

ARTICLE

Benchmark of Spallation Models

Jean-Christophe DAVID^{1,*}, Detlef FILGES², Franz GALLMEIER³, Mayeen KHANDAKER⁴, Alexander KONOBAYEV⁵, Sylvie LERAY¹, Guenter MANK⁴, Alberto MENGONI⁴, Rolf MICHEL⁶, Naohiko OTUKA⁴ and Yair YARIV⁷

¹CEA, Centre de Saclay, IRFU/Service de Physique Nucléaire, F-91191 Gif-sur-Yvette, France

²FZJ, IKP, Postfach 1913, D-52425, Jülich, Germany

³ORNL, P.O. Box 2008, MS-6466 Oak Ridge TN 37831-6466, USA

⁴Department of Nuclear Sciences and Applications, IAEA, P.O. Box 100 A-1400 Vienna, Austria

⁵INR, Karlsruher Institut für Technologie, Postfach 3640, 76021 Karlsruhe, Germany

⁶ZSR, Hannover University, D-30419 Hannover, Germany

⁷Soreq NRC, 81800 Yavne, Israel

Spallation reactions play an important role in many branches of physics. If the nuclear reaction induced by low energy neutrons as in nuclear power plants or research reactors can be based on cross section database, it is not the case for spallation reactions due to the energy range, from 100 MeV up to ~3 GeV, and to the candidate targets (all nuclei). Then these reactions have to be modeled. Developments of codes are linked to new experiment and this explains why fifteen years ago a first benchmark has been launched, restricted to particle production, followed by a second one on residues. Since then new data have been measured and computer tools are improved day after day, so a new benchmark has been organized, under the auspices of the IAEA, taken into account all types of results, particle and residue production, with a possibility to update the data (experimental and code calculation) via a web site. Results of the benchmark concerning both the analysis of the prediction capabilities of the seventeen model-participants and the first conclusions on the physics of spallation models are presented.

KEYWORDS: Monte Carlo, spallation, benchmark, neutron sources, intra-nuclear cascade models, evaporation-fission models

I. Introduction

Spallation reactions play an important role in a wide domain of applications: meteorite origin, simulation of detector set-ups in nuclear and particle physics experiments, radiation protection near accelerators or in space, neutron sources for basic science, material irradiation or ADS, radioactive ion beam, hadrontherapy, etc. The computing tools developed for the simulation of spallation-based facilities are high-energy transport codes in which nuclear model codes are used to compute the production yields and characteristics of all the particles and nuclei generated in these reactions. The nuclear model codes are generally Monte-Carlo implementations of Intra-Nuclear Cascade (INC) or Quantum Molecular Dynamics (QMD) models followed by de-excitation (principally evaporation/fission) models.

Two former benchmarks of spallation models were organized some fifteen years ago. The first one was limited to particle (neutron and proton) production¹⁾ while the second one was devoted to excitation functions of residue production.²⁾ At that time, very large discrepancies between the model-participants were found. These exercises also pointed out the necessity of more reliable and complete experimental data, in particular above 800 MeV and for residue production.

Since that time, in Europe, an important effort has been

devoted to the collection of high quality experimental data^{3,4)} concerning all emission channels, neutrons, light charged particles and residues. Meanwhile, model developers have worked a lot to improve existing models or propose new ones based on the newly measured experimental data. It was therefore the appropriate time to organize a new benchmark of the models developed by different groups in the world with respect to a complete set of experimental data. This was done under the auspices of the International Atomic Energy Agency (IAEA). The objectives were, first, to assess the prediction capabilities of the spallation models used or that could be used in high-energy transport codes in the different mass and energy regions or for the different exit channels; second, to understand the reason for the success or deficiency of the models so that some consensus on the most appropriate physics ingredients of the models could be reached.

The main technical aspects of this benchmark are briefly presented in Section II, and Section III is devoted to the global analysis (models / experimental data).

