Session No. 642

Dynamic Modeling—An Approach for the Design of Loss Prevention Programs

Shakirudeen O. Shakioye
Chevron Corporation
Houston, Texas

Joel M. Haight
The Pennsylvania State University
Department of Energy and Mineral Engineering
University Park, Pennsylvania

Abstract

The National Safety Council's records show that in 2004 alone, on-the-job injuries to workers constituted 35% of total recorded injuries in the United States. This generated an associated cost of about $142.2 billion. Unfortunately, the safety intervention programs enforced at work places to mitigate such losses are driven mainly by intuition and experience of involved safety personnel. This paper details implementing a computer program to furnish safety personnel with an empirical basis for designing loss prevention programs based on historical safety data. The computer tool is driven by a dynamic mathematical model which adapts itself to variations in data patterns and explains the correlation between historical incident rates and corresponding resources committed to interventions. This study empowers the industry with a tool that is capable of forming the core of optimizing the allocation of valuable human resources in safety program designs.

1. Background and problem statement

The Bureau of Labor Statistics records show that the natural resources and mining industry alone recorded 850 fatalities in 2003, 821 fatalities in 2004 and 873 fatalities in 2005 in the United States. In 2005 alone, the National Safety Council reported $156.6 billion due to on-the-job injuries to workers.

All above are indications of a genuine need for an unrelenting preventive safety intervention process or a system implemented by all industries. The importance of preventative measures or safety interventions can be estimated by the cost of workplace incidents resulting from worker's injuries and illnesses, fatalities, downtime and property loss. The cost associated with incidents includes dollars spent and dollars lost as income due to the incident.

The Health, Environmental and Safety (HES) function as a community is clearly faced with the problem of identifying whether implemented interventions have a positive effect or even if they do more harm than good. Intervention design becomes critical since resources are limited and must be put to judicious use [Shannon *et al.*, 1999]. Limited resources could be in the form of man-hours available for use in the different facets of a safety intervention program to yield the best results. In view of the fact that engineering processes/ programs become applicable as useful tools in any industry if they help make economic sense, it is therefore intended that some engineering practices will be introduced to solve some of these problems.

The research outlined in this paper includes a discussion of the building of a resource allocation tool to determine how to optimize the allotment of valuable man-hours to each safety program intervention element for incident rate minimization.

A five year study by the National Institute for Occupational Safety and Health [NIOSH, 1999] proves a decline in incident rate is a consequence of the level of applied intervention or its effectiveness. Exhibit 1 shows an exponential decaying relationship between incident rate and total man-hours applied to safety intervention processes. The graph indicates that at some point an additional allocation of man-hours does not necessarily impact incident rate reduction substantially [Haight *et al.*, 2001a].

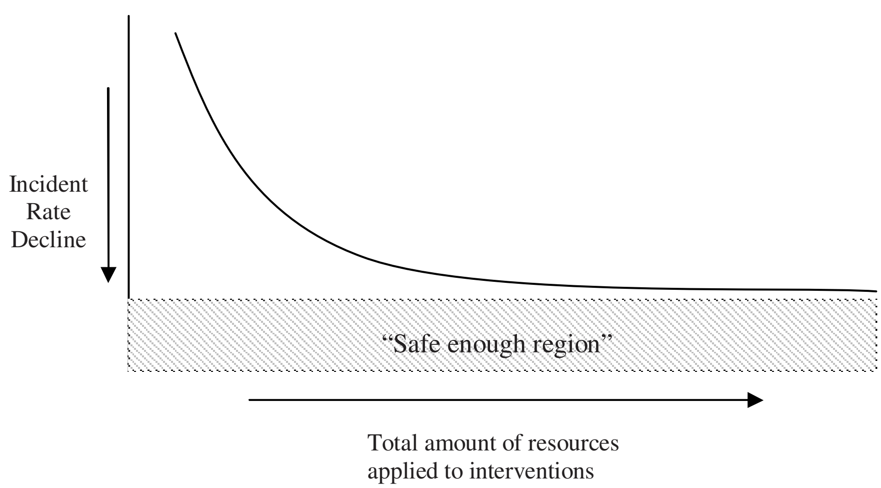

Exhibit 1. Incident rate decline curve.

