

EISCAT metadata and data formats

Carl-Fredrik Enell, EISCAT Headquarters

April 15, 2015

1 Introduction

The purpose of this document is to

1. describe the types and volumes of incoherent scatter radar and ionosonde data archived at the Headquarters of EISCAT Scientific Association (henceforth EISCAT);
2. summarize the sources, formats and volumes of metadata available for these data.

1.1 EISCAT systems

EISCAT operates

1. Three incoherent scatter radars
Out of these two are on the northern Scandinavian mainland: the UHF at Ramfjord near Tromsø, Norway (UHF or TRO) and the tristatic VHF radar with remote stations in Kiruna, Sweden and Sodankylä, Finland (VHF/EIS, KIR, SOD).
The third system is the EISCAT Svalbard radar (ESR) in Longyearbyen, Svalbard, with one transmitter and two antennas (LYR 32 m and LYR 42 m).
2. An ionospheric HF heater in Tromsø (with HF radar capabilities)
3. Two ionosondes, one at each radar transmitter site

1.2 Instrument coordination

When analysing radar data, there is often a need to take into account the operations of other instruments, both within EISCAT and by external parties. This is the case when the heater is running together with the mainland radars, when dynasonde data are needed in order to interpret radar data and sometimes when complementary instruments are operated at the EISCAT sites or elsewhere together with EISCAT operations.

In all these cases coordination of measurements is required. Data portals like ESPAS (<http://www.espas-fp7.eu>), ENVRI and the upcoming EISCAT_3D data system also specifically aim at finding simultaneous and spatially coincident data from multiple instruments, both ground-based and on satellites.

The coordination between instruments is facilitated by the availability of GPS time stamps. Thus, timing is the natural way to find overlapping complementary (meta)data. However, in future an interface to include notes and metadata e.g. of heater operations in radar metadata files would be important.

2 Data products and formats

2.1 Incoherent scatter radar data

EISCAT archives incoherent scatter data at levels 2 (correlated data) and 3 (analysed parameters), and to a limited extent at level 1 (voltage domain samples). See the Data level document https://www.eiscat.se/about/experiments2/eiscat-data-levels/at_download/file for a description of how EISCAT data levels are defined. The archival and data retrieval systems are entirely different for levels 2(+1) and level 3.

Data archives

1. Level 2 data files (which for some experiments also contain data at level 1). The files are in .mat format, compatible with Matlab® version 4, and are stored in a RAID archive. These files are indexed in a MySQL database. A web-based search and retrieval system is implemented in Apache + Python CGI.

Approximate volume of data since 1981: 50 TB

2. Analysed data (level 3) are stored in a Madrigal database, the de facto standard for distributing incoherent scatter data. Madrigal is a distributed system where each site stores data from its own instruments locally whereas metadata are shared between the sites. A Madrigal site is based on a directory structure with analysed and derived parameters stored in binary NCAR format files. HDF5 files are also created when inserting data but are so far a secondary product.

File indices and other local and shared metadata are in plain text files. The system has a comprehensive set of web services for data retrieval (Apache, CGI and APIs for several languages). See <http://www.openmadrigal.org> and the EISCAT install at <http://www.eiscat.se/madrigal>.

A new release of Madrigal will soon be available. It will use HDF5 directly as its only storage format. These HDF5 files will contain sufficient metadata to be self-explanatory.

Approximate volume of data since 1981: 110 GB

2.2 HF radar (heating)

The EISCAT Heater transmitter and receiver systems are nowadays similar to the incoherent scatter radar controllers and therefore use the same format of control and data files as the radars. Before the upgrade there were no radar capabilities, only transmission.

2.3 EISCAT ionosondes

The EISCAT ionosondes are developed from old HF radars and use the Dynasonde (DSND) pulse and analysis scheme developed at the University of Colorado by Bill Wright, Nikolay Zobotin and coworkers. The ionosonde data are important also as radar metadata, since the critical frequency of the F2 layer (foF2) is related to the electron density at the F2 peak and is often used for absolute calibration of the radar data.

Raw data The Tromsø ionosonde stores raw in-phase and quadrature (IQ)-sampled data from identified echoes on CDs and DVDs at the Heating site in Ramfjord. For the volume of discs see <http://www.eiscat.uit.no/heating/Dynasonde/logs/cdr-log/>.

