Stochastic Models for Analysis and Control of Air Pollution in a Manufacturing System

Jan, 1, 2005

Technical Report SOPTL-05-01

Missouri University of Science and Technology

Abhijit Gosavi gosavia@mst.edu

Acknowledgement. This research was supported in part by a grant DMI: 0114007 to the author from the National Science Foundation.

Abstract

This report presents stochastic models that can be used for controlling pollution in a manufacturing system. The models are developed to capture the harmful effects of exposure of pollutants to workers in a manufacturing environment.

1. Introduction

Pollution of air resulting from toxic wastes emitted by large manufacturing plants and power-generation plants is a much-studied topic. The issue of health hazards arising from water pollution (see Wagner, Shamir, and Marks, 1994), storage and transportation of hazardous materials, and genetically-engineered food products has been a topic of intense debate and considerable research for the last few years. A great deal of attention has been paid to *reprocessing* of wastes before they are discharged into the ground from industries.

An equally important topic is the pollution of air *inside* a manufacturing plant which typically also houses the source of pollution. This subject deserves close research attention because it is a matter of public health and also one of concern to the managers of almost every manufacturing plant. The persons who work in such plants are exposed to toxics and can suffer long-term ailments as a result of the exposure.

Air pollution has been identified as the root cause of a very large number of chronic diseases. Laws have been passed in almost all countries in the world in recent times to minimize the deleterious effects of air pollution. The evidence gathered to measure the effects of air pollution has played a significant role in the passage of such laws. These laws have had a significant impact on reducing the occurrence of related diseases and on improving the average health index of individuals in that country. Not surprisingly, operations-research models have provided important tools in measuring *and* controlling air pollution. An example is the formula of Grandell (1985) that is used commonly by meteorologists.

In this paper, we provide an operations-research perspective to address an important problem of air pollution in a *static* mass of air that occupies the *inside* of manufacturing plants. The problem of air pollution inside manufacturing plants is of serious public concern in the rapidly-growing semi-conductor industry and the plastics industry. In the semi-

conductor industry, unlike traditional manufacturing, most of the machining processes are of a chemical nature. The US Environmental Protection Agency (EPA) has identified at least 30 air toxics emitted in semi-conductor manufacturing. The semi-conductor industry produces "chips" that are required for electronic devices in computers, televisions, radios, CD players, etc. The production processes in semi-conductor manufacturing emit a large number of noxious pollutants into the air immediately surrounding the machines. Important examples, which constitute 90% of the emissions in semi-conductor manufacturing, are: hydrochloric acid, hydrofluoric acid, glycol ethers, methanol, and xylene (see EPA website, 2003 and World Bank Website, 1998). Exposure to these particles over extended period of times has serious implications for the health of the exposed individual. The plastic industry is also critical because a large number of pollutants are emitted into the air during the manufacture of reinforced-plastic composites. Some of the chief culprits in the processes are: open molding, pultrusion, centrifugal casting, and continuous lamination (see EPA website, 2003). Some examples of pollutants are: styrene and volatile organic compounds (VOCs).

A variety of pollutant-absorbing devices are now available in the market. Their use is imperative in order to minimize the effect of pollution. Sensors can be used to continuously monitor the quality of air and detect the presence of pollutants. If and when pollution levels reach unacceptable limits, the related production processes should be stopped immediately. The generation of pollutants is usually directly linked to production rates. In other words, high production rates, which are dictated by economic considerations, *drive us to increased pollution risk!* This is an important fact that deserves attention from policy makers. If the pollution risk is quantified in economic terms, it is possible to develop production schedules that will not only be economically feasible but will also meet pollution standards. A goal of this paper is to present models directed towards decision-making in the context of pollution prevention and public policy.

It needs to be emphasized that the *cost* of pollution — in any setting, but in particular in the manufacturing setting — cannot be ignored. Clearly, the society pays a heavy price for the harmful effects of pollution. The managers of both public and private firms pay medical compensations for people who are affected, although the effects are usually clear only years after the damage ensues. Hence today this is especially true of young industries, e.g., semi-conductor manufacturing.

The literature on open-air pollution primarily addresses an air system that moves over large areas with a non-zero velocity. Important work in this area is from Gibbs and Slin (1973), Grandell (1985), and Stein (1984). An important ingredient of these models — that we also make use of — is the notion of a source and a sink (scavenger) of pollutants. The idea of a source and a sink is commonly found in almost all theoretical / physical models of air pollution including indoor air pollution (see Guo, 1993). Indoor air pollution has also been of intense research interest recently. There is a significant body of literature that studies the effects of pollution resulting from polluted air from the outside (e.g., a nearby factory) seeping into a home (see Furtaw et al., 1996 and McBride et al., 1999) and the effects of pollution indoors from domestic sources, e.g., carpets and surface coatings (see Tichenor, Guo, and Sparks, 1993). See Guo (2002a) and Guo (2002b) for a nice review of indoor air quality (IAQ) models. Most of these models assume deterministic rates of arrival for the polluting particles.

