Probabilistic Decision-Making of Air Emission Control Alternatives

Evan Paleologos
Abu Dhabi University

Probabilistic Decision-Making of Air Emission Control Alternatives

Evan K Paleologos[1], **Mohamed Elhakeem**[1], **and Mohamed El Amrousi**[2]

[1] Department of Civil Engineering, Abu Dhabi University, Abu Dhabi 59911, UAE
[2] Department of Architecture, Abu Dhabi University, Abu Dhabi 59911, UAE
[3] Corresponding author: Dr. Evan K Paleologos evan.paleologos@adu.ac.ae

Abstract. The polluter pays principle that has been adopted in environmental regulations worldwide places the economic burden of prevention, control, and cost of damage and mitigation on the pollution generator. The United States and Canada have established monetary penalties for air emission violations based on formulas and notions that account, among others, for the harm done by a violation to the environment and to human health, the environmental history of the violator, and the economic benefits reaped as a result of noncompliance. Despite their legal completeness these regulations do not address the probabilistic nature of air pollution. The current paper recasts the issue of air pollution penalties in a Bayesian decision-making framework where the prior probability distribution function from historical data on emissions or experts' opinions can be constructed and used to establish optimal decisions regarding new emission controls. Our article advances the use of the loss function as an alternative risk analysis tool to the current deterministic penalty regulatory formulas, which can also be used as a public policy instrument to promote environmentally, friendlier air emission choices.

1. Introduction

The polluter pays principle (PPP) was initially adopted by the Organization for Economic Co-operation and Development (OECD) in 1972 as the fundamental economic principle based on which those that generate pollution bear also the cost of preventing, controlling, and mitigating its effects on human health and the environment. The PPP was reaffirmed in the OECD recommendation of November 1974, and since then it has become the fundamental principle of regulations regarding land, water, and air pollution, in practically every country [1]. The European Union adopted the PPP in its first Environmental Action Program (1973-1976), in the March 3, 1975 Recommendation regarding cost allocation and action by public authorities on environmental matters, and since 1987 in the Treaty of the European Communities. The PPP was also established as Principle 16 of the UN Declaration on Environment and Development in the 1992 Rio Declaration and it is being applied to emissions of greenhouse gases.

Establishing specific penalties on emission violations depends on the regulations of a country. Canada utilizes a formula where the penalty is composed of four components, a baseline penalty that depends on the type of violation and category of violator, a history of non-compliance penalty, which examines whether the violator had a history of pollution the last five years preceding the violation, an environmental harm penalty that assesses the harm done to the environment, and finally an economic gain component, which assesses whether the violator gained from non-complying with the environmental regulations. Maximum limits are imposed on all penalty components [2].

EPA has established detailed penalty policy violations of the Clean Air Act, depending on the harm done to the environment and the extent of deviation from emission standards. EPA has established conformity, so that similar cases are assessed proportional penalties, the penalties are judicially defensible, and the agency's employees apply uniformly its standards. The penalty policies are characterized by a deterrence component, both toward the violating party and to the general public to cease breach of the law and to deter it from violating activities; fair and equitable treatment; and swift judicial resolution. Mitigating factors, such as past environmental history, degree of cooperation, and wilfulness or negligence, etc. of the violator are also considered in the assessment of cases [3, 4].

Deterrence is composed of two parts: the benefit and the gravity components. The benefit component seeks to remove any economic benefits reaped by the violator as a result of noncompliance, such as delayed costs or any other costs avoided as a result of noncompliance. The gravity component addresses the damage done to the environment and aims to ensure that the violator is worse off economically than if he had complied with the law. The sum of the benefit and gravity components constitutes the preliminary deterrence amount, which is modified by mitigating factors to arrive at the initial penalty target to be used in settlement negotiations. Recently, the U.S. EPA adjusted for inflation its civil penalties, and in the case of emissions from stationary sources these were set, for the gravity component, at \$37,500 for each day of violation, the total gravity penalty not to exceed \$320,000 [5, 6, 7].

The current article addresses the issue of air pollution violations by recasting the above concepts in a probabilistic framework. Initially, the Bayesian decision theory is used to extend the standard cost-benefit analysis of investing on new air emission controls by addressing the probabilistic nature of air pollution incidents. Subsequently, we provide a mathematical alternative to air emission penalty formulas used in Canada and USA through the concept of loss function and illustrate how this can be used not only as an enforcement instrument, but also as a policy tool that favours environmentally friendly decisions.

