Assessing the Potential of Land Use Modification to Mitigate Ambient NO$_2$ and Its Consequences for Respiratory Health

Meenakshi Rao [1], Linda A. George [1,*], Vivek Shandas [2] and Todd N. Rosenstiel [3]

[1] School of the Environment, Portland State University, Portland, OR 97207, USA; mrao@pdx.edu
[2] Nohad A. Toulan School of Urban Studies and Planning, Portland State University,
Portland, OR 97207, USA; vshandas@pdx.edu
[3] Department of Biology, Portland State University, Portland, OR 97207, USA; rosensti@pdx.edu
[*] Correspondence: georgel@pdx.edu; Tel.: +1-503-725-3861

Abstract: Understanding how local land use and land cover (LULC) shapes intra-urban concentrations of atmospheric pollutants—and thus human health—is a key component in designing healthier cities. Here, NO$_2$ is modeled based on spatially dense summer and winter NO$_2$ observations in Portland-Hillsboro-Vancouver (USA), and the spatial variation of NO$_2$ with LULC investigated using random forest, an ensemble data learning technique. The NO$_2$ random forest model, together with BenMAP, is further used to develop a better understanding of the relationship among LULC, ambient NO$_2$ and respiratory health. The impact of land use modifications on ambient NO$_2$, and consequently on respiratory health, is also investigated using a sensitivity analysis. We find that NO$_2$ associated with roadways and tree-canopied areas may be affecting annual incidence rates of asthma exacerbation in 4–12 year olds by +3000 per 100,000 and −1400 per 100,000, respectively. Our model shows that increasing local tree canopy by 5% may reduce local incidences rates of asthma exacerbation by 6%, indicating that targeted local tree-planting efforts may have a substantial impact on reducing city-wide incidence of respiratory distress. Our findings demonstrate the utility of random forest modeling in evaluating LULC modifications for enhanced respiratory health.

Keywords: nitrogen dioxide; air pollution; land use regression; random forest; health; BenMAP

1. Introduction

Cities have increasingly become the nexus of air pollution due to emissions from anthropogenic activities within the cities [1]. The dispersion of these emissions is not uniform across the urban landscape leading to high spatial variation in ambient air pollution concentrations [2,3]. Local concentrations of air pollutants are affected not just by the strength of local emissions and formation of secondary air pollutant through atmospheric chemistry, but also by land use and land cover (LULC) features which influence air flow and hence transport of these pollutants. Urban zoning and other siting policies affect the proximity of a location to emission sources, while the quantity of emissions associated with each LULC category is controlled through permitting, policies, and urban planning. Atmospheric scientists capture the spatial variation of atmospheric pollutants within a city or region using a wide variety of techniques [2,4] such as land use regression (LUR) [5–7]; geo-statistical techniques including kriging [8–10]; modeling simulations such as dispersion models [11–13] and atmospheric chemistry & transport models (ACTMs) [14,15]; computational fluid dynamics models (CFDs) [16]; and other emerging techniques [17–20]. All these modeling techniques explicitly include LULC and terrain data as model inputs, acknowledging the influence of LULC on local ambient

concentrations of air pollutants. In our rapidly urbanizing world, where ambient air pollution is recognized as a leading environmental health risk [1], the role of land use in modulating ambient concentrations of air pollution gives rise to a timely and relevant question: to what extent can urban land use be managed to decrease local air pollution, and consequently, its impact on human health?

Globally, managing land use to reduce greenhouse gas emissions or air pollution is not a new concept. Hong Kong requires an air ventilation assessment for all publicly funded construction to mitigate the stagnant wind conditions that could allow accumulation of air pollutants or air-borne diseases like SARS [21] in its urban canyons. In California, schools are required to be located more than a quarter mile away from sources of potentially hazardous air pollutants [22]. Many cities actively seek to reduce vehicle miles traveled (VMT) to reduce CO_2 emissions and manage regional air quality [23–25]. However, even though LULC has been regulated to reduce or avoid exposure to emissions, the role of local LULC modifications on ambient concentrations through dispersion or deposition of air pollutants, has—with few exceptions [26–30]—not been systematically investigated.

Assessing the impact of local land use on air pollution presents some unique challenges. Simulation models, although theoretically ideal for studying the effect of LULC on ambient air pollutions, suffer from a lack of validated emission inventories at a fine spatial scale (~1 km) [31,32]. Additionally, ACTMs require a very high level of expertise and computational power to be run at this fine intra-urban resolution. Geo-statistical models such as LUR, on the other hand, have been successfully used to capture the fine spatial scale variation of urban air pollutants in cities across the world [33–37]. However, though these statistical regression techniques are excellent at developing predictive models, the correlated nature of most land use variables makes it difficult to use these same techniques to isolate the impact of individual LULC categories on local ambient air pollution concentrations [38,39]. Nevertheless, with rising urbanization and the growing recognition of the negative economic and health impacts of air pollution [40,41], there exists an urgent demand for developing techniques that can help us evaluate the role of LULC and LULC modifications on urban air pollutants, especially techniques that do not require a high level of expertise or computational power, and thus are readily accessible to all stakeholders engaged in managing or minimizing local concentrations of ambient air pollutants [42,43].

In this paper we use a technique called random forest that can be used to investigate the association of air pollution with individual LULC categories. Random forest [44,45] is a powerful ensemble-based data mining technique that makes minimal assumptions about the independence or underlying distribution of the predictor variables. It has been widely adopted and successfully applied in many domains, including bioinformatics and medical research [46–49], land use classification [50,51] and ecological modeling [52,53], and is being increasingly applied in the field of for air pollution as well [19,20,54]. The strengths of random forest lie in that fact that it makes minimal assumptions about the underlying data distribution or correlation of predictors, has the ability to deal with noisy and missing data, and can deal with the "large number of predictors, small number of observations" situation typical of urban air quality and other environmental studies in general. Further, as new environmental sensor technologies are more widely deployed [55,56] , providing streams of intra-urban observational data, the need for new methodologies to analyze the flood of incoming data and make it accessible and informative to scientists, planners and the public alike, will be intensified.

We exploit the ability of random forest to handle a large number of correlated predictors in order to examine the response of the US criteria pollutant nitrogen dioxide (NO_2) [57] to LULC and LULC modifications in the Portland-Hillsboro-Vancouver metropolitan region. We study NO_2 as it is one of the more easily measured of the US criteria pollutants. It is also a strong marker of anthropogenic combustion-related pollution, and a precursor to two other criteria pollutants, ozone and fine particulate matter [58,59]. In this study, we first use random forest to develop a model of NO_2, fitting NO_2 observations with LULC variables. Since random forest is not yet widely used to model high spatial resolution urban air pollution, we also assess the performance of the land use random forest (LURF) model. Leveraging the ability of random forest to handle correlated predictors,

Int. J. Environ. Res. Public Health **2017**, *14*, 750

3 of 19

we use the LURF model to investigate the association of ambient NO_2 concentrations with individual LULC categories. Further, by using the LURF NO_2 models in conjunction with BenMAP [60–62], a health impact assessment tool from the US Environmental Protection Agency (EPA), we estimate the relative health impacts associated with each LULC category through its association with NO_2. Finally, we investigate the potential of LULC modifications as a strategy for mitigating the health impacts of NO_2 by undertaking a sensitivity analysis, examining the change in NO_2 associated with the systematic variation of selected LULC categories, and using BenMAP to estimate the relative health benefits arising from the changes in NO_2 associated with these LULC modifications.

2. Materials and Methods

Our study area is the Portland-Hillsboro-Vancouver urban area, a mid-size metropolitan area located in the US Pacific northwest. Just over 1.8 million people reside within the study area, which encompasses 2350 km^2. This area has diverse terrain—two rivers, mountains, and, like other urban areas, a wide mix of current land use. Within the study area, 6% of land area is high density development, 7.5% is developed open space, and 13% is forested area, based on the 2011 National Land Cover Database (NLCD) [63] and its land use categories (Figure 1, Table S1). According to the National Emissions Inventory (NEI) 2011 data, there are only three facilities in the study area permitted to emit >500 tons NO_x annually. Annual average daily traffic (AADT) for the freeways and state highways in the area ranges from 169,500 on I-84 to 610 on an inner city access ramp [64]. A brief overview of our methodology is provided below, with more details in the following subsections.
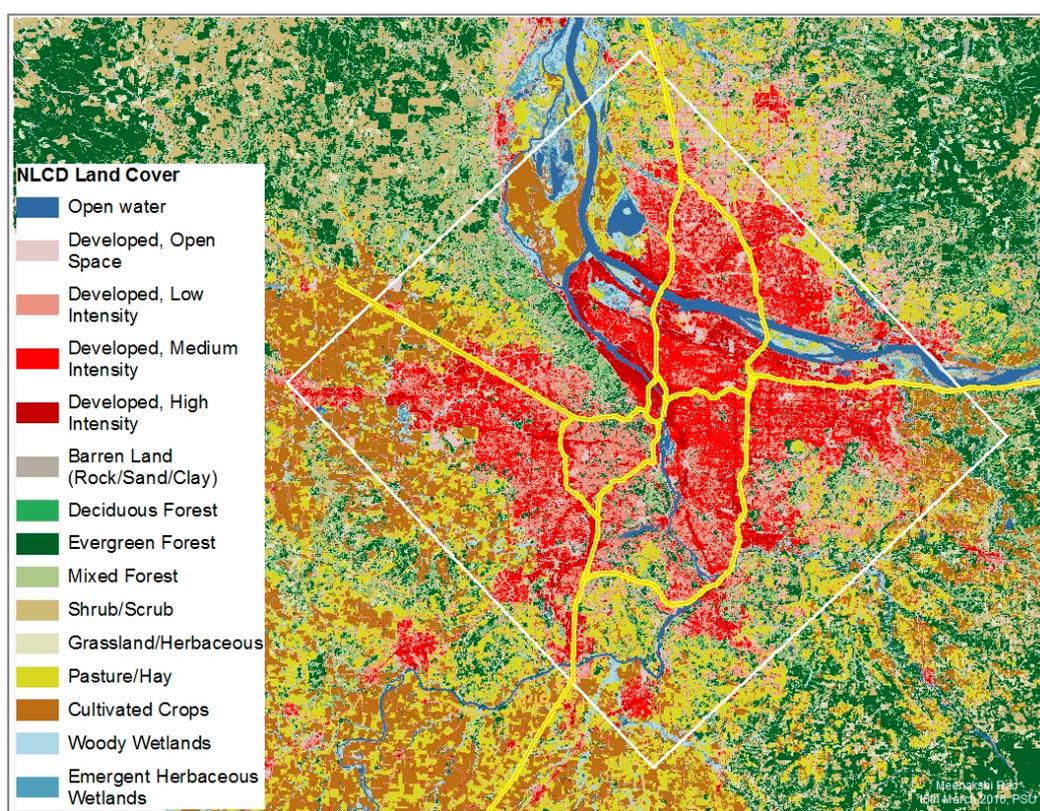

