Statistics and Applications {ISSN 2454-7395 (online)}
Volume 16 Nos.2, (New Series), pp. 21-35

Optimal Site Selection Strategy for Inspection Under the Chemical Weapons Convention

Debasis Sengupta and Kamrul Hassan
Indian Statistical Institute, Kolkata and East West University, Dhaka

Received 3 July, 2017; Revised: July 28, 2017; Accepted 28 July, 2017

Abstract

The Chemical Weapons Convention calls for random inspection of facilities, which manufacture chemicals that can be used for the production of chemical weapon ingredients. The scheme for random selection of sites partially depends on the probability of selection of a country on the basis of the number of inspectable facilities it has declared under the Convention. There has been debate among the parties to the convention, regarding the way in which this probability should depend on the number of sites. An objective criterion for choosing the probability function is proposed this article. It is shown that the currently used method is not very far from being optimal for this criterion, but there is room for improvement. Search for the optimal probability function is posed as a multi-variable optimization problem. It is shown that there is a unique solution to this problem. A computationally simple iterative algorithm is proposed. The resulting probability function has some interesting similarities and dissimilarities with the currently used function.

Key words: Probability of selection; random inspection; iterative optimization.

1. Introduction

The Convention on the Prohibition of the Development, Production, Stockpiling and Use of Chemical Weapons, commonly referred to as the Chemical Weapons Convention (CWC), is an international treaty that bans the development, production, possession or use of chemical weapons, and requires the destruction of existing weapons (Thakur, 2006). After years of negotiation among various countries in the world, the CWC was opened for signature on 13 January, 1993. It entered into force on 29 April 1997. As of 3 January 2008, there were 183 member countries (called states parties) to the CWC. The current agreement is administered by the Organization for the Prohibition of Chemical Weapons (OPCW). The OPCW has a Technical Secretariat for technical support.

In order to minimize the chances of violation, the states parties under the treaty agreed to accept inspections at declared industrial sites. The chemicals whose production comes under the ambit of inspection are (a)Schedule1 chemicals which have little or no use outside of chemical weapons, (b) Schedule 2 chemicals which have legitimate small-scale

Corresponding Author: Kamrul Hassan
E-mail: kamrul@ewubd.edu

applications other than chemical weapons, and (c)Schedule3 chemicals which have large-scale use apart from chemical weapons. As a further precaution, there is provision for on-site inspection of randomly selected facilities, which manufacture other chemicals, but can be used for the production of chemical weapon ingredients. These sites are called Other Chemical Production Facilities (OCPF). The verification efforts for OCPF plants are centered on those which produce phosphorus, sulfur or fluorine (referred to as PSF plants), while sites producing discrete organic chemicals (DOC) are also of interest.

OCPF sites are more numerous than the sites producing the scheduled chemicals. Between April 2000 and December 2007, the OPCW has inspected a total of 521 of the more than 4,560 inspectable OCPF sites (vide OPCW, 2008, p.5). The OPCW seeks to make the best use ofits limited resources for random inspection of these sites. The methodology for random selection of inspection sites has been a subject of much debate among the states parties over the past decade. According to the treaty (see OPCW, 1997, Verification Annex Part IX, Section B, Paragraph 11), "the Technical Secretariat shall randomly select plant sites for inspection through appropriate mechanisms, … on the basis of the following weighting factors: (a)equitable geographical distribution of inspections; (b) the information on the listed plant sites available to the Technical Secretariat, related to the characteristics of the plant site and the activities carried out there; and (c) proposals by states parties on a basis to be agreed upon in accordance with paragraph 25." The restrictions are: (a) "no plant site shall receive more than two inspections per year"; and (b) "number of inspections (for a state party) shall not exceed three plus 5 percent of the total number of plant sites declared by a State Party, or 20 inspections, whichever of these two figures is lower".

A possible form of the selection probability of the j^{th} site in the i^{th} state party is

$$\pi_{ij} = f\left(g_{ij}, t_{ij}, p_{ij}, \sum_{k=1}^{n_i} g_{ik}, \sum_{k=1}^{n_i} t_{ik}, \sum_{k=1}^{n_i} p_{ik}\right), \quad j = 1, \dots, n_i, \qquad i = 1, \dots, m, \tag{1}$$

where m is the number of states parties with one or more inspectable OCPF sites, n_i is the number of OCPF sites of the i^{th} state party, g_{ij}, t_{ij} and p_{ij} are probabilities of selection of the j^{th} site in the i^{th} state party computed exclusively from geographical, technical and proposal aspects, respectively, and f is a monotone function of its arguments, having range in [0,1]. There has been general consensus that g_{ij} should be an increasing function of n_i and should be the same for all j, and that t_{ij} should be an increasing function of r_{ij}, the technical score ('information points') given by the Technical Secretariat to the j^{th} site in the i^{th} state party, according to an agreed formula (e.g., the A-14 method, vide OPCW, 2008, p.56).

There has been disagreement among the states parties about how $f$and p_{ij} should be chosen, and how g_{ij} and t_{ij} should depend on $n_i$and r_{ij}, respectively. At the beginning of inspections in April 2000, the Technical Secretariat selected sites using the probability of selection

$$\pi_{ij} = \left(\sum_{k=1}^{n_i} g_{ik}\right) \times \frac{t_{ij}}{\sum_{k=1}^{n_i} t_{ik}}, \quad j = 1, \dots, n_i, \qquad i = 1, \dots, m, \tag{2}$$

where,

$$g_{ij} = \frac{n_i^{1/3}}{\sum_{l=1}^{m} n_l^{1/3}} \times \frac{1}{n_i}, \tag{3}$$

$$t_{ij} = \frac{r_{ij}}{\sum_{l=1}^{m} \sum_{k=1}^{n_l} r_{lk}}. \tag{4}$$

This selection process can be interpreted as a two-step process, where the state party is selected in the first step and the site (within the selected state party) in the second step. This two-step process is followed every year in respect of every selection, subject to the maximum of $\min\{3 + .05n_i, 20\}$ inspections per state party per year, and a maximum of two inspections per site per year. The total number of inspections per year is determined by the availability of budget, with nearly 100 inspections taking place in each of the last few years (vide OPCW, 2006, p.8 and OPCW, 2007, p.8).

