

VERIFYING NORTH KOREA'S NUCLEAR PROGRAMME: TECHNICAL OPTIONS AND POLICY CHOICES

James M. Acton¹

Carnegie Endowment for International Peace
1779 Massachusetts Ave NW, Washington DC, 20036

ABSTRACT

This paper considers the challenges and policy choices associated with verifying a North Korean declaration about its past nuclear activities. In terms of the plutonium programme, there are two key choices facing policy makers. First, in light of North Korea's reported statement that its October 2006 test involved only 2 kg of plutonium, is the IAEA's standard definition of a significant quantity appropriate? Second, using open source information, this paper argues that the uncertainty in the material unaccounted for, $\sigma(\text{MUF})$, will be between 5 and 10 kg, depending on the degree of cooperation afforded by North Korea. In order to ensure that verification produces a conclusive answer, the magnitude of $\sigma(\text{MUF})$ will necessitate some combination of decreasing the detection probability and increasing the false alarm rate from their standard IAEA values, or adopting entirely different criteria against which to assess North Korean compliance. The paper also includes a discussion of the process for verifying the absence or existence of a clandestine uranium enrichment programme by formulating it as a Bayesian inference problem. This framework, together with an analysis of the politics of inspections at undeclared locations suggests that, unless detailed intelligence pinpointing the location of a clandestine centrifuge facility is available, inspection rights at undeclared facilities may be of limited utility.

INTRODUCTION

Verifiable nuclear disarmament continues to be the principal goal of US policy towards North Korea. In spite of the recent collapse of the denuclearization process, culminating in North Korea's second nuclear test on 25 May 2009, the need to design and eventually implement a verification scheme has not evaporated. This paper considers, quantitatively, the challenges posed by verification in order to highlight the choices faced by policy makers.

There are two reasons for considering the design of a verification scheme well in advance of when it is needed. First, with a clear understanding of what verification can and cannot achieve, those tasked with negotiating with North Korea will be better placed to negotiate an effective agreement.

Second, as argued below, if standard International Atomic Energy Criteria (IAEA) safeguards criteria are applied, verification may be inconclusive. Even if North Korea acts in good faith, submits a complete and correct declaration and makes every effort to facilitate verification, it may be impossible for inspectors to verify with confidence that it has not, for example, secreted away a small stockpile of plutonium. Conversely, if North Korea has violated its obligations and retained some nuclear material illicitly, it could equally prove impossible for inspectors to prove non-compliance.

¹ E-mail: jacton@ceip.org

In the event that verification produces inconclusive results, it is likely that it will be the subject of intense arguments between different agencies in the US government, in Congress and the media, and between the states involved in the Six Party Process. These controversies are likely to be exacerbated if North Korea insists—as it has done to date—that the IAEA have no role in verifying its past nuclear activities, thereby eliminating the obvious ‘honest broker’.

Arguments over verification have the potential to be acrimonious and damaging. Although they are to some extent inevitable, they can be minimised if all parties are able to agree on procedures for dealing with ambiguities or on a verification standard that avoids the possibility of inconclusive results. To this end, there is an advantage in the US seeking to engage and educate all stakeholders—domestic and international—about the challenges of verification and trying to build agreement, in advance, on a way forward.

VERIFYING NORTH KOREA’S PLUTONIUM PROGRAMME

Even if the IAEA is to have no role in verification, its standard material accountancy criteria could still be adopted for verifying North Korea’s plutonium programme and are the natural starting point for an analysis of verification. The basis for any system of materials accountancy is the comparison of a state’s calculated inventory to its measured holdings. In the following, the quantity of plutonium produced by North Korea in its 5 MWe reactor and possibly in its IRT research reactor (both at Yongbyon) is denoted by PB. If North Korea has not retained any of this material clandestinely, its current holdings, PI, should be equal to this quantity less the material used in its two tests, T, together with any reprocessing (or other) losses, R. In practice, there is always some difference, known as the material unaccounted for (MUF), and given by

$$\text{MUF} = \text{PB} - \text{PI} - \text{T} - \text{R}.$$

The challenge facing inspectors is to decide whether the measured MUF can be attributed to measurement errors or is the result of a diversion. Whether the MUF is considered statistically significant depends on three parameters that must be chosen by the inspectorate:

- **Significant quantity (SQ):** The minimum quantity of diverted nuclear material that the safeguards system is designed to detect. The IAEA sets the SQ for plutonium to 8 kg.
- **Detection probability (1-β):** The probability that the diversion of 1 SQ of nuclear material will be detected. Where high confidence is required the IAEA sets 1-β = 0.9.
- **False alarm probability (α):** The probability that a diversion of nuclear material will be detected even though none has occurred. The IAEA typically chooses α = 0.05.

