# SOME SUGGESTIONS FOR STUDYING NUCLEAR PROLIFERATION

Paul R. Williamson
Global Vision, Inc.
11/95; revised 11/96 and 10/03

## 1. Introduction to the Problem

This proposal sets forth a possible research approach to the problem of anticipating when states will take steps to acquire nuclear, chemical, or biological weapons. The approach will be based on a model that is crudely "dynamic" in that it will seek to predict proliferation events in the relatively short-term, say on the order of a year, in the future; but one of the questions to be explored is how long a period is possible, in actuality. This approach will be based on techniques of artificial neural network modeling. The modeling also will draw on historical data which describe the characteristics of states and of relationships among them for each individual state or pair of states (dyad), year-by-year. The advantages of this approach include:

1) as we explain below, the proposed techniques may provide a way of dealing with several methodological difficulties, involving complex social systems with many variables affecting each other in unknown ways, which characterize proliferation modeling and prediction problems;

2) the modeling techniques have been previously applied to many problems of a similar complex nature;

3) many of the data are already developed by international relations experts, for every state or dyad, worldwide, in every year going back to well before the Second World War.

Since we do not discuss the original biological type of neural network, herein we drop the phrase “artificial”. In the next section we discuss some problems which have attended application of dynamic modeling to social systems but which appear to be capable of improvement using neural network modeling. Following that, we introduce the concept of neural networks, describe its proposed application to forecasting proliferation, and indicate some problems which will remain. Finally, we discuss the proposed variables.

## 2. Dynamic Modeling

In attempting to use numerical models to predict societal behavior, many complications arise concerning what form of model shall be tested. It is important to see what three of these complications are since they often have appeared to raise significant barriers to social system modeling, and because the approach that we outline appears to ease all of them. To see what they are, we need to formulate more precisely what we mean by “prediction.” Suppose one knows the values of N many variables of interest $x_i(t)$, $i = 1, \dots N$ at a referent time $t$ (where $t$ refers, say, to a given year). And suppose one knows how to compute, *entirely* as a function of one or more of the original values themselves, the changes $\Delta x_i(t)$ in each such value from that time to the later time (say the next year) $t+1$. Then one *also* can know the values of each of the variables at the later time: they are given by

$$x_i(t+1) = x_i(t) + \Delta x_i(t) \qquad (1).$$

In that event we would say that one can “predict” from the one time period to the next. It is in this sense that we use the term prediction.

By a "dynamic model" we mean a situation of the above type, where one has the recipe to get from the one time period to the next period, in the form of a function for the changes. (The model may give the differential form $dx_i / dt$ from which the changes $\Delta x_i$ are then estimated.)

In a dynamic model, the variables $x_i$ which go into the change recipe are called *state variables*. Since the changes use only the original values, the whole process is self-contained. Using equation (1), one "bootstraps" from period $t$ values to period $t+1$ values; one then uses the $t+1$ values to compute the changes for that period, then to bootstrap to the values of the next period; and so on. In this manner, starting in period $t=1$ and using the information for that period only, it is possible to compute the values for subsequent periods 2, 3, 4, ••• into the future. *If* the initial values were completely accurate and the change functions were correct, then all the future computations would be completely accurate also; but realistically, the initial values will contain errors which affect the accuracy of all subsequent calculations. Probably too, no change function would be entirely accurate. However, for some finite number of initial periods, the estimated future values might turn out to be of imperfect but tolerable accuracy. For weather prediction, that period of tolerable accuracy is a few days; for the solar system, it may be on the order of 100 million years (Murray 1993, p. 106.). How long it might be for global social phenomena requires empirical study; as mentioned, in the proliferation study we will begin by investigating the ability to predict one year into the future.

Let us consider the first of the three complications to which we just alluded. For social systems, quite likely one does not know exactly, or even approximately, what the change function ought to look like, or there may be a great many candidates among which it is difficult to choose; so one could spend a long time trying to find the "correct" choice. One aspect of the variety of possible choices deserves special mention. In the standard dynamic modeling approach referred to above, system change at time $t$ depends only on the values of the variables at that moment. Alternatively, system change at time $t$ might well depend on values of variables at present but also at previous times $t-1$, $t-2$, ••• $t-n$, •••, indefinitely far into the past. So the question is, how to address this multiplicity of possible schemes?

