EXFORTABLES-2.1

An experimental nuclear reaction database based on EXFOR

Arjan Koning

s

nds.iaea.org/talys

Typesetting: The Legrand Orange Book, LaTeX Template, Version 2.1.1 (14/2/16), downloaded from: www.LaTeXTemplates.com. Original author: Mathias Legrand (legrand.mathias@gmail.com) with modifications by: Vel (vel@latextemplates.com). License: CC BY-NC-SA 3.0 creativecommons.org/licenses/by-nc-sa/3.0/.

December 2025

About the author

Arjan Koning is a nuclear physicist with a Masters Degree in Theoretical physics at the Univ. of Amsterdam, a PhD in the Natural Sciences on Multi-step direct reactions at the Univ. of Groningen, and a Professorship at the Univ. of Uppsala on theoretical nuclear reactions.
Arjan is currently Head of the Nuclear Data Section at the IAEA in Vienna. Before that, he has worked at ECN/NRG Petten, the Netherlands, on nuclear reaction data for science and technology, and as guest scientist at CEA/Bruyères-le-Châtel and Los Alamos National Laboratory on the development and computational implementation of nuclear reaction models. He has led several students to PhD degrees, has coordinated and chaired various international nuclear data projects such as the OECD/NEA JEFF and WPEC projects, and has advised governments and international organisations on nuclear research and development. Among his scientific accomplishments are innovations in nuclear reaction physics, especially for the optical model and pre-equilibrium reactions, and initiating the '3 T's': the TALYS nuclear model code, Total Monte Carlo uncertainty propagation and the TENDL nuclear data library. As of 2025, his h-index is 62 with more than 26 000 citations.
Although Arjan is currently in a managerial role, he aims to keep his scientific creativity alive by maintaining and extending TALYS plus all software that emerges from that. Pleas from his friends to also spend time on other things are sometimes honoured.

Preface

EXFORTABLES is a directory-structured database with experimental nuclear reaction data. It is entirely based on the international experimental nuclear reaction database EXFOR but, unlike EXFOR, contains the data in an easy accessible *projectile/element/mass/reaction* directory structure. A code *exfortables.f90* has been written which reads in the entire EXFOR database and produces EXFORTABLES. During this database translation, statistical tests are performed on almost the entire EXFOR database and the results of these tests are written to various diagnostic files, which can be taken into account by the Nuclear Reaction Data Centres network to correct EXFOR.

The idea to make EXFORTABLES was born in 2006, when I argued that in these times of strongly increasing computer power, the mining of experimental data from EXFOR, and assessing their correctness, is rapidly becoming one of the main delaying factors in data evaluation. Fortunately, several others had the same opinion and soon the NEA WPEC Subgroup 30 on the "Quality improvement of the EXFOR database" took off to address this issue. The current database is one of the results of that.

At certain moments in time, a well-defined version of EXFORTABLES is frozen. You are now reading the tutorial of version 1.0. Until 2020, EXFORTABLES was called NEWBASE.

For some reason, it took me 10 years to release this software and tutorial, so meanwhile other initiatives and visions on the availability of EXFOR are starting to emerge. EXFORTABLES translates the entire EXFOR database at once while it may be more flexible to have user-defined EXFOR retrievals via a dedicated API. Until that has been accomplished, there is EXFORTABLES, and later the results from the two approaches can always be compared.

License, contact and reference

As mentioned on the first page and in the source code, EXFORTABLES falls in the category of MIT License software.

In addition to the MIT *terms* I have a *request*:

- When EXFORTABLES is used for your reports, publications, etc., please make a proper reference to the code. At the moment this is:
  When you refer to the application of this software:
  A.J. Koning, Bayesian Monte Carlo method for nuclear data evaluation, Eur. Phys. Journ. A51(12) 1 (2015).
  When you refer to something particular of this tutorial:
  A.J. Koning, EXFORTABLES-1.0: An experimental nuclear reaction database based on EXFOR, IAEA NDS Document Series IAEA(NDS)-235, December 2020

The webpage for EXFORTABLES is **nds.iaea.org/talys**.

Acknowledgements

I wish to thank a few persons who have contributed to the present state of EXFORTABLES:

- Viktor Zerkin, for his ongoing effort to translate the original EXFOR (X4) database into a computational format database XC4, and later C5.
- Present and past IAEA staff responsible for EXFOR, Naohiko Otsuka, Otto Schwerer, Svetlana Dunaeva, made an essential contribution by correcting the errors that are revealed by projects like this.
- Dimitri Rochman helped with the initial "total database retrievals" when I started this in 2006.
- The members of the NEA WPEC Subgroup 30 on *Quality improvement of the EXFOR database* are acknowledged for their helpful feedback.
- Amanda Lewis and Denise Neudecker for starting WPEC Subgroup 50 on an *Automatically readable, comprehensive, curated experimental reaction database*, which made me realize to (finally) release my database and software.
- NEA Data Bank, in particular Hans Henriksson, Nicolas Soppera and Emmeric Dupont have given important advice on how to proceed with the EXFOR checking process, and for funding an important part of this work.

Arjan Koning

Contents

1. Introduction

EXFOR[1] is by far the most important and complete experimental nuclear reaction database in the world. Over the past 50 years, experimental data have been added to EXFOR by various compilers who used different (and allowed!) formatting rules to store the data. The result is a database that contains, certainly for neutrons but also for other incident particles, the numerical data of almost the entire history of nuclear reaction measurements. The main question is now whether and how the user has access to all these data. What has been put in can not always easily be taken out. For the study of a few detailed reactions, web interfaces are helpful tools to retrieve a few data sets, but a genuine step forward in the production of nuclear data libraries for applications requires that *all* nuclear data that exist in EXFOR can be retrieved in an unambiguous manner, and this is lacking at the moment. Moreover, EXFOR is known to contain various errors which in the past maybe have not been reported back sufficiently by the user community to the Nuclear Reaction Data Centers (NRDC).

The fast increase in computer power and the automation of nuclear model codes, cross section plotting software, etc. have made the easy accessibility of experimental data more important than it was in the past. In fact, the retrieval of experimental data is now becoming a delaying factor in contemporary data evaluation.

For this reason, EXFORTABLES was created. EXFORTABLES is a directory-structured database, derived from the experimental nuclear reaction database EXFOR. It is a follow-up database from an earlier version called NEWBASE which was the result of WPEC Subgroup 30 on the Quality improvement of the EXFOR database [2]. There are two objectives for EXFORTABLES:

- To have all experimental nuclear reaction data readily available in logically ordered directories, with one file per experimental data set, to enable quicker and more flexible use for both humans and software.
- To automatically test all nuclear reaction data of the EXFOR database against global nuclear model calculations, in order to reveal errors in the database and to test those nuclear models at the same time.

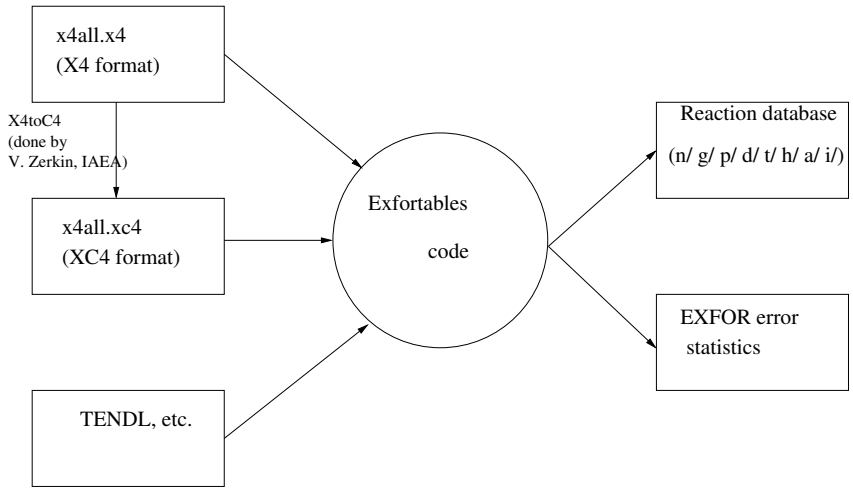

Figure 1.1: Databases produced out of EXFOR.

This is accomplished by the Fortran code *exfortables.f90*, which reads in three groups of large databases:

- the entire "mother" EXFOR database in X4 format, here used as one file *x4all.x4*
- the entire EXFOR database in extended computational XC4 format, here used as one file *x4all.xc4*
- The world nuclear data libraries TENDL-2023 [3], ENDF/B-VIII [4], JEFF-3.3 [5], JENDL-5.0 [6], CENDL-3.2 [7], IRDFF-2.0 [8], EAF-2010 [9], for all projectiles, target nuclides and energies.

The *exfortables.f90* code processes all these data, after which the EXFORTABLES database is produced in directories n/ g/ p/ d/ t/ h/ a/ and i/ (heavy ions), for the various projectiles.

In addition, a directory *stat/* is produced which contains all kinds on statistics of the EXFOR database. For example, there is an automatic C/E or χ^2 comparison with theoretical or evaluated data values for all cases where experiment could be compared with the world nuclear data libraries. Suspicious data are ordered in several output files.

The whole suite of databases we use to produce the experimental nuclear reaction database is presented in Fig. 1.1 as a flowchart. The central message is that we provide an experimental nuclear data library that is much easier to access, plot, and use in nuclear data evaluation. In the process, its quality is tested by large-scale comparison with TALYS, TENDL, and other world nuclear data libraries and statistical results on that quality becomes available. One thing is certain: The EXFOR database is so large and complex, that it is impossible to accomplish an entirely correct translation into a new database. Errors may pop up at the most unexpected places, and such errors will have to be corrected in future versions of EXFOR, XC4 or the translation code *exfortables.f90*. For EXFORTABLES-1.0, we may state it is particularly complete for cross sections, while for other data such as secondary angle or spectrum distributions, ratios etc. more development is required.

1.1 This tutorial

After this Introduction, you will find the following,

Chapter 2: Installation guide for EXFORTABLES, containing a general outline of the contents of the EXFORTABLES database, and also an explanation on how to (optionally) regenerate the database yourself.

Chapter 3: The experimental nuclear database. This Chapter contains a detailed description on how nuclear reaction data are stored in a structured way.

Chapter 4: Testing the EXFOR database. This describes the huge amount of statistical information, including errors and suspicious data, that is available after this analysis.
Chapter 5: Quality scoring of the EXFOR data sets.
Chapter 6: Input description, which is only needed if you want to reproduce EXFORTABLES yourself. In that case you need to run the *exfortables.f90* code.
Chapter 7: The reference guide with all input options.
Chapter 8: Outlook and conclusions.

2. Installation and getting started

2.1 The EXFORTABLES package

In what follows we assume EXFORTABLES will be installed on a Linux or MacOS operating system. In total, you will need about 3 Gb of free disk space to install EXFORTABLES (this relatively large size is due to the EXFOR and XC4 files). If you obtain the entire EXFORTABLES package as a tar file, you should do

- **tar zxf exfortables.tar**

and the total EXFORTABLES package will be automatically stored in the *exfortables/*

EXFORtables contains the following directories and files:

- *LICENSE* is the license file,
- *README* outlines the contents of the package,
- *install_exfortables.bash*, *code_build.bash* and *path_change.bash* are scripts that take care of the installation,
- *source/* is the Fortran source of the *exfortables* code that performs the entire database creation and checking of the data,
- *files/*: the entire EXFOR database in X4 (*X4all*) and C5 (*C5reac*) format, respectively, and other info such as abundance and discrete levels, original quality scoring etc,
- *n/ p/ d/ t/ h/ a/ g/ i/* are directories with the entire directory-structured database, produced by exfortables.f90,
- *doc/* contains the documentation: this tutorial in postscript and pdf format,
- *input/* contains the input file for *exfortables.f90* used to create the database delivered with this package,
- *quality* contains the quality scores as outline in Ref [10],
- *stat/* contains files with statistical (checking) information on all processed EXFOR entries,
- *special/* contains special data files such as thermal cross sections, MACS, etc.

Most users may now directly proceed to the next chapters in which the contents of the database are explained. **Thus, you do not need to further install or run anything and can use the database**

directly. However, if you want to regenerate the database yourself with the *exfortables.f90* code, continue with the following section.

2.2 Installation

The installation of EXFORtables is straightforward. You can download EXFORtables via either git
- **git clone https://github.com/arjankoning1/exfortables.git**

or by getting the tar file
- from **https://nds.iaea.org/talys/exfortables.tar**
- **tar zxf exfortables.tar**

Although most users will only be interested in the final EXFORtables database, there may be reasons to regenerate the entire database from scratch again, such as:
- your local changes in *exfortables.f90* which improve or extend the database, (of course I advise to share that with the author),
- newer versions of the mother databases,
- etc.

We here provide the necessary steps to do the installation, For a Linux or MacOs system, the installation is expected to be handled by the *install_exfortables.bash* script, as follows
- if necessary, edit *code_build.bash* and set the first two variables: the name of your compiler and its flags.
- **install_exfortables.bash**

An alternative installation option is
- **cd exfortables/source**
- **make**

If this does not work for some reason, we here provide the necessary steps to do the installation manually. For a Linux or MacOS system, the following steps should be taken:
- **cd exfortables/source**
- Ensure that EXFORtables can read directories with data. This is done in subroutine *machine.f90*. If *code_build.bash* has not already replaced the path name in *machine.f90*, do it yourself. We think this is the only machine dependence of EXFORtables. We expect no complaints from the compiler.
- **gfortran -c *.f90**
- **gfortran *.o -o exfortables**
- **mv exfortables ../bin**

If you run the EXFORtables code, it will overwrite all existing directories. The command for this is

exfortables < exfortables.inp

3. The experimental nuclear database

Until a general EXFOR-API comes along, we think that the most versatile use of experimental data for direct use is made if nuclear reaction data are stored directory-wise per reaction type, with a logical filenaming convention to make everything machine-readable. Then all experimental data is directly available to a computer. Therefore, the directory structure of EXFORTABLES is of the type *projectile/element/mass/reaction*. On the highest level, EXFORTABLES consists of directories *n/ p/ d/ t/ h/ a/ g/ i/*, for neutron-, proton-, deuteron-, triton-, helion-, alpha-, photon- and heavy ion-induced reactions, respectively. These directories contain all experimental nuclear reaction data which **so far** could be processed from the EXFOR database by *exfortables*.