The safe enough region under the decline curve could be a smaller area depending on how well a safety program performs. Although, having a zero tolerance for workplace incidents would be ideal, it can be quite challenging to have this safe enough region collapse to zero. This line of research is bringing about the ability to answer the question "what quantity is enough to reach the safe-enough region?" of a company [Iyer *et al*., 2005].

Unlike the conventional non-mathematical approaches in the industry, several studies show safety intervention programs can be mathematically designed by optimizing the representative model [Haight *et al.,* 2001b; Iyer *et al*., 2004]. For practicability and everyday use within the industry, a computer program was developed using a dynamic model that has the capability of repeatedly establishing an updated relationship between the intervention factors (independent variables) and the incident rate (dependent variable) as more data are added to the database from day to day operations of the company.

The goal is for the user to have a tool that generates suggestive man-hour allocation ratios by optimization (minimizing resource allocation while concurrently minimizing incident rate). This tool forms an additional resource from which the user can make decisions on future design of safety intervention programs. It enables safety personnel to determine where or on what interventions to put supplementary efforts in terms of valuable resources. Based on historic data and suggested man-hours, the model also has an ability to forecast incident rates.

2. Methods

Data were collected for a period of 30 weeks from a power distribution company. A weekly data sheet was used to record man-hours allocated to intervention activities peculiar to the company's safety program. The incident rates were also recorded accordingly but not much information was available to qualitatively measure the intervention process. The safety activities like job planning, safety trainings, technical trainings, audits and assessments, facilities inspections, preventive maintenance activities etc (referred to as interventions) were categorized into 4 main factors by expected similarity of effect.

Factor A $\rightarrow$ Safety Awareness and Motivation Activities (x_1)

Factor B $\rightarrow$ Skill Development and Training Activities (x_2)

Factor C $\rightarrow$ New Tools and Equipment Design Methods and Activities (x_3)

Factor D $\rightarrow$ Equipment Related Activities (x_4)

There have been several studies to show the possibility of quantitative and qualitative assessment of intervention efficiency, its optimization and design. These studies have been verified within the industry by implementing the suggested design (better mix of man-hour allocation to individual interventions) from optimizing a static mathematical model, which is a representation of the intervention program peculiar to that company [Iyer *et al*., 2004].

2.1 Bridging the gap

Any mathematical model that defines an intervention program should have the capability to adapt itself to changes in the characteristics of the intervention program which could be a direct effect of changes in company safety policies, employee attitude or other external effects such as more stringent legislative constraints. Taking another step further from a one time intervention optimization process to incorporating the concept into the daily decision making within the industry, a computation tool with the required simplicity then becomes inevitable. Also, to bridge the gap between research and industry, a computer tool with the ability to create a functional user friendly interface becomes critical.

2.2 Basic levels of intervention model/ algorithm

The initial intervention data collected over a 30 week period forms the initial database of the computer tool, and it is stored as a matrix in a text file to be accessed and updated at will. Exhibit 2.2 presents an illustration of the flow path of weekly data for each run of the tool.

