AIHAJ - American Industrial Hygiene Association

ISSN: 1529-8663 (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/uaah20

Markov Modeling of Contaminant Concentrations in Indoor Air

Mark Nicas

To cite this article: Mark Nicas (2000) Markov Modeling of Contaminant Concentrations in Indoor Air, AIHAJ - American Industrial Hygiene Association, 61:4, 484-491, DOI: 10.1080/15298660008984559

To link to this article: https://doi.org/10.1080/15298660008984559

Markov Modeling of Contaminant Concentrations in Indoor Air

AUTHOR
Mark Nicas

Center for Occupational and Environmental Health, School of Public Health, University of California, Berkeley, CA 94720

Most models for contaminant dispersion in indoor air are deterministic and do not account for the probabilistic nature of the pollutant concentration at a given room position and time. Such variability can be important when estimating concentrations involving small numbers of contaminant particles. This article describes the use of probabilistic models termed Markov chains to account for a portion of this variability. The deterministic and Markov models are related in that the former provide the expected concentration values. To explain this relationship, a single-zone (well-mixed room) scenario is described as a Markov chain. Subsequently, a two-zone room is cast as a Markov model, and the latter is applied to assessing a health care worker's risk of tuberculosis infection. Airborne particles carrying *Mycobacterium tuberculosis* bacilli are usually present in small numbers in a room occupied by an infectious tuberculosis patient. For a given scenario, the Markov model permits estimates of variability in exposure intensity and the resulting variability in infection risk.

Keywords: contaminant dispersion, *Mycobacterium tuberculosis*, Markov model

This research was funded by a grant from the National Institute for Occupational Safety and Health, Centers for Disease Control and Prevention, Award No. K01-OH00155-01. Additional support was provided by the Division of Occupational Safety and Health, California Department of Industrial Relations. The ideas expressed are solely the author's and do not necessarily represent the official views of the funding agencies.

To describe contaminant dispersion in indoor air spaces that are imperfectly mixed, various investigators have used a multizone modeling approach.[1–5] For example, if one considers a single room, the space may be divided into different zones. Air within each zone is perfectly mixed, but there is limited airflow between physically contiguous zones. Differential equations describe the rate of change in the contaminant level in each zone. These expressions account for contaminant sources being located in one or more zones, a possible contaminant source in the supply air, contaminant dispersion via interzonal airflows, and potential contaminant removal by mechanisms other than ventilation. The resulting set of concentration functions $C_j(t)$, where the subscript denotes the jth zone, are *continuous*; that is, the $C_j(t)$ do not make sudden "jumps" in value from one instant of time to another. The $C_j(t)$ are also *deterministic*; that is, in a given zone at a given time, a single concentration value is predicted, and variability in this value is not considered.

Where small numbers of contaminant particles are present in room air, deterministic models fail to account for considerable variability in the particle concentration that can pertain at a given room position and time. In the alternative, Markov models describe a portion of this variability and account for discontinuity in concentration values. The two approaches are related in that the deterministic model equations provide the expected concentration values. To explain this relationship and to introduce basic Markov chain techniques, a well-mixed room (single zone) is first considered. Subsequently, a two-zone room is cast as a Markov model, and the latter is applied to occupational tuberculosis transmission. Airborne particles carrying *Mycobacterium tuberculosis* bacilli usually are present in small numbers (on the order of 1 to 10) in a room occupied by an infectious patient.[6]

MARKOV MODELS

A Single Zone

Assume there is a well-mixed room with N_0 contaminant particles present at time zero. Given some room volume V (m^3), the initial concentration (#/m^3) at time zero is N_0/V. Let the room have supply/exhaust air rate Q (m^3/sec). For simplicity, let the only mechanism of particle removal be ventilation air; however, other mechanisms of particle removal can be accommodated. Let there be no additional particle sources; that is, new particles do not enter the room via the supply air or by in-room release. The aim is

to describe the decrease in particle concentration over time. As a basis for comparison, a deterministic model for the particle concentration, C(t), is:

$$C(t) = \frac{N_0}{V}\exp\left(-\frac{Q}{V}\cdot t\right) \quad (1)$$

where time t is in seconds. At $t_{1/2} = (\ln 2)(V/Q)$ sec after the start of the process, $C(t_{1/2}) = 0.5\ N_0/V$, or the particle concentration is one-half the initial concentration.

In the alternative, consider the fate of an individual particle after t = 0. The time it spends in the room is modeled by the exponential probability density function $\lambda\ e^{-\lambda t}$, where $\lambda = Q/V$ sec^{-1}. The parameter λ is the probability per second that the particle is exhausted from the room. Given that the particle is currently in the room, the probability that the particle is in the room 1 sec later (after a 1-sec time step) is $e^{-(Q/V)}$, whereas the probability that the particle has left the room is the complement $1 - e^{-(Q/V)}$. These quantities are termed single-step transition probabilities.

The term "transition" refers to a particle's movement between "states." In this system, a particle is either in the room (State 1) or exhausted from the room (State 0). A particle in State 1 can move in the next 1-sec time step to State 0 with probability $P_{10} = 1 - e^{-(Q/V)}$, or can remain in State 1 with probability $P_{11} = e^{-(Q/V)}$, where the notation P_{ij} indicates a transition from the current State i to State j. A particle in State 0 remains in State 0 with probability $P_{00} = 1$, and has probability $P_{01} = 0$ of entering State 1 because the system was initially defined not to permit particles to enter the room. State 0 is termed an "absorbing" state, because once a particle transitions into State 0, it never leaves. These single-step transition probabilities are described by a probability matrix, denoted **P**:

$$\mathbf{P} = \begin{array}{c} \\ 0 \\ 1 \end{array}\begin{array}{c} \begin{array}{cc} 0 & 1 \end{array} \\ \begin{bmatrix} 1 & 0 \\ 1 - e^{-Q/V} & e^{-Q/V} \end{bmatrix} \end{array}$$

The numbers outside the matrix help denote the subscript state numbers for the P_{ij}. The row number designates where the particle currently resides, whereas the column number designates the state it moves to in the next time step. The probability entries in a given row always sum to 1.

