Emission Permit Trading for Air Toxics Using Hazard Zones

WERNER ANTWEILER*
University of British Columbia

May 3, 2008

Abstract

Tradeable emission permits (and emission taxes) provide an efficient allocation of abatement activities across firms in the presence of informational asymmetry between firms and environmental regulator. While tradeable emission permits are well-suited for uniformly mixed pollutants such as greenhouse gases, theory suggests the use of tradeable ambient concentration contribution permits for non-uniformly mixed pollutants such as toxics. However, numerous practical obstacles impede the use of tradeable ambient concentration permits, most of all the potential illiquidity of the many small markets created by the necessary geographic segmentation. This paper develops an alternative trading system based on attainment and non-attainment zones and corresponding emission permit markets. Analytic solutions and simulation results demonstrate the practicability of the concept and its ability to approximate the first-best solution. The concept is applied to analyze emissions of metallic air toxics in Central Canada.

*Sauder School of Business, University of British Columbia, 2053 Main Mall, Vancouver, BC, V6T 1Z2, Canada. Phone: 604-822-8484. E-mail: werner.antweiler@ubc.ca. An earlier version of this paper was entitled 'Smart Emission Trading for Toxics: Spatial Heterogeneity and Hybrid Regimes.' This paper has benefited hugely from presentations at the University of British Columbia, the University of Kiel, the University of Cologne, York University, the University of Maryland, as well as the Canadian Economics Association conference in 2007. I am grateful to all the audiences and individuals who have provided feedback and comments.

1 Introduction

While emission permit trading systems have grown in use over the last number of years, their application potential remains somewhat limited because such trading systems work well only for uniformly mixed pollutants such as greenhouse gases, or at least widely dispersed pollutants such as sulfur dioxide. However, conventional emission permits for non-uniformly mixed pollutants such as toxics are inefficient. Since the work of Montgomery (1972) it is known that the theoretically ideal alternative employs "ambient concentration contribution permits" in which permits are contracted in fractions of emission concentrations; see Tietenberg (2006, chap. 4) for an extensive discussion. In the presence of strong spatial heterogeneity, defined by the location of emission points (plants) and receptor points (people), ambient concentration permit markets suffer from practical limitations. Most importantly, scuh a system requires a large number of markets to work efficiently. As Atkinson and Tietenberg (1987, p. 372) point out:

> An ambient permit market is complex, because it involves a separate permit for each monitored receptor where air quality equals the standard. Any source seeking to increase its emissions would have to purchase sufficient permits from *each* of these monitored receptors. Since the source's emissions would affect air quality differently in each of these markets, the interdependence of its decisions in these multiple markets creates a rather complex pattern of transactions.

The multiplicity of markets implies high transaction costs. It also fractures the trading volume across these markets. Liquidity in each of these markets could be reduced to such a point where the number of participants is simply too small to generate meaningful price signals. A further problem is the nature of the permit contracts. Unlike emission permits, ambient concentration contribution permits may be difficult to monitor and enforce as additional contributions to emission concentrations at a given receptor location may be impossible to track down to a particular source (i.e., permit violator). It is much easier to monitor emissions at the source.

In the presence of these challenges, finding suitable practicable alternatives to implement a market solution for non-uniformly mixed pollutants has been elusive. Many of the alternatives suggested in the literature appear to offer very complex solutions. The key to a practicable alternative is simplicity and objectivity. A trading system is simple if it keeps transaction costs low, and offers verifiable and enforceable contracts. It strives for objectivity if any firm-specific rules and targets are determined by measurable quantities that are relatively immune to challenge or lobbying. In an UNCTAD report Tietenberg et al. (1999, pp. 105-7) comment on design principles for a permit trading system:

> The emissions trading system should be designed to be as simple as possible. The historic evidence is very clear that simple emissions trading systems work much better than severely constrained ones. The

> transaction costs associated with implementing and administering an emissions trading system rise with the number of constraints imposed, and as transactions costs rise, the number of trades falls. As the number of trades falls, the cost savings achieved by the programme also decline. [...] Transaction costs play a key role in the success or failure of an emissions trading system. In the past, only emissions trading programmes with low transaction costs have succeeded in substantially lowering the cost of compliance.

This paper proposes a new practicable emission trading system for non-uniformly mixed pollutants. It involves identifying two hazard zones (a low-hazard attainment zone and a high-hazard non-attainment zone) and permit trading that is contracted in units of emissions rather than emission concentrations. The proposed solution differs significantly from other work that suggested the use of zones. Conventional zoning assigns each firm to a zone, and trading takes place only within each zone; there are no trades across zonal boundaries Tietenberg (1995). By comparison, the new zonal system proposed in this paper allows many firms to participate in both markets based on their relative location to the 'hot spots.' The proposed system, developed in section 3, appeals through it simplicity while retaining a good deal of the efficiency of the first-best solution.

Emission permit trading for toxics faces numerous challenges. Which toxics should be covered? Should the trading system cover only one or multiple pollutants? How does firm entry and exit affect the trading system? How can the trading system accommodate high variability in emissions over time? This paper will shed some light on these questions in section 4. Section 5 proceeds to illustrate the practicability of emissions trading for toxics by looking at twelve dangerous metallic air toxics in Central Canada (the provinces of Ontario and Quebec). The relative efficiency of the proposed hazard zone emission permit system is analyzed using numerical simulation techniques. Section 7 concludes.

2 The 'Hot Spot' Problem, Permit Market Multiplicity, and Transaction Costs

The key to understanding the problem with non-uniformly pollutants is the recognition that emission concentrations vary strongly across receptor locations. 'Hot spots' are localized areas with high concentrations of the pollutant. The efficacy of any policy intervention depends crucially on how it targets these hot spots. The more localized these hot spots, the more localized the intervention needs to be, as the largest environmental benefits are generated by reducing the ambient concentrations in these hot spots.

In the presence of information asymmetries between regulator and firms, emission permit trading systems (as well as pollution taxes) are demonstrably more efficient than conventional command-and-control regulation through emission standards. In the presence of spatial heterogeneity in emissions and non-uniformly mixed pollutants, ambient concentration permits are more efficient than simple emission permits. Nevertheless, the efficiency rank order of these regulatory interventions may be greatly affected by large differences in transaction costs.

The importance of transaction costs is well understood. Nevertheless, Stavins (1995) points out that even complicated trading systems with relatively high transaction costs will likely dominate command-and-control regulation because there is substantial heterogeneity in abatement cost (i.e., abatement technology) across plants. A trading system provides flexibility to plants in choosing the pollution abatement technology best suited for the particular environment.

Transaction costs for permit trading systems consist of a variety of specific costs, which in turn are driven by particular economic factors. Egenhofer (2003) discusses some of the main contributors to the overall transaction costs. *Search costs*, the cost of matching buyer and seller, are greatly reduced through organized exchanges, but crucially depend on the liquidity and transparency of the market. *Negotiation costs* arise from contracting and standardization of contracts through permit exchanges. These depend on the clarity of the property rights assigned by the contracts. *Monitoring costs* arise through verification of compliance, but are typically borne by the regulator. Similarly, *enforcement costs* arise in the case of non-compliance when the regulator needs to enforce compliance or fine violators. For individual firms there are also numerous types of *information costs*, for example the cost of monitoring permit markets, and the cost of involuntary sharing of private or confidential technical information. Firms may also incur *insurance costs* to allow for the technical risk of accidental non-compliance (e.g., unpredictable leaks, spills, or accidents). While some of these transaction costs are similar across different types of regulatory intervention, different types of emission permit markets differ most strongly with respect to search and information costs. Potential illiquidity of the market may reduce trading opportunities, and in the extreme case may turn such a market into bilateral bargaining.

Trading costs, combined with market and regulatory uncertainty, may reduce the efficiency of permit trading systems. Montero (1997) shows that these frictions can have a non-negligible impact on abatement outcomes: trade volume decreases, and total compliance cost increases. They also find that in the presence of transaction costs the initial allocation is not neutral with respect to efficiency. Gangadharan (2000) consider a variety of the aforementioned transaction costs in the context of the RECLAIM market in the Los Angeles basin. The presence of transaction costs influences the participation decision of plants. The author finds that various types of transaction costs can be identified as the reason for an overall 32% reduction in the probability of trading in the RECLAIM market in 1995.

Ultimately, the transaction cost problem boils down to a question about how many markets are needed. One one extreme is the efficient Montgomery (1972) solution with a large number of markets. On the other extreme is the conventional single cap-and-trade market. There is an implicit trade-off. Both transaction costs and environmental efficiency rise with the number of markets. Consequently, what is the most suitable number of markets? Going from one market to two markets facilitates large efficiency gains while keeping the increase in transaction costs manageable. The appeal of the hazard zones permit market presented in this paper is thus twofold: it approximates the efficient first-best solution quite well without making the trading system unncessarily complex or costly.

3 Model

3.1 Regulator's Objective

Consider the regulator's problem of controlling the ecosystem hazard emanating from emissions generated by plants $i = 1, .., I$. The regulator is concerned about the total ecosystem hazard H_j for regions $j = 1, .., J$. Ecosystem hazard includes, as a first approximation, the health risk of the population directly exposed to the toxics. However, the ecosystem hazard may also include the health risk to animals and plant, which in addition to their intrinsic value also matter because of their indirect effect on the health of the population through the food chain. Let P_j denote appropriate ecosystem weights for region j and define regions sufficiently small that emission concentrations can be thought of as homogeneous within that area. The ecosystem hazard is then defined as

$$H_j = P_j \left[E_j^{\square} + \sum_i d_{ij} E_i \right] \tag{1}$$

where E_i is the emission from plant i and d_{ij} is the dispersion factor between firm i and location j. $E_j^{\square}$ indicates ambient concentrations from natural or transboundary sources. While the medical literature typically stipulates non-linear (quadratic) dose-response functions as in Hoel and Portier (1994), environmental regulation targets ambient concentration ranges where the dose-response function can be approximated more suitably by a linear function.