The next level contains the isotope, e.g. *Fe054/*, for which experimental data exists. They are given in i3.3 format, e.g. *n/Fe054/* while data for natural elements are stored under the *Fe000/* directory. One level deeper are the reaction types. In this way, it is directly visible, by the number of data files, how often a certain reaction has been measured, and also particular nuclear reaction data can be found very fast with such a directory-structure. We will discuss the various classes of nuclear reaction quantities below.

3.1 Cross sections

First of all, cross section data are stored under the corresponding MT number as defined by the ENDF-6 format [11]. Although this may be a mysterious quantity for a "pure" nuclear physicist, it is helpful for nuclear data evaluation for applications to store data in this way. To serve all communities, the entire correspondence table of nuclear reaction channels and MT numbers is given in Table 3.1. Also in each subdirectory, there is a file, e.g. *n/Fe056/xs/n-Fe056.list* which gives the correspondence between an MT number and a reaction string.

As an example, all experimental cross section sets for ^{89}Y(n,p) reactions are stored in a subdirectory *n/Y089/xs/103/*. For cross sections to the ground state or an isomer, the MT number is extended by a 'g', 'm' or 'n'. Inside such a directory there are various files, one per experiment, for which the first authors name, the MT number, the subentry number and the year of publication are

used to construct the filename. In the case of ^{89}Y(n,p) we find the following files in *n/Y089/xs/103/*

```
n-Y089-MT103-Bayhurst-11462009.1961
n-Y089-MT103-Csikai-30115008.1967
n-Y089-MT103-Klopries-31532004.1997
n-Y089-MT103-Levkovskii-402230201.1969
n-Y089-MT103-LuoJunhuaa-32729004.2016
n-Y089-MT103-Molla-31494005.1998
n-Y089-MT103-Tewes-11504007.1960
```

To discuss the contents of such files, let us zoom in on one particular file, e.g. *n-Y089-MT103-Molla-31494005.1998* (to let the table fit on this page, we have cut it at column 80. Consult the database itself for the full table),

```
# header:
#   title: Y89(n,p)Sr89 cross section
#   source: EXFOR
#   user: Arjan Koning
#   date: 2023-12-17
#   format: YANDF-0.1
# exfor:
#   author: Molla
#   year: 1998
#   subentry: 31494005
#   X4 reaction: 39-Y-89(N,P)38-SR-89,,SIG
#   X4 source: IAEA-NDS C5 file, database version 2023-08-08
#   X4 link: https://nds.iaea.org/EXFOR/31494005
# target:
#   Z: 39
#   A: 89
#   nuclide: Y89
# reaction:
#   type: (n,p)
#   ENDF_MF: 3
#   ENDF_MT: 103
# residual:
#   Z: 38
#   A: 89
#   nuclide: Sr89
# datablock:
#   quantity: cross section
#   columns: 5
#   entries: 3
##       E             dE             xs             dxs       Normalization
##     [MeV]          [MeV]          [mb]           [mb]            []
   1.441000E+01   1.800000E-01   2.299853E+01   5.999600E+00   9.999360E-01
   1.463000E+01   1.600000E-01   1.999872E+01   4.999680E+00   9.999360E-01
   1.471000E+01   1.200000E-01   1.899878E+01   5.999600E+00   9.999360E-01
# reference:
#   author: N.I.MOLLA,S.BASUNIA,R.U.MIAH,S.M.HOSSAIN,M.RAHMAN, S.SPELLERBERG,S.M.QAIM
#   title: Radiochemical study of the Sc-45(n,p)Ac-45 and Y-89(n,p)Sr-89 reactions in the neutron
#   journal: Jour: Radiochimica Acta, Vol.80, p.189 (1998)
```

In addition, there is a file with the same name in *stat/n/Y089/xs/103/* which has in addition to the above info also statistical information, in particular how much it deviates fro nuclear data libraries and other measurements,

```
\begin{verbatim}
# statistics:
#   NEA score: T1
#   IAEA score:  1
#   E-min [MeV]:  1.441000E+01
#   E-max [MeV]:  1.471000E+01
#   F-value:  1.083605E+00
#   A-value:  1.066099E+00
#   chi-2:  3.218843E-01
#   p-value:  1.993185E-01
#   exp. data sets for p-value:      4
#   libraries:      5
##  Library          F              A            chi-2
# tendl.2021    1.065895E+00   1.050905E+00   2.471817E-01
# endfb8.0      1.090777E+00   1.073105E+00   3.572899E-01
# jeff3.3       1.090777E+00   1.073105E+00   3.572899E-01
# jendl5.0      1.078770E+00   1.061952E+00   2.986469E-01
# cendl3.2      1.089284E+00   1.071605E+00   3.490129E-01
```

All the output is in so-called YANDF format, which is explained in the appendix. Note that a lot of reaction information has been adopted from the original EXFOR database. All reaction identifiers, such as the target, projectile, etc, are at well-defined locations in our files for further automatic processing. Of course, the original EXFOR reaction identifier, the so-called ID-number, is also given. This is important for further judgement or treatment of the data, such as looking up precise experimental details in the original EXFOR database, quality flagging outside the EXFOR database, etc. The "X4 reaction" field contains the reaction string as given in EXFOR, so that we can check whether indeed the correct translation was carried out for this particular file. The complete bibliographic information from the original EXFOR database has been adopted and is given at the end of the file. All such non-numerical information is preceded by a '#', a character which is disregarded by most plotting packages. The only non-commented data are the 4 columns with the actual experimental data from EXFOR. Since most plotting packages, and also other codes, expect data in the x-y-dy form, we have used that format too. However, sometimes there is also an uncertainty on the incident energy available, so we have put dE in the 4th column.

Below the actual data we have added bibliographic information and statistics regarding the comparison with the world nuclear data libraries. Here, 'world' is an average over all libraries. The statistical quantities are explained in the next Chapter. This can be disregarded if one is only interested in the experimental data. It is added to reveal any possible problems with either those data libraries or EXFOR.

Some caution should be used for inelastic scattering to discrete levels, which are stored in MT51-90. Since EXFOR gives only the excitation energy, and not the level number, we have to use the TALYS (RIPL) discrete level database to estimate the number of the particular discrete level. This may not always give the correct answer, so it is possible that such data may end up in the wrong MT directory. A similar uncertainty holds for discrete level (n,p), etc. reactions that are stored in MT600-840.

Note that MT numbers 201-207 are used for total particle production cross sections, i.e. MT207 contains the $(n,x\alpha)$ cross section. At incident energies below about 20 MeV these data are equal to

MT107 ((n,α)). Hence, to compare calculated results with e.g. all (n,α) cross sections one may take, besides MT107, the low energy part of data from MT207 into account as well.

For the current version of EXFORTABLES, we are mostly interested in the actual data points, plus the leading metadata. In the future, we may include all original EXFOR information as well.

3.2 Residual production cross sections

Residual production cross sections are stored in the *residual/* subdirectory. Inside *residual/*, data files per residual product are stored in directories *ZZZAAA* with ZZZ the charge number and AAA the mass number of the product. Hence, in e.g. *p/Y089/residual/039088/* the various filenames for the reaction ^{89}Y(p,x)^{88}Y can be found. The contents have basically the same shape as that of the cross sections of section 3.1. As an example, the contents of *p/Y089/residual/039088/p-Y089-rp039088-Tarkanyi-D41670042.2004* are

```
# header:
#   title: Y89(p,x)Y88 cross section
#   source: EXFOR
#   user: Arjan Koning
#   date: 2023-12-17
#   format: YANDF-0.1
# exfor:
#   author: Tarkanyi
#   year: 2004
#   subentry: D41670042
#   X4 reaction: 39-Y-89(P,X)39-Y-88,,SIG
#   X4 source: IAEA-NDS C5 file, database version 2023-08-08
#   X4 link: https://nds.iaea.org/EXFOR/D41670042
# target:
#   Z: 39
#   A: 89
#   nuclide: Y89
# reaction:
#   type: (p,x)
#   ENDF_MF: 3
#   ENDF_MT: 5
# residual:
#   Z: 39
#   A: 88
#   nuclide: Y88
# datablock:
#   quantity: cross section
#   columns: 5
#   entries: 16
##       E             dE             xs             dxs        Normalization
##     [MeV]          [MeV]          [mb]           [mb]             []
   1.550000E+01   0.000000E+00   8.000000E-01   7.000000E-01   1.000000E+00
   2.640000E+01   0.000000E+00   2.802000E+02   3.040000E+01   1.000000E+00
   3.440000E+01   0.000000E+00   2.724000E+02   2.950000E+01   1.000000E+00
   3.690000E+01   0.000000E+00   2.154000E+02   2.340000E+01   1.000000E+00
   4.060000E+01   0.000000E+00   2.094000E+02   2.280000E+01   1.000000E+00
   4.270000E+01   0.000000E+00   1.701000E+02   1.850000E+01   1.000000E+00
.....................
```

Also here, reactions to the ground state are in subdirectories with the extension 'g', while cross sections to an isomer in 'm'.

3.3 Angular distributions

Angular distributions are stored in the *angle/* subdirectories.

3.3.1 Elastic scattering angular distributions

Elastic scattering angular distributions are stored in the *angle/002/* subdirectory. For example, neutron elastic scattering angular distributions for ^{89}Y can be found in *n/Y089/angle/002/*. Next to the authors name, MT number, subentry and the year and the term 'ang', the incident energy in MeV is used to construct the filename. The contents of *n/Y089/angle/002/* are

```
n-Y089-MT002-Becker-11511019-ang-E0003.200.1966
n-Y089-MT002-Bostrom-11130010-ang-E0003.670.1959
n-Y089-MT002-Bostrom-11130013-ang-E0001.450.1959
n-Y089-MT002-Cox-10332012-ang-E0000.889.1972
n-Y089-MT002-Hansen-12935007-ang-E0014.600.1985
.......
```

and for example the file *n-Y089-MT002-Hansen-12935007-ang-E0014.600.1985* looks as follows

```
# header:
#   title: Y89(n,el) angular distribution at              MeV
#   source: EXFOR
#   user: Arjan Koning
#   date: 2023-12-17
#   format: YANDF-0.1
# exfor:
#   author: Hansen
#   year: 1985
#   subentry: 12935007
#   X4 reaction: 39-Y-89(N,EL)39-Y-89,,DA
#   X4 source: IAEA-NDS C5 file, database version 2023-08-08
#   X4 link: https://nds.iaea.org/EXFOR/12935007
# target:
#   Z: 39
#   A: 89
#   nuclide: Y89
# reaction:
#   type: (n,el)
#   ENDF_MF: 4
#   ENDF_MT: 2
#   E-incident [MeV]:  1.460000E+01
# datablock:
#   quantity: angular distribution
#   columns: 4
#   entries: 15
##     Angle           xs             dxs       Normalization
##     [deg]         [mb/sr]        [mb/sr]           []
   9.300012E+00   6.078300E+03   4.261000E+02   1.000000E+00
   1.688000E+01   2.783800E+03   1.954000E+02   1.000000E+00
```

```
   2.406000E+01   9.718000E+02   6.850000E+01   1.000000E+00
   3.263999E+01   1.323000E+02   9.600000E+00   1.000000E+00
   3.910000E+01   1.029000E+01   8.300000E-01   1.000000E+00
   4.677000E+01   5.427000E+00   8.280000E-01   1.000000E+00
...................
```

Note that we have not yet performed an automatic library comparison for angular distributions in this version of EXFORTABLES.

For charged-particle elastic scattering angular distributions we produce, in addition to the files mentioned above, files with differential cross sections relative to the Rutherford cross section with the extra extension *.ruth*, see e.g. *p-Y089-MT002-Bertrand-O0293002-ang-E0061.500.1969.ruth* which has the same structure as the example above.

3.3.2 Inelastic scattering and other angular distributions

Inelastic scattering angular distributions are stored in the subdirectories such as *angle/051/* (for the first inelastic level). For example, neutron inelastic scattering angular distributions to the first level of ^{56}Fe can be found in *n/Fe056/angle/051/*. For example the file *n-Fe056-MT051-Hyakutake-20690005-ang-E0014.100.1975* looks as follows

```
# header:
#   title: Fe56(n,n)Fe57 angular distribution at                MeV
#   source: EXFOR
#   user: Arjan Koning
#   date: 2023-12-17
#   format: YANDF-0.1
# exfor:
#   author: Hyakutake
#   year: 1975
#   subentry: 20690005
#   X4 reaction: 26-FE-56(N,INL)26-FE-56,PAR,DA
#   X4 source: IAEA-NDS C5 file, database version 2023-08-08
#   X4 link: https://nds.iaea.org/EXFOR/20690005
#   level energy [MeV]  8.50000E-01
# target:
#   Z: 26
#   A: 56
#   nuclide: Fe56
# reaction:
#   type: (n,n)
#   ENDF_MF: 4
#   ENDF_MT: 51
#   level:
#     number: 1
#     energy [MeV]:  8.467780E-01
#     spin:  2.000000E+00
#     parity: 1
#   E-incident [MeV]:  1.410000E+01
# residual:
#   Z: 26
#   A: 57
#   nuclide: Fe57
# datablock:
```

```
#   quantity: angular distribution
#   columns: 4
#   entries: 20
##     Angle           xs             dxs       Normalization
##     [deg]         [mb/sr]         [mb/sr]          []
   2.000000E+01   1.200646E+01   2.099300E+00   1.000000E+00
   2.510000E+01   1.550394E+01   2.293100E+00   1.000000E+00
   3.000000E+01   1.429579E+01   1.309900E+00   1.000000E+00
   3.499999E+01   1.328320E+01   1.050300E+00   1.000000E+00
   4.000000E+01   1.320642E+01   9.249600E-01   1.000000E+00
# ...............
```

Note that, as for cross sections, we have guessed that this concerns the first inelastic level by comparing the excitation energy given in EXFOR with the energies of the TALYS (RIPL) discrete level file. We give both values near the top of the file, so they can be compared. The contents are not full-proof, since sorting the XC4 file for secondary distributions has not yet been done adequately. In other words, use angular information in EXFORTABLES with care.

