*PHYSOR 2004 -The Physics of Fuel Cycles and Advanced Nuclear Systems: Global Developments*
Chicago, Illinois, April 25-29, 2004, on CD-ROM, American Nuclear Society, Lagrange Park, IL. (2004)

# Box-Cox Transformation for Resolving the Peelle's Pertinent Puzzle in a Curve Fitting

Soo-Youl Oh* and Chul-Gyo Seo
*Korea Atomic Energy Research Institute, Yuseong, Daejeon 305-600, Korea*

Incorporating the Box-Cox transformation into a curve fitting is presented as one of methods for resolving an anomaly known as the Peelle's Pertinent Puzzle in the nuclear data community. The Box-Cox transformation is a strategy to make non-normal distribution data resemble normal distribution data. The proposed method consists of the following steps: transform the raw data to be fitted with the optimized Box-Cox transformation parameter, fit the transformed data using a conventional curve fitting tool, the least-squares method in this study, then inverse-transform the fitted results to the final estimates. Covariance matrices are correspondingly transformed and inverse-transformed with the aid of the law of error propagation.

In addition to a sensible answer to the Puzzle, the proposed method resulted in reasonable estimates for a test evaluation with pseudo-experimental $^6$Li(n,t) cross sections in several to 800 keV energy region, while the GMA code resulted in systematic underestimates that characterize the Puzzle. Meanwhile, it is observed that the present method and the Chiba-Smith method yield almost the same estimates for the test evaluation on $^6$Li(n,t). Conceptually, however, two methods are very different from each other and further discussions are needed for a consensus on the issue of how to resolve the Puzzle.

***KEYWORDS: Cross Section Evaluation, Peelle's Pertinent Puzzle, GMA, Box-Cox Transformation, Normality, Least-Squares Method***

## 1. Introduction

The least-squares method (LSM) is a well-developed, powerful tool for curve fitting along with the model parameter estimation. Nevertheless, an anomaly is observed such that the method sometimes yields strange, unacceptable estimates. In the nuclear data community, the anomaly is known as the Peelle's Pertinent Puzzle (PPP), which is quoted below from a secondary source:[1]

"Suppose we are required to obtain the weighted average of two experimental results for the same physical quantity. The first result is 1.5, and the second result 1.0. The full covariance matrix of these data is believed to be the sum of three components. The first component is fully correlated with standard error of 20% of each respective value. The second and third components are independent of the first and of each other, and correspond to 10% random uncertainties in each experimental result.

The weighted average obtained from the least-squares method is 0.88 ± 0.22, a value outside the range of the input values! Under what conditions is this the reasonable result that we sought to achieve by use of an advanced data reduction technique?"

Several studies have been devoted to investigate the reason for and resolution of the PPP.[1-5] Previous studies will not be reviewed here, but D. Smith's view[6] on this issue is briefly introduced. He showed seven distinct solutions to the PPP and noted that each method is unique in concept and each treats the available experimental information differently. He suggested solutions from the Bayesian approach as the rigorous one, in which a quite skewed, non-normal posterior probability distribution for

* Corresponding author, Tel. 42-868-2961, FAX 42-868-8341, E-mail: syoh@kaeri.re.kr

the quantity under evaluation is properly dealt with. His point is that the anomaly stems from the probability distribution of the observable that is non-normal but implicitly assumed as normal in, say, a least-squares method. Interpretations such that the PPP originates from an improper linearization[3] or from an improper treatment/calculation of the covariance of derived (in many cases of nuclear data evaluation, normalized) quantity[2,4,5] are essentially not far from Smith's viewpoint.

Unfortunately, the PPP is alive in a real world; for example, in an evaluation of $^{115}$In(n,n')$^{115m}$In cross sections.[1] Even though 'the Bayesian procedure (with due consideration on the governing probability distribution) is both appealing and rigorous in principle,' we face the question of how to deal with a real problem in a practical way because of the complexity of a real problem.[6] Chiba and Smith proposed[1] a practical, iterative procedure for resolving the PPP and implemented it in GMAJ as a derived version of the generalized least-squares method fitting code, GMA.[7] Their idea looks to work well in the evaluation, but is based on a more or less subjective interpretation of the fractional uncertainties usually provided along with the measured data. In their procedure, the absolute uncertainty of the raw data is computed as the reported fractional uncertainty multiplied by the 'true' value, not by the measured one. Without knowing the true value, the procedure needs iterations. We propose here a new procedure to properly deal with a, of implicit, non-normal distribution data.