II. Benchmark Specifications

A first workshop was held in Trieste in 2008 in order to agree on the benchmark specifications. It gathered model developers who presented the physics ingredients of the models⁵⁾ that could participate to the exercise and experimentalists specialized in spallation data measurements. The

*Corresponding author, E-mail: jean-christophe.david@cea.fr

participants agreed on a comprehensive set of experimental data to be calculated by the models, which covers all the emission channels (residues, neutrons, light charged particles from proton to alpha, and pions), an energy range from 20 to 3,000 MeV and different target masses. The chosen observables were generally double differential cross sections (DDXS) for particle production, but average multiplicities and multiplicity distributions have been added for neutrons. For residue production, the selected data include isotopic cross-sections, mass and charge distributions but also excitation functions. Different targets have been considered but with focus on Pb and Fe.

The results from the seventeen calculations, representing either different models or different combinations of intranuclear cascade and deexcitation models, have been received and compared to the selected set of experimental data. All the results were uploaded on a dedicated web site.⁶⁾ Tools (Fortran codes and Perl script) have been developed to draw all figures (around 10,000). Different deviation factors or figures of merit⁷⁾ have also been calculated and added on the website in order to help a more quantitative analysis.

Finally in February 2010 a second workshop devoted to the benchmark analysis was held in Saclay, during which global analyses of the results were presented and model developers summarized the qualities and shortcomings of their model revealed by the benchmark and discussed possible improvements.

A final report including all analyses and conclusions will be released at the end of 2010.

III. Analysis

The global analysis of the agreement between the models and the experimental data was based on a coarse eye-guided rating of all sets of data and done independently for double differential cross sections of neutrons and light charged particles (LCPs) and for mass, charge, and isotopic residue production. The rating shown in the upper part of **Table 1** was used for neutrons and residues, while results concerning LCPs were evaluated through the ratio between experimental data and the calculation, as indicated in the lower part of Table 1.

Obviously it is impossible to show the analysis in minute

Table 1 Ratings used to analyze the benchmark results: Upper part for neutron and residues and lower part for LCPs

Quality	Points
Good	2
Moderately good. minor problems	1
Moderately bad. particular problems	-1
Unacceptably bad, systematically wrong	-2

Acceptance band [eval/x : eval*x]	Points
x=5	1
x=3	2
x=2	3
x=1.4	4

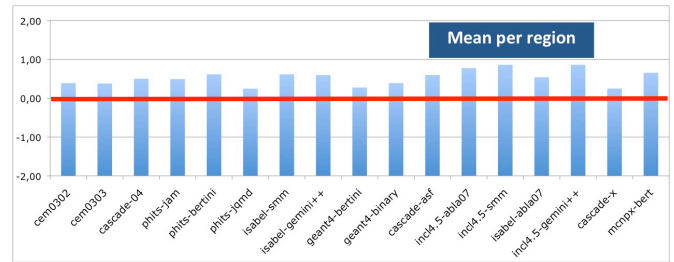


Fig. 1 Rating results obtained using the method given in the upper part of Table 1 for neutron DDXS

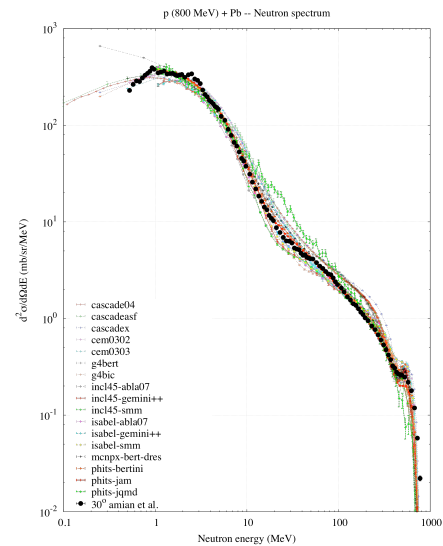


Fig. 2 Neutron spectra for the reaction p (800 MeV) + Pb (30°). All models are plotted.

detail in conference proceedings. However we give hereafter the main trends and some interesting points are mentioned to show the need of such a benchmark for code developers and end-users.