For a particle currently in the room, the probability that it is still in the room right after the nth time step is denoted P_{11}^n, which is the entry in the second row and second column of the matrix $\mathbf{P}^{(n)}$. The latter matrix is **P** raised to the nth power. For example, $\mathbf{P}^{(2)} = \mathbf{P} \times \mathbf{P}$; $\mathbf{P}^{(3)} = \mathbf{P} \times \mathbf{P} \times \mathbf{P}$; and so forth. For the two-state system considered here, it can readily be shown that:

$$\mathbf{P}^{(n)} = \begin{array}{c} \\ 0 \\ 1 \end{array}\begin{array}{c} \begin{array}{cc} 0 & 1 \end{array} \\ \begin{bmatrix} 1 & 0 \\ 1 - e^{-(Q/V)n} & e^{-(Q/V)n} \end{bmatrix} \end{array}$$

Given that the particle fates are mutually independent, the number remaining at time n = t (where t must be an integer number of seconds) is a binomial random variable N(t) with expected value $E[N(t)] = N_0 \times P_{11}^t$, and with variance $Var[N(t)] = N_0 \times P_{11}^t \times (1 - P_{11}^t)$. The binomial probability that k particles are present at time t, for $0 \le k \le N_0$, is:

$$P[N(t) = k] = \binom{N_0}{k}(P_{11}^t)^k(1 - P_{11}^t)^{N_0-k}$$

In turn, the particle concentration at time t, C(t), is the random variable N(t)/V. The expected value and variance of C(t), denoted E[C(t)] and Var[C(t)], respectively, and the percentage coefficient of variation, denoted %CV[C(t)], are given by the following expressions:

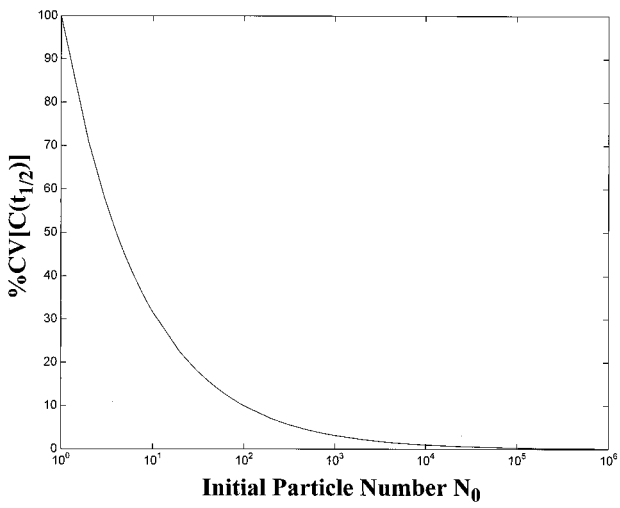

FIGURE 1. The relationship between the percentage coefficient of variation of the particle concentration at the half-time, %CV[C($t_{1/2}$)], and the initial number of particles, N_0. The curve is based on Equation 4 of the main text.

$$E[C(t)] = \frac{N_0 \times P_{11}^t}{V} \quad (2)$$

$$Var[C(t)] = \frac{N_0 \times P_{11}^t \times (1 - P_{11}^t)}{V^2} \quad (3)$$

$$\%CV[C(t)] = \frac{1}{\sqrt{N_0}}\sqrt{\frac{(1 - P_{11}^t)}{P_{11}^t}} \times 100\% \quad (4)$$

If N_0 particles initially are present, then after $t_{1/2} = (\ln 2)(V/Q)$ time steps (where it is assumed for simplicity that $t_{1/2}$ is an integer), $P_{11}^{t_{1/2}} = 0.5$. By Equation 2, the expected concentration is $0.5\ N_0/V$, the same value provided by the deterministic model Equation 1. However, the concentration value is explicitly treated as a random variable. At $t_{1/2}$, the variance is $0.25\ N_0/V^2$ by Equation 3, and the %CV is $(1/\sqrt{N_0}) \times 100\%$ by Equation 4. To illustrate, if $N_0 = 10$ particles and V = 25 m^3, then $E[C(t_{1/2})] = 0.2$ per m^3 and $\%CV[C(t_{1/2})] = 32\%$. The expected concentration is 0.2 per m^3, but there is a 5% chance that it might be ≥0.28 per m^3.

The $1/\sqrt{N_0}$ term in Equation 4 signifies that as N_0 increases, the %CV[C(t)] decreases. Figure 1 depicts $\%CV[C(t_{1/2})]$ for N_0 values ranging from 1 to 10^6. For $N_0 \ge 10^3$, $\%CV[C(t_{1/2})] \le 3.2\%$, which signifies little variability around the expected concentration value provided by the deterministic model Equation 1. Large N_0 values would pertain to the great majority of industrial hygiene scenarios. On the other hand, the $\sqrt{(1 - P_{11}^t)/P_{11}^t}$ term in Equation 4 signifies that as time passes subsequent to $t_{1/2}$, the %CV[C(t)] increases because P_{11}^t decreases. The toxicological significance of this increase in %CV is moderated in most cases, because relatively few particles are expected to remain. For example, when $N_0 = 10^3$ and $P_{11}^t = .001$, the %CV[C(t)] is 100%, but the expected particle number is 1.

In occupational exposure assessment, interest often focuses on the buildup in contaminant levels due to ongoing emissions. The single-zone Markov model can accommodate this process. Assume

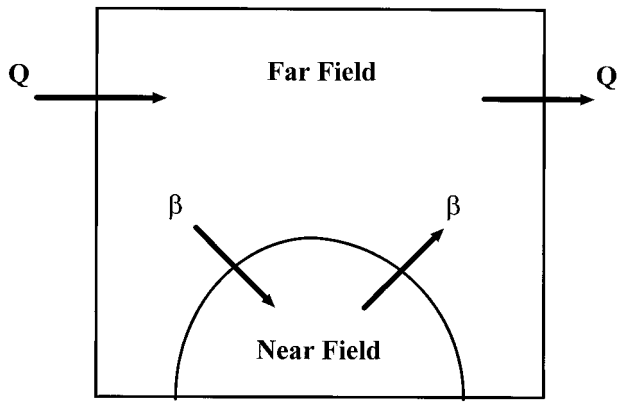

FIGURE 2. A hypothetical two-zone room containing a near-field zone around a contaminant emission source; the remainder of the room constitutes the far-field zone. Supply air enters and room air exits the far field at rate Q (m^3/sec). A balanced random air flow at rate β (m^3/sec) occurs between the near-field and far-field zones.

that at each value of time q in seconds (where q = 0, 1, 2, . . . , t), G_q particles are released, where the value of G_q is deterministic but may vary across the time steps. With the definitions $P_{11}^0 = 1$ and $P_{11}^1 = P_{11}$, the expected value and variance of C(t) are given by:

$$E[C(t)] = \frac{\sum_{q=0}^{t} G_q \times P_{11}^{t-q}}{V} \quad (5)$$

$$Var[C(t)] = \frac{\sum_{q=0}^{t} G_q \times P_{11}^{t-q} \times (1 - P_{11}^{t-q})}{V^2} \quad (6)$$

The above equations are derived in Appendix 1. This appendix also discusses the circumstance in which the G_q are probabilistic rather than deterministic quantities.