## 2. Proposed Method for Resolving the Puzzle

### 2.1 Outline

A least-squares method (LSM) assumes a normality of residuals, *i.e.* the differences between estimates and raw data, even though the derivation of the necessary formulas does not explicitly require the normality. A distribution of the residuals far from the normality implies that the fitting model, rather than the methodology, is inappropriate to such a kind of raw data. Remember that the PPP is perceived from biased residuals. There are three options in dealing with non-normally distributed residuals: re-interpret the covariance associated with the measurement, transform the 'independent' variable (model parameter in other words), or transform the 'dependent' variable to be a suitable type for the model. The second option is equivalent to a revision of the model. The last one, while keeping the original model, is the approach proposed in this study.

The fitting procedure proposed is as follows: transform the measured data and its covariance matrix, fit the transformed data using a conventional method, then inverse-transform the estimates and associated covariance into the space of the original data. The key of the proposal is a concept dealing with 'transformed' data in a curve fitting or, equivalently, in an evaluation of the model parameters. The Box-Cox transformation[8] is utilized as the tool for making non-normally distributed data resemble normally distributed data.

As a tool for a curve fitting itself, the generalized LSM in the GMA code is adopted for tests in this study. The concept, however, can be incorporated not only into a LSM method but also into any other methods such as a usual Bayesian approach and a maximum likelihood method as well. It is worth noting that the PPP is invoked by the data discrepant and strongly correlated with each other, not by a fitting methodology. In other words, even a Bayesian approach, if it assumes a normal distribution of the observable in its formulation, is not free from the PPP. In addition to the normality assumption, an assumption of the linearity of measured quantities with respect to the model parameters makes the Bayesian method (with non-informative priors) identical to the LSM.[9] This provides another basis that the concept of a variable transformation can be applied to a usual Bayesian method.

### 2.2 Box-Cox Transformation and Associated Formulas

Box and Cox proposed a transformation of a dependent (response) variable $y$ to $w$ by

$$w_i(\lambda)=\begin{cases}\left(y_i^{\lambda}-1\right)/\lambda, & \lambda\neq 0\\ \ln y_i, & \lambda=0\end{cases}\quad,\quad i=1,\ldots,N, \tag{1}$$

which makes the probability density function of $w$ rather close to the normal distribution.[8] The transformation is performed for all $N$ data points of a vector $\mathbf{y}=(y_1\, y_2 \ldots y_N)^t$ to the vector $\mathbf{w}$. The transformation parameter $\lambda$ is determined to maximize the log-likelihood function

$$\ln L(\lambda)=-\frac{N}{2}\ln\left[\frac{1}{N}\sum_{i=1}^{N}(\hat{w}_i-w_i)^2\right]+(\lambda-1)\sum_{i=1}^{N}\ln y_i\ , \tag{2}$$

where $\hat{w}_i$ is the estimate of $w_i$. A numerical solver, usually a grid search method, is used in determining $\lambda$ since $L(\lambda)$ in Eq. (2) is a recursive function of $w$. The optimum $\lambda$ is usually searched for in the range of [–2, 2]. Note that $\lambda=1$ implies no transformation in fact.

$\mathbf{V}_y$, the covariance matrix associated and provided along with $\mathbf{y}$, is transformed to $\mathbf{V}_w$ with the aid of the law of error propagation as follows:

$$\mathbf{V}_w=\mathbf{S}\,\mathbf{V}_y\,\mathbf{S}^t, \tag{3}$$

where $\mathbf{S}$ is a diagonal sensitivity matrix whose $(i, i)$ element is computed as

$$S_{i,i}=\frac{\partial w_i}{\partial y_i}=y_i^{\lambda-1}. \tag{4}$$

Then, a curve fitting method, the generalized (weighted, in other word) least-squares method (GLSM) in this study, yields the estimate $\hat{\mathbf{w}}$ and its associated covariance matrix $\mathbf{V}_{\hat{w}}$. The estimate $\hat{\mathbf{w}}$ is easily inverse-transformed into $\hat{\mathbf{y}}$ by

$$\hat{y}_i=\begin{cases}\left(\hat{w}_i\lambda+1\right)^{1/\lambda}, & \lambda\neq 0\\ exp(\hat{w}_i)\,, & \lambda=0\end{cases}\ . \tag{5}$$