During the calendar years 2002 to 2007, the OPCW continued with this selection process with the modification

$$g_{ij} = \frac{1}{m} \times \frac{1}{n_i}, \tag{5}$$

which results in equal probability of selection of each state party in the first step (vide OPCW, 2008, p.55).

There have been many proposals on alternative methodologies for OCPF site selection, put forward by different states parties and facilitators. One of these proposals[1]provided for a two-step selection process having selection probability

$$\pi_{ij} = \frac{\sum_{k=1}^{n_i} g_{ik} + \sum_{k=1}^{n_i} t_{ik} + \sum_{k=1}^{n_i} p_{ik}}{3} \times \frac{1 + t_{ij} + p_{ij}}{n_i + \sum_{k=1}^{n_i} t_{ik} + \sum_{k=1}^{n_i} p_{ik}}, \quad j = 1, \dots, n_i, \qquad i = 1, \dots, m, \tag{6}$$

with g_{ij} as in (3), a range of choices for t_{ij}, and the proposal probability p_{ij} determined by nomination points given by different states parties to the j^{th}OCPF site in the i^{th} state party. Another proposal[2]called for a single-step sampling with selection probability

$$\pi_{ij} = \frac{g_{ij} + t_{ij} + p_{ij}}{3}, \quad j = 1, \dots, n_i, \qquad i = 1, \dots, m, \tag{7}$$

with

$$g_{ij} = \frac{n_i^{1/2}}{\sum_{l=1}^{m} n_l^{1/2}} \times \frac{1}{n_i}, \tag{8}$$

$$t_{ij} = \frac{r_{ij}}{\sum_{l=1}^{m} \sum_{k=1}^{n_l} r_{lk}}, \tag{9}$$

---

[1]A US-Swiss proposal dated 12th March 2004.

[2]Presented on 4th May 2005 by the Dutch facilitator Johan O. Verboom, Deputy Permanent Representative of the Royal Kingdom of the Netherlands to the OCPF.

and p_{ij} determined by nomination points given by the states parties. Yet another proposal[3] had π_{ij} given by (7), with

$$g_{ij} = \frac{1 + \frac{1}{2} n_i^{1/2}}{\sum_{l=1}^{m} \left(1 + \frac{1}{2} n_l^{1/2}\right)} \times \frac{1}{n_i}, \tag{10}$$

$$t_{ij} = \frac{1 + \frac{1}{2} \left(\sum_{k=1}^{n_i} r_{ik}\right)^{1/2}}{\sum_{l=1}^{m} \left\{1 + \frac{1}{2} \left(\sum_{k=1}^{n_l} r_{lk}\right)^{1/2}\right\}} \times \frac{r_{ij}}{\sum_{k=1}^{n_i} r_{ik}}, \tag{11}$$

and p_{ij} determined by nomination points given by the states parties. The form of g_{ij} given in (10) is the same as that used for selection of inspection sites of Schedule 3 chemicals under CWC.

While the debate on nomination points and other issues continued, the Director General-OPCW, on 25th May 2007, informed the states parties of the adoption of a simplified single-step sampling scheme, with effect from 1st January, 2008 (vide OPCW, 2008, p.13 and p.57). According to this scheme, the selection probability of the j^{th} site of the i^{th} state party is

$$\pi_{ij} = \frac{g_{ij} + t_{ij}}{2}, \quad j = 1, \dots, n_i, \qquad i = 1, \dots, m, \tag{12}$$

with g_{ij} and t_{ij} defined by (10) and (11), respectively.

Methodologies for verification of various provisions of the CWC have been investigated by different experts (see, e.g., Trapp, 1993; Daoudi and Trapp, 2006). The mathematical problem of devising appropriate inspection strategies in the presence of attempts to evade detection has been considered from a game theoretic angle by Avenhaus and Canty (1996) and Avenhaus et al. (2006). However, the debate among the state parties of CWC have so far involved heuristic considerations only. Neither the strategies adopted by the OPCW (before and after 1st January 2008) nor the proposals placed before it have been demonstrated to have optimality according to any objective principle.

The purpose of the present article is to formalize some of the heuristic considerations used in determining selection probability of an OCPF plant site. An objective criterion for choosing the ‘geographical’ component of selection probability (g_{ij}) is proposed. The suitability of the formulae (3), (5), (8) and (10) to this criterion are examined, and further alternatives are explored. It is shown that, subject to the assumption of constant probability of detection, there exists a unique optimal probability function to determine g_{ij}.

All the computations made in this article are based on the data on the number of inspectable DOC/PSF sites for 77 states parties, declared till December 2006 (vide OPCW, 2007, p.34).

[3] Facilitator's proposal, communicated on 22nd May, 2007.