A fundamental model used in IAQ modeling is:

$$V\frac{d\tilde{\psi}}{dt} = \tilde{\lambda}_{source} - \tilde{\lambda}_{sink}, \tag{1}$$

where t denotes time, V denotes the volume of the air inside, $\tilde{\psi}$ denotes the concentration of particles (mass per unit volume), $\tilde{\lambda}_{source}$ denotes the rate of arrival (mass per unit time) of particles from the source and $\tilde{\lambda}_{sink}$ denotes the rate of absorption of pollutants by the sink. In most models (1) also has a term for penetration of ambient air. Our work in this paper is based on adapting this simple model for the manufacturing environment. Many of the models for source rates in the literature assume that the rate of arrival is dependent on mass of generating material remaining. Such models typically assume that the rate of arrival decays with time. In the context of the manufacturing system, e.g., semi-conductor manufacturing and manufacture of reinforced-plastics, this assumption is not reasonable because the source of pollution keeps getting replenished. Also, the rate of arrival is likely to be a random variable. Hence, in our models $\tilde{\lambda}$ in (1) will be treated as a random variable. Also, we will deal with a suitably-scaled parameter:

$$\lambda = \tilde{\lambda}/m_{avg}, \tag{2}$$

where λ denotes the rate of arrival of each particle, and m_{avg} denotes the mean mass of each particle. With this scaling, we will be able to deal with the arrival rate of individual particles — a parameter that is more commonly used in operations research. It must be noted that this will in no way affect the usefulness of any analysis that we will perform because via (2) one obtain all the quantities calculated on the scale of mass (in kilograms).

In our models, we will assume that the mass of air inside a manufacturing system is static in the following sense: the loss of air mass to the outside is negligible. This implies that particles that are emitted into the system do not disappear to the outside *fast enough.* If particles disappear (get dissipated into the air such that their concentration is below the danger level) *fast enough*, there is no reason to worry about pollution inside the manufacturing plant. The reality, however, is that particles rarely disappear fast enough — not even in open air systems. This has given birth to the science of studying air pollution. And particularly inside manufacturing plants, the danger is more acute (see World Bank Website, 1998). The assumption although realistic is critical for the validity of our modeling process. This paper addresses issues related to developing models that use the cost of pollution, analyze the net gains obtained from controlling pollution, and solve the problem of designing a safe environment inside a manufacturing plant.

Contributions of this paper: Our models differ in a number of important ways from those in the literature. Firstly, we employ the *cost* of pollution in our models. Secondly, we study the problem in the context of *manufacturing problems* and the standpoint of the production planner who has to develop production schedules and meet due-dates. Thirdly, most of the open-air pollution models are based on the Poisson arrival assumption, while indoor-air pollution models are based on deterministic arrivals. Our models assume arbitrary arrival patterns. Finally, we also show how simple (existing) results from queuing theory, renewal theory, and stochastic approximation can be exploited in analyzing pollution-control problems. A primary contribution of this paper is to show how the highly-developed theory

of stochastic processes can be exploited to generate useful models in the area of indoor-air pollution control. We also provide reinforcing numerical evidence to demonstrate the effectiveness of our models. An interesting conclusion from our study is that optimization techniques can play a useful role in solving several important problems in air-pollution control.

Section 2 analyzes the effect of cleaning-up operations on the production schedules. Section 3 develops an economic model for determining the optimal stopping time for cleaning up. Section 4 discusses interaction costs. Section 5 develops a model for designing the number of pollutant absorbers in a continuously-cleaned system. Section 6 concludes this paper.

2. Analysis of the effect of stoppages for cleaning

It might appear on first thought that shutting down machines for a significant period of time may be economically infeasible. The cost of shutting down machines that pollute should be weighed against the long-term cost of pollution to the society. Our analysis in this section is geared toward measuring the downtime that results from such stoppages. The analysis yields interesting insights on the cost of downtime due to lost production vis-a-vis the long-term cost of not adhering to pollution standards and exposing people to hazardous substances. In the rest of the paper, whenever machines resume work after a cleaning-up operation, a new *run* will be assumed to have started.

The length of each cleaning-up operation *and* its frequency will clearly impact the effective production rate of the plant. Both factors must be taken into account in performing this analysis. Also, we will assume that machines are shut down when biologically-determined thresholds are reached for any of the polluting particles. The pollutant particles arrive into the air randomly. Our analysis is performed under two assumptions, A and B, which are related to the rate of arrival.

Assumption A: Each particle has an exponentially-distributed inter-arrival time.

The following result will yield a mechanism for measuring the mean downtime of the system.