The above hazard model (1) is marked by one critical deviation from the existing literature on ambient emission level control. The total ecosystem hazard is not merely a function of ambient concentrations, but also a function of the number of people, plants, and animals that are exposed to the concentration. This implies a trade-off between higher ambient concentrations in losely populated areas against lower concentrations in densely populated areas. Different regulators may therefore rely on different weights P_j. For example, setting P_j equal to one implies equal weights on emission concentrations. Choosing an appropriate P_j is a non-trivial problem that requires a thorough understanding about the potential trade-offs, as well as any non-linearities in the dose-response function. It may be useful to think of P_j not necessarily as population, but more generally as a 'penalty factor' linked to ecosystem hazard.

The conventional approach lets the regulator target a specific ambient concentration level (Atkinson and Tietenberg, 1987, p. 379) and minimizes abatement cost with respect to that threshold. However, meeting an absolute standard may be technologically infeasible or may incur unrealistically high abatement cost.[1] Involving penalty factors P_j allows for a much more flexible treatment of the problem by allowing regions to modestly exceed desired ambient concentrations. The model proposed here also makes use of a threshold, but this threshold is not defined in terms of an ambient concentration level, but instead in terms of an ecosystem hazard level. Furthermore, this threshold is not defined as a binding constraint, but as as a level separating high-hazard from low-hazard regions. Let ξ denote the hazard threshold that assigns each region j to one of two zones: the attainment (low hazard) zone or the non-attainment (high hazard) zone. The hazard zone membership indicator

$$r_j(\xi) = 1\left(H_j^0 > \xi\right) \tag{2}$$

can be defined using the binary indicator function $1(\cdot)$ and the initial (pre-abatement) hazard level H_j^0 using initial emission levels (E_i^0), defined below. To simplify notation later, let $E^0 \equiv \sum_i E_i^0$ and $H^0 \equiv \sum_j H_j^0$ denote total initial emssions and total initial hazard, respectively.

The regulator is also concerned about the (expected) abatement cost $B = \sum_i B_i$ incurred by firms in meeting the desired environmental footprint. The regulator's objective then is to minimize the expected 'welfare loss' L from ecosystem hazard H and abatement cost B

$$L = \mathcal{E}\left\{\sum_i B_i + \gamma \sum_j H_j\right\} \tag{3}$$

Here, γ is the marginal cost of a unit of ecosystem hazard. Alternatively, γ may also be thought of as a political preference parameter indicating the regulator's weight on ecosystem hazard.

The structure of the model is a two-stage game. In stage one the regulator chooses the hazard threshold ξ that seperates the low-hazard (attainment) zone from the high-hazard (non-attainment) zone, indicated by superscripts L and H, respectively. In order to minimize the expected loss L the regulator also chooses

[1] Ambient concentration standards remain the preferred approach for many regulators. For example, for several decades the United States Environmental Protection Agency (EPA) maintained an ambient standard of 1.5 μg/m^3 for airborne lead. Lead can cause kidney damage and other serious health problems. While airborne lead has decreased by 97% over the last three decades due to the removal of lead as an additive in gasoline, 'hot spots' of airborne lead still exist around iron and steel foundries, copper smelters, mining operations, waste incineratrs, and concrete plants (Wald, 2008). A new standard proposed in 2008, 0.1–0.3 μg/m^3, implies a reduction by a factor of 5 to 15. At the old level of the standard, only two U.S. counties were out of compliance, whereas at the new level, probably over 20 counties would fail to comply. With this proposal the EPA follows a court order resulting from a lawsuit brought by environmentalists. This example illustrates that emission standards tend to be set cautiously; they tend to the close to the non-binding range due to the possibility that some plants may be technically or economically unable to commply, and the only remaining options is plant closure. An emission trading system allows regulators to set more ambitious ambient concentration targets while allowing non-compliant plants to catch up over time by investing in existing abatement technology, or developing new technolgy if necessary.

allocations A^H and A^L of emission permits. In stage two the plants choose their optimal abatement effort in response to the permit prices τ^H and τ^L. The game is solved through backward induction.

3.2 The Plant

Each plant's emissions affect the neighboring regions j through the dispersion factor d_{ij} so that $\sum_j d_{ij} \leq 1$. These dispersion factors are determined by measurement or modeling of atmospheric disperson processes. They need not add up to one if some of the emissions are transboundary.

Depending on whether a plant's neighboring regions are high hazard zones (hot spots) or low hazard zones, each plant will be required to buy a share $0 \leq s_i \leq 1$ of non-attainment zone permits and a share $1 - s_i$ of attainment zone permits for its emissions E_i. Prices in these two markets are τ^H for the high hazard zone and $\tau_L < \tau^H$ for the low hazard zone. This means that each firm faces an effective permit price

$$\tau_i = s_i \tau^H + (1 - s_i)\tau^L \tag{4}$$

that is a linear combination of the two prices with

$$s_i = \sum_j d_{ij} r_j(\xi) \tag{5}$$

Plant i's emissions are a function of its abatement effort $\theta_i \in [0, 1[$ so that emissions E_i are

$$E_i = (1 - \theta_i) z_i q_i = (1 - \theta_i) E_i^0 \tag{6}$$

with original unabated emission intensity z_i and output (scale) q_i fixed. Let E_i^0 denote total unabated emissions of plant i. Abatement costs are

$$B_i = b_i[-\ln(1 - \theta_i)]E_i^0 \tag{7}$$

with cost factor b_i. The crucial information asymmetry between regulator and firm is that the regulator cannot observe b_i; the regulator can only conjecture $\hat{b} = \mathcal{E}\{b_i\}$. The abatement cost function has the property that as $\theta \to 1$ (complete abatement), abatement costs rise to infinity. This implies that plants are unable to completely abate emissions. Furthermore, $B'(\theta) > 0$ and $B''(\theta) > 0$ imply that abatement costs are increasing in abatement effort at an accelerating rate: abatement effort becomes increasingly costly. At low levels of effort, the cost function is almost linear, but as the effort level approaches total elimination of emissions, abatement cost starts to rise steeply. Abatement costs rise proportional with output, and there are no increasing or decreasing returns to scale in abatement activity. Choosing a specific functional form for abatement costs generates a more tractable analytic solution while retaining the underlying economic intuition. To simplify the analysis further, the abatement decision is independent of the output decision (and thus market structure and plant entry/exit effects). This is not necessarily an unrealistic assumption when abatement costs are sufficiently small compared to overall costs. At later points in the analysis it will be useful to introduce $B_i^0 \equiv b_i z_i q_i$ as the 'reference' abatement cost, which according to (7) implies $\theta = (e-1)/e = 0.632$; this will merely simplify notation.

The firm's objective is to minimize its environment-related costs

$$C_i = \min_{\theta_i} \{B_i + \tau_i E_i\} \tag{8}$$

by choosing the optimal abatement effort θ_i in response to the effective price τ_i for buying permits for its (post-abatement) emissions E_i. The size of permit allocations A_i^H and A_i^L, and the method (auction or grandfathering) through which they are received, does not enter the decision problem for the optimal abatement effort. Using (6) and (7) and solving the minimization problem (8) yields the optimal abatement effort $\theta_i^* = 1 - b_i/\tau_i$, which in turn implies $B_i^* = B_i^0 \ln(\tau_i/b_i)$ and $E_i^* = B_i^0/\tau_i$. Here it is assumed that the permit price is sufficiently high ($\tau_i > b_i$) that it induces firms to engage in abatement effort. If the permit price is too low, or the firm's available abatement technology too costly, the firm will rather buy permits than engage in any abatement effort. If the distribution of b_i is wide, non-abatement cases where $b_i \geq \tau_i$ may account for a significant share of firms. In simulations it is important to allow for these cases. With the above results the regulator's objective function becomes

$$L = \mathcal{E}\left\{\sum_i B_i^0 \ln\left(\frac{\tau_i}{b_i}\right) + \gamma \sum_j P_j \left[E_j^{\square} + \sum_i d_{ij} \frac{B_i^0}{\tau_i}\right]\right\} \tag{9}$$

3.3 Optimal Interventions

Differentiating (9) with respect to τ_i yields the first-order condition for an optimal plant-specific intervention. The optimal plant-specific emission tax or permit price turns out to be

$$\tau_i = \gamma \sum_j P_j d_{ij} \equiv \gamma P_i \tag{10}$$

where P_i is introduced as the plant-specific penalty factor, defined as the average of the region-specific penalty factors weighted by each plant's dispersion factors. This optimal plant-specific tax does not depend on perfect knowledge of the abatement ability b_i. By comparison, the optimal emission standard $\bar{E}_i = b_i E_i^0/\gamma P_i$ depends on perfect knowledge of b_i. This replicates the standard result that emission standards are inefficient compared to emission taxes and emission trading in the presence of information asymmetry. Using optimal intervention, the total loss is

$$L^* = \sum_i B_i^0 \left[1 + \ln\left(\frac{\gamma P_i}{b_i}\right)\right] \tag{11}$$

which will serve as a useful benchmark to compare other solutions against. The optimal solution also replicates the Montgomery (1972) ambient concentration permit market. The allocations for this market are $A_j = E_j = \sum_i d_{ij} B_i^0/\tau_i$ where the effective plant-specific permit price is a weighted average of the regional permit prices τ_j so that $\tau_i = \sum_j d_{ij}\tau_j$. The regulator then chooses optimal emission concentration permit allocations

$$A_j = \sum_i \frac{d_{ij} B_i^0}{\gamma P_i} \tag{12}$$

The optimal emission tax (10) is an instrument that exactly replicates the first-best results from an ambient concentration permit market as defined in (12). If the optimal plant-specific emission tax was indeed a practicable instrument, this would be the point to stop and adopt this policy. Whereas the problem with an ambient concentration permit market is transaction costs, the problem with a plant-specific emission tax is the way emission concentrations are translated into hazard potentials by way of assigning region-specific penalty factors P_j. With plant-specific emission taxes these factors would become the target of intense lobbying, thus creating inefficiencies of a different type. In the hazard zone model proposed below the penalty factors effectively hide behind the definition of the hazard zone boundaries, which are perhaps easier to defend against lobbying. The key to the solution of the hot-spot problem is indeed assigning different prices to the emissions from different plants, depending on their location. How can this be achieved with an emission permit system where, necessarily, there can be only one price for a given type of permit?