Given these values and assuming a Gaussian error model, it can be shown that the MUF should be considered statistically significant if it is larger than some detection threshold, $D = \phi^{-1}(1-\alpha)\sigma(\text{MUF})$, where $\sigma(\text{MUF})$ is the uncertainty in the MUF and $\phi^{-1}(z)$ is the inverse standard normal cumulative distribution function [1]. With α = 0.05, $D = 1.6\sigma(\text{MUF})$. If the same detection threshold, D, is also to ensure that the detection probability is 1-β (or better) then a second condition must also be met: $\sigma(\text{MUF}) \leq \text{SQ}/[\phi^{-1}(1-\alpha)-\phi^{-1}(\beta)]$. For the standard IAEA values of α, 1-β and SQ, this translates to $\sigma(\text{MUF}) \leq 2.7$ kg.

There are two potential problems with applying this standard to North Korea. First, the IAEA defines an SQ as ‘the approximate amount of nuclear material for which the possibility of manufacturing a nuclear explosive device cannot be excluded’. North Korea has, however, reportedly claimed that the nuclear device it exploded in October 2006 contained only 2 kg of plutonium [2]. If this report is correct, a key choice facing inspectors is whether to use the standard value for an SQ or a lower one.

The second problem arises from the likely magnitude of $\sigma(\text{MUF})$, which is almost certain to be larger than about 5 kg—almost twice the size of the value it would need to be for standard IAEA safeguards parameters to be applied. The different contributions to $\sigma(\text{MUF})$ are shown in Table 1 and justified in more depth in the Appendix at the end of this paper. The estimate of $\sigma(\text{MUF})$ assumes, inter alia, that (i) North Korea eventually gives inspectors access to all separated plutonium (or else there is little point doing any verification frankly); (ii) that no more fuel is irradiated; and (iii) that North Korea does not perform any more nuclear tests. If assumptions (ii) or (iii) turn out to be wrong—and they may well be— $\sigma(\text{MUF})$ can be recalculated straightforwardly but will necessarily increase. The single most important factor in determining $\sigma(\text{MUF})$ is whether North Korea permits samples to be taken from the moderator of the 5 MWe reactor and analysed (something to which it objected in 2008 and 2009, according to press reports).

Source of error	Upper bound (kg Pu)	Lower bound (kg Pu)
Pu production, $\sigma(\text{PB})$	9	3
Current holdings, $\sigma(\text{PI})$	0	0
Usage in tests, $\sigma(\text{T})$	4	2
Reprocessing wastes, $\sigma(\text{R})$	6	3
$\sigma(\text{MUF})$	10	5

TABLE 1: Contributions to $\sigma(\text{MUF})$ for North Korea’s plutonium programme (all quantities are given to one significant figure and justified in the Appendix to this paper).

That $\sigma(\text{MUF}) > 2.7$ kg poses a serious problem for the decision making process in verification. This problem does arise in other contexts, such as applying IAEA safeguards to large bulk handling facilities (most notably, Japan’s Rokkasho Reprocessing Plant). In such facilities, however, additional containment and surveillance can be used to enhance confidence in non-diversion. This is, however, clearly not an option when verifying past nuclear activities, such as North Korea’s.

