Ensemble atmospheric dispersion modelling of volcanic species: interpreting ensemble data and applications

Cities on Volcanoes 12, Antigua Guatemala

L. Mingari [1] A. Folch [1] E. Vazquez [2] M. S. Osores [2] A. Costa [3]

[1]Geosciences Barcelona (GEO3BCN-CSIC), Barcelona, Spain

[2]National Weather Service, Buenos Aires, Argentina

[3]Istituto Nazionale di Geofisica e Vulcanologia, Sezione di Bologna, Bologna, Italy

February 13, 2024

Introduction and motivation

- Atmospheric dispersion models can provide realistic distributions of airborne volcanic ash and gases or tephra deposits
- Traditionally, operational forecast systems rely on volcanic ash transport and dispersal (VATD) models to produce deterministic forecasts
- Ensemble modelling poses new challenges: we explore approaches for dealing with large volumes of ensemble data, data interpretation and applications

Why ensemble modelling?

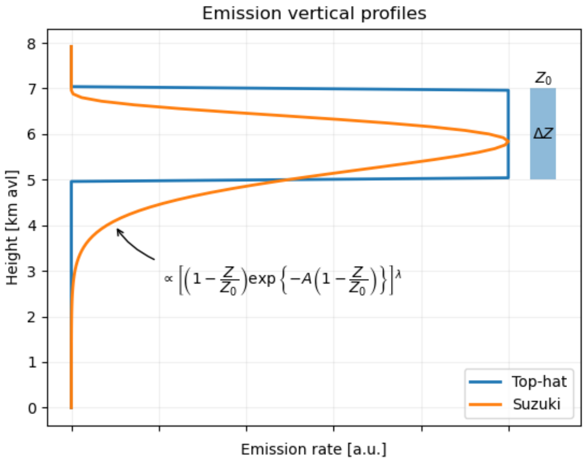
- **Uncertainty in model input parameters:** Deterministic models are highly sensitive to uncertain model input parameters (e.g. eruption source parameters) and meteorological fields. We can take into account these uncertainties using ensemble modelling
- **Quantification of model output uncertainty:** Ensemble-based modelling allows one to characterise and quantify model output uncertainties. In addition to traditional forecasting products, the associated errors can be provided
- **Improvement of forecast skill:** Real observations can be incorporated into dispersal models using ensemble-based data assimilation techniques
- **Source inversion:** Different techniques for source term inversion have been proposed based on ensemble modelling

Ensemble simulations

Numerical model:
- FALL3D: model for atmospheric transport and deposition of particles and aerosols

Ensemble construction:
- Emission source parameters
- Grain size distribution
- Aggregation parameters
- Meteorological fields
- Diffusivity coefficients

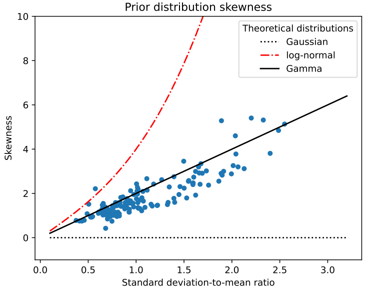
Outputs:
- Ash/gases concentration
- Column mass loading
- Top cloud height
- Deposit mass loading

