

# COOPEUS

## Deliverable 2.1.

---

Fact Finding Report on Incoherent Scatter Data

Prepared by EISCAT Scientific Association

---

The fact-finding was compromised by lack of communication with the US partners due to funding problems and a resulting delay of work. At present the facts can be summarized as follows.

### 1. EISCAT Observations:

The majority of the observation time is devoted to the Common Programme and to the Special Programme to about equal amounts ( $\approx 45\%$  each). The Common observation program is decided by EISCAT Council based on recommendation of the Science Oversight Committee. The Special Programme time is distributed by researchers in the EISCAT member countries, each of them being allocated observation time in proportion of the countries payments into the Association. The total amount of observation time within the Special Programme is roughly 1200 hrs per year. In addition a fraction of Special Programme time is granted based on open peer review to scientists from any country. The open Programme grants roughly 200 hrs per year in response to submitted proposals for of the order of 400 hours of observation time. Within the different observation programmes EISCAT currently runs several different observation modes and at least one, often several, analyzed data sets are permanently stored corresponding to a set of correlated data.

The Special Programme data are exclusively for the observers during the first year and then open to all researchers within the EISCAT member countries. EISCAT Scientific Association is currently planning to revise its statutes and to include free open access to all observational data after 3 years. This time span was assumed to be appropriate for protecting the work of individual scientists and especially of PhD students on specific topics.

The Common Programme data are immediately open to researchers from the member countries. Many observations within the Common Programme are carried out in coordination with other incoherent radar measurements around the globe. In addition it is under discussion how to provide a level of standard monitoring for immediate access to the public and how to fund the preparation of this data product.

EISCAT data are at present available directly from the EISCAT WebPage and through the Madrigal database.

## 2. Access to EISCAT data:

The EISCAT WebPage is only available for a limited number of users at the same time. EISCAT provides metadata descriptions for its data sets. Partly EISCAT also provides a catalogue of the data that is based on these metadata descriptions. The catalogue and the meta data are open to public and available in electronic format. The data, metadata and catalogue are available in ASCII format and searchable via the internet<sup>1</sup>. Terminologies to describe the data are partly standardized. The EISCAT data services are not registered at GEOSS. EISCAT data are partly open to public and specialised APIs for different programming languages are offered. EISCAT exclusively manages data which originates directly from the RI and there is no formal procedure to submit, review and appraise data submissions.

The EISCAT raw data are automatically transferred to the archive and the archive verifies each submitted data set for completeness and correctness. Further data products are derived from the raw data and it is well-defined what metadata must be saved when these data products are initially created. The actions performed during the ingest process are recorded. Data processing steps are indicated within the metadata. At present EISCAT regularly checks data sets for integrity. There is no regular update of data formats.

At present EISCAT does not follow a formalized data management workflow, but rather the data access developed during the 30-year lifetime of the association and through communication with users. A direct exchange with the EISCAT users takes place during the different meetings that are organized by the association<sup>2</sup>, but there is no formal feedback mechanism in place to ensure that the data formats meet the communities requirement. There is also no feedback mechanism in place, which allows users to report on errors in data or

---

<sup>1</sup> <http://www.eiscat.se/raw/schedule/schedule.cgi?A=on>

<sup>2</sup> EISCAT International Symposium is a biannual scientific meeting (previously EISCAT International Workshop for exchange between EISCAT users and scientists working in related fields. EISCAT radar school is an annual workshop for students and new users to get acquainted with the system. EISCAT\_3D user meeting is an annual workshop to engage the users in the design and planning for the new instrument.

metadata. The user activities are not logged. The archive has an authentication mechanism in place, but does not use licences or other means to claim IPR.

The data dissemination and access policy is under re-negotiation among the EISCAT Associates and EISCAT intends to build up a strategy which data sets are suitable for long-term archiving. It is envisioned to have different policies for real time data and archived data.

### 3. Access to Madrigal

Madrigal is an upper atmospheric science database used by atmospheric research groups throughout the world. Madrigal is a robust, World Wide Web based system capable of managing and serving archival and real-time data, in a variety of formats, from a wide range of upper atmospheric science instruments. Data at each Madrigal site is locally controlled and can be updated at any time, but shared metadata between Madrigal sites allow searching of all Madrigal sites at once from any Madrigal site.

Data can be accessed from a variety of Madrigal sites, including (but not limited to) Millstone Hill, USA, Arecibo, Puerto Rico, EISCAT, Norway, SRI International, USA, Cornell University, USA, Jicamarca, Peru, the Institute of Geodesy and Geophysics, the Chinese Academy of Sciences, and finally, the archival CEDAR<sup>3</sup> site. Data can also be accessed directly, using APIs which are available for several popular programming languages (Matlab, python, and IDL). The latest version of Madrigal and the remote API's may also be downloaded from Open Madrigal Web site.

The Madrigal system will be used as a basis for discussing data harmonization with the US partners. The system includes in principle the opportunity that the incoherent scatter data instruments even use globally the same raw data format.

### 4. Evaluation of data access:

The Madrigal data base is frequently used, but only by a fraction of the potential users: experienced EISCAT or incoherent scatter radar users. In addition, the data should be easily accessible to (a) new users that are interested in the incoherent scatter technique, to (b) environmental and space weather modellers and service providers, and to (c) occasional users interested in EISCAT measurements for short-duration research projects or as source of supporting data.

---

<sup>3</sup> Coupling, Energetics and Dynamics of Atmospheric Regions (CEDAR) is a focused Global Change program sponsored by the National Science Foundation

The current data documentation and the service provided requires considerable effort from new users and potentially discourages those users from the last two categories that are primarily interested in standard data products. Some of the aspects that make it complicated to use the data were found from discussion with a researcher who used EISCAT measurements to include them in atmospheric models (“an occasional user”). The following list may serve as an example for the problems that an occasional user encountered:

The content of the data:

- the data includes several measurements with different modes for the same day;
- the data are often deviating and no clear information is given which one is the best and which one to use;
- the data base contains uncorrected and corrected numbers for e.g. electron density without describing how the values were corrected;
- a user found up to 13 different measurements for the same time and place with values that differed exceeding error values;
- in some cases the viewing direction of the antenna during observation is not given;
- some data were collected during heating campaigns (i.e. active modification of the ionosphere by low HF radio waves) without this being documented.

The format of the data:

- data gaps are not properly marked;
- some files have wrong line-breaks;
- some data lack information on geometry
- some problems with time intervals, start and stop times of observations.