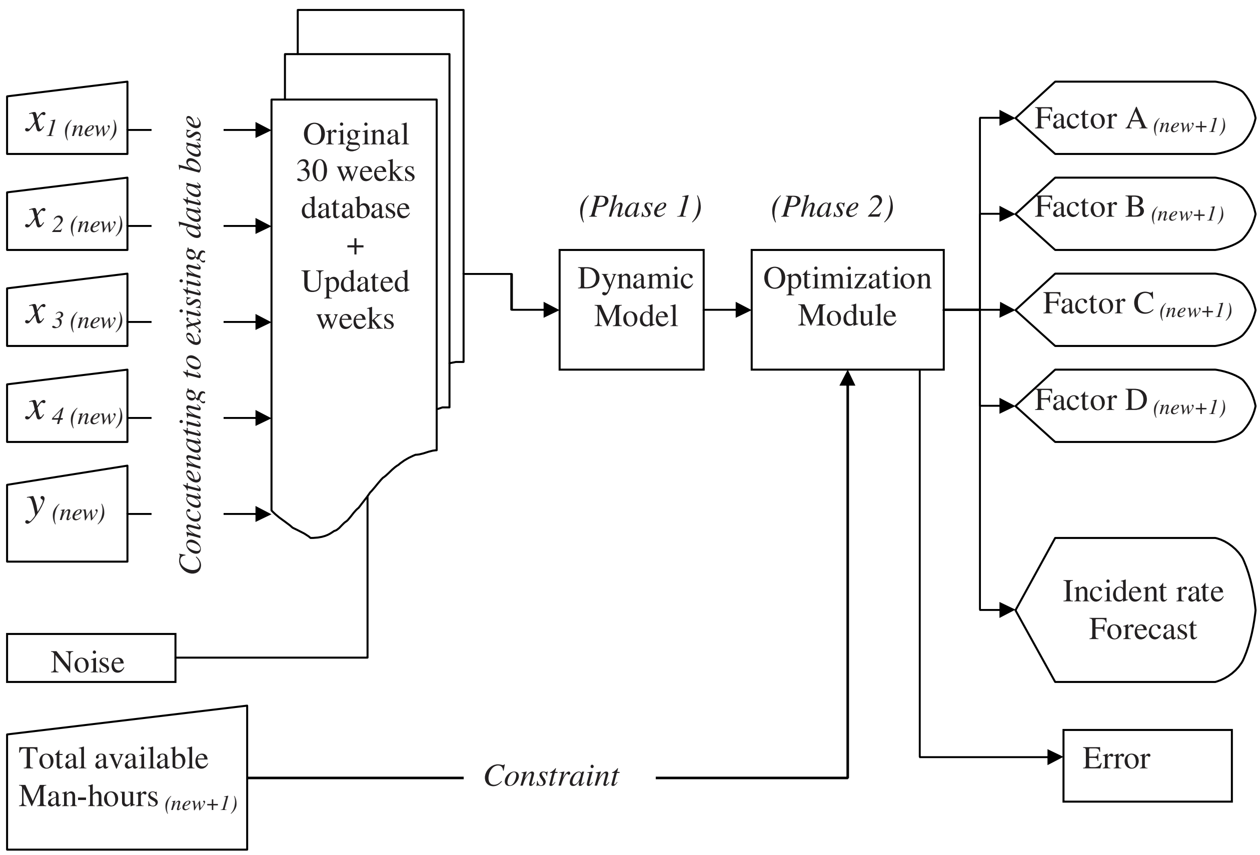

Exhibit 2.2 Flow path summary.

The tool as shown above is developed such that it continually accepts, updates and concatenates man-hours allocated to individual intervention factors (independent variables $x_{1...4}$) and its corresponding incident rate (dependent variable y) for that week onto the existing database. The database over time is expected to grow and the updated database carries with it the company's safety intervention history (accommodates expected changes in the characterization of safety programs over time).

For every run of the program, the computer updates the coefficients of the mathematical model using regression analysis; this in turn produces a model that better represents the system with consideration giving to the effects of the newly introduced data points. The coefficients of the

representative model are passed along to the optimization module. The optimization module then generates the better mix of apportioned man-hours for each of the factors that will yield a minimized incident rate ($y \rightarrow 0$). The optimization is done such that the total amount of available man-hours for the planning week is given and forms a constraint for the apportioning values to the suggested man-hours;
Available man-hours = Suggested: factor A + factor B + factor C + factor D

2.3 Source of errors

The organization of safety management is made up of two main components which are the technical and human systems as shown in Exhibit 2.3.

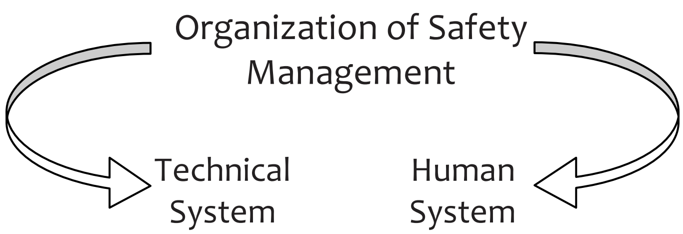

Exhibit 2.3 Organization structure.

