Estimation of North Korea's uranium production using Bayesian network

Yein Seo[a], Yonhong Jeong and SeungHo Jeong[*]
[a]*Center for Nuclear Nonproliferation Strategy & Technology, Korea Institute of Nuclear Nonproliferation and Control, 1418 Yuseong-daero, Yuseong-gu, Daejeon 34101, Republic of Korea*
[*]*Corresponding author: shjeong@kinac.re.kr*

1. Introduction

Uncertainty in estimating the total amount of the uranium used for nuclear weapons is one of the most difficult problems in the denuclearization verification. Especially, in the case of North Korea (NK), there is a lack of information about the amount of nuclear materials including enriched uranium (U) and plutonium (Pu) due to its secrecy. And even that, it could hardly be confirmed [1].

The Bayesian network (BN) would be one of the effective tools to estimate the range of the nuclear materials production in NK, considering the uncertainty. BN has the ability to make a fairly acceptable estimation providing a probability distribution, based on the imperfect information through the chain of the causal relationship.

Several existing studies have tried to estimate the amount of U production in NK using a BN model [2, 3]. However, the main focus of these researches is the capacity of the U enrichment and the production of Pu was not included. Also, there is little chance to find evidence related to the number of the gas centrifuges in NK which is one of the major variables in these studies. U enrichment facilities are relatively easier to be concealed than the others in the nuclear fuel cycle (NFC). Thus, in order to improve the estimation of NK's uranium production, it is necessary to take account of the other facilities and processes in the NFC together with Pu.

On the other hand, Von Hippel proposed the novel methodology to calculate NK's nuclear materials production including low-enriched uranium (LEU) and Pu in the whole NFC [4]. It evaluates NK's capacity of producing the natural U, which means the refined U such as a yellow cake in this paper, and U ores in two ways, 'Top-down' and 'Bottom-up'. Top-down starts to calculate from estimated inventories of enriched uranium (EU) and Pu, but Bottom-up does from the U mining capacity. In addition, variables and minimum, median, and maximum values for them in the estimation process were provided.

In this study, the total cumulative amounts of the U ores and natural U produced in NK were estimated through a BN model considering LEU, high-enriched uranium (HEU), and Pu in the whole NFC with reference to the 2-way (Top-down, Bottom-up) method. Also, this paper is a summarized and revised version with some additional contents of an analysis report under review [5].

2. Methods and Results

In this section, first, the basic concept of BN is explained. And methods used to establish the BN model structure and to assign states and probability are described. After quantification, the results demonstrate the estimation ability and applicability of the constructed BN model.

2.1 Bayesian Network

BN is one of the probability graphical models that probabilistically present the causal relationships among variables. It can both diagnose the reason of the events retrospectively and forecast the results prospectively through updating belief based on the Bayes' theorem expressed as an equation (1).

$$P(A_i|B) = \frac{P(A_i)P(B|A_i)}{\sum_{i=1}^{N} P(A_i)P(B|A_i)} \quad (1)$$

BN consists of nodes, links, and conditional probability table(CPT)s. Nodes represent the variables, links express the causal relationship between nodes, and CPTs contain the quantitative information of the probability.

In terms of estimation on the nuclear materials production, BN has the following advantages. First, it can show the uncertainty in the estimation in the form of the probability distribution. And its chain structure of causal relationships helps to derive the quite reasonable results from the imperfect information. Next, all kinds of data including quantitative statistics and qualitative expert opinions can be reflected in the BN model. It is also easy to modify, learn, and update models through new information. Lastly, its graphics and results can provide some help for information-based decision making.

2.2 Structure Development

The BN model in this study was basically developed based on Von Hippel's method of U production estimation. Variables used in Von Hippel's calculation were adopted to the nodes [4]. Links connecting nodes were drawn, reflecting the relationship between variables assumed in the independent reproduction of mentioned calculation [6].

The BN model comprises 11 parts to achieve the advantages in the quantification. There are 3 'Top-down' parts that estimate the amount of U ores and natural U, going through reprocess, enrichment, conversion, etc.

from Pu, LEU, and HEU, respectively. 6 'Bottom-up' Parts dividing 6 periods from 1945 to 2018 consider mining capacity and working ratio, export ratio of ore, U quality, and efficiency of processing. And 2 parts, one aggregates all top-down parts and the other does bottom-up parts, show the cumulative production results calculated in each way.

2.3 Assignment of States and Probabilities

For each node, the range of the values was divided by several intervals and states were assigned. Minimum, medium, and maximum values proposed in the existing research were referred to set the range of values [4].

