Integrated Transportation and Land Use Regression Modelling for Nitrogen Dioxide Mitigation

Aonghus Ó Domhnaill, Brian Broderick and Margaret O’Mahony

a. Department of Civil, Structural & Environmental Engineering
University of Dublin, Trinity College, Dublin 2, Ireland

Email addresses:
domhnaia@tcd.ie
brian.broderick@tcd.ie
margaret.omahony@tcd.ie (Corresponding author)

To cite:
Abstract

The transport sector has been identified as one of the main contributors to nitrogen dioxide (NO₂) pollution in Ireland. This research develops an enhanced Wind Sector Land Use Regression (WS-LUR) model to estimate NO₂ concentrations across Ireland, in areas where air pollution monitoring is not available. The model incorporates details of the vehicle fleet breakdown to weight vehicle type flows based on the emission rates of the vehicle type, differentiating routes with varying proportions of heavier emitting vehicles. In 2008, car taxation underwent a significant change from an engine size based system to a carbon dioxide (CO₂) emission rate based system resulting in a significant transition towards diesel fuelled vehicles. A mitigation strategy to remove diesel fuelled vehicles from the public service vehicle fleet was tested achieving predicted NO₂ reductions in the range of 0.3 μg/m³ to 1.9 μg/m³. The impact of COVID-19 on NO₂ concentration levels was also investigated.

Key words: land use regression, model, nitrogen dioxide (NO₂), vehicle type, mitigation.

Declarations of Interest: None

Highlights

- Developed wind sector land use regression model to estimate nitrogen dioxide (NO₂)
- Inclusion of varying vehicle type emission rates an improvement on existing models
- Method transferrable to land use regression models that include distance travelled
- Application of model to NO₂ mitigation strategies also demonstrated
- Removal of diesel from public service fleet achieved NO₂ reduction of 0.3-1.9 μg/m³.
1. Introduction

Short-term and long-term air pollution exposure can have significant impacts on health, with numerous health effects (cardiovascular and respiratory) linked to one or a number of air pollutants (WHO, 2021; US EPA, 2016). Nitrogen Dioxide (NO₂) has been identified as one of the main air pollutants which have sufficient evidence to confirm links to health effects. The transport and energy sectors as well as residential heating are common sources of NO₂ (WHO, 2021). According to the update of World Health Organisation guidelines in September 2021, levels of NO₂ exceed the World Health Organisation limits (WHO, 2021) at most monitoring locations and adverse weather conditions or increased emissions could lead to exceedances of the EU limits also (EPA, 2018). Therefore development of mitigation measures to reduce air pollution at specific locations is critical to achieve the revised limits and reduce health impacts.

The overall objective of the research presented in this paper is to develop a land use regression (LUR) model capable of identifying the environmental, meteorological and traffic related conditions which result in increased concentrations of NO₂ at various locations in Ireland. This analysis focuses on estimating concentrations within the Leinster, Cavan and Monaghan regions of Ireland. The emphasis of the work is the traffic related parameter within the LUR model that accounts for vehicle fleet breakdown on each route surrounding a study location. This model can estimate NO₂ concentrations and calculate changes in concentrations due to the implementation of mitigation measures such as adjustments in the vehicle fleet composition. The novel contribution is the potential of the proposed method to extend the analytical capability of LUR models to distinguish between the contributions of
different vehicle types, and allows locations with atypical traffic conditions to be more accurately represented.

Model validation was completed and the model was then tested on a characterisation of the meteorological, environmental and traffic conditions during the 1st COVID lockdown period. The model was then used to test a potential NO$_2$ mitigation strategy: the removal of diesel vehicles from the public service fleet.

The paper is structured along the following lines. The next section covers the background and relevant literature pertaining to the topic and this is followed by a description of the data and methods used. Results of both the validation and testing of the model are then presented followed by a discussion section. The main findings of the work are presented in the conclusions section.

2. **Background**

Nitrogen oxides (NOx) are formed by the combination of oxygen and nitrogen at high temperatures (WHO, 2010). Most nitrogen oxides are emitted as nitric oxide (NO) but in ambient conditions this is rapidly oxidised in air to form nitrogen dioxide (NO$_2$) which is considered a primary pollutant. In terms of ambient air quality, the main source of NO$_2$ in Ireland is road transport (EPA, 2019a), although other forms of transport and stationary sources also contribute to emissions. For road transport, diesel engines emit more NO$_2$ than petrol engines and their increased popularity in recent years is of concern, especially in urban areas where public exposure is highest. Long-range transport pollution also makes a significant contribution to NO$_2$ concentrations in Ireland (Donnelly et al, 2015).

Exposure to traffic pollutants has been linked to various cardiovascular and respiratory health issues in adults with some evidence also suggesting links to health impacts in children and birth outcomes (Health Effects Institute, 2022). Nitrogen dioxide has been
associated with adverse effects on hospital admissions for various diagnoses; decrements in
measures of lung function and lung function growth, increases in respiratory symptoms,
asthma incidence, adverse birth outcomes and mortality. (USEPA, 2016; WHO, 2013;
Brown, 2015, Liu et al., 2017). The evidence of associations of ambient concentrations of
NO2 with a range of effects on health has strengthened in recent years and the associations
are robust to adjustment for other pollutants including some particle metrics.

2.1 Monitoring of Nitrogen Dioxide (NO2)

The Irish EPA, working with local authorities and other public bodies has
established 96 air pollution monitoring stations, 18 of which were installed in 2020. (EPA,
2021). The National Ambient Air Quality Monitoring Programme was launched in 2017 to
expand the existing monitoring network to provide more local information on different
environments in built-up and rural areas (EPA, 2017). The historical database of
measurements represents a rich resource on air quality at different location types that
supports the development and validation of national air quality models.