2. A Criterion for Selection

The total 'geographical probability' of selection a state party is given by

$$g_i = \sum_{j=1}^{n_i} g_{ij}\,, \quad i = 1, \dots, m.$$

An apparent point of convergence among states parties in respect of the choice of g_{ij} has been that it should be of the form

$$g_{ij} = \frac{g_i}{n_i}, \quad j = 1, \dots, n_i, \qquad i = 1, \dots, m, \tag{13}$$

and that g_i should be of the form

$$g_i = \frac{h(n_i)}{\sum_{l=1}^{m} h(n_l)}, \quad i = 1, \dots, m, \tag{14}$$

where h is a positive-valued function. Each of the formulae (3), (5), (8) and (10) corresponds to a special case of (13)–(14), with h chosen as 1,$x^{1/3}$,$x^{1/2}$ and$1 + \frac{1}{2}x^{1/2}$,respectively.

For the purpose of comparing different choices of h, one can assume that plant sites are selected on the basis of the geographical consideration only. The choice $g_{ij} = g_i/n_i$corresponds to a two-step selection where the state party isselected in the first step with probability g_i, and a site within the chosen state party is selected in the second step with probability$1/n_i$, so that all sites within a state party have equal probability of being selected.

The idea of "equitable geographical distribution" is that there should be more inspections where there are more plant sites, that is, h should be a monotonically increasing function. One has to look for a reasonable criterion for choosing the function h, subject to this constraint of monotonicity. The role of the function h is to accentuate or to moderate the effect of a large value of n_i. If one chooses $h(x) = 1$, then all states parties have equal probability of selection. Another extreme situation is observed when h is a sharply increasing function (e.g., $h(x) = x$), so that states parties with a large number of sites are selected most of the time, while those with a handful of sites have negligible probability of being selected.

A possible way to strike a balance between these extreme possibilities is to *maximize the expected number of countries where violation of CWC is detected*. Note that this criterion is different from maximization of the number of violations detected. The former may be more attractive from the point of view of arms control, as resources need not be diverted towards detection of multiple violations by one state party. Detection of a single violation may in any case prepare the ground for more intensive inspections for the errant state party.

Let I be the total number of inspections planned. Let d be the probability that the inspection of a site does not lead to detection of any violation. It is assumed, for mathematical simplicity as well as political correctness, that this probability is the same for all states parties and all sites. It is also assumed that detection of violation at different inspections are independent events. If a state party receives j inspections, then the probability that these inspections would not lead to any detection of violation is d^j. The number of inspections received by the i^{th} state party is a binomial random variable with parameters I and n_i, subject to the maximum number of inspections $\min\{2n_i, [3 + .05n_i], 20\}$, according to the restrictions mentioned after equation (4). (Here, as well as in the sequel, $[3 + .05n_i]$ denotes the largest integer less than or equal to $3+.05n_i$.) Thus, the probability that the inspections do not lead to any violation in the i^{th} state party is

$$\sum_{j=0}^{I} d^{\min\{j,2n_i,[3+.05n_i],20\}}\binom{I}{j}\left(\frac{h(n_i)}{\sum_{l=1}^{m} h(n_i)}\right)^{j}\left(1-\frac{h(n_i)}{\sum_{l=1}^{m} h(n_i)}\right)^{I-j}.$$

It follows that the probability of detection of *at least one violation* by the i^{th} state party is one minus the above expression. Therefore, the expected number of countries where at least one violation is detected, is

$$V_{I,d}(h)=\sum_{i=1}^{m}\left[1-\sum_{j=0}^{I} d^{\min\{j,2n_i,[3+.05n_i],20\}}\binom{I}{j}\left(\frac{h(n_i)}{\sum_{l=1}^{m} h(n_i)}\right)^{j}\left(1-\frac{h(n_i)}{\sum_{l=1}^{m} h(n_i)}\right)^{I-j}\right]. \quad (15)$$

For fixed number of inspections (I) and probability of no detection of violation resulting from an inspection (d), this expected number depends on the function h. It would be reasonable to choose h in such a way that $V_{I,d}(h)$ is maximized.

3. Comparison of Some Specific Functions

Table 1 gives a summary of the values of the criterion $V_{I,d}(h)$for different combinations of the number of inspections (I) and probability of no violation (d), and for different choices of the function h. These values are computed from the December 2006 data on inspectable sites (vide OPCW, 2008, p.34). Apart from the choices1,$x^{1/3}$,$x^{1/2}$ and$1+\frac{1}{2}x^{1/2}$ implied by the formulae (3), (5), (8) and (10), respectively, the value of this criterion is also tabulated for the power function x^{α}, where α is chosen to maximize the criterion.

Table 1: Expected number of countries with at least one violation, for different inspection strategies

Choice of $h(x)$	Value of $V_{I,d}(h)$ for					
	$I=100$			$I=150$		
	$d=.99$	$d=.999$	$d=.9999$	$d=.99$	$d=.999$	$d=.9999$
1	0.948621	0.095349	0.00954	1.342081	0.135151	0.013525
$x^{1/3}$	0.977714	0.09862	0.009871	1.426645	0.144481	0.014467
$x^{1/2}$	0.978055	0.099221	0.009937	1.432119	0.146221	0.014653
$1+\frac{1}{2}x^{1/2}$	0.978363	0.098854	0.009896	1.43256	0.14546	0.014568
x^{α} (optimal α)	0.978965 ($\alpha=.43$)	0.099293 ($\alpha=.56$)	0.009947 ($\alpha=.56$)	1.434593 ($\alpha=.45$)	0.146237 ($\alpha=.49$)	0.014654 ($\alpha=.49$)

It is found that the optimal choice of α is generally about 0.5, and the selections $h(x)=x^{1/2}$ and $h(x)=1+\frac{1}{2}x^{1/2}$are both quite reasonable in comparison with this optimal power function, for the range of I and d considered here. These findings provide the following justification for the change in the selection strategy introduced by the Director general-OPCW since 1st January, 2008 (vide OPCW, 2008, p. 13 and p.57): *If the geographical aspect is used as the sole basis for selection, then the new strategy is expected* *to produce detection of violation by a larger number of states parties, in comparison of the two strategies followed in the past.*

Thus, the new strategy is found to be a step in the right direction. A natural question is: is there an even better strategy?