1. Neutrons

As mentioned previously, three observables have been studied for neutrons: double-differential cross-sections, average multiplicities and multiplicity distributions. The rating described in the upper part of Table 1 was used for DDXS divided into four energy bins (regions) representative respectively of the evaporation, pre-equilibrium, pure cascade and quasi-elastic regions.

Although all models can be further improved, the results (**Fig. 1**) are rather good and if some models are better than the others, the differences are not very large. Definitely, it can be said that the quality of the models has been considerably improved since the 1994 benchmark. The best results being obtained for projectile energies around GeV, where four orders of magnitude are well reproduced. (**Fig. 2**).

However, a closer look at the different data sets shows some specific discrepancies, as can be seen on examples of results displayed in **Figs. 3** and **4**.

Figure 3 shows the regions or energy bins where some models can still be improved, especially in forward direction. At very high energy the elastic peak has not always the right

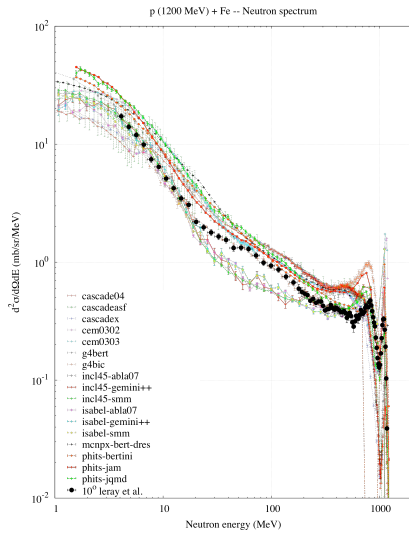


Fig. 3 Neutron spectra for the reaction p (1200 MeV) + Fe (10°). All models are plotted.

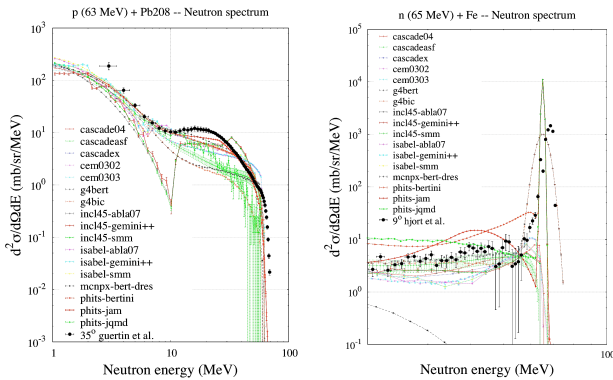


Fig. 4 Neutron spectra for the reaction p (63 MeV) + ^{208}Pb (35°) (left) and n (65 MeV) + Fe (9°) (right). All models are plotted.

height, which is also true for the quasi-inelastic peaks (given mainly by the Δ resonance), and the region around 100 MeV (pre-equilibrium) split the models in two groups: above and below the data. In this later case the question of a third phase, called pre-equilibrium, and taken into account by a specific mechanism like an exciton model, is open. The difficulty to prove the need of such a mechanism comes from the correlations that exist between the INC, pre-equilibrium and deexcitation stages.

At low incident energies, i.e. below ~ 100 MeV, the models have generally difficulties to fit all the details of the experimental data, which is not surprising since, at these energies, the physics hypotheses inherent to intranuclear cascade models are not valid (Fig. 4). However, since data libraries do not exist for all nuclei between 20 and 150 MeV, models have often to be used. Therefore it is important to check their reliability and, actually, some of them are not so bad (Fig. 4). It is interesting to note that the quasi-elastic peak (right part of Fig. 4) is only produced by two types of model, but in an ad-hoc way.

