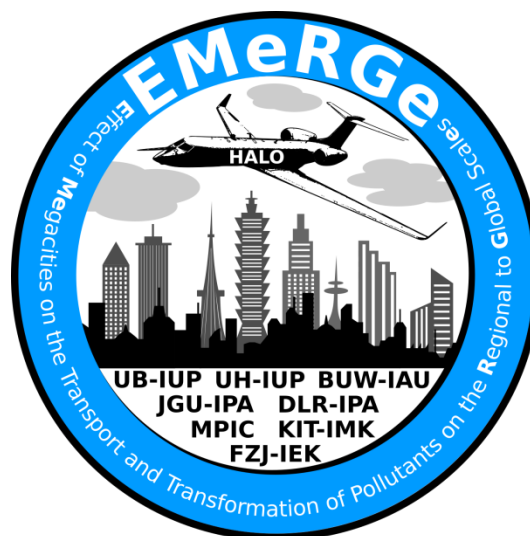


Data protocol

Effect of Megacities on the Transport and Transformation of Pollutants on the Regional to Global Scales (EMeRGe).

Mission EMeRGe Data protocol



Preamble

The number and size of major population centers (MPCs) is increasing worldwide. The investigation of the impact of transport and transformation of plumes from MPC has lately received much attention in the scientific community but our knowledge is still quite limited and inadequate.

EMeRGe exploits the unique capabilities of the new High Altitude and Long Range (HALO) aircraft research platform to investigate the impact of MPC emissions on air pollution at local, regional and hemispheric scales by making dedicated airborne measurement campaigns, coupled with interpretational and modelling studies of primarily the short lived climate pollutants, i.e. reactive gases, temporary reservoirs, and aerosol particles.

The transport and transformation of plumes from selected European and Asian MPC will be investigated with a special focus on the rate of formation of ozone and aerosols. Two field campaigns using an optimized payload of measurements on board the HALO platform based in Europe (summer 2017) and Asia (spring 2018) are planned.

In EMeRGe the composition of the plumes of pollution entering and leaving Europe and leaving Asia will be measured by performing optimized transects and vertical profiling. This strategy enables the outflow from a variety of MPCs, having different characteristics, to be studied in quasi-Lagrangian approach and thereby to identify differences and commonalities in the transport, and transformation of the outflow for European and Asian MPCs. Key scientific questions address the following topics: the dispersion and transport patterns of MPC plumes, the factors dominating the chemical transformation of MPCs emissions, the regional, hemispheric and global impact of European and Asian MPCs on atmospheric composition and pollution, the relevance and importance of emission from European and Asian MPCs for radiative forcing and climate change and the adequacy of chemical models for the simulation of transport and transformation processes of European and Asian MPC plumes.

An additional overarching objective of EMeRGe to the stand-alone EMeRGe objectives is to stimulate measurements and modelling studies, within an international EMeRGe research partnership with the European, Asian, and American science community. This will facilitate a much more comprehensive integrated analysis of all set of observational data products (i.e. Aircraft, ground and satellite based data), making EMeRGe a milestone in the study of the transport and transformation of plumes from MPCs.

EMeRGe aims to improve our knowledge and prediction of the transport and transformation patterns of European and Asian MPC pollutant outflows.

EMeRGe hypothesis is that the nature of the local emissions, coupled with the specific regional meteorology determines the transport and transformation pathways and patterns in the plumes from MPCs. To test this hypothesis, the following key scientific questions need to be answered:

- a) Which dispersion and transport patterns from the MPC outflows dominate in Europe and Asia for the selected time periods?
- b) What are the factors dominating the chemical transformation of MPC emissions?
- c) What is the regional, hemispheric impact of European and Asian MPCs in the change of atmospheric composition?
- d) What is the relevance of emission from European and Asian MPCs for radiative forcing and climate change?
- e) How adequate are chemical models for the simulation of transport and transformation processes of European and Asian MPC outflows?

The participants of the Effect of Megacities on the Transport and Transformation of Pollutants on the Regional to Global Scales campaign (EMeRGe)—in the remainder of this document this campaign will be referred to as *the mission*—agree on this data protocol.