The second complication is, the plausible change functions include those which are non-linear, for which usually there is no general solution algorithm (such as "least squares", etc. that exist for linear functions), thus there is no assurance of being able to use observed data to infer the parameters that best define the function. (As we intimated above, non-linearity may also give rise to chaotic phenomena which defeat all subsequent predictive attempts after finitely many periods, but how soon this happens should be answered by empirical study.)

Finally is the added practical difficulty of accommodating large numbers of state variables, especially when they may involve many disparate functional relationships. Multiple indicators may also complicate standard methods of estimating models to the extent that the latter require questionable statistical assumptions such as sufficient degrees of freedom, non-correlation between independent variables and residuals, etc.

While a good part of the choice thus is simply in the quite considerable number of plausible attribute and behavioral indicators available (see the Variable List at the end), again let us call attention to a particular aspect: Suppose we have a collection of indicators we believe relate to system change, where one of them is denoted by $x_1$. Now the question arising is, do we really mean $x_1$ is the state variable or do we mean its change $\Delta x_1$ instead? That is, perhaps the effective relationship is between the amount by which the original variable *changed* during the period as state variable contributing to *its* change, the second order change $\Delta^2 x_1$, as the dependent variable. Going in the other direction, perhaps the *cumulation* of $x_1$ over time is the appropriate state variable, with $x_1$ the dependent change variable. Corresponding state variable questions arise for many candidate indicators.

Neural network modeling technique has developed characteristics which may help circumvent these complications about choosing the form of the function, finding the correct parameters of nonlinear functions, and integrating a large number of variables (Karasik and Williamson 1994, Williamson and Karasik 1994). The circumvention is achieved through a process in which such a model uses empirical examples to converge through iteration toward suitable specifications of these factors. As in all empirical testing, many cases are better than few; and there are dangers of false generalization from too few cases. However this process works on any number of cases and allows all candidate state variables to be used at once, nor does it assume anything concerning mutual correlations among the latter. And the process allows one to introduce inputs from many past time periods at once and to calibrate the parameters of the model in multiple passes, in each one adding additional past years to search for a satisfactory tradeoff between parsimony and accuracy.

Our approach to predicting proliferation steps will make use of such characteristics and process. Again, our motivation for such an approach is the belief that this method can help overcome the indicated difficulties of explicit dynamic modeling. In addition, we will relax the definition of a dynamic model given above by seeking to develop a variant, in which some but not all of the variables are updated within the model. Let us call a variable $x_i$ "endogenous" if its change function (or differential) is computed within the model. In the previous definition every variable must be endogenous; this, of course, is the key to a dynamic model being self-contained in the sense of computing all its changes. Suppose the changes $\Delta x_i$ also depended on the values of some other variables $y_k$, for which one was *not* able to accurately compute changes from one period to the next? Let us call these latter variables "exogenous". Then, for predictions beyond one period, the initial period 1 input values would not be sufficient; one would need to wait until the future values of the exogenous variables were observed, before using them to help compute the subsequent round of $\Delta x_i$ values.

While the ultimate goal may be a fully dynamic model, we believe the realistic situation at this juncture for social systems requires many variables at first be treated as exogenous. In the proposed study, many or all the inputs other than the proliferation steps to be modeled will be exogenous. This limitation is a helpful simplification for this first modeling attempt. We do not expect it to encumber the goal of predicting on a one-year time scale, for the exogenous variables either are slowly changing (such as material capabilities) or can quickly be known (such as the outbreak of a war).

## 3. Artificial Neural Networks

By neural network is meant any of several schemes like those described in Hammerstrom (1993), Haykin (1994), and Rich and Knight (1991, chapter 18). We intend to use a "back propagation" scheme, and a network composed of "input" and "output layers", and one or more "hidden layers". We will describe this scheme in the case of one hidden layer.

In brief, the idea of a back propagation network is as follows. The input layer consists of a series of components, each of which is simply a counter holding the numerical value characterizing one of the pieces of input information. We will return to the identity of these inputs below, after discussing the mechanics of the modeling approach.