Theorem 1 *Let λ_i denote the mean rate of arrival for particle of type i, $N_i(t)$ denote the number of arrivals of type i by time t, K_i^* denote the biologically-determined threshold for particle of type i, and τ denote the time-interval after which machines are stopped. Then, under Assumption A, the mean stopping time $\bar{\tau}$ is:*

$$\bar{\tau} = \int_0^\infty t f_\tau(t) dt, \tag{3}$$

where the density function of the stopping time is:

$$f_\tau(t) \equiv \sum_{j=1}^{k} \prod_{i=1}^{k} R(i,j,t), \tag{4}$$

in which

$$R(i,j,t) = \begin{cases} g^i_{K_i^*}(t) & \text{if } i = j; \\ Pr[N_i(t) < K_i^*] & \text{otherwise.} \end{cases}$$

where

$$Pr[N_i(t) < K_i^*] = \sum_{j=0}^{K_i^*-1} e^{-\lambda_i t} \frac{(\lambda_i t)^j}{j!}, \tag{5}$$

and

$$g_n^i(t) = \lambda_i e^{-\lambda_i t} \frac{(\lambda_i t)^{n-1}}{(n-1)!}. \tag{6}$$

A direct offshoot of this result is a formula for finding the mean fraction of time the machines are down due to cleaning-up stoppages. Then, if $\bar{c}$ denotes the average time to clean the air, the mean fraction of time the system is down for cleaning can be expressed as:

$$\gamma = \frac{\bar{c}}{\bar{\tau} + \bar{c}}. \tag{7}$$

Proof The proof follows from the fact that the stopping time is dictated by the event of any one particle reaching its threshold. The expression in (4) is based on this argument. In (4), $g_n^i(t)$, which denotes the density function of the arrival time of the nth arrival of type i, is gamma-distributed with parameters n and λ. The latter follows from an elementary result (see Ross, 2003). The validity of (5) also follows from an elementary result. ∎

Note that K_i^* (for any i) can be determined from the threshold concentration, ψ_i, which is defined as the number of units of particle of type i in unit volume of air. Typically, pollution-control sensors and equipment measure the concentration of particles in the air and not the *absolute number* of particles. The formula for K_i^* is simply:

$$K_i^* = \psi_i V,$$

where V denotes the volume of air in the manufacturing system.

The downtime will naturally reduce the production rate of the system. And this is bound to result in expenditure for companies with expensive resources that have to be utilized almost on a 24-hour basis. However, this cost, in all probability, will be much smaller than the sum of the costs of damage to human health and the compensations paid by companies. To illustrate the use of (4) to evaluate (3), we next provide a simple numerical example.

Numerical Example: Consider the simple case with $k = 2$, $K_1^* = K_2^* = 2$, and $\lambda_1 = \lambda_2 = 0.01$. Then, (4) will be

$$f_\tau(t) = g_2^1(t) Pr[N_2(t) < 2] + g_2^2(t) Pr[N_1(t) < 2].$$

Then using (5) and (6), we can express ρ, after some algebraic manipulations, as:

$$\rho = \frac{\int_0^\infty t^2 e^{-t(\lambda_1+\lambda_2)}(\lambda_1^2 + \lambda_1^2\lambda_2 t + \lambda_2^2 + \lambda_2^2\lambda_1 t)dt}{\bar{c} + \int_0^\infty t^2 e^{-t(\lambda_1+\lambda_2)}(\lambda_1^2 + \lambda_1^2\lambda_2 t + \lambda_2^2 + \lambda_2^2\lambda_1 t)dt}$$

For larger values of k and K_i^*, deriving these expressions will get somewhat tedious, but the idea is conceptually simple. We illustrate this point using $k = 3$ and $K_1^* = K_2^* = K_3^* = 3$.

In this case, (4) will be given by:

$$\begin{aligned}
f_\tau(t) &= g^1_{K^*_1}(t)Pr[N_2(t) < K^*_2]Pr[N_3(t) < K^*_3] + g^2_{K^*_2}(t)Pr[N_1(t) < K^*_1]Pr[N_3(t) < K^*_3] \\
&\quad + g^3_{K^*_3}(t)Pr[N_1(t) < K^*_1]Pr[N_2(t) < K^*_2] \\
&= g^1_3(t)\left[e^{-\lambda_2 t} + e^{-\lambda_2 t}\lambda_2 t + e^{-\lambda_2 t}\frac{(\lambda_2 t)^2}{2!}\right]\left[e^{-\lambda_3 t} + e^{-\lambda_3 t}\lambda_3 t + e^{-\lambda_3 t}\frac{(\lambda_3 t)^2}{2!}\right] \\
&+ g^2_3(t)\left[e^{-\lambda_1 t} + e^{-\lambda_1 t}\lambda_1 t + e^{-\lambda_1 t}\frac{(\lambda_1 t)^2}{2!}\right]\left[e^{-\lambda_3 t} + e^{-\lambda_3 t}\lambda_3 t + e^{-\lambda_3 t}\frac{(\lambda_3 t)^2}{2!}\right] \\
&+ g^3_3(t)\left[e^{-\lambda_1 t} + e^{-\lambda_1 t}\lambda_1 t + e^{-\lambda_1 t}\frac{(\lambda_1 t)^2}{2!}\right]\left[e^{-\lambda_2 t} + e^{-\lambda_2 t}\lambda_2 t + e^{-\lambda_2 t}\frac{(\lambda_2 t)^2}{2!}\right].
\end{aligned}$$

