2006
2006	September 4, 2006	December 7, 2006
2007	April, 2007	May, 2007

**Table 3.** Eruptions that occurred after 1981, as used for the validation.

$$VI_1 = \frac{100}{C * p_{max}} \sum_{i=1}^N (C_i * p_i) \quad (6)$$

where  $C$  is the total number of pixels invaded by the lava-flow simulations,  $C_i$  and  $p_i$  are the number of pixels and the probability level, respectively, for the  $i$ -th hazard class,  $p_{max}$  is the highest observed value of hazard, and  $N$  is the total number of hazard classes. We obtained a value for  $VI_1$  of 17.3, which is not significant as it only depends on the extension of the highest hazard class, not on the expected distribution of all of the classes.

For this reason, we introduced another validation index,  $VI_2$ , that represents the mean distance between the observed and expected values for each class of probability. This is 0 if the values provide a perfect fit, while it increases if the discrepancy increases. This has been calculated as:

$$VI_2 = \frac{1}{N} \sum_{i=1}^N \left| \frac{C_i}{C} - p_i \right| \quad (7)$$

where  $C$  is the total number of pixels invaded by the lava-flow simulations,  $C_i$  and  $p_i$  are the number of pixels and the probability level, respectively, for the  $i$ -th hazard class, and  $N$  is the total number of hazard classes. We obtained a  $VI_2$  of 0.09.

## Conclusions

A lava-flow invasion hazard map defines the probability that a specific area will be reached by a lava flow during a given time window. Many different approaches have been developed for the generation of lava-flow hazard maps at Mount Etna volcano [Wadge et al. 1994, Favalli et al. 2009, Crisci et al. 2010, Cappello et al. 2011], which have confirmed that the main difficulties regard specifically not the elaboration of the map, but also the verification of its consistency. Ideally, the validation should include a further independent experiment that is run forwards in time and that uses a significant number of events to make the statistical tests valid. Since further forward studies are unworkable from a scientific point of view, the only chance to validate the map is to re-use the datasets from which it was derived.

The strategy described in the present study represents a possible approach that is based on a retrospective analysis of the hazard, as it looks at what happened in the past to predict what will happen in the future. We believe that this a plausible approach, because it exploits the only knowledge we have, which is the eruptions that occurred in the recent past.

Nevertheless our study must be considered as a preliminary study of how lava-flow invasion hazards can be validated in volcanic areas. The application of the retrospective validation shows that the two validation indices provide a quantitative evaluation that relates not really to the hazard map, but to the methodology developed to elaborate it. Moreover, the simulations used to generate the lava-flow invasion hazard map up to 1981 were performed by using a

digital elevation model updated to 2005. From a statistical point of view, this change should not cause any large variation in the results of the two validation indices. However, for some specific eruptions, it is possible that the solidified lava represented a limiting factor, or even a barrier, and therefore changed the time-space evolution of the simulated lava flows.

**Acknowledgements.** This study was performed with the financial support from the V3-LAVA project (INGV-DPC 2007-2009 contract). The authors thank the anonymous reviewer for the helpful and constructive comments.

## References

- Andronico, D. and L. Lodato (2005). Effusive activity at Mount Etna volcano (Italy) during the 20th century: a contribution to volcanic hazard assessment, *Nat. Hazards*, 36, 407-443.
- Behncke, B., M. Neri and A. Nagay (2005). Lava flow hazard at Mount Etna (Italy): new data from a GIS-based study, *Geol. S. Am. S.*, January 2005, 396, 189-208.
- Cappello, A., A. Vicari and C. Del Negro (2011). Assessment and modeling of lava flow hazard on Mt Etna volcano, *B. Geofis. Teor. Appl.*, 52 (2), 299-308.
- Connor, C.B., and B.E. Hill (1995). Three nonhomogeneous Poisson models for the probability of basaltic volcanism: application to the Yucca Mountain region, *J. Geophys. Res.*, 100 (B6), 10107-10125.
- Crisci, G.M., M.V. Avolio, B. Behncke, D. D'Ambrosio, S. Di Gregorio, V. Lupiano, M. Neri, R. Rongo and W. Spataro (2010). Predicting the impact of lava flows at Mount Etna, Italy, *J. Geophys. Res.*, 115, B04203; doi: 10.1029/2009JB006431.
- Damiani, M.L., G. Groppelli, G. Norini, E. Bertino, A. Gigliuto and A. Nucita (2006). A lava flow simulation model for the development of volcanic hazard maps for Mount Etna (Italy), *Comput. Geosci.-UK*, 32 (4), 512-526.
- Del Negro, C., L. Fortuna, A. Hérault and A. Vicari (2008). Simulations of the 2004 lava flow at Etna volcano by the MAGFLOW cellular automata model, *B. Volcanol.*, 70 (7), 805-812; doi: 10.1007/s00445-007-0168-8.
- Favalli, M., F. Mazzarini, M.T. Pareschi and E. Boschi (2009). Topographic control on lava flow paths at Mt. Etna (Italy): implications for hazard assessment, *J. Geophys. Res.*, 114, F01019; doi: 10.1029/2007JF000918.
- Giordano, D. and D.B. Dingwell (2003). Viscosity of hydrous Etna basalt: implications for Plinian-style basaltic eruptions, *B. Volcanol.*, 65, 8-14; doi: 10.1007/s00445-002-0233-2.
- Hérault, A., A. Vicari, A. Cirauco and C. Del Negro (2009). Forecasting lava flow hazards during the 2006 Etna eruption: Using the MAGFLOW cellular automata model, *Comput. Geosci.-UK*, 35 (5), 1050-1060; doi: 10.1016/j.cageo.2007.10.008.
- Lutz, T.M. and J.T. Gutmann (1995). An improved method