The hidden layer consists of a second set of components and a set of weights called synaptic weights. Let $j$ refer to the jth such hidden layer component. Each such component first computes a value $s(j)$ which is a weighted sum of the values of the input layer, minus a constant reference value: if $x(i)$ is the value of the ith input component, $w_1(j,i)$ is the synaptic weight applied to $x(i)$ , and $w_1(j,0)$ is the reference value (also considered a synaptic weight), then we define

$$s(j) = \sum_i w_1(j,i)x(i) - w_1(j,0) \ , \tag{2}$$

across all values of $i$ . Second, the jth component computes a new number based on a sigmoid (inverse logistic) function the argument of which is $s(j)$ . The value of this new function, defined as

$$h(j) = 1/[1+e^{-s(j)}] \ , \tag{3}$$

evidently approaches 0 as $s(j)$ becomes more strongly negative and 1 as $s(j)$ becomes more strongly positive, and $h(j) = 1/2$ for $s(j) = 0$ . Additional hidden layers would draw on the final computations $h(j)$ of the previous hidden layer in the same way as the first one draws on the values of the input layer.

Finally, the output layer consists of a third (or final) set of components, each of which corresponds to one of the empirically observed change variables, in a manner we will discuss later. Everything described in the hidden layer has a corresponding element in the output layer: A weighted sum defined as

$$y(k) = \sum_j [w_2(k,j)h(j)] - w_2(k,0) \ , \tag{4}$$

is computed across all values of $j$ . Then the sigmoid function

$$o(k) = 1/[1+e^{-y(k)}] \ , \tag{5}$$

is computed.

In the proposed use of this approach, each of the outputs $o(k)$ corresponds to a "desired" quantity $d(k)$ which is a suitable transformation of some empirically observed quantity. In our application, this observed quantity is the change $\Delta x_k$ in one of the input variables $x_k$ . The ideal of the model is to make each output equal its corresponding desired quantity; and any discrepancy between the two supplies a non-zero component to the error vector $e(k) = o(k) - d(k)$ . Qualitatively, the criterion of all such models is: find synaptic weights $w_1(j,i)$, $w_1(j,0)$, $w_2(k,j)$ , and $w_2(k,0)$ for all $i$ , $j$ and $k$ such that a measure of the variance of $e(k)$ is acceptably small. (A quantitative statement of the criterion is given by Haykin 1994.)

The question then becomes, can one find such numbers? Repeated empirical investigation as well as mathematical analysis suggests that under a very wide set of circumstances the answer is likely to be Yes, provided one is able to subject the model to a process of calibration called "training". In training, the initial synaptic weights are set by some plausible guess (possibly at "random") and the corresponding outputs are compared with the empirically observed values (criterion values) of a known outcome; then the synaptic weights are changed as a function of the error vector of output minus criterion values. When this process is repeated using additional criterion values from other known outcomes, the resulting error vectors typically become smaller. (One must be alert to the problem of solutions converging to local error minima which may be very inferior to the desired global minimum. Techniques are available for this purpose.)

Success with this process, even to the point of predictive accuracy, means something different than ordinary, however. Usually in modeling, one begins by positing an explicit functional relationship among data-- presumably reflecting a conception of the substantive relationships at work --*then* one tests the model. If it shows acceptable accuracy, presumably one's previous substantive conception provides an interpretation of that happy result. Neural network models, however, work quite well without anyone knowing the underlying function; so discovering the latter remains a task. In that sense, the usual modeling process is stood on its head; once determined, the effective synaptic weights will themselves be new data telling us about the structure of relationships among the old. And the model becomes a new way of generating, by simulation, the observable results of the functional relationships. Thus the neural network is also a measuring instrument, revealing previously unseen data. In the present context, explaining that data will then become part of the agenda for understanding and predicting proliferation.