Multiple Zones

Consider a room air space divided into multiple zones. Particles exhausted from the room are assigned State 0 if particle reentry is not permitted. A transition probability $P_{ij} > 0$ pertains only for two physically contiguous zones (states) i and j. That is, there is zero probability that a particle moves between two noncontiguous zones in one time step. If there are n zones, the transition probability matrix **P** is a (n + 1) × (n + 1) matrix as shown below:

$$\mathbf{P} = \begin{array}{c} \\ 0 \\ 1 \\ \cdots \\ n \end{array} \begin{array}{c} \begin{array}{cccc} 0 & 1 & \cdots & n \end{array} \\ \begin{bmatrix} 1 & 0 & \cdots & 0 \\ P_{10} & P_{11} & \cdots & P_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ P_{n0} & P_{n1} & \cdots & P_{nn} \end{bmatrix} \end{array}$$

Assigning values a priori for the P_{ij} is difficult where numerous zones are involved, because one needs to account for the random and advective components of airflow between the zones. Appendix 2 discusses these probabilities in greater detail. To illustrate the basic concepts, however, a relatively simple two-zone model is considered.

Figure 2 depicts a hypothetical room. A near-field zone with volume V_N contains the emission source, and the remainder of the room constitutes a far-field zone with volume V_F.[5,7] Supply air enters, and exhaust air exits, the far-field zone at rate Q (m^3/sec). Air is exchanged between the near- and far-field zones at rate β (m^3/sec), which is the product of the room random air speed and the free surface area of the near- and far-field interface, divided by two. Contaminant particles are emitted directly into the near-field zone. The reason for formulating this model is to better estimate exposure intensity for a worker located close to an emission source; in general, the worker will have a higher exposure intensity than predicted by a traditional single-zone (well-mixed room) model.

In terms of a Markov model, let State 0 denote a particle's fate of being exhausted from the room, let State 1 denote the far-field zone, and let State 2 denote the near-field zone. The transition probability matrix **P** for this system is:

$$\mathbf{P} = \begin{array}{c} \\ 0 \\ 1 \\ 2 \end{array} \begin{array}{c} \begin{array}{ccc} 0 & 1 & 2 \end{array} \\ \begin{bmatrix} 1 & 0 & 0 \\ (1 - e^{-((Q+\beta)/V_F)})\frac{Q}{Q+\beta} & e^{-((Q+\beta)/V_F)} & (1 - e^{-((Q+\beta)/V_F)})\frac{\beta}{Q+\beta} \\ 0 & 1 - e^{-(\beta/V_N)} & e^{-(\beta/V_N)} \end{bmatrix} \end{array} \quad (7)$$

To explain the formulation of the P_{ij}, State 0 (exhausted from the room) is an absorbing state such that $P_{00} = 1$. A particle in State 1 (the far-field zone) can move to State 0 with exponential rate parameter Q/V_F, or to State 2 (the near-field zone) with exponential rate parameter β/V_F. Therefore, a particle leaves State 1 with an overall exponential rate parameter $(Q + \beta)/V_F$. The probability that it remains in State 1 over one time step is $P_{11} = e^{-(Q+\beta)/V_F}$. The probability that it leaves State 1 is the complement $1 - e^{-(Q+\beta)/V_F}$. Given that the particle leaves State 1, the probability that it moves to state 0 is $Q/(Q + \beta)$, and the probability that it moves to State 2 is $\beta/(Q + \beta)$. Therefore, $P_{10} = (1 - e^{-(Q+\beta)/V_F})[Q/(Q + \beta)]$, and $P_{12} = (1 - e^{-(Q+\beta)/V_F})[\beta/(Q + \beta)]$.

With regard to a particle in State 2, it cannot enter State 0 in one time step (because it must first move through State 1), so $P_{20} = 0$. However, the particle can enter State 1 with exponential rate parameter β/V_N. Therefore, the probability that it remains in State 2 over one time step is $P_{22} = e^{-\beta/V_N}$, and the probability that it enters State 1 is $P_{21} = 1 - e^{-\beta/V_N}$.

Given that N_0 particles are in the near-field zone and none are in the far-field zone at time zero, and given that no additional particles are released, the expected value and variance of the far-field and near-field particle concentrations at time t are given by the following equations, where j = 1 for the far-field zone and j = 2 for the near-field zone:

$$E[C_j(t)] = \frac{N_0 \times P_{2j}^t}{V_j} \quad (8)$$

$$Var[C_j(t)] = \frac{N_0 \times P_{2j}^t \times (1 - P_{2j}^t)}{V_j^2} \quad (9)$$

For a buildup scenario, assume that at time q in seconds (where q = 0, 1, 2, . . . , t), G_q particles are released into the near-field zone, where the G_q values are deterministic but may vary across the time steps. The expected value and variance of the far-field and near-field particle concentrations at time t are given by the following equations, where j = 1 for the far-field zone and j = 2 for the near-field zone:

$$E[C_j(t)] = \frac{\sum_{q=0}^{t} G_q \times P_{2j}^{t-q}}{V_j} \quad (10)$$

$$Var[C_j(t)] = \frac{\sum_{q=0}^{t} G_q \times P_{2j}^{t-q} \times (1 - P_{2j}^{t-q})}{V_j^2} \quad (11)$$

APPLICATION TO TUBERCULOSIS TRANSMISSION

Infection by *Mycobacterium tuberculosis* (*M. tb*) usually occurs via inhalation of respirable particles carrying viable bacilli. In health care settings, such particles originate most often from the respiratory tract emissions of pulmonary tuberculosis (TB) patients. In turn, infection risk among health care workers (HCWs) depends on the airborne *M. tb* particle concentration to which they are exposed. To predict the latter concentration for purposes of risk assessment and management, deterministic models have been used.[1,8] However, *M. tb* apparently are released in small numbers (on the order of 1 to 10 per hour), in which case their concentration in any room zone can abruptly fluctuate between zero and positive values. Given that an HCW is usually in close proximity to the emission source when attending a TB patient, it is reasonable to apply the near-field/far-field Markov model to describe the HCW's *M. tb* exposure intensity.