The World Health Organisation’s Guidelines (WHO, 2005) and the European Union’s
Directive on Ambient Air Quality and Cleaner Air for Europe (2008/50/EC) (EU, 2008) have
set out short-term and long-term concentration limits for numerous air pollutants including
NO2. The WHO (2005) short-term limit (1-hour mean) of limit 200 μg/m³ was adopted by
the EU Directive which requires that the concentration level should not exceed this value
more than 18 times a year at any individual monitoring station. The long-term limit (annual
mean) of 40 μg/m³ was set out based on results from studies carried out on mixtures of air
pollutants containing NO2, which showed that people experienced health effects with
increasing NO2 concentrations (WHO, 2005). In September 2021 the WHO changed its
guideline values for NO2 to an annual mean of 10 μg / m³ and a 24-hour mean of 25 μg / m³.
(WHO, 2021). This reflected an observed increase in adverse health effects due to exposure
to air pollution and better information on the contribution of pollutants including NO₂ to the
global burden of disease in the fifteen years since the previous guidelines (WHO, 2021).

The annual mean limit value of 40 μg/m³ has been exceeded on five occasions in
Ireland, four of which occurred before 2010 and one in 2019. All of these occurred at Dublin
City monitoring stations, with the most recent occurring at a new monitoring location. In
response the regional local authorities cooperated to develop Air Quality Management Plans
to tackle the issue of increasing air pollution in the county (Dublin City Council, et al, 2021).
These highlight the need for improved local NO₂ modelling capability for air quality
assessment and management.

Nitrogen dioxide monitoring locations are distributed throughout Ireland in a variety
of urban, sub-urban and rural locations, as well as some locations near specific sources, such
as roadside monitors. However, fixed site monitoring is always limited to the number of
locations where measurements can be obtained. Air quality modelling can provide
information on concentrations of pollutants at a much larger number of locations limited only
by the availability of model input data, and computing/data processing resources. In this
project, air quality modelling is employed to calculate NO₂ concentrations at different
locations to investigate potential mitigation strategies. The land use regression modelling
approach employed makes use of monitoring data collected throughout Ireland and examines
the likely change in NO₂ concentrations expected due to changes in NO₂ source activities.

2.2 Land Use Regression Modelling

Land Use Regression (LUR) Modelling aims to produce a regression equation which
can estimate pollutant concentrations at any location based on the surrounding land use (EU,
2010; Donnelly et al., 2019; Aeroqual, 2021). Values for potential predictor variables
(including potential sources of pollution) at air quality monitoring locations are captured using Geographic Information Systems (GIS). Statistical associations between potential predictor variables and measured pollutant concentrations are determined using regression techniques. This determines the predictor variables which are linked to pollutant levels at the monitoring locations and produces regression coefficients which are applied to the predictor variables within the LUR equation.

The European Study of Cohorts for Air Pollution Effects (ESCAPE) project identified a standardised approach for the development of LUR models and analyzed the accuracy of models across 36 regions (Beelen et al., 2013). In the standardized approach, which was adopted in the development of the model employed in this study (Naughton et al, 2018) variables are only included in the final model if (EU, 2010; Beelen et al., 2013):

i. The adjusted $R^2$ increases by more than 1%;

ii. The direction of effect of the variable does not change;

iii. The variable does not change the direction of effect of previously included variables in the model

The cross-validation $R^2$ of LUR models identified by Briggs (2007) were between 0.45 and 0.7 and had standard errors less than 20%, which is similar to results that would be expected from advanced dispersion models (Briggs, 2007; Briggs, et al., 1997; Briggs, et al., 2000; Gilbert, et al., 2005; Briggs, 2005). The development of LUR models for 36 different regions in Europe within the ESCAPE project also generated strong results, with cross validation $R^2$ between 0.55 and 0.92 (Beelen et al., 2013).

Recent studies which utilised LUR models to estimate concentrations of air pollutants in various locations throughout the world also achieved significant results using different combinations of predictor variables. Jones et al. (2020) captured 66% of the spatial variability in Ultrafine Particles (UFPs) in Southern California using a combination of variables,
including inverse distance from LAX airport; NO$_2$ background concentrations from a spatiotemporal model; airport area within a 1km buffer; and total length of major highways within 50m of the study location. Liu et al. (2019) produced an annual NO$_2$ concentration LUR model for Xi’an in China and achieved an adjusted R$^2$ of greater than 0.85. The predictor variables for this study included green land use area within 500m, residential land use area within 1km, total road area within 3km and distance to the three nearest polluting factories.

A fine particulate matter concentration LUR model developed by Ross et al. (2007) for the New York City region captured approximately 64% of the spatial variability of the pollutant. The predictor variables within the model were vehicle kilometers per hour within 500m, population within 1km and industrial land use area within 300m. A study by Eeftens et al. (2016) developed LUR models for 10 regions in Switzerland that achieved adjusted R$^2$ values between 0.46 and 0.89. As the environments in each of the regions varied considerably, the unique combinations of predictor variables and buffer sizes were employed for each region.

A study of NO$_2$ estimation using an LUR approach completed by Shi, et al. (2020) along a transportation corridor in Mississauga, Canada captured 69% of spatial variability in NO$_2$ concentrations. The final model predictor variables included daily traffic flow within 200m, length of major roads within 50m, length of minor roads within 100m, distance to nearest major intersection, and areas of government/institutional and parks/recreational land use within 500m.

Previous LUR models share two common characteristics: they all require substantial monitoring data for calibration, and because this calibration is based on historical data they are rarely employed to predict future air quality. There is a need for an improved LUR modeling approach that supports the calculation of the changes in ambient concentrations
associated with future scenarios, such as proposed mitigation strategies, that can be implemented using available air quality and emissions information. This paper describes the development of a new LUR modeling approach that can account for changes in future vehicle fleet composition that can be utilised by professionals such planners and engineers to compare the performance of proposed transport policies, mitigation measures, and transport construction projects during the planning phase.

3. Methods

The research aimed to identify the main environmental, meteorological and traffic-related conditions which contribute to high NO\textsubscript{2} levels at various locations throughout Ireland. This required the synthesis of a considerable amount of data in the form of a land use regression-based (LUR) model that can be used to determine the NO\textsubscript{2} concentration at any location in Ireland. The LUR model was created by combining existing data on land use and traffic with historical air quality data, and it can be used to predict changes in air quality due to changes in land use or traffic characteristics.

