CRAF EMI and Spectrum Occupancy Database

ESF - CRAF secretariat

Netherlands Foundation for Research in Astronomy
P.O.Box 2, 7990 AA Dwingeloo, The Netherlands

8 April 2002

Contents

1. Introduction

Following a discussion during the 17th CRAF meeting (Madrid, 17-18 October 1994), CRAF decided to set up a CRAF database for EMI events. The reason for this was to provide Administrations and Regulatory Authorities with adequate quantitative information to evaluate the degradation of radio astronomy observations due to the interference environment. Not having such a facility was considered as one of the weaknesses of radio astronomy.

The format was adopted in the 18th CRAF meeting (Grenoble, 10-11 April 1995) and revised in the 23rd CRAF meeting (St.Petersburg, 24-25 April 1997; Spoelstra and Cohen, 1997, section 10.2). This 18th CRAF meeting, discussing the issues, noted one of the reasons to have such a database is to quantify interference to radio astronomy. Statistical completeness which is important for many cases, is not an issue for this subject. The type of interference and its effect on different observatories is important: e.g. GLONASS interference affected only observatories observing at 18 cm wavelength. Recalling problems quantitatively gives effect.

The idea behind the database format is to provide to the observatories with a compact and easy questionnaire providing an Interference Report, which electronically is sent to the CRAF clearing house.

During several meetings, CRAF also recognised the need for active spectrum monitoring. Such monitoring of spectrum occupancy is an important instrument in the discussions with the Regulatory Authorities on the one hand and scheduling of radio astronomy stations on the other hand, complementary to registering actual EMI events.

Apart from an EMI database, CRAF also maintains in a database information on spectrum occupancy collected by monitoring facilities at radio astronomy stations, such as operating at several European VLBI stations. The data formats of the EMI database and of the Spectrum Occupancy database are identical, but for practical reasons the databases are stored separately.

2. Database input file format

The file format for the CRAF database was defined by the 18th CRAF meeting and updated in the 23rd CRAF meeting on request by Nançay Observatory. This revised format is as follows:

Field	Field-Name	Format	Width (in bytes)	Format description
1	**DATE**	Character	8	yy-mm-dd yy = year – 1900 for year before 2000 and yy = year – 2000 for year after 1999. mm = number of month dd = day number in month
2	**STATION**	Character	10	up to the first 10 leading characters of the station name (first 10 leading characters of station name)

3	START	Character	5	hh:mm (UT) hh = hour mm = minute. Note that the minimum time resolution is 15 minutes.
4	END	Character	5	hh:mm (UT) hh = hour mm = minute. Note that the minimum time resolution is 15 minutes.
5	**ANTENNA**	Character	4	according to the following coding: *for EMI database*: according to coding: '[diameter in meters]m '; *for Spectrum Occupancy Database*: 'MON '. Note that the 4th character is usually blank (except for 100m telescope at Effelsberg).
6	RFIFREQ	Character	10	EMI centre frequency in MHz (accuracy 0.001 MHz) - format is ffffff.fff. Note that the minimum frequency resolution is 1 kHz.
7	**BANDWIDTH**	Character	10	EMI bandwidth in MHz (accuracy 0.001 MHz) (if known) - format is ffffff.fff. Note that the minimum frequency resolution is 1 kHz.
8	REP_INTERVAL	Character	4	Repetition interval of pulses (for radars): seconds
	NOTE: REP_INTERVAL = -1.0 if no repetition is observed			
9	INTENSITY	Character	6	Intensity of interference (decimals are allowed): if the intensity exceeds 999999, the number 999999 should be specified.
10	INT_UNIT	Character	2	Intensity unit: KE = Kelvin, JY = Jansky ***Note***: the intensity is uncalibrated: it can only be calibrated if it is known where in the antenna pattern the interfering transmission is received and the antenna pattern is known.
11	RFI_AZ	Character	3	Azimuth of EMI source in degrees (if available); 'AAA' if azimuth is not defined.
12	RFI_EL	Character	2	Elevation of EMI source in degrees (if available); 'EE' if elevation is not defined.
13	**TYPE**	Character	2	kind of observation: BR = broadband SP = spectral
14	**ANT_AZ**	Character	3	Azimuth of observation (in degrees); 'AAA' if azimuth is not defined
15	**ANT_EL**	Character	2	Elevation of observation (in degrees); 'EE' if elevation is not defined.
16	**DEG**	Character	3	degree of degradation in percent. For spectrum occupancy data: '000'.
20	EOR	bytes	1	specify '=' to identify end-of-record
** Total **			80 bytes	

The input to the database file is a simple ASCII file with a record-length of 80 characters, which is written into the database itself using a dedicated conversion program. Other formats than ASCII are not accepted.

The fields in this format form the EMI-questionnaire to be filled in by the radio observatory (e.g. by an operator checking the observation). Experience at Westerbork Radio Observatory showed that the handling of the questionnaire can be automated for the fields given in boldface characters. With a simple tool the other field can easily be filled in.