2. Probabilistic analysis of air pollution emission controls

Let us consider a waste incineration facility that was established some time ago under a certain regulatory framework that is due to be amended and become more stringent. The facility has to evaluate whether it can meet the new emission standards with its existing technology, or whether it needs to update some of its filters and/or the processes to guarantee compliance. At the simplest level the facility is facing two decisions:

$$\begin{aligned} &A_0\text{: } \textit{Continue business as usual and take no new measures} \\ &A_1\text{: } \textit{Install new emission systems and establish new processes} \end{aligned} \tag{1}$$

Of course decision A_0 exposes the facility to the benefit and gravity components of the penalty, which is designated by B for each day of violation. Let us define as Y the days of violation in a monitoring period of N days. Let us also define as C the daily cost of the new emission system that is considered by the facility, which includes its capital and operational cost.

The loss function [8] for the compliance of the facility to new emission standards can be given as follows:

$$L(A, Y) = \begin{cases} BY & if\ A = A_0 \\ CN & if\ A = A_1 \end{cases} \tag{2}$$

Deviation from the emission standards for each day j (j=1,2,…N) during the period N is described by the Bernoulli variable X:

$$X = \begin{cases} 1 & if\ there\ is\ violation \\ 0 & if\ there\ is\ compliance \end{cases} \tag{3}$$

where the probability not to comply with the emission standards any given day is P(X=1) = θ, and P(X=0) = 1- θ. The total number of days Y that violation of the emission standards takes place in a period of N days is given by

$$Y = \sum_{j=1}^{N} X(j), \tag{4}$$

and the random variable Y={y successes (violations) in N days} is described by a binomial probability distribution B(N, θ):

$$f_Y(y) = \binom{N}{y} \theta^y (1-\theta)^{N-y} \qquad where\ y = 0, 1, 2, \dots, N \tag{5a}$$

with:

$$\text{expected value } <\mathrm{Y}> = \mathrm{N}\theta \tag{5b}$$

$$\text{and variance } \mathrm{Var(Y)} = \mathrm{N}\theta(1-\theta). \tag{5c}$$

The binomial distribution assumes that Y is the sum of independent random Bernoulli variables. In the absence of perfect control systems one would like air emission violations to be random, of low probability of occurrence and environmental impact, and independent of each other. Thus, potential correlation of violations (for example, non-compliance during specific periods of operation or shifts) rather than being viewed as disqualifying the binomial probabilistic model can be seen as providing valuable information to investigate what distinct activities, which diverge from regular operations, took place during those periods, for which facility controls proved not to be efficient, and hence guide us to discontinue them.

The expected value of the loss function with respect to the random variable Y, the goal or risk function G(A, θ), is given for each action as:

$$G(A_0, \theta) = E_Y[L(A_0, Y)] = E_Y[BY] = B\ E_Y[Y] = BN\theta \tag{6a}$$

$$G(A_1, \theta) = E_Y[L(A_1, Y)] = E_Y[CN] = CN \tag{6b}$$

The two expressions indicate what would be the average loss depending on the decision (or action) undertaken. A standard cost-benefit analysis would then seek to find the decision that minimizes the goal function [9]. Figure 1 indicates that decision A_0 would be the optimal if the probability to violate emission standards any given day is below the ratio C/B, otherwise decision A_1 is preferable.

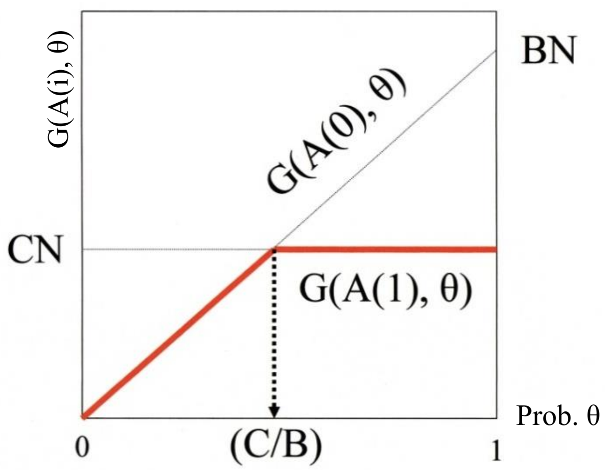

Figure 1. Regions of optimal goal function.

.

