Data uncertainty management in volcanic hazard assessment: review and examples

Alessandro Tadini

Bevilacqua A., Bisson M., Neri A.

Pisa – 06/09/2017

Presentation Outline

- Theoretical definition of Uncertainty (epistemic and aleatoric) in volcanology

- Examples of uncertainty quantification in volcanology (at Somma-Vesuvio)
  - Spatial uncertainty
  - Epistemic/aleatoric uncertainty in vent opening probability maps
  - Uncertainty in measuring eruptive parameters
  - Uncertainty after numerical modelling of volcanic phenomena

Uncertainty in volcanology

The *geological information* is often affected by a relevant **uncertainty**: particularly, volcanoes can be seen as **random systems** that must be assessed with **uncertain information**.

Hazard assessment, risk management and urban planning need to have a quantitative estimation of the reliability/effectiveness of thematic products related to hazard/risk assessment.

Uncertainty quantification (UQ) and probability forecasts of the future behavior of the volcanic systems require specific **statistical** approaches. Two types of Uncertainty are classically recognized:

I. the **aleatoric uncertainty** (from the Latin "alea" = dice) or **physical variability,** i.e. the intrinsic randomness of the system under study,

II. the **epistemic uncertainty** (from the Greek "episteme" = knowledge) due to the imperfect knowledge of the system.

Aleatoric uncertainty can be estimated better (or even quantified) but cannot be reduced through advances in theoretical or observational science (Woo, 1999). It is sometimes referred as a property of the system.

Epistemic uncertainty is instead a lack of knowledge which might be remedied in principle by further learning and experiment. It is sometimes referred as a property of the analyst.

The types of **epistemic uncertainty** are numerous and can reflect:

- unknown/uncertain **data** related to events occurred in the past
- degrees of belief of alternative **conceptual models**
- inability to **measure** all the variables of a complex system, or to provide unique and accurate values of such variables (sampling/measurements errors)
- Different **degrees of accuracy** of numerical models used to reproduce the phenomenon under investigation

From a practical viewpoint, it is **rare to encounter only one type of uncertainty**:

- pure aleatoric uncertainty would mean that *all relations and their parameters* which describe the random process are *exactly known*
- pure epistemic uncertainty would mean that a *deterministic process is considered* but the *relevant information cannot be obtained*, e.g. due to the inability to measure the relevant parameters.

In *volcanological contexts*, epistemic and aleatoric uncertainties are **sometimes indistinguishable** and often treated together

Some examples of uncertainties that need to be quantified or evaluated in volcanology:

- *Positional uncertainty* of present/past volcanic features
- Uncertainty in the *range distribution of some key eruptive parameters* (e.g. mass flow rate, volume of erupted materials, grain sizes,…) when they need to be used for numerical modelling and hazard assessment
- Uncertainty in the *range distribution of parameters that influences volcanic hazard maps*
- Uncertainty in the *reliability of different numerical models* when reproduced a volcanic phenomenon.

HOW TO TREAT UNCERTAINTY

In general, when forward models or statistical procedures are not available, the quantification of epistemic/aleatoric uncertainty must be based directly on the **expert opinion**: in such cases, **structured expert judgement** tehniques are the main source for the UQ.

As a consequence of this approach, all the probability estimates will have their own confidence intervals.

Even probability maps produced with such an approach will be affected by uncertainty: typically, the **mean, 5th and 95th percentiles** for the probability density functions values.

Forecasting models which are based on uncertain modelling choices and data are also called **doubly stochastic**.

Doubly stochastic models construction always includes **two steps**:

- the distribution of the output of interest is represented using one or more parameters;
- at a second stage, some of these parameters are treated as being themselves random variables.

Other ways to treat uncertainty in volcanic context (better explained in following slides), include: drawing of uncertainty areas, comparison among the calculation of eruptive parameters with different methods (and therefore identification of a range variability), validation of numerical models.

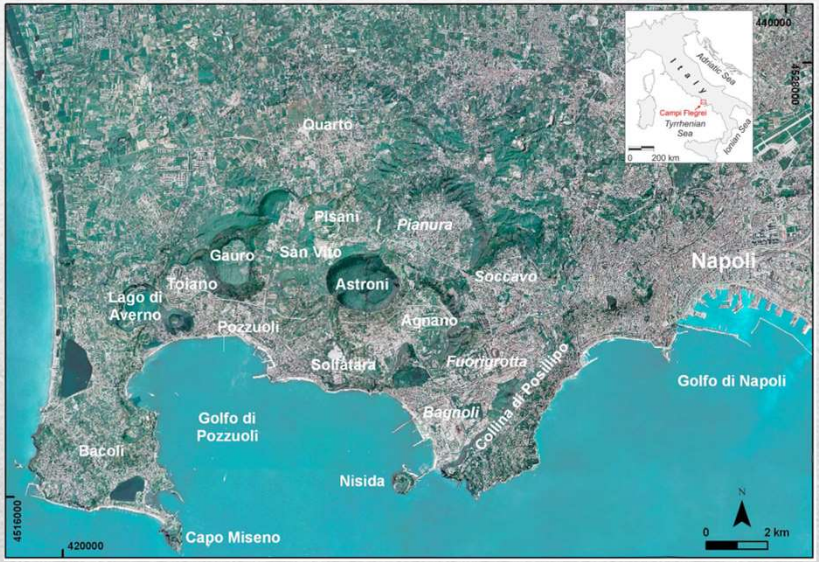
In the SV case study several uncertainties (mostly epistemic but also aleatoric) have been treated, analyzed and quantified (see also related poser on uncertainty treatment after DEM reconstruction).