In the cases that variables directly indicate the amount of nuclear material or their range of the value is relatively wide compared to the scale of the value, for example, U quality and mining capacity, the corresponding nodes were divided into 10 states, exceptionally 14 states for the amounts of produced natural U. The range of values was selected to include from zero to the maximum, and the intervals of the states are regular in scale.

The other nodes whose range of value is relatively narrow, such as tail assay and losses in the conversion, fabrication, and process, obtain 3 states, 'Low', 'Mid', and 'High'. The range from the minimum to the maximum value was uniformly divided into 3 equal parts to assign to the interval of each state.

As shown in Fig. 1, the marginal probability distribution of root nodes with no link pointing them was assumed to be a step distribution with the probability of 0.5 between the minimum and median, and between the median and maximum, respectively. But when only one real value was suggested for the minimum, median, and maximum, it was assumed that a corresponding value was put into all states and the only mid state had a probability of 1.0. The conditional probability for non-root nodes, variables with parent nodes, was entered by calculation formulas according to the relationship between variables. Through sampling based on these probability distributions and equations, the probability of each state was determined according to its interval.

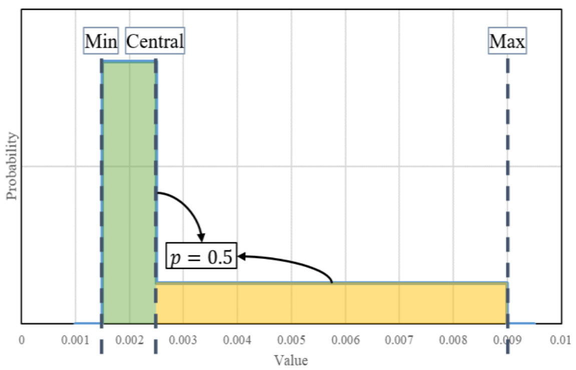

Fig. 1. Assumption on the probability distribution of variables by using a minimum, medium, maximum value.

2.4 Results of Quantification

Fig. 2 shows a part of the model after quantification. The estimated probability distribution of total U production in NK is displayed in Fig. 3. As a result, the amount of natural U estimated by top-down method was the average of 375, the median of 360, and the standard deviation of 142 tons. And the amount of U ore was estimated as the average of 2.31E+5, the median of 2.13E+5, and the standard deviation of 1.28E+5 tons. On the other hand, in the case of the bottom-up method, the amount of natural U was figured out to be the average of 1,411, the median of 1,382, and the standard deviation of 442 tons. Also, the amount of the U ore was calculated to be the average of 4.88E+5, the median of 4.98E+5, and the standard deviation of 1.09E+5 tons.

Table I tabulates the ratio of the estimated U production in this study to that in the existing study by Von Hippel [4]. Note that the minimum and the maximum value in this study are the 95% confidence interval. The results estimated by the BN model showed significantly different values, 0.32 ~ 2.89 times the results in the existing research. Except for the minimum value of natural U by top-down method, there was a tendency that minimums increased and maximums decreased. In other words, the range of the estimation became narrow.

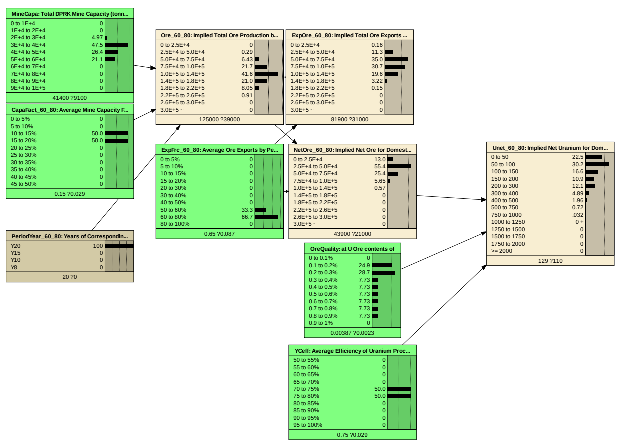

Fig. 2. Part of the quantified model which estimates from mining capacity between 1960 ~ 1980.