### 4. An Approach to Training a Proliferation Model

A list of proposed input and criterion data is given in the Variable List. As indicated there, many of these data already exist as machine readable files; others will need to be developed for this study. People with relevant skill and considerable prior experience in such data development are currently available for this task. The basic scheme that we will use is a follows: Each training case will consist of one year of experience for one state, where the range is all states in each year from 1946 to circa 1995; optionally, the experience of states which were potential or actual nuclear or chemical weapon developers during the Second World War also might be included for purpose of comparing findings. The basic output will be that a given state did or did not take a specified proliferation step in the year following the input data year. Several such steps are identified in the Variable List. For each case, inputs will be of five types: information concerning

1) characteristics and behavior of the referent state;

2) compounds in the form of various sums of data describing all states that are geographically contiguous to the referent state, including their relationships to the referent state;

3) each of five politically “major” powers (USA, Russia, China, Britain, France), including their relationships to the referent state;

4) compounds in the form of various sums of data describing all sovereign state dyads other than the dyads of items 2 and 3;

5) collective characteristics and behavior of the global system of states.

The inputs of items 2 through 4 are of the form of some quantity describing a state, multiplied by a factor connecting that state to the referent state. For example, suppose the referent state was Pakistan, the descriptive quantity was annual change in military personnel as a percent of total population, and the connecting factor was geographical contiguity (dichotomously = 0 or 1). Then annual change in military personnel of India as a percent of total population of India would be multiplied by 1, signifying that India and Pakistan are contiguous, and this change quantity would be summed into a quantity describing the change in military personnel of all states contiguous to Pakistan. This summed quantity would be input to the network as a descriptor of the environment of Pakistan. More complex connecting factors would also be used-- for instance factors describing the extent to which the referent state and the state being described have a recent history of mutual hostility, or mutual cooperation, etc. For example, the military personnel factor describing India would be weighted by the extent of past hostility between that state and Pakistan, which quantity would be summed into another input describing the environment of Pakistan. The choice of indicators in the variable list reflects various ideas by international relations students and others, concerning which kinds of such environmental factors might be systematically influencing proliferation or non-proliferation behavior.

As the above suggests, there is no shortage of interesting, relevant inputs and permutations thereof to a neural network proliferation modeling scheme. Given the large potential number of training cases in the form of many nation state-years, it will be possible to test a fair assortment of such variables in alternative combinations. However there is a limitation imposed by the temporal pattern of cases. The system membership criteria admit most or all entities that one would likely consider a nation state today; as one goes back to earlier years, however, the number of qualifying entities gradually declines. (For instance, Ukraine qualifies as a state from 1992 onward, but not before.) Correspondingly, the available data decline as well. (Some data are not clearly defined for non-state actors; others are uncollected or completely unknown.) Thus a data matrix of entities-variables versus years has a trapezoidal shape, with the longer base corresponding to the most recent year and the shorter one to the earliest year of interest. This implication is: while one might wish to test a model in which all, or many, states separately are providing inputs to each other, the trapezoidal data structure imposes a tradeoff between number of nations and number of years of data.

In the approach outlined here, the tradeoff is resolved in favor of many data years; the affect of the five major powers are separately represented in each case; and the affect of other nations is subsumed into the two groups of weighted sums of data from geographically contiguous states and from all others, respectively, referred to as items 2 and 4. In recognition of the possibility that the system has "memory" (as discussed in the prior section), data from several preceding years could be input; we will begin with, say, the 10 years preceding each data year; later training runs could try other time intervals. Annual cumulation and incremental versions of variables, discussed above, would be tried as well. In the first step, all candidate inputs will be used to estimate a single global model the synaptic weights of which are inferred from all data years. At the end of this first step, inputs will be deleted which appear to make less important contributions to the accuracy of the outputs and the model will be re-estimated using the reduced set of inputs. (At a later juncture, results from this initial inquiry may provide information to help choose an input structure that makes more efficient use of inputs coming from other states.)

## 5. Limitations

As mentioned above, the initial effort will stop short of a fully dynamic model. In addition, many conceivable inputs are omitted from the proposed study; these include variables describing interstate effects which have been missed in the compounding and summarizing process; internal attributes and behaviors, such as human attitudes; and environmental stresses and other non-societal elements which may impinge on state actions.

Many of the inputs which are incorporated are dichotomous, denoting the presence or absence of some characteristic, such as “major power status”, the occurrence or non-occurrence of some event such as the formation or termination of an alliance, and the proliferation steps themselves. For such variables, the more realistic approach to quantification may be to treat the event as random and estimate the underlying propensity or likelihood that the dichotomy will alter its value. The change in such an underlying, continuous variable would then be the corresponding model output. This suggests two kinds of issues. First is the problem of how to represent each such underlying propensity and to estimate its numerical values. Second is the need for a Monte Carlo -like element to translate the propensities into outcomes of events which actually do or do not occur in each time period. Instead, to get a start modeling proliferation, we would begin with the simpler approach of treating discrete events as deterministic processes, to be emulated directly by a neural network.