Land use regression (LUR) models use statistical relationships between air quality at a location and the emission sources in the vicinity. The emission sources can be specific sources, such as individual roads or commercial facilities, but generally they are grouped together depending on the predominant land use in an area, e.g. residential, industrial, agricultural; hence the name ‘land use regression’. LUR models are calibrated using historical pollution concentration measurements and data on the land uses and other emission sources around the monitoring site. Meteorological conditions are also usually taken into account. If sufficient data are available, a well-developed LUR model can capture all of the parameters that, on average, statistically influence the ambient concentrations of pollutants in a city, region or country. They can be used to calculate pollution concentrations at locations
where monitoring has not been carried out, and to develop air quality maps, so long as the required input data are available. They can also be employed in scenario analyses to model the effect on air quality of future changes in land uses or emission sources. In this project, nationwide LUR model input data was compiled for three years, 2016 to 2018, allowing the average annual NO\textsubscript{2} concentration to be calculated at any location in any of those years. Traffic data is based on modelled data (typically outputs from models maintained by the National Transport Authority) whilst data for the other variables (meteorological, commercial properties and land use) are all actual measurements. The model also includes a manual entry option that facilitates calculation of average NO\textsubscript{2} concentrations for other years, including future years.

3.1 Wind Sector-based LUR Modelling – Original Model

The modelling approach employed is based on an original methodology used in Naughton et al. (2018) that calculates NO\textsubscript{2} concentrations using a wind sector-land use regression (WS-LUR) model. The calibration of the model is based on analysis of hourly concentrations measured at each site in the national network of fixed monitoring sites and includes corrections for seasonal and diurnal bias. Wind sector-based regression was found to be the best option for modelling air pollution concentrations in areas with a complex spatial distribution of sources and where the prevailing wind varies considerably. The approach involves dividing the area around a monitoring location into eight wind direction sectors. For any modelled hour, predictor variable values are evaluated using only land uses and sources within the sector containing the hourly wind direction. As well as improving the specificity of the land use data, this also increases the number of independent sets of calibration data by a factor of eight. The position of different sources in the vicinity of the monitoring location is defined by further dividing the eight sectors into eight buffer zones with minimum and maximum sector radii varying between 25m and 5km, as shown in Figure 1. This ensures that
the proximity of a source to a receptor and its ability to influence concentrations at that receptor is taken into account.

Using this method, the WS-LUR model for the mean concentration at a specific location can be represented as (Naughton et al, 2018)

\[ C = \alpha_0 + \sum_{i=1}^{8} \sum_{j=1}^{M} W_{fi} \alpha_j P_{ji} \]  

in which \( C \) is the modelled pollutant concentration, \( P_{ji} \) (\( j = 1, M \)) are the values of the \( M \) selected predictor variables in sector \( i \) (\( i = 1, 8 \)), \( \alpha_j \) are the corresponding regression coefficients, and \( W_{fi} \) is the fraction of hourly wind directions within sector \( i \) (\( i = 1, 8 \)) throughout the whole sampling period.

Figure 1. WS-LUR wind sectors and buffers (Naughton et al., 2018)
Data were gathered for a large number of candidate predictor variables on a 50m x 50m grid describing local spatial distributions of pollutant sources, with a focus on variables relating to the traffic and background characteristics. The LUR model was created from these variables by adding the highest adjusted R² variable first, after which the next highest variables are added consecutively to the model and maintained only if (i) the R² of the model increased by 1% or more, (ii) the direction of effect of the variable is unchanged from the direction of effect from the initial check and (iii) the directions of effect of other variables already included in the model do not change. (If a variable was found to increase concentrations in the initial ranking but after including it with other variables the concentrations were reduced, this would be considered a change in the direction of effect of that variable and therefore it would no longer be included in the model). The variables which produced the highest adjusted R² value were included in the final version of model and are presented in Table 1.

Table 1. Original Model Predictor Variables (Naughton et al., 2018)

<table>
<thead>
<tr>
<th>Predictor Variable</th>
<th>Buffer Radii</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-</td>
</tr>
<tr>
<td>Inverse Distance Weighted Vehicle Kilometres Travelled (IDWVKT)</td>
<td>0.025 – 5km (25m, 50m, 100m, 250m, 500m, 1km, 2km, 5km)</td>
</tr>
<tr>
<td>Commercial Buildings</td>
<td>1km</td>
</tr>
<tr>
<td>Natural / Agricultural Land Use</td>
<td>1km</td>
</tr>
<tr>
<td>Average Wind Speed</td>
<td>-</td>
</tr>
<tr>
<td>Road Density</td>
<td>0.25km</td>
</tr>
</tbody>
</table>

For each wind sector, the inverse distance weighted vehicle kilometres travelled (IDWVKT) predictor variable, $P_1$, is calculated using

$$P_1 = \text{IDWVKT} = \sum_{b=1}^{B} \frac{1}{d_b} vkm_b$$

(2)
in which $vkm_b$ is the total distance travelled by all vehicles in buffer zone $b$ ($b = 1, B$), and $d_b$ is the radial distance from the monitoring or receptor location to the centre of the buffer zone. The calculation of $vkm_b$ using transport model outputs of annual average daily traffic flows is described in the following section.

3.2 Wind Sector-based LUR Modelling Modifications to Facilitate Policy Scenario Analysis

The land-use regression modelling procedure developed by Naughton et al (2018) was retained but the capability of the model to capture the effects of vehicle emissions was enhanced. This improvement was required to support the calculation of the changes in ambient concentrations associated with future scenarios, such as proposed mitigation strategies, that can be implemented using available air quality and emissions information. The new LUR modeling approach can account for changes in future vehicle fleet composition that can be utilised by professionals such as planners and engineers to compare the performance of proposed transport policies, mitigation measures, and transport construction projects during the planning phase.

The new modelling approach involved including additional data describing the national distribution of vehicle characteristics, including vehicle fleet breakdown, Euro Classifications and fuel types. These data were introduced to further strengthen the representation of local measured concentrations by the model and to enable the analysis of mitigation strategies that reduce emissions in specific locations, or from specific classes of vehicle.