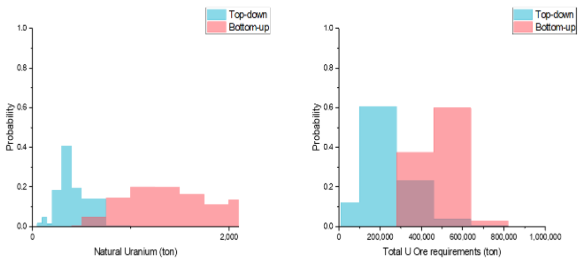

Fig. 3. Probability distribution of the total cumulative amount of nuclear materials produced in North Korea

Table I: Ratio of Results on Uranium Production to Existing Estimation

Type of Results	Estimating Method	Ratio of Results		
		Min	Median	Max
Natural U	Top-down	0.59	0.91	0.91
	Bottom-up	2.89	1.66	0.32
U Ore	Top-down	1.19	1.21	0.95
	Bottom-up	1.43	1.12	0.68

One of the reasons for the narrower range was excluding the almost unlikely estimate. In the existing estimation, the probability in the range was not considered at all. But results in this study were provided as the probability distribution, which could make it possible to exclude the values out of the 95% confidence interval with little chance.

The conjunction of the probability distributions in BN caused the change of median values and allowed more accurate calculation. It was guessed that the existing study assumed simply calculated value using the median of each variable as the median of the results. However, in this study, the median was obtained in the true sense of the word by performing the conjunction of the probability distribution.

Last, uniform sampling over the state interval was considered to be the cause of the exceptional decrease in the minimum value of natural U in top-down estimation. In the BN computation software used, samples are generated with the same probability within the interval of each state in the process of converting the input equation to the CPT. It could have the effect of distorting the calculation results, especially when the slope of the distribution is steep or at the end of the distribution.

It is discouraged to grant great significance to the number of results themselves because this model is based on insufficient data and preliminary for future work. However, it would be meaningful in that it has been confirmed that it can provide the probability distributions for results and narrow the range of the estimation, unlike the previous studies. It not only increases the reliability of the estimation, but also supports future improvement by grasping the accuracy and the precision of the current estimation. Besides, for the 2-way method with top-down and bottom-up, the likelihood that a true value exists in the corresponding area could be expected to be higher if the overlapping area with two distributions goes wider.

3. Conclusions

In this paper, the amount of North Korea’s U production was estimated, using the Bayesian network (BN) model developed based on the existing top-down and bottom-up method. It was confirmed that BN could present the uncertainty of estimation as the probability distribution and support to decrease it.

In order to improve the model and estimation, it is needed to increase the reliability of the prior probability for variables by supplementary literature surveys for further work. Also, the ability and applicability of the model should be verified by case studies with various scenarios. The results of this study are expected to be utilized to compensate for the uncertainty of the dynamic simulation model for estimating North Korea’s nuclear material production under development.

ACKNOWLEDGEMENT

This work was supported by the Nuclear Safety Research Program through the Korea Foundation Of Nuclear Safety (KoFONS) using the financial resource granted by the Nuclear Safety and Security Commission (NSSC) of the Republic of Korea (No. 1905008).

REFERENCES

[1] J.-H. Lee, J.-s. An, and I. J. Park, "Considerations for Verification of North Korea's Nuclear Program," Korea Institute of Nuclear Nonproliferation and Control, Daejeon. KINAC/TR-008/2011, 2011.

[2] J. E. Bistline, D. M. Blum, C. Rinaldi, G. Shields-Estrada, S. S. Hecker, and M. E. Paté-Cornell, "A Bayesian Model to Assess the Size of North Korea's Uranium Enrichment Program," Science & Global Security, vol. 23, no. 2, pp. 71-100, 2015.

[3] M. Croy et al., "North Korea WMD Assessment & Verification Project," in Briefing to KINAC/INSA, Daejeon, South Korea, 2019.

[4] D. F. Von Hippel, "Methods for Refining Estimates of Cumulative DRPK Uranium Production," Journal for Peace and Nuclear Disarmament, vol. 2, no. 2, pp. 555-585, 2019/07/03 2019.

[5] Y. Seo, S. Jeong, and Y. Jeong, "A Study on Application of Bayesian Network to Non-proliferation Analysis for Nuclear Fuel Cycle," Korea Institute of Nuclear Nonproliferation and Control, Nuclear Safety Technology Analysis Report. KINAC/TR-022/2020, TBD.

[6] S. Jeong, J.-H. Lee, Y. Jeong, H. Do, and Y. Seo, "Builing a System Dynamics Model for Estimating Cumulative Uranium Production form Denuclearization Perspective," Korea Institute of Nuclear Nonproliferation and Control. KINAC/TR-016/2020, 2020.