Figure 1. Land use in the Portland-Hillsboro-Vancouver area (based on NLCD 2011).

NO_2 was measured over two 2-week long field campaigns using relatively low cost passive chemical samplers made by Ogawa Co. (Pompano Beach, FL, USA). NO_2 was sampled once during summer (22 August–6 September 2013) at 174 sites in the Portland-Hillsboro-Vancouver area. Due to the logistical and resource intensity of resampling, a randomly selected subset of 82 sites was sampled

again in winter (13–27 February 2014) (Figure 2). Sites were chosen to capture the effect of roads, railroads and vegetation on ambient NO_2.The average measured NO_2 in summer was 11 ppb, with observed values ranging from 4 to 23 ppb, while the average measured winter NO_2 was 13 ppb, with observed values ranging from 3 to 29 ppb. Additional details about the field campaign can be found in Rao et al. [26,65].
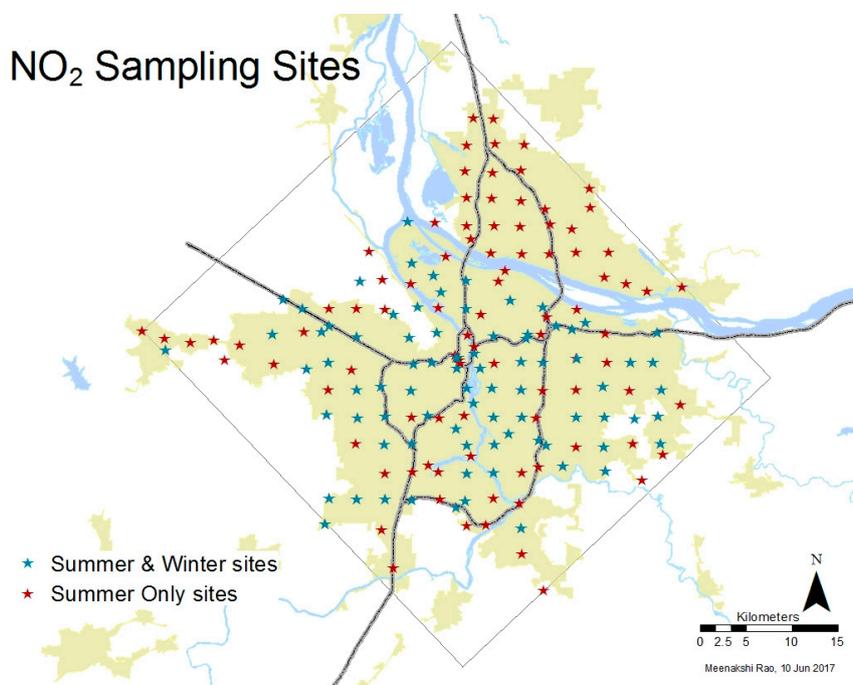

Figure 2. Summer and winter observation sites, with minimum bounding rectangle and urbanized area footprint. Blue stars represent sites monitored in summer and winter, red stars represent sites monitored in summer only.

Datasets described in Table 1 were used to determine LULC variables. Care was taken to find datasets that were closest to the period of the field campaigns. Our analysis followed the steps below:

1. Conditional inference random forest was used to develop summer and winter NO_2 models based on the NO_2. These observationally-based summer and winter LURF models of NO_2 were evaluated using both statistical measures as well as by comparison with LUR models, developed utilizing the same LULC variables.

2. A high spatial resolution, annual average NO_2 model (required to estimate health impact using BenMAP) was developed by applying the summer and winter LURF models to a grid of points 200 m apart covering the study area, then averaging the summer and winter NO_2 predictions at each grid point.

3. Next, the NO_2 associated with each individual land use category was estimated by applying the summer and winter LURF models to modified LULC variables, with the summer and winter NO_2 values being averaged to estimate annual average NO_2 for the 200 m grid. Further, BenMAP [56], the health benefits assessment tool from the US EPA, was used to assess the respiratory impact of each LULC category. The health impact is assessed as the change in incidence in health outcomes arising from the change in NO_2 between the NO_2 in step 2 and in step 3.

4. Finally, after we had identified the LULC categories that have the greatest impact on ambient NO_2 (from step 3) and identified ones that are amenable to change, a sensitivity analysis was undertaken in which these LULC categories were modified in the LURF model, and the change in NO_2 and the consequent respiratory health impact arising from each modification estimated.

These analyses are described in greater detail below. All spatial analyses were done in ESRI's ArcMAP (version 10.2, ESRI, Redlands, CA, USA), health impact analyses in BenMAP (version 4.0.35 US EPA, Research Triangle Park, NC, USA), and the statistical analyses in the open source statistical program R version (version 3.1.1, R Foundation for Statistical Computing, Vienna, Austria) (10 July 2014) [61,62].

Table 1. Land use/land cover categories used in analysis, data source, and spatial resolution.

Land Use/Land Cover	Data Source
Housing	US Census Bureau, 2010 (block level)
Land cover classes (developed open space, high intensity development, trees, shrub/scrub, grassland, pasture, cultivated crops)	National Land Cover Database (NLCD), USGS, 2011 (30 m)
Permitted NO_2 emissions	National Emissions Inventory, EPA, 2011 (point sources)
Elevation	USGS, 1/3 arc-second
AADT	NHPN (2010)
Roads (primary, secondary and local)	US Census Bureau, Tiger/Line (2013)
Latitude & Longitude	Google Earth, ArcMAP

2.1. Developing the LURF Models

2.1.1. Extracting Land Use Variables

LULC categories to be used as predictors in development of the LURF (and LUR) models were chosen either because they were known strong proxies for NO_2 (e.g., freeways) or identified based on a literature review and our prior campaigns in the Portland area [26,33–35,66–68]. Table 1 lists the LULC data sets used, the data source, and the spatial resolution of the data. LULC variables were extracted in 12 buffers, ranging from 100 to 1200 m in radius (in 100 m increments) for each land use category, at each site (174 sites, as the winter sites were a subset of the summer sites). Latitude, longitude and elevation were also associated with each site. The NLCD categories deciduous, evergreen and mixed forest were added together to create a "trees" category. In all, ~200 land use variables were associated with each site. A randomly selected 25% of observations (42/174 for summer, 20/82 for winter) were set aside as a "validation" data set for model evaluation prior to the start of model development to enable a hold-out validation assessment for the LURF models. All model development was subsequently done on the remaining 75% "training" data set.

2.1.2. Developing the LURF Model

Random Forest

Random forest is an ensemble statistical learning method based on regression trees. Regression trees [44] divide the p-dimensional predictor space into p-dimensional rectangles, such that the total of the residual sum of squares over all the rectangles is minimized. The prediction for any set of predictors P_i is the average of all observations that fall in the rectangle containing P_i. Regression trees tend to over-fit the training data, resulting in large variance, and hence potentially large prediction errors on unseen data. To address this issue, Brieman [45] developed the random forest methodology in which an ensemble of regression trees is created using bagging, that is by taking repeated samples with replacement from the training data set. Further, at each node for each tree in the forest, only a random subset of variables is considered for splitting, which results in decorrelated trees. Predictions are the average over all predictions for all trees in the forest for which the sample is out-of-bag. Strobl et al. [69] have further refined the methodology by using a conditional permutation scheme

that corrects for the inflated variable importance of correlated predictors in random forest [70], which they call conditional inference random forest. All random forest model development for this study was done using conditional inference random forests as implemented in the "party" package version 1.0–23 [70–72] in R [73].

Developing the LURF and LUR Models

Summer and winter random forest NO$_2$ models were developed using the ~200 LULC variables as predictors, in two phases. In the first phase, a conditional inference random forest, using the "party" package [70–72] in R [73], was used to identify the buffer size that was the most important predictor within each LULC category. Using only the most important buffer size for each land use category reduced the number of potential predictor variables from ~200 to ~20, and made the final model more interpretable. In the second phase, we again used conditional inference random forests, now with the reduced predictor set containing one buffer size for each LULC category, together with the point features latitude, longitude and elevation, to develop the observationally-based NO$_2$ LURF models for summer and winter. Random forest models take a random seed and two hyper-parameters, namely number of regression trees to include in the random forest (ntree), and the number of variables to consider for a split at each node in each regression trees in the forest (mtry). We systematically explored the ntree (500, 1000, 2000, 3000, 4000, 5000) × mtry (1, 2, 3, 4, 5, 6, 7, 8) space for a range of random seeds to identify a robust LURF model for each season.

The seasonal LUR models were developed used the same set of variables as the LURF seasonal models. We used a correlational matrix to narrow the variables to a smaller, less correlated subset. We next used stepwise regression, AIC, and VIF to identify four models, then used k-fold (k = 6) to identify the final LUR model for summer and winter.

These observationally-based summer and winter random forest models were each applied to points on a 200 m grid covering the study area. The 200 m resolution was chosen as a balance between computational time and intensity of extracting the LULC variables at each grid point, and the number of grid points required for higher spatial resolution. These seasonal fine-spatial scale NO$_2$ models for the Portland-Hillsboro-Vancouver area were then averaged to develop the annual average NO$_2$ LURF model.

2.2. Assessing the Performance of the LURF Model

Performance of the summer and winter LURF models was assessed using statistical performance measures for both model fit and predictive ability on unseen data. The statistical performance metrics used for model assessment are goodness of fit (R^2); normalized mean bias (=1/N * Σ {[modeled (NO$_2$) − observed(NO$_2$)]/[observed(NO$_2$)]}); and normalized mean error (=1/N * Σ {abs[modeled (NO$_2$) − observed(NO$_2$)]/[observed(NO$_2$)]}). The normalized mean bias is an estimate of systematic over- or under-estimation of the LURF models as compared to the observations, while the normalized mean error is the estimate of the average difference in the NO$_2$ predicted by the models and observations.

The predictive ability of the summer and winter LURF models was gauged by computing the root mean square error (RMSE) of the NO$_2$ predicted for the validation data with respect to the observations. Since the validation data sets are not used in model development, the validation RMSEs provide a good estimation of model performance on unseen data. The LURF validation data set RMSEs were further compared with RMSEs reported in the literature for LUR models, as well as the validation RMSEs of the LUR models developed using the same training, validation, and land use data sets as the LURF model.

2.3. Association of Current LULC, Ambient NO$_2$, and Respiratory Health

We estimated the influence of each LULC category on ambient NO$_2$ concentrations with a simple analysis: each land use category under consideration was set to zero over the entire study area, while keeping the remaining land use variables unchanged. Summer and winter NO$_2$ predictions for

the 200 m grid were averaged to estimate the annual impact of each land use category. The difference in modeled NO_2 concentrations between the annual model and the model with the LULC category set to zero was used as an indicator of the NO_2 associated with that LULC category. In essence, we simulated the annual average NO_2 concentrations across the urban landscape under the assumption that each LULC category was replaced by an NO_2-neutral land use; estimating, for example, the annual average NO_2 concentrations if there were no trees in the Portland-Hillsboro-Vancouver area or if there were no traffic on the highways. Once we had estimated the NO_2 associated with each land use category, we estimated the respiratory health impacts associated with this LULC-associated NO_2 in terms of incidence rates and economic valuations using BenMAP. Table S2 lists the health impact functions and valuation methods from BenMAP that were used in this evaluation. Population within each 200 m grid cell in 2013 was estimated using Popgrid, an ancillary program to BenMAP, based on block-group level population from the 2000 US Census, projected to 2013.

2.4. Evaluating the Mitigation Potential of LULC Modifications

Once we had determined the LULC categories with the strongest association with NO_2, We undertook a sensitivity analysis to estimate the relative impact of modifications to these LULC categories on ambient NO_2, and consequently, on respiratory health. We focused on four LULC categories, namely vehicle miles traveled on freeways, tree canopy, high intensity development, and open development. These categories were chosen as they were shown to be associated with NO_2 and are amenable to planned change: many city climate action plans [23–25] already incorporate targets for VMT, tree canopy and impervious areas; and previous research has shown that they have a discernible impact on ambient NO_2 concentrations [26,74]. For all these categories, we considered changes of $\pm 2\%$, $\pm 5\%$ and $\pm 10\%$ to the LULC feature. For tree canopy, high intensity and open development, the percentage change is based on the buffer size, so that a 2% increase results in an increase even at grid points that currently didn't have any of these land covers in their vicinity. In case of an increase, other LULC features in the vicinity of the point were proportionately decreased; while in the case of a decrease, the other LULC features were proportionately increased. However, care was taken that the modified land use did not go below 0% or above 100%. Health benefits (or dis-benefits), arising from the NO_2 changes associated with each LULC modification in the sensitivity analysis were estimated using BenMAP.
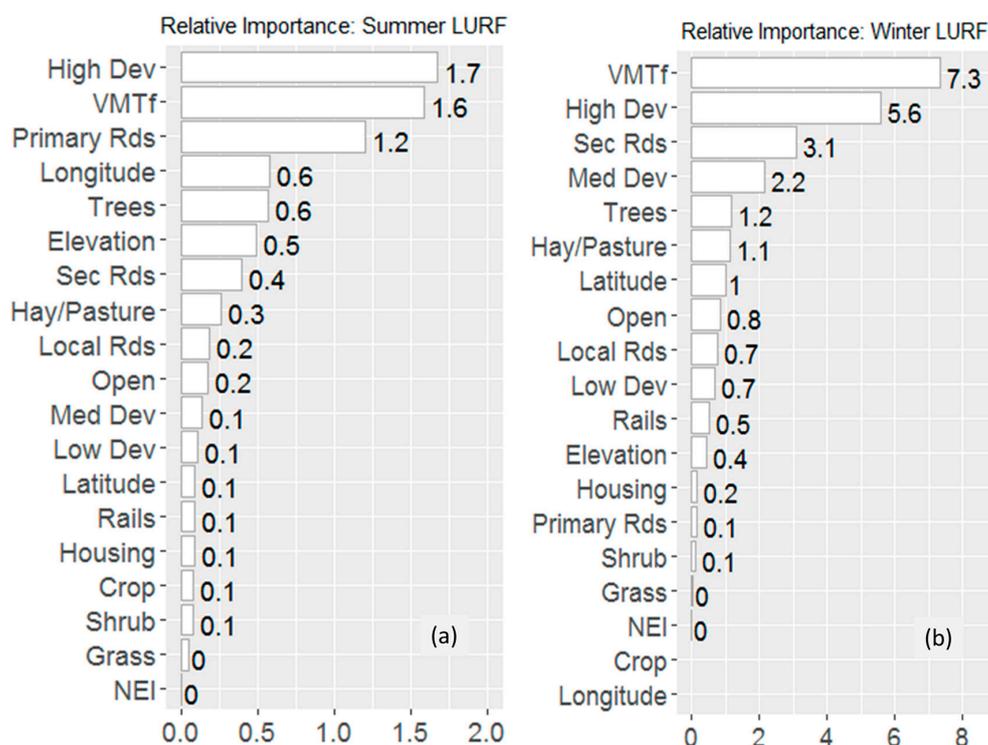
3. Results

3.1. Assessing the Performance of the LURF Models

The relative importance of the LULC predictors for the summer and winter LURF models can be found in Figure 3; and the LURF-derived 200 m-resolution map of annual average NO_2 in the Portland-Hillsboro-Vancouver area can be seen in Figure 4. Based on statistical measures using hold out validation, the summer and winter random forest models perform well with an R^2 of 0.80 and 0.83 respectively, indicating that a high degree of variance is captured by the models [7]. Both summer and winter LURF models show non-zero normalized mean bias and normalized mean error: the summer and winter LURF models show mean biases of 9% and 12% respectively; and mean errors of 20% and 24%. Thus, the LURF models systematically overestimate NO_2 concentrations as compared to the observations (Table 2). These findings are consistent with 10-fold validation as well (Table S3).

Table 2. Performance metrics for the summer and winter LURF and LUR models using hold out validation.

Season and Model	Goodness of Fit	Model Bias		Prediction Error	
	Adj R^2	Normalized Mean Bias	Normalized Mean Error	Validation MAE (NO$_2$ ppb)	Validation RMSE (NO$_2$ ppb)
Summer					
LUR	0.75	5%	20%	2.3	2.8
LURF	0.80	9%	20%	2.0	2.4
Winter					
LUR	0.80	5%	18%	2.5	3.4
LURF	0.83	12%	24%	2.8	3.8

Figure 3. Relative importance of the LULC predictors (**a**) in the summer and (**b**) winter LURF models.

The predictive ability of the LURF models, as gauged by the validation data RMSEs, indicates that the summer and winter models, on average, predict NO$_2$ concentrations within 2.4 ppb of the measured NO$_2$ in summer, and 3.8 ppb in winter (the higher winter validation RMSE being consistent with fewer winter observations). These RMSEs are consistent with RMSEs for NO$_2$ LUR models in the literature, in fact lying towards the lower end of the reported range of 1.4–34 ppb [75]. The RMSEs for the summer and winter LURF models of 2.4 ppb and 3.8 ppb are also on par with the validation RMSEs of 2.8 and 3.4 for summer and winter LUR models developed using the same data sets (Table 2). However, it is important to note that both the LUR and LURF models overestimate the annual average concentration of NO$_2$ at the DEQ monitoring station: 13.3 ppb and 13.4 ppb respectively, as compared to the annual average of 9.4 ppb based on the DEQ observations.