3.4 Conventional Emission Permit Trading

A useful benchmark case for comparison purposes is the conventional emission permit trading (cap-and-trade) regime with only one market in which firms receive an initial allocation A_i so that total allocation A equals total emission E:

$$A \equiv \sum_i A_i = \sum_i E_i \equiv E \tag{13}$$

Noting that with equilibrium price τ total emissions are $E = \sum_i E_i = \sum_i B_i^0/\tau = B^0/\tau$, it follows that $\tau = B^0/A$ and thus $\mathcal{E}\{\tau/b_i\} = E^0/A$.[2] Hence, the regulator's loss function (9) simplifies to

$$L^C = \hat{b}E^0 \ln(E^0/A) + \gamma H^0 A/E^0 \tag{14}$$

The regulator chooses A to minimize L and thus

$$A = E^0 \left[\frac{\hat{b}E^0}{\gamma H^0}\right] \tag{15}$$

[2]This assumes that all firms engage in some abatement. If there are some plants i' that abate and some plants i'' that merely buy permits but do not abate because $b_i > \tau$, the permit price will increase because $\tau = (\sum_{i'} B_i^0)/(A - \sum_{i''} E_i^0)$. Which plants fall into which category in turn depends on τ, and thus it is not possible to provide a closed-form solution in that case.

To reduce emissions, it must hold that $\hat{b}E^0 < \gamma H^0$: either the regulator's 'green' preference γ must be sufficiently large, or the expected abatement cost $\hat{b}$ must be sufficiently small.

Notwithstanding their relative inefficiency in the presence of emission concentration hot spots, simple cap-and-trade systems remain popular with regulators, in particular because of their perceived simplicity. Notably, under its *Clean Air Merucy Rule* the U.S. Environmental Protection Agency is establishing a cap-and-trade market for airborne mercury, due to be launched in 2010.

3.5 Hazard Zones Solution

The emission permit allocations for the two hazard zones are A^H and A^L, and as allocations must equal emissions in equilibrium,

$$A^H = E^H = \sum_i s_i E_i = \sum_i s_i B_i^0/\tau_i \tag{16}$$

$$A^L = E^L = \sum_i (1-s_i) E_i = \sum_i (1-s_i) B_i^0/\tau_i \tag{17}$$

The market value of both permit markets follows by multiplying (16) and (17) by τ^H and τ^L, respectively, and adding up using (4):

$$V \equiv \tau^H A^H + \tau^L A^L = \sum_i B_i^0 = B^0 \tag{18}$$

Adding up (16) and (17) also yields

$$A^H + A^L = E^H + E^L = \sum_i B_i^0/\tau_i \equiv B^1(\tau^H, \tau^L) \tag{19}$$

and therefore

$$A^H = \frac{B^0 - \tau^L B^1(\tau^H,\tau^L)}{\tau^H - \tau^L} \quad \text{and} \quad A^L = \frac{\tau^H B^1(\tau^H,\tau^L) - B^0}{\tau^H - \tau^L} \tag{20}$$

provide the allocations of permits for the high-hazard and low-hazard zones.

The first-order conditions for minimizing the loss function (9), without taking expectations first, are:

$$\frac{\partial L}{\partial \tau^H} = \sum_i B_i^0 \left[\frac{s_i}{\tau_i} - \gamma P_i \frac{s_i}{\tau_i^2}\right] = 0 \tag{21}$$

$$\frac{\partial L}{\partial \tau^L} = \sum_i B_i^0 \left[\frac{1-s_i\tau^L}{\tau_i} - \gamma P_i \frac{1-s_i}{\tau_i^2}\right] = 0 \tag{22}$$

Adding up (21) and (22) yields (23) below. Multiplying (21) and (22) by τ^H and τ^L, respectively, and summing up yields (24).

$$\sum_i B_i^0/\tau_i = \gamma \sum_i P_i B_i^0/\tau_i^2 \tag{23}$$

$$\sum_i B_i^0 = \gamma \sum_i P_i B_i^0/\tau_i \tag{24}$$

As τ_i is a function of τ^H and τ^H as defined in (4), equations (23) and (24) define a system of two equations in two unknowns. Not knowing the true b_i, the regulator uses $\hat{b}$, the expected average abatement ability, to calculate expected prices $\hat{\tau}^H$ and $\hat{\tau}^L$in order to determine the permit allocations A^H and A^L in (20), using predicted values of B^0 and B^1. Once permit allocations are known, the actual permit prices τ^H and τ^L are determined by the permit market and obey (20). Because the regulator uses imprecise allocations A^H and A^L, the effective (partial information) prices may be different from the optimal (full information) prices that can be obtained by solving (23) and (24) without taking expectations. For the purpose of decomposing the losses it is useful to calculate the total loss in (9) with both sets of prices, yielding $\hat{L}^Z$ for the effective and L^Z for the optimal hazard zone solution The difference between $\hat{L}^Z$ and L^Z can be attributed to the regulator's information asymmetry (moving from full information to partial information), and the difference between L^Z and L^* can be attributed to the 'pure' effect of moving from the optimal intervention to the hazard zone solution (both will full information).

Before the regulator can set allocations A^H and A^L, it is necessary to define the hazard threshold ξ that defines each firm's s_i shares. Because the loss function is not continuous in ξ, the corresponding minimization problem can only be solved numerically. Let ξ^* and $\hat{\xi}^*$ denote the hazard thresholds that minimize $L^Z(\xi)$ and $\hat{L}^Z(\xi)$, respectively.

3.6 Constrained Hazard Zones Solution

As simplicity is a key objective designing a pollution permit market, it may be useful to consider the permit system where the attainment market price is constrained to be zero ($\tau^L = 0$) so that permits are issued only for the non-attainment market. In this case the effective permit price for each plant becomes $\tau_i = s_i \tau^H$. As the allocation A^H must equal E^H in equilibrium, it holds that $A^H = B^0/\tau^H$. The single first-order condition for minimizing the loss function (9) implies that the optimal permit price is

$$\tau^H = \gamma \sum_j P_j \sum_{[i]} d_{ij} \left[\frac{B_i^0/B^0}{s_i} \right] \tag{25}$$

where the summation over i includes only the cases where $s_i > 0$. The notation $[i]$ for the summation subscript indicates this conditional summation.

Comparing solution (25) with the optimal intervention (10) reveals that the key difference is the adjustment factor $(B_i^0/B^0)/s_i$, the ratio of plant i's reference abatement cost share relative to its hazard zone contribution share. What drives up the permit price are large plants with small hazard zone contributions. The economic intuition is that in order to achieve the same amount of hazard reduction, large plants on the periphery of the non-attainment zone have to be induced more aggressively (with a higher price) compared to a small plant within the non-attainment zone.

Using the expected values of τ^H and B^0, the regulator allocates

$$A^H = \frac{\hat{b}E^0}{\gamma \sum_j P_j \sum_{[i]} d_{ij}(E_i^0/E^0)/s_i} \tag{26}$$

permits, and in turn the permit market price emerges as $\tau^H = (\sum_{[i]} B_i^0)/A^H$.

There may be conditions under which a constrained market is optimal. Allowing for complementary slackness in the first-order condition (22) so that $\partial L/\partial \tau^L > 0$ and $\tau^L = 0$ for minimizing the loss L implies that $\forall [i] : \tau_i > \gamma P_i$, and therefore

$$\tau^H > \gamma \max_{[i]} \left\{ \frac{P_i}{s_i} \right\} = \gamma \max_{[i]} \left\{ \frac{\sum_j d_{ij} P_j}{\sum_j d_{ij} r_j(\xi)} \right\} \tag{27}$$

When the permit price turns out to be sufficiently high relative to the highest penalty ratio P_i/s_i for any plant, or because the regulator puts a large preference γ on hazard avoidance, a separate attainment market with $\tau^L > 0$ becomes unnecessary.

What makes a constrained hazard zone model potentially appealing to the regulator is the implicit trade-off. On one hand, eliminating the attainment market reduces the transaction cost by reducing the number of firms participating in the permit system. This may also reduce political resistance to introducing a new permit system. On the other hand, the efficiency loss from eliminating the attainment market may be small as the non-attainment market, which targets the hazard hot spots, contributes much larger environmental gains.

It may also be noted that when the hazard threshold ξ is lowered to zero, the constrained hazard zone solution becomes increasingly similar to the simple cap-and-trade solution. However, even when $\xi = 0$ the markets will still differ. Some regions j are dropped from the analysis if their penalty factor P_j is zero (e.g., extra-jurisdictional territory, water surfaces). Whereas the simple cap-and-trade regime includes all emissions, the hazard zone system only incudes emissions that contribute to emission concentrations in regions j with $P_j > 0$. The simple cap-and-trade solution is equivalent to $\tau^H = 0$, $\tau^L > 0$, and $\xi \to \infty$ so that all $s_i = 0$.