The technical system forms a greater part of management planning and includes all controllable measures/ policies that are put in place that have a direct bearing on reduction of incident rates to a substantial degree. On the other hand, the human system consists of the human behavior which is quite complicated and cannot be predicted [Brauer, 1990]. Behavioral patterns of humans vary and are a function of physiological conditions, individual opinions, psychological state and so on [Brauer, 1990]. The unpredictable nature of the human factor contributes significantly as a source of error to the model thereby distorting actual correlation between the technical system (interventions) and incident rates [Brauer, 1990].

For the purpose of this research, another source contributing to the error term is the quality of intervention, which is not being accounted for by the model. It assumed that the quality of intervention is constant and acceptable all through this research. This is the case because interventions are measured in man-hours and this does not necessarily depict the quality of such intervention.

3. Results and discussion

3.1 Model building

Using the 30 weeks of data supplied by a power company, an attempt was made to establish a relationship that relates the interventions to the incident rates. The schematic in Exhibit 3.1 illustrates the model path that explains a set of weekly intervention application rates of four factors and their corresponding incident rate (for that specific week). The incident rate used in this study is the incident recorded in 200,000 hours per total hours worked. This enables a normalization of the incidents recorded taking into cognizance the size of labor.

However, interventions have been demonstrated to not just have instantaneous effects on their corresponding incident rate, rather their effects last for a while though not forever [Haight, 2001a]. This assumption was re-established and incorporated in this study by applying the model using several moving averages (i.e. leading weeks) of the incident rates (dependent variable).

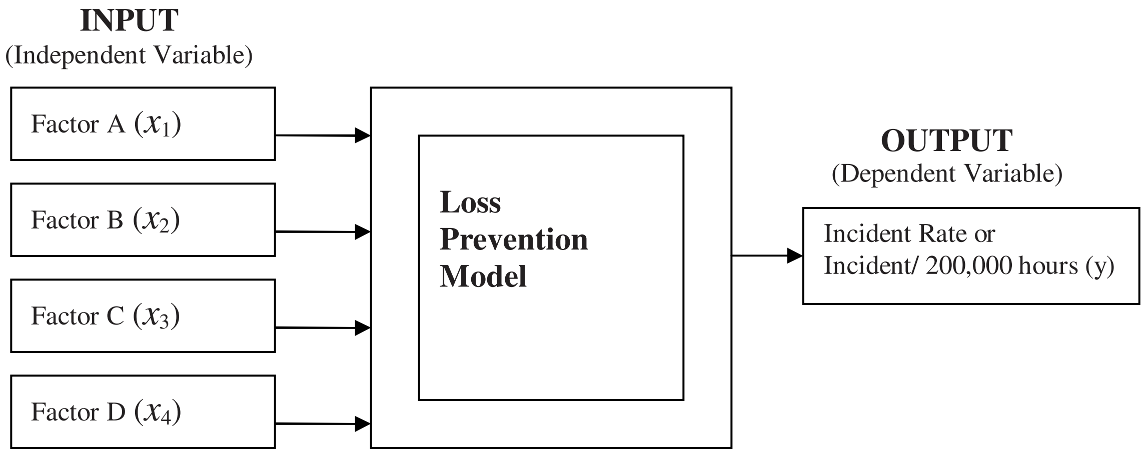

Exhibit 3.1 Representation of the Loss Prevention System Model (Adapted from Haight, 2001a).

Using the four main intervention factors and their interaction variables, stepwise regression was carried out for each moving average type to establish the statistically significant terms. The outcome of the regression indicated best correlation when a six-weeks moving average was employed, which suggests that the intervention for the power company typically lasts for a period of six weeks.

The initial model using four main factors and interaction factors was made up of 15 variables. After carrying out rigorous iterative experiments, other random variables were added and the variables with the least effects were dropped. A stable model was built to represent the fitted data.