This data protocol aims:

1. to ensure the fair use of the data,
2. to encourage the rapid dissemination of the scientific results,
3. to uphold the rights of the individual scientists,
4. to ensure the visibility and the integrity of the project,
5. to have all involved researchers treated equitably,
6. to encourage an orderly and timely analysis and publication of the data, and
7. to produce a central repository of the data to be released to the public domain.

This data protocol formulates the common scientific practice related to data—that has been collected in a joint effort by many scientists—as it has established in the field of atmospheric science for many decades.

§ 1 Definitions

In this document the term *principal investigator of the mission*—also called *Mission PI* or *PI of the mission*—refers to a person that is the scientific head of the mission.

A *principal investigator of an instrument*—also called *Instrument PI* or *PI of an Instrument*—is a person in charge of a scientific instrument.

A *principal investigator of a model* is a person in charge of a scientific model.

The term *participant of the mission* refers to a person that subscribed this document and was granted membership to the mission by the PI of the mission.

The term *quicklook* refers to a visualization of preliminary data that is produced shortly after the respective data set was created.

§ 2 Data management

(1) Data ownership: Data shall be property of the owner of the instrument which the data was created with.

(2) Data repository: The repository for the data of the mission is the HALO-database¹ operated by the *Institute of Atmospheric Physics* of the DLR.

(3) Citation of data: Reference to a persistent identifier such as a digital object identifier (DOI) shall be applied to the data and/or parts of it.

(4) Data format: The participants must provide the data in standard scientific data formats such as the NASA-Ames format or the NetCDF format. The usage of standard names for the observed entities according to the CF convention is encouraged.

(5) Data quality: The analysis of the scientific data has to be made following the *rules of good scientific practice*² defined by the Deutsche Forschungsgemeinschaft (DFG). This includes in

¹ <https://halo-db.pa.op.dlr.de>

particular that the analysis procedures have to be transparently documented and the scientific data and the documentation necessary to understand the data has to be stored for at least 10 years. A long time preservation of the data beyond 10 years is encouraged.

§ 3 Provision of Data/Data licensing and sharing

(1) Provision of data: Each group will provide *quicklooks* of their data to the EMeRGe HALO participants shortly after each flight. The data retrieved from the individual scientific instruments along with the corresponding documentation to understand the data shall be made available to all participants as soon as possible. In general data sets must contain qualifiers of the status of the analysis (preliminary, final, etc.).

- 6 months after the observations were made a preliminary data and the documentation must be made available to all EMeRGe participants.
- 18 months after the observations were made the a final data and the documentation must be made available to all EMeRGe participants.
- 36 months after the observations were the final data and documentation must be made available publically under DFG ethical rules.

Corrections and amendments to preliminary data must be made available to all participants as soon as possible.

(2) Data licensing and sharing: Until the end of the first 3 years (period of time) after the end of the mission the data is only available to the participants of the mission.

An individual scientist wishing to collaborate (Associated scientist) can apply to get access to the data of interest. The Associated scientist must provide a short written description of her/his aims to the PI of the mission. This will be considered by the principal investigator of the mission and access to the data will be granted after approval by the respective instrument PI.

During the period of time as set forth in this paragraph any publication or forwarding of data to public, non-participants of the mission or non-Associated scientists is prohibited. After the expiration of the period of time set forth in this paragraph § 3 (3) shall apply.

(3) After the period of time set forth in § 3 (2) the PI of an instrument must decide if the data, created with the instrument he/she is responsible for, shall be released to individual persons or to the public by one of the following options:

closed: A potential data user can request the data of the instrument by contacting the respective Instrument PI. The data can be made available to that person under the conditions defined in the agreement in Appendix A at the end of this data protocol. This agreement has to be signed by this person and returned to the respective Instrument PI. The final decision to share or not to share the data with that person remains by the respective Instrument PI.

or

open: The data is made publicly available under a suitable open license such as one of the Creative Commons licenses³ as recommended by the *Allianz der deutschen Wissenschaftsorganisationen*.

In either case the receiver(s) of the data are obliged to accept the provisions as stipulated in § 5 et seq.