The original land-use regression model developed by Naughton et al. (2018) considered only Annual Average Daily Traffic (AADT) flows within the Inverse Distance Weighted Vehicles Kilometres Travelled (IDWVKT) variable, $P_1$. The AADT is the average
daily total flow in both directions passing through a point on a route, based on a full calendar year (Transport Infrastructure Ireland, 2022).

When using AADT flow data, equation 2 becomes

$$P_1 = IDWVK = \sum_{b=1}^{B} \frac{1}{d_b} \sum_{r=1}^{R} AADT_r \cdot L_r$$  \hspace{1cm} (3)

in which $AADT_r$ is the Annual Average Daily Traffic flow on-road link $r$ of length $L_r$ with $R$ road links in buffer zone $b$. Link-by-link values of AADT flow are available as outputs from transport network models. These AADT flow values include, but usually do not distinguish between, different vehicle types which have considerably different properties such as engine sizes, vehicle weights and varying levels of emissions. Locations with atypical vehicle type distributions will therefore be less accurately represented within the WS-LUR model.

Moreover, over recent decades, vehicle emission standards have led to considerable improvements in vehicle technology to reduce emissions, while transport policies have had major impacts on the fuel type breakdown of the Irish vehicle fleet (EU, 1991; Department of Transport, Tourism and Sport, 2019). Ongoing changes in the vehicle fleet composition since the period when the model was originally developed could gradually become a major weakness when utilising the original regression coefficients in an analysis of future or past time periods outside of the original study period, which was between 2010 and 2012.

To address these issues, the method used to define traffic emission effects in the WS-LUR concentration formula was improved. The improvement involved splitting the AADT element of the IDWVK variable into separate components for each vehicle type (i.e. cars (PCs), LGVs, HDVs, etc.) to allow the application of emission weightings in the IDWVK variable, i.e.

$$AADT_r = \sum_{k=1}^{N} N_{k,r}$$  \hspace{1cm} (4)
in which \(N_{k,r}\) is the number of vehicles in category \(k\) (Euro Class or vehicle type) on road link \(r\).

The emission weighting applied to each vehicle category \(k\) is determined by defining a unit reference vehicle to which \(\text{NO}_2\) emissions from all vehicle types and Euro Classifications could be compared. Since the model regression coefficients were calibrated for the original AADT-based IDWVK variable, in which all vehicles are considered equal, the reference unit vehicle is a vehicle that emits the average emission for the period being studied. The emission weighting of each vehicle type / Euro Class is defined relative to this reference unit vehicle. Hence, equations 2 and 4 are modified to obtain equations 5 and 7:

\[
P_1 = IDWVK = \sum_{b=1}^{B} \frac{1}{d_b} \sum_{r=1}^{R} EAA\text{DT}_r \cdot L_r
\]  
\(5\)

and

\[
EAA\text{DT}_r = \sum_{k=1}^{N} E_k N_{k,r}
\]  
\(6\)

in which \(EAA\text{DT}_r\) is the emission weighted Annual Average Daily Traffic and \(E_k\) is the emission weighting for vehicle category \(k\), calculated as

\[
E_k = \frac{e_k}{e_A}
\]  
\(7\)

where \(e_k\) is the average emission from vehicle type \(k\) in a study period and \(e_A\) is the average emission from all vehicles in the same study period.

The European Monitoring and Evaluation Programme (EMEP) / European Environment Agency (EEA) Air Pollutant Emission Inventory Guidebook (European Environment Agency, 2019) identifies the average \(\text{NOx}\) emission rate in grams/km for each vehicle type and Euro Class, including pre-Euro vehicles classes, as well as a \(\text{NO}_2\) fraction (f-
NO\textsubscript{2}), for each fuel type, which determines the amount of NO\textsubscript{2} emitted based on the quantity of NO\textsubscript{X} emitted. This information was used to determine the typical NO\textsubscript{2} emission rate for each type of vehicle, which was then divided by the all-vehicle average emission rate during the time period being studied to determine the NO\textsubscript{2} emission weighting, \(E_k\), for the vehicle type. The Irish Bulletin of Vehicle and Driver Statistics e.g. Department of Transport, Tourism and Sport (2019) is published annually and collates data relating to the entire Irish vehicle fleet, such as year of registration, unladen weight, engine capacity and fuel type. These data were employed to determine the Euro Classification breakdown of the national vehicle fleet.

### 3.3 Data

Data on land use, traffic and meteorological data are required to define the values of the predictor variables and wind direction fractions employed in the WS-LUR model. Data defining the conditions at and surrounding fixed air quality monitoring sites are initially required for the development and calibration of the WS-LUR model. Subsequently, equivalent data defining conditions at and around any receptor location at which the ambient NO\textsubscript{2} concentration is to be calculated are also required. Details on data sources and data analysis methods are provided in the Supplementary Information Document.

The breakdowns of the vehicle fleet for each study year 2016-18 and the original study period, 2010 – 2012, were calculated, and from these data the average NO\textsubscript{2} emitted by a vehicle in each time period was calculated. This average emission value was used to determine the NO\textsubscript{2} emission weighting of every vehicle type / Euro Class and alter the IDWVKT variable in the WS-LUR model formula. Details of the fuel type, unladen weights, engine capacities and year when first licensed are available in the Irish Bulletin of Vehicle and Driver Statistics to determine the Euro Class breakdown of each vehicle category (i.e. Passenger Cars, LGVs, HDVs, etc.) as shown in the flow diagram in Figure 2.
3.4  Model Implementation

The original WS-LUR model (Naughton et al., 2018) was developed to create national maps of nitrogen dioxide concentrations, but its methodology did not distinguish between the characteristics of traffic on different roads or in different years. In the research presented here, the WS-LUR model was enhanced to take into account variations in vehicle types and fuels to facilitate the investigation of potential mitigation strategies.