A common alternative to the binomial distribution is the Poisson model. The Poisson is the limit that the binomial distribution tends to when the number of samples N is very large, i.e., N → ∞, the probability of non-compliance θ is very small, i.e., $\theta \to 0$, but the product $N\theta = \lambda$ remains constant. The Poisson distribution and its first two moments are given by:

$$f_Y(y) = \frac{\lambda^y}{y!} e^{-\lambda} \qquad y = 0,1,2,\dots,\infty \tag{7a}$$

$$< Y > = \lambda \qquad and \qquad Var(Y) = \lambda \tag{7b}$$

The goal function $G_P(A, \theta)$ under the Poisson probabilistic model is then given by:

$$G_P(A_0, \theta) = E_Y[BY] = B\, E_Y[Y] = B\lambda = BN\theta \tag{8a}$$

$$G_P(A_1, \theta) = E_Y[CN] = CN, \tag{8b}$$

that is the goal function for the two distributions appear to be the same. However, one must keep in mind that for Poisson to apply the probability of non-compliance must be infinitesimally small and the number N of monitoring days must be very large [10, 11]. This point is stressed because as will be seen in the subsequent analysis the choice of the probabilistic model may influence the selection of the optimum decision.

We proceed now with a Bayesian analysis to estimate the probability density function (pdf) of the unknown probability of non-compliance θ. This can be done by using historical enforcement and compliance data, such as those provided by U.S. EPA [12], or in the absence of those by utilizing qualitative information about the occurrence of air pollution violations. Thus, an expert statement that non-compliance would occur, for example, on the average, once a month can be used to construct the subjective prior pdf of θ, π(θ).

Extreme value distributions are used to model peak-over-threshold (POT) measurements, i.e. values above a threshold level. Among those, the Generalized Pareto Distribution (GPD) has been seen to fit permit violations data, such as for total ammonia nitrogen concentration violations in treated wastewater [13, 14]. For reasons of simplicity the Beta distribution will be employed here, because it can take different shapes allowing it to simulate data from various sources. The Beta distribution is described as:

$$\mathrm{B}(\mathrm{t}, \mathrm{r}; \theta) = \frac{(\mathrm{t}-1)!}{(\mathrm{r}-1)!\,(\mathrm{t}-\mathrm{r}-1)!} \theta^{\mathrm{r}-1} (1-\theta)^{\mathrm{t}-\mathrm{r}-1} \tag{9a}$$

with mean and variance given by:

$$\mathrm{m}_\theta = \frac{\mathrm{r}}{\mathrm{t}} \qquad \text{and} \qquad \sigma_\theta^2 = \frac{\mathrm{r}(\mathrm{t}-\mathrm{r})}{\mathrm{t}^2(\mathrm{t}+1)} \tag{9b}$$

Applying the qualitative information "non-compliance, on the average, once a month" gives the values r=1 and t=30, which yields the prior

$$\pi(\theta) = 29\,(1-\theta)^{28} \tag{10a}$$

and

$$\mathrm{m}_\theta = \frac{1}{30} \qquad \text{and} \qquad \sigma_\theta^2 \approx \frac{1}{30^2} \tag{10b}$$

The Bayes risk R(A) is defined as the expected value of the goal function with respect to θ:

$$\mathrm{R}(\mathrm{A}) = \mathrm{E}^\pi[\mathrm{G}(\mathrm{A}, \theta)] = \int \mathrm{G}(\mathrm{A}, \theta)\, \pi(\theta) d\theta, \tag{11}$$

which for our case gives:

$$R(A) = \begin{cases} E^\pi[G(A_0, \theta)] = E^\pi[BN\theta] = \dfrac{1}{30} BN \\ E^\pi[G(A_1, \theta)] = E^\pi[CN] = CN \end{cases} \tag{12}$$

The two actions are equivalent when $R(A_0)=R(A_1)$, which occurs when B/30=C, that is when the penalty is thirty times the daily cost C of the new emission system. When $R(A_0)>R(A_1)$, that is when $B/30 > C$ then decision A_1 is the best. On the opposite, when $B/30 < C$ then decision A_0 should be preferred.

For a loss function having a linear form, as assumed in this section, using the Poisson distribution in the place of the binomial does not alter the calculations about the goal function as well as for the Bayes risk, and hence, leads to the same decision. Indeed the Bayes risk for linear in θ goal functions does not depend on the prior distribution, but only on its mean. This situation would change in the subsequent section when a parabolic loss function is employed to penalize more severely the cases of repeated non-compliance.