This format applies to both the CRAF EMI database and to the Spectrum Occupancy database.

On the basis of practice with Westerbork database information it is considered useful a radio observatory sends each month the ASCII file by email to the CRAF clearing house (see section 2.1). This process can also easily be automated.

2.1 Transfer of data files to CRAF

The transfer of data files to CRAF is done via the ftp facility. For this, the following procedure should be used:

connect:	ftp	ftp.astron.nl
user:	ftp	
password:	*your complete eMail address*	
cd craf/incoming	*change to CRAF subdirectory*	
put <filename>	*transfer file*	
	Note that the filename should have the format: yymmddi_STAT,	
	where yy	= year – 1900 before 2000 and
	yy	= year – 2000 for the year after 1999.
	mm	= number of month
	dd	= day number in month
	i	= sequence number of data file of same station in day
	STAT	= first 4 (four) characters of the station name as given in field 2 of the data format.
ls	*to check the successful completion of the transfer*	
quit	*to disconnect*	

3. Database facilities

At present the CRAF EMI and Spectrum Occupancy databases are kept and maintained at ASTRON, Dwingeloo. To manipulate these databases, a set of tools have been developed, which may enable answers to a range of questions.

The following options can be selected:

To analyze the EMI database:

Option 1: - Interference intensity as a function of time of the day
Option 2: - Interference intensity as a function of days of the week
Option 3: -Interference intensity as a function of frequency
Option 4: - Development of interference intensity as a function of time
Option 5: - Observation degradation as a function of time of the day
Option 6: - Observation degradation as a function of days of the week
Option 7: - Observation degradation as a function of frequency
Option 8: - Development of observation degradation as a function of time
Option 9: - Interference occurrence as a function of time of the day
Option 10: - Interference occurrence as a function of days of the week
Option 11: - Interference occurrence as a function of frequency
Option 12: - Development of interference occurrence as a function of time

To analyze spectrum occupancy:

Option 13: - Signal intensity as a function of time of the day
Option 14: - Signal intensity as a function of days of the week
Option 15: - Signal intensity as a function of frequency
Option 16: - Development of signal intensity as a function of time
Option 17: - Signal occurrence as a function of time of the day
Option 18: - Signal occurrence as a function of days of the week
Option 19: - Signal occurrence as a function of frequency
Option 20: - Development of signal occurrence as a function of time

4. Software documentation

The software consists of a conversion routine to convert the incoming data to the internal format of the appropriate database and to add these to the database. For each facility, i.e. EMI analysis and Spectrum Occupancy analysis, a set of routines is available. Each set contains routines for the specific functions given in section 3 of this document. These routines are invoked via the CRAF website through a .HTML file for the specification of the input parameters.

4.1. Conversion routine

The conversion routine expects a straight ASCII file. Any input record containing a format or coding different from the format defined by CRAF is not accepted, but logged in an error log file when this record is recognized as containing possible correct but corrupted data. *Any deviation from this format must be corrected by proper editing.*

The conversion routine is invoked on through a .HTML file accessible via the CRAF website. The conversion routine asks for the kind of data to be converted, the name of the input file and the name of the output file. For the latter it will take the default database when no output file is specified. (The question for the kind of data is due to program security reasons and may be removed in future).

Only the CRAF website manager can write into the CRAF website directories. Therefore, for reasons of system security, the data conversion is performed on the CRAF ftp server.

The *input data* must reside in the directory: ***/users/ftp/craf/incoming***

The *output database file* resides in subdirectory: ***/users/ftp/craf/conv.***
If in that directory no file with the expected filename exists, the programs creates a new file. If a file with the expected filename does exist already, the converted records are added at the end of the file.

The *error log* is written in subdirectory: ***/users/ftp/craf/error***.
The file containing the error log contains all records for which an error is detected. After proper editing the error log file can be used as input file to add the corrected records to the database (note that for this operation, this error log file is put into the right directory, i.e. ***/users/ftp/craf/incoming).***

After running the conversion routine, the following UNIX command should be run on the server computer:

sort +0.0 -0.8 - o <name of output file> < name of input file>

Note: the <name of output file> and the <name of input file> may be the same.

After completion of the conversion the CRAF website manager should copy the updated database file to the appropriate directory to make it available for regular manipulation, i.e. to **~/fm/emi/** or **~/fm/spec/** for the EMI or spectrum occupancy database, respectively (see section 4.4).

4.2. EMI database software

The following table lists the routines for the different options:

Option	.HTML routine	PERL routine
Option 1: - Interference intensity as a function of time of the day	iday.htm	idday.pl
Option 2: - Interference intensity as a function of days of the week	iweek.htm	idweek.pl
Option 3: -Interference intensity as a function of frequency	ifreq.htm	idfreq.pl
Option 4: - Development of interference intensity as a function of time	itime.htm	idtime.pl
Option 5: - Observation degradation as a function of time of the day	dday.htm	idday.pl
Option 6: - Observation degradation as a function of days of the week	dweek.htm	idweek.pl
Option 7: - Observation degradation as a function of frequency	dfreq.htm	idfreq.pl
Option 8: - Development of observation degradation as a function of time	dtime.htm	idtime.pl
Option 9: - Interference occurrence as a function of time of the day	oday.htm	oday.pl
Option 10: - Interference occurrence as a function of days of the week	oweek.htm	oweek.pl
Option 11: - Interference occurrence as a function of frequency	ofreq.htm	ofreq.pl
Option 12: - Development of interference occurrence as a function of time	otime.htm	otime.pl

Via the CRAF website the desired specifications are done by the .HTML pages from which the appropriate .PERL routine is invoked.