The model has two modes of operation. The first is an automatic calculation where the modeller selects from one of the pre-set years included in the model (2016 to 2018), and the model calculates NO$_2$ concentrations using the data saved in the background. The other mode is the manual entry method in which the modeller enters the required data for a specific location and the resultant concentration is calculated. This mode supports the calculation of future or past concentrations. The step-by-step process to calculate the NO$_2$ concentration at a location is shown in Figure 3.
Figure 3. Improved WS-LUR model input step-by-step process

<table>
<thead>
<tr>
<th>Step</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Selection of Study Period/Analysis Method</td>
</tr>
<tr>
<td>2</td>
<td>Selection of Location</td>
</tr>
<tr>
<td>3</td>
<td>Selection of Triangulation Accuracy</td>
</tr>
<tr>
<td>4</td>
<td>Data Input</td>
</tr>
<tr>
<td>5</td>
<td>Selection of Inverse Distance Weighted Vehicle Kilometres Travelled Calculation Method</td>
</tr>
<tr>
<td>6</td>
<td>Calculation of NO₂ Concentrations</td>
</tr>
</tbody>
</table>

Once all the above steps are completed, the NO₂ concentration at the specified location is calculated using the WS-LUR equation and a summary of the total NO₂ concentration and the contributions from each wind direction sector are output.
4. Results

In this section, the model is validated against the original model by Naughton et al. (2018). This is followed by testing of the model on two scenarios 1) the impact of COVID-19 travel restrictions on NO$_2$ levels and 2) how the removal of diesel vehicles from the public service fleet would impact on NO$_2$ levels.

4.1 Model Validation

To validate the model, the results achieved in the original study by Naughton et al. (2018) were compared with those obtained using the new model, which accounts for the vehicle fleet breakdown. This was done by calculating the IDWVKT variable so that overall traffic emissions in the study years analysed in this project (2016-18) match those in the original study years analysed by Naughton et al (2010-12). Tables 2 and 3 identify the NO$_2$ emission weighting and vehicle type breakdown for each vehicle type for the 2010-12 and 2018 respectively. The resultant weighting factor represents the overall weighting applied to the IDWVKT variable. As this value is equal to 1 in both cases, the overall weighting of the IDWVKT variable is equal in the original and new study years. This supports the direct comparison of the concentration results obtained with the original and new models while introducing sufficient detail to identify particular vehicle types which contribute to elevated concentrations at a given location.

Table 2. Original model period 2010 - 2012 model validation

<table>
<thead>
<tr>
<th>Vehicle Type</th>
<th>NO$_2$ Emission Weighting</th>
<th>Vehicle Type Breakdown</th>
<th>NO$_2$ Emission Weighting x Vehicle Type Breakdown</th>
</tr>
</thead>
<tbody>
<tr>
<td>Passenger cars</td>
<td>0.712</td>
<td>82.85%</td>
<td>0.590</td>
</tr>
<tr>
<td>Light Commercial</td>
<td>2.515</td>
<td>12.81%</td>
<td>0.322</td>
</tr>
<tr>
<td>Heavy Duty Vehicles</td>
<td>3.498</td>
<td>1.24%</td>
<td>0.043</td>
</tr>
</tbody>
</table>
Table 3. Original model period 2018 model validation

<table>
<thead>
<tr>
<th>Vehicle Type</th>
<th>NO\textsubscript{2} Emission Weighting</th>
<th>Vehicle Type Breakdown</th>
<th>NO\textsubscript{2} Emission Weighting x Vehicle Type Breakdown</th>
</tr>
</thead>
<tbody>
<tr>
<td>Passenger Cars</td>
<td>0.821</td>
<td>82.94%</td>
<td>0.681</td>
</tr>
<tr>
<td>Light Commercial Vehicles</td>
<td>2.170</td>
<td>12.53%</td>
<td>0.272</td>
</tr>
<tr>
<td>Heavy Duty Vehicles</td>
<td>1.568</td>
<td>1.48%</td>
<td>0.023</td>
</tr>
<tr>
<td>Large Public Service Vehicles</td>
<td>2.966</td>
<td>0.43%</td>
<td>0.013</td>
</tr>
<tr>
<td>Motorcycles</td>
<td>0.051</td>
<td>1.58%</td>
<td>0.001</td>
</tr>
<tr>
<td>Small Public Service Vehicles</td>
<td>1.191</td>
<td>0.84%</td>
<td>0.010</td>
</tr>
<tr>
<td>Electric Vehicles</td>
<td>0</td>
<td>0.19%</td>
<td>0</td>
</tr>
<tr>
<td>Resultant Weighting</td>
<td></td>
<td></td>
<td>1</td>
</tr>
</tbody>
</table>

4.1.1 Monitored and Modelled Concentration Comparisons

The annual average NO\textsubscript{2} concentrations measured at EPA monitoring stations were compared to the concentrations calculated using the original model methodology and the new model methodology. For the purposes of the comparisons below, two new model options were compared with the original model, as defined below.
• **Original Model – AADT Only:** This option calculates the IDWVKT values as per the original model methodology by Naughton et al. (2018) in which only the AADT value is considered and all vehicle types are weighted equally.

• **New Model A – Original Vehicle Type Composition:** This option calculates the IDWVKT values using the vehicle type breakdown and NO\(_2\) emission weightings for the 2010 to 2012 study period.

• **New Model B – Vehicle Type and Euro Classification:** This option calculates the IDWVKT values using the vehicle type breakdown and NO\(_2\) emission weightings for one of the 2016-18 pre-set years.

Figure 4 compares the measured concentrations and various modelling concentrations at each monitoring station for 2018. (Similar figures for 2017 and 2016 are presented in the Supplementary Information Document). The difference between the measured and modelled concentrations varies more between sites than between model versions. At most locations the two values are relatively close, but at some locations (Ballyfermot, Winetavern St and St John’s Road) the differences are large. LUR models assume that the relationship between concentrations and nearby sources is the same throughout the modelled domain, which in this case is the whole country. Where large differences between measured and modelled concentrations are observed, this implies that some other effects that are not well captured by the model are locally important. These can include atypical stationary and non-stationary emissions sources and localised meteorological conditions. The model is able to distinguish between locations that experience high or low NO\(_2\) concentrations, implying that it is a suitable tool for evaluating the ability of potential mitigation strategies to improve air quality at specific locations.

The concentrations for the Original Model methodology and the New Model B methodology were consistently similar throughout. This reflects the fact that the differences
between the models affect only one of the predictor variables (IDWVKT) and the emissions weightings have been calibrated to maintain overall consistency with the original model. This similarity in results validates the ability of the new model method to retain the accuracy of the original model while introducing the potential to analyse the traffic variables in more detail.