The Svalbard ionosonde is an older system based on a hardware signal processing card and stores only analysed products.

SQL database A database (MySQL) stores parameters of identified volumetric radio echoes as well as a set of several automatically analysed ionospheric parameters including critical frequencies and density profiles. The total volume of data is around 1 TB.

3 Incoherent scatter radar metadata

Several disparate sources of metadata are required in order to fully specify and understand a radar experiment. In the following the incoherent scatter radar metadata sources and their formats and present volumes are described.

3.1 Radar schedule

The online radar schedule <http://www.eiscat.se/schedule/schedule.cgi> is the primary source of information on radar operations. It is also the preferred portal for level 2 (raw correlated and integrated) data and offers online reanalysis of these data.

Interface Web-based, Apache + CGI scripts mainly in Python, own development

Schedule database Directory structure with text files, size about 46 MB

Level 2 data index SQL database

Reanalysis Text file written when user selects reanalysis, cron script on analysis computer checks this file, calls Matlab® and emails a note to the user

3.2 Radar experiment and analysis specifications

The highest level of metadata (in addition to experiment scheduling) in an EISCAT radar experiment is the experiment configuration loaded at runtime by the EISCAT Realtime Operating System (EROS) and by the GUISDAP (originally Grand Unified Incoherent Scatter Design and Analysis Package) data analysis

software. Each experiment is defined by a number of files in an experiment directory and can be thought of as two separate fast and slow parts: the transmitted and received code, and the antenna scan pattern, respectively.

Location Information directories adjacent to the data, main site computers, GUIDAP analysis software distribution

Size Total some 2.5 GB

3.2.1 Old (pre-2000) system

The mainland radars, which were designed in the 1970's, used hardware correlators and Norsk Data SINTRAN minicomputers, storing data on magnetic tapes. These systems were in operation until 2000. The experiments were defined by the following files, most of which should be available in info directories together with the archived data (to the extent it has been possible to copy the tapes to the archive)

Configuration files

- Main experiment script: EISCAT Language (ELAN) script
- Definition of transmission and reception: Transmit and Receive Language (TARLAN) text file, compiled to binary code for the radar controllers
- Configuration of the correlator: Correlator Language (CLAN) file
- Graphics definition: GDEF file, text file describing format of data
- Information and metadata can often be found in experiment description (desc) and log files

Notes The experiments were usually named for intended Common Programme runs: cpN where N is a number. Each ELAN experiment loaded radar configurations and included scan patterns.

3.2.2 New system (Svalbard radar and mainland since 2000)

The Svalbard radar was designed in the 1990's using a system based on digital signal processing boards and SUN Sparc computers, and the mainland radars were upgraded with new digital receiver and transmitter hardware in 2000. The systems still use single VME board Sparc computers for channel board readout and parts of the signal processing, but the main SUN servers have been replaced with standard PC computers running Open Indiana, a derivative of Solaris. Some of the configuration files have the same structure as those of the old systems and a number of new configurations have been added.

Configuration files

- Main experiment script: Eiscat Language (ELAN) script, here based on extensions to the TCL script language interpreted by the **wishe** shell
- Antenna scan patterns: ELAN script subroutines

- Definition of transmission and reception: Transmit and Receive Language (TARLAN) text file, compiled to binary code for the radar controllers
- Exciter frequency setting (frq) files
- Channel board frequency setting (nco) files
- Configuration of channel boards and decoder software: Filter (fil) file, compiled to binary code
- Configuration of alternating code: code file
- Configuration of real time graph: Matlab file rtg_def.m

Notes

- The code programs have names many of which were originally acronyms: **beata**, **manda**, etc. Most of these are alternating codes, each designed for a certain time and range resolution. Several files are required to define a code experiment and the usual procedure is to write a Python program which generates those files.
- The antenna scan programs are usually named for their intended use: **cp1** for Common Programme 1 operation, etc.

Analysis configuration The analysis uses the Matlab® analysis software package GUIDAP. Analysis configurations, i.e. information on how data are arranged in the data dumps, are compiled from the experiment configurations. Importantly, the range and time resolution properties of the radar code are used to calculate so-called spectral ambiguity functions, which are used to convert theoretical scatter spectra to the lag domain.