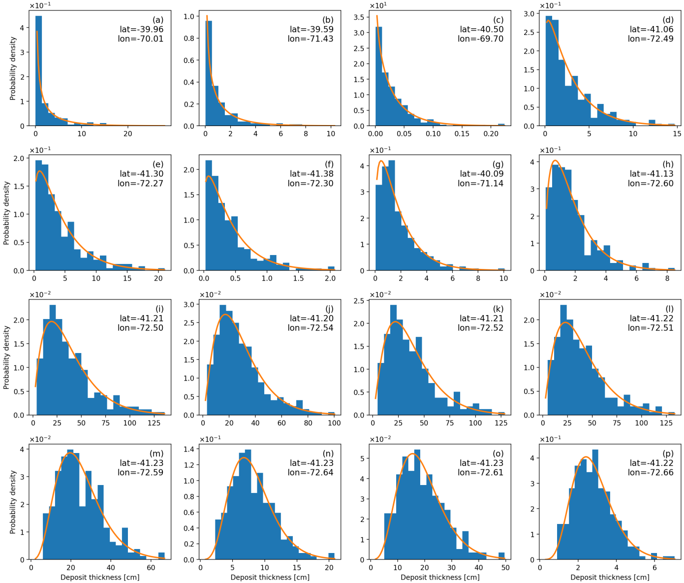


Ensemble simulations: Probability distribution

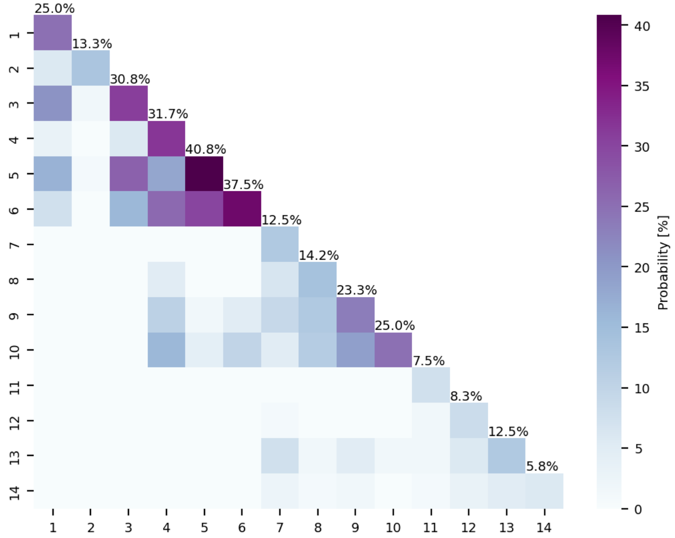
The gamma distribution provides a good approximation to the ensemble distribution:

- $\overline{y} < \sigma_y \rightarrow$ mode$=0$
- $\overline{y} > \sigma_y \rightarrow$ mode>0
- Right-skewed probability distributions





Ensemble data complexity reduction



Run configuration:

- Test case: **2015 Etna eruption**
- Ensemble size = 120 members

Traditional products:

- Ensemble mean
- Ensemble spread
- Exceedance probability

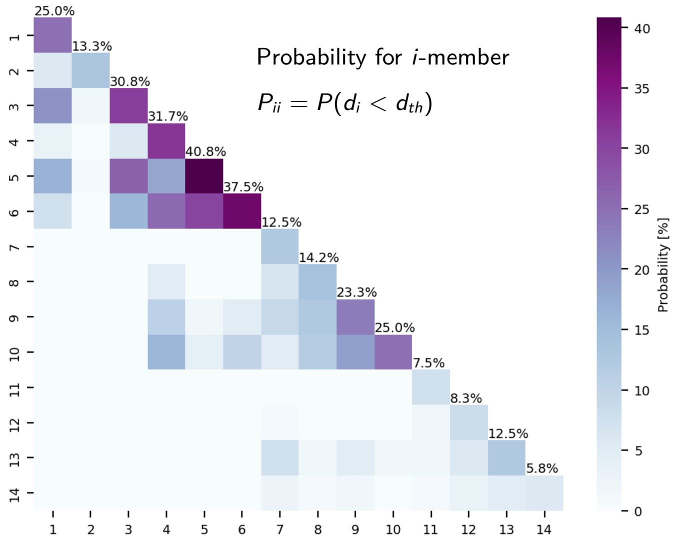
Complexity reduction:

- Remove redundancy in the ensemble data
- We need to measure the distance d_{ij} between two model states i and j
- Ensemble members with similar distances are grouped