Similar approach of UQ have been applied, for instance, in a study at Campi Flegrei focused on the development of vent opening probability maps (Bevilacqua et al. (2015), *Quantifying volcanic hazard at Campi Flegrei caldera (Italy) with uncertainty assessment: 1. Vent opening maps*) and Pyroclastic Density Currents probability invasion maps (Bevilacqua et al. (2015), *Quantifying volcanic hazard at Campi Flegrei caldera (Italy) with uncertainty assessment: 1. Pyroclastic Density Currents invasion maps*)

SOMMA-VESUVIO VOLCANIC COMPLEX

Poster 30-7: Angioletti A., Tadini A. & Bisson M., *Morphological evolution of Somma-Vesuvius caldera during the last century: integration between historical maps and airborne LiDAR survey*

CAMPI FLEGREI CALDERA



Poster 30-8: Bevilacqua A., Neri A., Bisson M., Esposti Ongaro T. , Flandoli F., Isaia R., Rosi M. & Vitale S., *Conditional effects of vent location, event scale and time forecasts on pyroclastic density currents hazard maps at Campi Flegrei caldera (Italy)*

Examples from Somma-Vesuvio volcano

Positional uncertainty in volcanic features/1

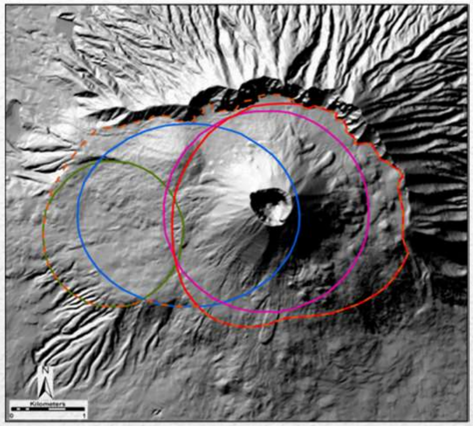
Positional uncertainty in volcanic features (especially vents and fissures) derive from

- incomplete knowledge and possible errors in acquistion (ambiguity of field data and paucity/lack of elements that help to recontruct the position of volcanological/structural features - epistemic uncertainty)
- intrinsic complexity during some eruptions (i.e. during caldera collapse after Plinian eruptions – aleatory uncertainty).

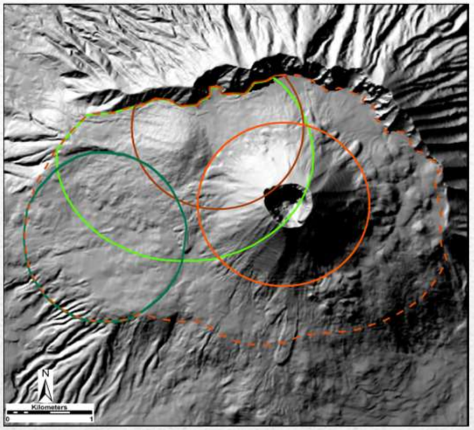
Quantification of uncertainty for the SV case:

- With respect to the definition of uncertatinty areas (aleatory/epistemic uncertainty). Areas have been defined according to:
  - **Morphological constraints**
  - **Temporal constraints/Field evidences**
  - **Resolution limits of interpolation profiles**
- With respect to the estimation of the amount of lost vents (i.e. cited in historical accounts but lost in the stratigraphic sequence)
- **8 datasets/variables**
  - **Distribution of PLINIAN/SUBPLINIAN I-II eruptions**
  - **Distribution of VIOLENT STROMBOLIAN (VS) to CONTINUOUS ASH EMISSION (AE) eruptions;**
  - **Distribution of EFFUSIVE eruptions (AD 1631>Age>AD 1944)**
  - **Distribution of DEEP FAULTS (Quaternary? – 3 faults crossing SV caldera);**
- **1 Homogeneous** distribution **(**Accounts for neglected factors and missing information**)**

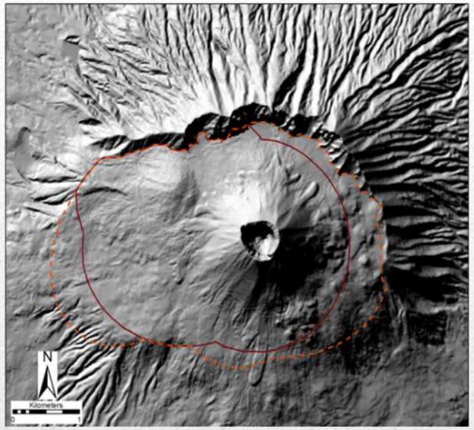
Positional uncertainty in volcanic features/2



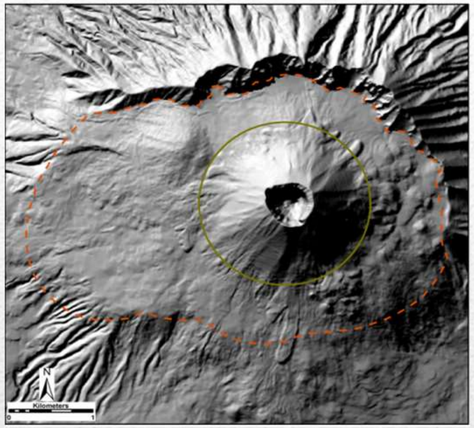
PLINIAN
22 ka BP<Age<AD 79



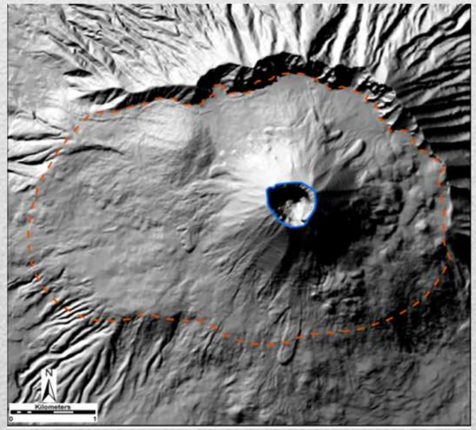
SUBPLINIAN
19 ka BP<Age<AD 1631



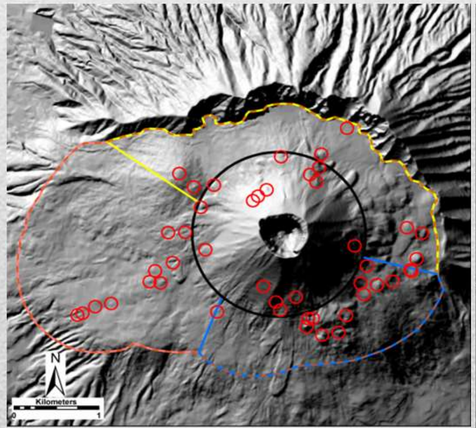
VS to AE – PREAVELLINO
22 ka BP<Age<4.3 ky BP