It can become quite difficult to integrate the function derived above to obtain $\bar{\tau}$ as required in (3). Clearly, therefore, a numerically-tractable technique is needed for solving real-world problems. Simulation is one such technique, although it will only provide approximations. An additional power of simulation is that we can relax Assumption A. Consider the next assumption, which is weaker than A.
Assumption B: The inter-arrival time of each particle is independent and identically distributed (i.i.d), whose distribution is arbitrary. The following result provides a mechanism for computing $\bar{\tau}$ approximately.

Theorem 2 *Let $(\Omega, \mathcal{F}, P)$ denote a probability space, and let $\omega^1, \omega^2 \ldots,$ denote i.i.d random variables (samples) generated (simulated) from P. Let $a(i, \omega^l)$ denote the lth sample of the time interval between the end of the last cleaning-up operation and the arrival of the K^*_ith particle of type i in the system. Then, under Assumption B, the mean stopping time, with probability 1, is:*

$$\bar{\tau} = \lim_{N\to\infty} \frac{\sum_{l=1}^{N} \inf_{i=1,2,\ldots,k} a(i, \omega^l)}{N}.$$

Proof The proof of the theorem follows easily from the strong law of large numbers (Durrett, 1995) and from noting that the system is stopped for cleaning when any one particle reaches its threshold. ■ ■

The above implies that for a given value of $\epsilon > 0$, there exists with probability 1, a finite value for N such that

$$\left|\bar{\tau} - \frac{\sum_{l=1}^{N} \inf_{i=1,2,\ldots,k} a(i, \omega^l)}{N}\right| < \epsilon.$$

Hence an approximate value of $\bar{\tau}$ can be derived from the following formula using a large value for N.

$$\bar{\tau} \approx \frac{\sum_{l=1}^{N} \inf_{i=1,2,\ldots,k} a(i, \omega^l)}{N}. \tag{8}$$

Table 1 illustrates the use of the results of this section to estimate the value of $\bar{\tau}$ and γ (see (7)). In all the experiments performed, $\bar{c}$ was set to 2 time units. The purpose of the numerical results is primarily to show that these results can be used easily by managers to

Experiment	λ_1	λ_2	λ_3	λ_4	λ_5	K_1^*	K_2^*	K_3^*	K_4^*	K_5^*	$\bar{\tau}$	γ %
1	0.1	0.2	0.3	0.1	0.2	10	25	35	15	25	86.9	2.2
2	0.9	0.2	0.7	0.4	0.1	10	25	35	15	25	11.16	15.19
3	0.1	0.5	0.9	0.6	0.4	15	12	30	11	12	16.66	10.71
4	0.1	0.5	0.15	0.6	0.4	15	12	30	11	15	16.9	10.51
5	0.1	0.15	0.35	0.2	0.25	11	26	38	20	21	75.46	2.5
6	0.01	0.02	0.32	0.1	0.15	28	32	40	21	18	109.31	1.79
7	0.15	0.07	0.03	0.7	0.6	22	31	62	30	17	28.01	6.6
8	0.3	0.05	0.01	0.1	0.2	11	17	45	51	39	36.79	5.2

Table 1: Results from experiments for a system with five pollutants. The values for γ (in percentage) and $\bar{\tau}$ for a variety of system settings (thresholds and arrival rates) are derived. In all experiments, $\bar{c} = 2$.

predict the economic implications of stoppages for cleaning the air. The arrival rates are Poisson, but via (8), tackling any other arrival rate is trivial.

We next develop an economic model for finding the optimal stopping time.

3. An economic model for optimal stopping time

The length of the stopping time clearly has implications on the production rate and the safety. In this section, we propose a simple model for determining the optimal stopping time. A cost model based on the stochastics of the system can be developed and then optimized. One has to be careful in the interpretation of the results obtained from optimization. The reason is as follows: The policy derived from economic considerations should not violate the policy dictated from environmental and biological considerations.

Before developing the cost model, we need some additional assumptions. They are as follows:

1. The cost of holding a pollutant particle in the air for unit time can be estimated.

2. The cost of cleaning-up and the loss of production due to stoppage can be estimated.

It is necessary to justify these assumptions. The cost of holding one particle of pollutant in the air for unit time can be measured from biological and economic considerations. If a person's health is jeopardized due to prolonged exposure, the company will ultimately have to compensate the individual. The cost of cleaning will depend on a number of factors such as the cost of the equipment needed for cleaning up and the cost of the resulting downtime.

We will use Assumption B for the inter-arrival time of particles. We will make an additional assumption, next.

Assumption C: There is no cost of interaction of two or more types of particles in the air.

Assumption C implies that additional costs that can arise out of the presence of multiple types of pollutants and their interactions are non-existent. The cost model is derived in the following result.

