Risk-based modeling of early warning systems for pollution accidents

W.M. Grayman* and R.M. Males**
* W.M. Grayman Consulting Engineer, 730 Avon Fields Lane, Cincinnati, Ohio 45229 USA
(E-mail: grayman@fuse.net)
** RMM Technical Services, Inc., 3319 Eastside Ave., Cincinnati, Ohio 45208 USA
(E-mail: males@iac.net)

Abstract An early warning system is a mechanism for detecting, characterizing and providing notification of a source water contamination event (spill event) in order to mitigate the impact of contamination. Spill events are highly probabilistic occurrences with major spills, which can have very significant impacts on raw water sources of drinking water, being relatively rare. A systematic method for designing and operating early warning systems that considers the highly variable, probabilistic nature of many aspects of the system is described. The methodology accounts for the probability of spills, behavior of monitoring equipment, variable hydrology, and the probability of obtaining information about spills independent of a monitoring system. Spill Risk, a risk-based model using Monte Carlo simulation techniques has been developed and its utility has been demonstrated as part of an AWWA Research Foundation sponsored project. The model has been applied to several hypothetical river situations and to an actual section of the Ohio River. Additionally, the model has been systematically applied to a wide range of conditions in order to develop general guidance on design of early warning systems.

Keywords Early warning systems; Monte Carlo simulation; simulation; uncertainty

Introduction
Early warning systems (EWS) are a mechanism for detecting, characterizing and providing notification of a source water contamination event, in the hopes that such early warning will allow for rapid mobilization to mitigate the impact of contamination. Many early warning systems established around the world were created in response to major spills. Examples include the 1986 Sandoz spill on the Rhine River, the 1976 carbon tetrachloride spill on the Ohio River, USA and the 1984 phenol spill on the River Dee, UK. Such major spills, which can have very significant impacts on raw water sources, are relatively rare events. Minor spills are much more common but generally have little impact on the water bodies. In fact, the absence of additional major spills has led to suggestions for curtailment of budgets for some existing early warning systems.

An EWS typically will consist of some mechanism for detecting a contamination event through automated monitoring, coupled with a model for predicting the movement of the contaminant in the water body, and an infrastructure for managing and disseminating this information to affected parties. Significant investments in monitoring equipment, staff, operations and maintenance, and organizational infrastructure are necessary for such systems.

Currently there are no rational methods for assessing and justifying the economics of early warning systems, nor for designing them. As a result, it is likely that early warning systems have not been developed in many situations where they would be justified and also many existing early warning systems are under attack as economically unjustified because there have not been any recent contaminant events on those rivers. There are no systematic methods for designing early warning systems that consider the highly variable,
probabilistic nature of many aspects of the system. These include probability of spills, spatial distribution of spills, behavior of monitoring equipment (including false positives and false negatives), variable hydrology, and the probability of obtaining information about spills independent of a monitoring system, through self-reporting or information from other agencies or the public.

In a recently-completed study sponsored by the AWWA Research Foundation entitled “Design of Early Warning and Predictive Source-Water Monitoring Systems”, the need to develop a risk-based perspective on the problem became clear. This led to the development of a Monte Carlo simulation model for analysis of early warning systems. This model, referred to as Spill Risk, is general in nature, and has been applied initially to the Ohio River, and to abstract situations to aid in development of EWS policies and design criteria. A complete technical description of the model is contained in the final report of the AWWA Research Foundation study (Grayman et al., in press).

**Model description**

**Overview**

The Spill Risk model is an event-driven Monte Carlo simulation model of an early warning system for a river reach. Spill events are randomly generated, and routed through a simplified transport model. Monitors determine the presence of the spill, and notify the treatment plant, which can respond with intake closures, or enhanced treatment. Different response policies are available for action at the treatment plant, depending upon the distance to the spill, and the desire to act before or after confirmation of the spill. The overall metric for system behavior is the population exposure over time to treated water with concentrations above maximum contaminant levels (MCL) or other user-specified acceptable levels. The model also allows for specification of values that relate to strengthened institutions, such as probability of self- or public reporting of spills.

Spill Risk is entirely data-driven, allowing for simple analysis of various configurations of monitoring systems and wide applicability. The model operates on desktop computers under Windows 95 and above operating systems, and is implemented as a C++ program that performs the simulation, working in conjunction with a Microsoft Access 97 database for data storage.