The inverse-transformation of $\mathbf{V}_{\hat{w}}$ to $\mathbf{V}_{\hat{y}}$ is performed similarly to Eq. (3) as

$$\mathbf{V}_{\hat{y}}=\mathbf{T}\,\mathbf{V}_{\hat{w}}\,\mathbf{T}^t, \tag{6}$$

where the $(i, i)$ element of the diagonal matrix $\mathbf{T}$ is computed as the derivative, $\partial\hat{y}_i/\partial\hat{w}_i$.

## 3. Test Evaluations

### 3.1 Solutions to the Peelle's Puzzle

3.1.1 Prerequisite: the Generalized Least-Squares Method

Before we advance, formulas for the GLSM are summarized below. With respect to a model defined as

$$\mathbf{y}=\mathbf{G}\,\mathbf{p}+\boldsymbol{\varepsilon}\ ,\ \boldsymbol{\varepsilon}\sim N(\mathbf{0},\mathbf{V}\sigma^2)\,, \tag{7}$$

the GLSM provides the estimate of the model parameter vector $\mathbf{p}$ and its associated covariance matrix as

$$\hat{\mathbf{p}}=\left(\mathbf{G}^t\mathbf{V}_y^{-1}\mathbf{G}\right)^{-1}\mathbf{G}^t\mathbf{V}_y^{-1}\,\mathbf{y}\ \text{ and} \tag{8}$$

$$\mathbf{V}_{\hat{p}}=\left(\mathbf{G}^t\mathbf{V}_y^{-1}\mathbf{G}\right)^{-1}. \tag{9}$$

$\mathbf{G}$ is the design matrix, $\mathbf{y}$ the measured data vector, and $\mathbf{V}_y$ the covariance matrix of $\mathbf{y}$. Then the estimate $\hat{\mathbf{y}}$ is obtained as

$$\hat{\mathbf{y}} = \mathbf{G}\,\hat{\mathbf{p}}\,, \tag{10}$$

while its covariance matrix $\mathbf{V}_{\hat{\mathbf{y}}}$ is calculated from the law of error propagation as

$$\mathbf{V}_{\hat{\mathbf{y}}} = \mathbf{G}\,\mathbf{V}_{\hat{\mathbf{p}}}\mathbf{G}^t\,. \tag{11}$$

3.1.2 GLSM Solution

With a single parameter $p$ representing the average of two data, the PPP is modeled as

$$\mathbf{y} = \mathbf{G}\,\mathbf{p} + \boldsymbol{\varepsilon} = \begin{pmatrix}1\\1\end{pmatrix}\cdot p + \boldsymbol{\varepsilon}\quad . \tag{12}$$

Measured data vector is $\mathbf{y} = (1.5\quad 1.0)^t$, and $\mathbf{V}_\mathrm{y}$ is given as

$$\mathbf{V}_\mathrm{y} = \begin{pmatrix}1.5^2(0.1^2+0.2^2) & 1.5\cdot 0.2\times 1.0\cdot 0.2\\ 1.5\cdot 0.2\times 1.0\cdot 0.2 & 1.0^2(0.1^2+0.2^2)\end{pmatrix} = \begin{pmatrix}0.1125 & 0.06\\ 0.06 & 0.05\end{pmatrix}.$$

Then the parameter $p$ is estimated from Eqs. (8) and (9) as $\hat{p}$=0.882±0.218. Since the model $\mathbf{G}$ is a unit vector, the final estimate $\hat{y}$ ( $=\hat{y}_1=\hat{y}_2$ ) is equal to $\hat{p}$, which is outside the range of the raw data of 1.5 and 1.0.