Overcoming these limitations will be possible points of departure for modeling efforts which follow the initial effort outlined here.

## 6. Variable List

Abbreviations:
End notes are numbered within parentheses: ( )
NPT = nuclear nonproliferation treaty.
MID = militarized interstate dispute (See end notes 1).
[item __] refers to the __ numbered item elsewhere in the text.
[C] means the indicated data are collected in machine-readable form. (See end note 2)
[P] means the indicated data are partly collected in machine-readable form.
[*] means the indicated data still need to be collected.

(0,1) denotes a dichotomous variable with values 0 and 1 ; these values correspond to the first and second options appearing on either side of a / character in the variable description.

Variables not so denoted are continuous.

All inputs are with reference to a specified year.

1. Input types by party:
self
contiguous neighbors
majors
  1 (US)
  2 (UK)
  3 (Frn)
  4 (Rus)
  5 (Chn)
non-contiguous non-major other system member inputs
global inputs

2. Self input subtypes

2.1. Self characteristics
material attributes [item 7.1]
possession of mass-destructive technology [item 7.2]
political regime [item 7.3]
diplomatic status [item 7.4]

2.2. Self behavior
Civil war is not / is in progress (0,1) [C]

2.3. Variable changes [item 10]

3. Neighbor input subtypes
Definition: a neighbor is a state contiguous to the referent state [item 11.1.1]

3.1. Couplings summed across all neighbors [item 8]
3.2. Weighted couplings summed across all neighbors [items 8, 9]
3.3. Variable changes [item 10]

4. Major state input subtypes; referent state in relation to Major 1 ... Major 5
4.1. Couplings between referent state and major [item 8]
4.2. Weighted couplings between referent state and major [items 8, 9]
4.3. Major has demanded that party discontinue a weapon program (0,1) [*]
4.4. Variable changes [item 10]

5. Non-contiguous non-major other system member input subtypes
5.1. Couplings summed across all others [item 8]
5.2. Weighted couplings summed across all others  [items 8, 9]
5.3. Variable changes [item 10]

6. Global input subtypes
6.1. A fission weapon has not / has been tested (i.e. pre- vs. post 7/45)  (0,1)
Comment: this coding will allow Second World War experience to be integrated or segregated as desired.
6.2. A radiation implosion fusion weapon has not / has been tested (0,1) [*]
6.3. The NPT has not / has taken effect (0,1) [*]
6.4. MID onset hazards summed across all dyads excluding referent state as a party [item 11.2.4]. (See end note 3.)
6.5. Number of parties having taken steps to acquire mass-destructive technology [item 7.2]
6.6. Number of parties claiming to have terminated a program [item 7.2.11]
6.7  Number of parties destroying weapons [item 7.2.12]
6.8. Variable changes [item 10]

Items 7 through 11 refer to details of inputs.

7. Characteristics of states

7.1. Material attributes

7.1.1 Size attributes
    Energy consumption [C]
    Iron / steel production [C]
    Military personnel [C]
    Military expenditure [C]
    Urban population [C]
    Total population [C]

7.1.2. Size attributes / total population

7.1.3. Non-population size attributes / urban population

7.2. Possession of mass-destructive technology [*]  (5)

Party has not / has (0,1)

    For:
        nuclear weapons ...

7.2.1. Acquired or refined weapon-quantities of uranium ore or metal
7.2.2. Acquired or separated weapon-grade concentration of U235
7.2.3. Acquired or synthesized Pu239
7.2.4. Acquired or refined weapon-quantities of neutron initiator materials

Comment: acquisition includes theft.

For:
C/B weapons ...

7.2.5. Begun weapon-quantity production of active material

For:
gun-type fission weapons ...
HE implosion fission weapons ...
Levitated-HE implosion weapons ...
TN boosted or radiation implosion weapons ...
Biological agent weapons ...
Chemical agent weapons ...