**Monte Carlo simulation**

Monte Carlo simulation is a well known technique for analyzing complex physical systems where probabilistic behavior is important. This technique is widely used in modeling probabilistic systems in the water resources area and many scientific fields. Events are represented as probabilistic occurrences (i.e. probability distribution of streamflows, probabilities of spills of different substances, magnitudes and duration, etc.). The relationships of these events (e.g. how a river responds to a spill) are embodied in the model, which is then run many times, with varying inputs based on the probabilities of the events. The results are recorded for each simulation run, and summarized statistically. In this fashion, the interacting probabilities result in statistics for the total system. A wide range of situations can be examined along with alternative designs and operations. The result is both an expected value and a distribution of results.

**River representation**

The model representation of a river is shown schematically in Figure 1. The model handles a single one-dimensional river reach, with seasonal hydrology. Multiple possible dischargers (referred to as spill generators) exist along the reach. Spill generators can be at fixed locations, representing known plant sites, and non-fixed (probabilistic) locations (e.g.
barges or other vessels traveling along the waterway, or spills on adjacent highways). A simplified advective-dispersion transport model routes contaminants from point of origin down the reach.

Monitors are located at known locations, and are characterized by sampling frequency and detection limits for various constituents. Different monitors can sample for different constituents, at different frequencies. One or more treatment plants are located along the reach. Removal rates are specified for each treatment plant, by constituent, under normal and “enhanced” treatment. Treatment plants are also characterized by the length of time that their intakes can be closed (to let a spill pass by).

Modeling framework

The model is typically run from the perspective of a single water treatment plant (WTP) with a river as the source of raw water for the water treatment plant. Various sources of contaminants exist (the spill generators). Spills take place probabilistically at these spill generators. Each spill is handled individually as a spill event. As a simplification, the assumption is made that there are no interacting spill events, due to the relative rarity of large spills.

Thus, the basic system being modeled consists of a one-dimensional reach of river (no major tributaries within the reach), a number of spill generators, a number of monitors located on the reach, and a water treatment plant (WTP) as the most downstream element of the system. River miles are used to designate the location of spill generators, monitors, and the WTP. Flow is seasonal (time varying) and probabilistic within a season, but constant in the reach at any time. Spills occur randomly. When a spill takes place, the entire set of activities and consequences associated with that spill is calculated. The constituent present in the spill is routed through the system. Pollutographs (curves of concentration vs. time at a location) are calculated based on the transport model results (and the flow in the river at that time), at each monitoring location and the WTP location.

The spill may be detected through one of three mechanisms: monitors, self-reporting by the discharger, or by “public” reporting, i.e. from other individuals or agencies that observe the spill on the waterway. If a spill is detected, the treatment plant is notified and may take some appropriate response consistent with the capabilities available at the plant – closing of intakes, enhanced treatment, or some combination of the two. The impact on the population served by the water treatment plant (WTP) is then based on the concentration of the contaminant in the finished water, during the period of exposure. At present, population impact is the only metric used in the model, i.e. there is no measure of environmental degradation.
Technical model description

Each spill is an event taking place at a given time. Once a spill event takes place, it is processed by the model. Basic model assumptions are: a single hydraulic reach (mainstem only, no tributaries), with constant flow in the reach; one-dimensional (laterally and vertically mixed) hydraulics; simple advection-dispersion transport model; seasonal flows (up to maximum of 12 seasons per year); flow taken as a constant for the duration of a given spill event; velocity in the reach is an exponential function of the flow, independent of season; each spill contains a single constituent; there is no seasonality in the occurrence of spills; and there are no interactions between spills.

Locations of spill generators, the water treatment plant, and monitors are defined using a river mile system along the reach. A set of possible constituents with known decay rate and MCL must be defined. Reporting of spills can be from monitors, discharger self-reports, or public reports. Once a spill is reported, a series of actions takes place on the part of the WTP or other organizations to confirm the spill, characterize the spill, and take remedial action. Confirmation of the spill is taken to be completely deterministic. That is, if a spill is reported, once confirmation has been carried out, then the WTP has complete information as to the pollutograph of the spill at the raw water intake.