3.1.3 Solution by the Chiba-Smith Procedure

The Chiba-Smith procedure is based on an assumption that the absolute uncertainties are computed in terms of 'fractional' uncertainties from the experiment and the 'true' value.[1] Instead of $\mathbf{V}_\mathrm{y}$, the GLSM deals with $\mathbf{V}_\mathrm{y}^*$ such that

$$\mathbf{V}_\mathrm{y}^* = \begin{pmatrix}y_0^2(0.1^2+0.2^2) & y_0\cdot 0.2\times y_0\cdot 0.2\\ y_0\cdot 0.2\times y_0\cdot 0.2 & y_0^2(0.1^2+0.2^2)\end{pmatrix} = y_0^2\begin{pmatrix}0.05 & 0.04\\ 0.04 & 0.05\end{pmatrix},$$

where $y_0$ is the true (unknown) value. Only one iteration is needed with any initial guess on $y_0$ because it is canceled in the evaluation of $\hat{p}$ using Eq. (8). The Chiba-Smith solution is $\hat{p} = 1.250 \pm 0.265$.

Even though this solution does not differ too much from the Bayesian solution of 1.21 ± 0.30 that Smith gave, it is argued that the procedure yields a most probable value instead of a mean value.[6]

3.1.4 Solution by the Proposed Procedure

Let the Box-Cox solution mean the GLSM solution with the Box-Cox transformation hereafter. The model is the same as Eq. (12), but the data that the GLSM deals with is different. For the Puzzle, the optimum value of the transformation parameter, which maximizes the log-likelihood function of Eq. (2), is zero. (See Table 1.) Using this value, data $\mathbf{w}$ and $\mathbf{V}_\mathrm{w}$ are computed as

$$\mathbf{w} = \begin{pmatrix}\ln y_1\\ \ln y_2\end{pmatrix} = \begin{pmatrix}0.405\\ 0.0\end{pmatrix} \text{ and } \mathbf{V}_\mathrm{w} = \begin{pmatrix}1/y_1 & 0\\ 0 & 1/y_2\end{pmatrix}\mathbf{V}_\mathrm{y}\begin{pmatrix}1/y_1 & 0\\ 0 & 1/y_2\end{pmatrix} = \begin{pmatrix}0.05 & 0.04\\ 0.04 & 0.05\end{pmatrix}.$$

The GLSM with the above $\mathbf{w}$ and $\mathbf{V}_\mathrm{w}$ yields the model parameter $\hat{p} = 0.203 \pm 0.212$ so that

$$\hat{\mathbf{w}} = \begin{pmatrix}1\\1\end{pmatrix}\cdot 0.203 \text{ and } \mathbf{V}_{\hat{\mathbf{w}}} = \begin{pmatrix}1\\1\end{pmatrix}0.212^2\,(1\ \ 1) = \begin{pmatrix}0.212^2 & 0.212^2\\ 0.212^2 & 0.212^2\end{pmatrix},$$

then the inverse-transformed final estimate is obtained as

$$\hat{\mathbf{y}} = \begin{pmatrix}1\\1\end{pmatrix}\cdot 1.225 \text{ and } \mathbf{V}_{\hat{\mathbf{y}}} = \begin{pmatrix}e^{0.203} & 0\\ 0 & e^{0.203}\end{pmatrix}\begin{pmatrix}0.212^2 & 0.212^2\\ 0.212^2 & 0.212^2\end{pmatrix}\begin{pmatrix}e^{0.203} & 0\\ 0 & e^{0.203}\end{pmatrix} = \begin{pmatrix}0.260^2 & 0.260^2\\ 0.260^2 & 0.260^2\end{pmatrix}.$$

This value of $\hat{y} = 1.225 \pm 0.260$ looks reasonable and is the same to the solution that Smith gave[6] as Method 2.

Interpret the proposed method applied to the Puzzle. Note that $\lambda = 0$ implies a logarithmic transformation. Normalization is regarded as the very origin of the PPP in most relevant studies, thus taking a logarithm that eliminates the non-normality due to the quotient seems to be reasonable and promising.