3. The loss function as an instrument for air emission compliance

The loss function that was used in equation (2) has a constant slope, and hence does not provide any incentive to decrease the rate of violations for a facility that has exceeded a certain threshold of air pollution incidents.

Figure 2 shows with $L_1(A(0),Y)$ the loss function that was used until now, and with $L_2(A(0),Y)$ one for which the penalty would double, equation (13), after a certain number of air pollution incidents are exceeded per year. Thus, an environmental agency based on historical data of air emission violation data can classify the number of violations in regions of variable severity, corresponding to different penalties, and even set a threshold, such as in equation (13), where the penalty M, after certain violations, becomes so large that a facility would find it beneficial to cease operations.

$$L_2(A(0),Y) = \begin{cases} BY & when\ 0 \le Y \le 2 \\ 2BY & when\ 2 < Y \le 5 \\ M & when\ Y \ge 6 \end{cases} \tag{13}$$

Hence, the shape and the slope of the loss function can be used by an agency as a policy instrument in order to penalize heavier after a fixed number of violations, deemed not to be consistent with the industry's record, has taken place. This way mathematical support can be provided to the gravity component of an agency's penalty and a facility would be able to decide based on an industry-wide comparison where its air emission performance stands.

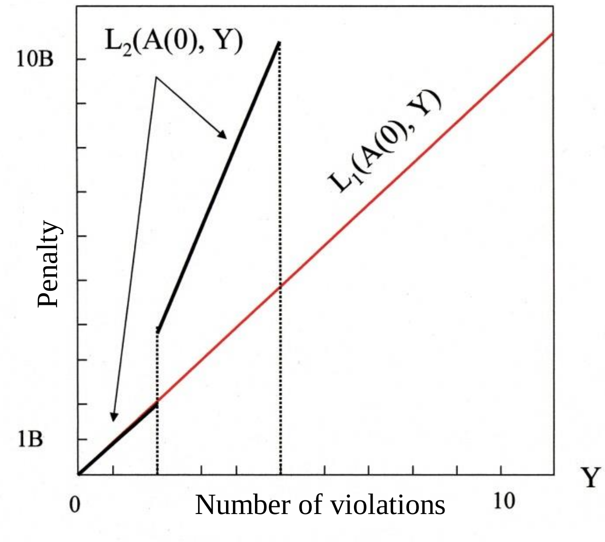

Figure 2. Loss functions with different slopes.

3.1. Effect of the choice of loss function on decision-making

We proceed now to investigate the effect that the shape of the loss functions has on the choice of the optimal action by using a parabolic function to describe the penalties:

$$L_3(A,Y) = \begin{cases} BY^2 & if\ A = A_0 \\ CN & if\ A = A_1 \end{cases} \tag{14}$$

The goal or risk function $G_3(A_0, \theta)$ now becomes:

$$G_3(A_0,\theta) = E_Y[L_3(A_0,Y)] = E_Y[BY^2] = B\,E_Y[Y^2] \tag{15a}$$

which based on the relation $E_Y[Y^2] = \mu^2 + \sigma_Y^2$ and the expressions for mean and variance for the binomial distribution, equations (5b) and (5c), gives:

$$G_3(A_0,\theta) = BN\theta\,[1+\theta(N-1)] \tag{15b}$$

It is for the parabolic loss function, which results in the term $E_Y[Y^2]$ in equation (15a) for the goal function that the results for the Poisson distribution diverge from those for the binomial. The goal function $G_3(A_1, \theta)$ remains the same:

$$G_3(A_1,\theta) = CN \tag{15c}$$

The Bayes risk for action A_0 now becomes:

$$R(A_0) = E^{\pi}[G(A_0,\theta)] = BE^{\pi}[N\theta + N(N-1)\theta^2] \approx$$
$$\approx B\left(\frac{N}{30} + N(N-1)\frac{2}{30^2}\right) = \frac{BN}{30}\left(1+\frac{N-1}{15}\right), \tag{16}$$

where in the second line the relation $E_Y[Y^2] = \mu^2 + \sigma_Y^2$ was used together with the mean and variance of the prior Beta distribution, equation (9b).