4. The Optimal Probability Function

Consider the objective function $V_{I,d}(h)$ given in (15). Although the problem is posed as that of optimization with respect to the function h, the objective function only depends on the values of this function at the discrete points $n_1, \dots, n_m$. Many countries have equal number of sites, and hence it is the set of *distinct* values of $h(n_i)$ that have to be adjusted in order to maximize the objective function. This is essentially a case of optimization with respect to multiple variables.

Let $1 = n_1^* < \cdots < n_s^*$ be the set of distinct positive values of the number of sites, and $m_1, \dots, m_s$ be the corresponding multiplicities (i.e., m_i is the number of countries havingn_i^* sites). Let $c = \min\{i: n_i^* \geq 340\}$. For$i = 1, \dots, s$, let us denote $h(n_i^*)/\sum_{l=1}^{s} h(n_l^*)$ by the simplified notation h_i. Then the objective function (15) can be rewritten as

$$\begin{aligned} V_{I,d}(h_1, \dots, h_s) = m_1 &\left[1 - \sum_{j=0}^{I} d^{\min\{j,2\}} \binom{I}{j} (h_1)^j (1-h_1)^{I-j}\right] \\ &+ \sum_{i=2}^{c-1} m_i \left[1 - \sum_{j=0}^{I} d^{\min\{j,[3+n_i^*/20]\}} \binom{I}{j} (h_i)^j (1-h_i)^{I-j}\right] \\ &+ \sum_{i=c}^{s} m_i \left[1 - \sum_{j=0}^{I} d^{\min\{j,20\}} \binom{I}{j} (h_i)^j (1-h_i)^{I-j}\right]. \end{aligned} \tag{16}$$

This function has to be maximized with respect to $h_1, \dots, h_s$subject to the constraints$0 < h_1 \leq \cdots \leq h_s$and $\sum_{i=1}^{s} m_i h_i = 1$.

The following proposition shows that if the function (16) is maximized over $(0,1)^s$subject to the constraint $\sum_{i=1}^{s} m_i h_i = 1$, the monotonicity constraint $0 < h_1 \leq \cdots \leq h_s$ is *automatically* satisfied.

Proposition 1. *The objective function (16) is maximized over* $(0,1)^s$ *subject to the constraint* $\sum_{i=1}^{s} m_i h_i = 1$ *only if* $0 < h_1 \leq \cdots \leq h_c = \cdots = h_s$. *When this choice is made, the above optimization problem is equivalent to maximizing*

$$W_{I,d}(h_1, \dots, h_c) = m - \sum_{i=1}^{c-1} m_i \varphi_{K_i}(h_i) - \varphi_{K_c}(h_c) \sum_{i=c}^{s} m_i \tag{17}$$

with respect to $h_1, \dots, h_c$, *subject to the conditions* $0 < h_1 \leq \cdots \leq h_c$ *and* $\sum_{i=1}^{c-1} m_i h_i + h_c \sum_{i=c}^{s} m_i = 1$, *where, for any integer* K *less than* I, φ_K *is the real-valued function defined over the unit interval by the relation*

$$\varphi_K(x) = d^K + \sum_{j=0}^{K} \left(d^j - d^K\right) \binom{I}{j} x^j (1-x)^{I-j}, \quad 0 \le x \le 1, \tag{18}$$

and

$$K_i = \begin{cases} 2 & for\ \ i = 1, \\ [3 + n_i^*/20] & for\ \ 2 \le i \le c, \\ 20 & for\ c \le i \le s. \end{cases}$$

The existence and the uniqueness of the solution is ensured through the next proposition.

Proposition 2. *There is a unique maximum of the objective function (16) over* $(0,1)^s$*, subject to the constraints* $0 < h_1 \le \cdots \le h_c = \cdots = h_s$ *and* $\sum_{i=1}^{s} m_i h_i = 1$.

In spite of the existence of a unique maximum, numerical optimization of (17) with respect to $h_1, \dots, h_{c-1}$may be difficult, because of the restriction in range. Even if the restrictions $h_i > 0$ for $i = 1, \dots, c-1$ are enforced through re-parametrization (e.g., by using $\log(h_i/(1-h_i))$as a variable instead of using h_i; see Maron, 1982), the restriction $1 - \sum_{i=1}^{c-1} m_i h_i \ge h_{c-1} \sum_{i=c}^{s} m_i$ may be more difficult to handle. The following proposition gives an interesting property of the solution, which can be utilized to simplify the computation.

Proposition 3. *The maximum of the objective function (16) over* $(0,1)^s$*, subject to the constraints* $0 < h_1 \le \cdots \le h_c = \cdots = h_s$ *and* $\sum_{i=1}^{s} m_i h_i = 1$*, is given by the set of unique solutions to the family of equations*

$$\varphi'_{K_i}(h_i) = k, \quad i = 1, \dots, c, \tag{19}$$

where φ'_{K_i} *is the first derivative of* φ_{K_i}*,* $K_1, \dots, K_c$ *are as given in Proposition1, and* k *is a constant in the range* $[-I(1-d), 0]$*, which corresponds to the constraint* $\sum_{i=1}^{c-1} m_i h_i + h_c \sum_{i=c}^{s} m_i = 1$.