Table 1 shows $\ln L(\lambda)$ and $\hat{y}$ varying with $\lambda$. Note that $\lambda = 1$ results in the same solution from the usual LSM, as was expected because $\lambda = 1$ implies no transformation in effect.

**Table 1** Log-Likelihood Function and Box-Cox Solution to the PPP Varying with λ

| $\lambda$ | $\ln L(\lambda)$ | $\hat{y}$ |
|---|---|---|
| -1.0 | 1.62 | 1.700 ± 0.421 |
| -0.5 | 2.16 | 1.475 ± 0.327 |
| -0.1 | 2.75 | 1.273 ± 0.271 |
| **0.0** | **2.79** | **1.225 ± 0.260** |
| 0.1 | 2.75 | 1.178 ± 0.250 |
| 0.5 | 2.16 | 1.017 ± 0.225 |
| **1.0** | 1.62 | **0.882 ± 0.218** |
| 1.5 | 1.44 | 0.776 ± 0.229 |

Table 2 presents solutions from various solutions with the same $y$ values but different uncertainties. Note that just scaling up or down both statistical and systematic uncertainties together does not alter the GLSM solution. Meanwhile, the estimate by the Chiba-Smith procedure does not change at all regardless of the magnitude of the uncertainties. With $\lambda = 0$, the estimate from the present procedure does not change either regardless of the uncertainties. It might be understood that the transformation parameter value of zero makes the impact of the inconsistency and the strong correlation between the raw data on the estimate ineffective.

**Table 2** Various Solutions to the Peelle's Puzzle with Different Data Uncertainties*

| Uncertainty (%) | | Generalized LSM (GMA method) | Chiba-Smith Procedure | Present Procedure** |
|---|---|---|---|---|
| Systematic | Statistical | | | |
| 20 | 20 | 1.071 ± 0.278 | 1.250 ± 0.306 | 1.225 ± 0.300 |
| **20**† | **10**† | **0.882 ± 0.218** | **1.250 ± 0.265** | **1.225 ± 0.260** |
| 10 | 10 | 1.071 ± 0.139 | 1.250 ± 0.153 | 1.225 ± 0.150 |
| 10 | 20 | 1.132 ± 0.202 | 1.250 ± 0.217 | 1.225 ± 0.212 |
| 10 | 5 | 0.882 ± 0.109 | 1.250 ± 0.133 | 1.225 ± 0.130 |
| 2 | 1 | 0.882 ± 0.022 | 1.250 ± 0.027 | 1.225 ± 0.026 |

* The rigorous Bayesian solution is 1.21 ± 0.30. [6]
** Optimum $\lambda$ for Box-Cox transformation is zero for all cases.
† Uncertainties in the original Peelle's Puzzle

### 3.2 Test Evaluation of $^{6}$Li(n,t) Cross Section

The cross sections of $^{6}$Li(n,t) reaction were evaluated using the GMA code, Chiba-Smith method, and present method, respectively. In all three methods, the fitting tool itself is the same, *i.e.*, the GLSM utilized in GMA. Five pseudo-experimental, but realistic data sets were utilized[10]: Lamaze *et al.* 1978, Fort and Marquette 1972, Fort 1970, Poenitz and Meadows 1972, and Friesenhahn *et al.* 1974. These sets were prepared for investigating the PPP during the IAEA Coordinated Research Project on the evaluation of the standard cross sections of light elements. Each set of Lamaze and of Friesenhahn includes data over the whole energy range from 2.5 keV to 800 keV, but these two sets are discrepant from each other. The Friesenhahn set, which is a shape data set originally, seems to cause the Peelle's problem in this evaluation.

Figure 1 compares the results from three methods, which are conceptually different from each other but utilize the same tool in the curve fitting itself. While the GMA solution runs below almost all the experimental data points, the solution from the present procedure (say, Box-Cox solution) passes through the data points. The optimum $\lambda$ for the transformation was determined as –0.07. Throughout the entire energy region, the Box-Cox cross sections are consistently higher than the GMA cross sections by 10~15 % and the Chiba-Smith's (denoted by GMAJ in the Figure) by 0.5~1.5%. Note that the solution by the Chiba-Smith procedure is not well distinguished from the Box-Cox solution in the Figure. Even so, there is a big philosophical difference: when inconsistent, correlated data sets are encountered during the evaluation, the Chiba-Smith approach supposes some misunderstandings in the reported uncertainties, but the present approach supposes an improper dimension or scale of the measured quantity. Similarly to the discussion in the case of the Peelle's problem, the optimum value of the transformation parameter near zero suggests that it is natural to expect the Gaussian behavior of the logarithmic transformed quantity rather than that of the reported quantity.