4 Practical Considerations

4.1 Determining Dispersion Factors

Precise knowledge of dispersion factors d_{ij} is essential for any regulatory intervention that involves non-uniformly mixed pollutants, regardless of the type of intervention. While pollutant concentrations can be measured at receptor locations, it requires an atmospheric dispersion model to link emissions to emission concentrations (Turner, 1994). Atmospheric dispersion is influenced by meteorological conditions (in particular wind speed and direction), pollutant source characteristics (in particular stack height and exit velocity), and regional geographic conditions (terrain). The basis for most models is a Gaussian dispersion equation where ambient concentrations are proportional to the product of flow rate, cross-wind horizontal dispersion factor, and vertical dispersion factor. Concretely, the horizontal dispersion factor follows a normal distribution density function with mean zero and standard deviation σ.

The U.S. Environmental Protection Agency (EPA) employs a number of more sophisticated atmospheric dispersion models. In particular, the EPA uses the Community Multiscale Air Quality (CMAQ) model to simulate the dispersion of air toxics. The CMAQ model has been widely adopted by other environmental agencies, including Canada's province of Ontario. The main application of CMAQ is to track multiple pollutants simultaneously across regions by projecting air plumes from multiple point sources on a spatial grid.

4.2 Multiple Pollutants

Environmental regulators have identified classes of toxic substances. For example, under the United States Clean Air Act of 1990, 188 substances were identified as hazardous air pollutants (HAPs) with the potential for causing adverse health effects, such as cancer, reproductive and neurological effects, immune system damage and birth defects. This list also includes the metallic air toxics that are a specific target in this paper.

An emission permit market has the potential to target multiple pollutants rather than just one. For example, greenhouse gas permit markets target carbon dioxide as well as methane and chlorofluorocarbons. Similarly, a metallic toxics market would include a list of substances with similar characteristics. Pollutants can be both complements and substitutes in production. Some pollutants are joint outputs, while others can be substitutes for each other by changing the production process. Regulating just one pollutant may cause undesirable substitution effects by shifting production to methods that use the unregulated pollutant. An emission permit system can be designed easily to cover multiple pollutants by assigning fixed toxicity factors to each pollutant by using measures such as the US-EPA's (2004) Risk-Screening Environmental Indicators (RSEI) of Chronic Human Health.

Another important reason to consider a multiple pollutant permit market is that many toxics are only emitted by a small number of plants, and thus single pollutant permit markets would face relatively low liquidity. By covering multiple pollutants, more plants and firms will participate in the market, thus increasing liquidity and efficiency of the market. Different pollutants may disperse differently over distance. Thus a practical limitation for grouping toxics is their similarity with respect to dispersion factors d_{ij}.

4.3 Bioaccumulation and Biomagnification

Bioaccumulation is the process by which chemical compounds accumulate or build up in an organism at a rate faster than they can be broken down. More concisely, environmental scientists draw a distinction between bioaccumulation and biomagnification. Bioaccumulation occurs within a trophic (food chain) level in an ecosystem and captures the increase in concentration of a substance in an individual's tissues due to ambient exposure to the substance. Biomagnification occurs across trophic levels and captures the increase in concentration of a substance when the substance moves from one trophic level to the next. For biomagnification to occur, a toxic must be long-lived (i.e., does not decay rapidly), mobile (e.g., airborne), soluble in fats, and biologically active (i.e., has an effect on the metabolic activity of living cells).

Many toxics have the tendency to bioaccumulate and biomagnify. For example, if emissions from a smelter in a rural and sparsely populated area expose nearby large herds of milk cows to arsenic, that arsenic may bioaccumulate in these cows and will biomagnify when consumed by city dwellers far away from the smelter. Hence, it is crucial that the ecosystem penalty factors P_j for each region j take into account that the health hazard may be transferred elsewhere. Just as atmospheric dispersion models project emissions from point sources to emission concentrations at receptor locations, perhaps it is necessary to model how emission concentrations at receptor locations are integrated into bioaccumulated and biomagnified health risk for population living elsewhere. Spatial models that take these types of second-order effects into account have not been developed yet, however.

4.4 Transboundary Pollution

A further complication arises from the problem that pollutants travel easily across national boundaries. Yap et al. (2005), a study commissioned by the Ontario Ministry of the Environment, concluded that of the estimated $9.6 billion in health and environmental damages from ground-level ozone and fine particulate matter in Ontario in 2003, 55 per cent is attributable to U.S. emissions. However, there are noticeable differences across the province of Ontario, with south-western Ontario experiencing a large proportion of damages attributable to U.S. emissions, and south-central Ontario (including the Greater Toronto Area) experiencing the greatest impact of air pollution originating in Ontario.

Transboundary pollution, which in the model is captured through $E_j^{\square}$, is a spatial fixed effect that does not impact the design of the permit market. The term drops out in the first-order conditions of the model. It only matters insofar as the regulator will set more aggressive emission reduction goals in order to compensate for the transboundary pollution. Local firms will bear the burdern of abatement that more efficiently could be borne by firms across the border.

4.5 Initial Allocation: Grandfathering or Auctions?

The method of allocating permits does not affect efficiency, except when transaction costs are significant and marginal transaction costs are decreasing in the volume of trades. However, whether permits are grandfathered or auctioned has important distributional effects that also touch upon the political feasibility of the permit system. Grandfathering, which allocates permits to existing polluters based on some formula, is often considered politically less onerous than an auction, which generates government revenue and looks more like taxation. On the other hand, grandfathering tends to lead to larger trading volume than auctions because the regulator's formula for allocating permits will tend to be biased. Consequently, allocating permits through auctions implies lower transaction costs. With a hazard zone emission permit system there can be one or two markets. In the latter case, it is possible to mix allocation methods, for example grandfathering for the attainment permits and auctioning for the non-attainment permits.

4.6 Emission Variability and Intertemporal Transfers Through Permit Banking and Borrowing

A stylized fact about toxic emissions is the high variability of releases across years relative to fundamental indicators of operational size, such as plant employment or value added. The evidence is that emissions exhibit great variability over time, which implies that emissions are not only generated through regular production activity, but are also caused by spills, leaks, and other unforeseen events or accidents. This variability of toxics releases necessitates a closer look at the intertemporal dimension of emissions trading.

A significant advantage of a tradeable permit system that targets emission levels rather than ambient emission concentrations is the ability to allow banking of permits (for use in future years), borrowing of permits (for current use), or both, across time periods. A permit system that is defined in terms of ambient concentrations cannot easily accommodate banking or borrowing because of potential temporal clustering of emissions. Temporal clustering would intensify 'hot spots' of very high peak concentrations, posing a significant health risk. The economic rationale for banking and borrowing is compelling: greater flexibility across time provides for a better allocation of abatement investments over time. Allowing consolidation of permits over a given time horizon, say two or three years, may contribute to intertemporal efficiency in the presence of high emission variability across years.

Even without banking, intertemporal efficiency can be enhanced by staggering the compliance schedule. As Gangadharan (2000, p. 603) reports for the RECLAIM permit market in Los Angeles, plants were randomly assigned to two compliance cycles (January 1 through December 31 and July 1 through June 30) and are permitted to trade with each other. This overlapping two-cycle system is meant to reduce the likelihood of permit shortages or surpluses at the end of a compliance cycle.

Permit banking may also become attractive when plants anticipate reductions in the number of permits issued in future periods. The economics of permit banking has been explored extensively in Rubin (1996), Cronshaw and Kruse (1996) and Schennach (2000). Through emission banking and borrowing firms can shift their emission stream through time. Banking is generally desirable when standards are becoming stricter over time. However, when standards are constant or easing over time, emission borrowing will increase environmental hazard while lowering firms' costs. Typically, regulators may want to limit the time horizon against which permits can be borrowed. Even though banking is generally desirable, Liski and Montero (2005, 2006) show that market power may distort the outcome, even though these market power effects are mitigated by pro-competitive effects from forward markets for permits. Encouragingly, they find that manipulating a permit banking program appears to be difficult even for large permit stockholders.

4.7 Plant Entry, Plant Exit, and Plant Relocation

The entry and exit of plants poses a slightly greater challenge to a hazard zone permit system than a conventional cap-and-trade permit system because the hazard threshold ξ that separates high-hazard from low-hazard regions depends on the activity of neighbouring plants. The closer plants are located to each other, the more they care about each others' actions. Entry of a new plant creates a negative externality on surrounding plants as the new plant will compete for emission permits and thus increase the existing plant's abatement costs. Plant entry may also amplify the 'hot spot' problem, prompting the regulator to adjust permit allocations and/or zonal boundaries in subsequent periods. Consequently, plants face an incentive to locate away from each other. This beneficial environmental effect may be offset, however, if firms face significant agglomeration benefits.

Whereas plant entry creates a negative externality for neighbouring plants, plant exit creates a corresponding positive externality. At the margin, due to the crowding effect that is priced through the permit trading system, a plant may find it less costly to pay another plant to either exit or relocate to a less crowded region rather than abate emissions. Such induced exits may come about through company mergers and plant consolidation. Induced relocation, on the other hand, is the equivalent of the higher smoke stack policy of the 1960s: spread the hazard more equally rather than reduce emissions.