Below (Equation 3.1) is a 10 variable final model expressed in the form:

$$y = \beta_o + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_1^2 + \beta_6 x_2^2 + \beta_7 x_4^{0.5} + \beta_8 x_1^2 x_2^2 + \beta_9 x_1^2 x_3 + \beta_{10} x_1^2 x_3^3 \qquad \text{Equation 3.1}$$

where;

y is a six weeks moving average of the incident rate,

β_0 is the model intercept,

$\beta_{1\text{-}10}$ are dynamic regression coefficients,

and ε_i *is* the model error term.

It is important at this point to stress that unlike the conventional regression coefficients and for the purpose of this study, the coefficients are dynamic and will be updated for every newly introduced data point (using computer programming with equation 3.1 implemented as a module). The model has an R^2 of 0.7, which is significant for this class of research as it receives a considerable influence from human factor.

3.2 Model estimates and predictions

The model was designed to have a reasonably small range of ordinary residuals ($y - y$). Where "y" implies actual response and "y", the estimated response calculated using the model. To investigate the predictive ability and further validate the model, the PRESS residual

$$(y_i - \hat{y}_{i,-i})$$

is considered. This option is valid because data splitting would not be practical for the original data (limited to 30 data points). The PRESS residual unlike the ordinary residual is generated by withholding an observation and using the other n-1 observations to carry out the regression that produces coefficients for the model variable. The omitted observation is then evaluated using the model to predict the response (please see plot as shown in Exhibit 3.2).

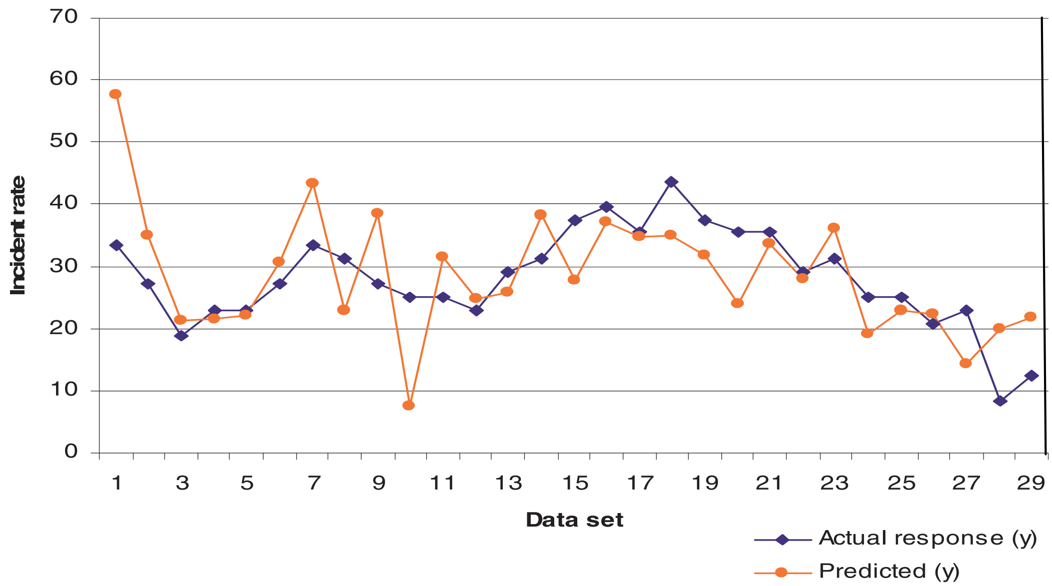

Exhibit 3.2 Actual incident rates versus model predicted values.

The results of the predictions were also close to the actual responses even though the errors as expected are higher than the ordinary residuals. On the other hand, the model has the downside of performing best within the range of values used to build it.

