Mapping the spatial distribution of tropospheric ozone and exploring its association with elevation and land cover over North Jordan

Salahuddin M. Jaber & Mahmoud M. Abu-Allaban

To cite this article: Salahuddin M. Jaber & Mahmoud M. Abu-Allaban (2017) Mapping the spatial distribution of tropospheric ozone and exploring its association with elevation and land cover over North Jordan, Journal of Spatial Science, 62:2, 307-322, DOI: 10.1080/14498596.2016.1277560

To link to this article: http://dx.doi.org/10.1080/14498596.2016.1277560

Published online: 01 Feb 2017.

Submit your article to this journal

Article views: 51

View related articles

View Crossmark data
Mapping the spatial distribution of tropospheric ozone and exploring its association with elevation and land cover over North Jordan

Salahuddin M. Jaber and Mahmoud M. Abu-Allaban

Department of Water Management and Environment, Faculty of Natural Resources and Environment, Hashemite University, Zarqa, Jordan

ABSTRACT
This study aims to map the spatial distribution of tropospheric ozone (O₃) in the northern parts of Jordan and explore the types of the relationships between the observed spatial patterns and elevation and land cover. Summer and daytime tropospheric O₃ data were procured following a stratified random sampling strategy. The collected data were corrected for temporal variations using piecewise linear regression. Empirical Bayesian Kriging showed that tropospheric O₃ concentrations are high in the west, where agricultural and urban areas dominate, relative to the east, where rangelands and barren lands dominate. One-way analysis of variance followed by means comparison using Tukey–Kramer HSD. test emphasized the previous observation suggesting a proportional relationship between agricultural and urbanization activities and high tropospheric O₃ concentrations. However, simple linear regression, polynomial regression, and geographically weighted regression showed that the relationship between elevation and tropospheric O₃ is complex and not clear and not amenable to direct explanation.

1. Introduction
Since the great London Smog of 1952, air pollution has been an active area of research due to its adverse environmental, social, and economic impacts (Davis et al. 2002). Tropospheric ozone (O₃) plays a key role in atmospheric chemistry, air quality, and radiative balance as an oxidant and a greenhouse gas. It occurs close to the earth’s surface in the troposphere as a result of photochemical reactions in the presence of sunlight and precursor pollutants, such as nitrogen oxides (NOx) and volatile organic compounds (VOC) (WHO 2006; EPA 2011).

Tropospheric O₃ is associated with adverse impacts on human health, including increased rates of hospital admissions, exacerbation of respiratory illness, acute and chronic health problems related to lung functions, asthma, and pulmonary infection, and eye, nose, and throat irritation (Desqueyrroux et al. 2002; Bell et al. 2004; Lu and Wang 2004; Oltmans et al. 2006). Exposure to high levels of tropospheric O₃ over extended periods of time increases...
mortality rates (O’Neill et al. 2004; Stedman 2004). Tropospheric O₃ also has significant destructive effects on vegetation and forests, such as altering the immune systems in plants, and reducing photosynthesis, growth rate, and crop yield, in addition to changing species composition within forested areas (Hogsett et al. 1997; Mauzerall and Wang 2001; Ashmore 2005). Moreover, tropospheric O₃ may cause cracks in rubber products by attacking the carbon double bond in natural rubber monomers (Bouble et al. 1994).

As recommended by the British Royal Society (RS 2008), significant investment in tropospheric O₃ research by national governments and international agencies is required to improve the evidence base and address the many science and research needs and gaps of tropospheric O₃. Nevertheless, tropospheric O₃ research in Jordan is unfortunately scarce. Consequently, the present study aims to achieve two major goals. The first goal is to map the spatial distribution of tropospheric O₃ in the northern parts of Jordan; results obtained from the very few previous studies that were conducted in Jordan (e.g. Weinroth et al. 2008; Aljaiatwah 2011; Alsawair 2011) showed that the concentrations of tropospheric O₃ in the northern parts of the country exceeded the standards recommended by the World Health Organization (WHO). The second goal is to explore the types of the relationships between the observed spatial patterns of tropospheric O₃ and elevation and land cover. The present study will fill gaps represented by poor tropospheric O₃ emission inventories in many developing countries. Policy and environmental mitigation and adaptation plans for controlling and minimizing the adverse impacts of tropospheric O₃ nationally and internationally will also benefit from the findings of this study.

2. Data and methods

2.1. Overall approach and study area

The overall approach to this study consisted of (1) developing elevation and land cover surfaces for the northern parts of Jordan, (2) procuring tropospheric O₃ data using a relatively inexpensive handheld portable O₃ monitor, (3) correcting the procured tropospheric O₃ data for temporal variations using regression analysis, (4) constructing a continuous surface showing the spatial distribution of corrected tropospheric O₃ data using kriging analysis, and (5) performing regression analyses for exploring the relationships between the observed spatial patterns of corrected tropospheric O₃ data and elevation and land cover.

The above-mentioned approach was implemented in the northern parts of Jordan – an area that covers approximately 3,849 km². It extends from the Jordan Valley in the west to Safawi City in the east between latitudes 32°05’ N and 32°45’ N and longitudes 35°30’ E and 37°00’ E (Figure 1).

2.2. Elevation data

Elevation data for the study area in metres were extracted from the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) Global Digital Elevation Model (GDEM) Version 2 data. They were developed and made available to the public by the Ministry of Economy, Trade, and Industry (METI) of Japan and the United States National Aeronautics and Space Administration (NASA) through the Earth Remote Sensing Data Analysis Center
The ASTER GDEM Version 2 is a very large product. It is comprised of 22,702 1º-by-1º tiles that cover the land surfaces of the earth between 83º N and 83º S. Each tile container accommodates two files; a Digital Elevation Model (DEM) file and a Quality Assessment (QA) file. The DEM data are stored as 16-bit signed integer in GeoTIFF format with geographic latitude/longitude coordinates and a 1 arc-second (approximately 30 m at the equator) grid of elevation postings (ASTER GDEM 2011).

First, two ASTER GDEM Version 2 tiles that cover the land surfaces of northern Jordan between latitudes 32º N and 33º N and longitudes 35º E and 37º E were downloaded from the ERSDAC website. The tiles were then mosaiced and re-projected into the Jordan Transverse Mercator (JTM) coordinate system and clipped according to the borders of the study area.

### 2.3. Land cover data

The relatively new Deimos-1 is a Spanish earth imaging satellite. Deimos-1 was successfully launched on 29 July 2009 into a sun-synchronous low earth orbit with an expected lifetime of five years. The satellite carries a multispectral imager, which delivers data in three spectral bands (i.e. green (0.520–0.600 μm), red (0.630–0.690 μm), and near infrared (0.770–0.900 μm)). One major advantage of Deimos-1 over the other operating multispectral satellites is its unprecedented wide swath of 625 km, which enables the rapid coverage and revisit of large
areas. Hence, Deimos-1 was tasked with collecting imagery for the study area. The image was acquired on 28 Jun 2012 between 11:06:51 am and 11:07:27 am (local time). The data consisted of radiometrically corrected data projected to the Universal Transverse Mercator (UTM) Zone 37 coordinate system and stored in 8-bit unsigned integer radiance values. It contained 6,720 pixels (columns) in the across-track direction and 5,314 pixels (rows) in the along-track direction with nominally a 22 m ground sampling distance (ELECNOR DEIMOS 2013).

Deimos-1 imagery was first re-projected into the JTM coordinate system and then was clipped according to the borders of the study area. The hybrid method combining the conventional unsupervised ISODATA and the supervised maximum likelihood algorithms was applied to classify the Deimos-1 imagery and produce a land cover map for the study area (Lillesand et al. 2008). To implement the ISODATA algorithm, the number of classes, the convergence threshold, and the number of iterations were set to be 30, 0.95, and 30, respectively. After 12 iterations, the 0.95 convergence threshold was achieved and the image was clustered into 30 spectral classes. In addition, an output signature set file comprising 30 signatures representing the 30 spectral classes was generated. Intensive ground truthing led to the adoption of the following information classes: (1) urban land; (2) agricultural land; (3) rangeland; (4) barren land; and (5) water. The definitions of these classes are based on the United States Geological Survey (USGS) Land Use Land Cover (LULC) Classification System (Anderson et al. 1976). After an in-depth evaluation and ground truthing, the 30 spectral classes were assigned to the five land cover information classes. However, moderate-to-heavy spectral mixing was noticed between some of the spectral classes. Therefore, only the spectral classes with little mixing were kept for further analysis. Hence, out of the 30 spectral classes, only nine spectral classes representing barren lands and water along with their associated signatures were kept. Based on experience in the study area and intensive ground truthing, a total of 205 training samples representing urban lands, agricultural lands, and rangelands were selected from the processed Deimos-1 imagery. In an iterative process, the 205 training samples were evaluated and merged and a total of 10 distinct signatures representing the three land cover classes were produced. Hence, a total of 19 distinct signatures representing all the five land cover classes were generated from the hybrid ISODATA clustering and training sample selection methodologies. Nine signatures (representing barren lands and water) were extracted from the ISODATA clustering technique and 10 signatures (representing urban lands, agricultural lands, and rangelands) were extracted from the training sample selection technique. These 19 signatures were plugged into the maximum likelihood algorithm to classify the image into 19 spectral classes. Each class corresponds to a signature. Further in-depth evaluation and ground truthing led to the combining of the 19 spectral classes into only the five land cover information classes. To remove the salt-and-pepper noise and enhance the quality and appearance of the resulting land cover thematic map, a 3 × 3-pixel majority spatial filter was applied.

Following a stratified random sampling strategy, a total of 403 reference pixels were generated to evaluate the accuracy of the land cover thematic map by constructing an error matrix using the five land cover classes (Jensen 2005). The original image, auxiliary data, and ground truthing were used to assess the accuracy of these reference pixels.
2.4. Tropospheric ozone data

Summer and daytime tropospheric O$_3$ data in parts per billion (ppb) were procured following a stratified random sampling strategy (Figure 1). First, the study area was randomly divided into 188 cells with sizes equal to 0.05° × 0.05°. During performing the fieldwork campaign, which lasted from 7 June 2012 to 11 July 2012, factors of logistics and accessibility conditions limited the number of sampled cells to only 158. In each sampled cell, one tropospheric O$_3$ observation was recorded using the handheld S-500L portable O$_3$ monitor (from Ozone Solutions, Inc.). Each tropospheric O$_3$ observation is merely the arithmetic mean of five consecutive tropospheric O$_3$ readings recorded at the same location during five minutes (i.e. one reading every one minute). The resulted variable was termed tropospheric O$_3$ before correction (OBC). It is noteworthy to mention that the location of each tropospheric O$_3$ observation inside each sampled cell was chosen in the field randomly and recorded using the handheld GPSMAP 60CSx portable global positioning system (GPS) device (from Garmin International, Inc.) with accuracy of less than 5 m.

It is well known that tropospheric O$_3$ varies not only spatially but also temporally (e.g. Tang et al. 2012). Therefore, the days and times (i.e. day of year in Julian days and time of day in minutes) of the 158 tropospheric O$_3$ observations were recorded. Actually, each day and time of each tropospheric O$_3$ observation is merely the arithmetic mean of the days and times of the five consecutive tropospheric O$_3$ readings recorded at the same location. The resulted variable was termed time (Time). Then, piecewise linear regression (PLR) analysis was applied to develop a relationship between OBC and Time. This relationship was used to correct OBC for temporal variations and report it as tropospheric O$_3$ after correction (OAC) using the day and time of the acquisition of the centre of the Deimos-1 imagery as the reference day and time. Piecewise linear regression (NCSS 2015) is a form of regression that allows multiple linear models to be fit to the data for different ranges of the independent variable. The values of the independent variable where the linear models meet are called knots or breakpoints. In the present study, the model’s parameters were estimated by non-linear regression using the standard method of least squares in an iterative manner.

Finally, Empirical Bayesian Kriging (EBK) was applied to produce a continuous surface with JTM coordinate system, which shows the spatial distribution of OAC in the study area. Empirical Bayesian Kriging (ESRI 2012; Krivoruchko 2012) is a geostatistical interpolation technique whereby the methodology of building a valid kriging model is automated with minimal user interaction through a process of subsetting and simulations. In the present study, during applying the EBK the default values for different inputs were kept; that is the subset size was set to 100, the overlap factor was set to 1, and the number of simulations was set to 100, with a power semivariogram type and no transformation and a smooth circular search neighbourhood type with smoothing factor of one.

2.5. Regression analysis

In order to explore the type of the relationship between OAC and elevation, multiple procedures of regression analyses were applied. They are: (1) simple linear regression; (2) polynomial regression; and (3) geographically weighted regression.

Simple linear regression (SLR) (e.g. Zou et al. 2003) explores the linear relationship between the dependent variable and single independent variable. In the present study, the standard
method of least squares was applied to estimate the regression parameters. Polynomial regression (PR) (e.g. Rawlings et al. 1998) is used when the relationship between the dependent variable and independent variable is curvilinear. In the present study, polynomial models in increasing order were fit successively and the significance of the regression coefficients were tested at each step of model fitting and the order kept increasing until the t-test for the highest order term is not significant. Geographically weighted regression (GWR) (Fotheringham et al. 2002) is a spatial regression technique applied between some sets of variables when spatial non-stationarity is assumed. It allows for different regression coefficients and consequently different regression diagnostics to be calculated for different observations using a moving kernel window approach. In the present study, an adaptive spatial kernel with a bandwidth determined by minimizing the corrected Akaike’s information criterion (AICc) was used.

The type of the relationship between OAC and land cover was explored by comparing the means of OAC for the four land cover types (i.e. urban land, agricultural land, rangeland, and barren land) by following a two-stage process (Abdi and Williams 2010). In the first stage, one-way analysis of various (ANOVA) was applied to answer the question: ‘Are all the means equal?’ In the second stage, when ANOVA gives a significant result indicating that at least one land cover differs from the other land covers, the Tukey–Kramer Honestly Significant Difference (HSD) test was applied to answer the question: ‘Which means are unequal?’

3. Results

3.1. Elevations variability in the study area

The developed DEM (Figure 2(a)) shows that the study area extends from the Jordan Valley in the west, with minimum elevations of −321 m below mean sea level, and passes through the mountains of Ajlun and Jarash, where maximum elevations reach about 1,224 m above mean sea level, until the elevations flatten in the east, with mean value of about 750 m above mean sea level.

3.2. Land cover characteristics of the study area

Land cover mapping produced a land cover thematic map shown in Figure 2(b). The overall accuracy of the map was about 85.9% and is shown in Table 1. Urban lands were found to have the highest user’s accuracy of about 89.3%, with a commission error of about 10.7%. Barren lands were found to have the highest producer’s accuracy of about 90.7%, with an omission error of about 9.3%. However, the lowest user’s accuracy was calculated for the water class of about 80.0%, with a commission error of about 20.0% and the lowest producer’s accuracy was calculated for the agricultural lands class of about 80.3%, with an omission error of about 19.7%.

As shown in Figure 2(b) and Table 2, more than half the area of the study area is covered by rangelands. They dominate the eastern parts of the study area. Agricultural lands cover about one-third the area of the study area and dominate its western parts. Barren lands cover about 9.3% of the study area. They exist primarily in its central and southern parts. Urban lands cover about 5.7% of the study area and constitute primarily Irbid and Mafraq cities and their suburbs. They exist primarily in the western parts of the study area. Water...
Figure 2. Elevation (a) and land cover (b) maps for the study area.

Table 1. Error matrix showing the accuracy of the developed land cover thematic map for the study area.

<table>
<thead>
<tr>
<th></th>
<th>Reference data</th>
<th>User's accuracy (%)</th>
<th>Commission error (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>UL</td>
<td>AL</td>
<td>RL</td>
</tr>
<tr>
<td>Classified data</td>
<td>25</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>AL</td>
<td>1</td>
<td>102</td>
<td>14</td>
</tr>
<tr>
<td>RL</td>
<td>1</td>
<td>23</td>
<td>172</td>
</tr>
<tr>
<td>BL</td>
<td>3</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>WA</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Column total</td>
<td>30</td>
<td>127</td>
<td>194</td>
</tr>
<tr>
<td>Producer's accuracy (%)</td>
<td>83.33</td>
<td>80.31</td>
<td>88.66</td>
</tr>
<tr>
<td>Omission error (%)</td>
<td>16.67</td>
<td>19.69</td>
<td>11.34</td>
</tr>
</tbody>
</table>

Note: UL: Urban Land, AL: Agricultural Land, RL: Rangeland, BL: Barren Land, WA: Water, Overall Accuracy = 85.86%.

Table 2. Land cover classes in the study area and their coverage areas and percentages.

<table>
<thead>
<tr>
<th>Land cover class</th>
<th>Area (km²)</th>
<th>Area (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Urban land</td>
<td>219.390</td>
<td>5.700</td>
</tr>
<tr>
<td>Agricultural land</td>
<td>1124.587</td>
<td>29.219</td>
</tr>
<tr>
<td>Rangeland</td>
<td>2145.962</td>
<td>55.756</td>
</tr>
<tr>
<td>Barren land</td>
<td>358.352</td>
<td>9.311</td>
</tr>
<tr>
<td>Water</td>
<td>0.565</td>
<td>0.015</td>
</tr>
<tr>
<td>Total</td>
<td>3848.856</td>
<td>100.001</td>
</tr>
</tbody>
</table>
covers only small portions of the study area (i.e. less than 0.02%) and exists mainly in the
dams and Jordan and Yarmouk rivers that form the western and parts of the northern borders
of the study area respectively.

3.3. The spatial distribution of tropospheric ozone in the study area

The values of OBC vary from 7.6 to 73.8 ppb with mean and standard deviation of
about 39.9 ppb and 11.9 ppb respectively. They were found to be normally distributed
(Figure 3(a)).

After 17 iterations, piecewise linear regression provided an acceptable significant model
at the 0.05 significance level, with a coefficient of determination \( R^2 \) of about 0.3320 and
root mean square error (RMSE) of about 9.6914 ppb (Equation (1) and Figure 3(b)). The model
defined the relationship between OBC and Time with three linear segments and two
breakpoints. The first breakpoint is at Time equals 165.6076 Julian days and the second is at
Time equals 179.4403 Julian days. The residuals associated with the model were close to
normal (Figure 3(c)), showed no clear signs of heteroscedasticity or non-independence
(Figure 3(d)), and had a mean (ME) close to zero (about –0.0031 ppb).

Using Equation (1) and knowing that the slope could be represented by Equations (2) and
(3) and using the day and time of the acquisition of the centre of the Deimos-1 imagery (i.e.
180.4633 Julian days) as the reference day and time, OBC was corrected by applying Equation
(4) and OAC was obtained.

\[
OBC = \begin{cases} 
-433.6850 + 2.8585 \times Time & \text{Time } < 165.6076 \\
-48.5245 + 0.5328 \times Time & 165.6076 \leq Time \leq 179.4403 \\
204.6961 - 0.8784 \times Time & \text{Time } > 179.4403 
\end{cases} 
\]  
(1)

\[
Slope = \frac{dOBC}{dTime} = \begin{cases} 
2.8585 & \text{Time } < 165.6076 \\
0.5328 & 165.6076 \leq Time \leq 179.4403 \\
-0.8784 & \text{Time } > 179.4403 
\end{cases} 
\]  
(2)

\[
Slope = \frac{\Delta OBC}{\Delta Time} = \frac{OAC - OBC}{180.4633 - Time} 
\]  
(3)

Figure 3. Diagnostic plots for the model developed for correcting OBC for time variations, which resulted
from applying piecewise regression with three linear segments. The solid lines in (a) and (c) are the 1:1
lines while the segmented solid line in (b) is the regression line.
The values of OAC were found to be relatively higher than OBC (Table 3). They vary from 27.5 to 105.1 ppb with mean and standard deviation of about 58.2 ppb and 18.7 ppb respectively. They were found to be close-to-normal distributed (Figure 4(a)).

Applying EBK on OAC produced an optimized model with RMSE of about 12.9997 ppb. The model had residuals that were normally distributed (Figure 4(b)) and showed no clear signs of heteroscedasticity or non-independence (Figure 4(c)) with ME close to zero (about –0.0140 ppb). The model also did not show clear signs of overestimation or underestimation as indicated by having root mean square standardized error (RMSSE) close to 1 (about 0.9687). The interpolated continuous surface for OAC along with its associated interpolated continuous surface for prediction standard errors are shown in Figure 5. It is clear that there is a general trend of increase of tropospheric O₃ concentrations in the study area towards the west from 42.5 to 90.6 ppb (Figure 5(a)). This trend of increase in tropospheric O₃ concentrations is associated with another general trend of increase of the prediction standard errors towards the west from 3.7 to 11.4 ppb (Figure 5(b)). That is, the higher the predicted tropospheric O₃ concentrations in the study area, the higher the standard errors of prediction, and vice versa. It is also clear that the accuracy of prediction increases in the study area across the areas surrounding the tropospheric O₃ sampling locations.

\[
OAC = \begin{cases} 
2.8585 \times (180.4633 - Time) + OBC & \text{if } Time < 165.6076 \\
0.5328 \times (180.4633 - Time) + OBC & \text{if } 165.6076 \leq Time \leq 179.4403 \\
-0.8784 \times (180.4633 - Time) + OBC & \text{if } Time > 179.4403 
\end{cases}
\]

Table 3. Summary statistics for OAC in the study area and in different land cover classes.

<table>
<thead>
<tr>
<th>Land cover class</th>
<th>Number of observations</th>
<th>Mean (ppb)</th>
<th>Standard deviation (ppb)</th>
<th>Coefficient of variation (%)</th>
<th>Minimum (ppb)</th>
<th>Maximum (ppb)</th>
<th>Range (ppb)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Urban land</td>
<td>23</td>
<td>66.0304</td>
<td>21.7915</td>
<td>33.0022</td>
<td>29.11</td>
<td>105.05</td>
<td>75.94</td>
</tr>
<tr>
<td>Agricultural land</td>
<td>36</td>
<td>73.0067</td>
<td>17.6733</td>
<td>24.2078</td>
<td>37.44</td>
<td>104.01</td>
<td>66.57</td>
</tr>
<tr>
<td>Rangeland</td>
<td>84</td>
<td>51.2120</td>
<td>14.4019</td>
<td>28.1221</td>
<td>27.45</td>
<td>91.88</td>
<td>64.43</td>
</tr>
<tr>
<td>Barren land</td>
<td>15</td>
<td>49.9460</td>
<td>11.1996</td>
<td>22.4233</td>
<td>32.89</td>
<td>77.44</td>
<td>44.55</td>
</tr>
<tr>
<td>Whole North Jordan</td>
<td>158</td>
<td>58.2148</td>
<td>18.6788</td>
<td>32.0861</td>
<td>27.45</td>
<td>105.05</td>
<td>77.6</td>
</tr>
</tbody>
</table>

Figure 4. Diagnostic plots for the model developed for OAC by applying Empirical Bayesian Kriging. The solid lines in (a) and (b) are the 1:1 lines.
3.4. The relationship between tropospheric ozone and elevation in the study area

Simple linear regression produced relatively weak but significant model at the 0.05 significance level with an $R^2$ of about 0.1431 and RMSE of about 17.3462 ppb (Equation (5)).

$$OAC = 73.3692 - 0.0240 \times \text{Elevation}$$ (5)

All the regression coefficients (i.e. the regression coefficient for elevation in addition to the intercept) were significant at the 0.05 significance level. The residuals associated with the model were close to normal (Figure 6(a)) and had a mean close to zero (about $8.32 \times 10^{-15}$ ppb). However, the residuals showed weak trend of decrease (Figure 6(b)), suggesting slight inadequacy in the model or the assumptions.

The diagnostics associated with the SLR model indicated that the model has room for improvement. Hence, applying PR showed that the relationship between OAC and elevation is not as simple as determined by SLR and also not linear. After fitting five polynomial models with increasing order from the second to the sixth order and testing the significance of each model, the polynomial model of fifth order shown in Equation (6) produced relatively the best results.

$$OAC = 59.8140 + (0.0751 \times \text{Elevation}) + (0.0003 \times \text{Elevation}^2) - (2 \times 10^{-6} \times \text{Elevation}^3)$$
$$+ (2 \times 10^{-9} \times \text{Elevation}^4) - (9 \times 10^{-13} \times \text{Elevation}^5)$$ (6)
The model was found to be significant at the 0.05 significance level. There was an increase in $R^2$ to about 0.2299 and a decrease in RMSE to about 16.6594 ppb. All the regression coefficients (i.e. the regression coefficients for Elevation raised to powers from one to five in addition to the intercept) were significant at the 0.05 significance level. The residuals associated with the model were close to normal (Figure 6(c)) and had a mean close to zero (about $2.81 \times 10^{-14}$ ppb). Furthermore, the weak trend noticed in the residuals resulted from the SLR (Figure 6(b)) was weakened more (Figure 6(d)) indicating better performance of the PR relative to the SLR.

Again, the diagnostics associated with the PR model indicated that the model still has room for improvement. The relatively low $R^2$ associated with SLR (about 0.1431) and PR (about 0.2299) and relatively high RMSE associated with SLR (about 17.3462) and PR (about 16.6594) could be overcome by assuming spatial non-stationarity by applying a spatial regression procedure. Consequently, GWR was applied. Geographically weighted regression produced an acceptable significant model at the 0.05 significance level. The global diagnostics of the GWR model showed significant enhancements if compared to the SLR or PR models. The GWR model had a higher $R^2$ of about 0.5710 and lower RMSE of about 12.1957 ppb. The residuals associated with the model were normally distributed (Figure 6(e)) and had a mean close to zero (about $-0.1373$ ppb). The weak trends noticed in the residuals from the SLR (Figure 6(b)) and PR (Figure 6(d)) disappeared (Figure 6(f)), indicating better performance of the GWR relative to SLR or PR. Furthermore, GWR produced local diagnostics that explored in further detail the relationship between OAC and Elevation in the study area (Figure 7). The Intercept coefficient (Figure 7(a)) increases from negative values in the central parts of the study area to positive values in the eastern, northern, and western parts. The Elevation coefficient (Figure 7(c)) also increases from negative values approximately in the central parts of the study area to positive values in the eastern and western parts. This indicates that in some areas the relationship between OAC and Elevation is negative and in other

Figure 6. Diagnostic plots for the models developed for exploring the relationship between OAC and elevation, which resulted from applying simple linear regression (a) and (b), polynomial regression (c) and (d), and geographically weighted regression (e) and (f). The solid lines in (a), (c), and (e) are the 1:1 lines.
areas the relationship is positive (ceteris paribus). The standard errors of the regression coefficients increase from the eastern and western parts of the study area to the central parts (Figures 7(b) and (d)). That is, small values of the regression coefficients are associated with large standard errors while large values of the regression coefficients are associated with small standard errors. The spatial distribution of Local $R^2$ (Figure 7(e)) shows somehow random pattern and increases from very low values (about zero) to moderate values of about 0.2057.

### 3.5. The relationship between tropospheric ozone and land cover in the study area

Table 3 shows summary statistics for OAC in the study area in different land cover types. The maximum tropospheric O$_3$ concentration of about 105.1 ppb was observed in urban areas, while the minimum concentration of about 27.5 ppb was recorded in rangelands. Urban areas also were associated with the highest tropospheric O$_3$ variability of about 75.9 ppb, while the lowest variability of about 44.6 ppb was associated with barren areas.

Running one-way ANOVA (Table 4) showed that there is a statistically significant difference at the 0.05 significance level between the means of OAC from one land cover type to another. Running means comparisons for all pairs using Tukey–Kramer HSD test (Table 5) showed that high tropospheric O$_3$ concentrations were found in agricultural lands, followed by urban lands and then rangelands and barren lands. The test also showed that the four levels of land cover could be grouped into two distinct homogenous groups of means within which
there are no statistically significant differences at the 0.05 significance level. The first group comprises agricultural and urban areas with relatively high tropospheric O₃ concentrations, while the second group comprises rangelands and barren areas with relatively low tropospheric O₃ concentrations.

4. Discussion, conclusions, and recommendations

Usually, studies related to exploring the spatial and temporal characteristics of tropospheric O₃ and its determinant environmental and social and economic factors (e.g. Metcalfe et al. 2002; Camalier et al. 2007; Lin et al. 2007; Kalabokas et al. 2008; Stedman and Kent 2008) make use of data collected from relatively expensive networks of tens of stationary tropospheric O₃ monitoring stations that are installed and maintained at certain heights close to the surface of the earth. These monitoring stations provide synchronized and continuous measurements that cover the days and nights and seasons of the year, which could be aggregated on hourly or daily or weekly or monthly basis. The results of such studies could be used by decision- and policy-makers to identify priority areas and design better plans for integrated tropospheric O₃ pollution control. However, no such networks for monitoring tropospheric O₃ are installed and actively maintained in Jordan. This might partly explain the poor tropospheric O₃ research and emission inventories in the country. From this came the logic behind the present study. Therefore, a relatively inexpensive single handheld mobile tropospheric O₃ monitor was used to collect tropospheric O₃ observations at different locations in the northern parts of Jordan following a stratified random sampling strategy. In order to synchronize the observations, they were corrected for temporal variations using PLR. However, every tropospheric O₃ observation represents the arithmetic mean of five consecutive readings recorded at specific location during a specified unit of time of not more than five minutes. Consequently, no hourly based aggregates can be obtained, which hinders the effectiveness of the results for exploring the areas of violations of exceedance as compared to the international standards. However, the observations could be used to explore the spatial characteristics of tropospheric O₃ levels and its determinant factors.

Table 4. One-way ANOVA table for OAC by land cover.