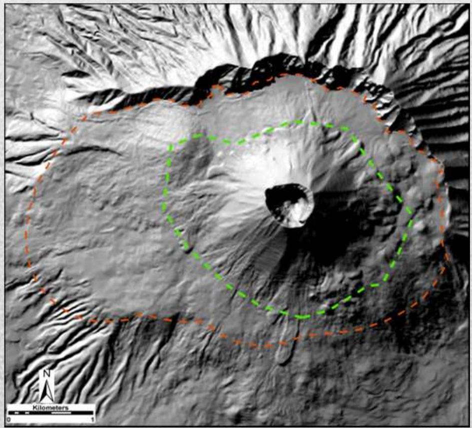
VS to AE – GRAN CONO
4.3 ka BP<Age<AD 1631



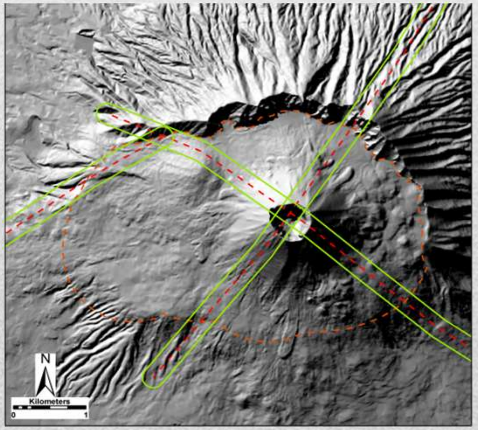
VS to AE – 1944 CRATER
AD 1631<Age<AD 1944



PARASITIC VENTS
AD 1631<Age<AD 1944



ERUPTIVE FISSURES
AD 1631<Age<AD 1944



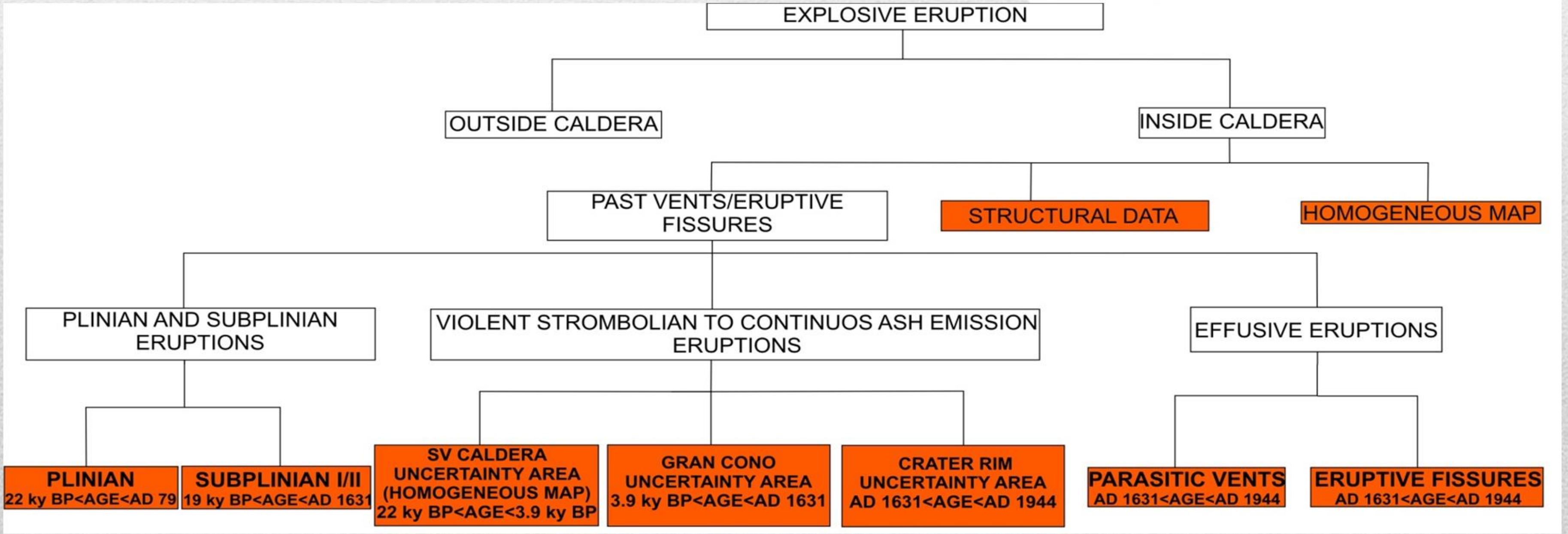
DEEP FAULTS
(Quaternary?)

Quantifiying uncertainty through expert elicitation/1

- Somma-Vesuvio case: final goal – production of a vent opening (VO) probability map

- 17 Experts were invited to provide their judgements on 16 seed items for calibration scoring purposes.

- Each expert responded then to 15 target items for providing weights that needs to be attributed the different *VO* maps related to previous variables according to the following logic tree.