The models have also been compared to experimental average multiplicities of low (< 20 MeV) (Fig. 5) and high ($>$

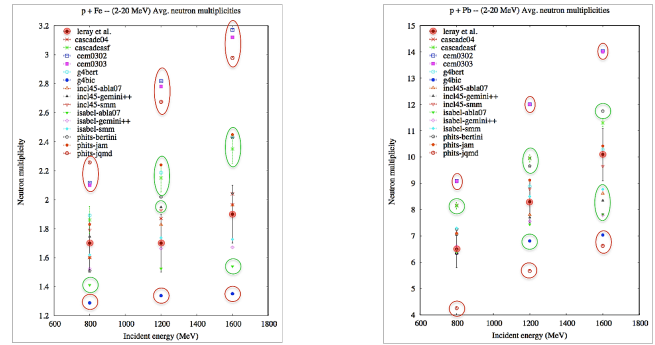


Fig. 5 Neutron average multiplicities for iron (left) and lead (right) targets at 3 energies (0.8, 1.2 and 1.6 GeV). Here are only the low energy neutrons (below 20 MeV). All models are plotted.

Table 2 Rating results obtained for light charged particle double differential cross sections

Emitted particle	Code-Data	Averaged rating
protons	2.4	i.e. within a factor 2-3
deuterons	1.2	i.e. within a factor 4
tritons	1.2	i.e. within a factor 4
He-3	0.9	i.e. within a factor 5
alpha	1.3	i.e. within a factor 4

20 MeV) energy neutrons data for two targets (Fe and Pb) at three energies (0.8, 1.2 and 1.6 GeV). Average multiplicity is a rather important quantity since it is directly related to neutron production in spallation targets. Lead is well reproduced by most the models, with a discrepancy of maximum 50% for the worst models. The situation is less good for iron, for which a larger number of models disagree with the data and the discrepancy can reach a factor 2. The conclusions are more or less the same for the high-energy average multiplicities (not shown here).

2. Light Charged Particles

Light charged particles (LCPs), i.e. proton, deuteron, triton, helium-3 and alpha, are abundantly produced in spallation reaction and are a concern for material damage issues. For instance, helium can be responsible of swelling in the structure materials in particular the window separating the target and the accelerator vacuum; tritium production is often an issue from the radioprotection point of view. Generally, the situation is different for proton production, which is rather well predicted, although less than neutron production, by most of the models, and for composite particles, for which considerable discrepancies between the different models can be found. This situation is confirmed by the global rating of the predictions of the models for the different types of particle using the procedure described in the lower part of Table 1. The results averaged over all models are given in Table 2. It shows that globally protons are reproduced within a factor 2 while LCP prediction is much more difficult. However, it has to be kept in mind that large differences exist from one model to another.

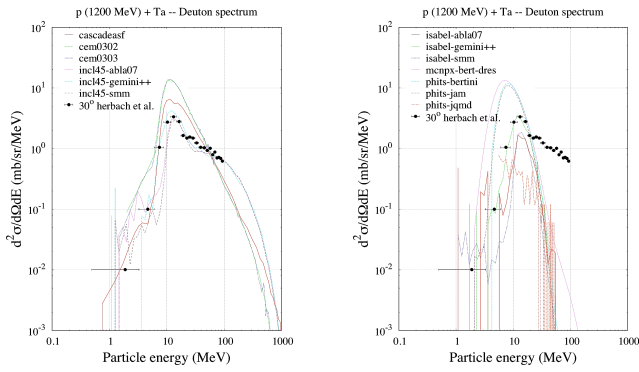


Fig. 6 Deuteron DDXS in p (1200 MeV) + Ta $\theta=30^\circ$ given by two groups of models. On the left models with a specific mechanism to emit composite particles in the INC phase and on the right models without.

