FLEXpart (flexible particle dispersion model) on EWC - FLEXWEB

Michael Blaschek[1], Marina Dütsch[1], Andreas Stohl[1]

1) Institut für Meteorologie und Geophysik, Universität Wien, Vienna, Austria

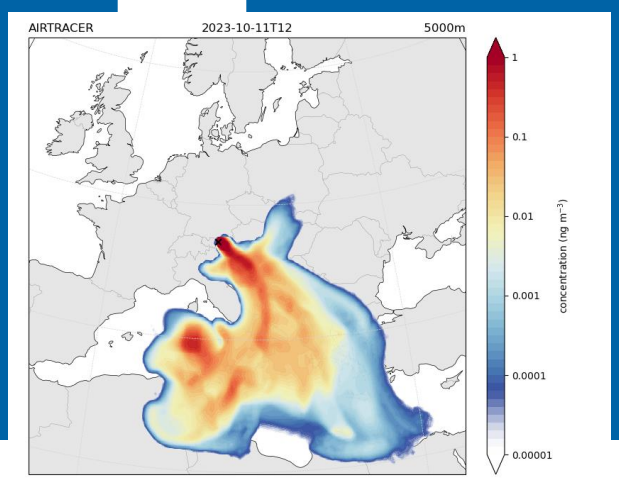
universität wien

Department of Meteorology and Geophysics

michael.blaschek@univie.ac.at

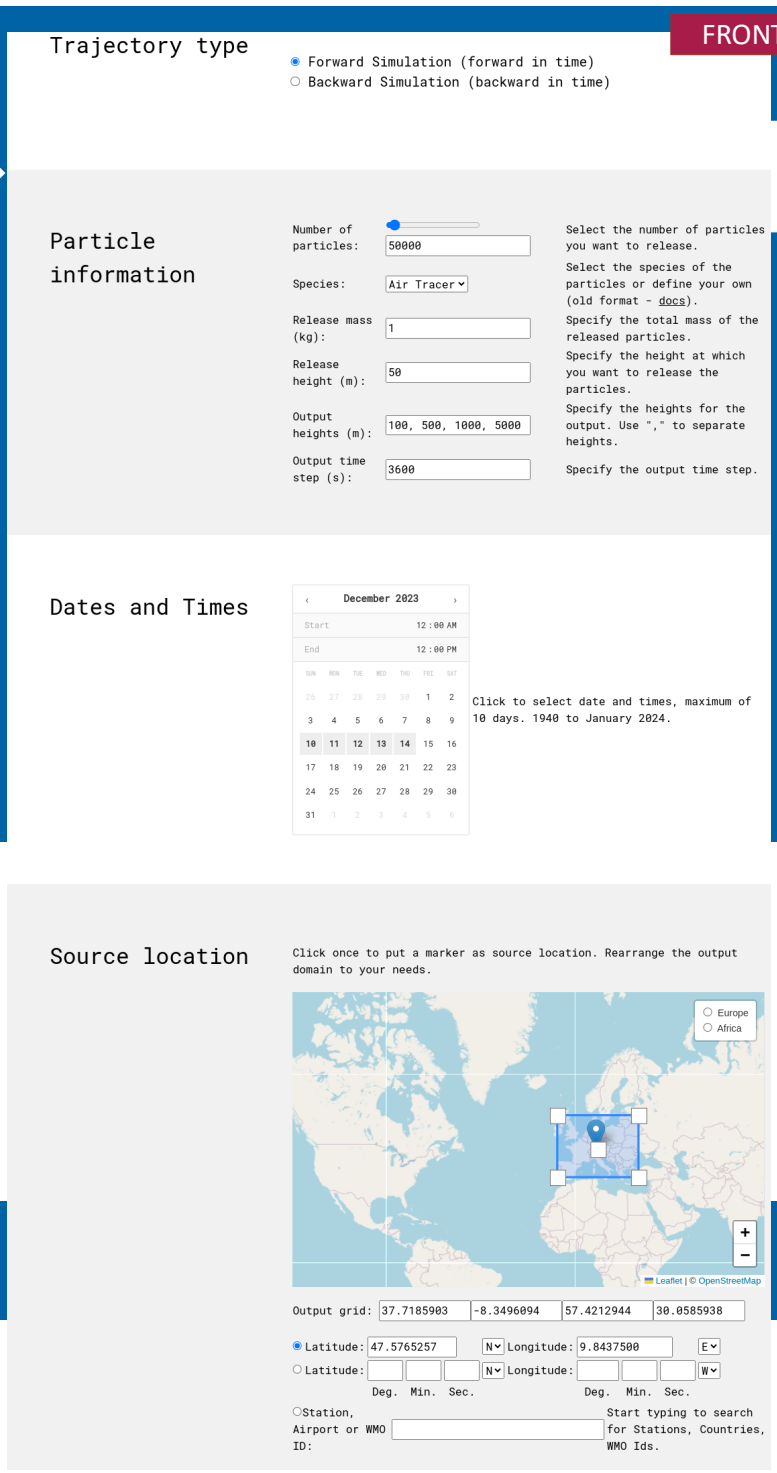
User

Design your simulation



SCAN ME

FRONTEND



Submit

DB

AGENT

- **Deploy on regular HPC**
- Use **SLURM**
- User credentials
- **Apptainer Flexpart Image**

KUBERNETES on EWC

- Rancher RKE2 cluster
- Scale according to DB length (jobs, <10)
- Retrieve Inputs from MARS (e.g. ERA5)
- Cache inputs and deliver only results
- Allow parallel retrieval and processing

TASK

Setup

Retrieve meteorological data

Simulation

Post process

1. Motivation

FLEXPART[1] is a numerical model that simulates the dispersion of gases and aerosols in the atmosphere. The goal is: ***allow a larger community (scientific, citizens, …) to focus on high-level analysis and interpretation of results, rather than managing the underlying computational challenges.***

5. Discussion / Take home messages

Large data sets are immobile by nature. **Computing closer to the data source is much more efficient.** Using agents or a cloud-like-service offered by HPC centres allows users to make their computational resources easier available to users, colleagues, citizens or the global South.

2. Setup

