WPEC Subgroup 30:

Quality improvement of the EXFOR database

Status report June 2009

Arjan Koning

NRG, Petten, The Netherlands

Introduction

The retrieval of experimental data from EXFOR runs the danger of becoming one of the main delaying factors in contemporary nuclear data evaluation. Therefore, SG-30 was founded, which is a subgroup that aims to establish EXFOR as an easily accessible and correct database. The ultimate objective could be formulated as an appropriate solution to the following two questions (in this order):

1. In the past decades, almost all nuclear reaction information has been put into EXFOR. Can we also get it out of EXFOR again?
2. If we can get the info out, are the data correct?

The most important spin-off would be a computational database that contains the entire EXFOR database in an easy-to-use tabular format. This will enable

- a huge increase in efficiency for the use of experimental data in data evaluation,
- easy and extensive validation of nuclear model codes,
- more feedback from users to the Data Centres to correct data.

To reach this goal, SG30 will focus on the following activities:

- attempt to translate (almost) the entire EXFOR database into computational format,
- solve the most obvious quantitative errors, using checking codes, plotting packages and comparisons with model codes,

- identify data which are stored incorrectly and attempt to harmonize the format.

In this document, the status of SG-30 after 2 years will be presented. For completeness, also the progress of the 2007-2008 period is retained in this document. Before I do that, here is the most important internet information:

- SG30 mailing list: **sg30@nea.fr**
- SG30 website at NEA: **www.nea.fr/html/science/wpec/SG30/**
- IAEA repository of entire EXFOR and computational format (XC4) databases: **www-nds.iaea.org/x4toc4-master**

Progress in 2007-2008 (copied from previous report)

On April 20, 2007, Subgroup 30 was endorsed at the WPEC meeting. After discussion, the most important change relative to the initial proposal was to remove quality flagging of individual data sets from the *official* SG30 activities. In the following month, initial emails to probe the interest in SG30 were sent. A lot of moral support for this was received. The mailing list **sg30@nea.fr** was created, which currently has about 40 members. About 150 emails have been sent to the list so far.

In June 2007, a first important step towards more EXFOR user-friendliness was taken by Viktor Zerkin (IAEA), who translated the entire EXFOR database into the so called XC4 format, a convenient extension of the C4 format that has been in use for years now. Use is made of X4toC4, a code maintained by Andrej Trkov, to perform this translation. The IAEA is now regularly providing updated versions of both the entire EXFOR master database and the computational database to SG30, so that plotting, checking and nuclear model codes of EXFOR users can frequently be re-run with a new version of the database. The databases are available at the IAEA website mentioned above. The latest version is from May 7 2008.

Out of this XC4 database, a directory structured database (see Appendix) can be created (Arjan Koning). In the process, a simple checking filter can be applied to find the most flagrant errors. A first list of such errors has been sent to the IAEA in July 2007, after which Otto Schwerer and Svetlana Dunaeva corrected all cases where this was necessary.

Various other error reports were sent to both the IAEA and SG-30, among others by Alexander Konobeyev, Stepan Mashnik, Helder Duarte, Emmeric Dupont, and Sylvie Leray. As mentioned above, some of these errors emerged for the first time, as a filtering product of the automatic checking of new computational databases. Others were encountered by users who were interested in particular data. Svetlana Dunaeva and Otto Schwerer have taken care of most of these errors in the meantime.

Emmeric Dupont (CEA) has performed various statistical checks on the entire collection of tabulated data. This includes calculation of data averages over many data sets and individual deviations from that average. This allowed various problems to be identified, mostly in the categories of dimension errors or wrong units. A set of tables with possible problems have been added as Appendix.

Nicolas Soppera and Hans Henriksson (NEA) have used the JANIS package to read in raw EXFOR data and produce C4 tabulated outputs. This is all available on the NEA SG-30 website. Recently the full EXFOR conversion score became available as well as an extensive list of (possible) problems. It is added here as an Appendix.