4.3. Spectrum occupancy database software

The following table lists the routines for the different options

Option	.HTML routine	PERL routine
Option 13: - Signal intensity as a function of time of the day	siday.htm	siday.pl
Option 14: - Signal intensity as a function of days of the week	siweek.htm	siweek.pl
Option 15: - Signal intensity as a function of frequency	sifreq.htm	sifreq.pl
Option 16: - Development of signal intensity as a function of time	sitime.htm	sitime.pl
Option 17: - Signal occurrence as a function of time of the day	soday.htm	soday.pl
Option 18: - Signal occurrence as a function of days of the week	soweek.htm	soweek.pl
Option 19: - Signal occurrence as a function of frequency	sofreq.htm	sofreq.pl
Option 20: - Development of signal occurrence as a function of time	sotime.htm	sotime.pl

4.4. On the directory structure

For the EMI climatology facility the following directory structure applies:

~/fm contains the files:
- the routine: conv.htm, index.html (for the general access to the EMI climatology facility.

~/fm/emi/ contains the files:
- the routines: dday.htm, dfreq.htm, dtime.htm, dweek.htm, iday.htm, ifreq.htm, itime.htm, iweek.htm, oday.htm, ofreq.htm, otime.htm, oweek.htm, index.html
- the EMI database named: dwl.db

~/fm/spec/ contains the files:
- the routines: siday.htm, sifreq.htm, sitime.htm, siweek.htm, soday.htm, sofreq.htm, sotime.htm, soweek.htm, index.html
- the Spectrum Occupancy database named: occ.db

~/fm/bin/ contains the files:
- the routines: conv.pl, idday.pl, idfreq.pl, idtime.pl idweek.pl, oday.pl, ofreq.pl, otime.pl, oweek.pl, siday.pl, sifreq.pl, sitime.pl, siweek.pl.

4.5. Algorithms used

The software has been written in PERL language.

The routines read the data from the database, check their reliability, take default values for some parameters not specified by the user, and do some administration for proper calculations. The calculations comply with the description in the following sections explaining the different options. The output is written in .HTML format and included in a web-page which is represented on the screen of the user.

4.5.1. averages and maximum values

To calculate the average and maximum value for the options 1 to 8 and 13 to 16, the calculation uses:

$$I_{bin} = (1/n) \sum_{k=1}^{n} I_{bin,k} \quad (1)$$

and

$$M_{bin} = \max(I_k, k=1 \text{ to } n) \quad (2)$$

where

I_{bin} = average value for a specific *bin* for interference intensity, signal intensity or degradation of the observation

bin = time resolution (i.e. used for time within day analysis, for day in week analysis, for function of time analysis) or frequency resolution (used for function of frequency analysis) within the specified time or frequency range as specified by the user

$I_{bin,k}$ = k-th value for interference found for a specific *bin*, signal intensity or degradation, where k runs from 1 to n

n = number of I-values, dependent on the choice of the users and/or availability of data (for n=0 a message is given)

M_{bin} = maximum value of I within each *bin*.

4.5.2. Occurrence of interference or signal as function of time or frequency

To calculate the occurrence interference or a signal for spectrum occupancy analysis for the options 9-12 and 17-20 the calculations uses:

$$O_{bin} = (100/T) \sum_{k=1}^{n} C_{bin,k} \quad (3)$$

where

O_{bin} = occurrence of EMI or signal (in percentage) for a specific *bin* in time or frequency. The occurrence shows the relative distribution of the reported EMI or spectrum occupancy data as function of the requested variable.

$C_{bin,k}$ = counter for events for EMI or spectrum occupancy within a specific *bin*,

bin = time resolution (i.e. used for time within day analysis, for day in week analysis, for function of time analysis) or frequency resolution (used for function of frequency analysis) within the specified time or frequency range as specified by the user,

n = total number of events for EMI or spectrum occupancy for a particular *bin*,
T = total number of events in the specified time or frequency range.

An *event* is defined as the identification of EMI or signal for a particular *bin* per telescope. More identifications for the same time moment and for the same telescope are taken as one single event for a time analysis of occurrence. More identifications for the same frequency band and for the same telescope are taken as one single event for a frequency analysis of occurrence. If such identification is made for different telescopes but for the same time moment or frequency band, the number of events for the related *bin* is increased by one for each telescope for which the identification is made.

NOTE:
It should be noted that the CRAF database facility is password protected. Username and password can be requested at the CRAF clearing house.