Spill generators and events

Spill events are associated with a spill generator. Each spill generator is named, and may be either at a point location (river mile), or occur probabilistically over a range of river miles (uniform probability of the spill occurring at any location within the range). Associated with each spill generator is a probability of self-reporting of the spill, and a maximum time after the spill until the spill is reported. These two data items are used to determine whether a self-report notification takes place, and, if so, at what specific time after the spill.

Each spill event is defined by the following: spill generator generating the spill event; constituent discharged; description of event; expected number of events per year; minimum/most likely/maximum magnitude of spill (pounds); minimum/maximum duration of spill (hours); probability of public or agency (outside) report of spill event; and minimum/maximum delay on public/agency report (hours).

Given the wide range of possible magnitudes of spills at a location, it is not possible to represent all of the spill events at a spill generator by a single distribution. To properly represent the nature of spills at spill generators, one or more probability distributions are defined for each contaminant. Thus, one distribution can describe common events, another distribution can describe large events, and yet another distribution can describe catastrophic, worst-case events. Generally common events are relatively low magnitude and medium frequency. Uncommon events are those of much higher magnitude and much lower frequency. The magnitude of the spill is defined probabilistically as a triangular distribution, giving the minimum, most likely, and maximum magnitudes for that particular spill event. The duration is defined by two parameters, the minimum duration and the maximum duration, and a uniform distribution between these two values is assumed.

Detection and notification

Notification of a spill comes about through one of three possible detection methods: self-reporting, monitoring, and/or public report. Behavior of monitors is described below. The probability of self-report is a function of the spill generator, not the particular constituent or spill, while the probability of public or agency report is a function of the spill event. This is designed to reflect the fact that certain types of spills (large magnitude, or highly visible spills such as oil spills) are more likely to be reported by the public, and that certain spill generators (e.g. industries) are likely to be more vigilant and honest than others.
Detection by monitors

Monitors exist at defined locations, with defined sampling frequencies and minimum detection limits (MDL) for each constituent. All monitors are treated as scheduled, i.e. sampled on a fixed, known frequency. The minimum sampling frequency is 1 hour, i.e. 1 sample per hour. Continuous monitors are treated as scheduled monitors sampling at a high rate (e.g. 1/hour). Monitoring is assumed to be carried out 24 hours a day, 365 days a year. There is no lag time for analysis; i.e., the monitor is assumed to be able to instantaneously detect a constituent, and report that detection. Communication time is, however, added to the reporting time, as described below, and can be used to account for the analysis time.

A monitor reports the presence of a constituent if the actual concentration at the monitor is greater than the MDL. Note that monitors simply report presence or absence, i.e. the magnitude is not determined. This is due to a simplifying assumption that, following notification through a monitor detect, the spill is eventually completely characterized by other means (e.g. detailed intensive sampling). The location associated with the report is the monitor location. False positive and negative reports by monitors are not considered.

This results in three possible pairs of notification information – detection time and river mile for each of the three detection methods. The earliest of these is taken as the notification mechanism. At present, confirming information present in the other notifications is not used within the model, but the presence of two or more notifications, or multiple notifications from monitors, could be used as additional characterization/confirmation information in a subsequent version of the model.

Associated with any notification method is a communication lag time, and a time delay associated with confirmation. It is assumed that, for each method of detection, the time and location associated with the detection are known. At present, detection accurately identifies only the presence or absence of a spill, not the magnitude of the spill. It is only after confirmation that the nature of the spill is known.

Spill communication, confirmation and characterization

The model uses the concept of characterization to reflect the fact that, in the real world, after notification of a potential spill, a variety of efforts are mobilized to confirm the report and obtain complete characterization of the spill. These efforts may involve checking with other monitoring locations on the river, sending out special sampling teams, taking repeat samples, etc. The characterization period is represented simply as a fixed time in hours. As well, the model makes use of a communication time, which is the time delay between detection by any of the methods, and communication of that detection to the water treatment plant. This lag time is used to represent delays in transmitting information, time for internal verification and approvals for self-reporting, or other factors that would cause delays in informing the treatment plant. For non-instantaneous monitors, it also includes the analysis time.

