Agent-based framework for modelling high resolution spatio-temporal air quality

# Introduction

Exposure to ambient air pollution is a serious human health risk. Epidemiological studies

have linked exposure to ambient air pollution to an increase in mortality.1–4 WHO attributes roughly 4.8 million deaths worldwide in 2016 to exposure to ambient air pollution.5 Among

the various pollutants, particulate matter is deemed the most potent. Exposure to PM2*.*5

has both long and short-term effects that are detrimental to human health.6–8 It is therefore important to measure the extent of exposure to humans to help assess its impact and take preventive action. Air quality is a dynamic entity and varies drastically with space and time. Understanding spatio-temporal variability in air quality is key to building a robust exposure assessment framework.9 Existing air quality monitoring infrastructure however is limited in its ability to capture intraurban spatial variation. Various modelling approaches have been suggested to fill such gaps in information. These include, but are not limited to, proximity modelling, interpolation techniques, dispersion modelling and land use regression (LUR) models among others. These classes of intraurban air quality modelling approaches have

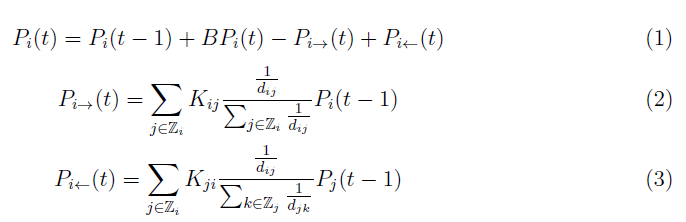
been studied extensively and have their own sets of merits and demerits.10–13 Reliable air pollution estimation and exposure assessment are done primarily to inform policy changes and strategies for mitigation and abatement. Bibbero, emphasizes the need for a systems approach to tackling control of air pollution.14 However, very little has been done in this area. Addressing the problem from a systems engineering perspective would allow us to complete the loop from measurement through control of air pollution. A robust model of the system under assessment is a fundamental requirement under this approach. Conventional dispersion models, considered more sophisticated and reliable, are expensive and computationally intensive. LUR models on the other hand offer reasonable estimates and are inexpensive.

However, they are limited by their area specificity and limited incorporation of physics[10].

We propose an agent-based framework to model spatio-temporal variations in air quality. It draws from the strengths of both physics based models as well as LUR models. Agent based modelling is an approach through which systems are described using a collection of individual entities, called agents, that interact among one another and their environment. The interactions are defined in terms of a predefined set of rules that could be either stochastic or deterministic. Each agent individually assesses its situation and makes decisions depending upon of the set of rules. Agents execute distinct behaviors or response depending upon the current state and the applicable rules for the system they represent. These interactions take place repeatedly between agents depending upon their state. In addition to this, agents are capable of evolving over time which can lead to complex emergent behaviour. In the current study, a framework is developed for estimating air quality over a well-defined space. The model uses a simple mass balance to quantify change in pollutant concentration. This framework avoids solving complex transport equations to predict the pollutant concentrations, while still maintaining the sanctity of fundamental physics. In addition, this framework allows for a straightforward extension to derive real world solutions and mitigation strategies in addition to just estimating air quality and its exposure.

# Model Description

This first step in this model is to consider a study area and discretize it into a finite number of distinct locations. These locations or nodes are the agents that interact with one another to update the state of the system. Each node covers a specific area. Features, such as land use, traffic information, meteorological information and pollutant concentration, to name a few, specified at each node’s domain become its property. The value that these features take at each node dictate the behavior of the agents. In this particular work, our focus is on pollutant concentration at each node and its exchange with other neighboring nodes. This exchange is dictated by a simple mass balance. The mass balance accounts for the total pollutant concentration at each node as a linear sum of generation/loss, transfer and accumulation of pollutants within each node at different times. Equation 1, describes the governing mass balance equation for pollutant concentration at each node ’i’, while equations 2 and 3 describe the transfer between the nodes. The mass balance is solved at each node individually and consolidated for all nodes to get a spatio-temporally variate, pollutant concentration map for the entire study area.



*Pi*(*t*) is the total pollutant at node *i* at time *t*. *Pi*(*t* − 1) is the pollutant at node *i* in the previous time instant. *BPi*(*t*) is the background pollutant addition/removal at node i between times (*t* − 1) and (*t*). *Pi*→(*t*) is the total pollutant transferred from node *i* to neighboring nodes between time (*t* − 1) and *t*. *Pi*←(*t*) is the total pollutant transferred to node *i* from neighboring nodes between time (*t* − 1) and *t*. Z*i* is the set of nodes that are in the neighborhood of *i*. *dij* and *dji* are the euclidean distances between nodes *i* and *j*, and *Kij* and *Kji* are the convective transfer coefficient for node *i* with respect to node *j* and

vice-versa, respectively.

Pollutant transfer across nodes is primarily convective and depends on the local convective forces, such as wind and other meteorological features. The coefficient K captures the effect of these convective forces at each node. Three modes of internodal pollutant transfer can be applied in this approach, which in turn, define the neighboring nodes, viz. 1) Adjacent: where the transfer occurs only with the immediately adjacent nodes, 2) Line of sight (LoS): where transfer occurs between such nodes that have no physical infrastructure that obstructs the transfer of pollutant and 3) All nodes: Where transfer can occur to all the nodes in the study area simultaneously.