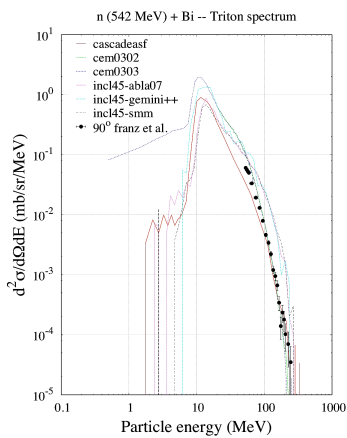


Fig. 7 Triton DDXS in n (542 MeV) + Bi at $\theta=90^\circ$ given by models with a specific mechanism to emit composite particles in the INC phase.

Indeed, one of the major difficulties of composite particle DDXS prediction concerns the high-energy tail observed in the experimental data. Actually, only models which have a specific mechanism to emit energetic clusters can reproduce this high-energy tail. **Fig. 6** shows clearly this fact with on the left models with and on the right model without a specific mechanism.

For neutron DDXS it was generally observed that a given model was behaving rather similarly for all the sets of experimental data (with some exceptions at low energies). For LCP DDXS on the contrary the same model was often found to agree rather well with some sets of data and disagree with some other sets, making it difficult to draw conclusions on the best model. For example in **Fig. 6** (left part) and **Fig. 7**, in the case of p (1,200 MeV) + Ta and n (542 MeV) + Bi, respectively, the best models are not the same.

3. Residues

Observables related to residue production are certainly the most difficult ones to predict by the models and very large discrepancies between the calculations have been observed. **Fig. 8** shows for instance the comparison of all models to the mass distribution measured in p (1 GeV) + ^{56}Fe . Some cal-

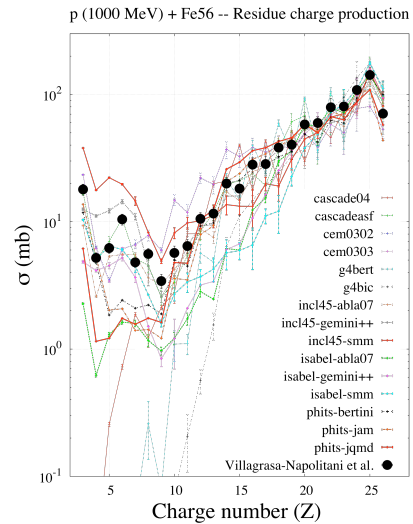


Fig. 8 Comparison of all models to the experimental charge distribution measured using reverse kinematics in the reaction p(1,000 MeV) + ^{56}Fe .

culations clearly disagree with the experimental data by sometimes orders of magnitude in some mass regions, especially for intermediate mass fragments in the case of iron and in the fission region in the case of lead (not shown here), while a few calculations give a reasonable agreement all along the mass range.

In order to look at the performance of the models, the rating of Table 1 upper part was used, distinguishing different product nuclide regimes: target-near products, spallation products with masses exceeding half the target mass, light products with masses (much) smaller than half the target mass, fission products (for lead and uranium). Separate ratings were done for mass and charge distributions on the one hand and for isotopic distributions on the other hand. It has to be emphasized that a good description of mass or charge distribution may be accompanied by deficits in the description of isotope distributions. Actually, the overall performance of the models and codes is generally worse for the isotope distributions than for mass and charge distributions. This is caused widely by the bad performance of many models for the heavy target elements at high energies.

This rating aims at giving a rough and average comparison between the models. According to the residue or group of residues studied the comparison can be different. Such detailed analyses were out the scope of this intercomparison campaign. If different mass regions were taken into account and if in other studies deviation factors or figures of merit were used to help,⁷⁾ the performance of a given model depends on what the user want to know or simulate. Some isotopes produced with low cross sections can lead to important consequences.