The fact that $\sigma(\text{MUF}) > 2.7$ kg means that there are some values of the MUF for which it is impossible to reach a conclusion about compliance with confidence. That is, there exists a range of MUF values that have more than a 10% probability of being observed under the assumption that North Korea has diverted one SQ of material *and* more than a 5% probability of being observed under the assumption that no diversion has taken place. The inconclusive MUF values lie in the range $D_1 < \sigma(\text{MUF}) < D_2$ where

$$D_1 = \text{SQ} + \phi^{-1}(\beta)\sigma(\text{MUF}) \text{ and } D_2 = \phi^{-1}(1 - \alpha)\sigma(\text{MUF}).$$

In fact, it may even be possible that the ambiguous situation is the most likely outcome. Using the standard values of α and $1-\beta$, Figure 1 shows the probability, p , of the results of verification being inconclusive for both the standard value, $SQ = 8$ kg, and a reduced value, $SQ = 2$ kg. In order to calculate, p , it is necessary to make an assumption about whether there has been a diversion (blue curves) or not (red curves). Since it is impossible to know in advance which assumption is correct these can be thought of as lower and upper bounds on p .

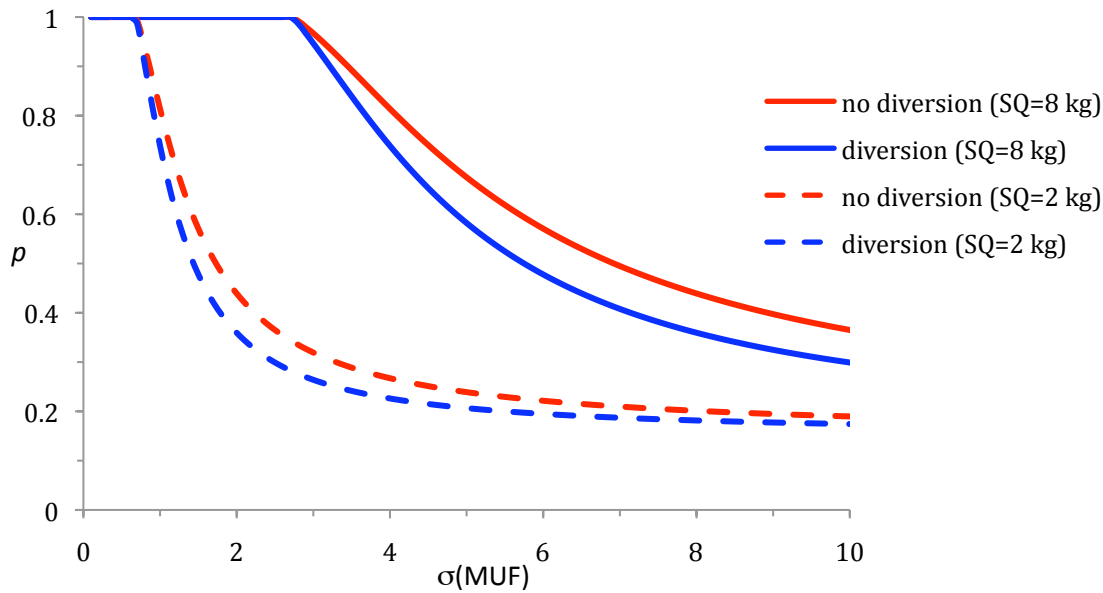


FIGURE 1: Plot of the probability, p , that standard IAEA safeguards parameters will lead to a conclusive result as a function of $\sigma(\text{MUF})$ result for various assumptions.

It can be seen from the graph that even if $\sigma(\text{MUF})$ is as small as it could credibly be at 5 kg then the probability of a conclusive result is between 58% and 67% when $SQ = 8$ kg. If the SQ is instead set to 2 kg then this probability drops to only about 20%. (In passing, it is interesting to note that as $\sigma(\text{MUF}) \rightarrow \infty, p \rightarrow \alpha + \beta$).

The second key choice faced by inspectors is, therefore, how to deal with the distinct possibility of inconclusive verification results. There are essentially two options given that a determination of compliance must ultimately be made. The first is to accept a lower detection probability ($1-\beta$) or a higher false alarm rate (α), or some combination of the two, so that, given the anticipated value of $SQ/\sigma(\text{MUF})$, verification will produce conclusive results (note that it is the ratio of SQ to $\sigma(\text{MUF})$ that is relevant here, not their individual values). Figure 2 shows values for α and $1-\beta$ that will lead to conclusive results for various different values of the parameter $\lambda = SQ/\sigma(\text{MUF})$. The cross marks the standard ‘operating point’ of IAEA safeguards ($\alpha = 0.05$ and $1-\beta = 0.9$ with $\lambda = 2.9$).