Theorem 3 *Let τ denote the stopping time (length of a run), λ_i the arrival rate of particle of type i, and S the stopping cost (sum of the cleaning cost and the cost of lost production time). Then, if c_i denotes the cost of holding one particle of type i for unit time, the expected long-run average cost per unit time, under Assumptions B and C, is:*

$$\rho = \frac{S}{\tau} + \sum_{i=1}^{k} \frac{c_i}{\lambda_i} \lceil \tau\lambda_i \rceil \frac{\lceil \tau\lambda_i \rceil - 1}{2\tau}.$$

In the above, $\lceil a \rceil$ denotes the smallest integer greater than or equal to a.

Proof The proof is based on a simple renewal theory argument. Consider a cycle at the end of which the production system is stopped for cleaning. Let $T_i(j)$ denote the time between the jth and the $(j+1)$th arrival of particle of type i and τ the length of the cycle. The expected number of arrivals of particle of type i is then $\lambda_i\tau$. Then, if C denotes the cost in one cycle of length τ,

$$\begin{aligned}
\mathsf{E}[C] &= S + \sum_{i=1}^{k} \mathsf{E}[c_i T_i(1) + 2c_i T_i(2) + \cdots + c_i(\lceil \tau\lambda_i \rceil - 1) T_i(\lceil \tau\lambda_i \rceil - 1)] \\
&= S + \sum_{i=1}^{k} c_i \mathsf{E}[T_i(.)] \left(1 + 2 + 3 + \cdots + (\lceil \tau\lambda_i \rceil - 1)\right) \\
&= S + \sum_{i=1}^{k} c_i \frac{1}{\lambda_i} \left(1 + 2 + 3 + \cdots + (\lceil \tau\lambda_i \rceil - 1)\right) \\
&= S + \sum_{i=1}^{k} \frac{c_i}{\lambda_i} \lceil \tau\lambda_i \rceil \frac{\lceil \tau\lambda_i \rceil - 1}{2}.
\end{aligned}$$

Then the result follows from the fact that the expected long-run cost would be

$$\rho = \frac{\mathsf{E}[C]}{\mathsf{E}[\tau]} = \frac{\mathsf{E}[C]}{\tau}. \quad \blacksquare$$

■

The result yields an approximate formula for computing the optimal value of τ that will minimize ρ. Simple calculus, treating $\lceil \tau\lambda_i \rceil$ to be a continuous variable, yields the optimal value of τ to be:

$$\tau^* = \sqrt{\frac{2S}{\sum_{i=1}^{k}(c_i\lambda_i - 1)}} \text{ because } \frac{d\rho}{d\tau} = 2S/\tau^3 > 0.$$

Numerical example: Let $S = \$7500$, $\lambda_1 = 0.1$, $\lambda_2 = 0.01$, $\lambda_3 = 0.25$, $\lambda_4 = 0.7$, $\lambda_5 = 0.02$, and $c_i = 100\ \forall i$, (unit of time being hours) then:

$$\tau^* = \sqrt{\frac{7500}{100(0.1 + 0.01 + 0.25 + 0.7 + 0.02) - 5}} = 8.53 \text{ hours.}$$

4. Interaction costs

Sometimes a chemical on its own does not produce any harm, but in combination with other chemicals becomes harmful. Our analysis so far has disregarded any costs arising from the simultaneous presence of two or more types of particles in the air. In fact, if chemicals become harmful *only* when they are present simultaneously, the analysis provided in the previous sections will not serve any purpose. In this section, our intent is to develop a model that accommodates these costs.

Let $D(i,j)$ denote the cost of carrying particles of type i and j *simultaneously* in the air for unit time. The expected time during which particles i and j are not present simultaneously is:

$$\begin{aligned} F(\lambda_i,\lambda_j,\tau) &= \int_0^\tau t(Pr[N_i(t)=0]Pr[N_j(t)>0]+Pr[N_i(t)=0]Pr[N_j(t)=0] \\ &\quad +Pr[N_j(t)=0]Pr[N_i(t)>0])dt. \end{aligned} \tag{9}$$

The cost due to simultaneous presence of particles of type i and j is:

$$\sum_{i=1}^{k}\sum_{j=1,i\neq j}^{k} D(i,j)[\tau - F(\lambda_i,\lambda_j,\tau)]. \tag{10}$$

To accommodate interactions costs, we can come with the following analogue of Theorem 3.

Theorem 4 *Assume Assumption C to be false. The average cost per unit time when the run length is τ, under Assumption A, is:*

$$\rho = \frac{S}{\tau} + \sum_{i=1}^{k}\frac{c_i}{\lambda_i}\lceil\tau\lambda_i\rceil\frac{\lceil\tau\lambda_i\rceil-1}{2\tau} + \sum_{i=1}^{k}\sum_{j=1,i\neq j}^{k} D(i,j)[1-\frac{\int_0^\tau t(e^{-\lambda_i t}+e^{\lambda_j t}-e^{-\lambda_i t}e^{-\lambda_j t})dt}{\tau}].$$

Proof The first two terms on the right-hand side of the equation above follow from the proof of Theorem 3. The last term follows from (10) in which (9) is derived from the Poisson-arrival assumption. ∎ ■

Under Assumption B, we have a similar result.