Because of the monotone increasing nature of φ'_{K_i} (see part(ii) of Lemma 1 given in the Appendix), the left hand side of (19) increases with h_i. Thus, a smaller or larger value of k would lead to a smaller or larger value, respectively, of the sum $\sum_{i=1}^{c-1} m_i h_i + h_c \sum_{i=c}^{s} m_i$. Therefore, the following iterative algorithm can be used for solving the optimization problem at hand.

Step I. For a set of initial values of the h_i, compute the average value of $\varphi'_{K_i}(h_i)$ and set it equal to k.

Step II. Compute updated values of the h_i's by solving the equations (19), one at a time.

Step III. Compute $\sum_{i=1}^{c-1} m_i h_i + h_c \sum_{i=c}^{s} m_i$. If this sum is greater than1, reduce k and repeat Steps I and II. If the sum is smaller than1, increase k and repeat Steps I and II. Stop if the sum is sufficiently close to1.

This algorithm replaces the s-dimensional search for solving the optimization problem (16) into a set of c one-dimensional searches nested in another one-dimensional search.

5. Results

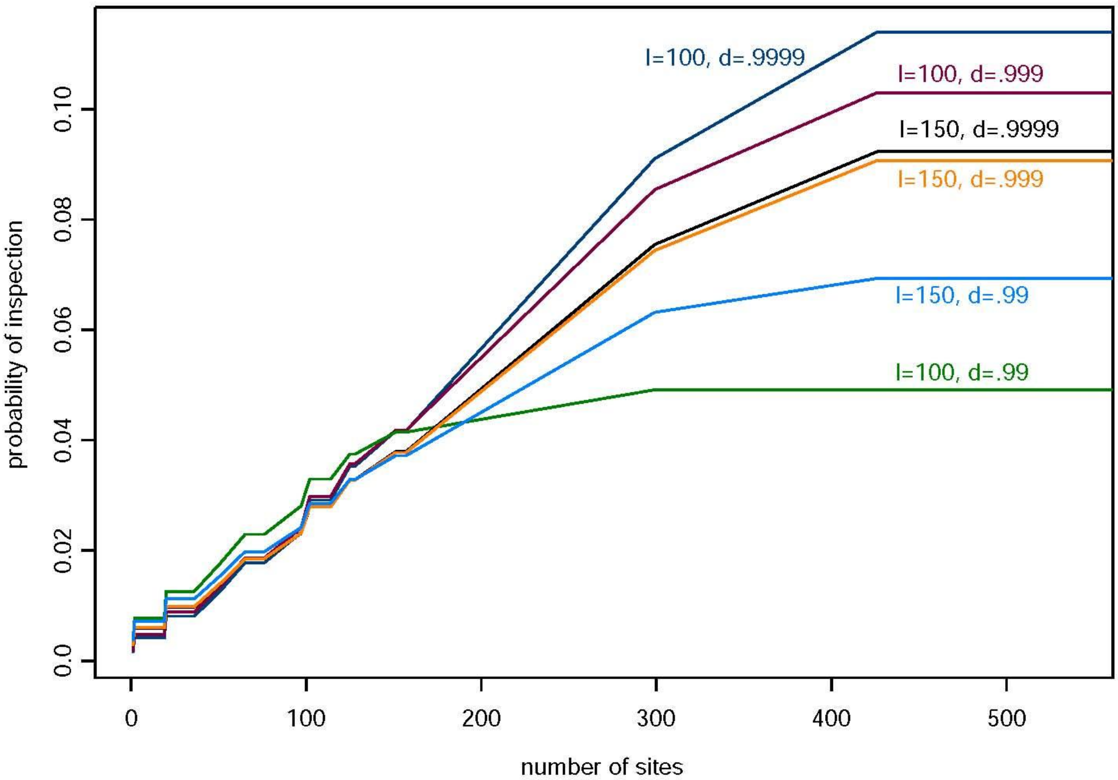
The optimal values of $V_{I,d}(h_1, \ldots, h_s)$ corresponding to some choices of $I$and d, for the inspectable DOC/PSF sites data of 77 states parties (vide OPCW, 2007, p.34), are shown in Table 2. The corresponding values of this criterion for the optimal power function (copied from Table 1) are reported alongside for ease of comparison. The criterion is seen to improve when the optimal selection probabilities are used.

Table 2: Expected number of countries with at least one violation, for optimal inspection strategy

Choice of $h(x)$	Value of $V_{I,d}(h)$ for					
	$I = 100$			$I = 150$		
	$d = .99$	$d = .999$	$d = .9999$	$d = .99$	$d = .999$	$d = .9999$
optimal	0.982064	0.099608	0.009985	1.445573	0.147957	0.014836
x^α (optimal α)	0.978965 ($\alpha = .43$)	0.099293 ($\alpha = .56$)	0.009947 ($\alpha = .56$)	1.434593 ($\alpha = .45$)	0.146237 ($\alpha = .49$)	0.014654 ($\alpha = .49$)

Figure 1 shows the values of the selection probabilities ($h(n_i^*)$) as a function of the number of sites (n_i^*), for$I = 100$ and $150$and $d = .99$, $.999$ and $.9999$. The optimal probability function levels off when the number of sites exceeds 340 (see Proposition1). This feature of the optimal solution is a consequence of Proposition1, and is in contrast with the power functions considered in the previous section.