If verification results in a $\sigma(\text{MUF})$ of 5 kg (the best case scenario) and an SQ is chosen to be a demanding 2 kg (so that $\lambda = 0.4$), a relatively manageable false alarm probability of 10% would

lead to a detection probability of a mere 19%. To obtain a detection probability of even 50% would require a false alarm probability of 34%. If an SQ is instead taken as 8 kg, while keeping $\sigma(\text{MUF})$ the same (so now $\lambda = 1.6$), then a false alarm probability of 10% would lead to a detection probability of 62%.

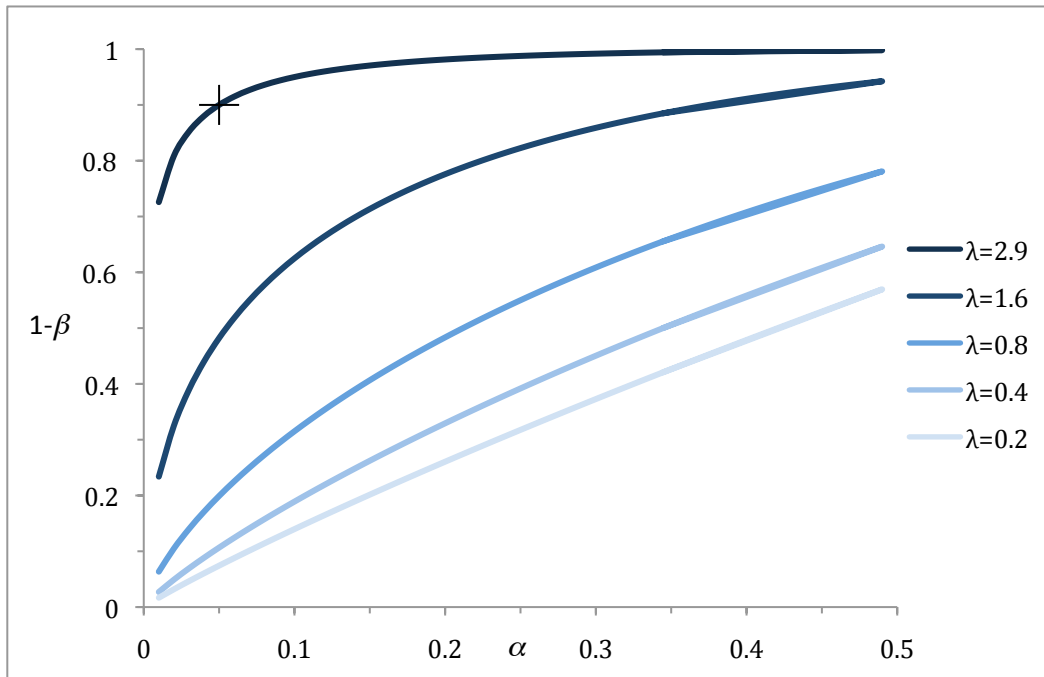


FIGURE 2: Contour plot of values of α and $1-\beta$ for different values of the parameter $\lambda = \text{SQ}/\sigma(\text{MUF})$ that will lead to a conclusive result from verification.

Changing the values of α and $1-\beta$ is a technically satisfying approach because it avoids the possibility of an inconclusive outcome to verification. However, agreeing values different from the IAEA standard amongst all stakeholders could be politically challenging, in particular because as an SQ becomes smaller or $\sigma(\text{MUF})$ becomes larger, it becomes necessary to accept an undesirably high false alarm rate or a disappointingly low detection probability.

The alternative approach is to accept that material accountancy may produce inconclusive results and, if it does, evaluate North Korean compliance on other grounds, such as the degree of transparency in its nuclear programme and the degree of cooperation afforded to inspectors. These were essentially the grounds that the IAEA used to evaluate South Africa’s initial declaration in 1992—3 when material accountancy could not rule out the existence of a clandestine HEU stockpile [3]. It may be easier to reach agreement on this approach because it defers difficult decisions until the analysis phase—but in doing so opens the possibility of acrimonious arguments over the results of verification with no clear method for resolving them.