An earlier study carried out by Iyer *et al.* [2004] using the same set of data as the one used for this present research resulted in a 15 variable model. This model (equation 3.1 – with 10 variables) is a successor to the model developed by Iyer *et al.* [2004]. R^2 denotes the portion of the data that is explained by the model and ranges from 0 to 1, the closer to 1 the better the fit. For both models, the R^2 values are 0.7 each. The new model has an adjusted-R^2 value of 0.54 compare to its predecessor with 0.36 (i.e. considering the degrees of freedom). R^2 has a weakness because for each additional variable added to a model, the R^2 value increases even in cases where the added variable reduces the efficiency of the model. To correct this problem, the adjusted-R^2 takes into account the degree of freedom. On the other hand, the adjusted-R^2 declines with the inclusion of additional variables that do not have much impact.

3.3 Optimization

One goal for any business is profit making. To achieve this objective, optimization models are used for carrying out complex business patterns to maximize profit. The safety field has several avenues where optimization models can be used judiciously for decision-making. One of these avenues is the loss prevention or safety intervention program. The health, environmental and safety (HES) departments of larger organizations may have a limited amount of resources allocated to them. The resources could range from financial to man-hours. The scope of this study was limited to the investigation of how to make the best use of the allocated man-hours due to the safety activities.

At this point, the mathematical model has been built to establish the relationship between the factors that make up our data and the incident rate (equation 3.1). The calculated coefficients (using regression) are passed on to the optimization model to generate the variables (percentage of man-hours for each factor) that correspond to minimal incident rate. Exhibit 3.3 shows a schematic of the described process.

Optimization Flow path

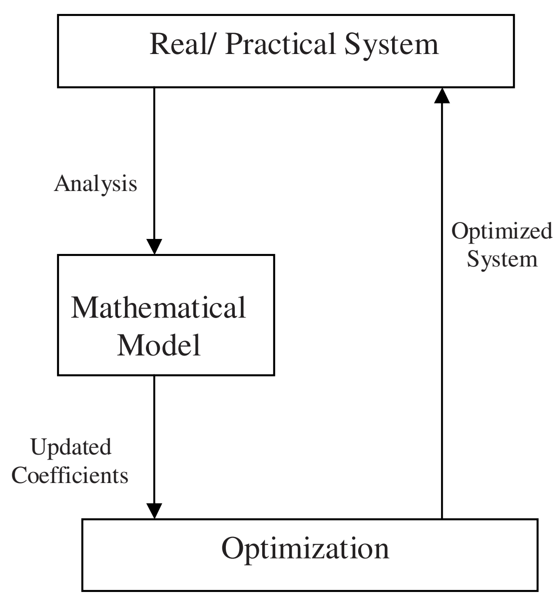

Exhibit 3.3 Optimization of safety system (adapted from Chinneck, 2006).

Studies by Haight *et al.* [2001a] investigated the relationship between the total man-hours/ resources dedicated to any functional safety intervention program and demonstrated that incident rate decline is defined by an exponentially declining curve.

3.3.1 Actual data behavior

Regression is employed to generate the exponential equation that defines the relationship between total applied intervention and incident rate, which will change with the dynamics of the data over time. For every additional datum point, running the model carries out regression to redefine this relationship and updates the coefficients of the exponential equation. Using the present available data, the curve created is as shown in Exhibit 3.3.1, and the governing equation of the form: y = α_1. + α_2 exp (-x). This relationship relates the total available man-hours to incident rates but does not solve the challenge of optimally apportioning the available man-hours to the individual factors that make up the intervention program. The coefficients and intercept obtained from equation 3.1 after regression has been carried out are passed onto the optimization module such that the values of x_1, x_2, x_3 and x_4 (i.e. factors A to D) are determined as y tends to zero (incident rate $\rightarrow 0$).