Reduced ensemble

Ensemble size reduction:
$120 \to 14$

Ensemble data complexity reduction



Run configuration:

- Test case: **2015 Etna eruption**
- Ensemble size = 120 members

Traditional products:

- Ensemble mean
- Ensemble spread
- Exceedance probability

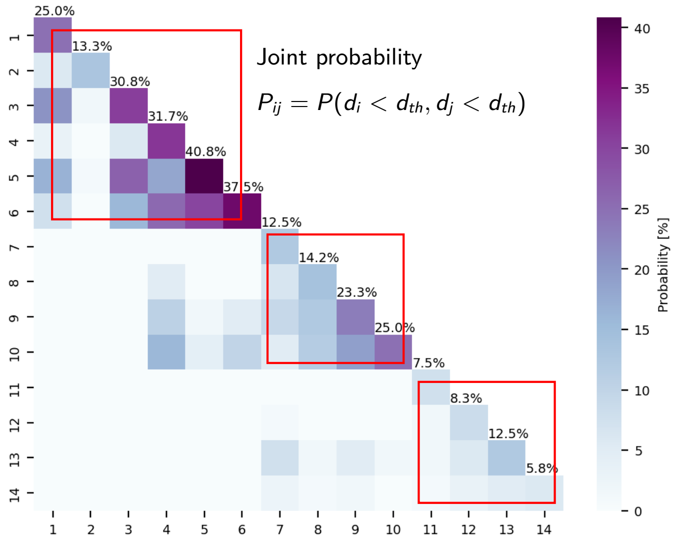
Complexity reduction:

- Remove redundancy in the ensemble data
- We need to measure the distance d_{ij} between two model states i and j
- Ensemble members with similar distances are grouped

Reduced ensemble

Ensemble size reduction:
$120 \to 14$

Ensemble data complexity reduction



Run configuration:

- Test case: **2015 Etna eruption**
- Ensemble size = 120 members

Traditional products:

- Ensemble mean
- Ensemble spread
- Exceedance probability

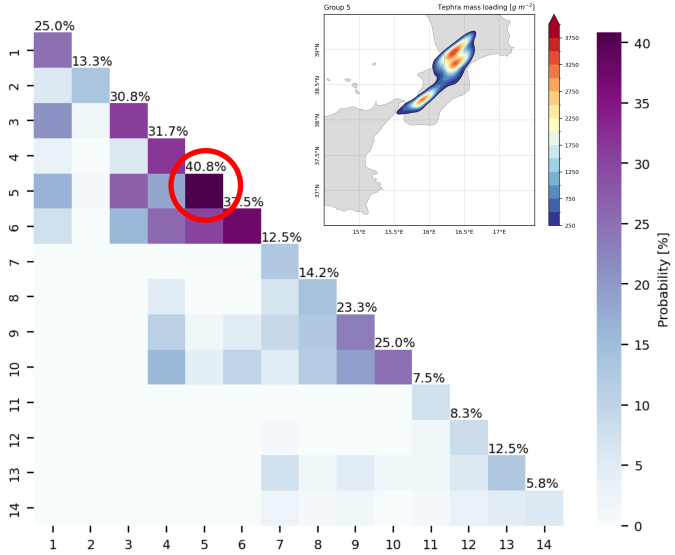
Complexity reduction:

- Remove redundancy in the ensemble data
- We need to measure the distance d_{ij} between two model states i and j
- Ensemble members with similar distances are grouped

Reduced ensemble

Ensemble size reduction:
120 → 14

Ensemble data complexity reduction



Run configuration:

- Test case: **2015 Etna eruption**
- Ensemble size = 120 members

Traditional products:

- Ensemble mean
- Ensemble spread
- Exceedance probability

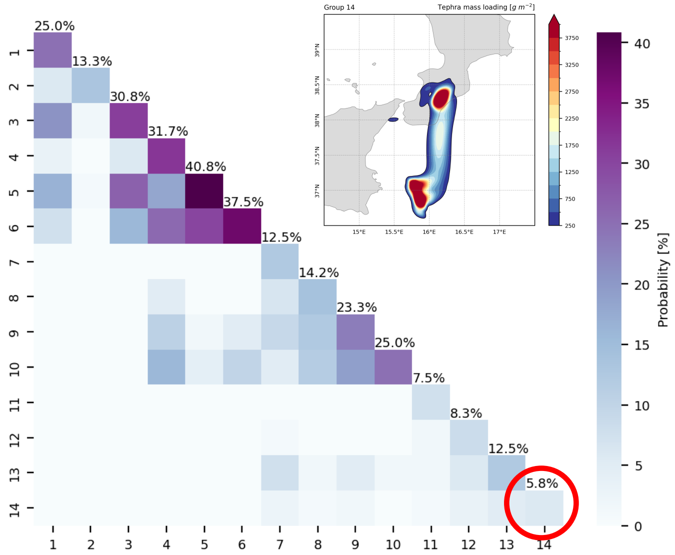
Complexity reduction:

- Remove redundancy in the ensemble data
- We need to measure the distance d_{ij} between two model states i and j
- Ensemble members with similar distances are grouped

Reduced ensemble

Ensemble size reduction:
$120 \to 14$

Ensemble data complexity reduction



Run configuration:

- Test case: **2015 Etna eruption**
- Ensemble size = 120 members

Traditional products:

- Ensemble mean
- Ensemble spread
- Exceedance probability

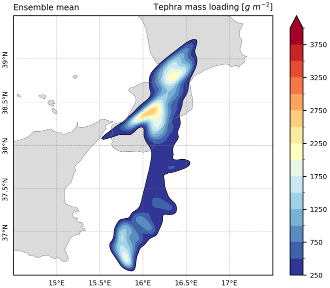
Complexity reduction:

- Remove redundancy in the ensemble data
- We need to measure the distance d_{ij} between two model states i and j
- Ensemble members with similar distances are grouped