FLEXWEB is Python, HTML and JavaScript code (FastAPI). **FLEXPART simulations require large amounts of input data and are CPU intense calculations, hence are by its nature HPC applications.** The input data is a postprocessed **ERA5** dataset running from 1940 to 2024 (84y) of about **420 TB**, residing on a department storage system (JETFS). For example, one **10-day forward simulation of 500k air tracer particles fetches 156 GB** of input data and takes **45 minutes on 20 CPU** cores with at **least 12 GB of memory**.

3. Flexpart in the Cloud (e.g. Azure)

A kubernetes setup requires a database, frontend and worker nodes (8/16 CPU, 16/32 GB). **The main cost factor is long-term storage in the cloud** (e.g. 197€/m=10TB). **Since the data needs to reside on some remote storage, just computing in the cloud is inefficient.**

4. Flexpart on EWC

The limitation/costs of other clouds can be avoided by fast **access to MARS**. Since the **input data** can be **retrieved** as efficiently as possible, **processed** and **cached** for some reuse, the only real challenge remains to **make the processing, retrieval and delivering of results as efficient as possible.**

6. Reference(s)

[1] Bakels, L., Tatsii, D., Tipka, A., Thompson, R., Dütsch, M., Blaschek, M., Seibert, P., Baier, K., Bucci, S., Cassiani, M., Eckhardt, S., Groot Zwaaftink, C., Henne, S., Kaufmann, P., Lechner, V., Maurer, C., Mulder, M. D., Pisso, I., Plach, A., Subramanian, R., Vojta, M., and Stohl, A.: FLEXPART version 11: improved accuracy, efficiency, and flexibility, Geosci. Model Dev., 17, 7595–7627, https://doi.org/10.5194/gmd-17-7595-2024, 2024.