Alexander Konobeyev (FZK) has done large-scale model versus EXFOR comparisons with the ALICE and TALYS nuclear model codes, for thousands of reaction channels. Obvious problems on the EXFOR-side have been reported to the NRDC.

On October 10-11 2007, the first SG-30 meeting was held at IAEA (see NEA/NSC/DOC(2007)25: Summary record (final version, 21 Dec 2007) prepared by A. Mengoni and A. Koning, see the Appendix (see also the SG30 page **www.nea.fr/html/science/wpec/SG30/**). All issues mentioned above were discussed. In addition, high-energy data testing is taking place at CEA (Jean-Christophe David and Sylvie Leray), showing consistency problems for the storage of residual yields and double-differential spectra. A final issue under study is the possibility to consistently distinguish between systematical and statistical errors (Otsuka, JAEA/IAEA)

The current claim is that 91% of the EXFOR database can be translated into the computational XC4 format. Out of this, about 80% has been translated into a directory-structured (reaction-by-reaction) database by Arjan Koning. One of the Appendices gives the current content statistics of EXFOR.

Translation now includes, cross sections, differential spectra, double-differential spectra, angular distributions, number of fission neutrons, fission yields, resonance parameters, and ratios. Not all of the resulting datasets have been tested however.

Progress in 2008-2009

Various error reports were sent to both the IAEA and SG-30, among others by Helder Duarte, Emmeric Dupont, Jean-Christophe David and Marina Mikhaylyukova. Naohiko Otsuka and Svetlana Dunaeva have taken care of most of these errors.

An email discussion was initiated by the IAEA on the issue of introducing quality flags in EXFOR, at either the basic level (X4) or at the derived level (C4). No firm conclusions have been reached however.

TENDL-2008 (NRG's complete nuclear data library for all particles, energies and nuclides) was released and also announced to the SG30 mailing list. This gives the possibility to do a complete EXFOR vs. ENDF library comparison. The IAEA made it available on their web-interface, allowing visual comparison with experimental data.

A global comparison between EXFOR and TALYS default calculations has been performed by Arjan Koning for all particles, energies and target nuclides. This concerns all quantities that a/ are available in computational X4 format (XC4) b/ can be calculated with TALYS. Obviously, this gives rise to a wealth of statistical information on the global quality of TALYS (not relevant here), on individual outliers (suspicious values) on wrong reaction identifiers, and a global picture of what can and what can not easily be retrieved from EXFOR. The current X4 library contains about 139000 subentries (different experimental data sets), of which 63000 could be translated into computational XC4 format, of which 22000 subentries (which are data sets, so many more data *points*) could be automatically comapred with TALYS. At the moment, this concerns cross sections only (no spectra, angular distributions etc.). The total amount of results is so large that a systematic plan to attack all observed *possible* problems has not yet been made. The first attempt of this global comparison has been added as an appendix.

Parallel to the TALYS-EXFOR comparison, Nicolas Soppera and Hans Henriksson have created an extension to Janis to do a systematic ENDF-EXFOR comaprison. The results are available on their website. The same statistical goodness-of-fit measures ('F' factor) as for the above-mentioned TALYS-EXFOR comparison were used, and results and recommendations were sent to the IAEA.

Viktor Zerkin has made new total X4 and XC4 databases available on the IAEA website.

Plans and possibilities for the third year

SG-30 has accomplished a lot, including unplanned results, but there are several essential unfinished problems left that deserve to be attacked. An important one is the testing of secondary information (i.e. other than cross sections), in particular single- and double-differential spectra. For this, the XC4 database needs another step forward (including sorting of secondary infomation such as emission energies). Another unfinished issue is a proper distinction between total, ground-state and isomeric data (important for activation cross sections).

If this is accomplished, a full overview of the quality of EXFOR can be given, including a complete list of individual items that deserve closer inspection, and a plan of attack for the coming years to solve these problems.

Therefore, an extension of 1 year is requested.