EXPLOSIVE ERUPTION

OUTSIDE CALDERA

INSIDE CALDERA

PAST VENTS/ERUPTIVE FISSURES

STRUCTURAL DATA

HOMOGENEOUS MAP

PLINIAN AND SUBPLINIAN ERUPTIONS

VIOLENT STROMBOLIAN TO CONTINUOS ASH EMISSION ERUPTIONS

EFFUSIVE ERUPTIONS

PLINIAN
22 ky BP<AGE<AD 79

SUBPLINIAN I/II
19 ky BP<AGE<AD 1631

SV CALDERA UNCERTAINTY AREA (HOMOGENEOUS MAP)
22 ky BP<AGE<3.9 ky BP

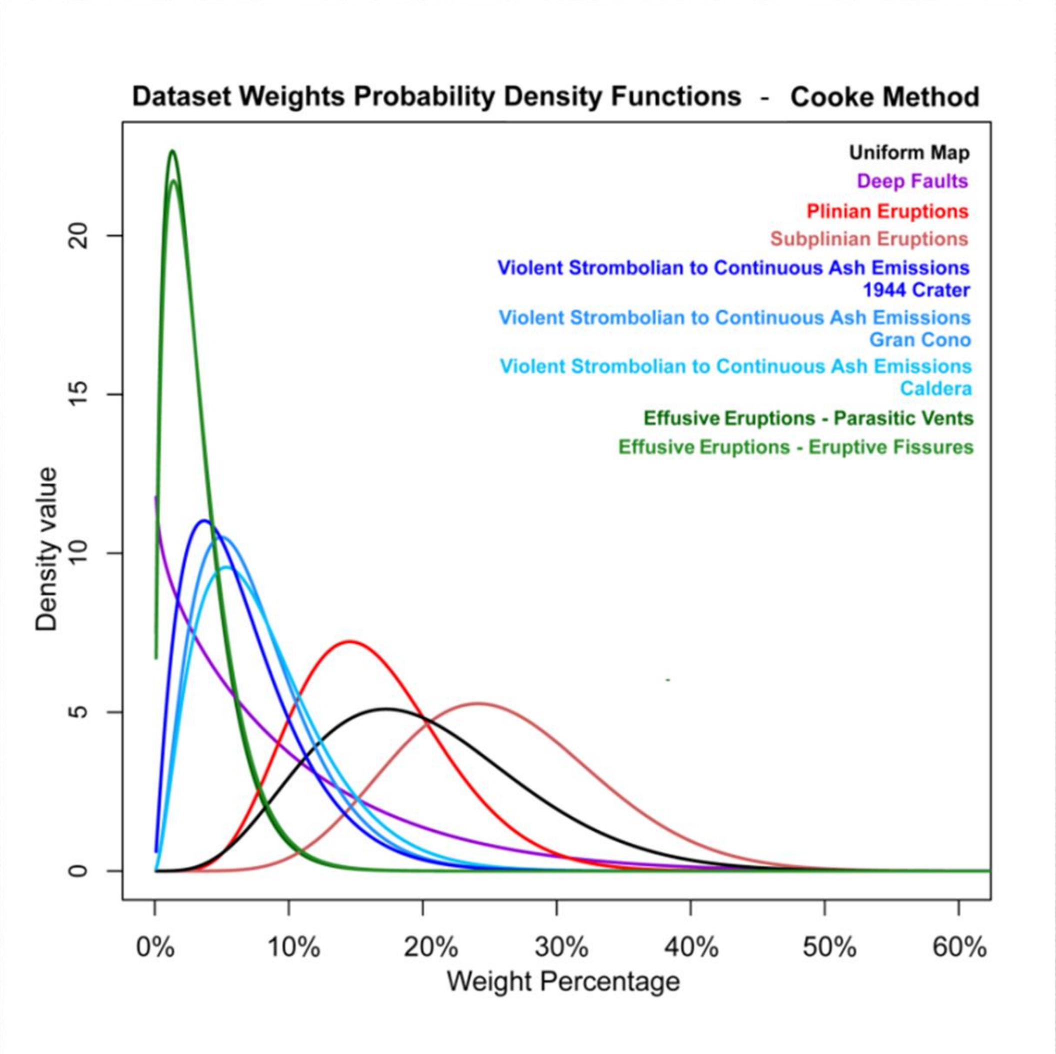
GRAN CONO UNCERTAINTY AREA
3.9 ky BP<AGE<AD 1631

CRATER RIM UNCERTAINTY AREA
AD 1631<AGE<AD 1944

PARASITIC VENTS
AD 1631<AGE<AD 1944

ERUPTIVE FISSURES
AD 1631<AGE<AD 1944

Quantifiying uncertainty through expert elicitation/2



- Three different judgement scoring procedures have been applied to the elicitation data: Classical Method (CM), Expected Relative Frequency method (ERF), and Equal Weights rule (EW).

- Sensitivity analyses were performed by removing controversial seed questions and by considering sub-groups of experts (Juniors/Seniors, Geologists/Modelers)

- Agreement between weights/maps among different scoring procedures and different sub-groups (weights differ maximum of 3%) – *CM maps with all the experts considered as a reference*

Explicit quantification of uncertainty in probabilistic maps/1

Vent opening maps display the estimated probability of vent opening at each point within the region of interest.

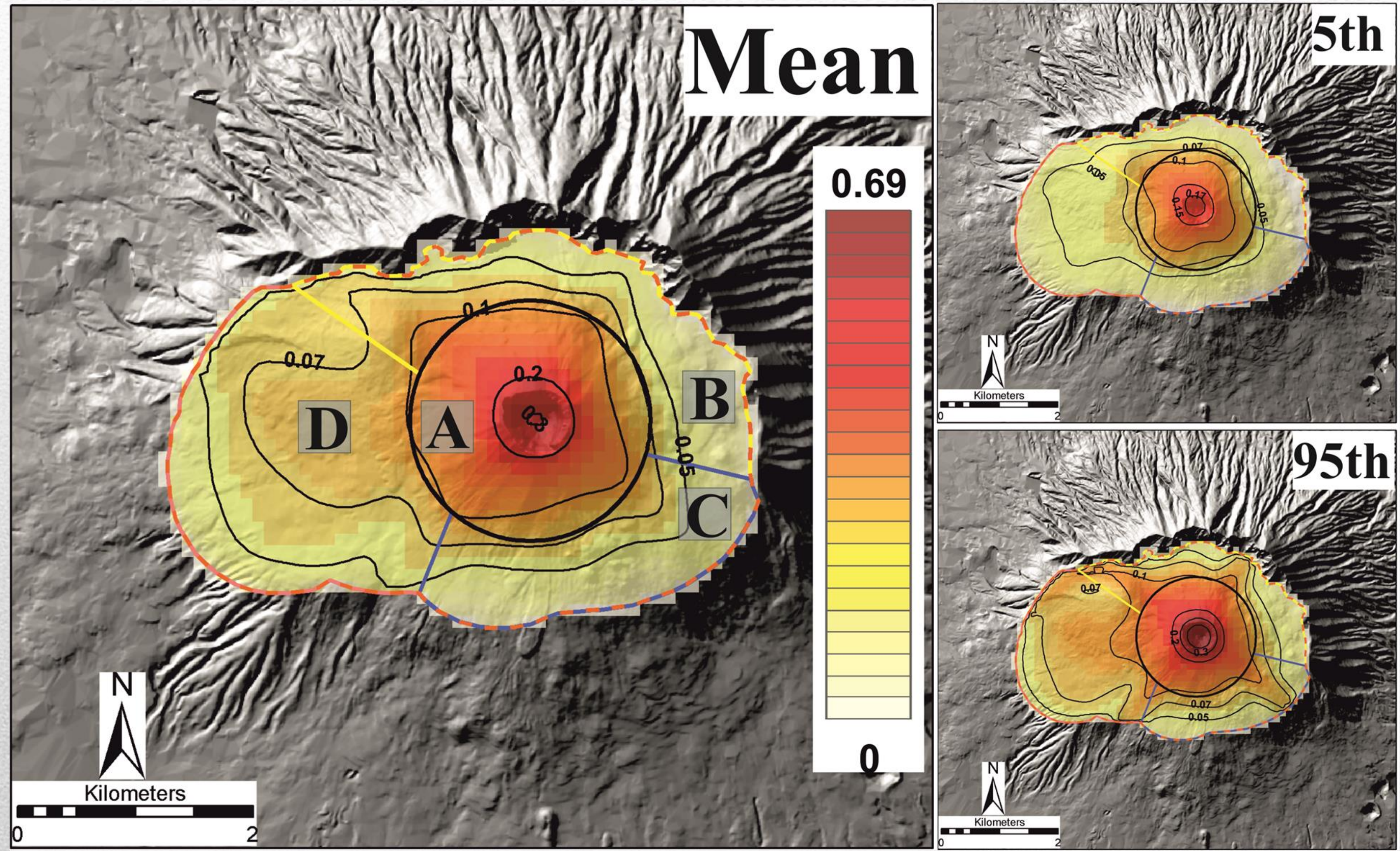
The example from the Somma-Vesuvio volcano derives from Tadini et al. [2017], *Assessing future vent opening locations at the Somma-Vesuvio volcanic complex: 2. Probability maps of the caldera for a future Plinian/sub-Plinian event with uncertainty quantification*