² http://www.dfg.de/en/research_funding/principles_dfg_funding/good_scientific_practicedatabase

³ <https://creativecommons.org>

§ 4 Model studies

- (1) Model results: Results of model studies using data of the mission must be made available to the participants as soon as possible.
- (2) Corrections: Corrections and amendments to preliminary results must be communicated to the participants of the mission as soon as possible.
- (3) Recent version of data: The principal investigator of a model must ensure that the data used in the modeling study is the best available at that time.

§ 5 Publications/Conference Contributions/Public meetings

- (1) Co-authorship: If data of the mission is used for publications (ex: scientific articles), digital media, conference contributions and/or in public meetings the lead author has to offer co-authorship to the creator(s) of the data.
- (2) Right to refuse: The creator(s) that created the data has/have the right to refuse the usage of the data in publications, conference contributions and/or in public meetings prior to her/his publication of that work.
- (3) Co-authors: The manuscript for any journal publication must be sent to the co-authors at least 3 weeks before submission with a copy to the principal investigator of the mission.

Abstracts to contributions to conferences and/or public meetings must be sent to the co-authors at least 1 week before submission with a copy to the principal investigator of the mission.

The lead author has to take into account substantial comments from the co-authors before the submission of the manuscript or abstract. The co-authors and the principal investigator of the mission shall be kept informed of the further progress of the review and publication.

- (4) Usage of data: The lead author must have the explicit agreement of the creator(s) that created the data for the usage of the data in publications, conference contributions and/or public meetings.
- (5) Recent version of data: The PI of an instrument must ensure that the data of this instrument used in publications, conference contributions and/or public meetings is the best available at that time.
- (6) Acknowledgement: The mission must be acknowledged in a suitable way.

§ 6 Press releases

- (1) Publications and/or press releases where the mission itself is the main subject—so called overview articles—must be approved by all PIs of the mission and all PIs of the instruments before submission.
- (2) Press releases relating to data, created during a mission, must be approved by the creator(s) of the data.

§ 7 Collaborations

Effective and productive collaborations between the participants of the mission and other projects are encouraged.

§ 8 Avoidance of disputes

(1) Any disputes about the use of the data in particular with respect to publications will be resolved by a committee of scientists suggested by the principal investigator of the mission and approved by the instrument PIs. For HALO EMeRGe this body is the EMeRGe scientific steering committee.

(2) During the periods defined in § 3, when the data is still not public, it is encouraged to hold scientific meetings on a regular basis where all participants of the mission can present their ongoing work and their plans for future work with the data of the mission. This shall be made to avoid redundant work and parallel efforts which are both sources of potential disputes.

§ 9 Constraints

National and international laws or regulations might constrain the agreements of this data protocol.

In that case the principal investigator of the mission has to inform the instrument PIs about those constraints. The mission PI and the instrument PIs shall find a mutual solution on a case-by-case basis.

The undersigned agrees to the conditions of the data protocol.

Signature of participant:

Date, place:

Name:

Position (PI, post-doc, student, etc.):

Address:

.....

E-mail:

The principal investigator of the mission grants the membership of the mission.

Signature of principal investigator of the mission:

Date, place:

Please, return to EMeRGe principal investigator:

Prof. J.-P. Burrows
 University of Bremen
 Institute of Environmental Physics (IUP)
 Otto-Hahn-Allee 1
 28359 Bremen, Germany
burrows@iup.physik.uni-bremen.de

Appendix A

Agreement to receive the data of an instrument that was licensed according to §§ 2 and 3

(closed version)

I will receive a copy of a data set of the mission Effect of Megacities on the Transport and Transformation of Pollutants on the Regional to Global Scales.

I agree to the provisions stipulated in the Data protocol of the mission „Effect of Megacities on the Transport and Transformation of Pollutants on the Regional to Global Scales“ and the following common principle:

“If data of the mission is used for a publication, conference contribution and/or public meeting, co-authorship has to be offered to the creators of the data.”

I will not distribute this data set or parts of it.

Signature of the receiver of the data set:

Date, place:

Name:

Affiliation and address:

.....

E-mail: