Comparison of remote sensing and fixed-site monitoring approaches for examining air pollution and health in a national study population

Genevieve Prud’homme a, Nina A. Dobbin a, Liu Sun a, Richard T. Burnett b, Randall V. Martin c,d, Andrew Davidson e,f, Sabit Cakmak b, Paul J. Villeneuve b,g,h, Lok N. Lamsal i,j, Aaron van Donkelaar c, Paul A. Peters k, Markey Johnson a,*

a Air Health Science Division, Health Canada, Ottawa, ON, Canada
b Population Studies Division, Health Canada, Ottawa, ON, Canada
c Department of Physics and Atmospheric Science, Dalhousie University, Halifax, NS, Canada
d Harvard-Smithsonian Center for Astrophysics, Cambridge, MA, USA
e Earth Observation Service, Agriculture and Agri-Food Canada, Ottawa, ON, Canada
f Department of Geography and Environmental Studies, Carleton University, Ottawa, ON, Canada
gh Institute of Health: Science, Technology and Policy, Carleton University, Ottawa, ON, Canada
h Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada
i Goddard Earth Sciences Technology and Research, Universities Space Research Association, Columbia, MD, USA
j NASA Goddard Space Flight Center, Greenbelt, MD, USA
k Health Analysis Division, Statistics Canada, Ottawa, ON, Canada

HIGHLIGHTS

- Remote sensing (RS) and regulatory monitoring (RM) were used to estimate air pollution.
- Pollution concentrations were assigned to homes in a national health study (N = 211,789).
- NO2 and PM2.5 were associated with adverse respiratory and allergic health outcomes.
- Risk estimates based on RS and RM were similar for participants living near monitors.
- RS pollutants were associated with adverse outcomes in remote/rural areas (p < 0.05).

ARTICLE INFO

Article history:
Received 27 March 2013
Received in revised form 25 June 2013
Accepted 9 July 2013

Keywords:
Air pollution
Satellite remote sensing
Regulatory monitoring
Asthma
Allergy
Population health
Epidemiology

ABSTRACT

Satellite remote sensing (RS) has emerged as a cutting edge approach for estimating ground level ambient air pollution. Previous studies have reported a high correlation between ground level PM2.5 and NO2 estimated by RS and measurements collected at regulatory monitoring sites. The current study examined associations between air pollution and adverse respiratory and allergic health outcomes using multi-year averages of NO2 and PM2.5 from RS and from regulatory monitoring.

RS estimates were derived using satellite measurements from OMI, MODIS, and MISR instruments. Regulatory monitoring data were obtained from Canada’s National Air Pollution Surveillance Network. Self-reported prevalence of doctor-diagnosed asthma, current asthma, allergies, and chronic bronchitis were obtained from the Canadian Community Health Survey (a national sample of individuals 12 years of age and older). Multi-year ambient pollutant averages were assigned to each study participant based on their six digit postal code at the time of health survey, and were used as a marker for long-term exposure to air pollution.

RS derived estimates of NO2 and PM2.5, were associated with 6–10% increases in respiratory and allergic health outcomes per interquartile range (3.97 μg m⁻³ for PM2.5 and 1.03 ppb for NO2) among adults (aged 20–64) in the national study population. Risk estimates for air pollution and respiratory/allergic health outcomes based on RS were similar to risk estimates based on regulatory monitoring for areas where regulatory monitoring data were available (within 40 km of a regulatory monitoring station). RS derived estimates of air pollution were also associated with adverse health outcomes among participants residing outside the catchment area of the regulatory monitoring network (p < 0.05).
1. Introduction

Exposure to ambient air pollution has been consistently associated with respiratory and cardiovascular morbidity and mortality (Brook, 2008; Brook et al., 2010; Brunekreef and Forsberg, 2005; Chen et al., 2008; Franchini and Mannucci, 2012; Krewski et al., 2003, 2005), and implicated in adverse allergic, metabolic, neurological, reproductive, and developmental health outcomes (Curtis et al., 2006; Genc et al., 2012; Health Effects Institute, 2010; Kampa and Castanas, 2008; Langer, 2010; Lewtas, 2007; Reidl, 2008; Saxon and Diaz-Sanchez, 2005; Sram et al., 2005). These effects have been demonstrated through a variety of epidemiologic studies examining long and short term associations at multi-city and intra-urban scales (Brook, 2008; Brook et al., 2010; Brunekreef and Forsberg, 2005; Dominici et al., 2003; Health Effects Institute, 2010; Peters et al., 2006; Ren and Tong, 2008; Weinmayr et al., 2010). However, large scale studies have typically been conducted in densely populated areas of developed nations due to the challenge and resource burden of assessing exposure to air pollution. There is an emerging interest within both scientific and regulatory communities to better understand risks associated with exposure to ambient air pollution among non-urban populations.

Ground-level regulatory monitoring networks in developed countries have often been used to estimate exposure in health studies (Kelly et al., 2012; Krewski et al., 2003, 2005; Laden and Neas, 2011; Miller et al., 2007; Wilson et al., 2005). However, because regulatory fixed-site monitors are primarily intended for surveillance, they are often restricted to assessing emissions from specific industrial sources and regional background levels in highly populated areas. This leads to sparse coverage throughout many rural areas of developed countries. Developing countries have minimal to no coverage. As a result, fixed-site monitors have limited utility in evaluating health effects in rural or developing areas.

Advancements in exposure science have led to improved characterization of air pollution exposure and reduced reliance on fixed-site monitoring to provide exposure estimates. Household level ambient measurements provide accurate data for small-scale studies and can be collected either by trained technicians (Breysse et al., 2005; Dietch et al., 2007; Wheeler et al., 2011; Williams et al., 2008) or by participants (Johnson et al., 2009; Petreas et al., 1988; Spengler et al., 1983; Sexton et al., 1986; Whitmore et al., 1999) but are too resource intensive to be used in large scale epidemiologic studies.

Air quality models have provided significant support for health studies and regulatory policy. Physical, mechanistic, and atmospheric air quality models use detailed source information, meteorology, and atmospheric chemistry to predict pollutant concentrations at a long-range spatial scale in both urban and rural areas (Arrandale et al., 2011; Bey et al., 2001; Bothe et al., 2005; Isakov and Ožkaynak, 2007; Jerrett et al., 2005; Kelly et al., 2012). Land use regression (LUR) and spatial interpolation models such as kriging or spline characterize the spatial distribution and health effects of air pollution at a local scale (Arrandale et al., 2011; Hoek et al., 2008; Jerrett et al., 2010; Johnson et al., 2010; Mejia et al., 2011). However, air quality models may exhibit high levels of uncertainty in areas where limited information is available. For example, natural emission sources such as biogenic soil NO_x or mineral dust remain uncertain. LUR and spatial interpolation models, which are most commonly used to assess exposure in health studies, require a dense monitoring network for model development and evaluation, and therefore are typically limited to urban areas.

Satellite remote sensing (RS) has emerged as a new tool for estimating ambient pollutant concentrations (Hidy et al., 2009; Hoff and Christopher, 2009; van Donkelaar et al., 2010). Pollutant concentrations based on satellite remote sensing utilize observational data that are collected daily at a global scale, providing consistently derived measurements in urban and rural areas in both developed and developing nations. Previous studies have demonstrated that long-term pollutant concentrations based on RS are significantly correlated with regulatory fixed site monitoring measurements (Liu et al., 2004; van Donkelaar et al., 2010; Lamsal et al., 2008; Kloog et al., 2012).

RS has previously been used to extend the capacity of LUR models. For example, RS has been used to develop national LUR models in the United States (Novotny et al., 2011) and Canada (Hystad et al., 2011) and to provide temporal refinement for LUR models (Kloog et al., 2011; Liu et al., 2009; Mao et al., 2012). RS has also been used to assess air pollution in developing countries where ground-based regulatory measurements are limited (Chu et al., 2003a, 2003b; Gupta et al., 2006; Kaiser et al., 2011; Mariano et al., 2010; Pereira et al., 2009; Xia et al., 2006).

However, few studies to date have examined the implications of using RS pollutant concentrations to estimate air pollution health effects. Crouse et al. (2012) demonstrated that long-term exposure to ambient PM2.5 based on RS was associated with increased cardiovascular mortality in the Canadian population. Kloog et al. (2012) reported that ambient PM2.5 concentrations during pregnancy based on RS were associated with adverse birth outcomes. Henderson et al. (2011) found that daily RS estimates of forest fire smoke exposure in British Columbia were positively associated with respiratory physician and hospital visits, although these associations did not reach statistical significance.

Here, we present a cross-sectional study that uses both satellite RS and ground-based regulatory monitoring data to examine respiratory health impacts of PM2.5 and NO_2 across Canada using multi-year ambient averages assigned to the home address of the study participant at the time of the health survey as a marker for long-term exposure to air pollution. The first objective was to examine the implications of using air pollution concentrations derived from RS to estimate exposure in a large-scale health study by comparing effect estimates based on RS and regulatory fixed-site monitoring. The second objective was to evaluate the utility of RS for assessing air pollution risk in remote and rural populations residing outside the catchment area for regulatory monitoring networks. We focused on respiratory disease due to the well-documented association between exposure to air pollution and respiratory health outcomes in urban areas (Health Effects Institute, 2010).
2. Methods

The study linked air pollution data from 1) satellite remote sensing and 2) ground-based regulatory monitors with residential location from the Canadian Community Health Survey (CCHS), a national sample of Canadians aged 12 and older. We pooled data from CCHS participants interviewed between 2001 and 2005. The household-level and person-level response rates for the CCHS survey years 2001–2005 ranged from 85 to 87% and 91 to 93%, respectively (Statistics Canada, 2013a,b,c). Multiple logistic regression models were used to assess associations between air pollution and respiratory and allergic health outcomes across Canada. The main analyses were restricted to respondents aged 20–64; however, results for younger and older groups were also examined.

2.1. Study population

To address the study objectives, we conducted analyses in three different populations for each pollutant of interest (PM$_{2.5}$ and NO$_2$). Due to differences in ground-based monitoring and RS coverage for different pollutants, the number of participants included in each group differed slightly for the two pollutants analyzed. The three populations were as follows:

1) **Respondents living within 40 km of a regulatory monitoring station** ($N = 123,039$ for PM$_{2.5}$; $N = 119,282$ for NO$_2$) were assigned estimates of pollution generated by both regulatory monitoring and RS. Analyses performed in this group were used to compare associations between air pollution and respiratory health outcomes based on RS versus ground-based regulatory monitoring.

2) **Respondents living further than 40 km from the nearest regulatory monitoring station** ($N = 88,750$ for PM$_{2.5}$; $N = 93,686$ for NO$_2$) lived outside the range recommended for regulatory monitoring coverage (Environment Canada, 2005; Environment Canada, 2011). This group was used to examine whether air pollution is associated with respiratory health effects in remote and rural areas (outside the catchment area of regulatory monitoring networks) using RS to estimate exposure.

3) **All respondents** ($N = 211,789$ for PM$_{2.5}$; $N = 212,968$ for NO$_2$) were used to examine associations between air pollution and respiratory/allergic health outcomes at a national scale using RS to estimate exposure.

2.2. Air pollution data

Fig. 1 shows the concentrations of PM$_{2.5}$ and NO$_2$ based on satellite RS as well as the locations of regulatory monitoring stations from the NAPS network during the study period. Estimates of PM$_{2.5}$ and NO$_2$ concentrations based on RS were assigned to CCHS participants by residential six digit postal code centroid in areas for which satellite-based concentrations were available. Pollution concentrations measured at regulatory monitoring stations were assigned to all respondents living within 40 km of the monitoring station. If more than one station was located within 40 km of the respondent, the concentration from the nearest monitor was assigned. Maps showing multi-year average RS concentrations of PM$_{2.5}$ and NO$_2$ for major cities in Canada are provided in the Supplemental Information (Suppl. Figs. 1 and 2).

2.2.1. Remotely sensed estimates of air pollution

Satellite RS concentrations for PM$_{2.5}$ were calculated using the methodology previously described by van Donkelaar et al. (2010). Ground-level PM$_{2.5}$ concentrations were derived from satellite measurements of aerosol optical depth (AOD) from the Moderate Resolution Imaging Spectroradiometer (MODIS) and the Multiangle Imaging Spectroradiometer (MISR) satellite instruments (Kahn et al., 2009; Levy et al., 2007) located on the Terra satellite which has been circumnavigating the globe and collecting measurements for approximately 10 years.

AOD is a measure of light extinction by atmospheric aerosols. AOD from MODIS (at 10 km resolution) and MISR (at 18 km resolution) were co-registered to a spatial resolution of $0.1^\circ 	imes 0.1^\circ$ (approximately 10 km × 10 km at mid-latitudes) using a within-grid area-weighted average to account for partial coverage of grid cells by individual satellite pixels. These gridded AOD were combined with linearly interpolated simulated aerosol vertical structure and scattering properties from the GEOS-Chem chemical transport model (www.geos-chem.org) to produce a global raster surface of PM$_{2.5}$. Estimates of PM$_{2.5}$ for each grid cell are based on at least 50 satellite retrievals averaged over the six-year period between 2001 and 2006. Yearly variation in sampling had less impact than the seasonal sampling cycles resulting from retrieval limitations over snow-covered regions. Both seasonal and yearly sampling bias were corrected using simulated sampling effects. The number of observations for each grid cell is typically 300–500 as shown in Fig. 7 of van Donkelaar et al. (2010). The global population-averaged average number of retrievals per grid cell (plus or minus standard deviation) is 297 ±202. The RS data were corrected for non-continuous sampling as described in van Donkelaar et al. (2010). RS derived long-term mean PM$_{2.5}$ values were in close agreement ($r = 0.8$, slope = 1.1, $n = 1057$) with in situ ground-based regulatory measurements in both Canada and the United States (van Donkelaar et al., 2010).

Satellite RS NO$_2$ concentrations were calculated by Lamsal et al. (2008, 2010) and inferred from tropospheric NO$_2$ columns retrieved from the Ozone Monitoring Instrument (OMI) on the Aura satellite, which has been collecting daily global measurements since 2004 with a resolution of 13 km × 24 km—26 km × 128 km. Data with a spatial resolution of finer than 19 km × 65 km were used here. Lamsal et al. (2008, 2010) derived ground level concentrations by applying local scaling factors from GEOS-Chem to tropospheric NO$_2$ measurements retrieved by OMI in 2005–2007 to produce a ground level NO$_2$ surface of long-term average NO$_2$ concentration with a spatial resolution of approximately $0.1^\circ 	imes 0.1^\circ$. Remotely sensed ground level NO$_2$ is significantly correlated with daily in situ measurements ($r = 0.3–0.9$; mean $= 0.7$, $N = 307$) with a tendency for higher correlations in polluted areas (Lamsal et al., 2008, 2010). The seasonal spatial correlation over these 307 sites is 0.8 (Lamsal et al., 2010).

2.2.2. Estimates derived from ground-level monitoring

Ground-level regulatory monitoring stations in the National Air Pollution Surveillance (NAPS) network were selected based on their availability of PM$_{2.5}$ and NO$_2$ data. A monitoring station was selected if at least half of the ground-level samples collected during the period of interest (2001–2006 for PM$_{2.5}$ and 2005–2007 for NO$_2$) were available for that station. Furthermore, stations used to estimate ambient concentrations had to provide approximately even coverage (approximately 50 ±10%) across warm and cold months (May to September, and November to March, respectively). Approximately 90 stations for PM$_{2.5}$ and 42 stations for NO$_2$ were excluded based on these criteria.

The closest station to the CCHS respondents — based on the latitude/longitude coordinate of their residential postal code centroid at the time of the health survey — was identified among the stations with valid measurements. The distance in kilometers between the ground-level monitoring station and the respondent’s
Postal code centroid was also determined for all the CCHS respondents.

As recommended by Environment Canada, the estimates of ambient air pollution derived from the ground-level monitoring stations were not considered representative of the respondents’ residential exposure if their postal code centroid was located further than 40 km from the closest station (Environment Canada, 2005; Environment Canada, 2011). In total, the estimates among respondents living within 40 km of a station originate from 120 different stations for NO₂ and 122 stations for PM₂.₅.

The calculation of PM₂.₅ averages has been described in depth by Crouse et al. (2012). Briefly, PM₂.₅ estimates (from 2001 to 2006) were calculated from the average of three samplers available at each station: TEOM, TEOM with Dryer only and BAM 35% RH. The values collected by the two TEOM samplers were adjusted for season and region, and daily averages were created only if 18 hourly samples or more were available.

Fig. 1. A,B Multi-year average concentrations of A) PM₂.₅ and B) NO₂ derived from RS in Canada. Maps depicting RS concentrations in major Canadian cities are provided in the Supplemental Information (Suppl. Figs. 1 and 2).
2.3. Health outcomes

Lifetime diagnosis of asthma, allergy and chronic bronchitis by a physician were self-reported by CCHS participants. Individuals who reported experiencing asthma symptoms in the 12 months preceding the survey were classified as having current asthma. Allergies were limited to non-food related allergies. For all health outcomes, missing values were classified as not having the adverse health outcome of interest.

2.4. Covariates

Individual level covariates were derived from the CCHS survey and were self-reported. Household level income adequacy and urban residence were derived using algorithms developed by Statistics Canada (2001). Income adequacy was based on household size and income, while participants residing in a continuously built up area with a total population greater than 1000 and a population density greater than 400 per square kilometer were classified as urban residents by Statistics Canada (2013a,b,c). All covariates, with the exception of body mass index (BMI) and household income adequacy, were modeled as binary variables. Missing BMI values were imputed as the median BMI for the respondent's age, sex, and ethnicity. Missing household income values (used to calculate income adequacy) were imputed as the median by age, sex, and household education. Missing values for binary covariates were assigned as the value corresponding to the largest sub-group. These imputations were made to minimize the impact of missing values on the calculation of air pollution health risks.

2.5. Neighborhood variables

Neighborhood level (ecological) variables were derived for each Dissemination Area (DA) based on socioeconomic and demographic characteristics from the 2001 Canadian census (Crouse et al., 2012). Values for the neighborhood level variables were assigned to participants based on the 2001 census DA corresponding to the residential postal code (Wilkins and Peters, 2012). These variables included the proportion of households in the highest and lowest income quintiles, proportion of homes in need of major repairs, proportion of the adult population in the following groups: currently unemployed, less than high school education, recent immigrants, recently moved, spoke English as a second language, or single parent. Neighborhood level variables did not contribute significantly to the model results and therefore were not included in the final models reported in this paper.

2.6. Exclusion criteria

Consistent with previous analyses reported by Crouse et al. (2012), the following groups were excluded from the analyses: 1) respondents living in the Yukon, Northwest Territories or Nunavut (N ~ 5000); 2) first generation immigrants (N ~ 51,000), and 3) respondents greater than 80 years old (N ~ 12,000). Respondents for whom immigration status was not available were also removed from the pool of CCHS respondents (N ~ 11,000). First, there is greater uncertainty about previous exposures among non-Canadian born respondents. Furthermore, first generation immigrants are more likely to live in large metropolitan areas with higher levels of ambient pollution, but also tend to have better health and health behaviors compared with the Canadian-born population (Ali et al., 2004; McDonald and Kennedy, 2004; Villeneuve et al., 2011). Second, the Yukon, Northwest Territories, and Nunavut were excluded because RS estimates of air pollution were not available for the northern areas of Canada. Finally, elderly respondents (80 years of age or older) were excluded because potential relocation to institutional or family assisted living introduced potential exposure misclassification for exposure based on residential location. The elderly were also excluded due to potential differences in their health profile (e.g., the survivor effect) due to attrition of less healthy individuals from the population.

2.7. Statistical analyses

Descriptive statistics for air pollution, health outcomes, and covariates were generated in each population of interest. We also examined correlations between household level ambient concentrations of PM2.5 and NO2 based on concentrations at the nearest monitoring station within 40 km and the RS concentration for the 10 x 10 km area grid occupied by the home.

Associations between ambient air pollution and chronic respiratory health outcomes were assessed using logistic regression models at the individual level. Multiple logistic regression models for adults aged 20–64 were adjusted for age, sex, ethnicity, BMI, urban residence, household socioeconomic status (SES variables included post-secondary education and income adequacy), and tobacco smoke (Equation (1)). Tobacco smoke variables included current smoking, past smoking, and exposure to second-hand smoke at home.

\[
\text{Health outcome} = \beta_0 + \beta_{\text{ap}} \times \text{ap} + \beta_{\text{age}} \times \text{age} + \beta_{\text{sex}} \times \text{sex} + \\
\beta_{\text{eth}} \times \text{eth} + \beta_{\text{edu}} \times \text{edu} + \beta_{\text{BMI}} \times \text{BMI} + \\
\beta_{\text{cs}} \times \text{cs} + \beta_{\text{ps}} \times \text{ps} + \beta_{\text{shs}} \times \text{shs} + \beta_{\text{ur}} \times \text{ur} + \epsilon
\]  

(1)

where ap = air pollution

age = age

sex = sex

eth = ethnicity

edu = education

ia = income adequacy

BMI = body mass index

cs = current smoker

ps = past smoker

shs = second hand smoke

ur = urban residence

Models for teenagers (aged 12–19) and older adults (aged 65–80) were adjusted for the same factors with the exception of BMI, which was not available for these age groups. Potential interactions of air pollution with gender, BMI, and income adequacy were also tested. All statistical analyses were conducted using SAS Enterprise Guide 4.2 and SAS 9.2 (Cary, North Carolina).

2.8. Multi-level spatial modeling

We used logistic-binomial models implemented by the paired-Poisson approach (Renjun et al., 2003) to examine spatial clustering of air pollution health risks in the total population (ages 20–64) at the census division level. The magnitude and significance of the associations between air pollution and health outcomes for multi-level models were similar to the results of the simpler multiple logistic regression models; therefore we presented results from the simpler models.

2.9. Sensitivity analyses

Sensitivity analysis of health outcome data was performed using individuals from the CCHS 2001 cycle for whom age of diagnosis
was available. These analyses were limited to respondents diagnosed with asthma in the 5 years preceding their interview.

We also conducted sensitivity analyses among respondents living within 10 km of the nearest regulatory monitoring station. These health models were limited to adults aged 20–64 in the 2001–2005 pooled dataset who lived within 10 km of the nearest regulatory monitoring station at the time of the CCHS interview.

### 3. Results

#### 3.1. Socio-demographic, health, and air pollution exposure in the study population

Table 1 provides descriptive statistics for respondents aged 20–64 in the total study population, and stratified by proximity to the nearest ground-based monitoring station. The stratification is based on proximity to PM$_{2.5}$ monitors; however, the same associations were observed when stratifying by NO$_2$ monitor location (results not shown).

Socio-demographic characteristics and disease prevalence varied by proximity to a ground-based monitor. The prevalence of both allergies and respiratory disease (with the exception of chronic bronchitis) was greater among respondents living within 40 km of a monitoring station compared with participants living further than 40 km from the nearest regulatory monitoring station ($p < 0.01$). Participants living within 40 km of a monitoring station were more likely to live in an urban area, be female, have some post-secondary education, and be at the extremes of income adequacy, while those living more than 40 km from a monitor were older, had higher BMI, and were more likely to be exposed to tobacco smoke ($p < 0.01$). Respondents living within 40 km of the nearest ground-level monitoring station were also exposed to higher levels of ambient pollution compared with those living further than 40 km from the nearest monitoring station ($p < 0.01$) as estimated with RS. This trend was consistent across pollutants. These results were consistent with expected differences between urban and rural populations.

The distribution of exposure to ambient pollution differed between estimates based on RS and nearest ground-level monitoring station. Briefly, estimates of PM$_{2.5}$ and NO$_2$ based on ground-level monitoring were normally distributed, whereas RS estimates were right-skewed. The range of ambient NO$_2$ concentrations based on ground-level monitoring was also greater than the range of estimates derived from RS (Table 1). RS mixing ratios averaged over roughly 100–1000 km$^2$ are substantially lower than those from point monitors, which tend to be located near emission sources to address regulatory objectives and which suffer interference from other reactive oxidized nitrogen species (EPA, 1975; Fehsenfeld et al., 1990; Dunlea et al., 2007; Steinbacher et al., 2007; Lamsal et al., 2008). Diurnal variation also plays a role since the RS measurement is made between 1 and 2 pm local time when NO$_2$ levels tend to be low due to photochemistry (Lamsal et al., 2008). Finally, the observed differences between RS and regulatory monitoring distributions may be due to the different spatial scales for these measurements — e.g., regulatory monitoring data represent point measurements, while RS values reflect average concentrations in 10 $\times$ 10 km grid cells.

In this study, we compared ambient concentrations assigned to each study home based on RS in a 10 $\times$ 10 km grid cell occupied by the home versus in situ measurements collected at the nearest monitoring station within 40 km of the home. Correlations between average household-level concentrations based on RS and measurements at the nearest regulatory monitoring were 0.73 and 0.70 for PM$_{2.5}$ and 0.58 and 0.53 for NO$_2$ at households within 10 km and 40 km, respectively, of the nearest regulatory monitoring station. The correlation between RS and in situ estimates decreased as the distance to the nearest monitor increased (Table 2). Scatterplots showing multi-year average concentrations based on RS versus regulatory monitoring at the regulatory monitoring sites are provided in the Supplemental Information (Suppl Figs. 3 and 4).

### Table 1

Socio-demographic characteristics, health outcomes, and household level pollution estimates for the study population (ages 20–64).

<table>
<thead>
<tr>
<th></th>
<th>Participants residing within 40 km of regulatory monitoring station ($N = 123,039$)</th>
<th>Participants residing more than 40 km from regulatory monitoring station ($N = 88,750$)</th>
<th>(Total) National study population ($N = 211,789$)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age (years)</strong></td>
<td>Mean (±SDV) 41.40 (±12.27)**</td>
<td>Mean (±SDV) 43.03 (±12.26)**</td>
<td>Mean (±SDV) 42.08 (±12.29)</td>
</tr>
<tr>
<td><strong>BMI</strong></td>
<td>25.90 (±4.94)**</td>
<td>26.60 (±5.11)**</td>
<td>26.19 (±5.03)</td>
</tr>
<tr>
<td><strong>Female</strong></td>
<td>66,170 (53.8%)</td>
<td>46,957 (52.9%)</td>
<td>113,127 (53.4%)</td>
</tr>
<tr>
<td><strong>Caucasian</strong></td>
<td>100,900 (82.0%)</td>
<td>72,577 (81.8%)</td>
<td>173,477 (81.9%)</td>
</tr>
<tr>
<td><strong>Post-secondary education</strong></td>
<td>86,201 (70.1%)</td>
<td>55,470 (62.5%)</td>
<td>141,671 (66.9%)</td>
</tr>
<tr>
<td><strong>Urban residence</strong></td>
<td>104,058 (84.6%)</td>
<td>48,720 (54.9%)</td>
<td>152,778 (72.1%)</td>
</tr>
<tr>
<td><strong>Income adequacy</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Lowest quintile</strong></td>
<td>17,281 (14.1%)**</td>
<td>19,162 (21.6%)</td>
<td>55,261 (26.1%)</td>
</tr>
<tr>
<td><strong>Highest quintile</strong></td>
<td>36,099 (29.3%)**</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Exposure to tobacco smoke</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Current smoker</strong></td>
<td>37,418 (30.4%)</td>
<td>28,994 (32.7%)</td>
<td>66,412 (31.4%)</td>
</tr>
<tr>
<td><strong>Past smoker</strong></td>
<td>51,573 (41.9%)</td>
<td>37,678 (42.5%)</td>
<td>89,251 (42.1%)</td>
</tr>
<tr>
<td><strong>Exposed to second-hand smoke</strong></td>
<td>12,442 (10.1%)</td>
<td>10,885 (12.3%)</td>
<td>23,327 (11.0%)</td>
</tr>
<tr>
<td><strong>Health outcomes</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Frequency (%)</strong></td>
<td>11,562 (9.4%)</td>
<td>3717 (8.1%)</td>
<td>18,779 (8.9%)</td>
</tr>
<tr>
<td><strong>Asthma (ever)</strong></td>
<td>6996 (5.7%)</td>
<td>4136 (4.7%)</td>
<td>11,322 (5.3%)</td>
</tr>
<tr>
<td><strong>Allergies</strong></td>
<td>37,766 (30.7%)</td>
<td>23,027 (26.0%)</td>
<td>60,793 (28.7%)</td>
</tr>
<tr>
<td><strong>Chronic bronchitis</strong></td>
<td>3801 (3.1%)</td>
<td>2708 (3.1%)</td>
<td>6599 (3.1%)</td>
</tr>
<tr>
<td><strong>NO$_2$ (ppb)</strong></td>
<td>Mean (±SDV) 11.24 (±4.72)</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td><strong>PM$_{2.5}$ (µg m$^{-3}$)</strong></td>
<td>8.74 (±2.19)</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td><strong>Remote sensing</strong></td>
<td>Mean (±SDV) 1.48 (±1.13)**</td>
<td>Mean (±SDV) 0.46 (±0.35)**</td>
<td>Mean (±SDV) 1.05 (±1.02)</td>
</tr>
<tr>
<td><strong>NO$_2$ (ppb)</strong></td>
<td>9.05 (±3.77)**</td>
<td>5.63 (±1.95)**</td>
<td>7.62 (±3.56)</td>
</tr>
<tr>
<td><strong>PM$_{2.5}$ (µg m$^{-3}$)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Significantly different at $p < 0.05$ level.

**Significantly different at $p < 0.01$ level.
Table 2
Correlation coefficients for ambient household level pollutant concentrations based on remote sensing and measurements at the nearest regulatory monitoring station by distance to the nearest monitoring station.

<table>
<thead>
<tr>
<th>Distance from home to nearest regulatory monitoring station (km)</th>
<th>Household level estimates of PM$_{2.5}$</th>
<th>Household level estimates of NO$_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 1</td>
<td>0.74 (N = 7002)</td>
<td>0.60 (N = 6954)</td>
</tr>
<tr>
<td>1–2</td>
<td>0.76 (N = 15,002)</td>
<td>0.60 (N = 15,242)</td>
</tr>
<tr>
<td>2–5</td>
<td>0.75 (N = 51,633)</td>
<td>0.60 (N = 49,287)</td>
</tr>
<tr>
<td>5–10</td>
<td>0.70 (N = 43,198)</td>
<td>0.54 (N = 41,275)</td>
</tr>
<tr>
<td>10–20</td>
<td>0.67 (N = 32,398)</td>
<td>0.44 (N = 29,899)</td>
</tr>
<tr>
<td>20–30</td>
<td>0.65 (N = 16,147)</td>
<td>0.36 (N = 17,608)</td>
</tr>
<tr>
<td>30–40</td>
<td>0.60 (N = 16,070)</td>
<td>0.25 (N = 15,886)</td>
</tr>
</tbody>
</table>

3.2. Air pollution health effects

Figs. 2–4 show risk estimates for respiratory and allergic health outcomes associated with PM$_{2.5}$ and NO$_2$ based on RS and ground-level monitoring. To facilitate comparison between risk estimates based on RS and ground level monitoring, all odds ratios were expressed per interquartile range (IQR) of the pollutant in the population of interest.

Fig. 2 shows adjusted odds ratios for PM$_{2.5}$ and NO$_2$ derived from both RS and regulatory measurements among participants aged 20–64 living within 40 km of a monitoring station. PM$_{2.5}$ and NO$_2$ exposures based on RS were associated with a 6–8% increased risk per IQR for respiratory health outcomes. Odds ratios and 95% confidence intervals for RS-based NO$_2$ were as follows: asthma: 1.06 (1.03–1.08); current asthma: 1.06 (1.02–1.09); allergies: 1.07 (1.06–1.09); and chronic bronchitis: 1.07 (1.03–1.12) all calculated over the IQR of 1.46 ppb. Odds ratios for RS-based PM$_{2.5}$ were: asthma: 1.07 (1.04–1.10); current asthma: 1.07 (1.03–1.11); allergies: 1.08 (1.06–1.10); and chronic bronchitis: 1.10 (1.02–1.13) all calculated over IQR 5.94 µg m$^{-3}$. Effect estimates based on ground-level monitoring were similar to RS-based estimates for asthma and current asthma, and lower for allergies and chronic bronchitis (calculated over the IQRs of 3.24 µg m$^{-3}$ for PM$_{2.5}$ and 5.32 ppb for NO$_2$).

Fig. 3 shows adjusted odds ratios for PM$_{2.5}$ and NO$_2$ derived from RS among participants living further than 40 km from a ground-based monitoring station over the IQR in pollutant concentration: for NO$_2$, the IQR was 0.33 ppb and for PM$_{2.5}$, 1.62 µg m$^{-3}$. Models were adjusted for age, sex, ethnicity, body mass index (BMI), urban residence, household SES (post-secondary education and income adequacy), and tobacco smoke (current smoking, past smoking, and exposure to second-hand smoke).

In the total national study population (Fig. 4), PM$_{2.5}$ and NO$_2$ based on RS were associated with a 6–10% increase per IQR in respiratory health outcomes. Odds ratios for NO$_2$ were: asthma: 1.07 (1.05–1.08); current asthma: 1.08 (1.06–1.10); allergies: 1.09 (1.08–1.10); and chronic bronchitis: 1.06 (1.03–1.08) over IQR 1.03 ppb.
Odds ratios for PM$_{2.5}$ were: asthma: 1.08 (1.06–1.10); current asthma: 1.09 (1.07–1.12); allergies: 1.10 (1.09–1.11); and chronic bronchitis: 1.07 (1.04–1.10) calculated over IQR 3.97 $\mu$g m$^{-3}$.

3.3. Sensitivity analyses

We also examined health models that were limited to respondents living within 10 km of the nearest monitoring station. Risk estimates based on both RS and regulatory monitoring were slightly (2–3%) higher than those found in the larger population living within 40 km of the nearest monitoring stations. However, risk estimates based on RS and regulatory monitoring for respondents living within 10 km and within 40 km of the nearest monitoring station were highly similar across pollutants and health outcomes (data not shown).

Associations between air pollution and asthma were similar for respondents diagnosed within 5 years of the survey compared with those with a lifetime history of doctor-diagnosed asthma (data not shown). Associations between air pollution and health outcomes were weaker among younger (12–19 year old) and older (65–80 year old) age groups for all study populations. This may be due to higher residential mobility in these age groups resulting in greater exposure misclassification (because home address was used to assign exposure), lower disease prevalence (resulting in less statistical power to identify associations), or decreased susceptibility to air pollution (e.g., due to a healthy survivor effect among elderly participants). We also examined the impact of clustering by census district through multi-level spatial modeling and the results were similar to those for the multivariate logistic models presented in the paper; therefore we presented results for the simpler models.

4. Discussion

Previous air pollution studies have primarily focused on urban populations because urban dwellers are typically exposed to higher levels of ambient air pollution, and because the resource burden associated with estimating exposure in rural and remote areas is prohibitive. This study examined the implications of using RS to examine associations between air pollution and allergic/respiratory health outcomes among Canadians living in both urban and rural areas.

4.1. Remote sensing versus in situ measurements of air pollution

Remote sensing estimates of ambient PM$_{2.5}$ and NO$_2$ have been validated in previous studies which reported strong correlations between ground-based in situ measurements and RS concentrations for both pollutants (Lamsal et al., 2008, 2010; van Donkelaar et al., 2010; Lee et al., 2012). RS measures of both PM$_{2.5}$ and NO$_2$ were significantly correlated with in situ measurements, with average correlations of 0.8 for PM$_{2.5}$ (Van Donkelaar et al., 2010) and 0.7 for NO$_2$ (Lamsal et al., 2008, 2010) in North America. As in previous studies, mean NO$_2$ concentration based on RS were significantly lower than in situ measurements. There are a number of methodological differences that contribute to this discrepancy. While the RS estimates are an area average, the ground-based estimates are based on fixed-site regulatory monitors, which tend to be preferentially located near roads and other locations with significant human activities where NO$_2$ concentrations are higher. The RS estimate is recorded between 1 and 2 pm local time when NO$_2$ levels tend to be low, while the ground-based measurements are averaged across the day. Ground-based monitors also use commercial chemiluminescence analyzers which overestimate true NO$_2$ due to interference from other reactive nitrogen species. In addition, RS estimates exhibit higher uncertainty for snow and ice conditions in winter, resulting in greater exclusion of winter values from long-term averages. NO$_2$ concentrations are typically higher in winter; therefore a higher rate of exclusion for winter values likely results in lower average concentrations for NO$_2$ based on RS. However, seasonal NO$_2$ concentrations based on RS and in situ concentrations were highly correlated (Lamsal et al., 2008, 2010).

We compared air pollution estimates assigned to study participants based on RS versus regulatory monitoring to examine the implications of estimating household level concentrations based on these alternate methods. As expected, estimates of ambient air pollution at participant homes based on RS and regulatory monitoring were highly correlated for participants living near a regulatory monitoring station where the concentrations measured at the monitoring site were more likely to reflect household level ambient concentrations. As the distance between the participant home and the nearest regulatory monitoring site increased, the correlation between exposure estimates based on RS and regulatory monitoring decreased. These results suggest that for homes located near a monitoring station both regulatory monitoring data and RS provided good estimates of household level ambient concentration. However, for homes located further from regulatory monitoring stations, RS provided a better estimate of household-level ambient concentrations for PM$_{2.5}$ and NO$_2$.

Although these analyses show a high correlation between household estimates based on remote sensing and regulatory monitoring, ground level in situ measurements collected at regulatory monitoring stations were not used to further evaluate RS because the spatial scales for RS and regulatory monitoring measurements were not directly comparable. Estimates based on
regulatory monitoring data were point measurements collected at a regulatory monitoring station that were applied to participants living up to 40 km from the a monitoring station, based on the self-described catchment area for these regulatory monitoring stations. In contrast, the RS estimates were averages calculated over areas of approximately 10 × 10 km (Fig. 1) and assigned to participants residing within those area grids.

4.2. Air pollution health effects

Long-term exposure to air pollution based on satellite RS was associated with a 6–10% increase in the prevalence of allergies, asthma, current asthma, and bronchitis per IQR for PM$_{2.5}$ and NO$_2$ among Canadian adults aged 20–64 years. These results are consistent with previous research demonstrating respiratory and allergic health impacts associated with air pollution (Brunekreef and Holgate, 2002; Byrd and Joad, 2006; Díaz-Sanchez et al., 2003; Gauderman et al., 2005; Health Effects Institute, 2010; McConnell et al., 2006; Samal et al., 2008; Samet, 2007; Sarnat and Holguín, 2007; Wong and Lai, 2004).

Absolute values for pollutant concentrations at the study homes differed between RS and ground-based monitoring. However, air pollution risk estimates scaled by pollutant IQR were similar for the two methods among the population residing near a regulatory monitoring station, because household level concentrations based on RS and regulatory monitoring were highly correlated. For both PM$_{2.5}$ and NO$_2$, odds ratios for asthma, current asthma and allergies were similar based on RS and ground level monitoring. Effect estimates for chronic bronchitis displayed greater variability, with significant associations only evident based on RS. However, effect estimates for chronic bronchitis displayed based on RS and regulatory monitoring were not significantly different based on overlapping confidence intervals. Risk estimates based on RS appeared to be slightly stronger than those based on regulatory monitoring, which may be due to the improved spatial scale of RS exposure estimates compared with concentrations measured at the nearest monitoring site. These results suggest that RS is a comparable metric to ground-based regulatory monitoring for estimating long-term exposure to ambient pollution in a health study.

Crouse et al. (2012) previously reported a significant increase in mortality associated with long-term exposure to air pollution based on RS. The findings of the current study suggest that RS can be used to assess air pollution impacts on morbidity as well as mortality.

Importantly, air pollution health effects in this study were not limited to urban populations with higher pollution levels. Although the prevalence of asthma and allergies was higher among study participants residing within 40 km of a ground-level monitoring station (in predominantly urban areas), the magnitude of air pollution risk estimates was similar among participants living further than 40 km from a ground-based monitoring station. This suggests that air pollution has adverse allergic and respiratory effects in rural and remote areas with lower ambient concentrations.

4.3. Limitations

Our results suggest that RS can provide a powerful tool for estimating long-term exposure to both gaseous and particulate air pollution. However, there are some limitations in the analyses.

The spatial resolution of the pollutant estimates based on remote sensing was limited by the current retrievals from available satellite measurements. The 10 × 10 km resolution for PM$_{2.5}$ was likely adequate for estimating exposure to ambient pollution based on the spatial homogeneity of PM$_{2.5}$. However, the 10 × 10 km resolution for NO$_2$ is problematic given the spatial heterogeneity and highly localized sources of NO$_2$. The risk estimates reported in this study can at best capture the effects of average NO$_2$ concentration at a community level, and therefore likely underestimate the “true” impact of NO$_2$ on these health outcomes. As remote sensing technology and retrievals improve, future analyses will be needed to examine air pollution health effects associated with ambient concentrations at a smaller spatial scale. The use of satellite retrieval is still a novel method for estimating ambient air pollution, and is being constantly improved. For example, enhancements in RS-based PM$_{2.5}$ have been observed in the southern prairies and western mountains due to the effects of surface reflectance on the retrieval. This issue has been addressed in recent retrievals (van Donkelaar et al., 2013). There was no evidence to suggest that surface reflectance impacted the analyses reported in this paper.

The goal of this paper was not to evaluate the catchment area for the regulatory monitoring values, but rather to examine the implications of using RS to estimate long-term exposure in comparison with regulatory monitoring values. We therefore used the relatively large 40 km catchment area prescribed by Environment Canada. However, we found that effect estimates for RS and regulatory monitoring were similar when the study population was restricted to participants residing within 10 km of the nearest monitoring station. This suggests that the results were not biased by the large catchment area used for the analyses reported in this paper.

Finally, long-term average pollutant concentrations used to predict exposure represented multi-year averages for PM$_{2.5}$ (2001–2006) and NO$_2$ (2005–2007) based on the availability of remotely sensed estimates of air pollution. These averaging periods overlapped (rather than preceded) the health survey period (2001–2005). The analyses assumed that relative spatial patterns in long-term ambient pollution concentrations were stable during the five-year study period. In other words, we assumed that areas with high pollutant concentrations were consistently high relative to low pollution areas during the health survey period.

This assumption was based on the robustness of multi-year pollution concentrations estimated using RS and on the coarse spatial resolution of the RS estimates. For example, Crouse et al. (2012) reported that the 1987–2001 PM$_{2.5}$ average based on regulatory monitoring in 11 of Canada’s largest cities was highly correlated with the 2001–2006 RS PM$_{2.5}$ ($r = 0.89$). Furthermore, associations between air pollution and health reported in this paper were consistent for individual years within the cohort (e.g., respondents surveyed in 2001, 2003, and 2005 survey cycles), suggesting that exposure misclassification generated by using pollutant concentrations measured during the health survey period did not change the results of these analyses.

Finally, exposure misclassification may have been introduced by the lack of residential history needed to accurately characterize historic exposure and inability to verify that exposure preceded diagnoses. However, analyses restricted to respondents diagnosed within the five-year period preceding the survey showed similar results.

Despite our efforts to evaluate the air pollution risk estimates reported in this study, these associations may be biased by self-reporting, exposure misclassification associated with residential mobility, and variation in spatial distribution of air pollution over time. However, the primary goal of this study was not to quantify the association between air pollution and respiratory disease, which has been well characterized in the literature, but rather to examine the utility of using RS to estimate exposure to air pollution in a health study.

5. Conclusions

Long-term exposure to ambient air pollution as estimated by ground-based regulatory monitoring and satellite RS was associated with increased prevalence of allergic and respiratory health
outcomes among Canadian adults. These associations were significant in predominantly urban areas with higher ambient concentrations and in rural and remote areas with lower concentrations. Effect estimates were also similar for models using satellite RS versus ground-based regulatory monitoring to estimate exposure.

The consistency between risk estimates based on RS and regulatory monitoring as well as the associations between air pollution and health outcomes in remote areas demonstrate the utility of air pollution estimates derived from RS for characterizing long-term ambient air pollution in rural communities and other areas for which ground-level air monitoring data are not available. RS provided a powerful tool for assigning consistently derived estimates of long-term exposure across both urban and rural populations.

Acknowledgments

The authors would like to thank Stan Judek for compiling the regulatory monitoring data and Dominic Odwai Atari for providing the neighborhood level variables, as well as Dan Crousean Hwa Shin for providing feedback.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.atmosenv.2013.07.020.

References