3.4 Single-differential particle spectra

Single-differential (i.e. angle-integrated) particle spectra are given in the *nspec/*, *pspec/*, etc. subdirectories. For example, neutron induced proton emission spectra for ^{56}Fe can be found in *n/Fe056/spectrum/p/*. The file *n-Fe056-MT203-Grimes-10827107-spec-E0014.800.1979* looks as follows

```
# Target Z    :  26
# Target A    :  56
# Target state:
# Projectile  : n
# Reaction    : (n, xp)
# E-inc       :   14.800 MeV
# Quantity    : Differential cross section
# Frame       : L
# MF          :    5
# MT          :  203
# X4 ID       : 10827107
# X4 code     : 26-FE-56(N,X)1-H-1,,DE
# Author      : Grimes
# Year        : 1979
# Data points :     22
# E-out(MeV)    xs(mb/MeV)  dxs(mb/MeV)     dE(MeV)
  1.75000E+00  2.50000E-05  0.00000E+00  2.50000E-01
  2.25000E+00  2.17000E-05  0.00000E+00  2.50000E-01
  2.75000E+00  2.32000E-05  0.00000E+00  2.50000E-01
  3.25000E+00  3.29000E-05  3.24000E-06  2.50000E-01
  3.75000E+00  3.79000E-05  0.00000E+00  2.50000E-01
.................
```

Here, the first column now contains the emission energy. The contents are not full-proof, since sorting the XC4 file for secondary distributions has not yet been done adequately. In other words, use spectra in EXFORTABLES with care.

3.5 Double-differential particle spectra

Double-differential particle spectra are given in the *ddx/n/*, *ddx/p/*, etc. subdirectories. For example, neutron induced double-differential neutron spectra for ^{56}Fe can be found in *n/Fe056/ddx/n/*. In this directory, the file *n-Fe056-MT201-Marcinkowski-12811003-ddx-E0025.700.1983* looks as follows

```
# Target Z    :  26
# Target A    :  56
# Target state:
# Projectile  : n
# Reaction    : (n, xn)
# E-inc       :   25.700 MeV
# Quantity    : Double-differential cross section
# Frame       : C
# MF          :    6
# MT          :  201
# X4 ID       : 12811003
# X4 code     : 26-FE-56(N,X)0-NN-1,,DA/DE
# Author      : Marcinkowski
# Year        : 1983
# Data points :     98
# E-out(MeV) xs(mb/MeV.sr) dxs(mb/MeV.sr) dE(MeV)
  1.25000E+01  3.97000E-06  1.30000E-07  5.00000E-01  2.47000E+01
  1.35000E+01  3.25000E-06  1.30000E-07  5.00000E-01  2.47000E+01
  1.45000E+01  2.90000E-06  1.20000E-07  5.00000E-01  2.47000E+01
  1.55000E+01  2.83000E-06  1.20000E-07  5.00000E-01  2.47000E+01
.....................
```

Here, the first column now contains the emission energy. The double-differential particle spectra could not yet be processed correctly. The emission angle is still in the final column and we need a better sorted XC4 file before a better file system can be made for this class of data.

3.6 Other types of data

3.6.1 Ratio data

Ratio data have been automatically stored under *ratio/* but are still to be further investigated and described.

3.6.2 Resonance data

Resonance data have been automatically stored under *resonance/* but are still to be further investigated and described.

3.6.3 Resonance integrals

Resonance integrals have been automatically stored under *resint/* but are still to be further investigated and described.

3.6.4 Fission yields

Fission yields have been automatically stored under *FY/* but are still to be further investigated and described.

3.6.5 Average number of fission neutrons

The average number of fission neutron have been automatically stored under *fission/*, for the total, delayed and prompt fission neutrons, but are still to be further investigated and described.

MT	Reaction	MT	Reaction	MT	Reaction
1	Total	34	(n,nh)	113	(n,t2α)
2	Elastic	35	(n,nd2α)	114	(n,d2α)
3	Non-elastic	36	(n,nt2α)	115	(n,pd)
4	Total (n,n')	37	(n,4n)	116	(n,pt)
5	(n,x)	38	4th-chance (n,f)	117	(n,dα)
11	(n,2nd)	41	(n,2np)	201	(n,xn)
16	(n,2n)	42	(n,3np)	202	(n,xγ)
17	(n,3n)	44	(n,n2p)	203	(n,xp)
18	Total (n,f)	45	(n,npα)	204	(n,xd)
19	1st-chance (n,f)	51-90	(n,n'$_1$) - (n,n'$_{40}$)	205	(n,xt)
20	2nd-chance (n,f)	91	Continuum (n,n')	206	(n,xh)
21	3rd-chance (n,f)	102	(n,γ)	207	(n,xα)
22	(n,nα)	103	(n,p)	600-640	(n,p$_0$) - (n,p$_{40}$)
23	(n,n3α)	104	(n,d)	649	Continuum (n,p)
24	(n,2nα)	105	(n,t)	650-690	(n,d$_0$) - (n,d$_{40}$)
25	(n,3nα)	106	(n,h)	699	Continuum (n,d)
28	(n,np)	107	(n,α)	700-740	(n,t$_0$) - (n,t$_{40}$)
29	(n,n2α)	108	(n,2α)	749	Continuum (n,t)
30	(n,2n2α)	109	(n,3α)	750-790	(n,h$_0$) - (n,h$_{40}$)
32	(n,nd)	111	(n,2p)	799	Continuum (n,h)
33	(n,nt)	112	(n,pα)	800-840	(n,α_0) - (n,α_{40})
				849	Continuum (n,α)

Table 3.1: The ENDF-6 MT numbers and corresponding reaction channels.

4. Testing the EXFOR database

While producing the directory-structured database described in the previous chapter, we can simultaneously check experimental data on some basic level of correctness and quality (in terms of reasonable values). For all nuclear reaction data for which this was possible, we have compared EXFOR data with a set of nuclear data libraries (NDL) or TALYS calculations. Since the C5 database is not yet well sorted and processable in terms of secondary distributions (angular distributions, spectra), the current comparison has only been done for cross sections. With future versions of C5 and the *exfortables.f90* code, we hope that more data can be tested.

For each experimental energy point we search for the corresponding energy point, using interpolation, in the NDL's and provide a measure for the deviation.

Besides the NDL comparison, we have also done a global study on the uncertainties of data points in EXFOR. This will be reported in a separate section.

This testing leads to a subjective quality scoring as outline in Ref. [10].

While we make the translation from the C5 computational database to our own directory-structured database, we do our NDL comparison, checks and statistical analyses on the fly. After about 1-2 hours, the conversion is done and all checking and statistical results are available. Though not present in this release, various TALYS options (global vs. local, microscopic vs phenomenological etc.) can also be tested against the entire EXFOR database, provided first the database with TALYS results is made, similar to those of the NDL's.

4.1 Global statistics

First of all, we like to keep track of how many entries EXFOR contains, and how many can be processed by our system. This information is given in various files in the directory *stat/total/*.

The main output file is *statistics*, which looks as follows:

```
EXFORTABLES Statistics
```

```
Date of XC4 file              : 20191212
Time of XC4 file              :   184918
Time of X4 file               : 20191211

              Entries        Subentries           Data points

EXFOR(NRDC)                     172946

XC4(NRDC)       16103           104970           8887277

EXFOR           23933           181369

XC4             16103 ( 67.3%)  104970 ( 57.9%)  8887277 (100.0%)

EXFORTABLES                      95995 ( 91.4%)  8574341 ( 96.5%)

Libraries                        67420 ( 70.2%)  3581289 ( 41.8%)
```

This table shows among others that only 2/3 of the EXFOR database has been translated into computational XC4 format, that almost the entire XC4 database has been translated into EXFORTABLES, and that we have been able to do an NDL comparison for less than half of the EXFOR data points.

Various other files are produced that give insight in our current ability to process EXFOR data, and which give a first indication of the quality of the data:

- *x4toxc4.yes*: entries which are in X4 and in XC4.
- *x4toxc4.not*: entries which are in X4 but not in XC4. This list could be judged by the person doing the X4toXC4 translation, to see how the translation rate could be increased.
- *xc4toexfortables.yes*: the EXFOR entries which have been translated from XC4 into EXFORTABLES.
- *xc4toexfortables.not*: the EXFOR entries which have not been translated from XC4 into EXFORTABLES. These remaining items give a clear indication on what classes of nuclear data still need to be considered in terms of processing by the *exfortables.f90* code.
- *warnings*: entries with values which are suspicious, on the basis of simple physics rules. As much as possible the reason for this suspicion is given in this file.
- *compare.yes*: the subentries of EXFORTABLES which could be compared with NDL's
- *compare.no*: the subentries of EXFORTABLES which could not be compared with NDL's

4.2 Goodness-of-fit estimators as an EXFOR test

The comparison done in EXFORTABLES is an EXFOR test and an NDL test at the same time. All statistical results averaged per data set are available in the directory *stat/comp/*. 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 or the TENDL library [3] provides very reasonable estimates for many reaction processes. Hence, with the exception of certain reactions, they should be able to give a reasonably good prediction of many reaction data, and obviously we will always try to extend such predictions to as many reactions as possible in future versions. At first sight, the problem is simple: If we know that an NDL or TALYS is usually within e.g. 30% of the

experimental data for a certain reaction channel, alarm bells should start ringing if the deviation of a data set for such a channel is suddenly much larger. We note that large deviations may also come from bad NDL or 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 3, close to threshold. In general the rule holds that the smaller the cross section, the larger the relative error. It is therefore important to judge 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 by comparing with a model code, which is why these EXFOR problems have not been discovered in the first place. The problems which are easiest to detect concern C/E values around 0.001 or 1000, (with C standing for an NDL or TALYS) suggesting the well known error of mistaking barns for millibarns. Unfortunately, the majority of cases is more difficult to judge. 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, which can be regarded as an "injustice" to experimentalists who have provided good-quality experimental data. In other words, if the NDL's are 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 NDL result with the correct quantity, and the EXFOR reaction identifier should perhaps be corrected.

4.2.1 Nuclear data libraries and TALYS

To judge a single experimental data point, one may compare it with various other estimates for that point:

- Other measurements for the same reaction and energy range,
- CENDL-3.2 [7]: Chinese Evaluated Nuclear Data Library (China), a general purpose library for neutrons,
- EAF-2010 [9]: European Activation File (UKAEA Culham/NRG Petten), a special purpose library for activation reactions,
- ENDF/B-VIII [4]: Evaluated Nuclear Data File (USA), a general purpose library for neutrons,
- IRDFF-2.0 [8]: International Reactor Dosimetry and Fusion File (IAEA), a special purpose library for a limited number of reaction channels,
- JEFF-3.3 [5] : Joint Evaluated Fission and Fusion file (NEA Data Bank), a general purpose library for neutrons,
- JENDL-5.0 [6]: Japanese Evaluated Nuclear Data Library (Japan), a general purpose library for neutrons,
- TENDL-2023 [3]: TALYS Evaluated Nuclear Data Library, a general purpose library for neutrons and all other incident particles,
- Nuclear model codes, in our case TALYS [12], with different options, e.g. phenomenological vs microscopic inputs, global versus local adjustment, etc.

The existing nuclear data libraries should be able to give a reasonably good prediction of many reaction data. It should of course be realised that the contents of these data libraries are already heavily dependent on the experimental data which are checked. Usually, they consist of nuclear model calculations tuned to EXFOR data, but often the experimental data are included, often through some least-squares fit, themselves as well. At first sight, the problem is simple: If it is known that libraries are usually reasonably close to the experimental data for a certain reaction channel, alarm bells should start ringing if the deviation of an experimental data set for such a channel is suddenly much larger.

4.2.2 Traditional goodness-of-fit estimators

To discover and classify problems, often a few well known goodness-of-fit estimators are used. If they are all very large, something is wrong somewhere. They are the F-factor

$$F = \mathrm{e}^{\sqrt{\frac{1}{N}\Sigma_i^N\left(\ln\frac{\sigma_T^i}{\sigma_E^i}\right)^2}}, \tag{4.1}$$

the χ^2 estimator,

$$\chi^2 = \frac{1}{N}\sum_i^N\left(\frac{\sigma_T^i-\sigma_E^i}{\Delta\sigma_E^i}\right)^2, \tag{4.2}$$

and the absolute deviation

$$\Delta = \frac{1}{N}\sum_i^N|\sigma_T^i-\sigma_{exp}^i|. \tag{4.3}$$

In these equations, the subscript T stands for theory, TALYS or NDL and E for experimental. In all cases, we average over the number of energy points, N, in each data set. Hence, each EXFOR subentry (data set) that contains a cross section excitation function, or only 1 point, is described by 3 average numbers: F, χ^2 and Δ, while we also keep track of all individual points F_i, χ_i^2 and Δ_i, in an extra column in the EXFORTABLES reaction database, see the previous chapter.

The F-factor is a kind of twisted C/E= σ_T/σ_E value. In fact, each individual component of the sum inside F contributes to C/E if it is larger than 1, and E/C if it is smaller than 1. This is a more appropriate quantity than the average C/E, since averaging C/E values over many points may not be very meaningful if the individual values cross unity at some point. Eq. (4.1) remedies this. A value of F=1.2 means that for the entire data set we are approximately 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 best and worst cases. Another standard indicator is of course χ^2, but then the extra complexity is that apart from the central values the uncertainties given in EXFOR need to be reliable as well. This is a separate issue which will addressed in a later section. Finally, large F or χ^2 values may be normal if the underlying quantities have a small value. To identify those cases, the absolute deviation in mb, Δ is helpful. In sum, it is best to look at all such indicators simultaneously.

4.2.3 The p-value, the ultimate goodness-of-fit estimator?

With the goodness-of-fit estimators outlined above, we test the deviation of one experimental data point at the time with various alternatives for NDL's and TALYS, also one by one, and there is an F-value for each NDL-experiment combination.

If we look at one particular experimental data point, we can also construct a probability distribution based on *all* alternative values, from NDL's and other measurements (at approximately the same energy).