The term BP, represents the background pollution added or removed at each node. Based on sign of BP, each node could act as either a source or a sink for the pollutant of interest. Here, the pollutant may be added to the system, for instance due to vehicular traffic, industrial or domestic sources. It may also act as a sink, for instance, if this node is in close proximity to trees or lakes or any other open body where absorption of a particular pollutant may happen. Approaches for obtaining K and BP are discussed later in this paper. The working of this model is explained in the following sections using a simulation study.

# Results And Discussion

In order to illustrate the agent-based air pollutant dispersion model, a case was simulated. The simulation results are presented just to elucidate the model and the implementation of the model in a real environment with actual data will described in a subsequent section. In this simulation, the region is envisioned to consist of 9 nodes that are distributed evenly across a 3x3 square grid. Figure 1 shows the arrangement of nodes in the grid. The figure also shows the paths for pollutant transfer and the areas that obstruct it. This particular run of simulation assumes the Line of Sight (LoS) mode of internodal pollutant transfer. The figure highlights the nodes that are in the line of sight of node 1.

The variation in concentration of an undefined, archetypical pollutant was considered across the grid over different times. Each time step corresponded to one hour of the day, starting from 0:00 hours, all the way through 24:00 hours. Figure 2 shows *Pi* varying as a function of *BPi* , *Pi*→ and *Pi*← for one particular run of the simulation. BP concentration in this case was obtained randomly from a normal distribution with a predefined mean and

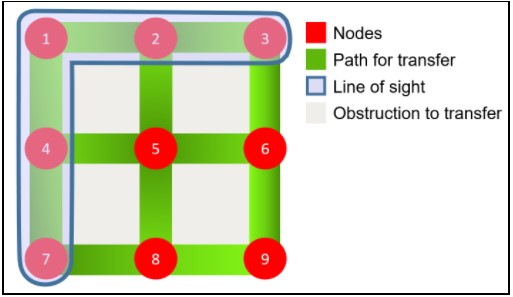


Figure 1: Arrangement of nodes for model simulation

standard deviation. In order to account for the source and sink cases, the mean for the normal distribution was chosen between -10 to 10 units with a standard deviation ranging from 3 to 5 units. Since K represents the portion of a node’s pollutant that gets transferred, it takes values from 0 to 1. In this particular run, K for all nodes is chosen to be 0.5.

The model, not only takes into account the state of the background pollution at each node but also that of it neighbouring nodes. This internodal interaction, through convective transfer ensures that spatial continuity in the model is preserved. This is evidenced at Node 5 between hours 19 and 23 as depicted in Figure 2. Between the 19th and 22nd hour, even though there is negligible increment in BP at node 5, the total pollutant concentration at the node keeps on increasing steadily until the 22nd hour. However, during the same time interval, a steep increment in the BP concentrations of nodes 2, 4 and 8(all three in node 5’s line of sight) can be observed. The increasing pollutant concentration at the neighbouring nodes spills over to node 5, thus contributing to the gradual increase in its total pollution. Subsequently, the total pollutant concentration decreases marginally in the 23rd hour owing to a decrease in the BP at node 5 as well as its neighbours.

One must not however, based on the aforementioned case, discount the impact of BP. At

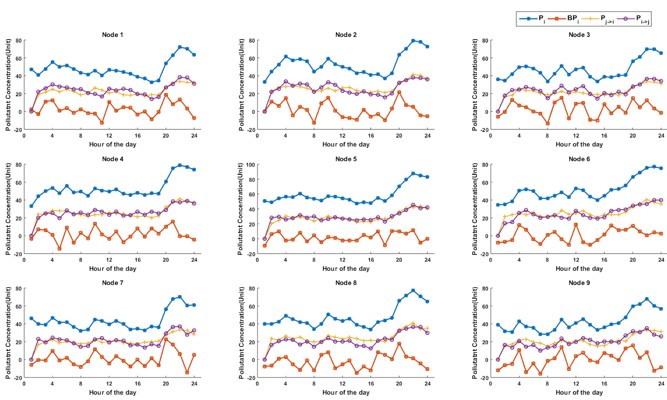


Figure 2: Variation of pollutant concentration at all nodes as a function of time

node 3, BP fluctuates drastically from hour 7 through 16, and has a conspicuous impact on the total pollutant concentration. A similar impact is apparent at node 7, at the 19th and 20th hour.

The relative impact of a node’s background pollution as against that of its neighboring node can be tweaked by tweaking the value of K. For higher values of K, the influence of convective transfer on each individual node would increase. Thus, K acts as a primary tuning parameter for tweaking the quantum of pollutant exchange. However, the total pollutant concentration within the system remains consistent and is dictated by BP.

The autoregressive nature of the model, which is contributed by the *Pi*(*t* − 1) term, preserves temporal continuity. The smooth transition in total pollutant concentration from one hour to the next, even in the presence of strong fluctuations in BP, can be attributed to this effect. Tuning the order of the autoregressive term provides control over the temporal effect in the model.