The similarity in results obtained with the original and new model methodologies in rural environments such as Seville Lodge, Emo Court and Kilkitt highlights that the composition of the vehicle fleet in these locations is close to the national average. In contrast, at a number of urban locations, such as St. John’s Road and Winetavern Street, higher concentration estimates are given by the new models, reflecting the much higher percentage of vehicles (LPSVs and HDVs) with above-average NO$_2$ emission rates. This indicates that the new model can be utilised to develop mitigation measures that target particular vehicle types, such as Euro Class restrictions on specific routes or migrating parts of the vehicle fleet to low emission vehicles.

The limitation in available traffic data, which was only available for the eastern region of Ireland, (Leinster and parts of Ulster) affected the scope of the model validation. Nevertheless, a complete set of 38 measurements obtained over a 3 year period at 16 different sites in this region were included in the comparative analysis. Measured concentrations were within the standard error range of the model for 50% of observations (19 measurements) using the original model methodology, 42.1% (16 measurements) using the New Model A methodology and 47.4% (18 measurements) using the New Model B methodology, as shown in Figure 4. These results do not suggest improved model accuracy but it is important that the modifications broadened the model's applicability (e.g. in mitigation analysis) without overall loss of model performance. In the original study, 68% of the measured concentrations were within the standard error range of the modelled concentrations.
However, even as the sample size has remained similar to the original study (16 locations compared to 15 locations) the locations have changed considerably with only 9 of the original 15 locations included in the 2016 to 2018 comparison. Moreover, 6 of these stations are located relatively close to each other in Dublin City, which limits analysis of spatial (environmental and meteorological) and temporal (traffic) variability in the study.

All model methodologies overestimated the NO$_2$ concentrations with two exceptions where the NO$_2$ concentrations are underestimated, both in 2018 at St. John’s Road and Ringsend. This was the first year the St. John’s Road monitoring site was active; further sampling at this site could confirm whether there are other factors, not captured by the model, which contribute to NO$_2$ concentrations. It should be noted that in 2019, the annual EU limit value for nitrogen dioxide at St. Johns Road was exceeded. The exceedance necessitated the preparation of the Dublin Region Air Quality Plan (Dublin City Council et al, 2021). The plan sets out 14 measures and a number of associated actions to address the exceedance of the NO$_2$ annual limit value.

The Ringsend monitoring site was only active in 2017 and 2018 during the study period. In 2017 the measured values agreed well with the modelled concentrations, whilst in 2018, they were marginally outside the standard error range. The meteorological conditions were very different in 2017 and 2018, with approximately 59% of the wind coming from a west / south-westerly direction and 16% from an east / south-easterly direction in 2017 compared to 45% west / south-westerly and 23.5% east/south-easterly proportions in 2018. This higher proportion of easterly winds in 2018 may have influenced the higher average concentration measured that year.

Spatial validation of the model was done by comparing locations in Dublin (2016-18) with those in the Rest of Ireland (2016-18) and temporal validation considered all locations in
each of the three years 2016, 2017 and 2018. The $R^2$ values for each are presented in Table 4. Some variation between the three years can be seen, probably as a result of different meteorological conditions, and the results confirm that the models perform better in rural conditions. However, there is almost no difference between the different versions of the model, indicating that the added capabilities of the new models can be employed without loss of accuracy in a range of settings.

**Table 4. Spatial and temporal validation**

<table>
<thead>
<tr>
<th></th>
<th>All Locations 2016</th>
<th>All Locations 2017</th>
<th>All Locations 2018</th>
<th>Dublin 2016-18</th>
<th>Rest of Ireland 2016-18</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Original Model</strong></td>
<td>0.922</td>
<td>0.861</td>
<td>0.624</td>
<td>0.357</td>
<td>0.841</td>
</tr>
<tr>
<td><strong>New Model A</strong></td>
<td>0.912</td>
<td>0.859</td>
<td>0.646</td>
<td>0.372</td>
<td>0.836</td>
</tr>
<tr>
<td><strong>New Model B</strong></td>
<td>0.917</td>
<td>0.86</td>
<td>0.633</td>
<td>0.36</td>
<td>0.84</td>
</tr>
</tbody>
</table>
Figure 4. 2018 Measured and modelled NO\textsubscript{2} concentrations with standard error high-low bars
The original model captured 78% of the spatial variability in NO\textsubscript{2} with a cross-validation R\textsuperscript{2} of 77.4% (Naughton et al., 2018). The cross-validation R\textsuperscript{2} was slightly lower when analysing the 2016 to 2018 measurements against the Original Model methodology at 75.44%, whilst the New Model A and New Model B methodologies were slightly better at 76.08% and 75.58% respectively, as shown in Figures 5.

![Figure 5. Measured vs modelled NO\textsubscript{2} concentrations (2016-18) for the original model, new Model A and new Model B methodologies.](image)

4.2 Impact of COVID-19 on NO\textsubscript{2} Levels

The improved WS-LUR model was also used to assess the impact of COVID-19 restrictions on NO\textsubscript{2} levels. Model data on traffic flows, commercial properties and meteorological conditions specific to the 1\textsuperscript{st} COVID lockdown period (28\textsuperscript{th} March 2020 to 17\textsuperscript{th} May 2020) were employed. The objectives were to:
1. Evaluate the capability of the WS-LUR model to calculate changes in ambient NO$_2$ concentrations for unique scenarios;

2. Model NO$_2$ concentrations at various locations during the 1$^{st}$ COVID lockdown period; (A comparison between modelled and measured concentrations for the 1$^{st}$ COVID lockdown period is included in the Supplementary Information document).

3. Characterise the meteorological, source locations and traffic conditions during the 1$^{st}$ COVID lockdown period;

4. Based on the performance of the model in this scenario, identify the effect of individual parameter changes experienced during the COVID lockdown on ambient NO$_2$ concentrations and the potential NO$_2$ reductions that can be achieved by mitigation measures that target these parameters.