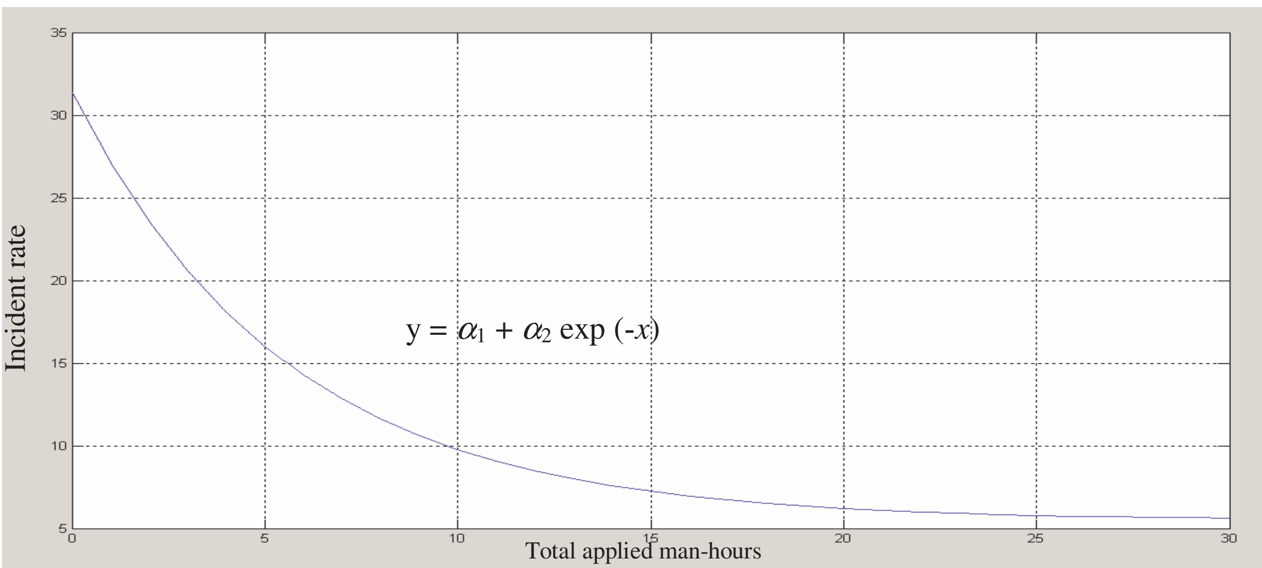

Exhibit 3.3.1 Fitted graph of actual data.

3.3.2 Optimization module and sample of generated results

A module of the computer program is called upon to perform a constrained linear least-square optimization on equation 3.1. An upper bound and a lower bound matrix are constructed to define the range within which the optimization takes place. Successive iterations are carried out until the program converges to the solution. This module receives the generated updated coefficients from the model (equation 3.1) that describes the relationship between the intervention factors and the incident rates (i.e. an outcome of the regression).

Carrying out a one-time optimization of the 10 variable model, an optimal combination of factors to yield the minimal incident rate was established as displayed on in Table 3.3.2.

	Factor A (% man-hours)	Factor B (% man-hours)	Factor C (% man-hours)	Factor D (% man-hours)	Total Suggested % man-hours
Present Model (equation 3.1)	6.0	1.0	0.002	3.9	**≈ 11 %**
Iyer *et al* [2004]	10.20	0.1	0.6	4.1	**15%**

Table 3.3.2 Comparison of suggested Optimization results.

In comparison to the study done by Iyer *et al* [2004] as shown in Table 3.3.2, the optimization of the new model suggests that a total of 11% of the company's man-hours is sufficient to adequately design their loss prevention program against the previously suggested 15%. The supplied validation data was analyzed and it was found that the power company hardly met the initial suggested 15% man-hour allocation but recorded a substantial decline in incident rates. Data sets with 0 and 1 incidents were sampled and 39.2% were as a result of implementing total man-hours between 8 and 10% while others were randomly represented. This is an indication that 11% is quite practical to fall within a comfort zone to minimize incident rates.

The ratios of man-hours suggested for allocation to each factor by the present model are quite similar to that of Iyer *et al* [2004]. Factors A and D are indicated to require substantial man-hour dedication relative to factors B and C. These results are not intended to be an absolute solution, but rather, they should complement or give a strong backing on intuition for decision-making in loss prevention design.

4. Conclusion and future work

This study further reinforces the earlier studies carried out by Haight *et al.,* [2001a] and Iyer *et al.,* [2004] on the possibility of mathematically modeling the relationship that exists between safety intervention activities and incident rates. This research resulted in building a more efficient model than its predecessor produced by Iyer *et al.,* [2004]. Investigations carried out in this study clearly indicate that for complex or multiple factored intervention analysis, the interaction effects are quite significant. Hence, isolating each factor and studying its direct effects on incidents will not be adequate.