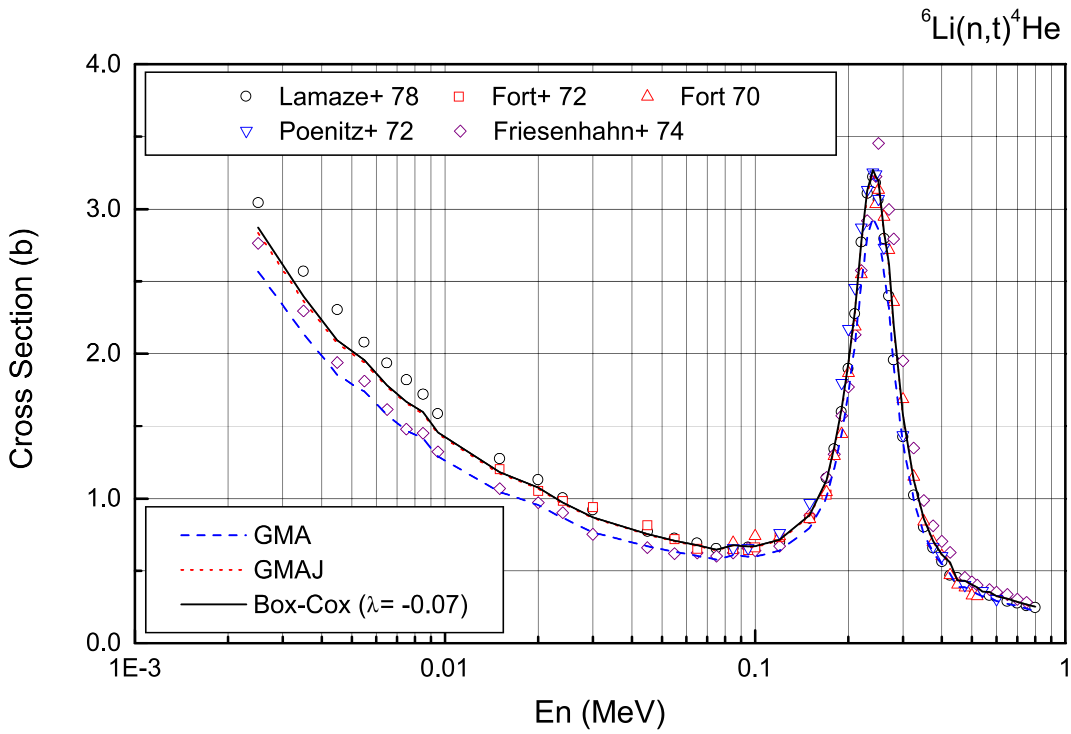


**Fig. 1** Evaluated $^6$Li(n,t) Cross Sections

Concerning the uncertainties of the evaluated cross sections (no detailed numbers are provided in this paper), the percent uncertainties in the Box-Cox cross sections are consistently smaller than those from the GMA by 2~6% and are comparable with those from the Chiba-Smith. However, the absolute uncertainties of the Box-Cox solution are larger than the GMA's.

Regarding the 'correlation' between the evaluated cross sections at different energy points, no significant difference is found between the three solutions except that the Box-Cox correlations are slightly larger than those from the GMA. In the case of 'covariance', however, the Box-Cox covariances are 12~15% larger than the GMA's very consistently.

The $\chi^2$ value per degree of freedom increases from 9.7 in the GMA to 11.2 in the Box-Cox solution with optimum $\lambda$ and to 11.0 with $\lambda = 0$, as was expected at the beginning. It is reminded, however, that the $\chi^2$ is less meaningful when the governing probability distribution deviates from the normal distribution.