Results of the crude rating for the agreement between theory and experiment for the isotope distributions are given in **Fig. 9**. The maximum number of available point was 28, the minimum -28. Clearly, none of the models and codes meets all the requirements, but there are significant differences demonstrating that some codes perform much better

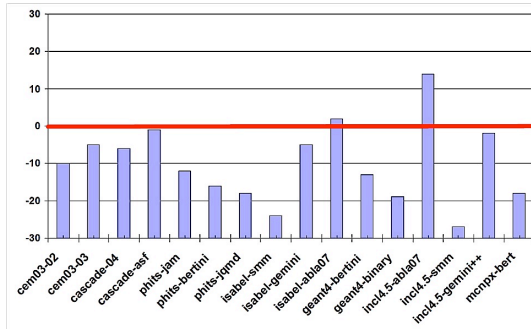


Fig. 9 Rating of the results of 16 participants for predicting the isotope distributions measured by inverse kinematics for iron, lead, and uranium at all energies (from 20 to 2,600 MeV)

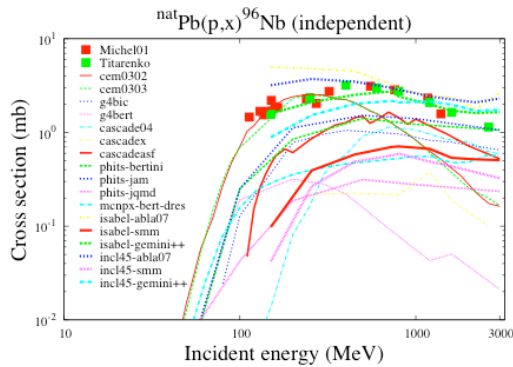


Fig. 10 Excitation function for the production of ^{96}Nb from a Pb target bombarded by protons

than others. There are also codes which generally perform badly and which need conceptual improvements. It is interesting to notice that calculations done with the same INC code coupled to different de-excitation models can obtain very different rates, showing the high sensitivity of residue production to the de-excitation stage.

The performance of models for the production of light complex nuclei ($A < 10$) could only be discussed on the basis of excitation functions in classical kinematics. Also the energy dependence can only be looked at in terms of the excitation functions. However, since the available excitation functions are biased due to our capabilities to measure the nuclides, they have to be looked at individually or exemplary. Here, a quantification using figures-of-merit or deviation factors may be helpful. Finally **Fig. 10** shows, as an example, the difficulty to reproduce an excitation function in shape and sometimes order of magnitude for some models, but also the danger to extrapolate results from a given projectile energy to another.

However, compared to the situation met in the former benchmark,²⁾ one can state that in many reaction regimes considerable progress has been made by the modelers during the past decade (**Fig. 11**).

IV. Conclusion

The new benchmark of spallation models, performed under the auspices of the IAEA, covers all the possible emission channels (neutrons, LCPs, residues and pions) with

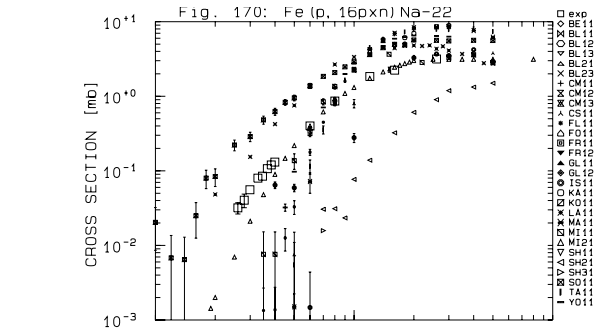
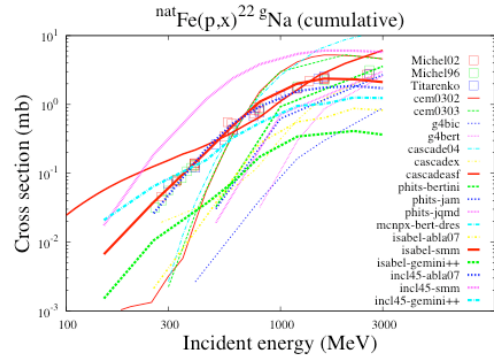