For discussion purposes, let the near-field zone be a 1-m radius hemisphere centered on a patient lying on a bed, in which case V_N = 2.1 m³ (i.e., $(2/3)\pi r^3$). The 1-m radius represents the distance between the TB patient and HCW. Let V_F = 47.9 m³, such that the total room volume ($V_N + V_F$) is 50 m³; the latter is a reasonable value for a patient isolation room. Let Q = .0833 m³/sec, which corresponds to six nominal room air changes per hour and adheres to ventilation guidelines recommended by the Centers for Disease Control and Prevention for TB patient isolation rooms.[9] Let β = 0.32 m³/sec, which corresponds to an air speed of 0.1 m/sec (20 ft/min); the latter represents relatively still air and corresponds to the median air speed measured in a survey of indoor workplaces.[10] Specifying V_N, V_F, Q, and β also specifies the P_{ij} values in the Equation 7 matrix.

Next, assume that no *M. tb* particles are present initially, but that at time zero, N_0 = 10 *M. tb* particles are released by a cough into the near-field zone, with no subsequent release; for simplicity, assume that no particle carries more than one bacillus. Therefore, the near-field *M. tb* particle concentration is 4.8 per m³ at time zero (N_0/V_N).

The solid lines labeled $E[C_2(t)]$ and $E[C_1(t)]$ in Figure 3 depict the expected values of the near-field and far-field *M. tb* concentrations, respectively, over the first 60 sec following the pulse release. There is a rapid decrease in $E[C_2(t)]$ from 4.8 per m³ to 0.21 per m³ by t = 40 sec, at which time it is close to $E[C_1(t)]$ = 0.19 per m³. The subsequent decrease in both expected concentrations is relatively gradual. The dotted line immediately above each solid line is the approximate 95th percentile value for the corresponding particle concentration. For example, at t = 25 sec, $E[C_2(t)]$ = 0.29 per m³, and the probability is 98% that $C_2(t) \leq 0.95$ per m³. The approximate 95th percentile values for $C_2(t)$ and $C_1(t)$ were found by computing the binomial probabilities for the possible particle numbers in each zone based on the values of P_{22}^t and P_{21}^t, respectively, and then finding that particle number in each zone for which the cumulative probability was closest to 95%.

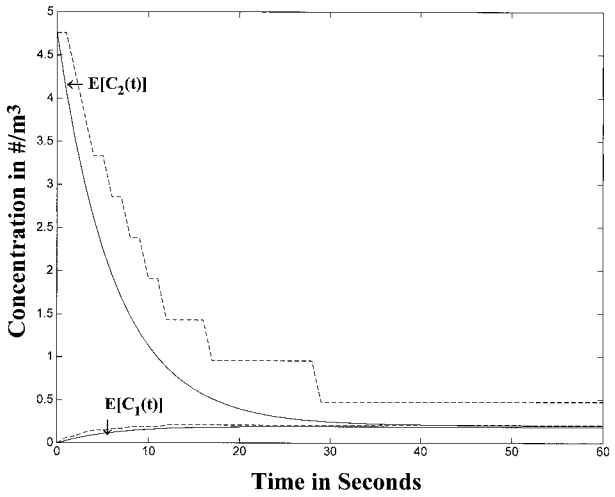

FIGURE 3. The *M. tb* particle concentrations in the near-field and far-field zones in the first 60 sec following a pulse release of 10 particles into the near field at time zero. The solid-line curve labeled "$E[C_2(t)]$" is the expected near-field particle concentration, and the solid-line curve labeled "$E[C_1(t)]$" is the expected far-field particle concentration. The dotted line immediately above each expected concentration curve is the approximate 95th percentile value for the corresponding particle concentration.

For example, at t = 25 sec, $Pr[N_2(25) \leq 1] = 0.88$, and $Pr[N_2(25) \leq 2] = 0.98$. The event $N_2(25) \leq 2$ corresponds to $C_2(25) \leq 0.95$ per m³, in which case 0.95 per m³ was plotted as the approximate 95th percentile value for the near-field *M. tb* concentration.

The health significance of variability in the near-field concentration is that it increases the variability in the risk of infection, which in turn increases uncertainty in selecting appropriate control measures. To explain this idea, assume that a susceptible HCW is in a TB patient's near-field zone over a T-sec interval. The worker's infection risk is a function of the expected number of *M. tb* particles that are inhaled and deposited in the pulmonary region during this interval, denoted D(T). In turn, D(T) is the product of the average *M. tb* particle concentration over the interval, the air volume inhaled by the worker over the interval, b (m³), and the inhaled particle deposition fraction in the alveolar region, f_D (0 to 1), which is the target site for infection:

$$D(T) = \frac{b \times f_D}{T} \sum_{t=0}^{T-1} C_2(t) \quad (12)$$

where $C_2(t)$ denotes the near-field *M. tb* particle concentration at time t. The summed concentrations run from t = 0 to t = T − 1, and not to t = T, because the concentration value $C_2(T)$ pertains only after the Tth time step. Dividing the summed concentrations by T yields the average concentration over the interval.