The maps rely on the **sites of past events** and on the **tectonic information** that may modulate dike rising (see in previous slides). Probability density functions with gaussian kernels have been linearly combined with weights defined through expert elicitation.

The probability assessments described in this study are conditional to the occurrence of an eruption and do not include any temporal assessment, unless specified.

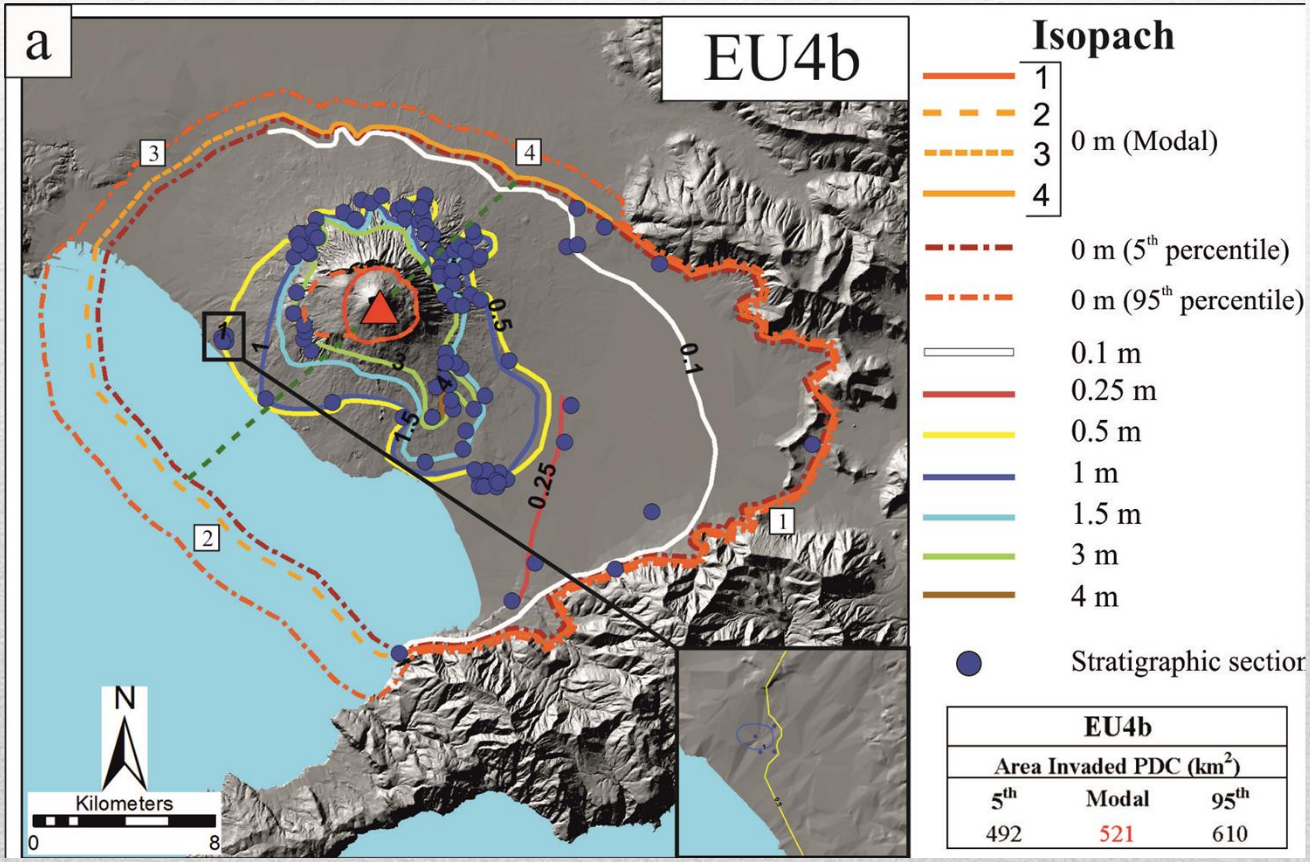
Epistemic and Aleatory uncertainties are here explicitly quantified and shown through the production of a set of three maps, representing a mean value and two upper (95° percentile) and lower (5° percentile) uncertainty bounds.

Explicit quantification of uncertainty in probabilistic maps/2



Uncertainty in measuring eruptive parameters/1

An important source of uncertainty in volcanology (but almost everywhere in the geosciences) is related to the quantification of eruptive parameters. Uncertainty in this case is both **epistemic** (data available for calculations are ether few, imprecise or heterogeneously disposed; methods for calculating the parameters have different assumptions) but also **aleatoric** (several parameters, if used as input data for numerical simulations, are influenced by the physical variability of the volcano).