One may typically have about 5 (correlated or independent) NDL's, a few other measurements, TALYS results from various options for the models etc. and, though not used here yet, other model codes, alternative systematics or Machine Learning estimates. This entire collection of information is sometimes enough to make statistically sound verdicts about outliers.

The proper quantity to do that is the p-value. In null hypothesis significance testing, the p-value is the probability of obtaining test results at least as extreme as the results actually observed, under

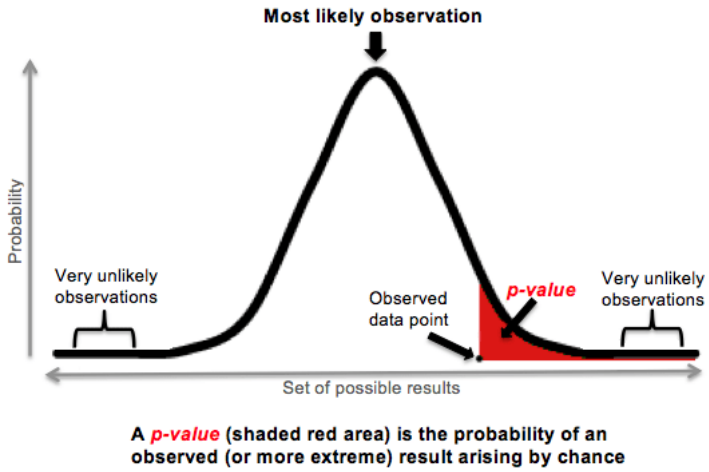

Figure 4.1: General notion of a p-value

the assumption that the null hypothesis is correct. Figure 4.1 explains the general notion of a p-value.

For the detection of outliers in EXFOR, we take the liberty to loosely translate this as follows: the p-value is the probability that the true cross section value is the observed data point or more extreme than that.

As an example, Figure 4.2 shows a plot of for the ^{31}P(n,p) reaction, for various measurements and NDL's. The lowest value at 14.8 MeV is from a measurement by Prasad(1971). Fig. 4.3 shows the probability distribution of the reaction at 14.8 MeV. Obviously, the more independent values can be found around the average (i.e. the peak of the distribution), the smaller the width of the distribution becomes and the worse the verdict is for the outlier. Information about data points inside an energy window close to the data point under consideration can be taken into account. For this particular example, the p-value, i.e. the probability that the true cross section is given by the value of Prasad *or lower than that* is 0.016.

4.3 Output of detailed statistical information

The statistical information can be further classified in various different ways, allowing for easier correction plans, and these are discussed below.

Most of the tables contain the F values for all reactions that have been compared with TALYS, with one F value per reaction, sorted by increasing F value. Other columns contain the EXFOR ID number, the number of points in each data set and the reaction string. Large F values may give an indication of problems in EXFOR, problems in TALYS, or an indication that incomparable quantities are compared. For the correction of EXFOR, the most interesting are of course the specific cases with (very) large deviation. This may range from F=2-3 up to $F > 10^{15}$. We note however, that the latter cases may not concern EXFOR errors at all! For example, TALYS is known to deviate from the experimental (n,He3) threshold by at least 1 MeV, resulting in a huge deviation near threshold. Often, for such cases the deviation is small on an absolute scale.

Eventually, we think the p-value will replace the F value as the leading goodness-of-fit estimator, but more development is needed for that. The p-values are given in the various statistical diagnosis files.

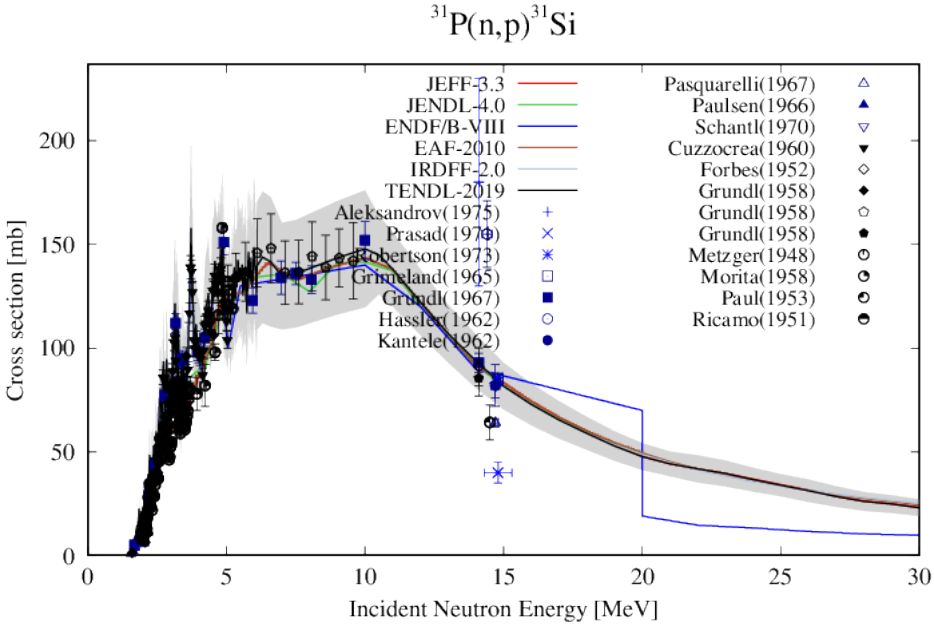

Figure 4.2: ^{31}P(n,p) reaction: experimental data and nuclear data libraries

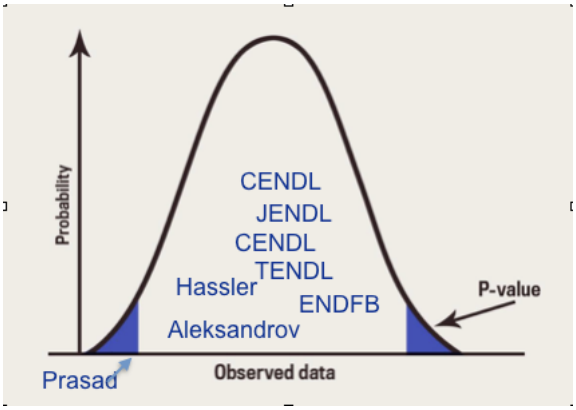

Figure 4.3: Probability distribution for ^{31}P(n,p) reaction at 14.8 MeV

4.3.1 F values per bin

To get a global view on the comparison between the world libraries and EXFOR, we store the F values for all measurements in numerical bins for each reaction. The results are given in *stat/comp/histo* for each MT number. With these files, the number of outliers per reaction channel can be estimated. As an example *n-MT001.histo*, for neutron total cross sections, looks as follows

```
#MT =   1 (n,tot)    #Sets:  2091 Reference: World
#  Fbin  #Sets  Cum. fraction   Average F(1-sigma):      1.54201 F(2-sigma):      6.15139
   1.000   279      13.343       1.031
   1.050   400      32.472       1.055
   1.100   235      43.711       1.073
   1.150   133      50.072       1.085
   1.200   100      54.854       1.097
   1.250    83      58.824       1.109
.............................
 127.893     8      99.857       2.803
 253.189     3     100.000       3.293
 503.187     0     100.000       3.293
1000.000     0     100.000       3.293
```

in which we list the width of the F bin, the number of sets falling in that bin, the accumulated fraction of all cases (which thus always runs to 100%), and the average F value calculated for all sets up to and including that bin.

We distribute the bins up to F=3 linearly and divide F=3-1000 over logarithmic equidistant bins. Hence, the first bin contains the NDL's which deviates between 0-5% from the experimental data, the second bin between 5-10%, and so on. All cases with F>1000 are put in the last bin. The high peak at the lowest bins means good news for the nuclear data libraries and EXFOR. The cases with very high F-values probably mean trouble for EXFOR (or XC4, or the it exfortables code). The cases in between mean trouble for either the libraries or EXFOR, or both. For error determination in EXFOR, 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 our 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, which are all stored in MT851, see e.g. *p-MT851.histo* for incident protons. Fig. 4.4 shows the distribution of the F-values for all (n,2n) and (p,n) reactions, in this case for TALYS, from files *n-MT016.histo*, and *p-MT004.histo*, respectively, that we managed to get out of the XC4 database.

4.3.2 F values per reaction

Another interesting check is to look at lists of sorted F values for one kind of reaction. These can be found in *stat/comp/MT/*, where e.g. the file *n-MT016.F* gives all (n,2n) reactions in EXFOR, sorted by increasing F value. This file looks as follows:

```
# Z   A   T   M  SUBENT      AUTHOR    YEAR  N            Reaction                F
  84 210   0  -1 41065002 Faddeev        1990  0 84-PO-210(N,2N)84-PO-209,,SIG  0.00
  40  88   0  -1 12763002 Prestwood      1984  0 40-ZR-88(N,2N)40-ZR-87,,SIG    0.00
  39  88   0  -1 127630042Prestwood      1984  0 39-Y-88(N,2N)39-Y-87,,SIG       0.00
  39  88   0  -1 31653002 HuangFeizengg 1990  0 39-Y-88(N,2N)39-Y-87,,SIG,,,DE 0.00
.....................
  69 169   0  -1 20802010 Dilg          1968  1 69-TM-169(N,2N)69-TM-168,,SIG  1.00
   3   6   0  -1 20794005 Mather        1969  1 3-LI-6(N,2N)3-LI-5,,SIG         1.00
```

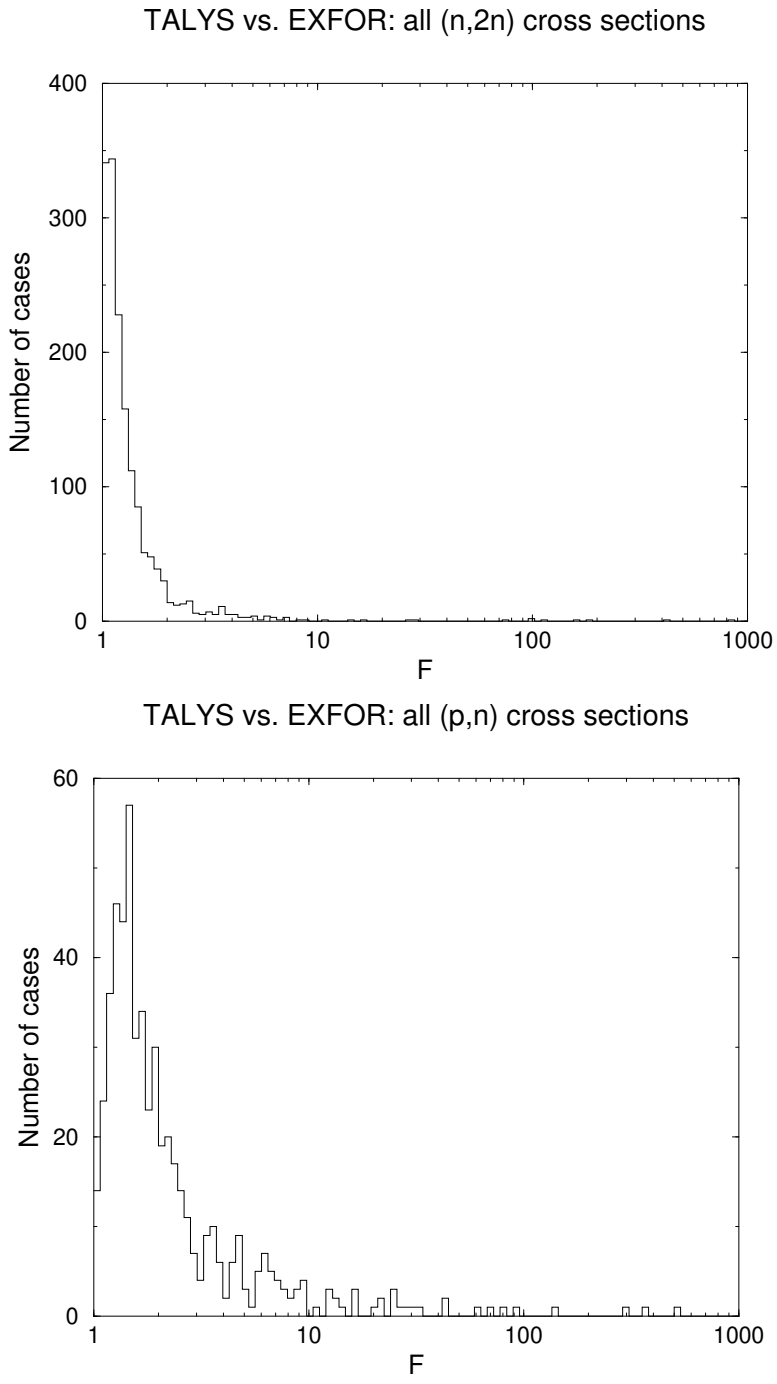

Figure 4.4: Frequency table for the F-values for (n,2n) reactions, and all (p,n) reactions.

```
  82 204   0  -1 40171005 Druzhinin      1972  1 82-PB-204(N,2N)82-PB-203,,SIG  1.01
  82   0   0  -1 40136025 Maslov         1974  1 82-PB-0(N,2N),,SIG             1.01
.....................
   5   0   0  -1 11632013 Ashby          1958  1 5-B-0(N,2N),,SIG               12.7
  20  40   0  -1 11520008 Arnold         1965  7 20-CA-40(N,2N)20-CA-39,,SIG    19.2
  14  28   0  -1 11520004 Arnold         1965  5 14-SI-28(N,2N)14-SI-27,,SIG    98.1
  25  55   0  -1 11684003 Nix            1961  1 25-MN-55(N,2N)25-MN-54,,SIG    781.
```

where one could expect that for F-values around 1000 we probably have a barn-millibarn error. Unfortunately, not all problems are that simple. On isomeric production, there may be an inconsistency between TENDL (most other NDL's do not contain isomers) and EXFOR on the definition of an isomer, leading to large discrepancies and it is also known that in several cases isomeric, ground state and total cross sections have been mixed in the EXFOR database.