The most error-free EXFOR database can probably be obtained by using as many different,

independent checking methods as possible: through automatic chi-2 based comparison with model codes, statistical methods, visual inspection, etc., using many different translation and database management codes. Some of the activities mentioned below have been announced by volunteers or have been suggested by others to be valuable additions to the working programme. We hope these can also be performed SG-30:

- Robin Forrest (UKAEA): general cross section input, on the basis of comparison with the EAF EXFOR-derived database, and advice on the issue of isomeric versus ground state data,
- Boris Pritychenko (BNL): Sigma web interface for testing the quality of EXFOR,
- Dorothea Wiarda (ORNL): database management in relation with EXFOR,
- David Brown (LLNL): python package x4i to read EXFOR and put it into python data structures. This route also finds EXFOR bugs,

Actual implementation of these plans would be very helpful, as well as additional independent tests not mentioned here.

As for EXFOR in computational format, a remaining issue is data sorting, to get e.g. all equal Legendre coefficients or emission energies in one data block, which seems to be the last major step for the creation of a full user-friendly database. With this, the issue of inconsistently stored reaction data, i.e. under the wrong EXFOR identifier, is of course not yet solved. However, almost all data are then at least accessible for testing. To solve the remaining format problems, the above mentioned list of different checking routes is indispensible.

Summary

SG-30 is well underway. Many milestones and deliverables foreseen in the original work plan have been met, and already now EXFOR is in a much better shape than it was before SG-30 was started. Now that almost the entire EXFOR database is available in computational format (although steps still need to be made, see above), the final year will mostly be devoted to extensive checking of the data in as many different ways as possible. The previous section lists some of the participants that could contribute to this. A continuously ongoing activity is the reporting of errors to the data centers, after which corrected versions of the entire database can be made available to SG-30.

As a final product, a significantly improved EXFOR database, and related computational databases, can be expected in June 2010. At the 2010 WPEC meeting a decision can be made about a follow-up, either as a long term subroup, a new subgroup, or a kind of long-term EXFOR update activity, or (hopefully not!) nothing at all.

Automatic test of EXFOR with TALYS: Attempt 1

Arjan koning

The objective of WPEC Subgroup 30 *Quality improvement of the EXFOR database* is to make EXFOR an easy accessible database of experimental nuclear reaction data, and to systematically check and correct EXFOR entries for errors. Significant progress has been made to make EXFOR available as a complete computational database. One obvious advantage of a computational database is that automatic comparisons with nuclear model codes can be made, to test the quality of either the data or the model code. This is exactly the topic of this document.

The objective is to compare two large databases:

- The entire EXFOR database in so called XC4 format, made available by the IAEA (Zerkin).
- A complete nuclear reaction database, created by default ("blind") TALYS calculations.

The term "complete" in the second item above means that we have stored all the output that Talys gives us, if we specify maximum output with the Talys keywords. In this first stage, we use only part of the library. For the current comparison with EXFOR we consider:

- Projectiles: gamma's, neutrons, protons, deuterons, tritons, helions and alpha's
- Targets: all stable isotopes in the Z=9-83 range
- Incident energies from thermal energies up to 200 MeV
- Cross sections: total, elastic, non-elastic, exclusive (i.e. per MT number) and residual production cross sections.

This exercise is an EXFOR test and a Talys test at the same time. This global comparison obviously does not replace a "true" evaluation for one particular isotope, which involves careful studying all experimental work, precise nuclear model fitting, etc. However, it has already been shown in many occasions that Talys provides very reasonable "blind" estimates for many reaction processes, with the thermal range and fission as notable exceptions (Talys can be used for that also, but only a "true" isotopic evaluation brings the results somewhere near the experimental data for those processes). Hence, with the exception of certain reactions, Talys should be able to give a reasonably good prediction of many reaction data, and obviously we will constantly try to extend such a prediction to as many reactions as possible. At first sight, the problem is