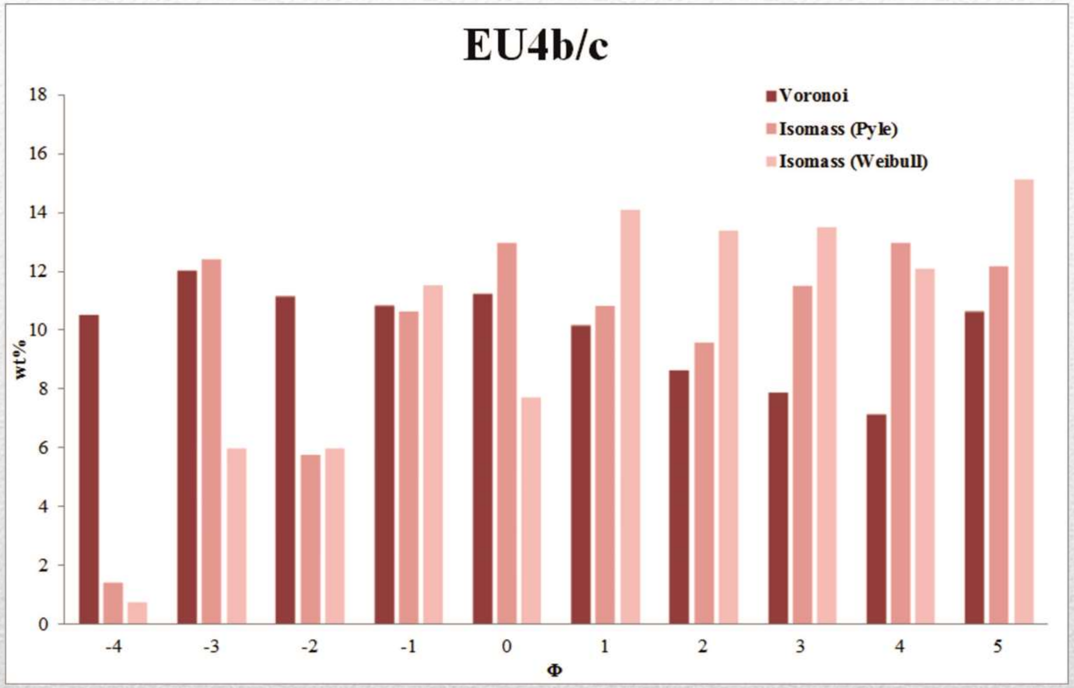


EU4b		
Area Invaded PDC (km²)		
5th	Modal	95th
492	521	610

Maximum Runout estimations performed by reconstructing the ideal PDC 0 m isopach (beyond which all the deposits do not show features typical of lateral transport). Three different outlines of PDC maximum runouts are presented for each unit (Modal, 5th percentile and 95th percentile - represent **epistemic uncertainty**). Modal maximum runout outline is composed by different segments which can be traced with different degrees of confidence: for each segment different percentiles are evaluated.

Uncertainty in measuring eruptive parameters/2

Total Grain-Size Distribution (TGSD)



- Voronoi tessellation works better if stratigraphic sections are more uniformly distributed
- Isomass method introduces a higher degree of subjectivity (ability and experience of the operator to draw isomass lines)

Volume

Method	**Percentile**	**Volume (km^3)**
TIN	5th	**0.292**
	Modal	**0.295**
	95th	**0.313**
TRAPEZOID	Modal	**0.364**
VORONOI	Modal	**0.560**
SECTORS (Slope Classes)	Modal	**0.507**

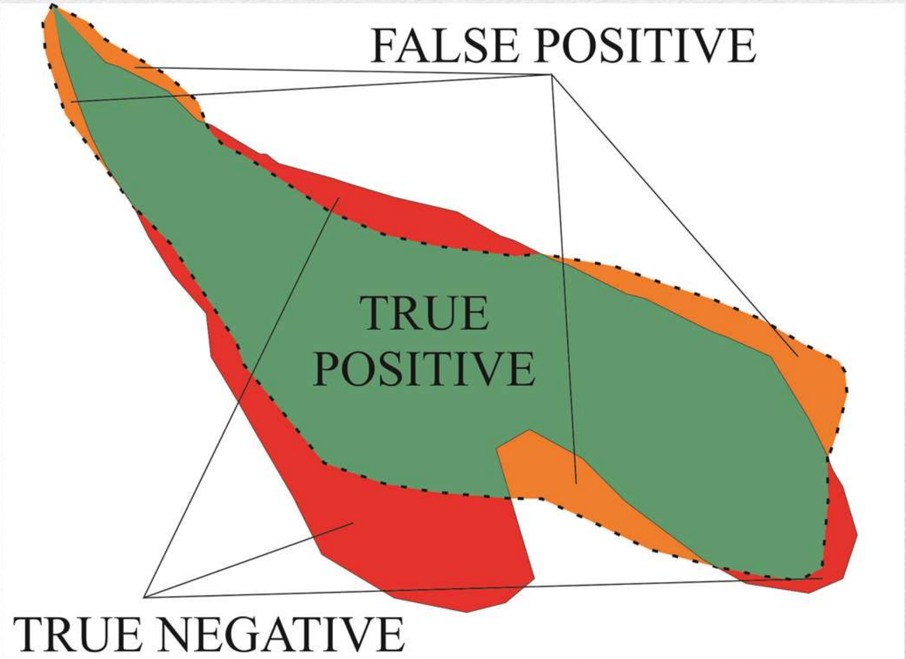
- Epistemic uncertainty derived because of different methods assumptions
- If such parameters are used as input data – range values derived from different methods might express the aleatoric uncertainty

Uncertainty after numerical modelling/1

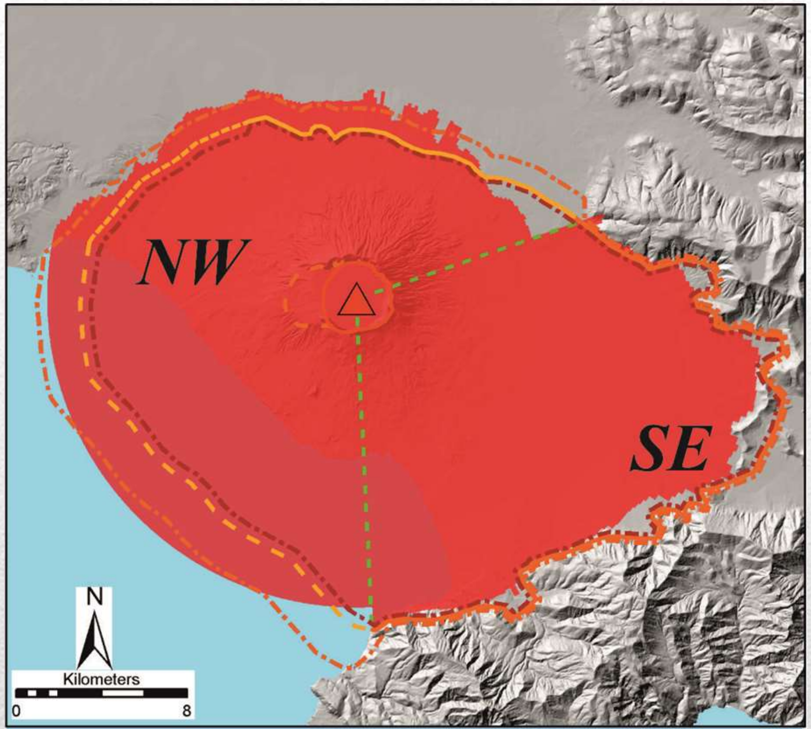
- Uncertainty in volcanic hazard assessment might arise when a numerical model is used to produce hazard maps
- This is due to the fact that
    - A single model is not capable of reproducing perfectly a volcanic phenomena due to approximation and/or simplifications
    - Different codes with different assumptions and/or equations produces different outputs
- Validation procedures might therefore quantify the degree of uncertainty associated with numerical models

