The AlpEnvironmental Data Analysis Center

alpendac.eu

DATA MANAGEMENT, OPERATING-ON-DEMAND AND BEYOND THE ALPINE ENVIRONMENTAL DATA ANALYSIS CENTRE (ALPENDAC)

Johannes Munke[1], Patrick Hannawald[2,3], Stephan Hachinger[1], Martin Schumann[1], Inga Beck[5], Alexander Götz[1], Oleg Goussev[3], Jana Handschuh[2,3], Helmut Heller[1], Roland Mair[4], Till Rehm[5], Bianca Wittmann[4], Sabine Wüst[3], Michael Bittner[2,3]

[1]*Leibniz Supercomputing Centre (LRZ) of the Bavarian Academy of Sciences & Humanities, Garching b. München, Germany*
[2]*University of Augsburg, Institute of Physics, Augsburg, Germany*
[3]*German Aerospace Center (DLR), Earth Observation Center, Weßling, Germany*
[4]*bifa Umweltinstitut GmbH, Augsburg, Germany*
[5]*Environmental Research Station Schneefernerhaus (UFS), Zugspitze, Germany*

INTRODUCTION OF THE AlpEnDAC

The **Alpine Environmental Data Analysis Center** (AlpEnDAC) is a "one-stop-shop" platform (www.alpendac.eu) for scientific data measured on high-altitude research stations in the alpine region and beyond. The platform provides **Research Data Management**, **analysis** and **simulation services** to support the research activities of the VAO community (Virtual Alpine Observatory, including the major European alpine research stations – www.vao.bayern.de/).

At the moment the system holds approx. **850 data products** and is accessed by various users and data contributors every day. With some new developments, we want to make environmental scientists profit from **near-real-time** (NRT) **data collection** and **processing**, as it is already an everyday tool e.g. in the Internet-of-Things sector and in commercial applications. Therefore some On Demand services like **Computing On Demand** and **Operating On Demand** are implemented in the platform.

The AlpEnDAC platform has been using infrastructure of the German Aerospace Center (DLR) and the Leibniz Supercomputing Centre (LRZ), major players in Europe's data and computing centre landscape.

RESEARCH DATA MANAGEMENT

With a strong focus on the **FAIR principles** [Wilkinson, M. et al., Scientific Data 3, 160018 (2016), https://doi.org/10.1038/sdata.2016.18]. the AlpEnDAC system offers a versatile platform for **Research Data Management** (Fig. 1). The platform is specialized on **timeseries** of climatological data, but also offers the possibility to upload other data types like e.g. **generic binary data** and **images**. Data products are always assigned with a sufficient set of **metadata** according to ISO 19115/39. The metadata set is also sufficient for obtaining Digital Object Identifiers (DOIs).