4.3.3 F values per reaction, nuclide and projectile

The most detailed sorting of the F values per reaction can be found in *stat/comp/reaction/*, which is probably the best directory to start working on erroneous cases. Here we have sorted the results in one file per target isotope, projectile, and reaction. This produces a large list of files, whereby each file directly shows the outliers from the average and from TALYS. If we look at e.g. *n-Y089-MT103.F* we have all ^{89}Y(n,p) reactions sorted, in increasing F-order,

```
# Z   A   T   M  SUBENT      AUTHOR    YEAR   N            Reaction           F
  39  89   0  -1 402230201Levkovskii    1969   1 39-Y-89(N,P)38-SR-89,,SIG  1.08
  39  89   0  -1 31532004 Klopries      1997  11 39-Y-89(N,P)38-SR-89,,SIG  1.15
  39  89   0  -1 31494005 Molla        1998   3 39-Y-89(N,P)38-SR-89,,SIG  1.16
  39  89   0  -1 32729004 LuoJunhuaa   2016   3 39-Y-89(N,P)38-SR-89,,SIG  1.19
  39  89   0  -1 11462009 Bayhurst     1961  18 39-Y-89(N,P)38-SR-89,,SIG  1.25
  39  89   0  -1 11504007 Tewes        1960   5 39-Y-89(N,P)38-SR-89,,SIG  1.50
  39  89   0  -1 30115008 Csikai       1967   1 39-Y-89(N,P)38-SR-89,,SIG  2.29
```

The interesting cases now concern different F values for similar energy ranges (listed in some columns), indicating problems for one or more data sets.

4.3.4 F values per reaction and mass unit

In *stat/comp/MTA/* we find the F-values per reaction channel and binned per mass number of the target. For example *n-MT016.F* looks as follows,

```
# Average F values per mass unit for neutron  and MT =  16 (n,2n)
# Z    A              World                  TALYS
#            |  F     #points #sets |  F     #points #sets
       1      1.24        189    46  0.00          0     0
       2      1.28        192    49  0.00          0     0
       3      1.28        192    49  0.00          0     0
       4      1.32        209    55  0.00          0     0
```

These files may be used to see if there is a trend of the libraries describing experimental data as a function of target mass.

4.3.5 F values per reaction and energy

In *stat/comp/MTE/* we find the F-values per reaction channel and binned per incident energy. For example *n-MT016.F* looks as follows,

```
# Average F values per energy bin for neutron  and MT =   4 (n,n')
# energy              World                 TALYS
#            |  F     #points #sets |  F     #points #sets
    0.00       1.40         11     7  0.00          0     0
   0.200       1.51         51    19  0.00          0     0
   0.500       1.76       4693    56  0.00          0     0
    1.00       1.89       7453    85  0.00          0     0
    1.50       1.65       5062    59  0.00          0     0
    2.00       1.86       5541    71  0.00          0     0
    3.00       1.28       1341    44  0.00          0     0
```

These files may be used to see if there is a trend of the libraries describing experimental data as a function of incident energy.

4.3.6 F values per reaction

In *stat/comp/MT/* we find the F-values per reaction channel, sorted by F value. For example *n-MT103.F* looks as follows,

```
# Z   A   T   M  SUBENT      AUTHOR      YEAR     N            Reaction
  55 133  0  -1 30175002 Borbely          1963     0 55-CS-133(N,P)54-XE-133,,S
  40  94  0  -1 11856006 Carroll          1966     0 40-ZR-94(N,P)39-Y-94,,SIG
  37  84  0  -1 V1001334 Mughabghab       2006     0 37-RB-84(N,P)36-KR-84,,SIG,
....
```

4.3.7 F values per nuclide and energy

In *stat/comp/nucE/* we find the F-values per nuclide, summed over reaction channels and binned per incident energy. For example *n-Fe056* looks as follows,

```
# Average F values per energy bin for neutron  +  Fe56
# energy              World                 TALYS
#            |  F     #points #sets |  F     #points #sets
  2.5300E-08  1.06         148    17  0.00          0     0
  1.0000E-06  1.07          61     6  0.00          0     0
  1.0000E-05  1.23          31     4  0.00          0     0
  1.0000E-04  1.12          13     4  0.00          0     0
  1.0000E-03  1.52          17     4  0.00          0     0
```

which may give a global profile of the description of the world libraries for a nuclide as a function of incident energy.

4.3.8 F values per nuclide and reaction

In *stat/comp/nucMT/* we find the F-values per nuclide, summed over reaction channels and binned per incident energy. For example *n-Fe056* looks as follows,

```
# Average F values per MT number for neutron  + Fe 56
# Z    A  MT Iso         World                  TALYS
#            |  F     #points #sets |  F     #points #sets
        1 -1  1.58       18600    13  0.00          0     0
        2 -1  1.08          19     9  0.00          0     0
        3 -1  1.10           2     1  0.00          0     0
        4 -1  1.98       14621    10  0.00          0     0
       16 -1  1.27          37     9  0.00          0     0
```

which may give a global profile of the description of the world libraries for a nuclide as a function of reaction channel.

4.3.9 F values per energy bin

In *stat/comp/parE/* we find the F-values per energy bin.

4.3.10 F values per entry

In *stat/comp/entry/* we find the F-values per EXFOR entry.

4.3.11 Total F values

In *stat/comp/total/* we find the total statistical numbers. The first are the average F values per MT number, together with the number of data points and data sets used in the comparison. For example, here is the top of *n-MT.F*

```
# Average F values per reaction summed over all nuclides for neutron for all
    28865 subentries with F < 3.00
# MT Iso            World               TALYS
#           |  F     #points #sets | F    #points #sets |
   1 -1      1.56    1827109 2046  0.00         0     0
   2 -1      1.26      27904  631  0.00         0     0
   3 -1      1.26        757  368  0.00         0     0
   4 -1      1.74      27012  250  0.00         0     0
   4  0      1.62         81    2  0.00         0     0
   4  1      1.42       1960  230  0.00         0     0
   4  2      2.32          7    3  0.00         0     0
.........................
```

Next, the file *n-nuc.F* has the F-values per nuclide

```
# Average F values per nucleus summed over all reactions for neutron for all
    28865 subentries with F < 3.00
# Z    A            World               TALYS
#           |  F     #points #sets | F    #points #sets
   3   6      1.29      25434  121  0.00         0     0
   3   7      1.41       6383   59  0.00         0     0
   4   9      1.30      30897  136  0.00         0     0
   4  10      1.03          2    2  0.00         0     0
.........................
```

The file *n-allreac*, is rather large since it contains *all* reactions sorted by F-value.
The file *all* has all EXFOR entries sort by F value.
The file *all.F* has the total numbers per incident particle:

```
# Average F values per projectile summed over all reactions
# particle          World               TALYS                tendl.2023
    endfb8.0             jeff3.3              jendl5.0             eaf.2010
                cendl3.2             irdff2.0
#           |  F     #points #sets | F    #points #sets
gamma        1.56      24035  781  0.00         0     0
neutron      1.57    2659349 28865 0.00         0     0
proton       1.62     114866 6909  0.00         0     0
deuteron     1.68      29688 2239  0.00         0     0
triton       2.28       1663   43  0.00         0     0
helium-3     1.80       5972  657  0.00         0     0
alpha        1.72      57435 3579  0.00         0     0
heavyion     0.00          0    0  0.00         0     0
```

```
Total:  1.58    2893008 43073 0.00        0     0
```

where the last line gives the total number of points and data sets used in the comparison. These tables show that one should not take the F values too literally if they are summed over *all* data. It is completely dominated by far outliers.

There is also a file *MT.sum* which gives the number of reactions considered per MT number. The top of *MT.sum* looks as follows

```
 Total number of reactions compared with TALYS
 (g,tot  )            4
 (g,el   )            2
 (g,non  )           99
 (g,n'   )          114
 (g,n'   ) isom=0     1
 (g,n'   ) isom=1     2
 (g,2n   )          119
...........................
```

4.4 Comparison with nuclear data libraries

Fig. 4.5 shows the distribution of all (n,2n) subentries over the various F values. Note that the term "data set", i.e. the sum over N, can apply to one EXFOR subentry, e.g. one excitation function, all subentries for the same (Z,A) nuclide and reaction channel (MT number), all subentries for the same (Z,A) nuclide, all subentries for the same reaction channel (MT number), all subentries for the same projectile, and finally to the entire EXFOR database, or at least the part that could be compared. For all this, average F-values are recorded. In addition, all these averages can be taken for each nuclear data library (i.e. ENDF/B-VIII, JEFF, etc.) separately, or averaged over all of them. For the purpose of checking EXFOR, the goodness-of-fit for one subentry, i.e. one experimental data set for one energy or a range of energies, averaged over all libraries, is used as the leading indicator. For statistics on all reaction channels we refer to Ref. [10].

The F values can also be binned per incident energy, averaged over all nuclides. This is displayed in Fig. 4.6, for the (n,2n) subentries compared with all libraries. For isomeric reactions, only libraries with a significant amount of, or effort in, isomeric reactions have been included in the plots. The energy scale for these figures has been shifted. If we would simply plot the F values as a function of incident energy, an insignificant scatter plot would show up, since the threshold is different for each reaction, and the deviation from nuclear models is largest around the threshold. To take Q-values and Coulomb barriers into account in an empirical way, we have determined for each reaction the incident energy E_{1mb} where the excitation function crosses the value of 1 mb. This corresponds to the cross section value around which several measurements have been attempted. Plotting F values as a function of $E - E_{1mb}$ then reveals some trends which are to be expected. First, around threshold, i.e. $E = E_{1mb}$ the deviation is relatively large, near the peak it is smaller, and in the tail of the excitation function, i.e. $E - E_{1mb}$ is around 15 MeV, the deviation from models or libraries increases again. Again, consult Ref. [10] for the other reaction channels. These trends are the basis for the prior uncertainties of global TALYS calculations as outlined in the Bayesian Monte Carlo approach [13].

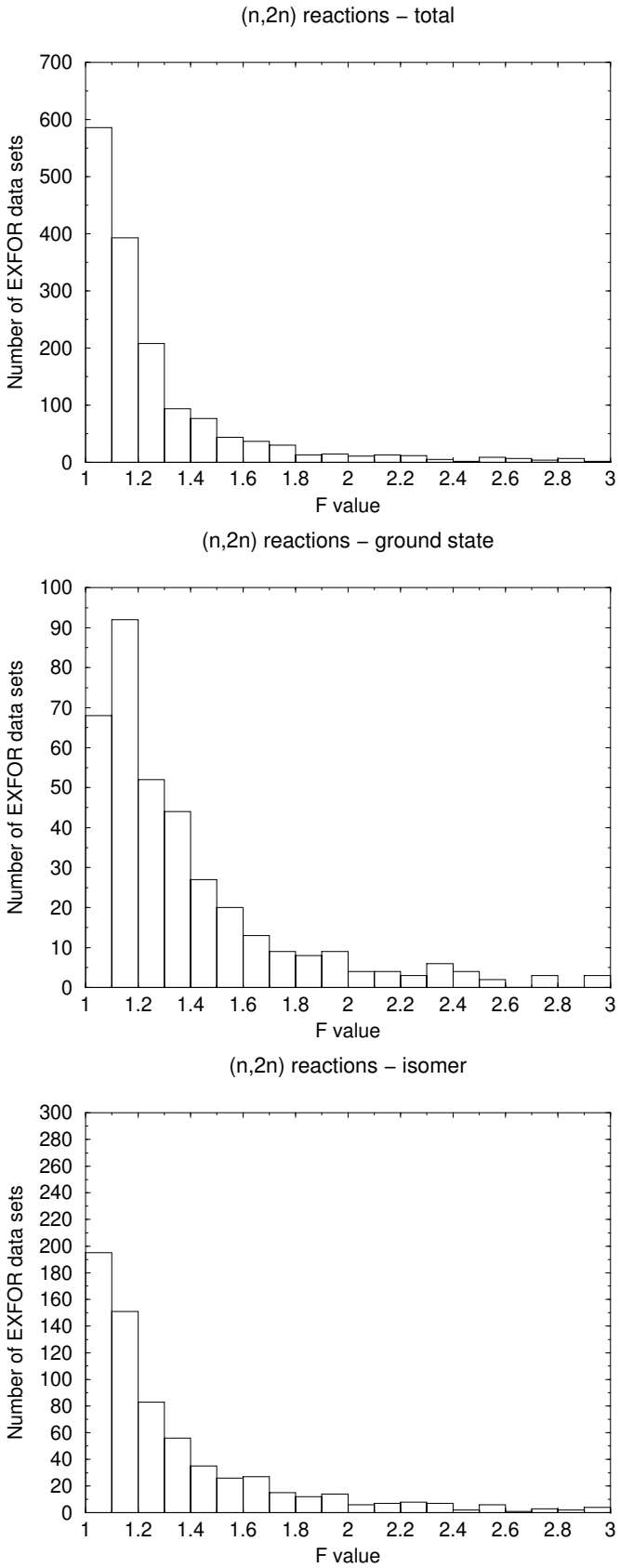

Figure 4.5: Distribution of F-values for the (n,2n) reaction.