Fig. 11 Excitation function for the production of ^{22}Na from a Fe target bombarded by protons. Results of this benchmark are in the upper part and results of the previous benchmark²⁾ in the lower part.

more than 40 sets of experimental data on 10 different targets induced by nucleons with energy going from 20 to 3,000 MeV. Seventeen models or model combinations have participated to the exercise. Data, calculation results and figures-of-merit are available on a dedicated web site⁶⁾. The global analysis, presented briefly in this paper, have shown that models are globally much more reliable than they were at the time of the two previous benchmarks,^{1,2)} thanks to both the availability of new and more constraining experimental data and to the efforts of the model developers. However, there is still a lot of room for improvement, in particular for the prediction of residues and composite light charged particles. If some models seem globally better than others it has to be stressed that all models have their strength and weaknesses and because of the complexity of the spallation reaction one should be careful with any kind of extrapolation.

From the overall analysis of the results, some general conclusions about the physics of the models can be drawn. For instance, it has been found that, although the hypotheses inherent to intranuclear cascade models are not valid below 150 MeV, INC+deexcitation models give acceptable results. Therefore, they can be used in transport codes when evaluated libraries are not available and as long as one is not looking at collective or detailed structure effects. In some codes a pre-equilibrium stage is added between the INC and deexcitation stages. It can be concluded that there is no clear advantage of having this additional stage or not. On the contrary, it can be stated that a specific process (coalescence for example) is necessary to reproduce high-energy tail of LCP

spectra. The benchmark also showed that results, in particular residue production, are very sensitive to the deexcitation stage and therefore that efforts should be made to improve evaporation-fission models. Finally, the benchmark gave the opportunity to test a more sophisticated model using Quantum Molecular Dynamics (QMD), which in principle contains better physics assumption and is expected to be more reliable, in particular at low energies. Although conclusions are a bit difficult to draw because of the lack of statistics and only one QMD model was tested, it seems that there is up to now no obvious advantage of using QMD instead of intranuclear cascade models in a transport code.

This benchmark has been very successful since in particular nearly all the models currently used in high-energy transport codes have participated. However, in the future new versions of the models or new models will be developed. Since the website and the tools to analyze the results are now fully operational and since this benchmark (initiative and results) has been appreciated by several communities (developers and accelerator shielding designers to mention only them), it therefore would be rather easy and useful to continue this benchmark in a dynamical way so that end-users of spallation models in transport codes have up-to-date information. This would mean that new versions of the models or new models compared to the benchmark set of data could be

added on the website. Also, if new interesting experimental data become available, calculations of these data with the participating models could be performed. Discussions with OECD/NEA are in progress and encouraging to get a framework for this benchmark next year.

References

- 1) M. Blann, H. Gruppelaar, P. Nagel, J. Rodens, *International Code Comparison for Intermediate Energy Nuclear Data*, NSC/DOC(94)-2, NEA/OECD (1993).
 - 2) R. Michel, P. Nagel, *International Codes and Model Inter-comparison for Intermediate Energy Activation Yields*, NSC/DOC(97)-1, NEA/OECD (1997).
 - 3) J.-P. Meulders, A. Koning, S. Leray, HINDAS EU Contract FIKW-CT-00031, final report (2005); http://www.theo.phys.ulg.ac.be/wiki/index.php/Cugnon_Joseph
 - 4) EUROTRANS/NUDATRA project, EU Contract FI6W-CT-2004-516529.
 - 5) D. Filges, S. Leray, Y. Yariv, A. Mengoni, A. Stanculescu, G. Mank (Eds.), *Proceedings of the Joint ICTP-IAEA Advanced Workshop on Model Codes for Spallation Reactions*, ICTP Trieste, Italy, 4-8 February 2008, IAEA INDC(NDS)-530 (2008).
 - 6) <http://www-nds.iaea.org/spallations>
 - 7) http://nds121.iaea.org/alberto/mediawiki-1.6.10/images/5/59/Deviation_factors.doc
-