Possible validation procedure:

- With respect to inundation areas of model/deposit (True Positive, True Negative, False Positive)
- With respect to thicknesses model/deposit with distance from vent area
- With respect to mass fractions of grain sizes of model/deposit with distance from vent area



Uncertainty after numerical modelling/2



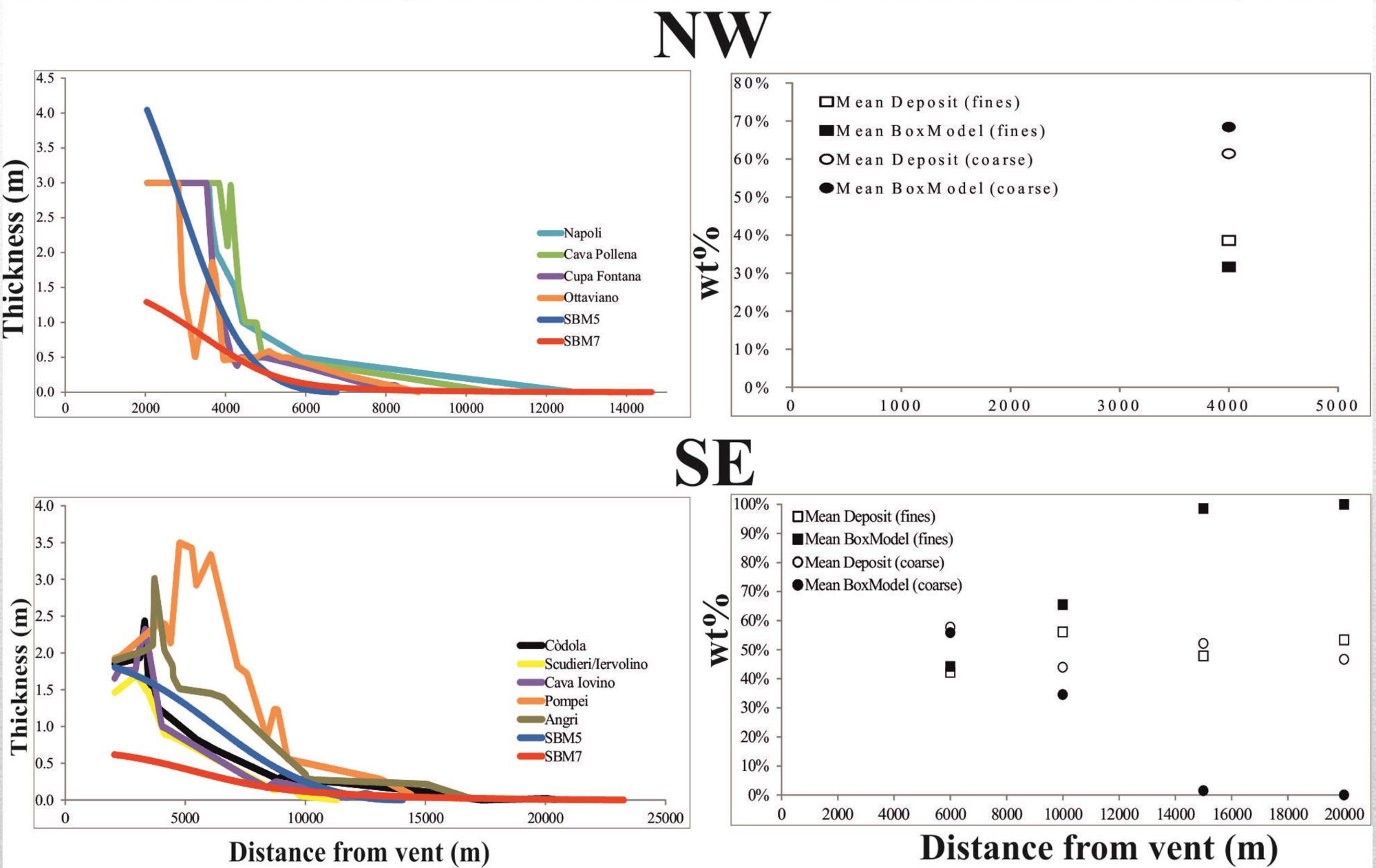
Percentile	True Positive	False Positive	True Negative
Modal	83.60%	12.19%	4.21%

BOX-MODEL

- Advanced version of the classical kinematic approach (e.g. Energy cone)
- The conservation of mass is obtained through equal area geometrical elements that stretches out through time
- Comparison between the topography and the decay of kinetic energy with distance («Energy conoid»)
- Capable of reproducing PDCs with volume fraction of solid particles from 0.5% to 5-6%