VERIFYING THE ABSENCE OF ENRICHMENT ACTIVITIES

A second distinct task is to verify whether North Korea is conducting undeclared enrichment activities—a claim it has consistently denied. The difficulty is that it is simply impossible to inspect every site in North Korea capable of hiding a small centrifuge enrichment plant. If a few sites are inspected and a programme is found then it is possible to conclude definitively that North Korea was conducting clandestine centrifuge activities. However, what happens if inspections do not uncover anything? At what point does the absence of evidence constitute evidence of absence [4]?

The US is currently confident, based on intelligence, that North Korea has or had an undeclared enrichment programme, albeit probably a very small one. Moreover, when negotiations were ongoing North Korea reportedly agreed to inspections at undeclared sites (with its consent). The challenge facing a US-led inspectorate is to decide where they should request inspections, given it probably has little idea where the centrifuge plant might be. After all, evidence of the existence of a programme is apparently based on procurement data (which rarely gives an indication of location) and because centrifuge plants have few distinguishing external features and produce negligible environmental emissions (if well run), it can be very hard to locate them. Inspectors can select sites at random and have a look. However, with many potential sites to choose, examining just a few would be unlikely to find evidence of a centrifuge programme, even if one existed. If the inspectorate conducts a small number of inspections and finds nothing it would be wrong to change its prior belief that North Korea has (or had) a clandestine centrifuge programme.

If, on the other hand, inspectors happen to have a very clear indication of where the plant might be (based on defectors or human sources, say), and an inspection of that site reveals nothing, then it would be rational to conclude that in fact North Korea never had a centrifuge programme.

Although verifying the absence of undeclared activities is normally viewed as a matter of qualitative judgement, it can be formulated as a quantitative problem using Bayes' theorem. If inspections fail to find any evidence of a clandestine centrifuge programme then the probability that such a programme nonetheless exists is given by

$$P(H_1|R) = \frac{P(R|H_1)P(H_1)}{P(R|H_1)P(H_1) + P(H_0)}$$

where H_0 denotes the hypothesis that a clandestine programme does not exist, H_1 the hypothesis that it does, and R the (negative) results of the inspections. The fact that $P(R|H_0) = 1$ has been used.

The limiting forms of the above expression when $P(R|H_1)$ is close to 1 and 0 correspond to the two cases described above. When there are many potential sites to choose from, and little idea which one might house a clandestine centrifuge facility, then $P(R|H_1)$ is close to unity and $P(H_1|R)$ can be shown to be just slightly less than $P(H_1)$. In other words, the effect of the inspections is only to reduce slightly confidence in the existence of a clandestine programme.

In contrast, if intelligence is able to pinpoint the suspected site of a clandestine programme then $P(R|H_1)$ is close to zero. Now, if inspections find nothing, $P(H_1|R)$ is also close to zero and the correct conclusion is that there is, in fact, no clandestine programme.

The controversy about the use of Bayes' theorem in verification is that it is necessary to assign a prior probability, $P(H_1)$, to non-compliance. 'Traditional' analysts argue that it is illegitimate to try to assign in advance some probability of cheating. However, as argued above, prior beliefs ought to affect compliance assessments. The principal advantage of formulating verification explicitly as a Bayesian inference problem is that it forces different analysts to make their assumptions explicit and therefore creates a framework for resolving disputes about interpretation. In addition, Bayes' theorem allows the detection of clandestine activities to be treated in a more quantitative way. Intelligence analysts are already effectively forced to quantify approximately the probability assigned to their judgements so they can correctly use estimative language (terms like 'very likely', 'likely', 'unlikely' etc. roughly relate to probabilities). It is only a small step from there to assigning values for the terms in Bayes' theorem. Moreover, expressions like $P(R|H_1)$, the probability that inspections will turn up no evidence of a clandestine programme even though one exists, can also be quantified under some circumstances. For example, if there are N possible sites for a clandestine centrifuge plant and inspectors are allowed to visit n of them then, in the absence of any data indicating which site might house the facility, $P(R|H_1)=1-n/N$.