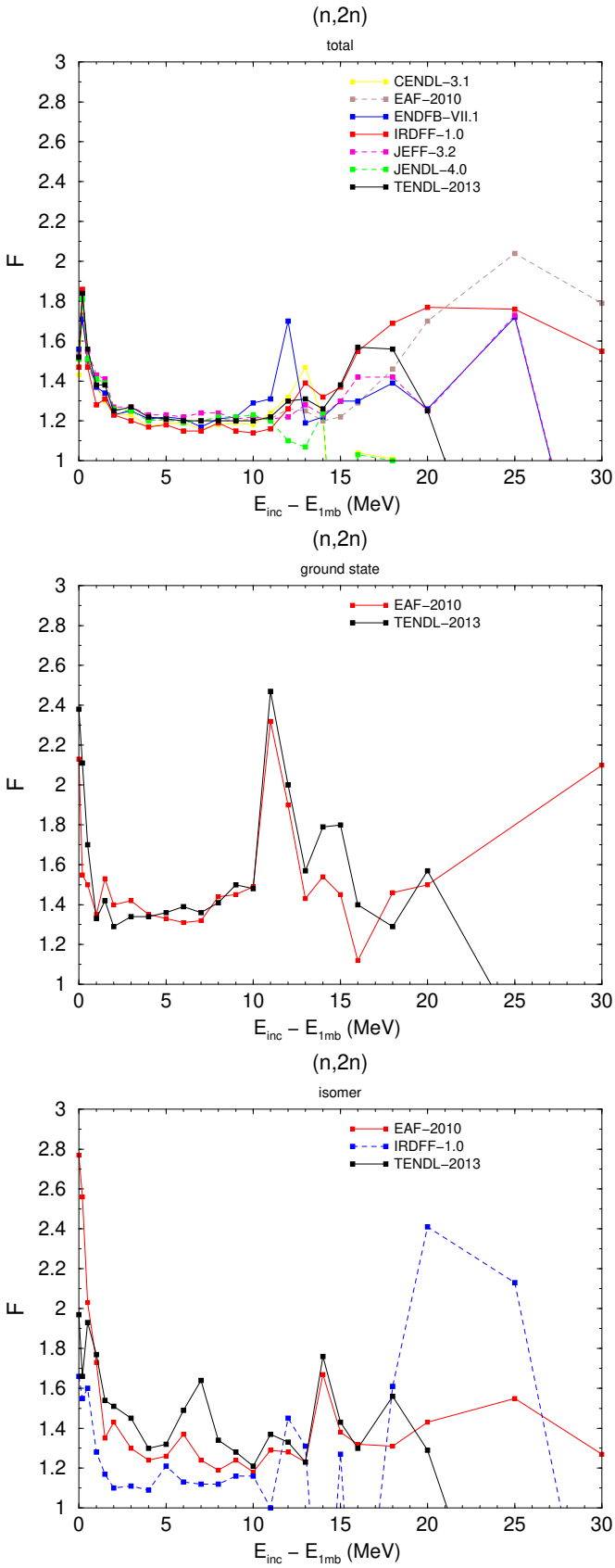

Figure 4.6: F values, averaged over all nuclides, for (n,2n) reactions as a function of incident energy, for the various nuclear data libraries. E_{1mb} is an estimate for the threshold energy and is the incident energy where the cross section crosses the 1 mb value.

4.5 Experimental uncertainties

Besides the comparison with the world libraries, we have also performed statistical tests on the experimental uncertainties. TENDL or other libraries are not needed for this: we simply analyze the uncertainties as given in the XC4 database. The results are in *stat/unc/*. The results are classified in different ways.

4.5.1 Uncertainties per MT number

Per reaction channel, the experimental uncertainties for all entries are averaged per data set, sorted in increasing uncertainty and printed. Unrealistically small or large experimental uncertainties can then be identified upon closer inspection of the EXFOR file. The results are stored in *stat/unc/MT/* where e.g. the file *n-MT018.unc* gives the average uncertainty per measured fission cross section data set. That file looks as follows:

```
 SUBENT  ....Rel. Err.  E-min     E-max    Abs. Err. Rel. Err.
14229020 ....  0.00     0.506      3.65      0.00      0.00
14229019 ....  0.00     0.505      3.86      0.00      0.00
12562003 ....  0.00     2.530E-08 2.530E-08  0.00      0.00
406700022....  0.00     0.130      7.40      0.00      0.00
10349004 ....  0.00     3.639E-07 1.700E-06  0.00      0.00
10670002 ....  0.00     0.700      2.98      0.00      0.00
14229013 ....  0.00     0.597      4.23      0.00      0.00
14229012 ....  0.00     0.891      4.94      0.00      0.00
14229011 ....  0.00     0.455      3.07      0.00      0.00
.............
103460022....  1.21      1.00      6.00      0.00      0.00
41303007 ....  1.22     0.500      14.9     7.675E-02  2.82
21520008 ....  1.22     2.530E-08 2.530E-08  0.00      0.00
40751006 ....  1.24     1.500E-04 4.500E-02 1.040E-03  12.8
41303011 ....  1.25     0.135      15.0     8.533E-02  5.44
20143006 ....  1.25     8.600E-09 7.170E-08  0.00      0.00
.............
20138002 ....  670.     2.005E-05 5.208E-05  0.00      0.00
10323002 ....7.763E+03 2.401E-05  4.68      0.00      0.00
23217003 ....2.472E+05  45.0      73.0      0.00      0.00
10266006 ....2.694E+06 2.001E-05 1.997E-04  0.00      0.00
20484002 ....2.483E+07 3.503E-06 2.975E-02  0.00      0.00
```

Again, the bottom of this file, with the largest average uncertainties, probably deserves close inspection. On the other hand also zero uncertainties (meaning no uncertainties reported or compiled), or very small uncertainties (too optimistic) may be suspicious.

4.5.2 Uncertainties per bin

The experimental uncertainties per reaction channel are also stored in bins, which gives info about the average experimental uncertainty for different reaction channels. The results are stored in *stat/unc/histo/* where e.g. the file *n-MT018.histo* gives the uncertainties per measured fission cross section data set binned per value. The top of that file looks as follows:

```
#Uncertainties for MT =  18 (n,f)
#  % bin      #sets
   0.00         0
   5.00       235
```

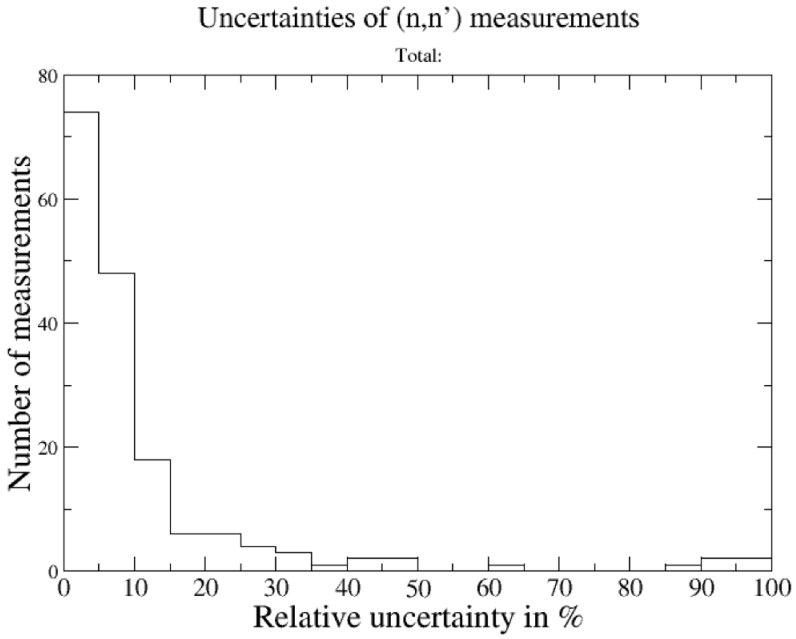

Figure 4.7: Uncertainties for the (n,n') reaction.

```
   10.0        123
   15.0         63
   20.0         36
   25.0         22
   30.0          8
   35.0          6
..............
   801.          0
   931.          0
 1.094E+03       0
 1.300E+03       4
```

which shows that most sets have an uncertainty, averaged over the entire set, within 5 % but there are also two cases with unrealistically large values. Closer inspection of the file mentioned above will then reveal the suspicious cases, and possibly directly the error. Fig. 4.7 shows the binned uncertainties for the (n,n') reaction. Such distributions could maybe be used to assign uncertainties to measurements where uncertainties were not reported.

4.5.3 All uncertainties

In *stat/unc/histo/* all uncertainties per projectile are given, sorted by uncertainty. These files, such as *n-unc* can be inspected to investigate the worst cases.

5. Quality scoring of EXFOR data sets

An extensive review of quality scoring of many data sets in EXFOR has been given in Ref. [10]. Many of these scores for subentries have been assigned automatically, by comparing the experimental values of the subentry with nuclear data libraries. Since the criteria for quality classes may change over time with new insights, the most important results of that study may be whether the data sets were correctly compiled in EXFOR or not, which was the result of a lot of manual work! The identifier for this, either a T, R, N or E, see below, will be maintained in the future no matter what the actual numerical quality score is.

For each subentry, a data block has been created including the basic information of the subentry review and the score. This database with all the scores is kept in the *files/quality.all* file. The T, R, N or E labels are read from there and combined with the quality score as produced by the last run of EXFORTABLES. The resulting files are in the *quality/* directory.

5.1 Scoring classes

As an initial classification, the data are categorized in three numerical classes: (1) close to, (2) reasonably close to or (3) far away from other data sources (usually evaluated data libraries). In addition, a symbol (T, R, N or E) is assigned to a data set to assign the review status.

5.1.1 Subentries which are not reviewed or not automatically compared (blank)

blank **Neither reviewed nor compared with evaluations.**
The subentry is not (yet) cross-checked with information from other measurements, libraries and/or calculations. This is the default score.

5.1.2 Subentries which are automatically compared with data libraries (T)

T1 **Automatically compared with libraries: small deviations.**
The subentry contains (very) probably the reaction and data measured by the author, and although the associated publication has not been checked by the reviewer, the quantities

have central values and uncertainties which are close to other measurements, libraries and/or calculations.

T2 **Automatically compared with libraries: questionable deviations**
The subentry contains maybe the reaction and data measured by the author, and the associated publication has not (yet) been checked by the reviewer. The quantities have central values and uncertainties which are deviating to some extent from other measurements, libraries and/or calculations.

T3 **Automatically compared with libraries: strong deviations**
The subentry contains probably not the reaction and data measured by the author, and the associated publication has not (yet) been checked by the reviewer. The quantities have central values and uncertainties which are strongly deviating from other measurements, libraries and/or calculations.

5.1.3 Subentries which are reviewed by checking the publication (R or N)

R1 **Paper reviewed: small deviations.**
The subentry contains certainly the reaction and data measured by the author, since the associated publication has been checked by the reviewer. The quantities have central values and uncertainties, which are close to other measurements, libraries and/or calculations.

R2 **Paper reviewed: questionable deviations**
The subentry contains certainly the reaction and data measured by the author, since the associated publication has been checked by the reviewer. The quantities have central values and uncertainties which are deviating to some extent from other measurements, libraries and/or calculations.

R3 **Paper reviewed: strong deviations**
The subentry contains certainly the reaction and data measured by the author, since the associated publication has been checked by the reviewer. The quantities have central values and uncertainties which are strongly deviating from other measurements, libraries and/or calculations.

N1 **Automatic score T1, but pdf of paper not available for checking**

N2 **Automatic score T2, but pdf of paper not available for checking**

N3 **Automatic score T3, but pdf of paper not available for checking**

Subentries which contain errors and require specified action (E)

E1 **Error: subentry contains other quantity or wrong values - small deviations.**
Although the quantities have central values and uncertainties which are close to other measurements, libraries and/or calculations, the subentry does not contain the reaction or data measured by the author, but either another quantity or a slightly different numerical value. Obviously, these errors are hardest to find, since these subentries initially get a 'T1' score. Action: confirmation and correction by Data Centres.

E2 **Error: subentry contains other quantity or wrong values - questionable deviations.**
The subentry does not contains the reaction or data values measured by the author, while the quantities have central values and uncertainties which are deviating to some extent from other measurements, libraries and/or calculations. These are errors in subentries which initially received a 'T2' score. The associated publication has been checked by the reviewer, and the values found are wrong. Action: confirmation and correction by Data Centres.

E3 **Error: subentry contains other quantity or wrong values - strong deviations.**

The subentry contains reaction and data that do not agree at all with other measurements, libraries and/or calculations. The associated publication has been checked by the reviewer, and often the values found are wrong. Sometimes, no origin of the value or alternative meaning for the value could be found. Action: further analysis, confirmation and correction by Data Centres.

5.2 Various stages of quality assignment

Stage 1
All EXFOR entries that can be automatically compared with nuclear data libraries get a score T1, T2 or T3. A score T1 will probably not be changed anymore. Only if unexpected new information comes to surface, from either the experimental or modelling side, this may change. Hence, the result of this stage is:

- T1: Definite assignment in database
- T2+T3: Could or should be reviewed.

Stage 2
The papers of subentries with score T2 and T3 are reviewed. In the course of time, the boundary between T1, T2 and T3 may be altered. This depends on the number of false alarms in the T2 class, which determines whether the decision for a paper review should be more or less strict. After paper review, a subentry with an initial score of T2 will end up in R2, N2, or E2, while a subentry with an initial score of T3 will end up in R3, N3, or E3. Hence, the result of this stage is:

- R2+R3: Definite assignment in database
- E1+E2+E3: Should be corrected
- N3 (and maybe N2): should have priority for acquiring the pdf file of the paper, so that it can be reviewed.

Stage 3
Cases with score E1, E2, E3 result in a message to the Data Centres with a recommended correction. After this correction, these subentries will be reviewed again after which they may be upgraded to R1, R2 or R3 in the next EXFOR update. Hence, the result of this stage is:

- R1+R2+R3: Definite assignment in database

Stage 4
The final scores, i.e. after all corrections, will be either blank, T1, N1, N2, N3, R1, R2 or R3.

It is noted that the scoring classes may be subject to change in the future. One could for example decide to use other numerical indicators, for example real numbers instead of just the integer 1, 2 and 3. What is most important now is the assignment of an 'R', specifying confirmation that the paper contains indeed the compiled quantity and value, even if there is a (large) discrepancy from other measured values or values from nuclear data libraries.

5.3 Quality scores

The goodness-of-fit estimators have already been described in Chapter 4, and represent an average deviation between data libraries and an experimental data set. Another factor we will use for quality scoring indicates the relative magnitude of the cross section itself, namely

$$Q_i = \frac{\sigma_T^i}{\sigma_{non}^i} \tag{5.1}$$

with σ^i_{non} the theoretical nonelastic cross section, and σ^i_T the theoretical partial reaction cross section.

Now that the goodness-of-fit estimators have been defined, the reviewing classification can be defined. The F values for all experimental data sets in a reaction channel have been ranked from small (close to 1) to large. For certain reaction channels the F values are relatively smaller than for others. This can have several reasons:

- The measurements are easier to perform, so that extreme outliers are generally not expected.
- The reaction channel is easier to model.
- There are more measurements per reaction channels, so that experimentalists are more influenced by previous work.
- etc.

This also means that so called "suspicious" F values are lower for some reaction channels than for others. For each reaction channel, the F value is determined for which "1-sigma", i.e. 68.27 % of all F values fall inside this particular F value. The similar F value for "2-sigma", i.e. 95.45 %, is determined. This leads to the values given in Table 5.1.

The quality classes assigned to an EXFOR data set are now as follows

Class 1 : (T1, N1, R1 and E1): $1 \leq F \leq F_{1\sigma}$

Class 2 : (T2, N2, R2 and E2): $F_{1\sigma} < F \leq F_{2\sigma}, \chi^2 < 30, Q_i < 0.10$
$F > F_{1\sigma}, \chi^2 < 30, Q_i < 0.05$

Class 3 : (T3, N3, R3 and E3): $F_{1\sigma} < F \leq F_{2\sigma}, \chi^2 > 30, Q_i > 0.10$
$F > F_{2\sigma}, \chi^2 > 30, Q_i > 0.05$

We think that class '3' should always be reviewed by checking the associated documentation (publication), while class '2' should be reviewed if, despite the more favorable numerical criterion, the visual fit is bad, also compared to other experiments.

All papers with an 'R' classification have been reviewed, while those with an 'E' classification have been sent to NRDC for correction. When these corrections have been applied, the 'E' category for these reactions should be turned into an 'R'. Similarly, the 'N3' category should disappear for these reactions, since after retrieval of the pdf papers from other sources, the papers have been reviewed and the code should change into 'R3'. Note that there are only a few 'E's, especially when compared with the 'R's. This can have several meanings:

- By far the majority of experiments in these reaction classes have been correctly compiled.
- A more efficient reviewing criterion is needed, only a few % of the suspicious values turn out to be compilation errors.
- A remarkable large amount of experimental data has been published which deviate considerably from the norm.

Reaction	$F_{1\sigma}$	$F_{2\sigma}$	#Sets
(n,n')	1.40	5.8	213
(n,n')m	1.52	3.5	251
(n,2n)	1.25	2.50	1675
(n,2n)g	1.41	2.45	397
(n,2n)m	1.28	3.6	719
(n,3n)	2.27	23.6	84
(n,n_1)	2.6	20.	491
(n,n_{2-40})	3.3	80.	312
(n,p)	1.31	3.55	1846
(n,p)g	1.82	8.0	188
(n,p)m	1.70	7.0	426
(n,d)	3.0	15.0	45
(n,t)	2.1	27.0	137
(n,a)	2.0	12.0	1068
(n,a)g	2.4	13.0	78
(n,a)m	2.75	13.0	196
(n,np)	7.5	121.7	148
(n,na)	7.2	6.75	54
(n,xp)	1.43	2.99	87
(n,xt)	2.00	159.	23
(n,xa)	2.07	9.07	164

Table 5.1: $F_{1\sigma}$ and $F_{2\sigma}$ values per reaction channel. A fraction of 68.27 and 95.45 %, respectively, of all F values fall inside the given boundaries. To indicate the statistical significance, also the number of included experimental data sets to come to these values is listed. Only channels with at least 20 experimental data sets have been assigned such boundaries.

6. Input description

The communication between EXFORTABLES and its users resembles strongly that of TALYS: It works with keywords which should obey some basic rules (see the TALYS tutorial at www.talys.eu). You may run EXFORTABLES with the input file below and you will get the same ready-to-use database as was delivered with this package. If you want to have a different version of the database you will have to use the keywords which are described in Chapter 7. The input file for EXFORTABLES to make the standard database is as follows

```
statistics y
tables y
x4 y
pointcomp all
qualitycomp all
lib y
talys n
xseps 0.001
Fmax 3.
dexp n
#maxentry 10000
#libspath /Users/koning/libraries/
#pointcomp tendl
#particle n
#Zmin 92
#Zmax 92
#Amin 235
#Amax 235
#Zmax 40
#eaf n
#endfb y
```

```
#jendl n
#tendl y
#jeff n
#irdff n
#cendl n
#outprocess y
#x4 n
```

Note that we have left some '#'-commented, and thus not used, keywords in this input file. These were used for some temporary tests, such as doing the database only up to 10000 EXFOR entries, with a specific path to the data libraries, comparing to TENDL only, for neutrons only, for U-235 only, only for TENDL and ENDF/B-VIII, to give detailed output for the processing, and to leave out detailed EXFOR input formation. The first 10 lines of the above input file drive the database creation and statistical comparison delivered with this package.

7. Reference Guide

In this part, all keywords will be described, one per page. The description of each keyword is as follows:

- Name of the keyword
- Explanation
- Examples
- Range of allowed values
- Default value
- Comments (optional), when we feel that some extra warnings or explanation for proper use is appropriate.

7.1 EXFORTABLES keywords

In this Section, we will explain all the possible keywords. For each keyword, we give an explanation, a few examples, the default value, and the theoretically allowed input range.

Amax

Maximal A value to process, used for quick tests.

Examples

Amax 208
Amax 40

Range

0 $\leq$ **Amax** $\leq$ **400** and **Amax** $\geq$ **Amin**

Default

Amax 400

Amin

Minimal A value to process, used for quick tests.

Examples

Amin 208
Amin 40

Range

0 $\leq$ **Amin** $\leq$ **400** and **Amin** $\leq$ **Amax**

Default

Amin 0

cendl

Flag to include or exclude CENDL from library average.

Examples

cendl y
cendl n

Range

y or **n**

Default

cendl y

dexp

Flag to use experimental uncertainty in F factor.

Examples

dexp y
dexp n

Range

y or **n**

Default

dexp y

eaf

Flag to include or exclude EAF from library average.

Examples

eaf y
eaf n

Range

y or **n**

Default

eaf y

Emin

Minimum energy (MeV) for comparison of EXFOR data.

Examples

Emin 0.001
Emin 50.

Range

0. <= **Emin** <= **1000.** and **Emin** <= **Emax**.

Default

Emin 0 MeV

Emax

Maximum energy (MeV) for comparison of EXFOR data.

Examples

Emax 200.
Emax 50.

Range

0. <= **Emax** <= **1000.** and **Emax** >= **emin**.

Default

Emax 1000 MeV.

endfb

Flag to include or exclude ENDF/B from library average.

Examples

endfb y
endfb n

Range

y or **n**

Default

endfb y

expo

Flag to use exponential root-mean-square factor instead of power of 10.

Examples

expo y
expo n

Range

y or **n**

Default

expo y

eview

Flag to make ECISVIEW files.

Examples

eview y
eview n

Range

y or **n**

Default

eview n

filespath

Path for the X4 and nuclear structure files.

Examples

filespath /home/koning/newfiles/

Range

filespath should exist.

Default

Default: **filespath ∼/exfortables/files/**

Fmax

Maximal F value per point taken into account.

Examples

Fmax 3.
Fmax 10.

Range

1. $<=$ **Fmax** $<=$ **1.e38**.

Default

Fmax 1.e38.

group

Flag to group resonance data.

Examples

group y
group n

Range

y or **n**

Default

group y

irdff

Flag to include or exclude IRDFF from library average.

Examples

irdff y
irdff n

Range

y or **n**

Default

irdff y

jeff

Flag to include or exclude JEFF from library average.

Examples

jeff y
jeff n

Range

y or **n**

Default

jeff y

jendl

Flag to include or exclude JENDL from library average.

Examples

jendl y
jendl n

Range

y or **n**

Default

jendl y

lib

Flag to compare EXFOR data with nuclear data libraries.

Examples

lib y
lib n

Range

y or **n**

Default

lib y

libspath

Path for the data libraries. You should have this path hardwired in subroutine *machine.f*, but it may be helpful to easily change between different versions of the cross section database.

Examples

libspath /home/koning/libraries

Range

libspath should exist.

Default

Default: **libspath ∼/libraries/**

MT

MT number to be included.

Examples

MT 2
MT 102

Range

1 <= **MT** <= **851**.

Default

MT is not used, i.e. all MT numbers are considered

maxentry

Number of EXFOR entries that are processed. Put this temporarily to a lower number (e.g. 1000) if you want to test whether a new version of EXFORTABLES, or new input case, works.

Examples

maxentry 2
maxentry 10

Range

1 $<=$ **maxentry** $<=$ **1000000000**.

Default

maxentry 1000000000, i.e. continue to the end.

outprocess

Flag for more extensive output on processing the EXFOR subentries.

Examples

outprocess y
outprocess n

Range

y or **n**

Default

outprocess n

particle

Particles which are included in the processing.

Examples

particle n
particle g n p a

Range

particle can be any combination of **g, n, p, d, t, h** and **a**.

Default

Include all possible particles, i.e. **particle g n p d t h a**

pointcomp

Reference for pointwise comparison.

Examples

pointcomp jeff
pointcomp tendl

Range

pointcomp should be equal to one of **all, talys, tendl, endfb, jeff, jendl, eaf, cendl, irdff**.

Default

pointcomp talys

qualitycomp

Reference library for quality assignment. For the 'quality; of each EXFOR subentry a weighted average of nuclear data libraries can be taken or a single library can be chosen.

Examples

qualitycomp jeff
qualitycomp tendl

Range

qualitycomp should be equal to one of **all, talys, tendl, endfb, jeff, jendl, eaf, cendl, irdff**.

Default

qualitycomp all, i.e. all libraries are included in the weighing.

remove

Flag to remove the previous database before creating a new one.

Examples

remove y
remove n

Range

y or **n**

Default

remove y

statistics

Flag for statistics of the EXFOR data.

Examples

statistics y
statistics n

Range

y or **n**

Default

statistics y

tables

Flag to produce new cross section database before doing statistics.

Examples

tables y
tables n

Range

y or **n**

Default

tables y

talys

Flag to do a TALYS/TENDL comparison for all the EXFOR data.

Examples

talys y
talys n

Range

y or **n**

Default

talys y

talysemin

Minimum energy (MeV) for TALYS comparison.

Examples

talysemin 0.1
talysemin 5.

Range

0. $<=$ **talysemin** $<=$ **1000.** and **talysemin** $<=$ **talysemax**.

Default

talysemin 0.001 MeV.

talysemax

Maximum energy (MeV) for TALYS comparison.

Examples

talysemax 200.
talysemax 50.

Range

0. $<=$ **talysemax** $<=$ **1000.** and **talysemax** $>=$ **talysemin**.

Default

talysemax 1000. MeV.

talyspath

Path for the TALYS database.

Examples

talyspath /home/koning/tendl

Range

talyspath should exist.

Default

talyspath /Users/koning/drip/

tendl

Flag to include or exclude TENDL from library average.

Examples

tendl y
tendl n

Range

y or **n**

Default

tendl y

uncertainty

Flag to check the uncertainties of the EXFOR data.

Examples

uncertainty y
uncertainty n

Range

y or **n**

Default

uncertainty y

x4

Flag to read the original EXFOR database for more complete translation statistics.

Examples

x4 y
x4 n

Range

y or **n**

Default

x4 y

xseps

Minimum cross section (mb) for TALYS + library comparison.

Examples

xseps 1.
xseps 10.

Range

1.e-10 $<=$ **xseps** $<=$ **1000**.

Default

xseps 0.1 mb.

xsonly

Flag to process only cross sections

Examples

xsonly y
xsonly n

Range

y or **n**

Default

xsonly n

Zmin

Minimal Z value to process, used for quick tests.

Examples

Zmin 92
Zmin 40

Range

0 $\leq$ **Zmin** $\leq$ **150** and **Zmin** $\leq$ **Zmax**

Default

Zmin 0

Zmax

Maximal Z value to process, used for quick tests.

Examples

Zmax 92
Zmax 40

Range

0 $\leq$ **Zmax** $\leq$ **150** and **Zmax** $\geq$ **Zmin**

Default

Zmax 150

Table 7.1: The keywords of EXFORTABLES.

Keyword	Range	Default	Page
Amax	0-400	400	52
Amin	0-400	0	53
cendl	y,n	y	54
dexp	y,n	y	55
eaf	y,n	y	56
Emax	0 - 1000.	1000.	58
Emin	0 - 1000.	0.001	57
endfb	y,n	y	59
eview	y,n	y	61
expo	y,n	y	60
filespath	filename	/exfortables/files/	62
Fmax	0 - 1.e38	1.e38	63
group	y,n	y	64
irdff	y,n	y	65
jeff	y,n	y	66
jendl	y,n	y	67
lib	y,n	y	68
libspath	filename	/libraries/	69
maxentry	0-1000000000	1000000000	71
MT	0-851	not used	70
outprocess	y,n	n	72
particle	g,n,p,d,t,h,a	g n p d t h a	73
pointcomp	all, talys, etc.	talys	74
qualitycomp	all, talys, endfb, etc.	all	75
remove	y,n	y	76
statistics	y,n	y	77
tables	y,n	y	78
talys	y,n	y	79
talysemax	0 - 1000.	1000.	81
talysemin	0 - 1000.	0.001	80
talyspath	filename	/Users/koning/drip/	82
tendl	y,n	y	83
uncertainty	y,n	y	84
x4	y,n	y	85
xseps	1.e-10 - 1000.	0.1	86
xsonly	y,n	n	87
Zmax	0-150	150	89
Zmin	0-150	0	88

8. Outlook and conclusions

This tutorial describes EXFORTABLES-2.0, a directory structured database derived from EXFOR, containing both experimental nuclear reaction data and checks on these data. All experimental data are presented in uniform x-y-dy tables, and classified according to projectile, target nucleus, and reaction. From an automation point of view, this is probably the best until a versatile EXFOR-API is developed.

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 C5 database first need to be sorted.

There are two obvious reasons to create EXFORTABLES:

- To identify problems in EXFOR and to correct them.
- To have an experimental database which is directly available to modern data evaluation using scripts and automation.

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 F-values close to 1 (where 'close' 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, and this is flagged in our database. This could be compared to other analyses of EXFOR data, e.g. via Machine Learning or simple visual inspection of the data. Hence, we can work towards 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. WPEC Subgroup 50 on an *Automatically readable, comprehensive, curated experimental reaction database* was launched in 2020 to address this issue.

Bibliography

[1] N. Otuka et al. “Towards a more complete and accurate experimental nuclear reaction data library (EXFOR): International collaboration between nuclear reaction data centres (NRDC)”. In: *Nuclear Data Sheets* 120 (2014), pp. 272–276.

[2] Arjan Koning et al. *Quality Improvement of the EXFOR database*. Tech. rep. NEA/WPEC-30, International Evaluation Cooperation, Vol. 30. NEA/NSC/WPEC/DOC, 2010.

[3] A.J. Koning et al. “TENDL: Complete Nuclear Data Library for innovative Nuclear Science and Technology”. In: *Nuclear Data Sheets* 155 (2019), p. 1.

[4] D.A. Brown et al. “ENDF/B-VIII.0: The 8th Major Release of the Nuclear Reaction Data Library with CIELO-project Cross Sections, New Standards and Thermal Scattering Data”. In: *Nuclear Data Sheets* 148 (2018), p. 1.

[5] A. J. M. Plompen et al. “The joint evaluated fission and fusion nuclear data library, JEFF-3.3”. In: *European Physical Journal A* 56 (2020), p. 181.

[6] K. Shibata et al. “JENDL-4.0: A New Library for Nuclear Science and Engineering”. In: *Journal of Nuclear Science and Technology* 48.1 (2011), pp. 1–30.

[7] Z.G. Ge et al. “The Updated Version of Chinese Evaluated Nuclear Data Library (CENDL-3.1)”. In: *J. Kor. Phys. Soc.* 59 (2011), p. 1052.

[8] A. Trkov et al. “IRDFF-II: A New Neutron Metrology Library”. In: *Special issue of Nuclear Data Sheets* 163 (2020), pp. 1–108.

[9] J.-Ch. Sublet et al. *Activation File: EAF-2010 neutron-induced cross section library*. Tech. rep. EASY Documentation Series CCFE-R (10) 05. UKAEA, 2010.

[10] A.J. Koning. *Statistical Verification and Validation of the EXFOR database: (n,g), (n,n’), (n,2n), (n,p), (n,a) and other neutron-induced reaction cross-sections*. Tech. rep. OECD/NEA Data Bank report NEA/DB/DOC(2017)1. 2017.

[11] A. Trkov, M. Herman, and D.A. Brown. *ENDF-6 Formats Manual, Data Formats and Procedures for the Evaluated Nuclear Data Files ENDF/B-VI, ENDF/B-VII and ENDF/B-VIII*. CSEWG Document ENDF-102, Report BNL-203218-2018-INRE, SVN Commit: revision 215. 2012.

[12] A.J. Koning and D. Rochman. "Modern nuclear data evaluation with the TALYS code system". In: *Nuclear Data Sheets* 113 (2012), p. 2841.

[13] AJ Koning. "Bayesian Monte Carlo method for nuclear data evaluation". In: *The European Physical Journal A* 51.12 (2015), p. 184.

A. Yet Another Nuclear Data Format: YANDF

YANDF stands for 'Yet Another Nuclear Data Format'. As it name suggests, it is inspired by the YAML markup language, as YAML is the serialization format which in my view is closest to being human and computer readable at the same time. YANDF is an attempt to unify the nuclear data structure for data sets which come from either TALYS, ENDF nuclear data libraries, EXFOR or other basic nuclear data sources. The format is aimed to be relatively simple. Once data are stored in YANDF, processing can be done independently from the particular source of the data into JSON, ENDF, GNDS etc. data formats. This may be useful for processing of data from any of these categories for numerical operations, plotting, ML applications etc. The main reason for its construction was a consistent output for the TALYS nuclear model code to enable easier processing towards various applications, but at the same time we took ENDF and EXFOR along in the process. Also, it is used for a compilation of all resonance parameters. The serialization of YANDF has a key-value schema which is not as non-descriptive as ENDF and not as heavy as GNDS. It aims to give a compact description of a nuclear reaction in terms of metadata. The source of the data, TALYS, ENDF or EXFOR may give rise to some keywords which are different, but in general the structure is the same for each of these categories. The metadata and associated keywords are supposed to be a direct classification of a nuclear reaction as defined by nuclear physics, such as found in textbooks or journal articles. This means that nuclear reaction observables are leading in the description, and not the ENDF format with its MF/MT numbers and neither the EXFOR format with its emphasis on experimental methods and details. As the role of a nuclear model code is to provide an estimate of all nuclear reaction observables as commonly defined in nuclear physics, we have taken TALYS as the basis for the schema. As EXFOR stores experimental observables, the step from EXFOR to YANDF is almost only a format change: the main keywords are the same as the one of TALYS and all experimental details are stored in their original format for the moment. We note that we only focus on the actual data, not the complete metadata of EXFOR, which can be obtained from the original EXFOR entries. (The main challenge for EXFOR is to determine

which EXFOR entries correspond to the reaction channels as defined in ENDF or in TALYS output files. Several different EXFOR categories of data may have to be included or excluded.) In ENDF libraries, only some data are observables (cross sections, nubar), but most of the data need to be processed into observables using operations on data in different parts of the data file.

Before we describe the format in more detail we give 3 examples of a typical YANDF file, for the same nuclear reaction ^{235}U(n,f). This is the file coming from TALYS

```
# header:
#   title: U235(n,f) cross section
#   source: TALYS-2.0
#   user: Arjan Koning
#   date: 2023-11-24
#   format: YANDF-0.1
# target:
#   Z: 92
#   A: 235
#   nuclide: U235
# reaction:
#   type: (n,f)
#   ENDF_MF: 3
#   ENDF_MT: 18
# datablock:
#   quantity: cross section
#   columns: 2
#   entries: 24
##       E             xs
##     [MeV]          [mb]
   1.000000E-11   0.000000E+00
   2.530000E-08   0.000000E+00
   2.000000E-07   0.000000E+00
   1.000000E-06   3.049857E+05
   1.000000E-05   9.592149E+04
.....
```

This is the file for the ENDF-B/VIII.0 data library

```
# header:
#   title: U235(n,f) cross section
#   source: ENDF
#   user: Arjan Koning
#   date: 2023-11-26
#   format: YANDF-0.1
# endf:
#   library: endfb8.0
#   author: IAEA CIELO Collaboration
#   year: 2017
# target:
#   Z: 92
```

```
#   A: 235
#   nuclide: U235
# reaction:
#   type: (n,f)
#   Q-value [MeV]:  1.934054E+02
#   E-threshold [MeV]:  1.000000E-11
#   ENDF_MF: 3
#   ENDF_MT: 18
# datablock:
#   quantity: cross section
#   columns: 4
#   entries: 333
##       E             xs            xslow          xsup
##     [MeV]          [mb]           [mb]           [mb]
   1.000000E-11   0.000000E+00   0.000000E+00   0.000000E+00
   2.250000E-03   0.000000E+00   0.000000E+00   0.000000E+00
   2.250000E-03   2.634378E+03   2.397892E+03   2.870864E+03
   2.250014E-03   2.668097E+03   2.428584E+03   2.907610E+03
   2.250056E-03   2.769988E+03   2.521328E+03   3.018648E+03
....
```

This is the file for one of the experimental data sets in EXFOR

```
# header:
#   title: U235(n,f) cross section
#   source: EXFOR
#   user: Arjan Koning
#   date: 2023-09-18
#   format: YANDF-0.1
# exfor:
#   author: Moore
#   year: 1978
#   subentry: 10629004
#   X4 reaction: 92-U-235(N,F),,SIG
#   X4 source: IAEA-NDS C5 file, database version 2023-07-18
#   X4 link: https://nds.iaea.org/EXFOR/10629004
# target:
#   Z: 92
#   A: 235
#   nuclide: U235
# reaction:
#   type: (n,f)
#   ENDF_MF: 3
#   ENDF_MT: 18
# datablock:
#   quantity: cross section
#   columns: 5
```

```
#   entries: 3777
##       E              dE             xs             dxs        Normalization
##     [MeV]          [MeV]          [mb]           [mb]             []
   1.625000E-06   0.000000E+00   1.304000E+04   1.539000E+02   1.000000E+00
   1.675000E-06   0.000000E+00   1.256000E+04   1.510000E+02   1.000000E+00
   1.725000E-06   0.000000E+00   1.270000E+04   1.549000E+02   1.000000E+00
   1.775000E-06   0.000000E+00   1.226000E+04   1.527000E+02   1.000000E+00
.....
```

Obviously, for EXFOR we have several files which in metadata only differ from TALYS or ENDF in the **exfor** keyword. The metadata in the above files completely defines the U235(n,f) reaction.

A.1 Format

The YANDF format is almost equal to the well-known YAML format. If the '#' is removed from the first columns of the metadata header of the above file, we almost have a YAML file. The difference is that we do not quote strings and that the data can be given in multi-column format. This means that indentation of the key-value pairs is essential, which is the price that YAML pays for not having to include computational symbols such as '', '[', ',' as in e.g. JSON The above files show the most general keywords without any indentation, while sub-keywords are indented by two spaces, subsub-keywords by 4 spaces, etc. As there are many users who want to use numerical data directly from the file, as in gnuplot or other software, we have chosen to use a '#' at the start of every metadata line. It is not too difficult to remove the '#' and parse the above file to JSON with either your own script or, as we already verified, with an AI assistant.

A YANDF file only contains what is relevant. For example, in the above case there is no specification of any isomeric level in either the target or residual nucleus. Hence, we have decided to leave all 'inactive' metadata out. Parsers will have to take this feature into account.

A.2 Keywords and values

The main keywords should be general enough to describe nuclear reaction observables from at least TALYS, ENDF or EXFOR, but also for additional quantities such as e.g. level densities, photon strength functions, radioisotope yields, etc. as written by TALYS to output files. Also, compiled data for e.g. thermal neutron cross sections, resonance parameters, Maxwellian-averaged cross sections etc. fit well into this schema.

A.2.1 header

All YANDF files start with the same keywords:

- **header**:
  - **title**: the title, generally constructed from the other metadata, enabling the user to see directly which nuclear reaction this concerns
  - **source**: the source of the datafile, this is often TALYS, ENDF or EXFOR. If calculated uncertainties and covariance data are available, this can also be another source, like e.g. TASMAN
  - **user**: the name of the person who produced this file (e.g. in TALYS you can change the hard-wired name into your own)
  - **date**: the date of the production of this file in yyyy-mm-dd format
  - **format**: version of the YANDF format

A.2.2 endf

When the source is an ENDF library, we have the keywords

- **endf**:
  - **library**: one of the NDL's such as ENDFB8.1, JENDL5.0, JEFF3.3, TENDL-2023, CENDL3.2 etc,
  - **author**: the author of the evaluation as extracted from the ENDF file
  - **year**: the year of the evaluation as extracted from the ENDF file

A.2.3 exfor

When the source is EXFOR, we have the keywords

- **exfor**:
  - **author**: first author of the experimental work
  - **year**: the year of the publication of the measurement
  - **subentry**: the EXFOR subentry number
  - **X4 reaction**: the particular EXFOR reaction code as extracted from EXFOR, for checking purposes
  - **X4 source**: the version of EXFOR, and the specific computational form of starting database
  - **X4 link**: the https link to the EXFOR subentry, for all experimental details
  - **level energy [MeV]**: only for discrete levels: the level energy as given by EXFOR

A.2.4 target

The first part of a nuclear reaction specification is the target nucleus.

- **target**:
  - **Z**: the charge number
  - **A**: the mass number
  - **nuclide**: the nuclide name

The above keywords are always present. In addition, isomeric level information can be provided by the **level** keywords described below.

A.2.5 reaction

The nuclear reaction may have several keywords for a complete description.

- **reaction**:
  - **type**: the nuclear reaction channel
  - **Q-value [MeV]**: the Q-value (only specified when appropriate)
  - **E-threshold [MeV]**: the incident energy threshold (only specified when appropriate)
  - **ENDF_MF**: the ENDF MF number for specification of the type of data
  - **ENDF_MT**: the ENDF MT number for specification of the reaction channel

A.2.6 residual

Often, but not always, a nuclear reaction leads to a well-defined residual nucleus.

- **residual**:
  - **Z**: the charge number
  - **A**: the mass number
  - **nuclide**: the nuclide name

The above keywords are always present when **residual** is present. In addition, isomeric level information can be provided by the **level** keyword described below.

A.2.7 datablock

Before we read the data, we need to know what we are reading and in what format.

- **datablock**: Description of the data block that follows below.
  - **quantity**: the physical quantity that we are reading
  - **columns**: the number of columns
  - **entries**: the number of entries

Below these keywords always follow 2 lines starting with '##', one with the quantities and the other one with the units.

A.2.8 Keyword: level

The **level** keyword describes the data of a discrete level. It can appear as an keyword under

- **target**, when the target is in an isomeric state
- **reaction**, for scattering off a discrete level
- **residual**, when the residual nuclide is in an isomeric state, or for gamma-ray transitions between discrete states

It is described by

- **level**: Description of discrete level
  - **isomer**: the isomeric number
  - **number**: the level number
  - **energy [MeV]**: the level energy
  - **spin**: the level spin
  - **parity**: the level parity
  - **half-life [sec]**: the half life

In general, the indentation for **level** is 2 spaces, i.e. one below the main keyword, but for discrete level gamma-ray transitions the final level is specified at 4 spaces.

A.2.9 parameters

This is a TALYS-specific keyword. It contains the nuclear models and parameters used in the calculation. It starts with

- **parameters**:

after which various parameters can be given. Here is an example for a level density output file

```
# parameters:
#   ldmodel keyword: 5
#   level density model: Hilaire-Goriely tables
#   Nlow: 8
#   Ntop: 17
#   ctable: -1.621100E-01
#   ptable: -5.763700E-01
```

A.2.10 observables

This is a TALYS-specific keyword. It contains the observables estimated by TALYS used in the calculation. It starts with

- **observables**:

after which various observables can be given. Here is an example for a level density output file

```
# observables:
#   experimental D0 [eV]:  1.200000E+01
#   experimental D0 unc. [eV]:  1.300000E+00
#   theoretical D0 [eV]:  1.245919E+01
#   Chi-2 D0:  1.247688E-01
#   C/E D0:  1.038266E+00
```