Reduced ensemble

Ensemble size reduction:
$120 \rightarrow 14$

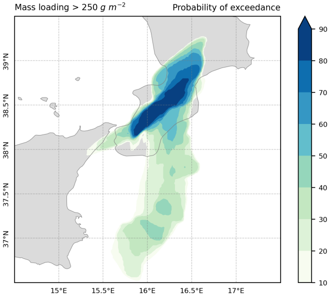
Comparison of model products

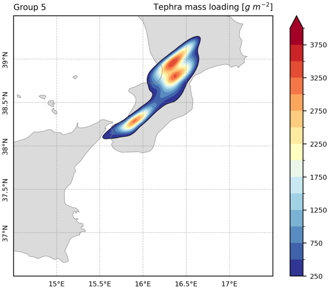


Traditional products

← Ensemble mean

Probability of exceedance →

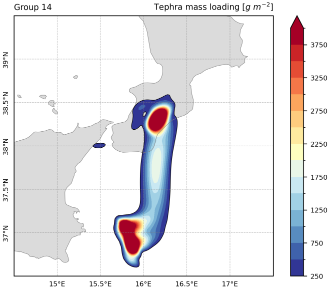




Reduced ensemble

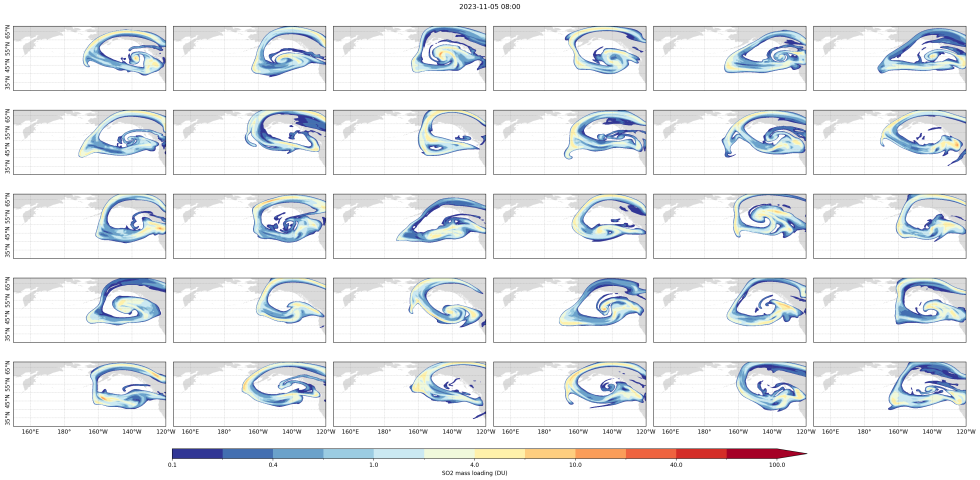
← High probability state (41%)

Low probability state (6%) →

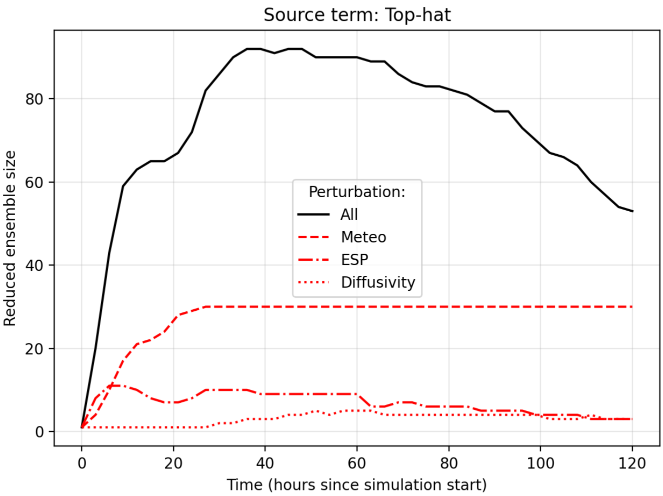


Sensitivity study

- Test case: SO2 cloud of the **2023 Klyuchevskoy eruption** (Kamchatka Peninsula)
- Multiple **ensemble simulations** were performed perturbing single model parameters
- **Size of the reduced ensemble** as a measure of the variable sensitivity
- Ensemble size = 120 members



Sensitivity study



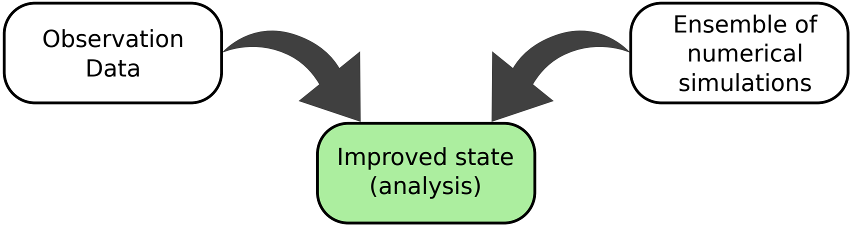
- Minimum ensemble size required
- Sensitivity

	Short-term forecasting	Long-term forecasting
Meteo	Medium	High
Eruption Source Parameters	Medium	Medium
Diffusivity	Low	Low

Ensemble-based data assimilation

GNC method:

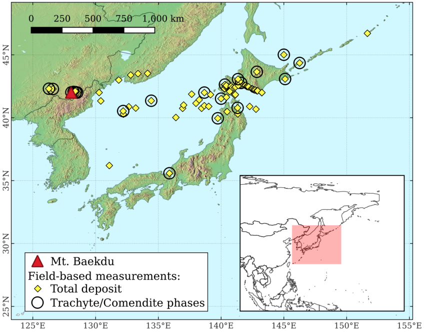
- We propose a heuristic first-order data assimilation method for lower-bounded (positive) variables
- The Gaussian with non-negative constraints (GNC) method assumes a multi-dimensional Gaussian probability distribution and a linear observation operator
- A non-negative quadratic programming problem is solved using an iterative approach
- This method leads to better results than other DA techniques, including the classical Ensemble Kalman Filter (EnKF)