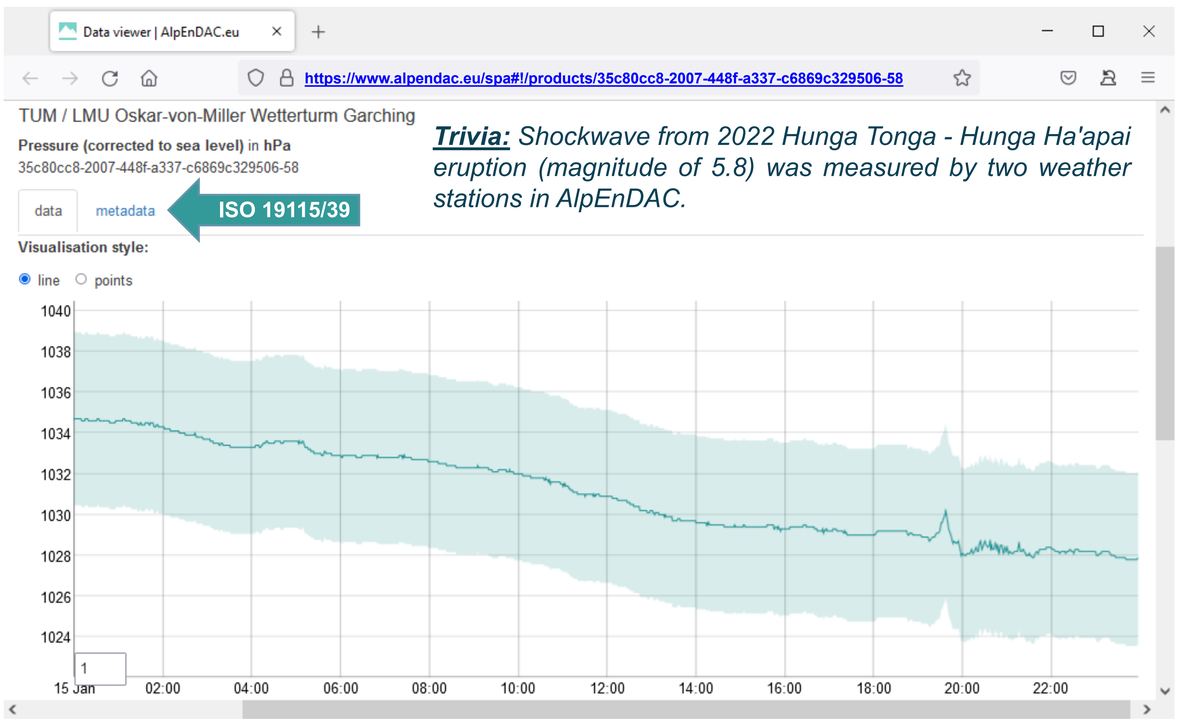

Figure 1: Screenshot of the visualization of an exemplary data product (data: LMU Meteorological Institute).

COMPUTING ON DEMAND

The AlpEnDAC platform offers a **Computing On Demand** service which enables to trigger the simulation of trajectory models. **Trajectory models** are very helpful tools for the investigation of atmospheric transport processes. For instance, they can be used for identifying source regions of air. There are numerous models for the computing of trajectories. Some of it can be used here for individual purposes. [www.alpendac.eu/cod]

The following models are available:

- **FLEXTRA**: Air mass trajectories [A. Stohl et al., flexpart.eu]
- **FLEXPART**: Particle dispersion [A. Stohl et al., flexpart.eu]
- **HYSPLIT**: Air mass trajectories [NOAA, ready.noaa.gov/HYSPLIT.php]

In addition, we offer a framework for computing infrasound propagation:

- **HARPA/DLR**: Sound ray trajectories [cf. Pilger & Bittner, J. Atmos. Sol.-Terr. Phys., 71(8), 816 (2009), https://doi.org/10.1016/j.jastp.2009.03.008]