A current simplification within the model is the assumption that, following the characterization period, the water utility has an accurate knowledge of the pollutograph at the intake. Although the full pollutograph is available, the model makes use of five different values derived from the pollutograph: time of arrival of leading edge of spill at treatment plant; time of arrival of trailing edge of spill at treatment plant; value of peak concentration in raw water at treatment plant; time of arrival of point at which post-normal treatment concentration exceeds MCL (on rising limb of pollutograph); and time of arrival of point at which post-normal treatment concentration falls below MCL (on falling limb of pollutograph). These values are used in calculations internal to the model to determine treatment plant response.
Treatment plant response

When a spill takes place, the treatment plant operators may have various levels of knowledge related to the spill: a) unaware of the spill; b) informed of a spill at some distance such that there is sufficient time for full characterization; and c) informed of a spill nearby, with no time to characterize the spill before taking action. In addition, the model can calculate responses based on assumed perfect knowledge of the spill, in order to provide a baseline against which to test other approaches.

The operational decisions available at the treatment plant are taken to be setting the time of intake closure and re-opening and/or setting the time of start and end of enhanced treatment. The maximum allowable closure time is a property of the treatment plant, and is specified in the input data (in hours). For each run of the simulation, one of five possible action policies, and one of four possible response actions, must be specified. The five action policies are: 1) DO NOTHING: take no action in event of a spill notification; 2) ACT ON DETECT: act immediately upon notification; 3) ACT ON DETECT IF WITHIN CRITICAL DISTANCE: act immediately upon notification if the presumed spill location is within some number of miles of the intake (user input); 4) ACT AFTER CHARACTERIZATION: act only after the spill has been confirmed and characterized; and 5) PERFECT KNOWLEDGE: act with perfect knowledge of the spill situation. The two “act on detect” policies are used to initiate a response immediately – the termination of the response (end intake closure or enhanced treatment) is assumed to take place only after characterization or after the maximum intake closure time is reached.

The four possible response actions are: 1) do nothing; 2) close the intake for the required length of time, or until the maximum closure time has been reached; 3) start enhanced treatment for the required length of time; and 4) first close the intake, then, if required, use enhanced treatment. The four actions determine which of the operational variables (times to start/end closure and enhanced treatment) can be set, and the policies determine when they are set. User input to the model selects one of the actions and one of the policies for each simulation run. Once a notification time is determined, the response is calculated according to the input policies and actions, as described above. This results in setting of the times for the start and end of intake closure and enhanced treatment. Based on these values, and the enhanced treatment removals, the population impact is calculated, as described below, by transforming the pollutograph at the intake by the factors associated with the closure and enhanced removals. This results in a post-treatment curve of concentration over time, from which the impact is calculated. Thus, for each spill, the action impact and no action impact are calculated, with the difference between the two, the impact reduction, as the measure of effectiveness of the overall early warning system.

Population impact

Population exposure to treated water with concentrations above the maximum contaminant level (MCL) is taken as the metric for overall system behavior and is calculated by the following equation where \( C_t \) is the concentration of the contaminant at the treatment plant at time \( t \):

\[
\text{Population Exposure} = \text{population} \times \sum_{\text{Leading Edge Time}}^{\text{Trailing Edge Time}} \Delta t \cdot (C_t - \text{MCL}) / \text{MCL}
\]

The process is shown graphically in Figure 2, in which three curves are shown: raw water concentration over time at the treatment plant, normal treatment finished water over time at the treatment plant, and finished water with intake closure followed by enhanced treatment. The “no action” impact is the area between the horizontal MCL line and the normal treat-
ment curve, while the action impact is the area between the early warning curve and the MCL.

Simulation results

The calculations are performed for each year of the simulation, and for the user-defined number of iterations. Statistics are generated for each iteration and for the scenario as a whole. Results are stored in a database for comparison between scenarios, and text format files are generated containing information on each iteration, and optional detailed spill data.

A sample of results data for a 50 iteration run is shown below (Figure 3). In this example, on average, there were 100.4 spills per iteration, with a maximum of 116 spills and a minimum of 88 spills. The average impact reduction was 113,700,559 person-days, but with a comparably high standard deviation of 123,555,547 person-days. This is typical of simulations of rare events, in which many of the iterations have few spills.