The annual average NO$_2$ predicted by the LURF differs from the bias-corrected LURF model by 4% (Figure S1). We next compare the LURF and LUR models for the Portland-Hillsboro-Vancouver metropolitan area: these are is highly correlated, with a Pearson correlation coefficient of 0.85 (Figure S1). The relationship between LUR and LURF predicted NO$_2$ can be summarized by the best

fit regression line which has an adjusted R^2 of 0.72 (Figure S1), indicating an effective linear mapping from LUR predicted NO_2 to LURF predicted NO_2 concentrations. The annual NO_2 predictions using random forest show a systematic overestimation as compared to the LUR predictions, with a 13% mean bias. This bias can be explained by the difference between the regression and random forest methodologies. Regression determines the best-fit slopes or coefficients based on minimizing distance from a curve in space, while random forest (an ensemble of regression trees) works on similarity, assigning an outcome based on the average of "similar" observations. Thus, in a regression model, predicted values can lie outside the observed range of values, while in a random forest model, predicted values are restricted to lie within the observed range. This results in an overestimation by LURF, as compared to LUR, at the lower end, and an underestimation at the upper end of predictions, as can be clearly seen in Figure S1. The mean error is 22%; that is, on average there is a 22% difference between the NO_2 concentrations predicted by the LURF and LUR models (Table 2), a difference which is on par with the mean error of both the LUR and LURF models with respect to observations.

3.2. Association of Current LULC, Ambient NO_2, and Respiratory Health

Figure 4 shows the local variation in annual average NO_2 in the Portland-Hillsboro-Vancouver area, as well as the local spatial variation of NO_2 associated with three LULC categories: VMT*f*, high-intensity development, and trees. Modeled, annual average concentrations of NO_2 within the study area range from 7 to 19 ppb, well below the US EPA standard (53 ppb) [76], though near the World Health Organization standard (20 ppb) [77].
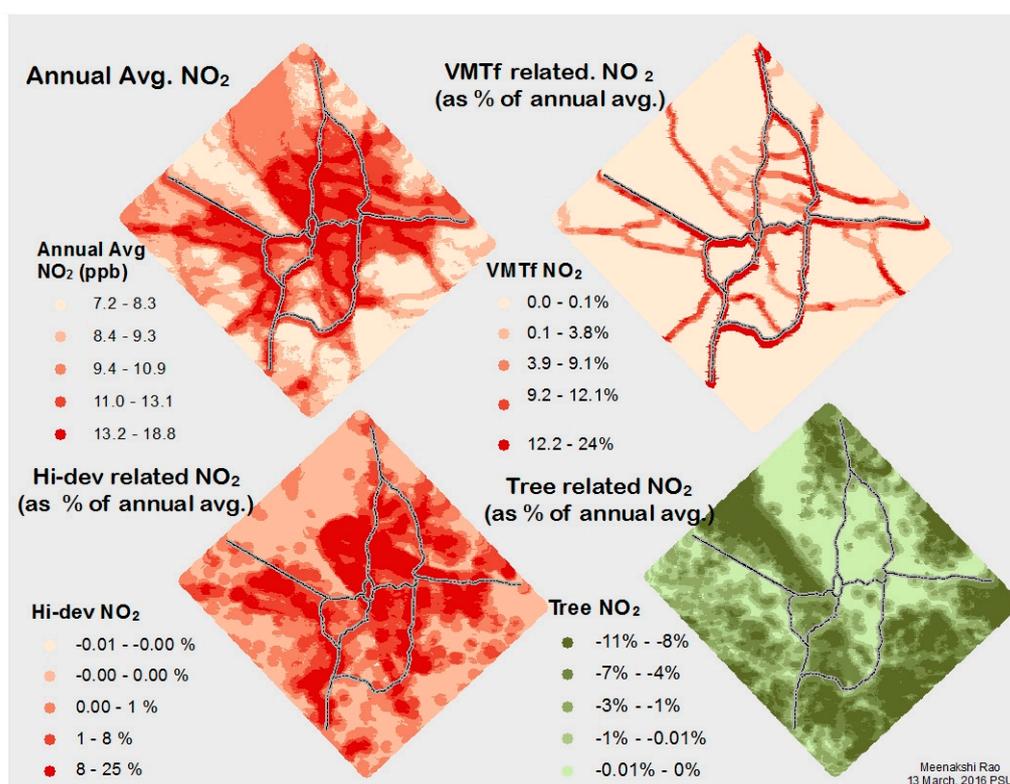

Figure 4. Percentage change in annual average NO_2 when a land use/land cover category is replaced with a neutral land use.

Of the 11 land use categories considered, eight (high intensity development, VMT*f*, primary, secondary & local roads, railroads, housing density, and permitted NO_2 emissions) contribute to increasing ambient NO_2. The remaining three land use categories (developed open spaces, trees, shrubs) are associated with decreasing ambient NO_2 concentrations (Table 3). The changes in ambient

NO_2 range from a decrease of 0.4 ppb associated with trees to an increase of 0.9 ppb associated with roadways, when averaged over the study area.

Table 3. Estimated association of land use and annual average NO_2 concentrations, averaged over the study area, as well as average land use values within the model buffers.

LULC Category	NO_2 (ppb) Associated with Land Use	Range NO_2 (ppb)	Typical LULC Values within Model Buffer	Range LULC Values within Model Buffer
Development, high-density	0.7	0–3.8	0.76 km^2	0–7.9 km^2
Roadways	0.9	0–6.2		
Vehicle Miles travelled on highways	0.4	0–3.5	133,916	0–1,329,013
Primary Roads	0.1	0–0.9	1.7 km	0–20 km
Secondary Roads	0.2	0–1.9	4.6 km	0–44 km
Local Roads	0.2	0–0.81	70 km	1.5–242 km
Railroads	0.1	0–0.6	4.3 km	0–60 km
Housing	0.1	0–0.7	42,917	5–311,582
Permitted NO_2 emissions	0.0	0–0.0	19 tons/year	0–1064 tons/year
Developed open space	−0.3	−0.9–0	0.24 ha	0–3 ha
Trees	−0.4	−1.0–0	6.7 ha	0–50 ha
Shrub/Scrub	−0.1	−0.2–0	24 ha	0–495 ha

Estimated annual incidence of respiratory health problems linked to LULC-related NO_2 concentrations are shown in Table 4, while the annual economic valuation of these health impacts can be found in Table S4. The health burden of NO_2, based on the health impact functions in the BenMAP database, appears to fall disproportionately on children under 14 years old, primarily in the form of asthma exacerbation and missed school days. For instance, there are 42,000 incidents of asthma exacerbation for every 100,000 4–12 year-olds, of which approximately 8% may be associated with VMT*f*, and −3% with tree canopy.

Table 4. Estimated annual incidence of respiratory problems per 100,000 individuals associated with LULC due to local influence on ambient NO_2, in the Portland-Hillsboro-Vancouver urban area.

Health Impact	Annual Incidence Rate (per 100,000) Associated with LULC Category						
	All NO_2	VMT*f*	Sec. Rds	High Intensity Dev.	Med. Intensity Dev.	Open Dev.	Trees
Asthma Exacerbation, Missed school days (4–12 year olds)	14,455	1109	1322	2393	1587	−583	−472
Asthma Exacerbation, One or More Symptoms (4–12 year olds)	42,171	3220	3837	6950	4606	−1692	−1369
Cough (7–14 year olds)	12,070	926	1108	1989	1338	−503	−414
Emergency Room Visits, Asthma (75 years and older)	22	2	2	3	2	−1	−1
Hospital admissions, Asthma (younger than 30 years)	1	0	0	0	0	0	0
Hospital admissions, Asthma (30 years and older)	1	0	0	0	0	0	0
Hospital admissions, Chronic Lung Disease (less Asthma) (65 years and older)	64	6	6	11	6	−2	−2
Hospital admissions, All Respiratory (65 years and older)	137	12	13	23	13	−5	−4