The Bayes risk for action A_1 remains the same:

$$R(A_1) = CN \tag{17}$$

The criterion for the best decision based on the Bayes risk depends on the ratio $R(A_0)/R(A_1)$. When this ratio is greater than one then the optimal decision is A_1; when the ratio is less than one the optimal decision is A_0. Using equations (16) and (17) ratio $R(A_0)/R(A_1)$ gives:

$$\frac{R(A_0)}{R(A_1)} = w\frac{B}{C} \qquad where \qquad w = \frac{1}{30}\left(1+\frac{N-1}{15}\right), \tag{18}$$

which shows that the ratio of air pollution penalty and daily cost of air filters depends on the number of samples N collected by the environmental inspection and monitoring agency. This, at first glance, unanticipated result will be further investigated in the next section.

The assumption of "one violation per month" based on an expert's opinion that allowed us to construct the prior pdf of θ can be construed to correspond to one violation out of a sample of N=30, which when entered in equation (18) yields for the parabolic loss function the criterion:

$$\frac{R(A_0)}{R(A_1)} = w\frac{B}{C} \qquad where \qquad w \approx \frac{1}{10} \tag{19a}$$

Equivalently for the linear loss function, equation (12), one obtains:

$$\frac{R(A_0)}{R(A_1)} = w\frac{B}{C} \qquad where \qquad w \approx \frac{1}{30} \tag{19b}$$

This means that if for example the daily cost of the new air emission system is $1,000 the use of a linear loss function will require a penalty of $30,000 to tilt the decision to A_1, i.e. to take action to consider a new technology, whereas a parabolic penalty function will require a much smaller penalty of $10,000 to promote a more environmentally friendly policy.

3.2. Effect of monitoring sample size on decision-making

Equation (18) gave an unexpected result that the decision to select A_0 or A_1 does not depend only on the penalty for non-compliance and the cost of a new emission system, but also on the number of samples taken. The significance of this result is explored now.

Let us consider that instead of "one violation per month" the assessment was that "two violations bi-monthly" occurred. This returns the same ratio of 1 to 30 as previously, and hence the same mean in equation (9b), but the variance now is different. The variance for r=2 and t=60 is given:

$$\sigma_\theta^2 = \frac{2\,(58)}{60^2(61)} \approx \frac{2}{60^2} = \frac{1}{2\,(30^2)} \tag{20}$$

that is the variance in θ is reduced by half compared to that in equation (10b), or in other words the uncertainty in estimating the probability of non-compliance θ is halved when the size of the sample doubles.

For this case the Bayes risk for action A_0 becomes:

$$R(A_0) = E^{\pi}[G(A_0,\theta)] = BE^{\pi}[N\theta + N(N-1)\theta^2] \approx$$
$$\approx B\left(\frac{N}{30} + N(N-1)\frac{3}{2(30^2)}\right) = \frac{BN}{30}\left(1 + \frac{N-1}{20}\right) \qquad (21)$$

and the ratio $R(A_0)/R(A_1)$ now becomes:

$$\frac{R(A_0)}{R(A_1)} = w\frac{B}{C} \qquad where \qquad w = \frac{1}{30}\left(1 + \frac{N-1}{20}\right), \qquad (22a)$$

which upon substitution of N=60 gives:

$$\frac{R(A_0)}{R(A_1)} = w\frac{B}{C} \qquad where \qquad w \approx \frac{1}{7.5} \qquad (22b)$$

Equation (22b) indicates that if as before the daily cost of the new air emission system remains fixed (say at $1,000) a smaller penalty of $7,500 will favour decision A_1 to install new air emission filters.

An assessment of 4 violations quarterly yields:

$$\frac{R(A_0)}{R(A_1)} = w\frac{B}{C} \qquad where \qquad w = \frac{1}{30}\left(1 + \frac{N-1}{24}\right), \qquad (23a)$$

and upon substitution of N=120 gives:

$$\frac{R(A_0)}{R(A_1)} = w\frac{B}{C} \qquad where \qquad w \approx \frac{1}{5} \qquad (23b)$$

Finally for 8 violations in N=240 one obtains:

$$\frac{R(A_0)}{R(A_1)} = w\frac{B}{C}, \;\; with = \frac{1}{30}\left(1 + \frac{3(N-1)}{80}\right),$$
$$and\ for\ N = 240 \qquad \frac{R(A_0)}{R(A_1)} = w\frac{B}{C}\ , where \;\; w \approx \frac{1}{3} \qquad (24)$$

Summarizing these results in a table gives:

Table 1. Influence of sample size on decision-making.