3.3 Level 1+2 data

This archive is on a RAID server located at EISCAT Headquarters in Kiruna. It is behind a NAT firewall and thus does not have a public IP address, but file download is possible through the schedule interface on <http://www.eiscat.se>. A backup copy is kept on a similar system at the Kiruna receiver site.

3.3.1 Metadata at experiment and storage level

- Directory names
The raw data directories are sorted as a three-level tree:
 1. Year
 2. Experiment name: Pulse code, antenna scan, version, and country code (for accounting of EISCAT hours per associate country)
 3. Date and hour
- File names: 8 digit integers, seconds since 1 January 00:00 of the year + **.mat.bz2**
- Experiment setup, from name and version of experiment, EROS configuration on site server

3.3.2 SQL databases

The index of the data archive consists of two MySQL databases as follows.

- `tape_archive`, the active database of experiment data
- `tape_archive_tapes`, index of historical tapes that have been converted to files on RAID

The databases index: experiments run: name, country and instrument, start and stop times and time used, and directories at the hourly level. The table structure is identical:

experiments

experiment_id	experiment_name	country	antenna	comment
Number	String	2 char country code	3 char name	

resource

resource_id	experiment_id	start	end	comment	type	account
Number	Number	Date	Date			string

storage

location	resource_id	priority	bytes	comment
String	Number	Number	Number	

tape_comments

tape_nr	comment
Number	String

3.3.3 Data and metadata in the archive files

The EISCAT archive `.mat` files are compatible with Matlab® since version 4 and can be opened with all compatible software. In addition to Matlab®, libraries are available for several common programming languages such as Python and GNU R. The files contain the following data vectors. The indexing of these data vectors is different for different experiments and it is often most convenient to work out the formats from the corresponding `rtg_def` or `gdef` files.

1. `d_ExpInfo` A short text string, name of experiment
2. `d_raw` Level 1 samples (16-bit complex integers), **saved only for special experiments.** `d_raw` is most often used for transmitter samples. For dedicated experiments, such as meteors and space debris, also in-phase and quadrature signals after digital mixing to base-band and decimating filters, data rate 15/M MHz where M is the decimation factor)
3. `d_data` Data dumps: Correlated signal, lag matrices of all possible lags decoded from one or more sets of pulse codes and frequencies, including calibration signals (noise injected at antennas). Ion and plasma line receiver data as well as ion line data from the two antennas of ESR may be combined. Complex floating point values (complex double). Indexing of data in `d_data`
4. `d_parb1` Parameter block. The parameter block is a vector of up to 128 numbers. The format is described in Sect. 6

3.4 Level 3 data

The calibrated level 2 autocorrelation domain data are analysed by the software package GUISDAP. GUISDAP runs under Matlab® and stores the results temporarily in Matlab version 6 or higher `.mat` files. These contain much information but are not archived once the results have been converted to NCAR files for Madrigal. GUISDAP also stores information files and an overview plot for each run.

The Madrigal database is the authoritative archive of level 3 data. It is highly advisable to use the Madrigal APIs or web GUI to retrieve (meta)data from the Madrigal system.

3.4.1 Sources of level 3 metadata

- Experiment and analysis setups as described above
- File names, should always contain
 - Date
 - Experiment name
 - Integration time
 - Instrument name
- Madrigal index files and APIs
- NCAR (and derived HDF5, future only HDF5) files in the Madrigal experiment directory structure. These files contain parameters like
 - Date and time
 - Antenna pointing (azimuth, elevation)
 - System temperature
 - Power
 - Fitted parameters with error limits, as function of range
 - Derived parameters like scatter angle, ground altitude
- Overview plots of
 1. electron density
 2. electron temperature
 3. ion temperature
 4. line of sight velocity
 5. the following parameters:
 - System temperature
 - Power
 - Azimuth
 - Elevation

4 Heating (HF radar) metadata

The authoritative source of metadata on heater operation is a hand-written paper logbook. Information can also be typed into the web-based information system of the transmitter site but this is rarely used of little use.

Scanned copies of the logbooks are available online at (??)