Theorem 5 *Assume Assumption C to be false. Consider a probability space $(\Omega,\mathcal{F},P)$. Using P, i.i.d. samples, $\omega^1,\omega^2,\ldots$, can be generated in the sample space. Let $b(i,j,\tau,\omega^l)$ denote the time for which particles i and j are simultaneously present in the lth sample of a run of length τ. Then with probability 1, the average cost per unit time when the run length is τ, under Assumption B, is:*

$$\rho = \frac{S}{\tau} + \sum_{i=1}^{k}\frac{c_i}{\lambda_i}\lceil\tau\lambda_i\rceil\frac{\lceil\tau\lambda_i\rceil-1}{2\tau} + \sum_{i=1}^{k}\sum_{j=1,i\neq j}^{k}\frac{D(i,j)}{\tau}\lim_{N\to\infty}\frac{\sum_{l=1}^{N} b(i,j,\tau,\omega^l)}{N}. \tag{11}$$

Proof The first two terms in the equation above can be accounted for by Theorem 3. The last term follows from the strong law of large numbers. ■ ■

The results in Theorems 4 and 5 can be used to find the optimal value for τ via a gradient-descent approach. The steps in the iterative algorithm required are described next.
Step 1: Let τ^m denote the value of τ in the mth iteration. Set $m = 0$ and τ^m to some arbitrary value. Set $m_{\max}$ to some large value.
Step 2: Update τ^m according to the following equation.

$$\tau^{m+1} \leftarrow \tau^m - \mu \left. \frac{d\rho(\tau)}{d\tau} \right|_{\tau=\tau^m},$$

where μ denotes a step size (a small number less than 1), $\rho(\tau)$ indicates that ρ is a function of τ, τ^m denotes the value of τ in the mth iteration, and the derivative is obtained with a finite-difference approximation that uses central differences:

$$\left. \frac{d\rho(\tau)}{d\tau} \right|_{\tau=\tau^m} \approx \frac{\rho(\tau^m + h) - \rho(\tau^m - h)}{h}$$

in which h is small number.
Step 3: Increment m by 1. If m is less than $m_{\max}$, return to Step 2. Otherwise, stop.

The final value of τ^m that is returned by the algorithm can be accepted to be an approximation of a local optimum. In practice, the algorithm is run several times starting at different values for τ^0 in order to minimize the chances of getting trapped in a poor local optimum. Note that $\rho(.)$ is evaluated from (11), as has been explained in the context of Equation (8).

Table 2 presents the results of optimization of (11). The values of λ_i used in these optimization experiments are those given in Table 1. Also, $c_i = \$50$ for all i, $D(i, j) = 100$ for all i and j, and $S = 7500$.

Experiment	τ^*	Optimized cost
1	18.00	696.59
2	12.07	1190.08
3	9.14	1301.58
4	13.08	1020.65
5	16.89	762.41
6	22.35	545.82
7	13.9	953.19
8	21.31	578.56

Table 2: Here τ^* denotes the optimal value for the stopping time in hours and the optimized cost is the cost with the optimal stopping time.

5. Number of sinks

Thus far we have assumed that the manufacturing process has to be shut down for the cleaning operations. We now analyze a system in which cleaning is done continuously. We assume that pollutant sinks (absorbers) are placed inside the plant, and that they continuously absorb pollutants. In other words, we do not have a *run* here during which pollutant particles can accumulate. The problem considered here is to design the number of sinks to minimize the cost of pollution.

We first consider Assumption A. We assume that the absorbing rate is a deterministic constant. A sink can absorb only one type of particle. Let the rate of absorption of particle of type i be denoted by μ_i per unit time. Then, the average cost of operating a system — in which (i) the sink for particle of type i costs d_i, (ii) there are a_i sinks for particle i, and (iii) the cost of holding an average inventory of unit particle of type i in the long-run is c_i — is

$$\sum_{i=1}^{k}\left(c_i\frac{(\frac{\lambda_i}{a_i\mu_i})^2}{1-\frac{\lambda_i}{a_i\mu_i}}+d_ia_i\right). \tag{12}$$

The above follows from the well-known Pollatschek-Khinchine formula for average inventory in a queue with exponential inter-arrival times and arbitrary service rates.

Strictly speaking the expression in (12) cannot be differentiated with respect to a_i and the optimization problem is an integer program. Since the optimization space for all practical purposes is not very big (of the order of a few thousand at the most), the optimization can be performed with an exhaustive search.

For the next result, we will use the notation: $\vec{x}=\{x_1,x_2,\ldots,x_k\}$ where x_i will denote the ith element of a vector $\vec{x}$ with k dimensions.