5 Metallic Air Toxics in Central Canada

Metallic air toxics account for some of the most significant airborne health hazards. These substances are considered developmental and reproductive toxicants. Using the US-EPA (2004) toxicity scores, table 1 shows the total cumulative emissions of these air toxics for the five-year period 2001-2005 across all plants in Canada. As is shown in the table, Central Canada accounts for large fractions of the point sources. Because Ontario and Quebec (which constitute the region of Central Canada) contain Canada's most populated areas along a west-east corridor, it is useful to concentrate on this geographically well-defined area. Emissions are highly concentrated. The top four polluters account for just over half of the entire toxic load, the top 34 polluters account for 90% of the toxic load, the top 59 polluters account for 95% of the toxic load, and the top 177 polluters account for 99% of the toxic load.[3]

Table 1: Metallic Toxics Emissions in Canada, 2001-2005

Toxics*	Emission [tons]	Toxicity Score	Toxic Load	
			Canada	ON+QC
Nickel	2,072.4	36,000	74,607	39.7 %
Manganese	1,211.7	36,000	43,622	52.3 %
Chromium	292.1	86,000	25,121	91.7 %
Arsenic	736.3	31,000	22,825	82.6 %
Cadmium	198.0	90,000	17,821	31.2 %
Lead	1,851.4	8,800	16,292	64.7 %
Cobalt	53.2	90,000	4,792	58.9 %
Copper	2,070.5	750	1,553	72.6 %
Selenium	272.6	3,600	981	100.0 %
Antimony	60.7	9,000	546	94.1 %
Zinc	4,574.0	51	233	41.3 %
Mercury	28.4	6,000	170	33.4 %
Total	13,421.3		208,564	55.6 %

* and their compounds

There are over 800 plants in Central Canada that emit one or more of the twelve toxics into the air; just over 400 of them have emissions with a toxic load greater than one. Included in this category are releases through the stack, from storage or handling, from fugitive sources such as leaks from valves, seals and connections, spills and other non-point air releases. Table 2 shows the composition of emissions from these plants by industry and pollutant. Metal ore mining accounts for the bulk of arsenic, cadmium, cobalt, copper, lead, nickel, and selenium emissions. Electricity generation is contributing the largest share of mercury emissions. Foundries are responsible for large shares of antimony, arsenicm chromium, and lead emissions. A look at these twelve metallic air toxics will illustrate the geographic dimension of the proposed emissions trading system.

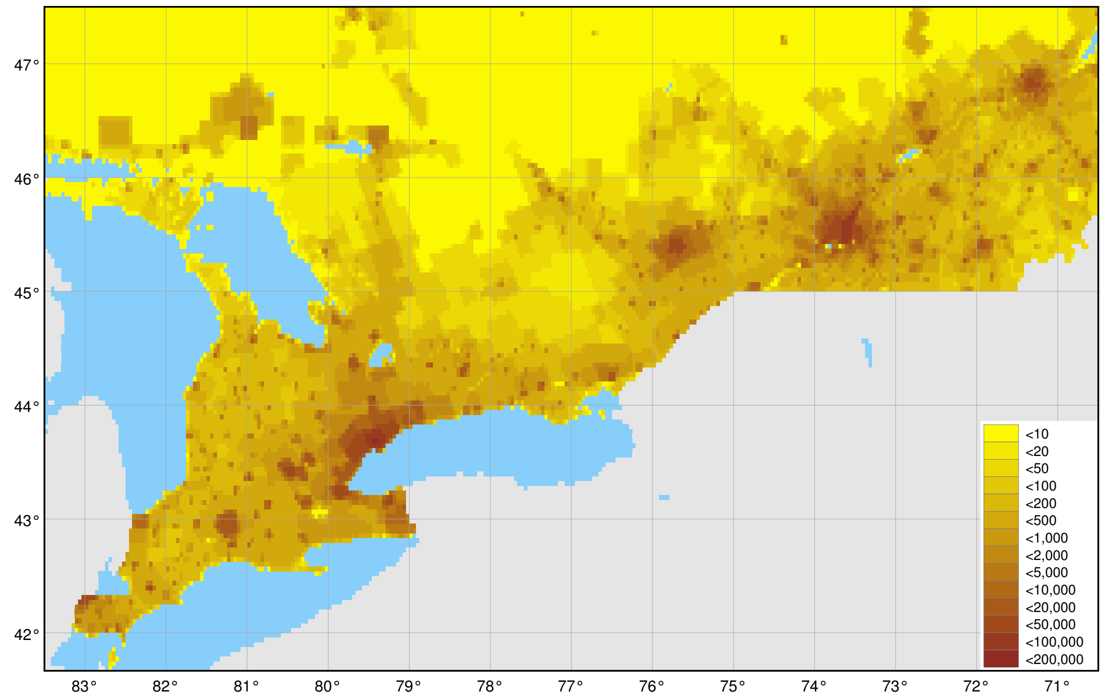
Key to modeling the effect of emissions from points sources on receptors at a distance is an understanding of the physical processes that determine dispersion and absorption. In practical terms it requires that plant-specific emissions are translated into surrogate emission concentrations at reception points. Dispersion is the process by which emissions spread spatially in terms of direction, distance, and density. Dispersion of air pollutants is influenced primarily by atmospheric conditions and topography. Absorption is the process by which surrogate emission concentrations at a given location are translated into exposure hazard. Absorption is influenced by factors such as natural decay and vegetation.

[3] The single largest emitter is Vale-Inco's Copper Cliff smelter near Sudbury, accounting for nearly 25% of the total toxic load in Central Canada. 'PollutionWatch,' a web site maintained by *Environmental Defence* and the *Canadian Environmental Law Association*, ranked the Copper Cliff smelter first or second across Canada for air releases of various respiratory toxicants.

Table 2: Composition of Emissions by Industry and Toxic Load (Percentage Shares of Emissions in Central Canada)

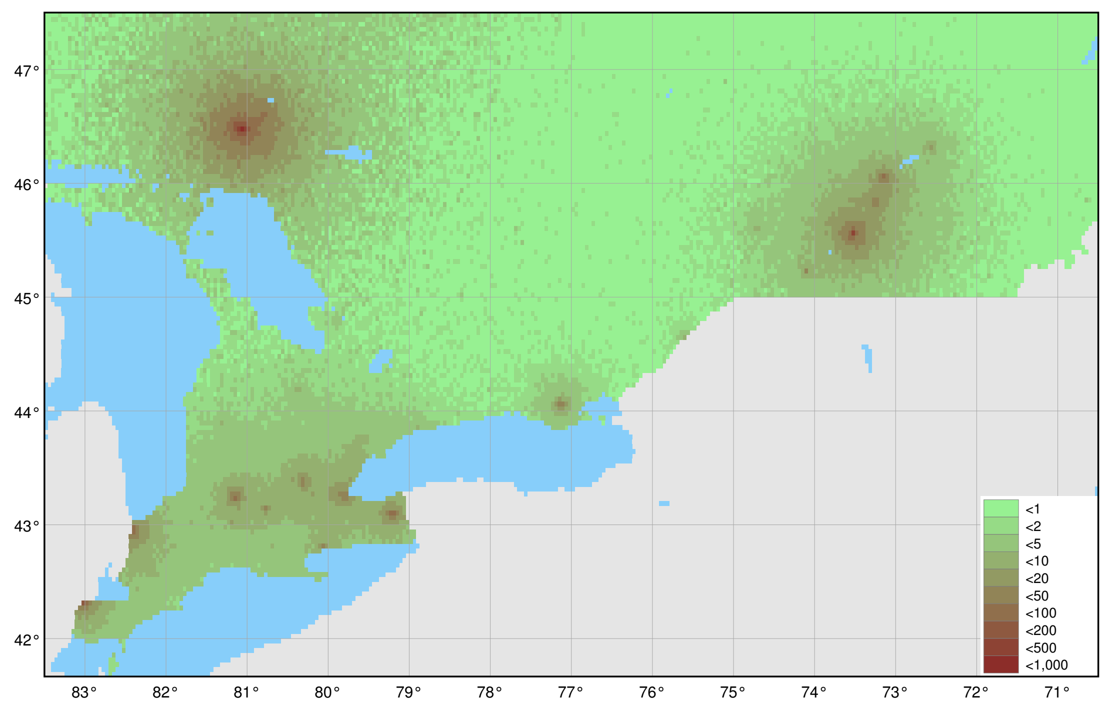
Industry	Antimony	Arsenic	Cadmium	Chromium	Cobalt	Copper	Lead	Manganese	Mercury	Nickel	Selenium	Zinc
Metal Ore Mining	0.5	45.8	61.4	8.9	43.3	49.1	25.1	2.3	2.8	60.3	42.7	13.4
Electricity Generation, Transmission & Dist.		0.7	0.7	1.8	6.7	0.7	0.4	1.2	25.7	0.8	17.2	1.5
Water, Sewage & Other Systems		<0.1	0.2	<0.1		<0.1	<0.1	<0.1	9.9			<0.1
Textile & Fabric Finishing & Fabric Coating	8.5											
Veneer, Plywood & Engin. Wood Product Mfg.			0.1				<0.1	3.6	0.1			
Pulp, Paper & Paperboard Mills	52.2	1.4	5.0			<0.1	1.0	11.9	3.0			2.5
Petroleum & Coal Products Mfg.		<0.1	0.5		<0.1	<0.1	0.2	0.1	1.0	23.2		1.4
Basic Chemical Mfg.	0.3	<0.1	<0.1	1.5	3.3	0.5	<0.1	<0.1	0.9	<0.1		1.2
Paint, Coating & Adhesive Mfg.				0.2	2.5	<0.1	0.2	<0.1	<0.1			<0.1
Plastic Product Mfg.	1.3	<0.1	0.1	<0.1		<0.1	<0.1		<0.1			<0.1
Clay Product & Refractory Mfg.				1.8								
Cement & Concrete Product Mfg.		0.1	0.7	10.7		0.3	8.7	8.7	10.6	0.1	25.1	5.4
Iron & Steel Mills & Ferro-Alloy Mfg.	<0.1	0.2	1.2	12.8		0.2	2.0	19.7	15.1	0.3	0.2	20.6
Steel Product Mfg. from Purchased Steel		<0.1		0.5	0.2	<0.1	0.2	1.1	0.1	0.4		14.1
Alumina & Aluminum Production & Processing		<0.1	<0.1	<0.1	<0.1	<0.1	<0.1	5.1	0.6	0.3	<0.1	0.4
Non-Ferrous (exc. Al) Production & Processing	<0.1	3.4	14.6	0.7		29.3	21.3	0.6	0.8	0.5	2.5	23.8
Foundries	36.8	48.4	15.0	45.7	2.8	14.0	37.2	16.4	10.5	1.0	12.4	9.2
Forging & Stamping				0.4			<0.1	1.4	<0.1	0.1		<0.1
Coating, Engraving & Heat Treating Act.	<0.1		<0.1	0.2		<0.1	<0.1			<0.1		1.9
Other Fabricated Metal Product Mfg.		<0.1	<0.1	2.0	39.4	1.0	<0.1	0.3	0.4	6.9		<0.1
Agr., Construction & Mining Machinery Mfg.				8.4	1.3	<0.1	<0.1	1.0		5.0		0.2
Ventilation, Heating, AC & Refrig. Equip. Mfg			<0.1				<0.1	5.0	<0.1			<0.1
Electric Lighting Equipment Mfg.							2.2		3.8			
Motor Vehicle Parts Mfg.	<0.1	<0.1	<0.1	2.2	<0.1	3.1	0.3	20.2	<0.1	0.3	<0.1	2.3
Waste Treatment & Disposal		<0.1	0.1	<0.1		<0.1	<0.1	<0.1	13.6	<0.1		0.3