Because it is thought that the infectious dose is one *M. tb* bacillus, infection risk R is modeled by the one-hit expression: $R = 1 - e^{-D(T)}$.[11,12] However, if D(T) is a small number, say, ≤.05, then a good approximation is: $R \cong D(T)$. Because D(T) is a function of the random variable $C_2(t)$, D(T) and R are also random variables. The expected infection risk, E[R], and the variance of the infection risk, Var[R], are approximately equal to E[D(T)] and Var[D(t)], respectively. The latter quantities are defined by:

$$E[D(T)] = \frac{b \times f_D}{T} \sum_{t=0}^{T-1} E[C_2(t)] \quad (13)$$

$$Var[D(T)] = \left(\frac{b \times f_D}{T}\right)^2 \left[\sum_{t=0}^{T-1} Var[C_2(t)] + 2 \sum_{t=0}^{T-2} \sum_{s>t}^{T-1} Cov[C_2(t), C_2(s)]\right] \quad (14)$$

The $E[C_2(t)]$ and $Var[C_2(t)]$ terms correspond to previous Equations 8 and 9, respectively, but a concise analytical expression is not available for $Cov[C_2(t), C_2(s)]$. Instead, the value of $Var[D(T)]$ can be approximated by Monte Carlo simulation as follows.

In a given simulation run, each particle is followed for $T-1$ one-sec time steps. For each particle at each time step, a uniform random variate U between 0 and 1 is generated. If the particle is in State 2, at the next time step the particle moves to State 1 if $U \leq P_{21}$, or remains in State 2 if $P_{21} < U \leq 1$. If the particle is in State 1, at the next time step the particle moves to State 0 if $U \leq P_{10}$, remains in State 1 if $P_{10} < U \leq (P_{10} + P_{11})$, or moves to State 2 if $(P_{10} + P_{11}) < U \leq 1$. If the particle is in State 0, the particle remains in State 0. The number of particles in State 2 at time step t is divided by the near-field volume to obtain $C_2(t)$. To compute the expected pulmonary dose, the T successive C_2 values are summed and divided by T to yield the TWA value for C_2, which is multiplied by the factor $(b \times f_D)$ to give D(T) for the simulation run. For this analysis, 10^5 simulation runs were performed, and the variance of the 10^5 values for D(T) was computed.

Returning to the scenario depicted in Figure 3, assume that the HCW's exposure starts at time zero when $N_0 = 10$ *M. tb* particles are released into the near-field zone (with no subsequent release), and lasts for T = 300 sec. Further assume that the HCW inhales b = .083 m³ in 300 sec, which corresponds to inhaling 1 m³/hour. Because the aerodynamic diameter of an *M. tb* particle is approximately 3 μm,[11] let $f_D = 0.3$.[13] Based on Equation 13, E[D(T)] = .0069, in which case E[R] ≅ .0069. Based on the Monte Carlo simulation, $Var[D(T)] \cong 2.8 \times 10^{-6}$ (the %CV of the estimate is <2%), in which case the standard deviation SD[R] ≅ .0017, and the associated %CV[R] ≅ 25%. These values signify that while the expected infection risk is .0069 for the exposure period, there is a reasonable chance that it can be as high as .0097 (the approximate 95th percentile value: E[R] + 1.645 SD[R]). Figure 4 is a histogram of 10^5 simulation outcomes of the approximate R value for this scenario; the outcomes reasonably conform to a normal distribution. The mean is .0069, the standard deviation is .0017, and 5.9% of the values exceed .0097. Note that while an infection risk on the order of .01 (1%) may seem small, experiencing just 10 similar exposures produces a cumulative infection risk of approximately 10%, or: $1 - (1 - .01)^{10} = .096$.

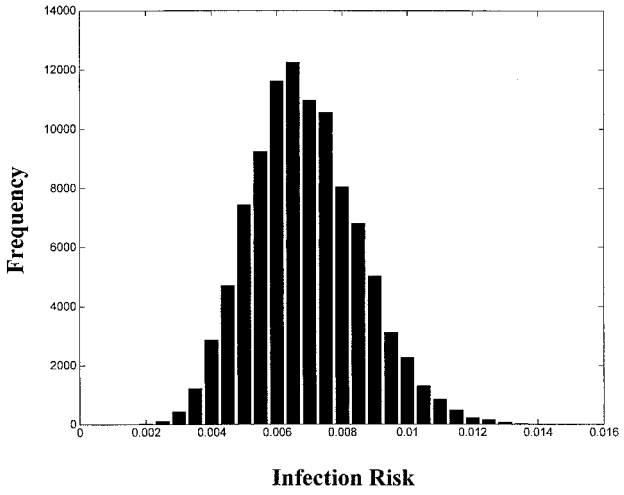

FIGURE 4. The distribution of *M. tb* infection risk values across 10^5 simulations of a scenario in which a health care worker is in the near field of a tuberculosis patient for 5 min immediately after the patient emits 10 *M. tb* particles. For this distribution, E[R] ≅ .0069 and SD[R] ≅ .0017.

DISCUSSION

For a scenario in which relatively small numbers of contaminant particles are involved, the failure to account for the probabilistic nature of particle dispersion and fate may lead to formulating an exposure control strategy that is inconsistent with the intended goal. For example, the tuberculosis standard proposed by the Occupational Safety and Health Administration (OSHA) requires that when a room or area is vacated by a TB patient, the room must be ventilated to achieve a 99.9% "removal efficiency" before employees may enter without respiratory protection.[14] Therefore, if a room receives six nominal air changes per hour, the ventilation system must be run for 69 min, at which time the *M. tb* particle concentration is expected to be 0.1% of the initial concentration according to Equation 1. Because OSHA did not propose an acceptable concentration of *M. tb* in air, OSHA may have intended that 99.9% "removal efficiency" also signify a 99.9% probability that no *M. tb* particles are present in room air on entry. However, requiring 99.9% "removal efficiency" fails to provide 99.9% probability that no *M. tb* particles are present.

To illustrate, consider a room having V = 50 m³ and Q = 300 m³/hour, but let $N_0 = 40$ *M. tb* particles. Such a scenario might involve a bronchoscopy on a pulmonary TB patient who emits 250 *M. tb* particles per hour over the course of a 1.5 hour procedure, as reported in a case study;[15] it can be shown that the number of *M. tb* particles in room air would be approximately 40 by the end of the bronchoscopy in the room just specified. For simplicity, if the room is assumed to be well-mixed, and if the dilution ventilation system operates for 69 min after the procedure, the probability that any given particle remains is .001. The binomial probability that none of the 40 *M. tb* particles remain is 0.961, or only 96.1%.