Under Assumption B, the average cost that is to be minimized can be defined as:

$$\rho(\vec{a})=\sum_{i=1}^{k}\left(d_ia_i+h_i\mathsf{E}\left[\limsup_{T\to\infty}\frac{1}{T}\int_0^T I_i(t)dt\right]\right). \tag{13}$$

If an arrival or a departure is viewed as an *event*, we will assume that a *new* "epoch" starts whenever an event occurs. Using the strong law of large numbers, as has been done before in this paper, the above expression can be evaluated in a simulator (generating independent samples of the system) using the following formula with "sufficiently" large values for N_1 and N_2:

$$\rho_{N_1,N_2}(\vec{a})=\sum_{i=1}^{k}\left(d_ia_i+h_i\sum_{l=1}^{N_1}\frac{1}{N_1}\frac{\sum_{e=1}^{N_2}I_i^e(\omega^l)t^e(\omega^l)}{\sum_{e=1}^{N_2}t^e(\omega^l)}\right),$$

where $t^e(\omega^l)$ denotes the duration of the eth epoch of the lth sample, and $I_i^e(\omega^l)$ denotes the number of particles of type i in the eth epoch of the lth sample. Since this is a discrete optimization problem, we note the following interesting property associated with the problem structure (which is a generalization of a related property in Gosavi and Ozkaya, 2003).

Theorem 6 *Let $\mathcal{S}_i$ denote the finite set that denotes all possible values of a_i. Consider the following definitions.*

$$\min_{a_i \in \mathcal{S}_i; \forall i} \rho(\vec{a}) \equiv \rho(\vec{y^*}) \text{ and } \min_{a_i \in \mathcal{S}_i; \forall i} \rho_{N_1,N_2}(\vec{a}) \equiv \rho(\vec{y}_{N_1,N_2}).$$

Then, with probability 1, there exist finite values for N_1 and N_2 such that,

$$\vec{y^*} = \vec{y}_{N_1,N_2}.$$

The result assures us that the global optimum with the actual cost function is identical to that found with the simulated-based evaluation of the cost function with sufficiently large values of N_1 and N_2.

Proof Since $\mathcal{S}_i$ is finite, the set of solutions associated with minimizing (13) is also finite. Rank this finite number of solutions in the ascending order of the value of the average cost such that solutions with the same objective function value have the same rank. Let ρ^r denote the actual (exact) average cost (as defined in expression (13)) of the solution that was assigned the rth rank. Let R denote the number of ranks.

The ranking will be such that: $\rho^1 < \rho^2 < \cdots < \rho^R$. Then, define δ as follows:

$$\delta = \frac{\rho^i - \rho^1}{M} \text{ where } M \geq 1.$$

Also, $\rho^i_{N_1,N_2}$ will denote the simulation-based average cost of the solution that has the ith rank according to the scheme described for the actual average cost. Let us define the simulation error in the average cost as follows:

$$e^i_{N_1,N_2} = \rho^i - \rho^i_{N_1,N_2},$$

for $i = 1, 2, \cdots, R$. Then, from the strong law of large numbers, with probability 1, for a given value of $\epsilon > 0$, there exist an N_1 and an N_2 such that for any $i = 1, 2, \ldots, R$,

$$|e^i_{N_1,N_2}| < \epsilon/2. \tag{14}$$

Then, it follows that for any $i \geq 2$,

$$\begin{aligned} \rho^i - \rho^1 &= \rho^i_{N_1,N_2} - \rho^1_{N_1,N_2} + e^i_{N_1,N_2} - e^1_{N_1,N_2} \\ &\leq \rho^i_{N_1,N_2} - \rho^1_{N_1,N_2} + |e^i_{N_1,N_2} + e^1_{N_1,N_2}| \\ &\leq \rho^i_{N_1,N_2} - \rho^1_{N_1,N_2} + |e^i_{N_1,N_2}| + |e^1_{N_1,N_2}| \\ &\leq \rho^i_{N_1,N_2} - \rho^1_{N_1,N_2} + \epsilon \end{aligned} \tag{15}$$

The last inequality follows from (14). Then, by selecting a suitable value for M, we can produce a δ such that $\epsilon = \delta$. Then, we have that for any $i \geq 2$:

$$\rho^i - \rho^1 \leq \rho^i_{N_1,N_2} - \rho^1_{N_1,N_2} + \delta \leq \rho^i_{N_1,N_2} - \rho^1_{N_1,N_2} + \frac{\rho^i - \rho^1}{M}.$$

The above implies that for every $i \geq 2$,

$$\rho^i_{N_1,N_2} - \rho^1_{N_1,N_2} \geq (1 - \frac{1}{M})(\rho^i - \rho^1) \geq 0.$$

This means that with N_1 and N_2 sufficiently large, $\rho^1_{N_1,N_2}$ is a minimum cost from which the result follows. ∎ ■

Table 3 shows the results of optimization performed for designing the number of sinks in a continuously-cleaned system. Parameters used in these sample experiments are: $c_i = 500$, for all i, and $\mu = 0.3$.