7.2.6. tested warhead or acquired test data
7.2.7. produced warheads
7.2.8. tested delivery vehicle or acquired test data
7.2.9. produced delivery vehicles
7.2.10. deployed weapons to launch sites
7.2.11. claimed to terminate weapon development program
7.2.12. destroyed weapons

Comment: acquisition of test data includes espionage.

7.3. Political regime characteristics [C]

7.4. Diplomatic status
7.4.1. party is not / is an international system member (0,1) [C]
Comment: "system membership" is based on certain specific population and diplomatic criteria.
7.4.2. party is not / is major state (0,1) [C]
7.4.3. party has not / has ratified the NPT (0,1) [*]

8. Couplings between referent party and named other party

8.1. Dichotomies (0,1)
8.1.1 The party other than the referent party is not / is an international system member [C]
8.1.2 With respect to each of several relationships the parties do not / do show that relationship [item 11.1]

8.2 Continua
8.2.1. Time since prior change in 8.1.1 (4)
8.2.2. Several other couplings [item 11.2]

9. Weighted couplings
Definition: weighted coupling = coupling to other × characteristic of other

types of other-party characteristics:

9.1. party has not / has had an unconstitutional change of government in the referent year (0,1) [item 7.3]
9.2. party has not / has taken steps to acquire mass-destructive technology (0,1) [any of items 7.2]
9.3. military personnel / total population [item 7.1.2]
9.4. military expenditure / urban population [item 7.1.3]

10. Variable changes
10.1. Dichotomous variable has shifted from 0 to 1 in referent year (0,1) *
10.2. Dichotomous variable has shifted from 1 to 0 in referent year (0,1) *
10.3. Increment (+ or -) in continuous variable compared with prior year value
*For instance, in referent state civil war has not / has
begun in referent year (0,1)
ended in referent year (0,1)

11. Relationships / behaviors between pairs of states (dyads)

11.1. Dichotomies; in the referent year dyad members do not / do ... (0,1)
11.1.1. have land boundary or narrow water body in common [C]
11.1.2. maintain mutual military alliance, neutrality, or consultative treaty [C]
11.1.3. maintain mutual diplomatic representation [P]
11.1.4. participate on opposing sides in an ongoing interstate war [C]
11.1.5. participate as allies in an ongoing interstate war [C]
11.1.6. participate on opposing sides in an ongoing non-war MID [C]
11.1.7. participate as allies in an ongoing non-war MID [C]

11.2. Continua
11.2.1. Time since prior change in items 11.1.1 - 11.1.7 (4)
11.2.2. Number of mutual military treaties [C]
11.2.3. Number of common inter-governmental organization memberships [P]
11.2.4. Event hazard, onset of mutual opposition in MID
Comment: hazard determined by waiting times analysis of MID onsets. (See end note 3.) Status of input data is [C]; but analysis needs to be re-done.

End notes:

(1) MID data are described by Gochman and Maoz 1984, pp. 587-589. Data record all instances of “threat to use force”, “display of force”, “use of military violence short of war”, and “war.”

(2) Existing data sets include indicators of the material capabilities and cultural characteristics of interstate system members; their dates of entry to and exit from the system, and classification to major or minor power status; their mutual diplomatic, alliance, and organizational ties; their involvement in inter-state and civil wars; and the presence or absence of geographical contiguity for each dyad (Singer and Small 1972, Small and Singer 1980). Additional data have been developed, having the same domain, which describe the political characteristics of the governing regimes of these system members (Gurr 1994). Additional uses of such data in neural network global system modeling are suggested in Williamson (1995).

(3) By “hazard” is meant roughly the rate at which the event in question, in this case an MID onset, is expected to occur per unit time, given the conditions yielding the hazard. Consider a group of entities, each of which is at “risk” to experiencing a given event. Call this group the “risk set”; and whenever one of the entities in question experiences the given event, consider this entity to be removed from the risk set. Let s denote the elapsed time since the entities in question were first members of the risk set and let f (s) be the proportion of the original membership still in the set at time s; by definition f (0) = 1. Then the hazard rate h (s) is defined

by $h(s) = -d \ln f(s) / ds$ (Tuma and Hannan 1984). This definition meshes with other statistical concepts; for instance $h(s)$ = constant defines a Poisson distribution for the event.