Since 20xx (??) the heater system is also controlled by EROS configuration files which can be found in their own experiment directories.

5 Ionosonde metadata

The ionosondes store both metadata and data in one SQL database per site, *tromso* and *svalbard*. A search interface is available at <http://dynserv.eiscat.uit.no>.

6 Appendix: Parameter block format

This appendix describes the format of `d_parb1` in the EISCAT `.mat`-compatible archive files. Not all parameters are used at all radar sites and parameters have been introduced as needed since operations started. See also the Notes following the list of parameters.

6.1 Old (pre-2000) format

Entry	Value
1	Site code
2-4	Time of dump
2	$(year - 1900) * 100 + month$
3	$day * 100 + hour$
4	$min * 100 + second$
5	UHF commanded azimuth $int(az * 10. + 0.5)$
6	UHF current true azimuth
7	UHF uncorrected hardware azimuth $int((az - 360) * 100. + .5)$
8	UHF commanded elevation $int(el * 10. + 0.5)$
9	UHF current true elevation
10	UHF uncorrected hardware elevation $int((el - 360) * 100. + .5)$
11	UHF local range $int(rn * 10. + 0.5)$ in 0.1 km to center of common volume
12	UHF height of center of common volume in 0.1 km
5	VHF phase steering index for "W": -17...0...+17
6	VHF phase steering index for "E"
7	VHF angle W (segment 1) $elevation * 10$
8	VHF hardware angle W $elevation * 10$
9	VHF angle E (segment 3) $elevation * 10$
10	VHF hardware angle E $elevation * 10$
11	VHF beam mode
12	VHF nvstat
14	VHF azimuth 1
15	VHF azimuth 2
14	UHF polarisation phase 0-409 degrees
15	UHF polarisation amplitude ratio, -127-127 in 0.25 dB steps
13	first local oscillator frequency (MHZ*10)
16	signal path switch
17-18	signal attenuators of path X and Y, 0-63 dB
19-26	second local oscillator frequency for channel 1-8 (kHz*10)
27-34	channel attenuator settings of channel 1-8, 0-63 dB
35-42	filter bandwidth settings of channel 1-8, (kHz*10)
43-44	Noise injection control
45	Correlator program start address
46-61	correlator APB stack
62-77	correlator APM stack
78-85	ADC sampling intervals for channels 1-8
86	High-precision synthesizer channels for LO2
87-88	filter type used in channels 1-8

89	Pulse propagation delay offset, microseconds
90	Pulse propagation delay, local relative to Tromsø, antenna position dependent
91	Counter for outermost DO-loop in ELAN program
92	Radar controller program number
93	Pulse repetition period, unit 10 microseconds
94	Integration time, seconds
95	Status word
96	Average transmitter power, UHF, VHF klystron A, kW
97	Average transmitter high voltage, UHF, VHF klystron A, kV
98	100/(RF duty cycle)
99	Derived peak power UHF, VHF klystron A
100	Average transmitter power, VHF klystron B, kW
101	Derived peak power, VHF klystron B
102	Reserved for outside temperature
111–125	User parameters
126	Line number of last executed ELAN statement when this dump was done
127	Source
128	Parameter block version

Notes

- Parameter 1: 1=Kiruna, 2=Tromsø, 4=Sodankylä
- Parameters 2–4: end of dump to nearest second
- VHF parameter 5–6: corresponding to -21.3 deg to +21.3 deg in steps of about 1.25
- UHF parameters 11–12: if range or height $> 2^{15} - 1$ their value is replaced by $-(value - 2^{15})$
- VHF parameter 11:
 1. setting not defined to EROS
 2. single beam (all horizontal dipoles connected to one klystron, all vertical dipoles to the other, polarization depends on radar controller program)
 3. dual beam (each klystron feeds one antenna half, only circular polarization possible)
- VHF parameter 12:
 - 0=access ok, all segments under computer control and in position,
 - 1=interface sequence read error, status undefined
 - < -1 =port access error, SINTRAN operating system error, code=abs(nvstat)
 - > 0 = segment related bit pattern bit N =segment N in manual or slave mode ($N = 1..4$)
 - bit $N + 4$ =segment N moving ($N=1..4$)
 - bit 9=segments 1/2 ("W") misaligned, see status bit 3
 - bit 10= segments 3/4 ("E") misaligned, see status bit 4