Model parameterization

The Spill Risk model includes a large number of parameters. Some of the parameters are readily derivable from historical records. Other parameters may require more subjective decisions or consensus among a panel of local experts. For those parameters that have the greatest inherent uncertainty, sensitivity analyses can be used by varying the parameters over a reasonable range.

Generally the physical parameters (flow distributions, velocity, monitor minimum detection limits) can be derived from analysis of historical measurements and instrument data. The parameters relating velocity to streamflow can be based on hydrologic models or field measurements. Parameters characterizing the distribution of spill events are generally

<table>
<thead>
<tr>
<th>Example Simulation Output</th>
</tr>
</thead>
<tbody>
<tr>
<td># Of Spills</td>
</tr>
<tr>
<td>NA Impact</td>
</tr>
<tr>
<td>Impact</td>
</tr>
<tr>
<td>Impact Reduction</td>
</tr>
<tr>
<td>Total Closure Time</td>
</tr>
<tr>
<td>Enhanced Treat. Time</td>
</tr>
</tbody>
</table>

Figure 2 Graphical representation of exposure calculations

Figure 3 Sample results of 50 iteration run
more difficult to estimate. Potential sources of historical spill information may include records available from the water utilities and local, state and federal environmental agencies. Van Baardwijk (1994) describes a risk analysis procedure developed in The Netherlands as part of the discharge permitting process that uses a structured approach for identifying the size and frequency of discharges. A technique that is frequently employed in parameterizing risk models is the use of a panel of local experts familiar with various aspects of spill situations for the river of interest. Discussion amongst the experts can lead to a consensus on appropriate ranges of values that may be used in the Spill Risk model.

Model application
The model has been applied to a 322 km (200 mile) industrialized stretch of the Ohio River, and used to develop general guidance on design of early warning systems. General guidance was developed based on abstract hypothetical data, while the application of the model to the Ohio River served to provide a large-scale demonstration of the application of the Spill Risk model, and to demonstrate issues associated with parameterizing the model.

In the Ohio River application, the model was used to study the effectiveness of the present early warning system on three water intakes. The Ohio River Valley Water Sanitation Commission (ORSANCO) operates an early warning system composed of 15 gas chromatographs on the Ohio River mainstem and tributaries. ORSANCO also provides coordination, integrates self reports and public reports of spills, maintains a spill simulation model of the river and provides communications with governmental agencies and water utilities. The effectiveness of the early warning system was found to vary significantly among the three intakes based on the following factors: the proximity of the intake to upstream spill generators; the location of monitors relative to the intakes; the sampling frequency and minimum detection limit for the instruments; the length of time that the water intakes could be closed; the type of normal treatment routinely applied at the water treatment plant. An effectiveness index, defined as [Impact reduction / Impacts associated with no actions] was used to rate the effectiveness of the early warning systems. Values for impact reduction and impacts associated with no actions were calculated by the model for each intake based on a simulation of 3000 iterations of one-year duration each. Based on the application of the Spill Risk model, the effectiveness was found to vary between 0.94 and 0.51 for the three intakes.

Findings and summary
Many of the factors associated with an early warning system are highly probabilistic in nature. Design or assessment of an early warning system must properly account for the probabilistic nature of these factors. A risk-based approach has been developed and applied as a demonstration of the incorporation of probabilistic factors in analyzing early warning systems. The Spill Risk model uses Monte Carlo simulation as a first-cut examination of spills and early warning systems in a riverine environment. The model requires input data, some of which can be developed analytically, but much of which, notably estimates of spill frequencies, magnitudes, and durations, will require expert judgment. The model should be useful in the initial examination of alternative early warning system strategies. It is also intended as a demonstration of the usefulness and viability of risk based analytical techniques to a wider range of drinking water issues that have a probabilistic component.

Acknowledgements
The work described in this paper was sponsored by the AWWA Research Foundation as part of a project entitled “Design of Early Warning and Predictive Source-Water Monitoring Systems”. The support and assistance of the Foundation, our project Manager,
Albert Ilges, our co-Principal Investigator, Rolf Deininger, the Project Advisory Committee and the participating agencies and utilities is gratefully acknowledged. We acknowledge the invaluable assistance of the Ohio River Valley Water Sanitation Commission (ORSANCO) and the Cincinnati Water Works in providing information used in the application of the model to the Ohio River.

References