3.3. Evaluating the Mitigation Potential of LULC Modifications

Figure 5 and Table S5 present the results of the sensitivity analysis, showing how local NO_2 concentrations change in response to modifications in four LULC categories: VMTf, high intensity development, open development and tree canopy. Table S5 summarizes the percent change in local annual NO_2 (averaged over the entire study area) in response to ±2%, ±5% and ±10% changes in these LULC categories. We see that changes in VMTf have relatively little impact on the region-wide average of NO_2, while increasing tree cover by 10% decreases local concentrations of NO_2 by 3%, on average, across the study area. The reason for the relatively small change in NO_2 in response to an increase in VMTf (0.1%) versus the comparatively larger changes in regionally averaged NO_2 in response to increases in high intensity development (2–11%), trees (1–3%), and open development (1–3%) (Table S5) is partly methodological: high intensity, tree, and open areas are increased as a percentage of the buffer area, and not as a percentage of the LULC category within the buffer. For example, a grid point with 0% tree cover, tree canopy is increased to 2%, 5% and 10% in the sensitivity analysis, similarly for a point with 20% tree cover, the increases are to 22%, 25% and 30%. This method of changing LULC categories results in effects observed across the entire study area.
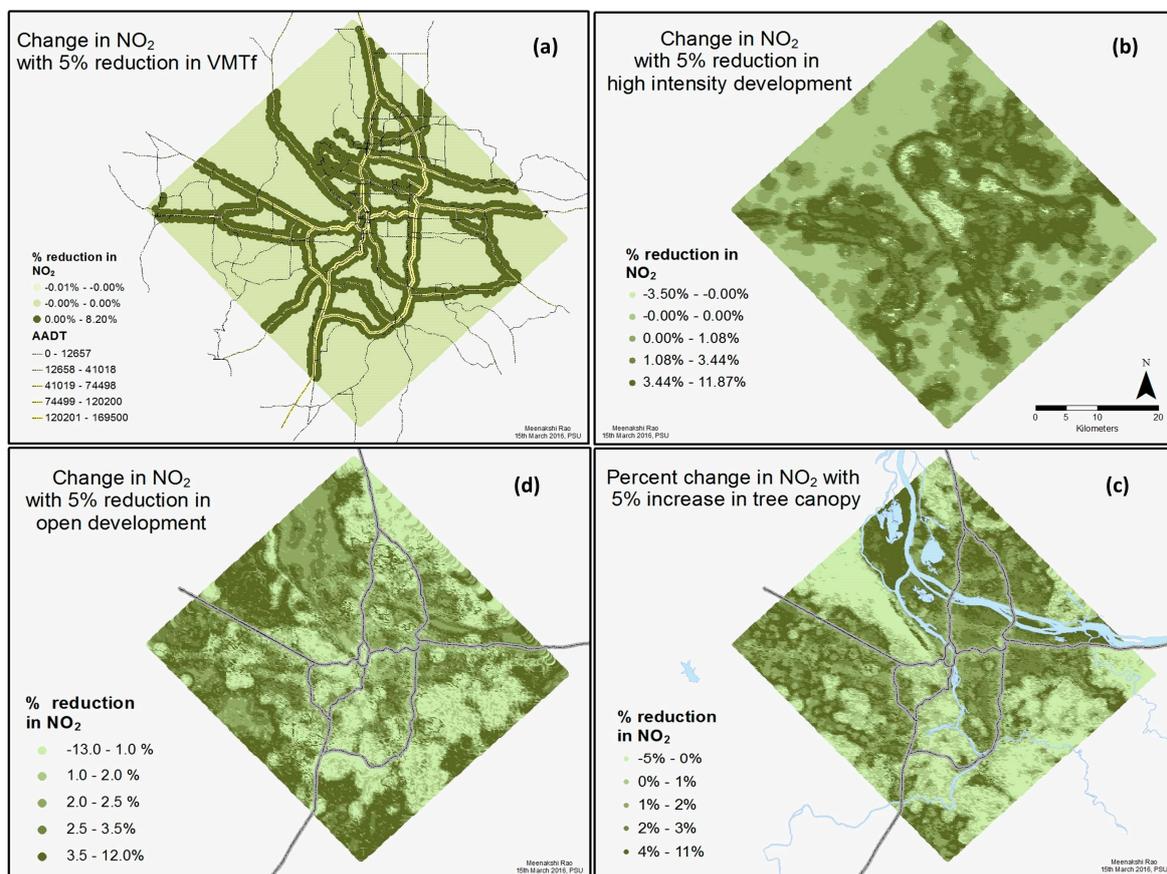

Figure 5. The spatial distribution and magnitude of the change in modeled NO_2 concentrations in response to a ±5% change in (clockwise from top left) (**a**) VMTf (**b**) high intentsity development (**c**) tree canopy (**d**) open development.

Figure 5 shows the spatial pattern and relative magnitude of NO_2 change associated with a 5% change in each of the four LULC categories. We see that the spatial pattern of NO_2 response to LULC modifications is distinct for the four LULC categories considered. The NO_2 change in response to changes in VMTf is constrained to a narrow buffer around the freeways, consistent with the observed drop-off in NO_2 concentrations with distance from roads [3]. The response of NO_2 to modifications

in the other three categories is much more spread out across the study area. For high intensity development, the greatest percentage reduction in NO_2 is towards the center of the study area, where the high density development is most intense, possibly reflecting the association of combustion processes and development. For developed open space, on the other hand, there is an increase in NO_2 on the periphery of the study area, consistent with treed areas on the outskirts being replaced with open development; and a decrease in NO_2 towards the center, due to high intensity development being replaced by open development. Similarly, we observe a small increase to no change in NO_2 concentrations in changing tree canopy in areas which already have high percentage of tree cover (see Figure 1), and a decrease in NO_2 in areas with previously low tree canopy. Local impact of a 5% change in VMT*f*, high intensity development, open development and tree canopy on NO_2 is as high as 8%, 12%, 12% and 11% respectively. Together these indicate that different LULC modification strategies might be optimum in different parts of the urban area for mitigating local NO_2.

With this modeled estimate of the change in NO_2 in response to LULC modifications, we next estimated the health benefit associated with this NO_2 change using BenMAP. We focused on change in incidence rates of asthma exacerbation in 4–12 year olds arising from the change in local NO_2 corresponding to changes in VMT*f* and tree canopy. Table 5 shows the percent change in annaul incidence rates of asthma exacerbation (as compared to the incidence rate for all NO_2-related asthma exacerbation of ~42,000, Table 3), arising from the change in local NO_2 associated with modifications to VMT*f* and tree canopy. Results shown are averaged over the study area, as well as averaged over just the area lying within the worst NO_2 quintile (Figure S2). We see that decreasing VMT*f* even by 10% has very little impact on asthma incidence rates, while increasing tree cover is associated with up to an 11% decrease in the incidence of all NO_2-related asthma (total NO_2-related asthma being about ~42,000 incidents/100,000, Table 4). Although this result may initially seem surprising, it is a consequence of the fact that health outcomes depend on the distribution of the air pollutant as well as the population. Although decreasing VMT*f* by 10% decreases NO_2 substantially in a 700 m around the freeways, the population of 4–12 year-olds in this narrow buffer is small, leading to a small change incidence rates.

Table 5. Estimated change in the incidence of NO_2-related asthma exacerbation associated with modifications to the two LULC categories VMTf and trees.

% Change in NO_2-Related Asthma Exacerbation Symptoms in 4–12 Year Olds Due to Changes in NO_2 Associated with LULC Modifications				
LULC Category/LULC Change	VMT*f*	VMT*f* (in Worst NO_2 Quintile)	Trees	Trees (in Worst NO_2 Quintile)
10% decrease	−0.5%	−0.8%	2%	1%
5% decrease	−0.2%	−0.4%	2%	1%
2% decrease	−0.1%	−0.2%	1%	1%
2% increase	0.1%	0.1%	−3%	−3%
5% increase	0.2%	0.3%	−6%	−6%
10% increase	0.4%	0.7%	−10%	−11%

4. Discussion

We showed that the random forest ensemble learning technique performed well in capturing the fine spatial scale variation of NO_2—the LURF model performed well based on statistical performance metrics and predictive ability, on par with the widely used LUR methodology. Although the LURF model overestimates NO_2 concentrations (Table 2), the ability of LURF to explore impacts of individual LULC variables suggests that the random forest technique can be added to the repertoire of new and established statistical such as structural equation modeling [78], generalized boosting models [19], and neural networks [18,79] that are used today to better understand air pollution in our urban environments; and particularly to explore how cities can modify land use to reduce

NO_2 and improve respiratory health. We should remain cognizant of the fact that the LURF models, like the LUR models, are likely not to transfer well between cities, especially cities in different parts of the world [80,81].

We found—based on the BenMAP health impact functions—that even in the Portland-Hillsboro-Vancouver metropolitan area, a city in compliance with US EPA and WHO standards, there still exists a significant respiratory burden, borne predominantly by children under the age of 12, resulting from the ubiquitous urban pollutant NO_2. Utilizing the LURF NO_2 model, we were able to examine the relative impact and spatial pattern of the different urban LULC categories on incidence rates of respiratory health issues. For example, the LURF model showed that NO_2 associated with VMT*f* was linked to an increase in respiratory health issues (~3200 per 100,000 increase in asthma exacerbation symptoms in 4–12 year olds), and further, this impact was spatially clustered close to freeways and highways. The overall effect of urban tree canopy was smaller (~1369 per 100,000 decrease in asthma exacerbation in 4–12 year olds) and but more widely spatially distributed over the study area.

Today, municipalities often engage in city-wide LULC modifications such as VMT reduction and "greening" campaigns, both to reduce emissions and improve human health. However, no clear and accessible assessment methodology has existed to help estimate the benefits accruing from these campaigns, or alternately, to identify optimum strategies from a range of strategies. Our sensitivity analysis using the LURF model to study the response of ambient NO_2 to LULC modifications showed that planting trees locally (which has the effect of reducing developed areas and increasing tree canopy) may be a good strategy to reduce local NO_2 concentrations and improve respiratory health. Furthermore, model outcomes suggest that a 10% increase in canopy coverage city-wide may reduce the incidences of childhood asthma by an order of magnitude more than a 5% or even 10% reduction in VMT*f*.

To the extent that the Portland-Hillsboro-Vancouver metropolitan area is representative of other mid-size cities in the USA, we can expect a similar burden of respiratory health due to NO_2, borne disproportionately by children, in other cities as well [82–85]. Recent studies show that childhood exposure to traffic-related air pollution may lead to impaired lung function in early adulthood [82,86] and that exposure to air pollution in childhood is linked to poorer performance in school [87] which in turn could lead to lower earnings potential in adulthood. Given the increasing number of studies that suggest the role of the urban forest in promoting physical and mental well-being [88–92], it seems likely that small-scale strategic tree planting campaigns in either high NO_2 areas and/or near roadways, or city-wide greening campaigns, may well play an important role in improving human health, which may come in some small part due to the mitigating of respiratory distress associated with NO_2.