Intuitively one may adopt zero for $\lambda$ to take a logarithm. It is equivalent to linearizing a quantity in a form of a quotient or multiplication of two or more sub-quantities. Cross section from a ratio measurement is an example of such quantity. However, in this test evaluation on the $^6$Li(n,t) reaction, the Box-Cox procedure with $\lambda = 0$ results in cross sections smaller than those with the optimum $\lambda = -0.07$ by 1% consistently throughout the entire energy region. Very likely, a theoretician may prefer zero to other values near zero. Nevertheless, the 1% difference in the evaluation is not negligible, so it is recommended to search for an optimum value that might be near zero but not zero. A two-digit precision might be enough for most applications including the $^6$Li(n,t) evaluation.

Figure 2 shows one of the sensitivity calculation results. Only two data sets, Lamaze and Friesenhahn, are utilized for the evaluation. The original systematic uncertainties are 1.6% and 2.7% for the Lamaze cross sections and Friesenhahn's, respectively. Each set of the evaluated cross sections from the GMA and Box-Cox procedure with these two sets is not very different from those in Figure 1, respectively. Now let a test problem assume 10% systematic uncertainties for each experiment set. As shown by a dotted line in the lower part of the figure, GMA resulted in meaningless cross sections. However, the present procedure, with the optimum $\lambda = -0.002$, produced reasonable estimates again. Indeed, it is not distinguished from the Box-Cox solution with the original uncertainty data (optimum $\lambda = -0.04$); the differences in the two sets of the evaluated cross sections are less than 0.3%. In short, with a $\lambda$ near zero, the proposed procedure is insensitive to given data uncertainties.

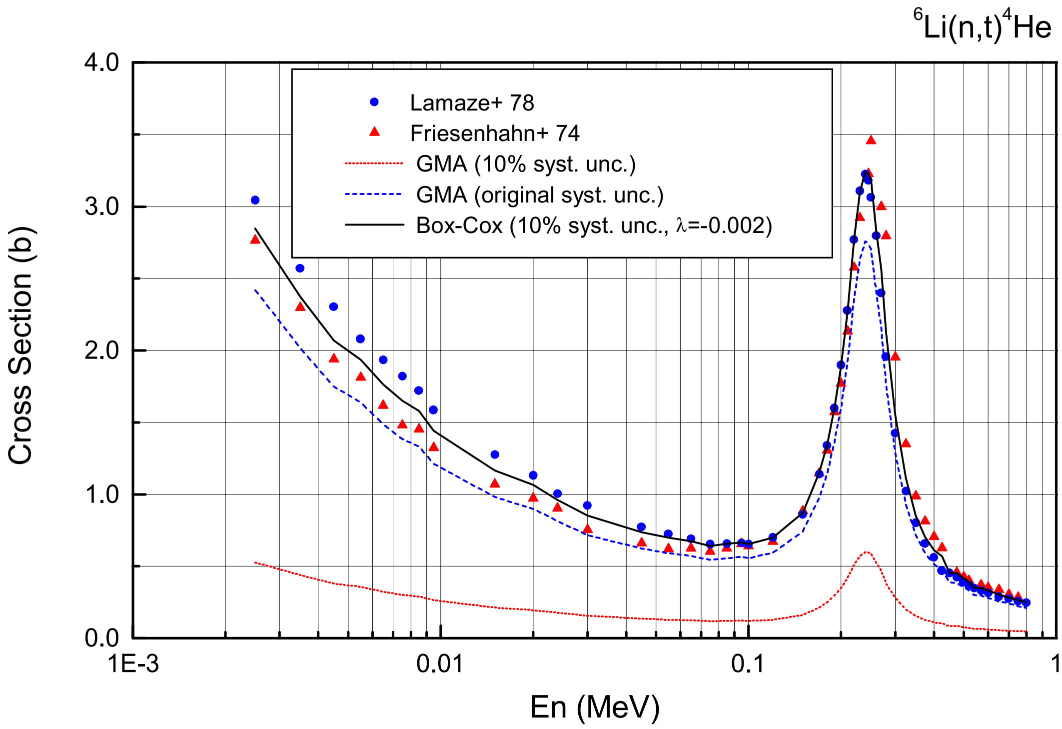


**Fig. 2** Effect of the Magnitude of Systematic Uncertainty on the Test Evaluation of $^6$Li(n,t) Cross Sections

## 4. Conclusion and Remarks

For resolving the Peelle's Pertinent Puzzle, a method was proposed that adopts the Box-Cox transformation in combination with a usual curve fitting method. The method when incorporated into a generalized least-squares method shows a good performance in a test, but realistic evaluation of the $^6$Li(n,t) reaction cross section as well as in the Peelle's Puzzle.