- Parameter 16:
 - 0: X to CH 2,4,6,8, Y to 1,3,5,7
 - 1: Y to CH 1,2,3,4,5,6,7,8
 - 2: X to CH 1,2,3,4,5,6,7,8 3: X to CH 1,3,5,7, Y to CH 2,4,6,8
- Parameters 43–44: Always 0, noise injection is fixed and activated by radar controller
- Parameters 78–85: unit 0.05 microseconds for channels 1–2, 0.1 microseconds for channels 3–8
- Parameter 86: Bits 1 to 8 are used, corresponding to LO2 channels 1–8. If the bit is set, then a high precision oscillator is used on the channel. For those channels, errors in the frequency setting are not indicated by bit 5 of the status word, as the synthesizers cannot be read back.
- Parameters 87–88: 4-bit code,
 - 1=Butterworth, 2=linear, 3=wideband filter.
 Par 88 contains codes for channels 1–4 with the channel 1 code in the least significant 4 bits. Par 87 contains channels 5–8 with the channel 8 code in the most significant 4 bits.
- UHF parameter 95:
 - bit 0: UHF transmitter off
 - bit 1: azimuth not in position
 - bit 2: elevation not in position
 - bit 3: polarizer phase not in position,
 - bit 4: polarizer amplitude not in position
 - bit 5: receiver settings differ from commanded values
 - bit 6: error in correlator- or DMA dump,
 - bit 7: communication to Tromsø interrupted (transmitter status bit 0 is then meaningless in KIR and SOD)
 - bit 8: heating facility transmitter in standby (Tromsø only)
 - bit 9: heating facility transmitter on (Tromsø only)
 - bit 10: heating facility feed lines arcing (Tromsø only)
- VHF parameter 95:
 - bit 0: VHF transmitter RF off
 - bit 1: antenna "W" half not in position or not accessible
 - bit 2: antenna "E" half not in position or not accessible
 - bit 3: antenna segments "W" (1/2) misaligned
 - bit 4: antenna segments "E" (3/4) misaligned
 - bits 5–10 as UHF

- Source parameter 127:
 - bit 0: 1=VHF antenna, 0=UHF antenna
 - bit 1: 1=spectrum analyser
 - bit 2: 1=special device
 - bit 3: 1=VHF correlator, 0=UHF correlator
 - bit 4: 1=passive experiment, 0=active experiment
- Version parameter 128:
 - 6=after 1984-06-13
 - 7=after 1985-01-01
 - 8=after 1985-12-01
 - 9=after 1987-08-17
 - 10=after 1992-03-01

6.2 Current (post-2000) format

Entry	Value	Introduced
1	Dump end year	
2	Dump end month	
3	Dump end days	
4	Dump end hours	
5	Dump end minutes	
6	Dump end seconds	
7	Integration time, sec	
8	Combined output power, W	
9	Elevation, degrees	
10	Azimuth, degrees	
11	Dump end time, secs since 1970	
12	Dump sequence number	
13	% power tx1 klystron a (of 62.5 kW)	
14	% power tx1 klystron b	
15	% power tx2 klystron a	
16	% power tx2 klystron b	
17	% power tx3 klystron a	
18	% power tx3 klystron b	
19	% power tx4 klystron a	
20	% power tx4 klystron b	
21	Noise injection calibration, K	
22	Pre-integration factor	
23	% power tx5 klystron a	
24	% power tx5 klystron b	
25	% power tx6 klystron a	
26	% power tx6 klystron b	
27	% power tx7 klystron a	
28	% power tx7 klystron b	
29	% power tx8 klystron a	
30	% power tx8 klystron b	
31	rx frequency, channel 1 (MHz)	1999-08-19
32	rx frequency, channel 2 (MHz)	1999-08-19
33	rx frequency, channel 3 (MHz)	1999-08-19
34	rx frequency, channel 4 (MHz)	1999-08-19
35	rx frequency, channel 5 (MHz)	1999-08-19
36	rx frequency, channel 6 (MHz)	1999-08-19
37	rx frequency, channel 7 (MHz)	1999-08-19
38	rx frequency, channel 8 (MHz)	1999-08-19
39	rx frequency, channel 9 (MHz)	1999-08-19