4.2.1. **Comparison of Pre-COVID and COVID Lockdown Concentrations**

The modelled concentrations display greater differences between the pre-COVID and COVID scenarios than are observed in the measured concentrations, as shown in Figure 6. The overestimation of the change in concentrations at the majority of the stations could be due to a number of factors relating to the unique conditions being examined.

The commercial properties predictor variable mainly reflects the number of commercial properties surrounding a study location, but an unknown proportion of this variable captures the effects of the other variables, such as residential property numbers, which were not included as independent predictor variables as they did not satisfy the selection conditions. This suggests that for the COVID analysis a more accurate result would be obtained if the modelled reduction in commercial properties were less than the full reduction experienced during the lockdown period to allow for the enduring implicit contributions of other influences such as residential properties. This would lead to smaller
differences between modelled pre-COVID concentrations and modelled COVID concentrations, particularly in more urban environments.

Considering the temporary nature of the conditions being examined, it is relevant that during the COVID lockdown period weather conditions were significantly different to those experienced during the model calibration period and in the earlier study years of 2016-18. The predominant wind direction during the lockdown period was from an easterly direction, whilst during the other periods (including the pre-COVID period), the predominant wind direction was from a westerly / south-westerly direction. A pre-dominant westerly / south-westerly wind would transport fresh air in from the Atlantic Ocean whereas an easterly wind direction would transport air from continental Europe and the United Kingdom, which could contain higher concentrations of pollutants (Donnelly et al. 2019). As a consequence, the measured reductions in concentrations during the lockdown period are likely to have been smaller than would have been experienced with annual average wind directions. This indicates that the full impact of reduced local emissions is not evident in the measured difference values shown in Figure 6 due to a concurrent increase in background concentrations.

The WS-LUR model is not able to capture the effect of atypical short-term meteorological conditions on average background concentrations. An increase in background concentrations should be reflected in a greater value for the constant within the NO$_2$ concentration regression equation, but the calibration process provides just a single annual average value that is applied uniformly across the entire country. On the other hand, the WS-LUR model did capture the effects of the COVID period wind directions on influences of the predictor variables within the model as these were weighted based on the wind direction proportions experienced during the particular study period being analysed.
Figure 6. Measured and modelled differences between pre-COVID and COVID lockdown scenario

4.3 Removal of Diesel Vehicles from the Public Service Fleet.

From 2010 to 2019, the percentage of small public service vehicles fuelled by diesel (which was already high) increased even further from 59% in 2010 to 82% in 2019 (Department of Transport, Tourism and Sport, 2020), while approximately 99.9% of all large public service vehicles throughout this period were diesel-fuelled. As diesel vehicles are associated with higher NO₂ emissions, cities in the United Kingdom have introduced all-electric powered large public service vehicles to improve air quality (Department for Transport, 2021). In Ireland, guidance on vehicles considered acceptable for use as small
public service vehicles specify limits on size and age (National Transport Authority, 2021). These age limits provide a basis for determining the potential timeline required for a fleet-replacement mitigation measure to be fully implemented.

The specific mitigation measure considered is a change from diesel to greener fuelled options within the small (cars) and large (buses) public service vehicles fleets. As a larger proportion of public service vehicles operate in the major cities, changes to newer Euro Class vehicles or greener fuel options would improve air quality in urban areas most.

The pre-mitigation scenario considered the 2019 national vehicle fleet. In the post-mitigation scenario the 2019 vehicle breakdown was altered to remove diesel-fuelled vehicles from the small and large public service vehicle fleets and replace them with electric-powered vehicles. This change was reflected in small or large changes in the IDWVKT values for different roads, depending on the number of public service vehicles travelling on each road, and proportionate changes in modelled NO$_2$ concentrations in the vicinity of these roads.

Figure 7 shows the predicted changes in NO$_2$ concentrations at monitoring station locations within Dublin due to the mitigation measure. All locations display improved air quality, with the largest reductions in ambient NO$_2$ concentrations experienced at locations within the Canal Corden (Pearse Street, Ringsend, St. John’s Road and Winetavern Street) where NO$_2$ reductions ranging from 1.0 to 1.8 μg / m$^3$ were observed. The Five Cities Demand Management Study showed that higher proportions of public service vehicles operate in these areas compared with the national average (Department of Transport, Tourism and Sport & Systra, 2020a and Department of Transport, Tourism and Sport & Systra, 2020b). Areas further away from the city centre experienced smaller reductions between 0.3 and 0.9 μg / m$^3$, in line with the smaller number of public service vehicles in use in rural and suburban locations.
Figure 7. Modelled concentration changes due to removal of diesel vehicles from the small and large public service vehicle fleet

5. Discussion

The research presented in this paper has developed a new method for including details of vehicle fleet breakdown within the traffic variable of an LUR model. This method can be utilised in any LUR model which has determined that total AADT flows or flows from a particular vehicle type are linked to ambient concentrations but which do not account explicitly for the contributions of all of the different vehicle types. The new method retains the link to the original WS-LUR model and ensures that the model can account for every vehicle type flow when estimating pollutant concentrations. This method also identifies the
potential to modify an existing LUR model by altering an existing variable or introducing a new variable without affecting the existing regression coefficients. This approach offers a number of benefits when estimating NO$_2$ concentrations at a study location.

1. It provides a more accurate representation of the traffic flows surrounding a study location as it segregates the AADT flows by vehicle type, and weights each vehicle type based on its emission rate relative to that of the average vehicle in the fleet.

2. It preserves model calibration by adjusting the value of the predictor variable representing traffic flows while leaving the values of the regression coefficients unchanged. An alternative approach involving the determination of separate regression coefficients for each vehicle type is unlikely to be supported by available data. The proposed method is transferrable to other LUR models which employ a traffic variable with limited detail on vehicle fleet composition.

3. The vehicle type weightings used to calculate the traffic flow predictor variable can be chosen in different ways depending on the modelling objectives. Hence, they can be chosen to preserve the average contribution of traffic emissions to that at the time of model calibration, or to capture trends in this contribution with time due to technological and regulatory changes, or to investigate proposed mitigation strategies.