Figure1: Graph of the optimal function h for different choices of I and$d$



6. Concluding Remarks

The optimal function h, which is commensurate with the chosenobjective (maximizing the expected count of countries where at least one violation is detected), depends on both I and d. The number of inspections, I, is generally determined by the available budget. Regarding the probability of an inspection leading to detection of violation, d, there is the question as to which value of it should be assumed. An indication in this regard is given by the fact that out of the 521 inspections of OCPF sites which took place between April 2000 and December 2007, not a single violation has been detected (vide OPCW, 2008, p.5). Another indicative fact is that the shape of the optimal functions for $d = .999$ and $d = .9999$do not differ much from one another, particularly when $I = 150$. Thus, once the number of inspections is decided, the choice of the function$h$ for different reasonable choices of d are about the same.

The shape of the optimal probability function is somewhat different from the probability functions that have been suggested to, or actually used by, OPCW in the first decade of its inspections. The optimal function becomes flat when the number of sites is larger than 340. This implies that there is no difference between the selection probabilities of different countries if they have more than 340 sites.

While the form of the optimal probability function is not expressed mathematically through a neat algebraic expression, it can be obtained easily through a program. The main computational hurdle of the multi-variable optimization has been overcome in the algorithm developed in Section4. The S-Plus implementation of the algorithm takes only a couple of seconds to run in a machine with Intel Centrino processor.

The criterion proposed in this article may be used for determining a suitable mapping from the site-specific information points r_{ij}'s) to technical probabilities (t_{ij}'s). For this problem, the r_{ij}'s would assume the roles of n_i's, while the values of t_{ij}'s for every distinct value of r_{ij} would be a variable for optimization. Optimization of the expected count of countries with at least one detection, an expression for which can be computed explicitly, may be the subject of another study.

Acknowledgments

The research for this work was funded in part by a project sponsored by the Department of Chemicals and Petrochemicals of the Ministry of Chemicals and Fertilizer, Government of India. The authors thank Dr. Sharmodeep Bhattacharya of Oregon State University, Corvallis, for helpful discussions and initial computations.

References

Avenhaus, R. and Canty, M.J. (1996). *Compliance Quantified: An Introduction to Data Verification*. Cambridge University Press, Cambridge (ISBN: 978-0521019194).

Avenhaus, R., Kyriakopoulos, N., Richard, M. and Stein, G. (editors). (2006). *Verifying Treaty Compliance: Limiting Weapons of Mass Destruction and Monitoring Kyoto Protocol Provisions*, Springer, Berlin (ISBN: 978-3540338543).

Daoudi, M. and Trapp, R. (2006). Verification under the chemical weapons convention. In *Verifying Treaty Compliance: Limiting Weapons of Mass Destruction and Monitoring Kyoto Protocol Provisions*, R. Avenhaus, N. Kyriakopoulos, M. Richard and G. Stein (editors),Springer, Berlin.

Maron, M.J. (1982). *Numerical Analysis: A Practical Approach*. Macmillan, New York (ISBN: 978-0534123727).

OPCW (1997). *Convention on the Prohibition of the Development, Production, Stockpiling and Use of Chemical Weapons and on Their Destruction*, Organization for the Prohibition of Chemical Weapons, http://www.opcw.org.

OPCW (2006).*Report of the OPCW on the Implementation of the Convention on the Prohibition of the Development, Production, Stockpiling and Use of Chemical Weapons and on their Destruction in2005*}, presented to its Eleventh Session of OCPF held on 5–8 December, 2006, http://www.opcw.org/docs/csp/csp11/en/c1104(e).pdf.

OPCW (2007). *Report of the OPCW on the Implementation of the Convention on the Prohibition of the Development, Production, Stockpiling and Use of Chemical Weapons and on their Destruction in2006*}, presented to its Twelfth Session of OCPF held on 5–9November, 2007,http://www.opcw.org/docs/csp/csp12/en/c1206(e).pdf.

OPCW (2008). *Review of the Operation of the Chemicals Weapon Convention since the First Review Conference*, Note by the Technical Secretariat of OPCW to the Second Review Conference held on 7–18 April, 2008, http://www.opcw.org/docs/csp/rc2/en/rc2s01+(e).pdf.

Takayama, A. (1985). *Mathematical Economics*, Cambridge University Press, Cambridge (ISBN: 978-0521314985).

Thakur, R., Haru, E. (2006). *The Chemical Weapons Convention: Implementation, Challenges and Opportunities*, United Nations University Press, Tokyo (ISBN:978-9280811230).

Trapp, R. (1993). *Verification under the Chemical Weapons Convention: On-site Inspection in Chemical Industry Facilities*. SIPRI Chemical and Biological Warfare Studies no. 14, Oxford University Press, Oxford (ISBN: 978-0198291602).

Appendix: Proofs of the propositions: The following lemma would pave the way for proving the propositions.

Lemma 1. *The function φ_K defined in (18) has the following properties.*

(i) *The function is monotonically decreasing over $(0,1)$, with $\varphi_K(0) = 1$ and $\varphi_K(1) = d^K$.*

(ii) *The function is strictly convex over $(0,1)$.*

(iii) *If the integers K and L satisfy $K < L$, then φ_L is dominated by φ_K.*

(iv) *If the integers K and L satisfy $K < L$, then the function $\varphi_K - \varphi_L$ is monotonically decreasing over $(0,1)$.*

Proof.

