Quality Improvement of the EXFOR database

WPEC subgroup 30

Arjan Koning
NRG Petten, the Netherlands
Contents

* EXFOR
* Objective
* Problems and challenges
* Plan
* A short history of SG-30
* List of problems
EXFOR

* Most complete nuclear reaction database:
  – More than 15,500 experiments
  – Practically complete for neutrons
  – More and more complete for charged particles
* Maintained/updated by the Data Centers and compilers
* An important treasure for nuclear science and technology.
Possible issues for improvement

* Completeness (not considered here)
* Database management and EXFOR format (not considered here, apart from format harmonization)
* Data Retrieval (completeness)
* Quality (correctness)
Objective

* Make EXFOR an easy accessible and correct database, available in computational format. This enables:
  – More efficient data evaluation
  – Easy and extensive validation of model codes
  – More feedback from users to EXFOR maintainers.
EXFOR
"MOTHER"
DATABASE

LOSS

COMPUTATIONAL
DATABASE

LOSS

PRIVATE
DATABASE

X4toC4 (IAEA, BNL)
NEA’s code
JANIS (?)

At least 2 versions:
- BNL–IAEA
- NEA

Forrest–Kopecky (EASY)
Koning (directory structured)
FZK
Directly in code (EMPIRE)
Problems

* Too much flexibility for compilers to use the EXFOR format?
  – Various nuclear reactions are stored with 2 or more different identifiers (format harmonization needed!)
  – Not all data can be consistently processed into normal x-y-dy format
  – Many entries unprocessable for current database conversion codes
    Are some data “lost” forever?

* Two main problems:
  – Errors in, and wild grow of, format
  – Errors in values themselves
Why is it a problem?

* All other aspects of nuclear data evaluation are well automated:
  1. Robust nuclear model codes
  2. Plotting software
  3. Checking, processing software
  4. ‘Scriptwise’ nuclear data evaluation
* Exp. data retrieval should not become the delaying factor
* Evaluator needs access to all data
* Covariance evaluation requires access to all data
* Experimentalist deserves appropriate credit.
**Approach**

* Use format converters and checking routines
  - Record how many % is converted
  - E.g. check for negative cross sections
  - E.g. check for xs > 4 barns for $E > 0.1$ MeV
* Some errors are obvious just by plotting the values
* Use the power of nuclear model codes:
  - Chi-square checks
  - Visual inspection
See www.talys.eu

SG-30, IAEA, October 10-11, 2007
Deliverables

* EXFOR database in computational format. Annual, or more frequent, releases in increasing quality and completeness (also for “mother” database)
  – Responsible: Data Centers
  – Input: Users

* Final report:
  – Status in 2007
  – Description of improvements and tools
  – Final status in 2009
Milestones

* 6 months:
  – Collect and compare all software that processes EXFOR(X4toC4, JANIS, etc.) and merge this into one strategy for conversion into computational format.
  – Correction of EXFOR with first lists of errors (Forrest, Koning, etc.)

* 12 months:
  – Computational library #1: All cross sections
Milestones

* 24 months:
  – Computational library #2: All cross sections + angular distributions + single- and double-differential energy spectra + everything else
Conclusions

* Steps to take:
  – Make the entire EXFOR database available in computational format (First step!!!)
  – Repair the errors
* This will enable:
  – Much more efficient evaluation of data files
  – Minimal delay between publication of experiments and their adoption in data files
  – More efficient nuclear model code validation
A short history of SG-30

* April 20 2007: SG-30 approved at WPEC meeting (compromise: remove quality flagging)
* June 2007: Initial emails sent:
  – Lots of moral support
  – Mailing list created: sg30@nea.fr
  – About 30-40 members of mailing list
* June/July 2007: Extended C4 format by V. Zerkin. Bilateral communication with A. Koning
A short history of SG-30

* July 2007: Correction/investigation of AK’s first list of errors by S. Dunaeva and O. Schwerer.
* September 2007:
  – Statistical tests by Emmeric Dupont
  – Preparing JANIS for SG-30 by NEA
  – October 10-11 2007: first SG-30 meeting at IAEA
A list of problems

* Obvious (?) dimensional errors: barns instead of millibarns, eV instead of MeV, etc.
* More than one identifier to store data: e.g. (n,inl) and (n,x)0-NN-1, etc. (to be solved in X4 or X4toC4?)
* Reporting (n,inl) as (n,n’gamma) data and vice versa.
* Storage of fission yields as (n,f) (MF3,MT18) data, e.g. entries O0777 and O0020
* Ratio’s given as cross sections (e.g. entry 21863)
  • Incomprehensible reaction strings in X4 (e.g. entry O1004, Heinz 2003, 1 GeV p + U238)
A list of problems

• How to identify the level for (n,inl) to a specific level automatically?
• High energy proton reactions: residual products given as many, many subsections
• X4toC4 translation problems if uncertainties are suddenly missing inside a data block
• Total cross sections labeled as isomeric and vice versa.
• Change of nuclear reaction inside a data block and NO change of sub ID (e.g. entry O0290)