4. LUR models represent average conditions throughout the model domain. They can be used to identify the relative contributions of different source types, including traffic emissions, at different locations. The proposed method extends this analytical capability to distinguish between the contributions of different vehicle types, and allows locations with atypical traffic conditions to be more accurately represented.

5. The proposed method allows the expected change in concentrations due to local changes in traffic flows and/or fleet composition to be calculated. This was tested for
the unique conditions of the COVID lockdown period. More practically, it allows the potential reductions in ambient concentrations due to proposed mitigation strategies to be quantified and compared.

6. The proposed method simplifies the process of analysing pre- and post-mitigation measure scenarios. Once pre-mitigation measure variable values are included in the model only the traffic variable values affected by the mitigation measure need to be revised when assessing a post-mitigation measure scenario, which significantly reduces pre-processing time and data errors.

The focused study on the unique situation presented by the COVID lockdown period compared modelled and measured concentrations to evaluate model performance under abnormal conditions. This provided insight into the use of the model to estimate concentrations for broad variations in source and meteorological conditions. The analysis also provided an opportunity to assess the accuracy of the model to predict changes in concentrations due to changes in individual predictor variables and to understand the magnitude of the potential reduction in ambient NO₂ concentrations associated with substantial changes in traffic conditions.

The development of the WS-LUR model to include vehicle fleet breakdown details supports the assessment of the impacts of mitigation measures that target particular vehicle types. This model can be a useful tool for planners and engineers in the transport sector, particularly during the planning phase of projects, as the LUR model now has the capability to assess air pollution changes due to the implementation of mitigation measures or due to the construction / expansion of the transport network. The assessment of the removal of diesel public service vehicles focused on modelled locations within the Greater Dublin Area as representative of the urban and suburban locations where the majority of these vehicle types
operate. This mitigation measure achieved reductions in the range of 0.3 μg/m$^3$ and 1.9 μg/m$^3$ in NO$_2$ concentrations. The St. John’s Road modelled location experienced the largest reduction in the mitigation measure analysis, while the other city centre locations (e.g. Winetavern Street and Pearse Street) were amongst the five largest concentration reductions. The analysis demonstrates the model’s capability in assessing mitigation measures that would be useful in ensuring compliance with the Directive 2008/50/EC limits (EU, 2008) and achieving the reduced limit set by the World Health Organisation (WHO, 2021). The removal of diesel fuelled vehicles from the public service fleets has the potential to considerably reduce NO$_2$ concentrations across urban areas where concentrations are typically higher and population exposure is greater, whilst smaller reductions are experienced in suburban and rural areas where public service fleet numbers and the distance travelled by these vehicles are less. The analysis shows that although public service vehicles represent only 1.2% of the national vehicle fleet, significant reductions in NO$_2$ concentrations could be achieved by promoting greener fuel options within a small proportion of the vehicle fleet that are predominantly business owned or government funded vehicles.

6. **Conclusions**

The research presented in the paper developed an enhanced WS-LUR model that can account for vehicle fleet breakdown within the traffic variable to estimate NO$_2$ concentrations at any location in Ireland. The weighting method employed offers opportunities to assess mitigation strategies to support policy development. The method is transferrable to any LUR model that includes vehicle distance travelled as a predictor variable. The emission weighting can be chosen to match average vehicle emissions at the time of the calibration or future average emissions due to technological and regulatory developments. The research demonstrates the importance of accounting for vehicle fleet breakdown on routes with larger numbers of HGVs and large public service vehicles that generate higher levels of NO$_2$ than
other routes. An analysis was carried out on the accuracy of the model, by estimating concentrations for a unique scenario / environment, specifically the 1st COVID lockdown.

The model offers the potential to policy makers to test NO₂ mitigation measures such as Low Emission Zones, the creation of industrial hubs, increased levels of working from home, etc. The paper demonstrates the capability of testing such mitigation measures by analysing the removal of diesel vehicles from the public service fleet resulting in NO₂ reductions in the range of 0.3 μg/m³ to 1.9 μg/m³.

The first limitation of the work is that potential predictor variables that make only minor contributions to NO₂ concentrations, or that make significant contributions to NO₂ concentrations at only a small number of locations, are excluded from the model. However, this is a potential weakness in all LUR models developed according to the recommended procedure that seeks to identify variables making a significant contribution to ambient concentrations (Beelen et al., 2013, Naughton, et al., 2018). The second limitation is that as transport model outputs for the COVID lockdown period were not available, traffic count data sourced from the Transport Infrastructure Ireland traffic counter database (Transport Infrastructure Ireland, 2022) were employed. These counters continuously record traffic flows on all major routes and provided observed data for the entire COVID lockdown period considered in this analysis. A corresponding full dataset of minor routes flows was not available for this period.

The research described makes an important scientific contribution by proposing and demonstrating a different method to the computation of predictor variables in LUR models. The proposed method makes novel use of standardised emission data, obtained independently from the ambient concentration data originally employed to develop the LUR model, while preserving the regression model calibration. This potential benefits offered by this approach
of integrating multiple datasets to strengthen the capability of LUR modelling can be explored in future research projects by the transport and air quality modelling communities.

Acknowledgements

The authors would like to thank the Irish Environmental Protection Agency for funding the research.

References


Department of Transport, Tourism and Sport & Systra. 2020b. Technical Note: Fleet Profile in Dublin City. Department of Transport, Tourism and Sport, Dublin, Ireland.


EPA, 2019b. EPA GIS Department, Environmental Protection Agency, Johnstown Castle, Ireland.


European Space Agency. 2020. Detecting Pollution from Individual Ships from Space.  
https://www.esa.int/Applications/Observing_the_Earth/Copernicus/Sentinel-5P/Detecting_pollution_from_individual_ships_from_space [Accessed: 28 June, 2022].


National Transport Authority. 2020. 2016 Eastern Region National Transport Model.

National Transport Authority, Dublin, Ireland.


https://www.who.int/news-room/fact-sheets/detail/ambient-(outdoor)-air-quality-and-
health#:~:text=Guideline%20values,-
NO&text=The%20current%20WHO%20guideline%20value.effects%20of%20gaseous%20nitrogen%20dioxide [Accessed 28 June, 2022]