(i) The derivative of φ_K is given by

$$\begin{aligned}\varphi_K'(x) = \sum_{j=1}^{K}\left(d^j - d^K\right)\frac{I!}{(j-1)!\,(I-j)!}x^{j-1}(1-x)^{I-j}\\ -\sum_{j=0}^{K}\left(d^j - d^K\right)\frac{I!}{j!\,(I-j-1)!}x^{j}(1-x)^{I-j-1}\\ =\sum_{j=0}^{K-1}\left(d^{j+1} - d^K\right)\frac{I!}{j!\,(I-j-1)!}x^{j}(1-x)^{I-j-1}\\ -\sum_{j=0}^{K-1}\left(d^j - d^K\right)\frac{I!}{j!\,(I-j-1)!}x^{j}(1-x)^{I-j-1}\\ =-\sum_{j=0}^{K-1}d^j(1-d)\frac{I!}{j!\,(I-j-1)!}x^{j}(1-x)^{I-j-1} < 0.\end{aligned} \tag{20}$$

The values of φ_K at 0 and 1 are easy to check.

(ii) It follows from (20) that the second derivative of φ_K is

$$\begin{aligned}\varphi_K''(x) = -\sum_{j=1}^{K-1}d^j(1-d)\frac{I!}{(j-1)!\,(I-j-1)!}x^{j-1}(1-x)^{I-j-1}\\ +\sum_{j=0}^{K-1}d^j(1-d)\frac{I!}{j!\,(I-j-2)!}x^{j}(1-x)^{I-j-2}\\ =-\sum_{j=0}^{K-2}d^{j+1}(1-d)\frac{I!}{j!\,(I-j-2)!}x^{j}(1-x)^{I-j-2}\\ +\sum_{j=0}^{K-1}d^j(1-d)\frac{I!}{j!\,(I-j-2)!}x^{j}(1-x)^{I-j-2}\\ =\sum_{j=0}^{K-1}d^j(1-d)^2\frac{I!}{j!\,(I-j-2)!}x^{j}(1-x)^{I-j-2}\\ +d^{K-1}(1-d)\frac{I!}{(K-1)!\,(I-K-1)!}x^{K-1}(1-x)^{I-K-1} > 0,\end{aligned} \tag{21}$$

which shows that the function φ_K is strictly convex for $x \in (0,1)$.

(i) It is enough to prove the result for $L = K + 1$. In this case,

$$\varphi_K(x) - \varphi_{K+1}(x) = d^K(1-d) \sum_{j=K+1}^{I} \binom{I}{j} x^j (1-x)^{I-j} > 0.$$

(ii) It is enough to prove the result for $L = K + 1$. As one can see from the above expression, $\varphi_K(x) - \varphi_{K+1}(x)$ is proportional to the upper tail probability of a binomial distribution, which is increasing in the parameter x.

Proof of Proposition 1. It is easy to see that

$$\begin{aligned}
\varphi_K(x) &= d^K + \sum_{j=0}^{K} \left(d^j - d^K\right) \binom{I}{j} x^j (1-x)^{I-j} \\
&= \sum_{j=0}^{I} d^K \binom{I}{j} x^j (1-x)^{I-j} + \sum_{j=0}^{K} \left(d^j - d^K\right) \binom{I}{j} x^j (1-x)^{I-j} \\
&= \sum_{j=0}^{I} d^{\min\{j,K\}} \binom{I}{j} x^j (1-x)^{I-j}.
\end{aligned}$$

This identity, together with the fact $\sum_{i=1}^{s} m_i = m$, leads to the following simplification of (16).

$$\begin{aligned}
V_{I,d}(h_1, \dots, h_s) &= m - m_1\varphi_2(h_1) - \sum_{i=2}^{c-1} m_i \varphi_{[3+n_i^*/20]}(h_i) - \sum_{i=c}^{s} m_i \varphi_{20}(h_i) \\
&= m - \sum_{i=1}^{s} m_i \varphi_{K_i}(h_i).
\end{aligned} \tag{22}$$

Consider a set of numbers $h_1, \dots, h_s$ satisfying theconstraints $h_i \geq 0$ for $i = 1, \dots, s$ and $\sum_{i=1}^{s} m_i h_i = 1$. Let $h_i > h_{i+1}$ hold for some $i \in \{1,2, \dots, s-1\}$. Define

$$h_l^* = \begin{cases} h_l & \text{for } 1 \leq l < i \text{ and } i+1 < l \leq s, \\ (m_i h_i + m_{i+1} h_{i+1})/(m_i + m_{i+1}) & \text{for } i \leq l \leq i+1. \end{cases}$$

Since $K_i \leq K_{i+1}$and $h_i \geq h_{i+1}$, it follows from part (iv)of Lemma 1 that

$$\varphi_{K_i}(h_{i+1}) - \varphi_{K_{i+1}}(h_{i+1}) \leq \varphi_{K_i}(h_i) - \varphi_{K_{i+1}}(h_i).$$

Consequently,

$$
\begin{aligned}
&m_i\varphi_{K_i}(h_i)+m_{i+1}\varphi_{K_i}(h_{i+1})\\
&\qquad=\frac{m_i^2}{m_i+m_{i+1}}\varphi_{K_i}(h_i)+\frac{m_im_{i+1}}{m_i+m_{i+1}}\varphi_{K_i}(h_i)+\frac{m_im_{i+1}}{m_i+m_{i+1}}\varphi_{K_{i+1}}(h_{i+1})\\
&\qquad+\frac{m_{i+1}^2}{m_i+m_{i+1}}\varphi_{K_{i+1}}(h_{i+1})\\
&\qquad\geq\frac{m_i^2}{m_i+m_{i+1}}\varphi_{K_i}(h_i)+\frac{m_im_{i+1}}{m_i+m_{i+1}}\varphi_{K_i}(h_{i+1})+\frac{m_im_{i+1}}{m_i+m_{i+1}}\varphi_{K_{i+1}}(h_i)\\
&\qquad+\frac{m_{i+1}^2}{m_i+m_{i+1}}\varphi_{K_{i+1}}(h_{i+1})\\
&\qquad\geq m_i\varphi_{K_i}\left(\frac{m_ih_i+m_{i+1}h_{i+1}}{m_i+m_{i+1}}\right)+m_{i+1}\varphi_{K_{i+1}}\left(\frac{m_ih_i+m_{i+1}h_{i+1}}{m_i+m_{i+1}}\right)\\
&\qquad=m_i\varphi_{K_i}(h_i^*)+m_{i+1}\varphi_{K_{i+1}}(h_{i+1}^*).
\end{aligned}
$$