Keeping in mind that correlation does not imply causation, it is prudent that we seek to better understand the mechanisms through which land use modifications, including tree plantings, affect ambient NO_2 concentrations. For example, trees have been shown to remove NO_2 from the atmosphere through dry deposition [93,94]; yet the rates of dry deposition in the urban environment [28,95,96], species-specific dry deposition rates [97], seasonal variations in dry deposition, and other questions have yet to be studied extensively. We hope that this paper inspires further research, both statistical and mechanistic, into how urban land use design and modifications can be used to mitigate the health effects of urban air pollution.

5. Conclusions

In today's rapidly urbanizing world, where, according to the World Health Organization, air pollution has become "the single largest environmental health risk" [1] there is an urgent need to design cities that promote cleaner atmospheres. Critically, the random forest technique applied in this paper is robust in handling noisy and missing data, a not uncommon feature of dense sensor networks, making it ideally suited for analyzing the flood of data from sensor technologies that are currently on the horizon. Since it is relatively easy to use, does not require intense computational support,

Int. J. Environ. Res. Public Health **2017**, *14*, 750

14 of 19

and the output models are readily interpreted, the use of this technique has the potential to include a wide range of stakeholders, including planners, citizens, and agencies, in the process of better characterizing and managing local LULC to optimize air quality.

The results presented here serve to highlight the need for future research to better understand the mechanisms that determine how different LULC categories shape the intra-urban patterns of air pollution within our cities. Combining the sophistication of new sensor technologies with advanced modeling techniques, including random forest, will contribute to a better understanding of the linkages between land use and urban air pollution and lead to creating healthier cities and more sustainable urban atmospheres for all.

References

1. World Health Organization. Air Pollution Rising in Many of the World's Poorest Cities. Available online: http://www.who.int/mediacentre/news/releases/2016/air-pollution-rising/en/ (accessed on 14 June 2016).

2. Jerrett, M.; Arain, A.; Kanaroglou, P.; Beckerman, B.; Potoglou, D.; Sahsuvaroglu, T.; Morrison, J.; Giovis, C. A review and evaluation of intraurban air pollution exposure models. *J. Expo. Anal. Environ. Epidemiol.* **2005**, *15*, 185–204. [CrossRef] [PubMed]

3. Karner, A.A.; Eisinger, D.S.; Niemeier, D.A. Near-roadway air quality: Synthesizing the findings from real-world data. *Environ. Sci. Technol.* **2010**, *44*, 5334–5344. [CrossRef] [PubMed]

4. Ozkaynak, H.; Baxter, L.K.; Dionisio, K.L.; Burke, J. Air pollution exposure prediction approaches used in air pollution epidemiology studies. *J. Expo. Sci. Environ. Epidemiol.* **2013**, *23*, 566–572. [CrossRef] [PubMed]

5. Briggs, D.J.; de Hoogh, C.; Gulliver, J.; Wills, J.; Elliott, P.; Kingham, S.; Smallbone, K. A regression-based method for mapping traffic-related air pollution: Application and testing in four contrasting urban environments. *Sci. Total Environ.* **2000**, *253*, 151–167. [CrossRef]

6. Gilbert, N.L.; Goldberg, M.S.; Beckerman, B.; Brook, J.R.; Jerrett, M. Assessing spatial variability of ambient nitrogen dioxide in Montréal, Canada, with a land-use regression model. *J. Air Waste Manag. Assoc.* **2005**, *55*, 1059–1063. [CrossRef] [PubMed]

7. Ryan, P.H.; LeMasters, G.K. A Review of Land-use Regression Models for Characterizing Intraurban Air Pollution Exposure. *Inhal. Toxicol.* **2007**, *19*, 127–133. [CrossRef] [PubMed]

8. Miller, L.; Lemke, L.D.; Xu, X.; Molaroni, S.M.; You, H.; Wheeler, A.J.; Booza, J.; Grgicak-Mannion, A.; Krajenta, R.; Graniero, P. Intra-urban correlation and spatial variability of air toxics across an international airshed in Detroit, Michigan (USA) and Windsor, Ontario (Canada). *Atmos. Environ.* **2010**, *44*, 1162–1174. [CrossRef]

9. Beelen, R.; Hoek, G.; Pebesma, E.; Vienneau, D.; de Hoogh, K.; Briggs, D.J. Mapping of background air pollution at a fine spatial scale across the European Union. *Sci. Total Environ.* **2009**, *407*, 1852–1867. [CrossRef] [PubMed]

10. Laurent, O.; Hu, J.; Li, L.; Cockburn, M.; Escobedo, L.; Kleeman, M.J.; Wu, J. Sources and contents of air pollution affecting term low birth weight in Los Angeles County, California, 2001–2008. *Environ. Res.* **2014**, *134*, 488–495. [CrossRef] [PubMed]

11. Borrego, C.; Amorim, J.H.; Tchepel, O.; Dias, D.; Rafael, S.; Sá, E.; Pimentel, C.; Fontes, T.; Fernandes, P.; Pereira, S.R.; et al. Urban scale air quality modelling using detailed traffic emissions estimates. *Atmos. Environ.* **2016**, *131*, 341–351. [CrossRef]

12. Lobdell, D.T.; Isakov, V.; Baxter, L.; Touma, J.S.; Smuts, M.B.; Özkaynak, H. Feasibility of assessing public health impacts of air pollution reduction programs on a local scale: New Haven case study. *Environ. Health Perspect.* **2011**, *119*, 487–493. [CrossRef] [PubMed]

13. Levy, J.I.; Spengler, J.D.; Hlinka, D.; Sullivan, D.; Moon, D. Using CALPUFF to evaluate the impacts of power plant emissions in Illinois: Model sensitivity and implications. *Atmos. Environ.* **2002**, *36*, 1063–1075. [CrossRef]

14. Tuccella, P.; Curci, G.; Visconti, G.; Bessagnet, B.; Menut, L.; Park, R.J. Modeling of gas and aerosol with WRF/Chem over Europe: Evaluation and sensitivity study. *J. Geophys. Res. Atmos.* **2012**, *117*, 1–15. [CrossRef]

15. Joe, D.K.; Zhang, H.; DeNero, S.P.; Lee, H.-H.; Chen, S.-H.; McDonald, B.C.; Harley, R.A.; Kleeman, M.J. Implementation of a high-resolution Source-Oriented WRF/Chem model at the Port of Oakland. *Atmos. Environ.* **2014**, *82*, 351–363. [CrossRef]

16. Sanchez, E.Y.; Lerner, J.C.; Porta, A.; Jacovkis, P.M. Accidental release of chlorine in Chicago: Coupling of an exposure model with a Computational Fluid Dynamics model. *Atmos. Environ.* **2013**, *64*, 47–55. [CrossRef]

17. Reyes, J.M.; Serre, M.L. An LUR/BME framework to estimate PM2.5 explained by on road mobile and stationary sources. *Environ. Sci. Technol.* **2014**, *48*, 1736–1744. [CrossRef] [PubMed]

18. Fernando, H.J.S.; Mammarella, M.C.; Grandoni, G.; Fedele, P.; Di Marco, R.; Dimitrova, R.; Hyde, P. Forecasting PM10 in metropolitan areas: Efficacy of neural networks. *Environ. Pollut.* **2012**, *163*, 62–67. [CrossRef] [PubMed]

19. Reid, C.E.; Jerrett, M.; Petersen, M.L.; Pfister, G.G.; Morefield, P.E.; Tager, I.B.; Raffuse, S.M.; Balmes, J.R. Spatiotemporal prediction of fine particulate matter during the 2008 Northern California wildfires using machine learning. *Environ. Sci. Technol.* **2015**, *49*, 3887–3896. [CrossRef] [PubMed]

20. Champendal, A.; Kanevski, M.; Huguenot, P.-E. Air pollution mapping using nonlinear land use regression models. In *Computational Science and Its Applications—ICCSA 2014, Part III, Proceedings of the 14th International Conference on Computational Science and Its Applications, Guimaraes, Portugal, 30 June–3 July 2014*; Springer: Cham, Switzerland, 2014; Volume 8583, pp. 682–690.

21. Ng, E. Policies and technical guidelines for urban planning of high-density cities—air ventilation assessment (AVA) of Hong Kong. *Build. Environ.* **2009**, *44*, 1478–1488. [CrossRef]

22. School Facilities Planning Division: California Department of Education. School Site Selection and Approval Guide. Available online: http://www.cde.ca.gov/ls/fa/sf/schoolsiteguide.asp (accessed on 16 November 2015).

23. Bureau of Planning and Sustainability. 2015 Climate Action Plan: The City of Portland Oregon. Available online: https://www.portlandoregon.gov/bps/66993 (accessed on 1 December 2015).

24. City of Burlington Vermont Burlington VT. Climate Action Plan. Available online: https://www.burlingtonvt.gov/Sustainability/CAP (accessed on 1 December 2015).

25. Seattle Office of Sustainability and Environment. Seattle Climate Action Plan. Available online: http://www.seattle.gov/Documents/Departments/OSE/2013_CAP_20130612.pdf (accessed on 1 December 2015).