Although the two-zone Markov model presented in this article accounts for concentration variability at different room locations as well as time, it does not fully describe the variability that pertains. The reason is that the transition probabilities P_{ij} have been treated as stationary (constant), whereas they might vary with time. For the two-zone model, nonstationarity signifies that β and/or Q are variable. If the great majority of room supply air is mechanically delivered (as opposed to entering via infiltration), and if the mechanical supply rate is designed to be constant, it is reasonable to treat Q as constant. However, it is not realistic to treat β as stationary if the air speed at the interface of the near- and far-field zones exhibits substantial variability. In general, the impact of nonstationarity is to increase the %CV of the particle concentration.

Markov modeling is widely used in the physical sciences,[16] but to the author's knowledge it has had little application in describing contaminant dispersion in indoor air. A recent paper on indoor air

modeling presented an elegant discussion of Markov chain techniques, although the investigators used a continuous-time Markov model, and the buildup of the contaminant concentration due to ongoing emission was not considered.[17] A continuous-time model yields equations of the form $P_{ij}(t)$, which represents the probability that a particle in State i is in State j at subsequent time t, where time is continuous. The two-zone discrete-time model using a 1-sec time step behaves similarly to the analogous continuous-time model for air speeds $\leq$ 0.25 m/sec (50 ft/min). As the air speed increases beyond 0.25 m/sec, the discrete-time model starts to substantially overpredict the near-field concentration relative to the continuous-time analog. For high air speeds, say, 5 m/sec (1000 ft/min), the discrete-time model using a 1-sec time step displays unrealistic oscillatory behavior. These potential errors should not be important in practice, because air speeds in indoor environments tend to be <0.25 m/sec, and one can use a smaller time step to accommodate higher speeds. However, if there is strong directional airflow near an emission source due to, say, the location of an air supply register or operation of a fan, the two-zone model would be inappropriate.

Markov models are complementary to other techniques for indoor air modeling. Simple turbulent diffusion models have been used to describe the gradient in contaminant concentration near an emission source.[18,19] Diffusion models can also incorporate an advective flow component,[20] as well as contaminant reflection by wall boundaries and removal by room ventilation.[21] Computational fluid dynamics (CFD) models are increasingly being applied to indoor air quality studies,[3] and specifically to workplace settings.[22,23] As techniques are refined, CFD models may provide highly accurate descriptions of contaminant dispersion in a wide array of indoor environments. The main advantages of the Markov modeling approach may be its relative simplicity and explicit probabilistic treatment of contaminant dispersion and removal processes.

ACKNOWLEDGMENTS

The author thanks Dr. William Nazaroff for his insights on the probability structure of a box model, and Dr. Jim Pitman for his elucidation of Markov chains and related stochastic processes.

REFERENCES

1. **Riley, R.L., and S. Permutt:** Room air disinfection by ultraviolet irradiation of upper air: Air mixing and germicidal effectiveness. *Arch. Environ. Health 22*:208–219 (1971).
2. **Wadden, R.A., and P.A. Scheff:** *Indoor Air Pollution: Characterization, Prediction and Control*. New York: John Wiley & Sons, 1983.
3. **Etheridge, D., and M. Sandberg:** *Building Ventilation: Theory and Measurement*. New York: John Wiley & Sons, 1996.
4. **Kee-Chiang, C.:** Airborne contaminant exposure control in a partitioned work environment by exhaust ventilation systems. *Am. Ind. Hyg. Assoc. J. 59*:346–352 (1998).
5. **Nicas, M.:** Estimating exposure intensity in an imperfectly mixed room. *Am. Ind. Hyg. Assoc. J. 57*:542–550 (1996).
6. **Riley, R.L., and F. O'Grady:** *Airborne Infection: Transmission and Control*. New York: MacMillan, 1961. pp. 89–92, 95–125.
7. **Furtaw, E.J., M.D. Panadian, D.R. Nelson, and J.V. Behar:** Modeling indor air concentrations near emission sources in imperfectly mixed rooms. *J. Air Waste Manage. Assoc. 46*:861–868 (1996).
8. **Nicas, M., and S.L. Miller:** A multi-zone model evaluation of the efficacy of upper-room air ultraviolet germicidal irradiation. *Appl. Occup. Environ. Hyg. 14*:317–328 (1999).
9. **Centers for Disease Control and Prevention:** Guidelines for preventing the transmission of *Mycobacterium tuberculosis* in health-care facilities, 1994. *Morb. Mort. Wkly. Rep. 43(RR–13)*:1–132 (1994).
10. **Baldwin, P.E.J., and A.D. Maynard**: A survey of wind speeds in indoor workplaces. *Ann. Occup. Hyg. 42*:303–313 (1998).
11. **Wells, W.F.:** *Airborne Contagion and Air Hygiene*. Cambridge, MA: Harvard University Press, 1955. pp. 105–141.
12. **Riley, R.L., and E.A. Nardell:** Clearing the air: the theory and application of ultraviolet air disinfection. *Am. Rev. Resp. Dis. 139*:1286–1294 (1989).
13. **Hinds, W.C.:** *Aerosol Technology*. New York: John Wiley & Sons, page 220.
14. "Occupational Exposure to Tuberculosis—Proposed Rule." *Federal Register* 62 (17 October 1997), pp. 54160–54308.
15. **Catanzaro, A.:** Nosocomial tuberculosis. *Am. Rev. Resp. Dis. 125*: 559–562 (1982).
16. **Gillespie, D.T.**: *Markov Processes: An Introduction for Physical Scientists*. San Diego: Academic Press, Inc., 1992.
17. **Peng, S.H., and L. Davidson:** Towards the determination of regional purging flow rate. *Building Environ. 32*:513–525 (1997).
18. **Roach, S.A.:** On the role of turbulent diffusion in ventilation. *Ann. Occup. Hyg. 24*:105–132 (1981).
19. **Wadden, R.A., J.L. Hawkins, P.A. Scheff, and J.E. Franke:** Characterization of emission factors related to source activity for trichloroethylene degreasing and chrome plating processes. *Am. Ind. Hyg. Assoc. J. 52*:349–356 (1991).
20. **Scheff, P.A., R.L. Friedman, J.E. Franke, L.M. Conroy and R.A. Wadden:** Source activity modeling of Freon® emissions from open-top vapor degreasers. *Appl. Occup. Environ. Hyg. 7*:127–134 (1992).
21. **Drivas, P.J., P. Valberg, B.L. Murphy and R. Wilson:** Modeling indoor air exposure from short-term point source releases. *Indoor Air 6*:271–277 (1996).
22. **Andersson, I.M., and S. Alenius:** A comparison between measured and numerically calculated styrene concentrations in hand lay-up moulding. *Occup. Hyg. 3*:399–415 (1996).
23. **Buchanan, C.R., and D. Dunn-Rankin:** Transport of surgically produced aerosols in an operating room. *Am. Ind. Hyg. Assoc. J. 59*:393–402 (1998)
24. **Ross, S.M.:** *Introduction to Probability Models, Fifth Edition*. San Diego, CA: Academic Press, 1993, pp. 231–232.
25. **Taylor, H.M. and S. Karlin:** *An Introduction to Stochastic Modeling, Third Edition*. San Diego, CA: Academic Press, 1998, p. 72.