The second inequality follows from the convexity of the functions$\varphi_{K_i}$and $\varphi_{K_{i+1}}$, as stated in part (ii) of Lemma1. Hence, from (22),

$$V_{I,d}(h_1,\dots,h_s)\leq V_{I,d}(h_1^*,\dots,h_s^*)$$

Thus, the criterion can only be improved by rectifying violation of the increasing order of the adjacenth_i's, and an optimal choice of the h_i's must satisfy this order.This fact is a consequence of the increasing order of the K_i's.Since $K_c=\cdots=K_s$, any optimal choice of the h_i's must satisfy $h_c=\cdots=h_s$.

Proof of Proposition 2. Proposition1 ensures that it wouldsuffice to maximize $W_{I,d}(h_1,\dots,h_s)$ over $[0,1]^c$ subjectto the constraint $\sum_{i=1}^{c-1}m_ih_i+h_c\sum_{i=c}^{s}m_i=1$. The latter constraint can be written as

$$h_c=\frac{1-\sum_{i=1}^{c-1}m_ih_i}{\sum_{i=c}^{s}m_i}. \tag{23}$$

If this expression is substituted in (17), the problem reduces to a $(c-1)$-variable optimization problem with respect to $h_1,\dots,h_{c-1}$. The set of feasible values of $\{h_1,\dots,h_{c-1}\}$ is the subset of $[0,1]^{c-1}$, which satisfies the order restrictions $h_1\leq\cdots\leq h_{c-1}\leq h_c$. It is easy to see that the set of feasible values is a convex set.
Note from (23) that

$$\frac{\partial h_c}{\partial h_i}=-\frac{m_i}{m_c+\cdots+m_s}.$$

Differentiation of (17), together with the above relation, gives

$$\frac{\partial W_{I,d}}{\partial h_i}=-m_i\{\varphi'_{K_i}(h_i)-\varphi'_{K_c}(h_c)\},\quad i=1,\dots,c-1,$$

$$\frac{\partial^2W_{I,d}}{\partial h_i^2}=-m_i\varphi''_{K_i}(h_i)-\frac{m_i^2}{m_c+\cdots+m_s}\varphi''_{K_c}(h_c),\quad i=1,\dots,c-1,$$

$$\frac{\partial^2 W_{I,d}}{\partial h_i \partial h_j} = -\frac{m_i m_j}{m_c + \cdots + m_s} \varphi''_{K_c}(h_c), \quad i,j = 1,\ldots,c-1;\ i \neq j.$$

If the vector $\boldsymbol{h}$ is defined as $(h_1 \quad \cdots \quad h_{c-1})^T$, then the above expressions can be summarized as

$$-\frac{\partial W_{I,d}}{\partial \boldsymbol{h}} = \begin{pmatrix} m_1\{\varphi'_{K_1}(h_1) - \varphi'_{K_c}(h_c)\} \\ \vdots \\ m_{c-1}\{\varphi'_{K_{c-1}}(h_{c-1}) - \varphi'_{K_c}(h_c)\} \end{pmatrix}, \tag{24}$$

$$\begin{aligned} -\frac{\partial^2 W_{I,d}}{\partial \boldsymbol{h} \partial \boldsymbol{h}^T} = &\begin{pmatrix} m_1 \varphi''_{K_1}(h_1) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & m_{c-1}\varphi''_{K_{c-1}}(h_{c-1}) \end{pmatrix} \\ &+ \frac{\varphi''_{K_c}(h_c)}{m_c + \cdots + m_s} \begin{pmatrix} m_1 \\ \vdots \\ m_{c-1} \end{pmatrix} (m_1 \quad \cdots \quad m_{c-1}). \end{aligned} \tag{25}$$

It follows that

$$\boldsymbol{h}^T \left(\frac{\partial^2 W_{I,d}}{\partial \boldsymbol{h} \partial \boldsymbol{h}^T} \right) \boldsymbol{h} = -\sum_{i=1}^{c-1} m_1 h_i^2 \varphi''_{K_1}(h_i) - \frac{\varphi''_{K_c}(h_c)}{m_c + \cdots + m_s} \left(\sum_{i=1}^{c-1} m_i h_i \right)^2$$

Part (ii) of Lemma 1 implies that the right hand side is strictly negativeover $(0,1)^{c-1}$. Therefore, the function $W_{I,d}$ is strictly concave over the feasible set (already shown to be convex). The statement of the proposition follows (see Takayama, 1985, p.87–88).

Proof of Proposition 3. The equations given in (19) are obtained by setting the gradient vector (24) equal to zero. According to part (ii) of Lemma 1, the function φ'_K is increasing, and the expression (20) shows that the minimum and maximum values of φ'_K, which occur at 0 and 1, respectively, are $-I(1-d)$ and 0, respectively. Each of the equations given in (19) has a unique solution because of the monotonicity of φ_K.