Figure 1: Population by Latitude-Longitude Squares



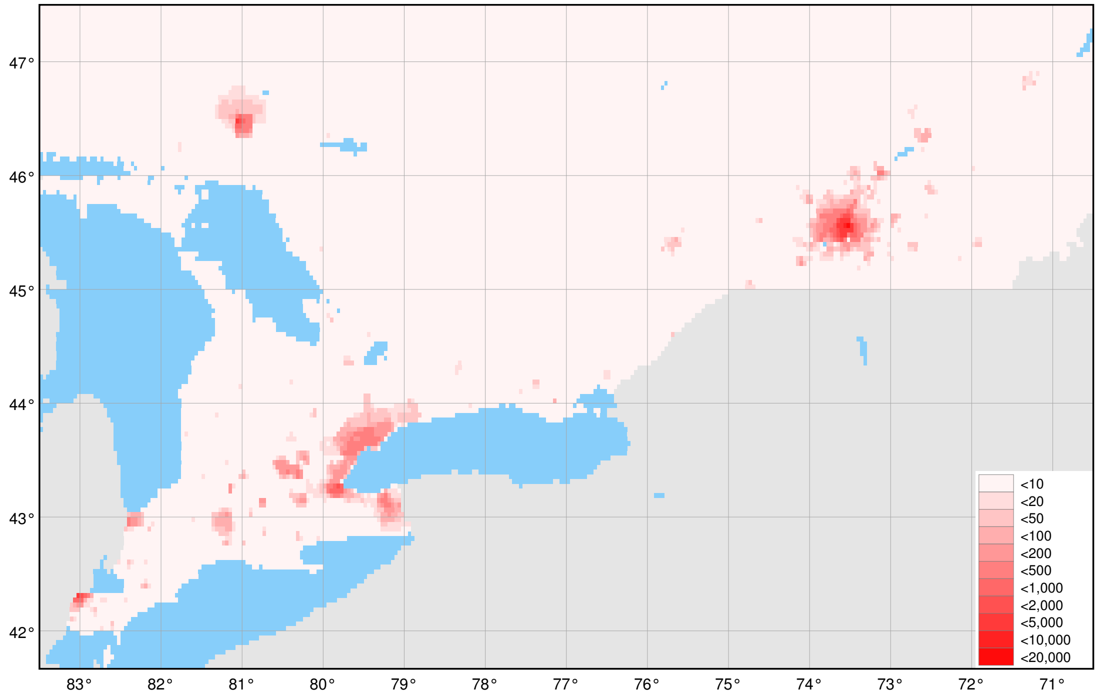
Note: Population figures in absolute numbers based on year 2000 reference data. Source: CIESIN (2005).

Figure 2: Pollution Concentration Plumes by Latitude-Longitude Squares



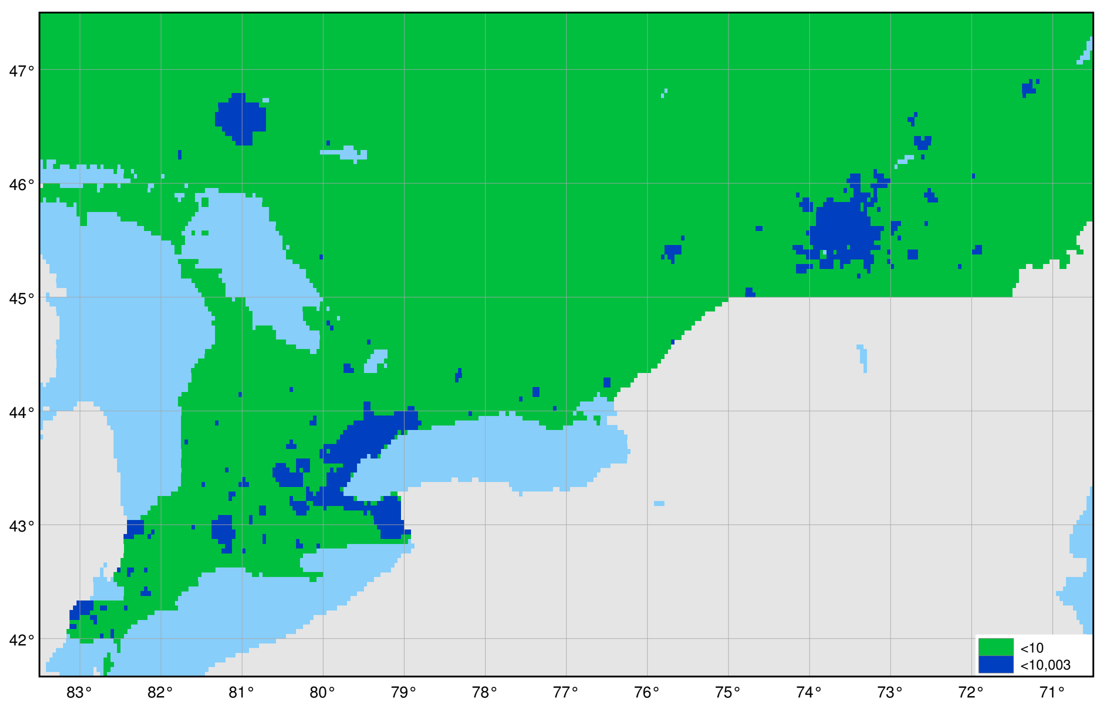
Note: Air pollution concentration plumes are computed using Gaussian plumes. Numbers are toxicity scores based on an aggregation of twelve metallic toxics using emission data from 2001–2005.

Figure 3: Hazard by Latitude-Longitude Squares



Note: Hazard (exposure risk) is calculated using equation (1) and utilizing the population and emission data as shown in figures (1) and (2).

Figure 4: Hazard Zones Allocation: Attainment Zone vs. Non-Attainment Zone



Note: Hazard threshold is based on a typical simulation run.

A suitable geographic grid for modeling air toxics is provided by the CIESIN (2005) database of the *Gridded Population of the World* (version 3), which projects the population of the world into grid squares of 2.5 arc minutes of longitude and latitude. At 45 degrees latitude, a typical grid cell covers about 15 square kilometers. Specifically, the field between 70°30″ and 83°30″ Western longitude and 41°40″ and 47°30″ Northern latitude covers 312 by 140 grid squares and about 17 million people, roughly half of Canada's population. Figure 1 shows the population density along the Ontario-Quebec corridor, including all major urban areas from west to east: Windsor, Waterloo/Kitchener, Hamilton, Toronto, Kingston, Ottawa, Montreal, and Quebec City). The map is in logarithmic scale, identifying the dense urban areas in red.

Emission dispersion can be modeled by projecting emissions from individual plants into air concentration plumes that decrease in intensity from the point of origin. One particularly simple version of this approach assumes that the plumes decrease exponentially in density and equally in all directions. With suitable knowledge of meteorological conditions such as wind speeds and wind directions it is possible to develop quite realistic air concentration plume projections. Such sophisticated modeling is outside the scope of this paper. Modeling a symmetric radial exponential air plume requires a single parameter r, the mean travel distance of pollution particles. (See Appendix B for details.)

Figure 2 simulates the emission distribution of toxics released into the air, using actual data from Canada's National Pollutant Release Inventory for the years 2001-2005. The metallic toxics were combined with the toxicity factors described earlier, and the mean dispersion distance was assumed (probably quite generously) as 100km. Dispersion does not account for differences in stack heights, wind direction and other relevant factors. The green areas indicate low emission concentrations, and the brown areas indicate high emission areas. The emission concentration scale is logarithmic. A small number of plants is responsible for very large emission loads. Very noticeable is the area around Sudbury, a major location for nickel mining and processing. A second hotspot is the Montreal area, where numerous metallurgical smelters and mills are located.

Figure 3 combines the population densities and projected emission concentrations into health/ecosystem hazards using the linear hazard function (1). What is immediately apparent is that the hazard potential is highly locally concentrated around major urban areas (the greater Montreal area and the 'golden horseshoe') and hotspots of smelter activity such as Sudbury. On the logarithmic scale of figure 3 virtually all other areas of Ontario and Quebec face very small health hazard due to metallic toxics. Lastly, based on simulation runs explained below, figure 4 shows a typical allocation of hazard zones, identifying non-attainment zones and attainment zones.