simple: If we know that Talys is usually within e.g. 30% of the experimental data for a certain reaction channel, alarm bells should start ringing if the deviation for such a channel is suddenly much larger. We note that large deviations may also come from bad Talys performance, even if the visual agreement on linear scale is good. For example, for threshold reactions the difference between Talys and experiment may easily be a factor of 10, close to threshold, and there may be experimental and theoretical reasons for that. In general the rule holds that the smaller the cross section, the larger the relative error. It is therefore important to watch not only the calculation/experiment (C/E) values, but also the absolute deviation. In several cases, it turns out that there are problems in EXFOR, and many of them can not so easily be detected with ways other than with a model code (which is why these problems are still in EXFOR in the first place). The problems which are easiest to detect are C/E values around 0.001 or 1000, suggesting the well known error of mixing barns and millibarns. Unfortunately, the majority of cases is more difficult to judge. This is an attempt to categorize possible problems of EXFOR (or Talys). The current comparison may also help to solve one of the largest problems of EXFOR: reaction identifiers which are assigned in wrong, inconsistent or even multiple ways. If Talys is expected to give a reasonably good prediction for a reaction and we obtain a large deviation, it may be that we are not comparing the Talys result with the correct quantity.

To discover and classify problems, we use 4 goodness-of-fit estimators. If they are all very large (or in the case of R very small), something is wrong somewhere. They are the F-factor

$$F = 10^{\sqrt{\frac{1}{N}\sum_i^N \left(\log \frac{\sigma_T^i}{\sigma_E^i}\right)^2}}, \tag{1}$$

the C/E value, also called R,

$$R = \frac{1}{N}\sum_i^N \frac{\sigma_T^i}{\sigma_E^i}, \tag{2}$$

the χ^2,

$$\chi^2 = \frac{1}{N}\sum_i^N \left(\frac{\sigma_T^i - \sigma_E^i}{\Delta\sigma_E^i}\right)^2, \tag{3}$$

and the absolute deviation

$$\Delta = \frac{1}{N}\sum_i^N |\sigma_T^i - \sigma_{exp}^i|, \tag{4}$$

where the subscript T stands for theory or Talys and E for experimental. In all cases, we average over the number of energy points, N, in each set. Hence, each EXFOR subentry (data set) that contains a cross section excitation function (or only 1 point) is described by four numbers: F, R, χ^2 and Δ, while we also keep track of all individual components F_i, R_i, χ_i^2 and Δ_i for e.g. C/E plots.

The F-factor is a kind of twisted R value. In fact, each individual component of the sum inside F is equal to R if it is larger than 1, and 1/R if it is smaller than 1. This is a helpful quantity, since averaging R values over many points may not be very meaningful if the values cross 1

at some point. A value of F=1.2 means that for the entire data set we are roughly 20% off on average. We use F as the leading indicator in our statistical study, i.e. we sort our results in order of increasing F to identify the worst cases.

In our data evaluation system, the most versatile use of experimental data is made if nuclear reactions are stored directory-wise per reaction type. For example, all experimental data sets for (n,2n) reactions on ^{54}Fe are stored in a subdirectory n/fe/054/016, inside which there are various files, one per experiment, e.g.

```
......
Bahal1984.tot
Bormann1976.tot
Cross1963.tot
Csikai1965.tot
......
```