Experiment	λ_1	λ_2	λ_3	λ_4	λ_5	$(a_1^*, a_2^*, a_3^*, a_4^*, a_5^*)$
1	0.1	0.11	0.18	0.20	0.12	(2,3,4,4,3)
2	0.12	0.28	0.11	0.23	0.5	(3,5,3,4,7)
3	0.36	0.12	0.13	0.21	0.29	(6,3,2,4,5)
4	1.0	0.13	0.1	0.36	0.47	(12,3,2,6,7)
5	0.8	1.0	0.11	0.19	0.27	(10,12,3,4,5)

Table 3: a_i^* denotes the optimal value for a_i

.

6. Conclusions

Although air pollution is a much-studied topic, the implications of air pollution control on *production control* have not been studied in the literature, i.e., models for the resultant downtime, the economic stopping times, and the design of an appropriate number of sinks. This can perhaps be attributed to the fact that generally metal-based (traditional) manufacturing does not cause pollution in the air immediately surrounding the machines. It is the new generation of industries, e.g., semi-conductor and reinforced plastics, that causes extensive pollution. Hence it is necessary to develop mathematical models that study the inter-relationship between pollution-control measures and their impact on the production schedules. Also important is the problem of designing the number of sinks to absorb pollutants. This paper, to the best of our knowledge, is the first attempt to develop a model for examining these issues. The theory of stochastic approximation, queuing, and renewal processes is extensively developed, and the paper provides a mechanism to exploit well-known results from this field to the important problem of air pollution inside manufacturing systems.

Some possible extensions of this work are: (i) the elimination of constraining assumptions e.g., considering a non-stationary mass of air within the production system so as to accommodate the transient effects of pollution, and (ii) the use of time-varying rates of arrival of particles. It is expected that pollution control in production systems (see EPA website, 2003 and World Bank Website, 1998) will be an important issue in production systems design in the coming years given the polluting nature of the new generation of manufacturing industries.

References

M. Bazaara, H.D. Sherali, and S. Shetty (1992). Non-linear Programming, Second edition, Wiley.

J.F. Bonnans and A. Shapiro (2000). *Perturbation Analysis of Optimization Problems.* Springer Verlag. New York, NY.

R. Durrett (1995). *Probability: Theory and Examples*, Second Edition, Duxbury Press.

Environmental Protection Agency (2003). URL: http://www.epa.gov/ttn/oarpg

E.J. Furtaw, M.D. Pandian, D.R. Nelson, and J. Behar (1996). Modeling indoor air concentrations near emission sources in imperfectly mixed rooms. *Journal of Air Waste Management Association*, 4, pp 861-868.

A.G. Gibbs and W.G.N. Slinn (1973). Fluctuations in trace gas concentrations in the troposphere. *Journal of Geophysics Research*, 78, pp. 574-576.

A. Gosavi (2003). *Simulation-based optimization: Parametric optimization techniques and reinforcement learning*, Kluwer Academic Press, Boston.

A. Gosavi and E. Ozkaya (2003). Network Revenue Management Using Simulation-based Optimization. *Under Review.*

J. Grandell (1985). *Stochastic Models of Air Pollutant Concentration (Lecture Notes in Statistics).* Springer-Verlag, Berlin, Germany.

Z. Guo (1993). On validation of source and sink models: problems and possible solutions. *Modeling of Indoor Air Quality and Exposure.* American Society for Testing and Raw Materials, Philadelphia, PA, pp. 131-144.

Z. Guo (2002a). Review of indoor emission source models. Part1. Overview. *Environmental Pollution*, 120, pp. 533-549.

Z. Guo (2002b). Review of indoor emission source models. Part2. Parameter Estimation. *Environmental Pollution*, 120, pp. 551-564.

S. McBride, A. Ferro, W. Ott, P.Switzer and L. Hildemann (1999). Investigations of the proximity effect for pollutants in the indoor environment. To appear in *Journal of Exposure Analysis and Environmental Epidemiology.*

S. Ross (2003). *Introduction to Probability Models*, Eighth Edition, Academic Press.

M. Stein (1984). System parameters governed by jump processes: A model for removal of air pollutants. *Advances in Applied Probability,* 16, pp. 603-617.

B.A. Tichenor, Z. Guo, and L.E. Sparks (1993). Fundamental mass transfer model for indoor air emissions from surface coatings. *Indoor Air*, 3, pp. 263-268.

J. M. Wagner, U. Shamir, and D.H. Marks. (1994). Containing Groundwater Contamination: Planning Models using Stochastic Programming with Recourse", J.M. Wagner. *European Journal of Operational Research,* Vol.77(1994), pp,1-26.

World Bank Report (1998). Pollution Prevention and Abatement Handbook of World Bank. URL: http://wbln0018.worldbank.org/essd/essd.nsf/Docs/PPAH