26. Rao, M.; George, L.A.; Rosenstiel, T.N.; Shandas, V.; Dinno, A. Assessing the relationship among urban trees, nitrogen dioxide, and respiratory health. *Environ. Pollut.* **2014**, *194*, 96–104. [CrossRef] [PubMed]

27. Borrego, C.; Martins, H.; Tchepel, O.; Salmim, L.; Monteiro, A.; Miranda, A.I. How urban structure can affect city sustainability from an air quality perspective. *Environ. Model. Softw.* **2006**, *21*, 461–467. [CrossRef]

28. Nowak, D.J.; Hirabayashi, S.; Bodine, A.; Hoehn, R. Modeled PM$_{2.5}$ removal by trees in ten U.S. cities and associated health effects. *Environ. Pollut.* **2013**, *178*, 395–402. [CrossRef] [PubMed]

29. Morani, A.; Nowak, D.J.; Hirabayashi, S.; Calfapietra, C. How to select the best tree planting locations to enhance air pollution removal in the MillionTreesNYC initiative. *Environ. Pollut.* **2011**, *159*, 1040–1047. [CrossRef] [PubMed]

30. Cabaraban, M.T.I.; Kroll, C.N.; Hirabayashi, S.; Nowak, D.J. Modeling of air pollutant removal by dry deposition to urban trees using a WRF/CMAQ/i-Tree Eco coupled system. *Environ. Pollut.* **2013**, *176*, 123–133. [CrossRef] [PubMed]

Int. J. Environ. Res. Public Health **2017**, *14*, 750

16 of 19

31. Brioude, J.; Angevine, W.M.; Ahmadov, R.; Kim, S.W.; Evan, S.; McKeen, S.A.; Hsie, E.Y.; Frost, G.J.; Neuman, J.A.; Pollack, I.B.; et al. Top-down estimate of surface flux in the Los Angeles Basin using a mesoscale inverse modeling technique: Assessing anthropogenic emissions of CO, NOx and CO_2 and their impacts. *Atmos. Chem. Phys.* **2013**, *13*, 3661–3677. [CrossRef]

32. Streets, D.G.; Canty, T.; Carmichael, G.R.; De Foy, B.; Dickerson, R.R.; Duncan, B.N.; Edwards, D.P.; Haynes, J.A.; Henze, D.K.; Houyoux, M.R.; et al. Emissions estimation from satellite retrievals: A review of current capability. *Atmos. Environ.* **2013**, *77*, 1011–1042. [CrossRef]

33. Beelen, R.; Hoek, G.; Vienneau, D.; Eeftens, M.; Dimakopoulou, K.; Pedeli, X.; Tsai, M.-Y.; Künzli, N.; Schikowski, T.; Marcon, A.; et al. Development of NO_2 and NO_x land use regression models for estimating air pollution exposure in 36 study areas in Europe—The ESCAPE project. *Atmos. Environ.* **2013**, *72*, 10–23. [CrossRef]

34. Mavko, M.E.; Tang, B.; George, L.A. A sub-neighborhood scale land use regression model for predicting NO_2. *Sci. Total Environ.* **2008**, *398*, 68–75. [CrossRef] [PubMed]

35. Henderson, S.B.; Beckerman, B.; Jerrett, M.; Brauer, M. Application of Land Use Regression to Estimate Long-Term Concentrations of Traffic-Related Nitrogen Oxides and Fine Particulate Matter. *Environ. Sci. Technol.* **2007**, *41*, 2422–2428. [CrossRef] [PubMed]

36. Kashima, S.; Yorifuji, T.; Tsuda, T.; Doi, H. Application of land use regression to regulatory air quality data in Japan. *Sci. Total Environ.* **2009**, *407*, 3055–3062. [CrossRef] [PubMed]

37. Liu, C.; Henderson, B.H.; Wang, D.; Yang, X.; Peng, Z. A land use regression application into assessing spatial variation of intra-urban fine particulate matter ($PM_{2.5}$) and nitrogen dioxide (NO_2) concentrations in City of Shanghai, China. *Sci. Total Environ.* **2016**, *565*, 607–615. [CrossRef] [PubMed]

38. Graham, M.H. Confronting multicollinearity in ecological multiple regression. *Ecology* **2003**, *84*, 2809–2815. [CrossRef]

39. Cohen, J.; Cohen, P.; West, S.G.; Aiken, L.S. *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*; Lawrence Erlbaum Associates: Mahwah, NJ, USA, 2003.

40. Brauer, M.; Amann, M.; Burnett, R.T.; Cohen, A.; Dentener, F.; Ezzati, M.; Henderson, S.B.; Krzyzanowski, M.; Martin, R.V.; Van Dingenen, R.; et al. Exposure assessment for estimation of the global burden of disease attributable to outdoor air pollution. *Environ. Sci. Technol.* **2012**, *46*, 652–660. [CrossRef] [PubMed]

41. Fann, N.; Lamson, A.D.; Anenberg, S.C.; Wesson, K.; Risley, D.; Hubbell, B.J. Estimating the national public health burden associated with exposure to ambient $PM_{2.5}$ and ozone. *Risk Anal.* **2012**, *32*, 81–95. [CrossRef] [PubMed]

42. United States Environmental Protection Agency (US EPA). AirNow. Available online: http://www.airnow.gov/ (accessed on 1 December 2015).

43. MarcoPolo-Panda. Air Quality Forecast. Available online: http://www.marcopolo-panda.eu/forecast/ (accessed on 1 December 2015).

44. James, G.; Witten, D.; Hastie, T.; Tibshirani, R. *An Introduction to Statistical Learning*; Casella, G., Fienberg, S., Olkin, I., Eds.; Springer: New York, NY, USA, 2006.

45. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]

46. Gray, K.R.; Aljabar, P.; Hammers, R.A.; Hammers, A.; Rueckert, D.; Alzheimer's Disease Neuroimaging Initiative. Random forest-based similarity measures for multi-modal classification of Alzheimer's disease. *Neuroimage* **2013**, *65*, 167–175. [CrossRef] [PubMed]

47. Jiang, P.; Wu, H.; Wang, W.; Ma, W.; Sun, X.; Lu, Z. MiPred: Classification of real and pseudo microRNA precursors using random forest prediction model with combined features. *Nucleic Acids Res.* **2007**, *35*, W339–W344. [CrossRef] [PubMed]

48. Chen, X.-W.; Liu, M. Prediction of protein–protein interactions using random decision forest framework. *Bioinformatics* **2005**, *21*, 4394–4400. [CrossRef] [PubMed]

49. Svetnik, V.; Liaw, A.; Tong, C.; Culberson, J.C.; Sheridan, R.P.; Feuston, B.P. Random forest: A classification and regression tool for compound classification and QSAR modeling. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 1947–1958. [CrossRef] [PubMed]

50. Rodriguez-Galiano, V.F.; Ghimire, B.; Rogan, J.; Chica-Olmo, M.; Rigol-Sanchez, J.P. An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS J. Photogramm. Remote Sens.* **2012**, *67*, 93–104. [CrossRef]

Int. J. Environ. Res. Public Health **2017**, *14*, 750

17 of 19

51. Pal, M. Random forest classifier for remote sensing classification. *Int. J. Remote Sens.* **2005**, *26*, 217–222. [CrossRef]

52. Cutler, D.R.; Edwards, T.C.; Beard, K.H.; Cutler, A.; Hess, K.T.; Gibson, J.; Lawler, J.J. Random forests for classification in ecology. *Ecology* **2007**, *88*, 2783–2792. [CrossRef] [PubMed]

53. Prasad, A.M.; Iverson, L.R.; Liaw, A. Newer classification and regression tree techniques: Bagging and random forests for ecological prediction. *Ecosystems* **2006**, *9*, 181–199. [CrossRef]

54. Brokamp, C.; Jandarov, R.; Rao, M.B.; LeMasters, G.; Ryan, P. Exposure assessment models for elemental components of particulate matter in an urban environment: A comparison of regression and random forest approaches. *Atmos. Environ.* **2017**, *151*, 1–11. [CrossRef]

55. Hunter, K. Aclima, Google, and the EPA Discuss Their Partnership for Planetary Health. Available online: https://blog.aclima.io/aclima-google-and-the-epa-discuss-their-partnership-for-planetary-health-b0f5402bbed3 (accessed on 30 November 2015).

56. Air Quality Egg. Available online: http://airqualityegg.com/ (accessed on 30 November 2015).

57. United States Environmental Protection Agency (US EPA). Criteria Air Pollutants. Available online: https://www.epa.gov/criteria-air-pollutants (accessed on 23 July 2016).

58. Seinfeld, J.H.; Pandis, S.N. *Atmospheric Chemistry and Physics: From Air Pollution to Climate Change*; John Wiley & Sons, Inc.: New York, NY, USA, 1998.

59. Finlayson-Pitts, B.J.; Pitts, J.N., Jr. *Chemistry of the Upper and Lower Atmosphere: Theory, Experiments and Applications*; Academic Press: San Diego, CA, USA, 2000.

60. United States Environmental Protection Agency (US EPA). Benefits Mapping and Analysis Program (BenMAP). Available online: http://www2.epa.gov/benmap/benmap-40 (accessed on 12 October 2015).

61. Abt Associates Inc. *BenMAP Environmental Benefits Mapping and Analysis Program User's Manual 2010*; U.S. Environmental Protection Agency: Washington, DC, USA, 2010.

62. Abt Associates Inc. *BenMAP Environmental Benefits Mapping and Analysis Program User's Manual Appendices*; U.S. Environmental Protection Agency: Washington, DC, USA, 2010.

63. Homer, C.G.; Dewitz, J.A.; Yang, L.; Jin, S.; Danielson, P.; Xian, G.; Coulston, J.; Herold, N.D.; Wickham, J.D.; Megown, K. Completion of the 2011 National Land Cover Database for the conterminous United States-Representing a decade of land cover change information. *Photogramm. Eng. Remote Sens.* **2015**, *77*, 758–762.