where e.g. the Bormann1976.tot file looks as follows

```
# Target  : Z= 26 A= 54       Isomer=   Projectile=n
# Reaction: MT= 16 (n,2n   )  Isomer=
# Quantity: Cross section          Frame: L
# X4 ID: 20614002  C4: MF   3 MT  16 X4 Code: 26-FE-54(N,2N)26-FE-53,,SIG
# Author  : M.Bormann+                Year: 1976
#    E(MeV)       xs(mb)      dxs(mb)      dE(MeV)       TALYS         F       R(C/E)       Chi-2      Delta(mb)
  1.40500E+01  1.01000E+01  1.00000E-01  9.00000E-02  3.41057E+00   2.961      0.3377       4475.       6.689
  1.45500E+01  1.75000E+01  1.20000E-01  1.00000E-01  1.10903E+01   1.578      0.6337       2853.       6.410
  1.48100E+01  2.25000E+01  1.50000E-01  1.00000E-01  1.65837E+01   1.357      0.7371       1556.       5.916
  1.50800E+01  2.36000E+01  1.50000E-01  1.00000E-01  2.23218E+01   1.057      0.9458       72.61       1.278
  1.53500E+01  2.90000E+01  1.70000E-01  1.00000E-01  2.81392E+01   1.031      0.9703       25.64      0.8608
  1.56100E+01  2.95000E+01  2.00000E-01  1.10000E-01  3.37412E+01   1.144       1.144       449.7       4.241
  1.59200E+01  3.70000E+01  2.10000E-01  1.10000E-01  4.04205E+01   1.092       1.092       265.3       3.421
  1.64300E+01  4.55000E+01  2.40000E-01  1.10000E-01  4.92008E+01   1.081       1.081       237.8       3.701
  1.69400E+01  5.60000E+01  3.00000E-01  1.10000E-01  5.75702E+01   1.028       1.028       27.39       1.570
  1.75600E+01  5.62000E+01  3.10000E-01  1.10000E-01  6.44940E+01   1.148       1.148       715.8       8.294
  1.82300E+01  5.81000E+01  3.00000E-01  1.20000E-01  7.06126E+01   1.215       1.215       1740.       12.51
#M.Bormann,  H-K.Feddersen,  H.-H.Holscher,
# W.Scobel, H.Wagner
#Jour. Zeitschrift fuer Physik, Section A
#Vol.277, p.203, 1976
#-(N,2N) EXCITATION FUNCTIONS FOR FE-54, GE-70, SE-74,
#RB-85, SR-86,88, Y-89, MO-92, AND HG-204 IN THE
#NEUTRON ENERGY REGION 13-18 MEV-. (GERMAN)
```

For this particular study, we simply print the blind, default, unadjusted (i.e. be gentle, please) Talys result, and the statistical factors for each point in the table, while we keep the average values for this set in another summary table. While we make the translation from the XC4 computational database to our own directory-structured database, we do our Talys comparison, checks and statistical analyses on the fly. After about 10 minutes, the conversion is done and all checking and statistical results are available.

We show a few examples here. Fig. 1 shows the distribution of the F-values for all (n,2n) reactions that we managed to get out of the XC4 database. We logged the number of cases per F-bin, whereby we distributed the range F=1-1000 over 100 logaritmically equidistant bins. Hence, the first bin means that Talys is inside 7% of the experimental data, the second bin

between 7-15%, the third bin 15-23%, and so on. All cases with F>1000 are put in the last bin. The high peak at the lowest bins probably means good news for Talys. The cases with very high F-values probably mean trouble for EXFOR (or XC4). The cases in between mean trouble for either Talys or Exfor, or both. For SG-30, the interesting cases are in the tail of the distribution and it is probably best to start checking and working on the highest values. Note that there is always the possibility of an erroneous XC4 interpretation from my side, leading to false alarm, and hopefully this improves over time. These distributions are available for all MT numbers, and also for residual production cross sections. Fig. 1 also shows the result for all neutron-induced reactions.

The most interesting are of course the specific cases with large deviation. The conversion process produces a file called *x4.sort* which has all EXFOR entries sorted by F-value. The bottom of that list looks as follows

EXFOR	C4 New	#C4	#New	Reaction	F	R	chi-2	Delta
.....								
11491007	1 1	1	1	25-MN-55(N,HE3)23-V-53,,SIG	1.080E+15	9.262E-16	0.00	0.420
11590082	1 1	1	1	81-TL-205(N,HE3)79-AU-203,,SIG	1.448E+15	6.908E-16	0.00	7.000E-02
30106012	1 1	1	1	57-LA-139(N,2P)55-CS-138,,SIG	1.523E+15	6.568E-16	2.25	3.000E-02
11491015	1 1	1	1	40-ZR-94(N,HE3)38-SR-92,,SIG	4.892E+15	2.044E-16	0.00	0.400
30152014	1 1	1	1	52-TE-130(N,HE3)50-SN-128,,SIG	5.786E+15	1.728E-16	3.52	1.500E-02

i.e. we find F-values above 10^{15}, and they even are probably not EXFOR errors! (Talys is known to miss the (n,He3) threshold by at least 1 MeV, resulting in a huge deviation near threshold. Note the small deviation in millibarns in the last column). More interesting is perhaps to filter out one kind of reaction. If under Linux, we do *grep 'N,2N' x4.sort* the bottom of that list looks as follows.