40	d_parbl version number	1999-11-12
41	antenna ID (1-8) see notes	1999-11-12
42	Remote antenna intersection range, m	2003-02-28
43-62	User parameters	2003-07-23
63	High voltage reading (V)	2003-09-01
64	Loop counter	2004-02-01
65	ESR Peak power read from power meter (W)	2004-02-20
66	ESR RF duty cycle calculated from RC binary	2004-03-20
67	ESR SPEAR tx status	2007-03-22
68	ESR LO settings	2007-05-14
69	ESR CHI attenuator settings (dB)	2007-05-14
70	ESR CHII attenuator settings (dB)	2007-05-14
71	ESR Peak power in wave guide 32m antenna (kW)	2008-05-24
72	ESR Peak power in wave guide 42m antenna (kW)	2008-05-24
65	UHF Peak power read from wave guide (W)	2004-02-20
66	UHF RF duty cycle read from wave guide	2004-03-20
67	UHF Power status on Tromsø systems	2006-11-01
65	VHF antenna elevation panel 1 (deg)	2005-05-21
66	VHF antenna elevation panel 2 (deg)	2005-05-21
67	VHF antenna elevation panel 3 (deg)	2005-05-21
68	VHF antenna elevation panel 4 (deg)	2005-05-21
69	VHF IF system setup see note 1	2005-05-21
70	VHF Peak power read from wave guide (W)	2006-10-30
71	VHF RF duty cycle read from wave guide	2006-10-30
72	VHF Power status on Tromsø systems	2006-11-01
73	VHF CHI attenuator settings (dB)	2008-04-21
74	VHF CHII attenuator settings (dB)	2008-04-21
75	VHF Average power read in wave guide(kW)	2008-12-11

Notes

- Antenna ID parameter 41:

1. 32m ESR
2. 42m ESR
3. VHF
4. UHF
5. Kiruna
6. Sodankylä
- 7.
8. 32p ESR

- User parameters 43–62

These are set by calling the UPAR command in ELAN files. Some experiments make use of these parameters in an almost standardized way — the following are used by **beata**:

- 43 (upar 1): ESR filter width

47–49 (upar 5–7) plasma line frequency offsets

57 (upar 15): experiment version

- ESR parameter 67
 - 0 SPEAR all tx off 1 SPEAR low power radar 2 SPEAR high power radar 3 SPEAR heating
- ESR parameter 68
 - 0 Lower plasma line LO1 set to 492 MHz and upper plasma line LO1 to 502 MHz
 - 1 Lower plasma line LO1 set to 496 MHz and upper plasma line LO1 to 502 MHz
 - 2 Lower plasma line LO1 set to 492 MHz and upper plasma line LO1 to 506 MHz
 - 3 Lower plasma line LO1 set to 496 MHz and upper plasma line LO1 to 506 MHz
- UHF parameter 67 (not available at UHF remotes) and VHF parameter 72 give the power status on the Tromsø systems as follow:
 - Bit 0 UHF RF on
 - Bit 1 UHF HV on
 - Bit 2 UHF power on
 - Bit 3 VHF RF on
 - Bit 4 VHF HV on
 - Bit 5 VHF power on
 - Bit 6 Heating RF on
 - Bit 7 Heating power on
- VHF parameter 69 is decoded as follows:
 - bit 0:1 Antenna phasing
 - 0 allB
 - 2 allA
 - 3 split
 - bit 2:5 IF setup
 - bit 2 lo1/chI 0 298 MHz 1 290 MHz
 - bit 3 lo1/chII 0 298 MHz 1 290 MHz
 - bit 4 lo2/chI 0 84 MHz 1 78 MHz
 - bit 5 lo2/chII 0 84 MHz 1 78 MHz

7 Appendix: Documentation of EROS experiment configurations

Links

- https://www.eiscat.se/groups/Documentation/UserGuides/usersguide/annot_experiment.html
- <http://sgo.fi/~jussi/eiscat/erosdoc/index.html>
- http://sgo.fi/~jussi/eiscat/Eros_talk_050816.pdf