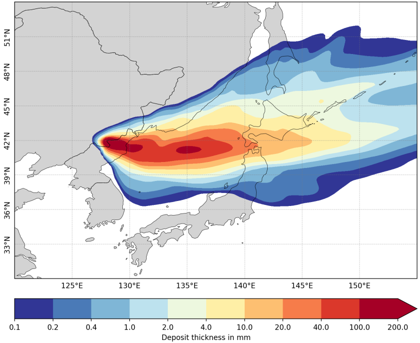


Ensemble-based data assimilation

GNC method:

- We propose a heuristic first-order data assimilation method for lower-bounded (positive) variables
- The Gaussian with non-negative constraints (GNC) method assumes a multi-dimensional Gaussian probability distribution and a linear observation operator
- A non-negative quadratic programming problem is solved using an iterative approach
- This method leads to better results than other DA techniques, including the classical Ensemble Kalman Filter (EnKF)



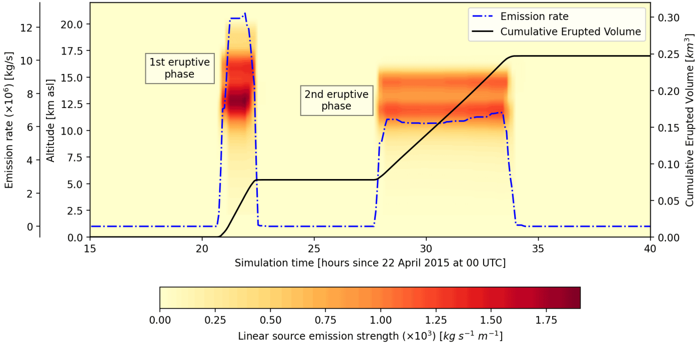


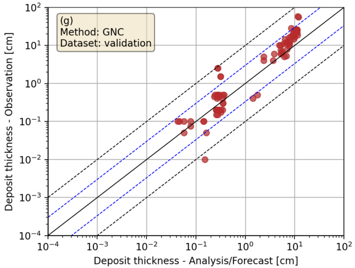
Millennium eruption

Reconstruction of the tephra fallout deposit of the 946 CE Millennium eruption of Changbaishan volcano assimilating deposit thickness data

Source term inversion

- A technique for emission source inversion based on the GNC method can be used to estimate the space–time distribution of the source
- Valid for problem with weak non-linearity effects





Calbuco eruption

Time evolution of emission rate profiles for the 2015 Calbuco eruption according to the source term inversion approach based on the GNC method

Conclusions

- When ensemble simulations are useful?
  - Whenever the ESP or meteorological conditions have large uncertainties
  - If forecast errors must be quantified
  - If observations are available to be assimilated
  - When the emission source needs to be characterised
- How should the ensemble be constructed? Which variables should be perturbed? What's the minimum ensemble size required?
  - We gave a preliminary insight into which parameters are the most sensitive
  - What is the minimum ensemble size corresponding to each case
- How interpret and deal with large volumes of ensemble data? Which probabilistic output products are relevant?
  - We proposed a complexity reduction technique in order to identify qualitatively different states in the ensemble
  - We obtained approximate solutions of the physical model, including the state that is most likely to be sampled, and assigned a probability to each state
  - These physically consistent states can be compared directly with satellite images
- Are the traditional assimilation methods suitable for the atmospheric dispersion of volcanic species?
  - Traditional data assimilation methods lead to suboptimal performance in the case VATD models
  - We propose a new ensemble-based data assimilation method which outperforms the classical EnKF method when is applied to VATD models