EXFOR	C4 New	#C4	#New	Reaction	F	R	chi-2	Delta
.....								
303220122	1 1	1	1	49-IN-115(N,2N)49-IN-114-M1,,SIG	107.	107.	3.368E+05	1.161E+03
22662014	1 1	6	6	58-CE-140(N,2N)58-CE-139-M,,SIG	140.	141.	1.175E+07	793.
23018004	1 1	6	6	80-HG-196(N,2N)80-HG-195-G,,SIG	638.	642.	2.852E+07	582.
23018002	1 1	6	6	80-HG-196(N,2N)80-HG-195,,SIG	702.	703.	7.502E+07	1.578E+03
23018003	1 1	6	6	80-HG-196(N,2N)80-HG-195-M,,SIG	748.	754.	6.158E+07	996.
11684003	1 1	1	1	25-MN-55(N,2N)25-MN-54,,SIG	786.	786.	0.00	785.
23018007	1 1	6	6	80-HG-198(N,2N)80-HG-197-G,,SIG	880.	881.	8.661E+07	1.007E+03
31454005	1 1	1	1	72-HF-179(N,2N)72-HF-178-M2,,SIG	908.	908.	1.039E+07	290.
20033003	1 1	1	1	37-RB-85(N,2N)37-RB-84,,SIG	960.	1.042E-03	399.	1.010E+06
22969007	1 1	4	4	72-HF-176(N,2N)72-HF-175,,SIG	961.	975.	1.119E+08	1.194E+03
23018014	1 1	6	6	80-HG-204(N,2N)80-HG-203,,SIG	993.	994.	4.301E+08	2.135E+03
23018005	1 1	6	6	80-HG-198(N,2N)80-HG-197,,SIG	994.	995.	2.696E+08	2.075E+03
23018006	1 1	6	6	80-HG-198(N,2N)80-HG-197-M,,SIG	1.139E+03	1.140E+03	2.872E+08	1.069E+03
22347006	1 1	3	3	53-I-129(N,2N)53-I-128,,SIG	1.450E+03	1.454E+03	1.164E+08	1.640E+03

Here we see the F-values around 1000 suggesting the b-mb problem. Indeed, the (n,2n) cross sections of the various Hg isotopes were a factor of 1000 too small and this has recently been corrected. Note that the R-factor indicates that 37-RB-85(N,2N) concerns kilobarns instead of barns (hence the deviation of 1.010E+06 mb). Unfortunately, not all problems are that simple. There may be an inconsistency between Talys and Exfor on the definition of an isomer, leading to large discrepancies and the list above may also suggest that isomeric, ground state and total cross sections may have been mixed.