6 Simulation Results

The numerical complexity of a simulation is quite considerable. In Central Canada there are over 800 possible participants in an emission trading system and 22,796 populated longitude-latitude squares, forming an dispersion matrix with over 17 million entries. Without much loss in accuracy it is possible to constrain the computational complexity somewhat by including only the top-400 polluters with toxic loads exceeding 1 (covering 99.9% of the toxic load) and the 15,056 population squares with 50 or more people (covering 99% of the population).

To perform a simulation it is necessary to generate some heterogeneity across plants in terms of their pollution abatement ability. This information is unobservable to researchers and regulators alike. In the absence of such information it is necessary to adopt a plausible assumption about the distribution of b_i. Concretely, let u_i denote a random draw from a uniform [0,1] distribution, and assign $\tilde{b}_i = \beta^{u_i - 0.5}$, where $\beta > 1$ describes the range, or heterogeneity, of abatement ability and abatement costs.[4] Table 3 shows the results of the simulation with 1000 repetitions each for three abatement cost ranges β: 2, 5, and 10.[5]

Five regimes are compared: the optimal intervention (given perfect knowledge of the abatement ability of each plant), a conventional cap-and-trade regime with a single permit price, the hypothetical zonal permit system with full information, the actual zonal permit system with regulator-defined permit allocations, and the corresponding constrained permit system in which only the non-attainment market is operational ($\tau^L = 0$). The optimal policy intervention is designed to reduce the total loss by about half, reducing hazard by over 70% and introducing abatement costs of just over 20% of the pre-abatement loss. The optimal policy reduces emissions to a level of about 70-74% of the pre-abatement level. A number of 'stylized results' emerge from the simulation:

a) The conventional cap-and-trade regime where all firms face an identical permit price is the least efficient intervention. This regime requires aggressive overall emission reductions (about 50-60%) in order to reduce the hazard level in the emission hot spots. This drives up abatement costs, even though the hazard cost remains relatively high (40-50% relative to 27-29% with the optimal intervention.

b) The hazard zone permit system fares considerably better than the conventional cap-and-trade regime. It is clearly not as good as the optimal intervention, but it is by a fair margin not as inefficient as a simple cap-and-trade system.

---

[4]By construction, the median of $\tilde{b}$ is one, although the mean is $\hat{b} = (\beta - 1)/(\ln(\beta)\sqrt{\beta}) > 1$. For the β parameters 2, 5, and 10 used in the simulation, the corresponding $\hat{b}$ are 1.020, 1.111, and 1.236, respectively. The chosen distribution creates greater cost variation in the upper half of plants and smaller cost variation in the lower half of plants.

[5]Without actual knowledge of plant-level abatement abilities it is impossible to say much about the abatement cost distribution or its range. However, the chosen abatement cost ranges from 2 to 10 cover plausible scenarios. Allowing for an even wider cost range will inevitably lead to scenarios that are driven by which plants participate and reduce emissions ($\tau_i > b_i$), and which plants only buy permits and do not reduce emissions ($\tau_i \leq b_i$).

Table 3: Simulation Results for Central Canada

Policy Intervention	Emission	Policy Costs			Permit Prices		Permit Allocations		Limit	Participation	
	$E(\%)$	$L(\%)$	$B(\%)$	$\gamma H(\%)$	τ^H	τ^L	$A^H(\%)$	$A^L(\%)$	$\rho(\xi)$	n^+	n^-
Panel A: abatement cost range 2											
Optimal Policy Intervention	73.6	49.3	22.0	27.4						265	91
	(0.7)	(2.0)	(1.2)	(1.0)						(2.0)	(2.0)
Conventional Cap-and-Trade	40.8	76.8	36.0	40.8		2.51		40.8		356	
	(0.4)	(3.9)	(2.4)	(3.6)		(0.16)		(0.4)			
Full-Information Hazard Zones	62.4	57.0	25.1	31.9	41.29	0.13			0.696	214	142
	(6.5)	(2.2)	(3.4)	(4.2)	(57.35)	(0.20)			(0.045)	(39.4)	(39.4)
Dual Market Hazard Zone System	63.5	58.6	24.6	34.0	105.27	0.38	1.8	61.7	0.737	230	126
	(3.6)	(2.7)	(2.6)	(3.4)	(172.9)	(0.48)	(1.0)	(4.5)	(0.086)	(47.9)	(47.9)
One Market Hazard Zone System	60.9	58.0	26.9	31.2	30.50		2.7		0.676	219	137
	(2.0)	(2.4)	(2.2)	(1.4)	(3.81)		(1.4)		(0.033)	(9.5)	(9.5)
Panel B: abatement cost range 5											
Optimal Policy Intervention	72.0	49.7	21.8	27.9						264	92
	(3.3)	(4.6)	(2.8)	(2.4)						(3.8)	(3.8)
Conventional Cap-and-Trade	44.4	76.6	32.1	44.5		2.50		44.4		352	4
	(1.1)	(9.2)	(4.7)	(8.9)		(0.34)		(1.1)		(10.5)	(10.5)
Full-Information Hazard Zones	69.7	57.0	19.7	37.2	118.33	0.56			0.746	219	137
	(10.2)	(5.7)	(4.8)	(7.6)	(142.7)	(0.55)			(0.145)	(47.8)	(47.8)
Dual Market Hazard Zone System	68.8	58.7	20.1	38.7	154.69	0.59	1.2	67.7	0.792	194	162
	(5.2)	(5.3)	(3.7)	(4.0)	(175.5)	(0.47)	(1.2)	(6.2)	(0.095)	(47.7)	(47.7)
One Market Hazard Zone System	63.6	57.6	22.8	34.8	25.21		4.7		0.639	218	138
	(5.3)	(5.5)	(5.3)	(3.3)	(9.21)		(4.4)		(0.095)	(24.0)	(24.0)
Panel C: abatement cost range 10											
Optimal Policy Intervention	70.6	50.4	21.7	28.7						259	97
	(5.2)	(6.9)	(3.9)	(3.8)						(4.7)	(4.7)
Conventional Cap-and-Trade	49.4	76.7	26.4	50.2		2.43		49.4		310	46
	(1.7)	(13.4)	(6.3)	(13.1)		(0.56)		(1.7)		(34.4)	(34.4)
Full-Information Hazard Zones	69.5	57.5	19.4	38.1	143.96	0.63			0.764	215	141
	(10.0)	(8.0)	(4.6)	(8.1)	(155.5)	(0.55)			(0.142)	(37.8)	(37.8)
Dual Market Hazard Zone System	71.8	60.1	18.1	42.0	134.37	0.50	1.0	70.9	0.807	153	203
	(5.3)	(7.9)	(6.4)	(4.8)	(156.8)	(0.33)	(1.2)	(6.0)	(0.073)	(48.4)	(48.4)
One Market Hazard Zone System	65.0	58.2	20.6	37.6	26.65		4.2		0.656	204	152
	(5.5)	(7.9)	(5.9)	(4.3)	(26.15)		(3.4)		(0.076)	(28.4)	(28.4)

Note: Avg.: Average of 1000 simulation runs. S.D.: corresponding standard deviation. $E(\%)$ are total emissions relative to unabated emissions. $L(\%)$, $B(\%)$, and $\gamma H(\%)$ are total loss, abatement cost, and hazard cost in percent relative to the no intervention case. In the case of abatement cost, the reference point is 'normal abatement cost' $\sum_i B_i$. Permit prices τ^H and τ^L for the non-attainment and attainment markets are to be understood in relative terms, not absolute terms, as abatement costs are not known. Permit allocations A^H and A^L are in percent of the total unabated emissions. $\rho(\xi)$ is a function of the hazard threshold so that $\xi = \bar{\xi}^\rho$ with $\bar{\xi} \equiv \max_j\{P_j E_j\}$. This expresses the hazard threshold on a logarithmic scale relative to the highest hazard of any region j. n^+ and n^- are the number of participating and non-participating plants, respectively.

c) The permit price range τ^H/τ^L in the hazard zone permit system tends to be relatively large, reflecting the fact that the permit allocation for the attainment market (about 60% of original emissions) is much larger than the permit allocation for the non-attainment market (about 2% of the original emission).
d) The system with only a single non-attainment market performs remarkably well compared to the two-market hazard zone system. The performance difference between the two systems is not statistically significant. This result is perhaps not surprising given that the attainment permit price τ^L tends to be very small relative to the non-attainment permit price τ^H. The additional gains from having the extra market for the (low-hazard) attainment zone tend to be very small. Compared to the two-market system, the allocation of permits for the single non-attainment market would be somewhat larger as the boundary of the high-hazard zone is expanded (i.e., the threshold ξ is lowered).
e) The grid search for the hazard threshold ξ that separates the attainment from the non-attainment zones reveals that the loss function L^Z is not monotonous with respect to ξ. The large standard errors in the optimal permit prices are caused by different ξ thresholds delivering very similar losses.
f) As the abatement cost range increases from 2 to 5 to 10, it becomes apparent that the abatement participation constraint ($\tau_i > b_i$ for non-zero effort) is triggered with increasing frequency. This issue is often ignored in theoretical models of pollution abatement. In practice, however, many plants may find it less costly to buy permits than to reduce emissions.

The simulation results have to be interpreted with great caution as they are based on particular assumptions about the abatement ability and abatement cost function of plants. In addition to defining hazard reduction targets, the regulator's most significant concern is how to split the allowable total emissions into attainment and non-attainment permit allocations. The simulation suggests that the non-attainment market will likely be very small in magnitude, although participation among plants remains widespread.