This project takes a step further in building a dynamic model rather than the conventional static mathematical model. The implementation of the mathematical model as a driver of a computer program helps create a working and applied tool that empowers the user in making safety intervention planning decisions. Although the computer model adapts itself to reasonable changes within specific ranges of values, long term periodic adjustments should be required in case of huge changes in company policies that have direct bearing on intervention activities.

The results from the analysis of the newly built model highlight significant prospects in mathematically optimizing a loss prevention system. This is important because limited resources

(available man-hours) are optimally allocated to each intervention activity to minimize incident rates. The resources allocated to safety intervention programs and the consequential magnitude of incident rates are both cost drivers. Hence, optimizing the system invariably manages these associated costs.
Consistent with Iyer *et al,.*[2004] and based on the 30 weeks of data, optimizing the model in the developed computer coder indicates that Factors A and D (*behavior, awareness, motivation* and *equipment related activities* respectively) require the most allocation of resources to minimize incident rates.

There is a wide range of opportunity presented by applying this concept of automation for future work. The automation of the mathematical relationship that exists between the quantified applied intervention and incident rates is critical to these opportunities. Such opportunity includes incorporating quality and cost measures into model building and investigating how applicable a single model can be to all companies.

It is imperative to stress that the model is applicable only to the power company whose data was used to construct it. With continual collection of data from different companies, future studies will result in a more generalized model that can be applicable to several companies within an industry with quite standard or common intervention factors. This may be achieved by building a computer program capable of restructuring the mathematical model by selecting a combination of variables.

References

Brauer, R. L. (1990). Safety and Health for Engineers.

Department of Labor: US factory inspection legislation http://www.dol.gov/oasam/programs/history/mono-regsafepart02.htm (Accessed December 18th 2006).

Haight, J. M., R. E. Thomas, L. A. Smith, R. L. Bulfin Jr. and B. K. Hopkins (2001a). "Intervention Effectiveness Research: Phase 1 Developing a Mathematical Relationship Between Interventions and Incident Rates for the Design of a Loss Prevention System." American Society of Safety Engineers Technical Forum: 38-44

Haight, J. M., R. E. Thomas, L. A. Smith, R. L. Bulfin Jr. and B. K. Hopkins (2001b). "Intervention Effectiveness Research: Phase 2 Design, Optimization & Verification of the Loss Prevention System & Analysis Models." Professional Safety - The Journal of The America Society of Safety Engineers : 33-37.

Haight, J. M., Thomas, R. E. (2003). "Intervention Effectiveness Research: A Review of the Literature on Leading Indicators." Division of Chemical Health and Safety of the American Chemical Society: 21-25

Iyer, P.S., J. M. Haight, E. Del Castillo, B. W. Tink, P. W. Hawkins (2004). "Intervention Effectiveness Research: Understanding and Optimizing Industrial Safety Programs Using

Leading Indicators." Division of Chemical Health and Safety of the American Chemical Society: 9-19.

Iyer, P. S., J. M. Haight, E. Del Castillo, B. W. Tink, P. W. Hawkins (2005). "A Research Model-Forecasting Incident Rates from Optimized Safety Program Intervention Strategies." Journal of Safety Research 36: 341-351.

John W. Chinneck. (2006). Practical Optimization: A gentle Introduction

Myers, R. H. (1986). Classical and Modern Regression with Application.

National Safety Council, Safety Intervention Evaluation: A Systematic Approach http://www.acgih.org/events/ControlBand/Thomas_SafetyIntervention.pdf. (Accessed January 10th 2007).

NIOSH, US Department of Health and Human Services: Performance Highlights of the National Institute for Occupational Safety and Health. 1995 – 1999.

Occupational Safety & Health Administration (OSHA) Challenge Pilot Program http://www.osha.gov/dcsp/vpp/challenge.html.(Accessed January 15th 2007).