No. of violations per sample size	Variance	Weight w
1 out of 30	$\sigma=1/(30^2)$	1/10=0.10
2 out of 60	σ/2	1/7.5=0.13
4 out of 120	σ/4	1/5=0.20
8 out of 240	σ/8	1/3=0.33

Table 1 shows that as the sample size increases and the uncertainty around the mean of the probability to violate decreases, i.e., as the effectiveness of the filters becomes clearer, large penalties are not critical to convince a facility to upgrade its emission system. This makes sense if the probability is viewed as measure of our knowledge of a situation [10], i.e., the clearer it becomes that compliance may be problematic the easier a decision to upgrade environmental protection measures becomes.

4. Summary and Conclusions

The introduction of the pollution pays principle as the cornerstone of allocating the financial burden of pollution on the person or entity that generated it created the need for monetary penalties for noncompliance to environmental regulations. The U.S.A. and Canada have developed detailed penalty amounts that are calculated based on the harm done by a violation to the environment and to human health, the environmental history of the violator, the economic benefits as a result of noncompliance, and the wish of the regulator for the penalty to be a deterrent in order to abstain from certain activities.

Our current paper extends these concepts by addressing the probabilistic nature of air pollution. Penalties are placed in a Bayesian decision-making framework and the prior probability distribution of the pollution probability is constructed from historical data of industry-wide emissions, or in the absence of data, as done in this paper, from experts' opinion, and is used to define optimal decisions regarding emission controls. In addition, our article advances the use of the loss function as an alternative risk analysis tool to the current deterministic penalty regulatory formulas, which can also be used as a public policy instrument to promote environmentally, friendlier air emission choices.

5. **References**

[1] Organization for Economic Co-Operation and Development (OECD) 1992 *The Polluter Pays Principle* (Paris: OECD Analyses and Recommendations Environment Directorate OCDE/GD(92) 81) p 49

[2] Canada Gazette Government of Canada (June 2 2017) Environmental Violations Administrative Monetary Penalties Regulations http://www.gazette.gc.ca/rp-pr/p2/2017/2017-06-14/html/sor-dors109-eng.php (accessed November 29 2017)

[3] Lawrence B M 1992 EPA's civil penalty policies: Making the penalty fit the violation *Environmental Law Reporter* 22 ELR 10529 http://elr.info/sites/default/files/articles/22.10529.htm

[4] U.S. EPA Dec. 6, 2013 *Amendments to the U.S. Environmental Protection Agency's Civil Penalty to Account for Inflation* Memorandum 12; together with Environmental Protection Agency (Nov. 6, 2013. 40 CFR Part 19: *Civil Monetary Penalty Inflation Adjustment Rule* Federal Register **78(215)** 66643-48

[5] U.S. Code 42 USC § 7413 Jan. 4 2012 *Title 42-The Public Health and Welfare Chapter 85-Air Pollution and Control Subchapter I-Programs and Activities Part A-Air Quality and Emission Limitations § 7413* Federal Enforcement p 9

[6] Caiazzo F, Ashok A, Waitz I A, Yim S H L and Barrett, S R H 2013 Air pollution and early deaths in the United States. Part I: Quantifying the impact of major sectors in 2005 *Atmospheric Environment* **79** 198-208

[7] U.S. EPA Jan. 2002 *EPA Air Pollution Control Cost Manual* (Washington D.C.: Sixth Edition EPA/452/B-02-001) p 752

[8] Berger J O 1985 *Statistical Decision Theory and Bayesian Analysis* (New York: Springer-Verlag 2nd Edition Springer Series in Statistics) p 617

[9] Lerche I and E K Paleologos 2001 *Environmental Risk Analysis* (New York: McGraw-Hill Professional Engineering Series) p 436

[10] Jeffreys H Sir 1948 *Theory of Probability* International Series on Monograms in Physics (Oxford: Clarendon Press 2nd Edition) p 412

[11] Jaynes E T 2003 *Probability Theory: The Logic of Science* (Cambridge: Cambridge University Press) p 758

[12] U.S. EPA Enforcement and Compliance History Online (ECHO) *Facility Search-Enforcement and Compliance Data* https://echo.epa.gov/facilities/facility-search (accessed 20 November 2017)

[13] Weirich S R, Silverstein J and Rajagopalan B 2015 Resilience of secondary wastewater treatment plants: Prior performance is predictive of future process failure and recovery time *Environ Eng Sci* **32** 222-231

[14] Suchetana B, Rajagopalan B and Silverstein J 2017 Assessment of wastewater treatment facility compliance with decreasing ammonia discharge limits using a regression tree model *Sci Total Environ* **598** 249-257