APPENDIX 1

The Particle Concentration Buildup Equations for A Single-Zone Room

Consider the deterministic number G_0 particles released at time q = 0. The number remaining at time t (i.e., after t time steps) is a binomial random variable with expected value $G_0 \times P_{11}^t$ and variance $G_0 \times P_{11}^t \times (1 - P_{11}^t)$, where P_{11}^t is the entry in the matrix $\mathbf{P^{(t)}}$. Next, consider the G_1 particles released at time q = 1 sec. The number remaining at time t (i.e., after t − 1 time steps) is an independent binomial random variable with expected value $G_1 \times P_{11}^{t-1}$ and variance $G_1 \times P_{11}^{t-1} \times (1 - P_{11}^{t-1})$, where P_{11}^{t-1} is the entry in the matrix $\mathbf{P^{(t-1)}}$. The assumption made is that the fates of all particles are mutually independent. Analogous terms are written for the particles released at each time step. The total number of particles remaining at time t is the sum N(t). Its expected value and variance are as follows:

$$E[N(t)] = \sum_{q=0}^{t} G_q \times P_{11}^{t-q} \qquad \textbf{(A1)}$$

$$Var[N(t)] = \sum_{q=0}^{t} G_q \times P_{11}^{t-q} \times (1 - P_{11}^{t-q}) \qquad \textbf{(A2)}$$

Because $C(t) = N(t)/V$, Equations 5 and 6 in the main text for $E[C(t)]$ and $Var[C(t)]$, respectively, immediately follow. If $G_q = G$ for all $q = 0, 1, \ldots, t$, it can be shown that as $t \to \infty$, $E[C(t)]$ approaches the same steady-state value G/Q provided by the traditional deterministic buildup equation: $C(t) = G/Q[1 - e^{-(Q/V)\cdot t}]$, where G is the number of particles continuously emitted per second. Further, as $t \to \infty$, $Var[C(t)] \cong G/(2VQ)$. Note that buildup Equations 10 and 11 of the main text for the near-field/far-field Markov model have an analogous derivation.

Where the number of particles released into a well-mixed room at the qth time step is a probabilistic rather than deterministic quantity, the value of $Var[N(t)]$ will usually be greater than indicated by Equation A2. Consider the simplest case where the G_q are independent and identically distributed (IID) random variables with $E[G_q] = \mu_G$ and $Var[G_q] = \sigma_G^2$. The parameter $E[N(t)]$ is given by:

$$E[N(t)] = \mu_G \sum_{q=0}^{t} P_{11}^{t-q} \cong \frac{\mu_G V(1 - P_{11}^{t+1})}{Q} \qquad \textbf{(A3)}$$

$E[N(t)]$ follows from Equation A1 by taking expectations and simplifying the sum using the power series expression: $1 + x + x^2 + \ldots + x^t = (1 - x^{t+1})/(1 - x)$, where $x = P_{11}$. Further, $1 - P_{11} \cong Q/V$, because: $P_{11} = e^{-(Q/V)} \cong 1 - (Q/V)$, for $Q/V \ll 1$.

$Var[N(t)]$ is derived by conditioning. Let R_q denote the number of particles released at time q that remain at time t. Given $G_q = g$, R_q is a binomial random variable with parameters g and P_{11}^{t-q}. The conditional expectation and variance are: $E[R_q | G_q = g] = g \times P_{11}^{t-q}$, and $Var[R_q | G_q = g] = g \times P_{11}^{t-q} \times (1 - P_{11}^{t-q})$. The unconditional expectation of R_q is: $E[R_q] = E[G_q \times P_{11}^{t-q}] = \mu_G \times P_{11}^{t-q}$. The unconditional variance of R_q is:

$$\begin{aligned} Var[R_q] &= E[Var[R_q | G_q]] + Var[E[R_q | G_q]] \\ &= E[G_q \times P_{11}^{t-q} \times (1 - P_{11}^{t-q})] \\ &\quad + E[(G_q \times P_{11}^{t-q} - \mu_G \times P_{11}^{t-q})^2] \\ &= \mu_G \times P_{11}^{t-q} \times (1 - P_{11}^{t-q}) + \sigma_G^2 \times (P_{11}^{t-q})^2 \end{aligned}$$

Due to the independence assumptions, $Var[N(t)]$ is the sum of the $Var[R_q]$ terms over the range $q = 0, 1, \ldots, t$, or:

$$\begin{aligned} Var[N(t)] &= \sum_{q=0}^{t} [\mu_G \times P_{11}^{t-q} \times (1 - P_{11}^{t-q}) + \sigma_G^2 \times (P_{11}^{t-q})^2] \\ &\cong \frac{\mu_G V(1 - P_{11}^{t+1})}{Q} + \frac{(\sigma_G^2 - \mu_G) V(1 - (P_{11}^{t+1})^2)}{2Q} \end{aligned} \qquad \textbf{(A4)}$$

The sum is simplified using the previous power series expression and the approximation for $e^{-(Q/V)}$. In Equation A4,
$Var[N(t \to \infty)] \cong (\mu_G + \sigma_G^2)V/2Q$,
in which case
$Var[C(t \to \infty)] \cong (\mu_G + \sigma_G^2)/2VQ$.