OPERATING ON DEMAND

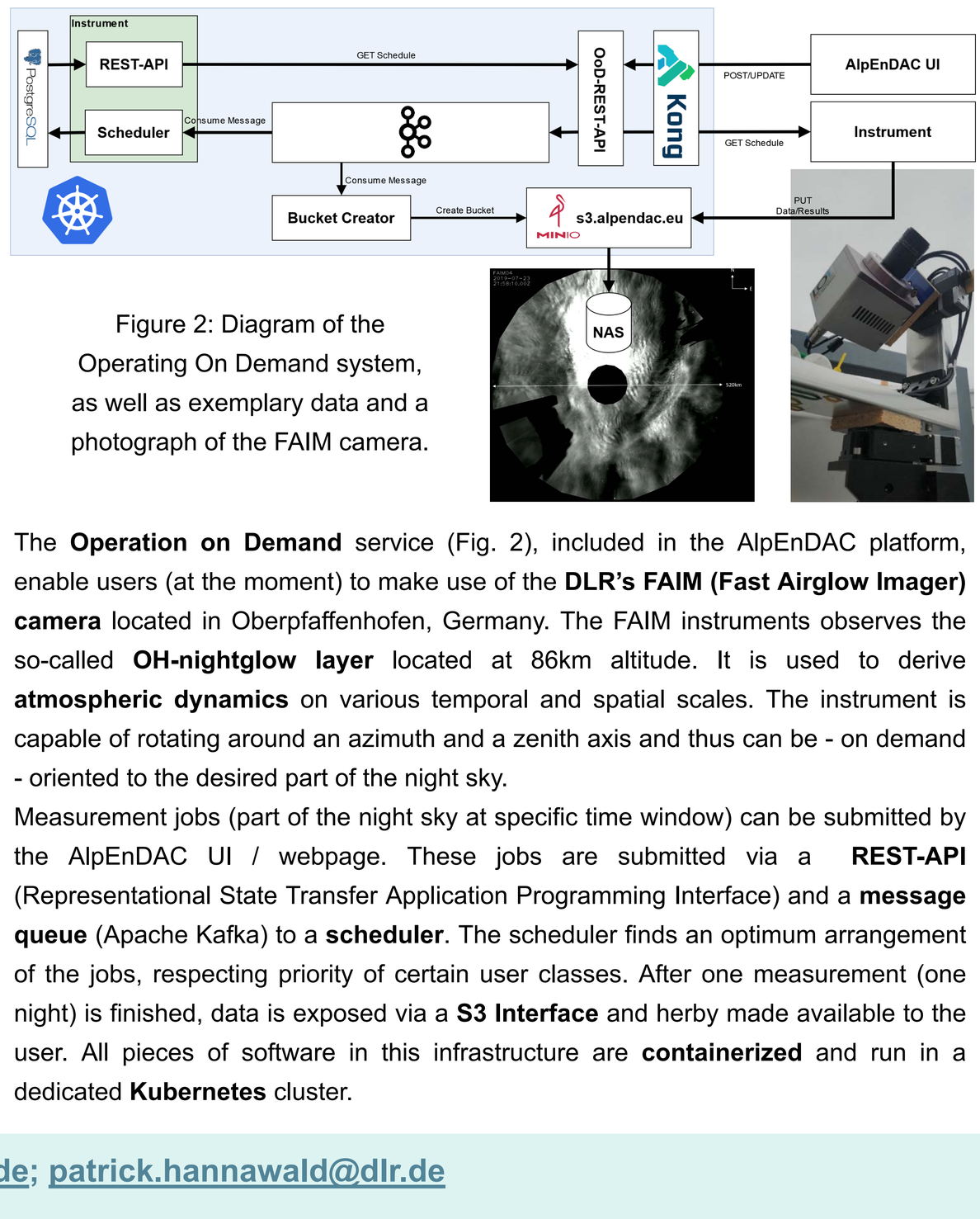

Figure 2: Diagram of the Operating On Demand system, as well as exemplary data and a photograph of the FAIM camera.

The **Operation on Demand** service (Fig. 2), included in the AlpEnDAC platform, enable users (at the moment) to make use of the **DLR's FAIM (Fast Airglow Imager) camera** located in Oberpfaffenhofen, Germany. The FAIM instruments observes the so-called **OH-nightglow layer** located at 86km altitude. It is used to derive **atmospheric dynamics** on various temporal and spatial scales. The instrument is capable of rotating around an azimuth and a zenith axis and thus can be - on demand - oriented to the desired part of the night sky.

Measurement jobs (part of the night sky at specific time window) can be submitted by the AlpEnDAC UI / webpage. These jobs are submitted via a **REST-API** (Representational State Transfer Application Programming Interface) and a **message queue** (Apache Kafka) to a **scheduler**. The scheduler finds an optimum arrangement of the jobs, respecting priority of certain user classes. After one measurement (one night) is finished, data is exposed via a **S3 Interface** and herby made available to the user. All pieces of software in this infrastructure are **containerized** and run in a dedicated **Kubernetes** cluster.

Contact: johannes.munke@lrz.de; patrick.hannawald@dlr.de

financed by Bavarian State Ministry of the Environment and Consumer Protection