64. US Department of Transportation National Highway Planning Network: 2011 Archive. Available online: http://www.fhwa.dot.gov/planning/processes/tools/nhpn/2011/ (accessed on 6 December 2015).

65. Rao, M. Investigating the Potential of Land Use Modifications to Mitigate the Respiratory Health Impacts of NO2: A Case Study in the Portland-Vancouver Metropolitan Area. Ph.D. Thesis, Portland State University, Portland, OR, USA, 2016.

66. Wang, M.; Beelen, R.; Basagana, X.; Becker, T.; Cesaroni, G.; de Hoogh, K.; Dedele, A.; Declercq, C.; Dimakopoulou, K.; Eeftens, M.; et al. Evaluation of Land Use Regression Models for NO_2 and Particulate Matter in 20 European Study Areas: The ESCAPE Project. *Environ. Sci. Technol.* **2013**, *47*, 4357–4364. [CrossRef] [PubMed]

67. Ross, Z.; English, P.B.; Scalf, R.; Gunier, R.; Smorodinsky, S.; Wall, S.; Jerrett, M. Nitrogen dioxide prediction in Southern California using land use regression modeling: Potential for environmental health analyses. *J. Expo. Sci. Environ. Epidemiol.* **2006**, *16*, 106–114. [CrossRef] [PubMed]

68. Kendrick, C.M.; Koonce, P.; George, L.A. Diurnal and seasonal variations of NO, NO_2 and $PM_{2.5}$ mass as a function of traffic volumes alongside an urban arterial. *Atmos. Environ.* **2015**, *122*, 133–141. [CrossRef]

69. Strobl, C.; Hothorn, T.; Zeileis, A. *Party on! A New, Conditional Variable-Importance Measure for Random Forests Available in the Party Package*; University of Munich: Munich, Germany, 2009.

70. Strobl, C.; Boulesteix, A.-L.; Kneib, T.; Augustin, T.; Zeileis, A. Conditional Variable Importance for Random Forests. *BMC Bioinform.* **2008**, *9*, 307. [CrossRef] [PubMed]

71. Hothorn, T.; Buehlmann, P.; Dudoit, S.; Molinaro, A.; Van Der Laan, M. Survival Ensembles. *Biostatistics* **2006**, *7*, 355–373. [CrossRef] [PubMed]

72. Strobl, C.; Boulesteix, A.-L.; Zeileis, A.; Hothorn, T. Bias in Random Forest Variable Importance Measures: Illustrations, Sources and a Solution. *Bioinformatics* **2007**, *8*. [CrossRef]

73. R Core Team. R: A language and Environment for Statistical Computing. 2014. Available online: https://www.r-project.org/ (accessed on 28 June 2017).

74. Yin, S.; Shen, Z.; Zhou, P.; Zou, X.; Che, S.; Wang, W. Quantifying air pollution attenuation within urban parks: An experimental approach in Shanghai, China. *Environ. Pollut.* **2011**, *159*, 2155–2163. [CrossRef] [PubMed]

75. Hoek, G.; Beelen, R.; de Hoogh, K.; Vienneau, D.; Gulliver, J.; Fischer, P.; Briggs, D. A review of land-use regression models to assess spatial variation of outdoor air pollution. *Atmos. Environ.* **2008**, *42*, 7561–7578. [CrossRef]

76. United States Environmental Protection Agency (US EPA). Technology Transfer Network: National Ambient Air Quality Stansards (NAAQS). Availble online: http://www3.epa.gov/ttn/naaqs/criteria.html (accessed on 1 December 2015).

77. World Health Organization. *WHO Air Quality Guidelines for Particulate Matter, Ozone, Nitrogen Dioxide and Sulfur Dioxide*; WHO: Geneva, Switzerland, 2005.

78. Sanchez, B.N.; Budtz-Jørgensen, E.; Ryan, L.M.; Hu, H. Structural equation models: A review with applications to environmental epidemiology. *J. Am. Stat. Assoc.* **2005**, *100*, 1443–1455. [CrossRef]

79. Di, Q.; Kloog, I.; Koutrakis, P.; Lyapustin, A.; Wang, Y.; Schwartz, J. Assessing $PM_{2.5}$ Exposures with High Spatiotemporal Resolution across the Continental United States. *Environ. Sci. Technol.* **2016**, *50*, 4712–4721. [CrossRef] [PubMed]

80. Allen, R.W.; Amram, O.; Wheeler, A.J.; Brauer, M. The transferability of NO and NO_2 land use regression models between cities and pollutants. *Atmos. Environ.* **2011**, *45*, 369–378. [CrossRef]

81. Poplawski, K.; Gould, T.; Setton, E.; Allen, R.; Su, J.; Larson, T.; Henderson, S.; Brauer, M.; Hystad, P.; Lightowlers, C.; et al. Intercity transferability of land use regression models for estimating ambient concentrations of nitrogen dioxide. *J. Expo. Sci. Environ. Epidemiol.* **2009**, *19*, 107–117. [CrossRef] [PubMed]

82. Chen, Z.; Salam, M.T.; Eckel, S.P.; Breton, C.V.; Gilliland, F.D. Chronic effects of air pollution on respiratory health in Southern California children: Findings from the Southern California Children's Health Study. *J. Thorac. Dis.* **2015**, *7*, 46–58. [PubMed]

83. Lemke, L.D.; Lamerato, L.E.; Xu, X.; Booza, J.C.; Reiners, J.J.; Raymond, D.M., III; Villeneuve, P.J.; Lavigne, E.; Larkin, D.; Krouse, H.J. Geospatial relationships of air pollution and acute asthma events across the Detroit-Windsor international border: Study design and preliminary results. *J. Expo. Sci. Environ. Epidemiol.* **2013**, *24*, 1–12.

84. Penard-Morand, C.; Raherison, C.; Charpin, D.; Kopferschmitt, C.; Lavaud, F.; Caillaud, D.; Annesi-maesano, I. Long-term exposure to close-proximity air pollution and asthma and allergies in urban children. *Eur. Respir. J.* **2010**, *36*, 33–40. [CrossRef] [PubMed]

85. Son, J.-Y.; Kim, H.; Bell, M.L. Does urban land-use increase risk of asthma symptoms? *Environ. Res.* **2015**, *142*, 309–318. [CrossRef] [PubMed]

86. Gauderman, W.J.; Vora, H.; McConnell, R.; Berhane, K.; Gilliland, F.; Thomas, D.; Lurmann, F.; Avol, E.; Kunzli, N.; Jerrett, M. Effect of exposure to traffic on lung development from 10 to 18 years of age: A cohort study. *Lancet* **2007**, *369*, 571–577. [CrossRef]

87. Clark-Reyna, S.E.; Grineski, S.E.; Collins, T.W. Residential exposure to air toxics is linked to lower grade point averages among school children in El Paso, Texas, USA. *Popul. Environ.* **2016**, *37*, 319–340. [CrossRef] [PubMed]

88. Donovan, G.H.; Butry, D.T.; Michael, Y.L.; Prestemon, J.P.; Liebhold, A.M.; Gatziolis, D.; Mao, M.Y. The relationship between trees and human health: Evidence from the spread of the emerald ash borer. *Am. J. Prev. Med.* **2013**, *44*, 139–145. [CrossRef] [PubMed]

89. Donovan, G.H.; Michael, Y.L.; Butry, D.T.; Sullivan, A.D.; Chase, J.M. Urban trees and the risk of poor birth outcomes. *Health Place* **2011**, *17*, 390–393. [CrossRef] [PubMed]

90. Takano, T.; Nakamura, K.; Watanabe, M. Urban residential environments and senior citizens' longevity in megacity areas: The importance of walkable green spaces. *J. Epidemiol. Community Health* **2002**, *56*, 913–918. [CrossRef] [PubMed]

91. Ulrich, R. View through a window may influence recovery from surgery. *Science* **1984**, *224*, 420–421. [CrossRef] [PubMed]

92. Maas, J.; Verheij, R.A.; Groenewegen, P.P.; de Vries, S.; Spreeuwenberg, P. Green space, urbanity, and health: How strong is the relation? *J. Epidemiol. Community Health* **2006**, *60*, 587–592. [CrossRef] [PubMed]

93. Wesely, M.L.; Hicks, B.B. A review of the current status of knowledge on dry deposition. *Atmos. Environ.* **2000**, *34*, 2261–2282. [CrossRef]

Int. J. Environ. Res. Public Health **2017**, *14*, 750

19 of 19

94. Baldocchi, D.D.; Hicks, B.B.; Camara, P. A canopy stomatal resistance model for gaseous deposition to vegetated surfaces. *Atmos. Environ.* **1987**, *21*, 91–101. [CrossRef]

95. Nowak, D.J.; Crane, D.E.; Stevens, J.C. Air pollution removal by urban trees and shrubs in the United States. *Urban For. Urban Green.* **2006**, *4*, 115–123. [CrossRef]

96. Nowak, D.J.; Civerolo, K.L.; Rao, S.T.; Sistla, G.; Luley, C.J.; Crane, D.E. A modeling study of the impact of urban trees on ozone. *Atmos. Environ.* **2000**, *34*, 1601–1613. [CrossRef]

97. Takahashi, M.; Higaki, A.; Nohno, M.; Kamada, M.; Okamura, Y.; Matsui, K.; Kitani, S.; Morikawa, H. Differential assimilation of nitrogen dioxide by 70 taxa of roadside trees at an urban pollution level. *Chemosphere* **2005**, *61*, 633–639. [CrossRef] [PubMed]