<table>
<thead>
<tr>
<th>Source</th>
<th>Sum of squares</th>
<th>Degrees of freedom</th>
<th>Mean square</th>
<th>F-ratio</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between groups</td>
<td>14426.6</td>
<td>3</td>
<td>4808.86</td>
<td>18.3532</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Within groups</td>
<td>40350.6</td>
<td>154</td>
<td>262.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>54777.2</td>
<td>157</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5. Connecting letters report resulted from applying means comparisons for all pairs using Tukey–Kramer HSD test.

<table>
<thead>
<tr>
<th>Land cover class</th>
<th>Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agricultural land</td>
<td>A</td>
</tr>
<tr>
<td>Urban land</td>
<td>A</td>
</tr>
<tr>
<td>Rangeland</td>
<td>B</td>
</tr>
<tr>
<td>Barren land</td>
<td>B</td>
</tr>
</tbody>
</table>

Notes: Confidence quantile: $q^* = 2.5973$ and $\alpha = 0.05$. Levels not connected by same letter are significantly different.
Hence, as obtained from EBK, the relatively high levels of tropospheric O\(_3\) concentrations, which were observed in the western parts of the study area (Figure 5(a)), as compared to the relatively low levels of tropospheric O\(_3\) concentrations which were observed in the eastern parts of the study area (Figure 5(a)), could be partly explained by the anthropogenic agricultural and urbanization activities that characterize the western parts of the study area (Figure 2(b)). These activities often generate tropospheric O\(_3\) precursor pollutants that react in the troposphere in the presence of sunlight and lead to the formation of tropospheric O\(_3\). This could be exacerbated by winds carrying these precursor pollutants from different sources (inside or outside Jordan) hundreds of kilometres, causing tropospheric O\(_3\) to occur also in the less populated regions as the case in the rangelands and barren lands in the eastern parts of the study area (Figure 2(b)) where agricultural and urbanization activities are minimum. In fact, this was one of the main results obtained from the present study; that is, one-way ANOVA followed by comparison between means using Tukey–Kramer HSD. test showed that high tropospheric O\(_3\) concentrations could be expected in agricultural and urban areas and low concentrations could be expected in rangelands and barren lands (ceteris paribus).

However, the role that elevation plays for explaining the spatial variability of tropospheric O\(_3\) concentrations within the study area is complex and not clear and not amenable to direct explanation as the case for land cover. Earlier studies (e.g. Zaveri et al. 1995; Bronnimann et al. 2000; Chevalier et al. 2007) found that tropospheric O\(_3\) concentration increases with altitude in the first kilometre of the troposphere. This could partly be explained by the fact that tropospheric O\(_3\) concentration in any given area is affected by a combination of formation, transport, destruction, and deposition factors, where tropospheric O\(_3\) near the surface is eroded by deposition and titration; processes common in the boundary layer of the troposphere (e.g. RS 2008). This, however, does not conform to what has been obtained from the present study by applying SLR, PL, or GWR. That is, the relationship between tropospheric O\(_3\) concentration and elevation could be complex where in some areas it is positive and in others it is negative (ceteris paribus). Actually, the formation of tropospheric O\(_3\) is a complicated process. It needs precursor pollutants, such as NOx and VOCs, and also other critical weather conditions, such as high temperature, sunshine, low wind, and high humidity. This might partly explains the relatively low to moderate values of \(R^2\) obtained from applying SLR, PL, or GWR indicating that elevation alone might not be enough for exploring the spatial variability of OAC. Moreover, there might be some philosophical difference between the terms ‘altitude’ and ‘elevation’. The term ‘altitude’, which was used in the previous studies mentioned above, means a distance measurement in the vertical direction from the mean sea level of a point in the atmosphere. However, the term ‘elevation’, which was used in the present study, means a distance measurement in the vertical direction from the mean sea level of a point on the surface of the earth.

Therefore, still there is a serious need for further studies aimed at exploring the spatial and temporal behaviour of tropospheric O\(_3\) with more details and pinpointing the different sources and risk factors of tropospheric O\(_3\) in north Jordan.

**Acknowledgements**

This material is based upon work supported by the Deanship of Scientific Research of the Hashemite University in the Hashemite Kingdom of Jordan. The grant title is ‘Studying the spatial distribution of tropospheric ozone over North Jordan using remote sensing and geographic information systems.’
Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Salahuddin M. Jaber http://orcid.org/0000-0002-8510-2203

References