Another stochastic process that may apply in some scenarios is a Poisson arrival process with variable amplitude and exponential decay.[24] Assume that emission events in a well-mixed room occur according to a Poisson process with rate λ (number per sec), where λ is small, such that one expects a limited number of events per hour. The expected number of emission events from time 0 to t (in sec) is $\lambda \times t$. The number of particles released per event is an IID random variable A with mean μ_A and variance σ_A^2. Particle removal from the room is modeled by the density function $\alpha e^{-\alpha t}$; if ventilation is the only particle removal mechanism, $\alpha = Q/V$. The number of particles present at time t, N(t), is a random sum:

$$N(t) = \sum_{i=0}^{M} A_i \times P_{11}^{t-s_i} \qquad \textbf{(A5)}$$

where M is a Poisson variable with mean $\lambda \times t$, A_i is the number of particles released at the ith emission event, $P_{11} = e^{-Q/V}$, and s_i denotes the integer time step at which the ith emission event occurs, $0 \le s_i \le t$.

Based on the equations for the mean and variance of a random sum,[25] the mean and variance of N(t) are:

$$E[N(t)] \cong \frac{\lambda \mu_A V(1 - P_{11}^{t+1})}{Q} \qquad \textbf{(A6)}$$

$$Var[N(t)] \cong \frac{\lambda(\mu_A^2 + \sigma_A^2) V(1 - (P_{11}^{t+1})^2)}{2Q} \qquad \textbf{(A7)}$$

In Equation A7, $Var[N(t \to \infty)] \cong \lambda(\mu_A^2 + \sigma_A^2)V/2Q$, in which case $Var[C(t \to \infty)] \cong \lambda(\mu_A^2 + \sigma_A^2)/2VQ$. Note that the discrete-time process just described is not strictly a Poisson arrival process, because the latter applies to continuous time. However, if λ is defined as the probability of an emission event at each discrete time point, and if λ is small, say, $<.02$, Equation A7 is an adequate approximation.

The preceding processes are now used to demonstrate the increase in variance where the number of particles released or where the time of emission is probabilistic rather than deterministic. Consider the expressions for $Var[C(t \to \infty)]$ for the process in which: (1) a constant number μ_G particles is released at each time step (note: $\sigma_G^2 = 0$); (2) the number of particles released at each time step is IID with mean μ_G and $\sigma_G^2 > 0$; and (3) emissions follow a Poisson arrival process and the number of particles released per event is IID with mean μ_A and $\sigma_A^2 > 0$:

Process (1): $Var[C(t \to \infty)] = \mu_G/2VQ$

Process (2): $Var[C(t \to \infty)] = (\mu_G + \sigma_G^2)/2VQ$

Process (3): $Var[C(t \to \infty)] = \lambda(\mu_A^2 + \sigma_A^2)/2VQ$

In process (3), let $\mu_A = \mu_G/\lambda$, in which case the values of $E[C(t \to \infty)]$ for all three processes equal μ_G/Q. For simplicity of comparison, in process (3) let $\sigma_A^2 = \sigma_G^2/\lambda$. To explain, if the number of emission events over t time steps equals the expected number $\lambda \times t$, and if $\mu_A = \mu_G/\lambda$, the expected number of particles emitted is $\lambda \times t \times \mu_A = t \times \mu_G$, which corresponds to the expected number of particles released over t time steps in both processes (1) and (2). Because the variance of the number of particles released over t time steps in process (2) equals $t \times \sigma_G^2$, setting $\sigma_A^2 = \sigma_G^2/\lambda$ causes the variance of the total number of particles released in process (3) to equal that in process (2), on average.

Next, let $\mu_G = 10^3$ particles per emission event, $\sigma_G^2 = 2.5 \times 10^5$, $\lambda = 1.67 \times 10^{-3}$/sec (or 6 events per hour), $V = 50$ m^3, and $Q = 8.33 \times 10^{-2}$ m^3/sec (or 6 air changes per hour for $V = 50$ m^3). For all three processes, $E[C(t \to \infty)] = 1.2 \times 10^4$ per m^3. For the three processes, the respective values of $Var[C(t \to \infty)]$ are: (1) 1.2×10^2 per m^6; (2) 3.0×10^4 per m^6; and (3) 7.2×10^7 per m^6. Finally, the values of the percentage coefficient of variation, or $(\sqrt{Var[C(t \to \infty)]}/E[C(t \to \infty)]) \times 100\%$, for the three processes, are:

Process (1): $\%CV[C(t \to \infty)] = 0.091\%$;

Process (2): $\%CV[C(t \to \infty)] = 1.4\%$;

Process (3): $\%CV[C(t \to \infty)] = 70\%$;

APPENDIX 2

Transition Probabilities for a Multiple-Zone Markov Model

In a multiple-zone model, one might conceptually divide the room into n three-dimensional zones by perpendicular planes running parallel to the room's length, width, and height axes. The transition probability matrix **P** is $(n + 1) \times (n + 1)$ as presented in the main text. Probabilities $P_{ij} > 0$ pertain only for physically contiguous zones i and j, and the P_{ij} must account for the random and advective components of airflow. Consider an interior zone (one not bordering a room surface), denoted i with volume V_i, which shares one surface with each of six contiguous zones, $k = 1, 2, \ldots, 6$. Let the random and advective components of airflow from i to the six contiguous zones be denoted, respectively, β_{ik} and $f_{ik} \times Q$, where $0 \leq f_{ik} \leq 1$. At least one of the f_{ik} must equal zero, and it is possible that all equal zero.

A particle in zone i leaves that zone with an overall exponential rate parameter λ given by:

$$\lambda = \frac{\sum_{k=1}^{6} \beta_{ik} + \sum_{k=1}^{6} f_{ik} \times Q}{V_i}$$

The probability that a particle in zone i remains in zone i over one time step is:

$$P_{ii} = \exp\left[-\left(\sum_{k=1}^{6} \beta_{ik} + \sum_{k=1}^{6} f_{ik} \times Q\right) \Big/ V_i\right]$$

The probability that a particle in zone i moves to a contiguous zone k′ over one time step is:

$$P_{ik'} = \left\{1 - \exp\left[-\left(\sum_{k=1}^{6} \beta_{ik} + \sum_{k=1}^{6} f_{ik} \times Q\right) \Big/ V_i\right]\right\} \times \left(\frac{\beta_{ik'} + f_{ik'} \times Q}{\sum_{k=1}^{6} \beta_{ik} + \sum_{k=1}^{6} f_{ik} \times Q}\right)$$