After *grep 'N,TOT' x4.sort* we find near the bottom

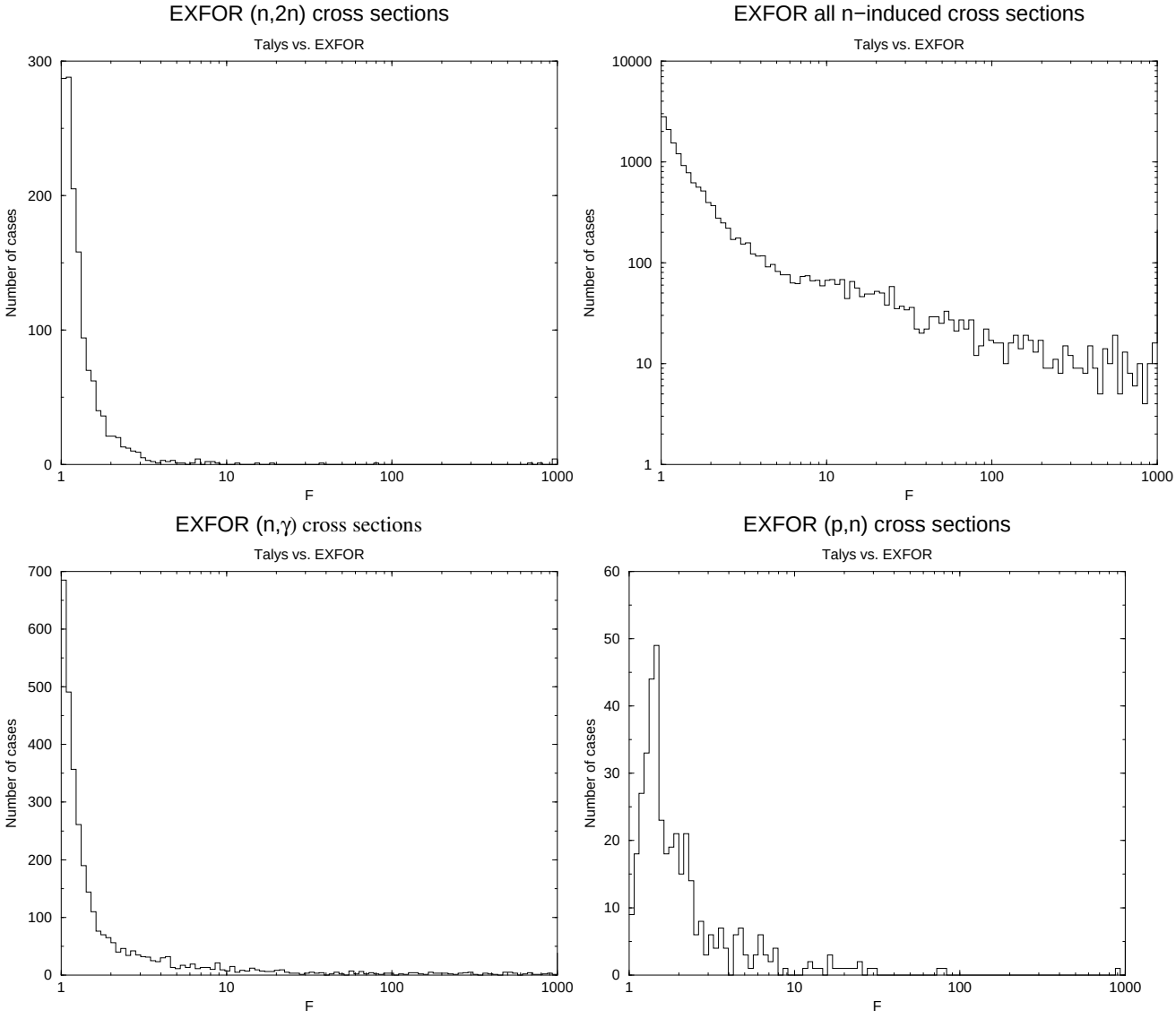

Figure 1: Frequency table for the F-values for (n,2n) reactions, all neutron-induced reactions, (n,γ), and (p,n) reactions.

```
  EXFOR    C4 New  #C4  #New     Reaction                                F          R        chi-2     Delta
.......
 21676005   1  1     1     1 23-V-51(N,TOT),,SIG,,AV                   914.       914.      1.442E+10  4.924E+06
 11026015   1  1     1     1 23-V-0(N,TOT),,SIG                        920.       920.      4.948E+09  4.924E+06
 22007002   1  1   121   121 14-SI-0(N,TOT),,SIG                       937.       983.       0.00      5.151E+05
 31468004   1  1     7     7 24-CR-0(N,TOT),,SIG                       972.       973.      9.116E+08  3.790E+03
 20602009   1  1     1     1 23-V-51(N,TOT),,SIG                       974.       974.      5.870E+10  4.846E+06
 31468006   1  1     6     6 28-NI-0(N,TOT),,SIG                     1.010E+03  1.010E+03  7.236E+08  3.603E+03
 31468005   1  1     6     6 26-FE-0(N,TOT),,SIG                     1.013E+03  1.014E+03  8.609E+08  3.642E+03
 31468007   1  1     2     2 29-CU-0(N,TOT),,SIG                     1.031E+03  1.031E+03  6.671E+08  3.442E+03
 31468003   1  1     7     7 23-V-0(N,TOT),,SIG                      1.039E+03  1.040E+03  9.202E+08  3.987E+03
 11204004   1  1    12    12 4-BE-9(N,TOT),,SIG                      1.045E+03  1.058E+03   0.00      1.528E+06
.......
```