- for determining and characterizing alignments of point-like features and its implications for the Pinacate volcanic field, Sonora, Mexico, *J. Geophys. Res.*, 100 (B9), 17,659-17,670.
- Martin, A.J., K. Umeda, C.B. Connor, J.N. Weller, D. Zhao and M. Takahashi (2004). Modeling long-term volcanic hazards through Bayesian inference: an example from the Tohoku arc, Japan, *J. Geophys. Res.*, 109, B10208; doi: 10.1029/2004JB003201.
- Salvi, F., R. Scandone and C. Palma (2006). Statistical analysis of the historical activity of Mount Etna, aimed at the evaluation of volcanic hazard, *J. Volcanol. Geoth. Res.*, 154, 159-168.
- Scifoni, S., M. Coltelli, M. Marsella, C. Proietti, Q. Napoleoni, A. Vicari and C. Del Negro (2010). Mitigation of lava flow invasion hazard through optimized barrier configuration aided by numerical simulation: the case of the 2001 Etna eruption, *J. Volcanol. Geoth. Res.* 192, 16-26.
- Smethurst L., M.R. James, H. Pinkerton and J.A. Tawn (2009). A statistical analysis of eruptive activity on Mount Etna, Sicily, *Geophys. J. Int.*, 179 (1), 655-666.
- Tanguy, J.-C., M. Condomines, M. Le Goff, V. Chillemi, S. Le Delfa and G. Patané (2007). Mount Etna eruptions of the last 2,750 years: revised chronology and location through archeomagnetic and Ra-230Th dating, *B. Volcanol.*, 70, 55-83.
- Vicari, A., A. Hérault, C. Del Negro, M. Coltelli, M. Marsella and C. Proietti (2007). Modeling of the 2001 lava flow at Etna volcano by a cellular automata approach, *Environ. Modell. Softw.*, 22, 1465-1471.
- Vicari A., A. Cirauda, C. Del Negro, A. Hérault and L. Fortuna (2009). Lava flow simulations using effusion rates from thermal infrared satellite imagery during the 2006 Etna eruption, *Nat. Hazards*, 50 (3), 539-550; doi: 10.1007/s11069-008-9306-7.
- Vicari, A., G. Ganci, B. Behncke, A. Cappello, M. Neri and C. Del Negro (2011). Near-real-time forecasting of lava flow hazards during the 12-13 January 2011 Etna eruption, *Geophys. Res. Lett.*, 38, L13317; doi: 10.1029/2011GL047545.
- Wadge, G., P.A.V. Young and I.J. McKendrick (1994). Mapping lava flow hazards using computer simulation, *J. Geophys. Res.*, 99 (B1), 489-504.
- Weller, J.N., A.J. Martin, C.B. Connor, L.J. Connor and A. Karakhanian (2006). Modelling the spatial distribution of volcanoes: an example from Armenia, In: H.M. Mader, S.G. Coles, C.B. Connor and L.J. Connor (eds.), *Statistics in Volcanology*, Special Publications of IAVCEI no. 1, Geological Society of London, 77-88.

---

\*Corresponding author: Annalisa Cappello,  
Istituto Nazionale di Geofisica e Vulcanologia, Sezione di Catania,  
Osservatorio Etneo, Catania, Italy; email: annalisa.cappello@ct.ingv.it.