Knowing when each MID involving interstate system members began, as one does from available data, allows one to generate a corresponding series of hazard rate estimates for each dyad in each year that both parties were members (Williamson, Warner and Hopkins 1988). To estimate the total expected rate of onsets, one would multiply the hazard times the number of dyads at risk to dispute or, for different rates, sum across dyads.

(4) Elapsed time may contribute in two ways. First, elapsed time since a most recent prior event is a surrogate for the hazard rate of recurrence of the event. Plausible values of the latter vary in a regular way with elapsed times after previous such events (but, given sufficient resources, we would prefer to estimate the hazard rates of these events and use them directly as model inputs). Second, times since 2nd and previous occurrences may be surrogates for unobserved factors that contribute to hazards and other variables. (For instance, the hazard of an MID onset may reflect the combined effect of how long it has been since the most recent *several* such onsets.)

(5) Parties Involved with Nuclear Weapons:

Most nations in most years have not, of course, exhibited proliferation behaviors; so there is no shortage of negative examples. The question then is whether there are enough examples of such behavior to provide sufficient many positive cases. The following list of known or alleged programs, which concerns only nuclear weapons, suggests there is; and chemical / biological weapons will provide several additional cases.

Tried, succeeded, announced:
USA
UK
USSR / Russia
France
China
India
Pakistan.

Tried, succeeded, unannounced:
Israel

Tried, program underway:
Iran
North Korea

Tried, stopped by military defeat:
German Third Reich
Imperial Japan
Iraq

Tried, program terminated by choice (?):
South Korea
Taiwan
South Africa
Brazil
Argentina
Kazakhstan*
Ukraine*
Belarus*
Libya

*weapons acquired upon dissolution of USSR

In addition, there seem to be a handful of states that have explicitly chosen not pursue a nuclear development program (at least for now).

Chose not to try:
Canada
German Federal Republic
post-1945 Japan.

**Citations**

Gochman, Charles S. and Zeev Maoz (1984). "Militarized Interstate Disputes, 1816-1976: Procedures, Patterns, and Insights." *J. Conflict Resolution* 28 (Dec.): 585-616.

Gurr, Ted (1994). “Peoples Against States: Ethnopolitical Conflict and the Changing World System.” *International Studies Quarterly* 38, No. 3 (Sept.): 347-377.

Hammerstrom, Dan (1993). “Neural Networks at Work”. *IEEE Spectrum* (June) , pp. 26-32.

Haykin, Simon (1994). *Neural Networks. A Comprehensive Foundation*. New York: Macmillan.

Karasik, Myron S. and Paul R. Williamson (1994). "Modeling Complex Human Social Dynamics Using Neural Networks of Fuzzy Controllers," *Proceedings*, World Congress on Neural Networks, San Diego.

Murray, Carl (1993). “Is the Solar System Stable?” in Hall, Nina (ed.) *Exploring Chaos: a Guide to the New Science of Disorder*. New York: W. W. Norton (pp. 96-107).

Rich, E. and K. Knight (1991). *Artificial Intelligence*. New York: McGraw Hill.

Singer, J. David and Melvin Small (1972). *The Wages of War, 1816-1965: A Statistical Handbook*. New York: Wiley.

Small, Melvin and J. David Singer (1980). *Resort to Arms: International and Civil Wars, 1816-1980*. Beverly Hills, CA: Sage.

Williamson, Paul R. and Myron S. Karasik (1994). "A Proposal for Developing and Implementing a Global Information and Forecasting Service." Presented at Chicago Area Sigma Xi Clubs Forum on Sustainable Development, Loyola University, Chicago. Ann Arbor, MI 48106-4394, P. O. Box 4394: Global Vision, Inc.

Williamson, Paul R. (1995). “Capabilities Concentration, Dispute Contagion, and Global War: Findings, Synthesis, and Speculations.” Ann Arbor: Correlates of War Project, Dept. Pol. Sci., University of Mich. Unpublished.

Williamson, Paul R., John Warner, and Stephen A. Hopkins (1988). "A Model of International Dispute Onsets With Preliminary Application to the Impact of Nuclear Weapons." Ann Arbor, MI: Correlates of War project, Dept. Pol. Sci., Univ. of Mich. Unpublished.