The emission reductions suggested by the simulation are relatively large. But even emissions reductions of 10–20% would necessarly lead to significant hazard reductions through the hazard zone system because these reductions would be focused on the emitters with the largest contributions to the non-attainment zone.

Perhaps the most surprising result of the simulations is the discovery that a single market hazard zone system works virtually as well as a dual-market hazard zone system. Adopting a single-market system is very appealing from a political feasibility point of view. A significant number of firms would not need to participate in the system, thus reducing transaction costs dramatically. The firms that do participate would face varying contribution shares s_i that are determined by their proximity to the non-attainment zone. As a permit system, the permit contracts are in units of emissions rather than contributions to ambient concentrations, which are much more difficult to verify or enforce. All considered, as long as the contribution shares s_i can be determined with a sufficient degree of objectivity, the single-market hazard zone system may well turn out to be one of the most practicable solutions to dealing with hazard hot spots. Introducing the non-attainment zone contribution shares s_i is a simple but powerful innovation that deals effectively with the 'hot spot' problem.

7 Conclusion

Emission permit trading for air toxics is made difficult by the fact that the dispersion of these toxics is geographically confined and impacts regions of varying population density or ecosystem importance. Addressing the resulting spatial heterogeneity in emission concentrations is a formidable task for market-based instruments. High transaction costs make it infeasible to operate a large number of regional ambient concentration contribution permit markets as envisioned in the pioneering work of Montgomery (1972). While numerous alternatives have been explored, a solution that is sufficiently simple to become a politically feasible option has been elusive. The hazard (attainment/non-attainment) zone model introduced in this paper attempts to present a solution that is simple, makes use of as few markets as possible, facilitates contracts in emissions rather than emission concentrations, while at the same time preserving a high level of economic (and environmental) efficiency. In order to increase efficiency and liquidity in the proposed permit system, it is worthwhile to include multiple pollutants with similar characteristics that are emitted by a large numbers of plants. It is possible to aggregate the emission of air toxics into a toxic load measure through suitable toxicity scores that capture their adverse health effects.

The proposed permit trading system using hazard zones is analyzed by simulating the likely consequences of operating such a system for metallic air toxics in the provinces of Ontario and Quebec, which covers roughly half of the population of Canada. The simulation results suggest that the proposed system is significantly superior to a conventional cap-and-trade system because it can focus more directly on emission concentration hot spots. A surprising result is that the constrained version with only a non-attainment permit market may perhaps lead to more efficient outcomes than a dual market system with an attainment and a non-attainment market. The simplicity of the single non-attainment permit market system is particularly appealing.

While second-best approaches to dealing with the spatial heterogeneity problem of air toxics may be suitable approximations of the first-best approach, numerous practical challanges remain. Detailed modeling of the air toxics dispersion is needed to determine plant-region specific dispersion factors. Large variability of emissions over time may require suitable lengthening of the permit period or the introduction of banking and borrowing instruments. Bioaccumulation and biomagnification may necessitate treating air toxics as stock pollutants rather than flow pollutants, and allow for the possibility that the location where emission concentrations first impact the ecosystem may become detached from the location where health impacts are realized. There remains a large knowledge deficit in exploring the empirical dimension of all of these practical problems.

References

Atkinson, Scott E., Thomas H. Tietenberg. 1987. Economic implications of emission trading rules for local and regional pollutants. *Canadian Journal of Economics* **20**(2) 370–386.

CIESIN. 2005. Gridded population of the world, version 3. At www.ciesin.org.

Cronshaw, Mark B., Jamie B. Kruse. 1996. Regulated firms in pollution permit markets with banking. *Journal of Regulatory Economiics* **9**(2) 179–189.

Egenhofer, Christian. 2003. The compatibility of the kyoto mechanisms with traditional environmental instruments. Carlo Carraro, Christian Egenhofer, eds., *Firms, Governments and Climate Policy: Incentive-Based Policies for Long-Term Climate Change*, chap. 1. Edward Elgar, Cheltenham, 17–82.

Gangadharan, Lata. 2000. Transaction costs in pollution markets: An empirical study. *Land Economics* **76**(4) 601–614.

Hoel, David G., Christopher J. Portier. 1994. Nonlinearity of dose-response functions for carcinogenicity. *Environmental Health Perspectives* **102**(Suppl 1) 109–113.

Liski, Matti, Juan-Pablo Montero. 2005. A note on market power in an emission permits market with banking. *Environmental and Resource Economics* **31**(2) 159–173.

Liski, Matti, Juan-Pablo Montero. 2006. On pollution permit banking and market power. *Journal of Regulatory Economics* **29**(3) 283–302.

Montero, Juan-Pablo. 1997. Marketable pollution permits with uncertainty and transaction costs. *Resource and Energy Economics* **20**(1) 27–50.

Montgomery, David W. 1972. Markets in licenses and efficient pollution control programs. *Journal of Economic Theory* **5**(3) 395–418.

Rubin, Jonathan D. 1996. A model of intertemporal emission trading, banking, and borrowing. *Journal of Environmental Economics and Management* **31**(3) 269–186.

Schennach, Susanne M. 2000. The economics of pollution permit banking in the contect of title IV of the 1990 Clean Air Act Amendments. *Journal of Environmental Economics and Management* **40**(3) 189–210.

Stavins, Robert N. 1995. Transaction costs and tradeable permits. *Journal of Environmental Economics and Management* **29**(2) 133–148.

Tietenberg, Thomas H. 1995. Tradeable permits for pollution control when emission location matters: What have we learned? *Environmental and Resource Economics* **5**(2) 95–113.

Tietenberg, Thomas H. 2006. *Emissions Trading: Principles and Practice*. 2nd ed. RFF Press, Washington, DC.

Tietenberg, Thomas H., Michael Grubb, Axel Michaelowa, Byron Swift, ZhongXiang Zhang. 1999. *International Rules for Greenhouse Gas Emissions Trading: Defining the principles, modalities, rules and guidelines for verification, reporting and accountability*. UNCTAD, Geneva.

Turner, D. Bruce. 1994. *Workbook of Atmospheric Dispersion Estimates: Second Edition*. 2nd ed. Lewis Publishers, Boca Raton.

US-EPA. 2004. EPA's Risk-Screening Environmental Indicators (RSEI) Chronic Human Health Methodology, Version 2.1.2, Appendix A. Tech. rep., United States Environmental Protection Agency. At www.epa.gov/opptintr/rsei/pubs/.

Wald, Matthew L. 2008. E.P.A. proposes new limits on lead in the air, the first revision in 30 years. *New York Times* 2 May 2008.

Yap, David, Neville Reid, Gary de Brou, Robert Bloxam. 2005. Transboundary air pollution in Ontario. Tech. rep., Ontario Ministry of the Environment.

A Emission Trading in Canada

While emission permit trading has been successfully used in numerous jurisdictions including most prominently the United States, to date Canada has very limited experience with emission trading systems. This experience consists mostly of analysis and consultations, two voluntary trial programs implemented as public-private partnerships, and private sector trades. The two experimental programs are the Pilot Emission Reduction Trading (PERT) project and the Greenhouse Gas Emission Reduction Trading (GERT) project. These credit systems were launched in 1996 and 1998, respectively. PERT provides tradeable credits to firms in Ontario that reduce emissions by more than required by regulation. Ownership of registered credits can be contractually transferred between parties. PERT covers NOx, VOC, CO_2 and SO_2 emissions. Curiously, the value of these credits stems from the ability to sell the credits as offsets to foreign companies, to the extent that these foreign markets allow offsets. GERT was a similar program involving six other Canadian provinces. It was operated through 2001 and only involved greenhouse gases. In 2000, PERT was supplanted by *Clean Air Canada Inc.*, a non-profit organization, formed by the original private sector members in PERT. Currently, most activity regarding emission permit trading focuses on the implementation of the Kyoto protocol to limit greenhouse gas emissions. No trading system has ever been considered for toxics.

B Emission Dispersion Simulation

In this paper, emission dispersion is simulated using a simple single-parameter spatial process defined by r, the mean distance for pollutants to be carried through the air. A given emission load v can be allocated to grid cells or polygons by splitting the load v into $n \geq 10,000$ partial loads of weight v/n, and then distributing these partial loads at impact locations drawn from a suitable dispersion distribution, described below. Partial loads v/n are added up within each grid cell or polygon. This procedure amounts to spatial integration.

Assume that emissions disperse symmetrically around a given source point, not taking into account that the dispersion plume is influenced by wind speed and direction and stack height. To determine depositions by grid cell or polygon, draw a pair of random numbers (ω, d) comprised of compass angle ω (clockwise from North) and distance d, given average radial dispersion of r. Angle ω is drawn from the uniform distribution $[0°, 360°[$, and distance d is drawn from the exponential distribution with cumulative distribution function $\Omega(d) = 1.0 - \exp(-d/r)$. The exponential distribution function has mean r, the average impact distance. The median impact distance is $\ln(2)r$. In practical terms the distance d can be generated by first drawing z from a uniform $[0, 1[$ distribution and then transforming z through the inverse cumulative distribution function $d = -r \ln(1 - z)$. Given (ω, d) and the location of the emission source (ϕ_0, λ_0) of latitude ϕ_0 and longitude λ_0, the impact location (ϕ_1, λ_1) can be calculated through loxodrome approximation, which is sufficiently accurate for small distances:

$$
\begin{aligned}
\phi_1 &= \arcsin\left(\sin(\phi_0)\cos(d/D) + \cos(\phi_0)\sin(d/D)\cos(\omega)\right) \\
\lambda_1 &= \lambda_0 + \arcsin\left(\sin(\omega)\sin(d/D)/\cos(\phi_0)\right)
\end{aligned}
$$

where D is the earth radius of 6371km.