In any case, there are significant policy implications to the fact that it may still be correct to conclude that North Korea has a clandestine centrifuge programme, even if inspections do not uncover it. In this eventuality, North Korea would undoubtedly claim a significant diplomatic victory and urge the US to apologise for its accusations. China and Russia, amongst others, could find it politically convenient to back North Korea, regardless of whether such a conclusion was warranted. Politically, it may be hard for the US to sustain the argument that North Korea is continuing with clandestine centrifuge activities.

This kind of scenario is far from hypothetical. Twice in the 1990s the US, to its cost, demanded inspections at undeclared sites. In 1993 the US navy held the *Yinhe*, a Chinese ship suspected of carrying illicit chemical agents to Iran, at sea for a month until China agreed to an inspection by Saudi Arabia. The inspection turned up nothing and the whole incident significantly damaged Sino-US relations [5]. In 1999, North Korea eventually agreed to US inspections of some caves at Kumchang-ri in return, inter alia, for \$177 million in food aid. Again nothing was found [6].

Given the political problems associated with the use of inspections at undeclared sites, there is, in practice, likely to be a strong reluctance against using them in North Korea, absent very strong intelligence information highlighting a suspect site [7]. Moreover, even negotiating undeclared inspections and not using them has risks because it would allow North Korea to demand that the US conduct an inspection or cease making accusations. (Although the threat of inspections of undeclared sites can have a significant deterrent value in some scenarios, it does not here as the concern relates to an extant clandestine programme.)

All of this has implications for US strategy towards North Korea. Recognizing the real limits to the effectiveness of inspections at clandestine sites would be useful in advance of negotiations. For instance, absent excellent intelligence, it may be worth spending political capital on measures other than mandatory inspections at undeclared sites, such as interviews with North Korean scientists, or even to obtain concessions in other areas, unrelated to verification.

CONCLUSIONS

The main purpose of this paper is to highlight the challenges and key policy choices faced in constructing a verification regime for North Korea. In the case of verifying North Korea's plutonium programme to ascertain whether it has retained any material clandestinely, there are two key issues.

First, IAEA safeguards are designed to detect the diversion of one significant quantity (SQ) or more of nuclear material. For plutonium, an SQ is set to 8 kg. Given that North Korea has claimed the device detonated in October 2006 only contained 2 kg, should the definition of an SQ be lowered in this case?

Second, because of inevitable measurement errors, if standard IAEA safeguards criteria are applied, there may be a significant chance that the results of verification will be inconclusive. How should this be dealt with? Should a lower detection probability or higher false alarm rate be adopted? Or should North Korean compliance be assessed on other grounds, such as its cooperation with inspectors?

When it comes to verifying the absence or existence of a clandestine enrichment programme, inspections at undeclared sites are of very limited utility unless good intelligence pinpointing the location of a suspect facility is available. Policy makers should therefore consider whether it is worth spending limited political capital on negotiating such inspections.

APPENDIX: CALCULATION OF σ (MUF)

This appendix outlines in detail the calculation of σ (MUF) for North Korea's plutonium programme [8], as summarized in Table 1.

Uncertainty in the total quantity of plutonium produced

Based on satellite imagery and other open source data, the Institute for Science and International Security (ISIS) estimates the total quantity of plutonium produced by North Korea in the IRT and 5 MWe reactors to within an uncertainty of about 9 kg [9]. This is a reasonable upper bound for σ (PB). Analysis of the operating records given to the US by North Korea in 2008 could reduce this uncertainty somewhat. However, in order to reduce it significantly, it will be necessary to take and analyze samples from the graphite moderator of the reactor(s). Using GIRM (Graphite Isotopic Ratio Method), the total uncertainty in plutonium production should be reduced to around 5% [10]. Assuming North Korea has produced roughly 60 kg of plutonium, a reasonable lower bound for σ (PB) is, therefore, 3 kg.

Uncertainty in North Korea's current holdings

North Korea has reportedly restarted the Radiochemical Laboratory and is currently reprocessing its remaining unseparated fuel. It seems sensible to assume that this will be completed before any future denuclearization agreement. North Korea reportedly claims that it will not permit inspectors access to weaponized material. Although it might make sense to delay verification of this material, there would be little point in attempting verification if North Korea never permits the quantity of weaponized plutonium to be measured. It is assumed here, therefore, that inspectors will be given access to all separated plutonium. The error in estimating the mass of this material is negligible compared to the other sources of error and can be neglected.