In addition to the conclusion above, several findings, along with future works, are summarized as follows.

(1) The degree of impact of the inconsistency and the correlation between the raw data on the resulting estimate is controlled by the magnitude of $\lambda$ in the Box-Cox transformation. For instance, a logarithmic transform makes the estimate insensitive to the magnitude of the uncertainties in the PPP.

(2) The optimum $\lambda$ is rather small, but not zero, in the $^6$Li(n,t) cross section evaluation. Taking zero for $\lambda$ may look reasonable and even physical for, for instance, normalized cross sections. However, blindly taking zero is not recommended. On the other hand, a further study is needed to deal with data types other than the type of an absolute cross section, such as the sum of several cross sections.

(3) When an evaluator encounters strange estimates, he/she will review the raw data at hand as a first action. However, doing that is beyond the role of the method proposed here: the method deals with the data sets as they are given, neither adds something to the raw data nor interprets given data subjectively. Besides the Chiba-Smith approach, Badikov's systematic search[11] for the uncertainties unrecognized, thus excluded from the reporting might be one of the resolutions for the PPP. Such an approach is the other version of the first option mentioned in Section 2.1 of this paper.

**Acknowledgements**

This study was done under the auspices of the IAEA Research Contract No. 12025, year 2002.

**References**

1) S. Chiba and D. L. Smith, "A Suggested Procedure for Resolving an Anomaly in Least-Squares Data Analysis Known as Peelle's Pertinent Puzzle and the General Implications for Nuclear Data Evaluation," ANL/NDM-121, Argonne National Lab. (1991).
2) Z. Zhao and F. G. Perey, "The Covariance Matrix of Derived Quantities and Their Combination," ORNL/TM-12106, Oak Ridge National Lab. (1992).
3) F. H. Froehner, "On assignment of uncertainties to scientific data," Extended version of a paper in Proc. of Int'l Conf. on Reactor Physics and Reactor Computations, p. 287, Tel Aviv, Jan. 23-26, 1994 (1994).
4) S. Chiba and D. L. Smith, "Impacts of data transformations on least-square solutions and their significance in data analysis and evaluation," J. Nucl. Sci. Technol., **31**, 770 (1994).
5) G. D'Agostini, "Probability and Measurement Uncertainty in Physics – a Bayesian Primer," DESY 95-242, Univ. "La Sapienza" and INFN, Rome (1995).
6) D. L. Smith, "Probability, Statistics, and Data Uncertainties in Nuclear Science and Technology," American Nuclear Society, La Grange Park, U.S.A., p.205 (1991).
7) W. P. Poenitz and S. E. Aumeier, "The Simultaneous Evaluation of the Standards and Other Cross Sections of Importance for Technology," ANL/NDM-139, Argonne National Lab. (1997); PSR-367, Radiation Safety Information Computational Center, Oak Ridge National Lab.
8) G. E. P. Box and D. R. Cox, "An analysis of transformations," J. Royal Statistical Soc., Series B, **26**, 211 (1964).
9) N.M. Larson, "Proof that Bayes and least squares give exactly equivalent results for arbitrary number of data sets assuming linearity," in "Summary Report of the First Research Co-ordination Meeting on Improvement of the Standard Cross Sections for Light Elements," INDC(NDS)-438, IAEA, p.91 (2003).
10) V. G. Pronyaev, "Test and inter-comparisons of data fitting with general least squares code GMA versus Bayesian code GLUCS," *ibid*., p.159 (2003).
11) S. A. Badikov and E. V. Gai, "Some Sources of the Underestimation of Evaluated Cross Section Uncertainties," *ibid*., p.172 (2003).