Appendix B: Combining Different Types of Information Using Bayesian Statistics

George Lindsey

Nearly all of the published material concerning the technology and operational procedures employed for verifying and monitoring arms control agreements has described methods of obtaining information from a single means of collection. The subject of this report is to investigate the use of synergy to exploit two or more **different** sources of information. This Appendix introduces an example of the problem of synthesizing information of different types in a mathematically logical manner, obtained by different methods.

The essence of synergy, as applied to monitoring for verification, is the synthesis of different sets of knowledge (data, opinions, assessments, etc., all of which are subject to error or imprecision), in order to arrive at the best available inference as to the true situation.

Classical statistics allows us to obtain estimates of numbers, based entirely on measurements, together with an indication of the accuracy of the estimates. Usually these measurements are all of the same type, although statistical techniques allow for cases in which some measurements differ from others in their accuracy.

However, while conducting a verification of an arms control agreement, a situation may arise in which auxiliary information, possibly not based on measurements, perhaps not even quantitative at all, is available. It may be desirable to make use of this information, probably incorporating it into the eventual conclusion as to what the real numbers are.

There is a theory involving subjective probabilities and Bayesian statistics (which is not universally accepted by statisticians) that can be applied to problems of the type described above.

This can be illustrated using the example of repeated sampling, by some sensor, of a large area in which an arms control agreement allows the deployment of $N_0 = 200$ mobile weapons. We suppose that the sensor can only observe a small fraction (S = 2%) of the area each day, under conditions in which it will detect the presence of every weapon actually deployed within its limited field of view. Statistical theory indicates that after 100 days of observation of a deployment of 200 weapons, the number N_E estimated to be there will have a normal probability distribution with a standard deviation σ of 10. This permits calculation of the probability that the estimated number will be N_E. For example, the probability that the estimate N_E will be exactly 200 is only 4%; but the probability that N_E will fall somewhere between 177 and 223 is 98%.

If the actual number of weapons is larger than N = 200, the standard deviation of N_E will be larger, i.e., the accuracy of the estimate will be lower.

Classical ("frequentist") statistics makes no prior assumptions as to the actual number N of weapons deployed. It is assumed that the actual number can be 0, 1, 2, ..., with no upper limit.

Suppose that, as well as the observations, there is evidence from completely different sources suggesting that, after a long period during which the deployment was at the permitted level of $N_0 = 200$, additional weapons were being deployed. To offer an example, suppose that a factory has been observed shipping 100 crates of a size and type which could each contain one new weapon of the type limited by the treaty. Call this hypothesis C, i.e., now N = 300. Suppose, further, that a new military unit of the type operating the weapons and of a size appropriate for 50 additional weapons is known to have been formed, yielding hypothesis B, that the actual number deployed is now N = 250. However, earlier experience has given little indication of any breakout, and many analysts continue to believe that only the permitted 200 weapons are deployed. Call this hypothesis A.