Uncertainty in reprocessing losses

It is difficult to estimate the likely operating losses from reprocessing, let alone the uncertainty in this figure. Albright and O'Neill report that North Korea claims its losses during 'hot testing' were around 25% [11]. It is very likely, however, that losses were reduced significantly as North Korea gained operating experience. Based on US government data, the average operating losses from US military reprocessing, over its entire history, were around 1.5% [12]. North Korean losses are almost certain to be larger than this. If operating losses are between 2% and 20%, then an upper bound for $\sigma(R)$ is roughly 6 kg (assuming about 60 kg of fuel has been reprocessed). Analysis of the hold up in the reprocessing plant, operating records and possibly wastes may allow $\sigma(R)$ to be reduced somewhat. For illustrative purposes, we assume a lower bound of 3 kg.

Uncertainty in the quantity of plutonium used in testing

As mentioned above, North Korea has reportedly claimed that it used just 2 kg of plutonium in its 2006 test. Most analysts typically estimate that North Korean devices probably contain 4 to 6 kg (few, if any, argue they are likely to contain a full SQ). Reflecting this range, an upper bound for the uncertainty in the amount of material used in a single test might be about 2 kg. Recognizing that the amount of material used in the first and second tests should probably not be treated as independent variables, an appropriate upper bound for $\sigma(T)$ is, therefore, about 4 kg. It might be possible to reduce $\sigma(T)$ if North Korea were to provide records of its weaponization activities and give inspectors access to its arsenal (and, just about conceivably, by drilling back into the test cavities). A lower bound for $\sigma(T)$ might be about 2 kg.

REFERENCES

1. Rudolf Avehaus and Morton John Canty, *Compliance Quantified: An Introduction to Data Verification* (Cambridge: Cambridge University Press, 1996), pp. 153—162.
2. 'Rice Presses for North Korean Nuclear Disarmament', *Global Security Newswire*, http://www.globalsecuritynewswire.org/gsn/GSN_20080630_BC4C70E1.php, 30 June 2008.
3. George Perkovich and James M. Acton, *Abolishing Nuclear Weapons* in Perkovich and Acton (eds.) *Abolishing Nuclear Weapons: A Debate* (Washington, DC: Carnegie Endowment for International Peace, 2009), Section 1, pp. 69—70.
4. James Acton, 'Definitely Maybe: Verifying North Korean Denuclearization', *Jane's Intelligence Review*, vol. 20, no. 8, August 2008, pp. 48—53.
5. Patrick Tyler, *A Great Wall* (New York: Public Affairs, 1993), pp. 394—400.
6. 'Uncovering The Truth About North Korea's Alleged Underground Nuclear Facility: The Kumchang-ri Controversy', James Martin Center for Nonproliferation Studies, http://cns.miis.edu/archive/country_north_korea/uncover.htm.
7. Indeed, in this regard it is notable that the special inspection provisions of INFCIRC/153 have only been publicly invoked once by the IAEA while the challenge inspection provisions of the Chemical Weapons Convention have never been used.

8. For an excellent overview of North Korea's nuclear weapons programme see Mary Beth Nikitin, 'North Korea's Nuclear Weapons', CRS Report for Congress, RL34256, 12 February 2009.
9. David Albright and Paul Brannan, 'The North Korean Plutonium Stock, February 2007', Institute for Science and International Security (ISIS), 20 February 2007.
10. Thomas W. Wood, Bruce D. Reid, John L. Smoot and James L. Fuller, 'Establishing Confidential Accounting for Russian Weapons Plutonium', *The Nonproliferation Review*, vol. 9, no. 2, Summer 2002, p. 126—137.
11. David Albright, 'Inconsistencies in North Korea's Declaration to the IAEA' in David Albright and Kevin O'Neill (eds.), *Solving the North Korean Nuclear Puzzle*, ISIS Reports (Washington DC: ISIS, 2000), p. 88.
12. US Department of Energy, 'Plutonium: The First 50 Years: United States Plutonium Production, Acquisition, and Utilization from 1944 through 1994', DOE/DP-0137, February 1996.