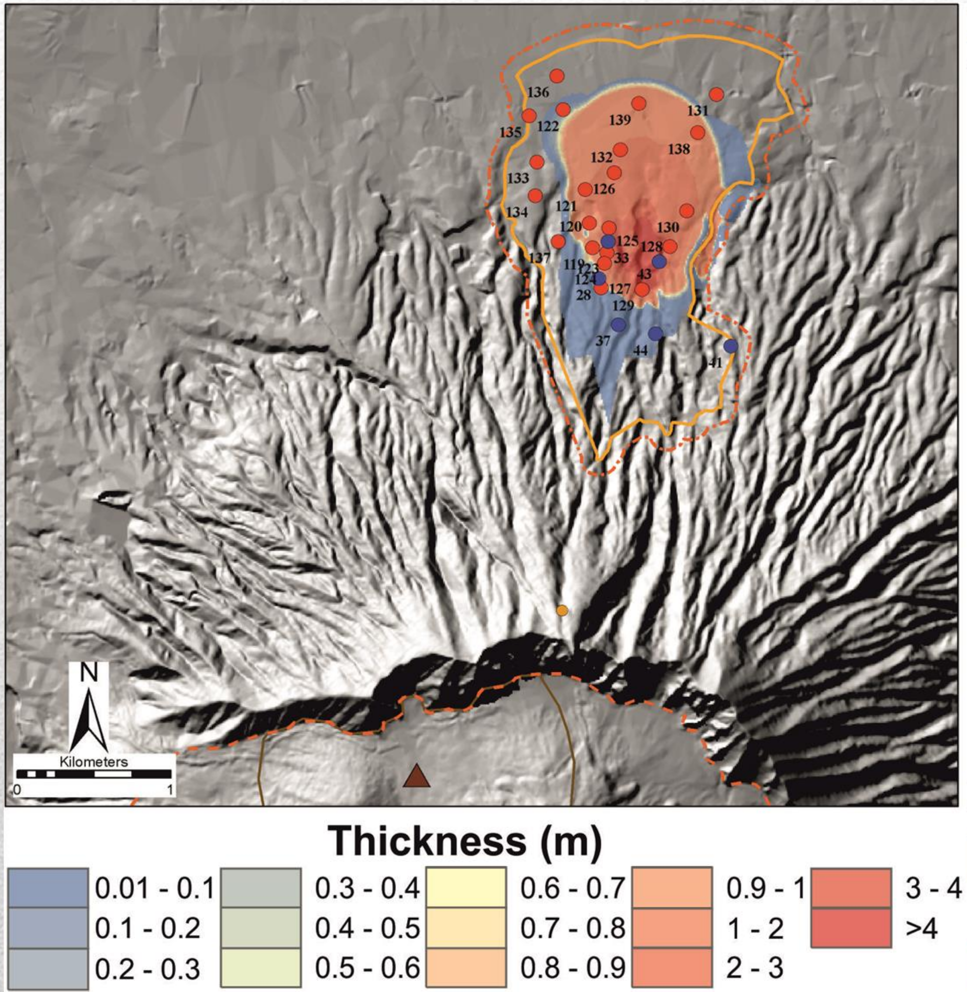
- Directional collapses to better capture the true deposit

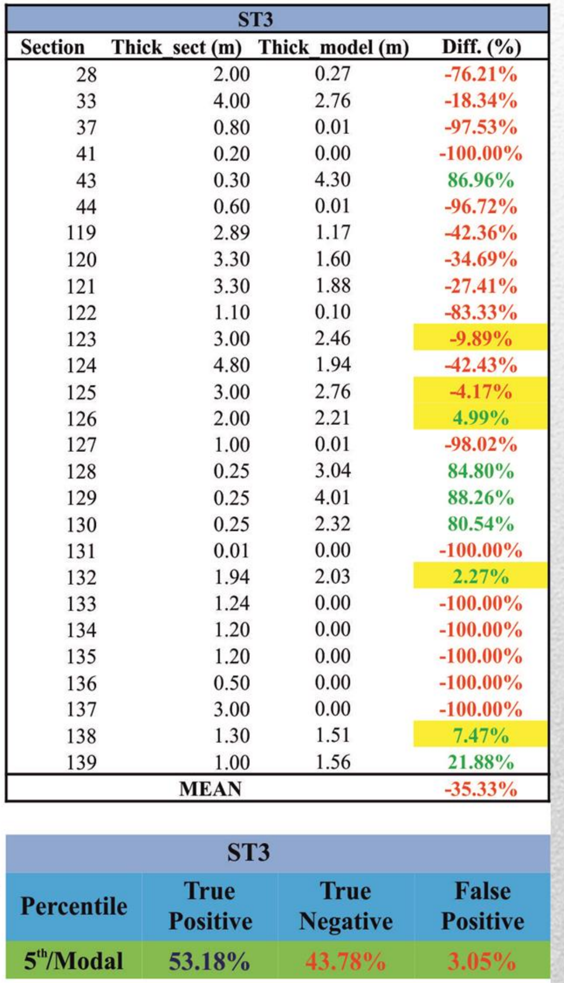
Uncertainty after numerical modelling/3



Uncertainty after numerical modelling/4

TITAN2D (Depth-averaged approach with shallow water-derived governing equations)





ST3			
Section	Thick_sect (m)	Thick_model (m)	Diff. (%)
28	2.00	0.27	-76.21%
33	4.00	2.76	-18.34%
37	0.80	0.01	-97.53%
41	0.20	0.00	-100.00%
43	0.30	4.30	86.96%
44	0.60	0.01	-96.72%
119	2.89	1.17	-42.36%
120	3.30	1.60	-34.69%
121	3.30	1.88	-27.41%
122	1.10	0.10	-83.33%
123	3.00	2.46	-9.89%
124	4.80	1.94	-42.43%
125	3.00	2.76	-4.17%
126	2.00	2.21	4.99%
127	1.00	0.01	-98.02%
128	0.25	3.04	84.80%
129	0.25	4.01	88.26%
130	0.25	2.32	80.54%
131	0.01	0.00	-100.00%
132	1.94	2.03	2.27%
133	1.24	0.00	-100.00%
134	1.20	0.00	-100.00%
135	1.20	0.00	-100.00%
136	0.50	0.00	-100.00%
137	3.00	0.00	-100.00%
138	1.30	1.51	7.47%
139	1.00	1.56	21.88%
MEAN			-35.33%

ST3			
Percentile	True Positive	True Negative	False Positive
5th/Modal	53.18%	43.78%	3.05%

Conclusions

- Uncertainty in geosciences (and in volcanology) needs to be quantified in order to provide a quantification of the reliability/effectiveness of thematic products related to hazard/risk assessment

- Classical subdivision of uncertainty in epistemic/aleatoric (often indistinguishible and treated together)
  - Epistemic – related to incomplete/erroneous knowledge (might be reduced)
  - Aleatoric – related to intrinsic randomnes of the system (might not be reduced but just quantified)

- Ways to treat and quantify uncertainty in some volcanic issues
  - Uncertainty areas
  - Expert elicitation techniques
  - Doubly stochastic approaches
  - Comparison of parameter estimations made with different methods (and definition of a range of values)
  - Validation of numerical models

Grazie per l'attenzione!

Questions?