of which some (maybe all?) are indeed mb-b problems. Note however that at very low energies Talys may be a factor 1000 off (while above a few MeV Talys should be within 1-3 %).

In the future, this procedure should also be applied to secondary distributions such as angular distributions and (double-)differential spectra. Before this can be done, the secondary energies and angles in the XC4 database first need to be sorted (with the sortC4 code).

As stated above, the first obvious use of all this information is to identify problems and to correct them. However, with all these results available, it is now also possible to set-up some ”zeroth-order” quality flagging. Although we can never be 100% sure, it is very probable that the subentries with small F-values (where ”small” depends on the reaction channel) represent indeed the type of quantities that are reported in EXFOR. In other words, the reaction identifier assigned by the compiler for these subentries is correct.

Note that here I am not yet talking about renormalization of wrongly interpreted experimental conditions (there has been some recent discussion about this and that may be very valuable too). The quality flagging mentioned above would simply ensure that e.g. a cross section reported as being (n,2n) in EXFOR is indeed just that, and not e.g. a compound (n,n') + (n,2n) cross section (which would place it further down the tail of our F-distribution since Talys is being compared with the wrong quantity). In this way we would obtain a large ”verified” set of EXFOR data, while ”validation” of the data would involve a more precise study of the detailed experiment and possible renormalization. To start with, such a quality list could exist of only the EXFOR subentry number and e.g. a 1 for a verified set and a 0 for a non-verified set. This requires some further study and discussion.

Finally, all experimental uncertainties of EXFOR were also sorted per reaction or MT number and stored in files, e.g. *mterr.102*. The largest errors for capture are obtained with *sort -g -k 3,3 mterr.016*:

```
.............
 10340007  52-TE-130(N,G)52-TE-131,,SIG              354.     %
 11507090  74-W-180(N,G)74-W-181,,SIG,,MXW           400.     %
 13742004  38-SR-88(N,G)38-SR-89,,SIG,,RAW           619.     %
 400070051 34-SE-80(N,G)34-SE-81,,SIG                749.     %
 40975014  68-ER-170(N,G)68-ER-171,,SIG             1.538E+03 %
```

which may or may not be wrong, closer inspection is required. The largest errors in the entire list are

```
.............
 20668004  74-W-180(N,P)73-TA-180-M,,SIG             960.     %
 21264010  42-MO-92(N,N+P)41-NB-91,,SIG           1.344E+03 %
 22640003  8-O-18(N,P)7-N-18,,SIG                 1.478E+03 %
 40975014  68-ER-170(N,G)68-ER-171,,SIG           1.538E+03 %
 21288007  20-CA-48(N,2N)20-CA-47,,SIG            5.565E+03 %
 20909003  6-C-12(N,TOT),,SIG                     1.629E+04 %
 21288005  20-CA-44(N,A)18-AR-41,,SIG             1.714E+04 %
 41163008  52-TE-126(N,P)51-SB-126-G,,SIG         2.000E+04 %
 C0774002  39-Y-89(P,N)40-ZR-89,,SIG              7.334E+04 %
 12940004  40-ZR-0(N,TOT),,SIG                   9.071E+04 %
```

Only results for incident neutrons were given here, but the same is done for other projectiles. More refined test are possible and I hope to find the time to perform them in the near future. All results described in this document, i.e. all resulting numerical files, have been sent to the IAEA for further inspection.