Remotely Sensed Population Displacement and Return

Behnam Nikparvar[1,2*], Jean-Claude Thill[3,4], Jacob Aronson[5], Jim Walsh[6]

[1]The William States Lee College of Engineering, University of North Carolina at Charlotte, Charlotte, NC, USA
[2]Department of Biology, Center for Infectious Disease Dynamics, Penn State University, University Park, PA, United States (e-mail: bzn5190@psu.edu)
[3]Department of Geography and Earth Sciences, University of North Carolina at Charlotte, Charlotte, NC, USA
[4]School of Data Science, University of North Carolina at Charlotte, Charlotte, NC, USA
[5]Center for International Development and Conflict Management, University of Maryland, College Park, MD, USA
[6]Department of Political Science and Public Administration, University of North Carolina at Charlotte, Charlotte, NC, USA

Reliable information about large population displacement and return resulting from wars and armed conflicts is fundamental to plan and deliver humanitarian aids and post-war recovery services. However, field data collection is usually not possible due to the potential risks. In this study, we set out to assess the potential of remotely sensed satellite imagery to estimate the population displacement and return before and after a conflict. Although extensive research has been carried out using remote sensing to estimate population growth and economic development, no previous study has investigated the potential of remotely sensed data to estimate population movement in unstable situations. We collected two sets of very high resolution and nighttime light images to estimate population change during and after an armed conflict in Mosul, Iraq, between 2014 and 2018. Using Xception deep convolutional neural network and very high-resolution satellite images, we classified the study area into four land use categories of high-residential, mixed-residential, non-residential, and background. The area of these land use classes and nighttime light emission were used as independent variables of a Geographically Weighted Regression (GWR) to estimate the population. The GWR model estimates a population with an R-squared 98.9%, a mean absolute error of 102.88. To evaluate our results, we used displacement and return data from the displacement tracking matrix (DTM) survey in Mosul between 2017 and 2018. The spatial cross-correlation of the actual and predicted patterns is 0.404 for returnee populations in 2018. Such results show that a combination of remotely sensed land use and nighttime light can effectively predict the population spatial and temporal change. The remotely sensed estimation of population change has implications for understanding why and how people are displaced and return or what factors are involved in their choice to seek refuge or return after an armed conflict.

Keywords: Armed conflicts, Population displacement, and return, remote sensing, Geographically Weighted Regression

Introduction

 The global population of forcibly displaced individuals due to conflict, violence, and human rights violations reached over 82 million people in 2020 (UNHCR, 2020). Timely and reliable

information on population displacement and movement is crucial for effective planning and delivery of humanitarian assistance and post-conflict recovery services (Cotter, 2019). Currently, addressing displacement is often reactive, relying on the establishment of camps or allowing movement and return to resolve itself. However, developing a low-latency method to detect population movement and destinations could enable proactive intervention and safer management of population displacement. Unfortunately, existing low-latency methods require on-the-ground presence, which is often limited in conflict situations where displacement is most prevalent.

Remote sensing could solve this problem by providing an estimate of current population across geographic areas using high resolution land use and nighttime light products[1]. On the one hand, Longitudinal land use products derived from very high-resolution images distinguish human residential areas from other land cover (e.g., vegetation) and land use (e.g., industrial zones) and recognizes potential changes due to damage or development. On the other hand, nightlight reveals the presence and density of the population.

During armed conflicts, land use and nightlight can work together to capture population dynamics effectively. In certain situations, buildings with damaged top-floors may not be recognized as residential areas in land use products, but people might still inhabit the lower floors (for instance, in Mosul, many returning shopkeeper families resided in the lower floors of buildings with damaged top-floors[2]). In such cases, nightlight data can capture the presence of population. Conversely, land use information can differentiate between residential neighborhoods and non-residential areas (such as industrial zones and infrastructure) that exhibit similar artificial light reflectance. This complementary use of land use and nightlight enhances the understanding of population patterns, particularly in conflict-affected regions.

Solving the above problem requires models that consider the functional relationship between population and remote sensing products (Mennis, 2009). Specifically, models must be able to differentiate residential neighborhoods from other types of land use and land cover as well as above mentioned concerns specific to the times of conflicts. Moreover, previous studies show light emission also depends on affluence and economic structure (Elvidge et al., 1997 a, b; Sutton et al., 2001). Existing solutions are limited by heroic assumptions about above heterogeneities, and by lack of fine grain data that can be used to validate their estimates, especially for some parts of the world like middle east (Briggs et al., 2007, Zeng et al., 2011).

To meet these requirements, our approach involves three main steps. In the first step, we use a deep learning classifier to create high resolution land use products from high-resolution images of Mosul. These products distinguish between residential and non-residential areas. In the second step, we map population and nightlight to these areas by employing a regression-based dasymetric approach (Mennis & Hultgren, 2006). Dasymetric mapping is necessary because a fundamental requirement to define functional relationship between variables is availability of their values for

---

[1] https://apps.dtic.mil/sti/citations/AD1085619
[2] https://unhabitat.org/the-initial-planning-framework-for-the-reconstruction-of-mosul-0

the same geographic units, which is rarely the case for population and remote sensing data (Briggs et al., 2007). We account for spatial non-stationarity sourced from unknown drivers using residual analysis and local spatial autocorrelation. In the final step, a geographically weighted regression estimates the relationship between population and nightlight within each land use category. We calibrate the model for a baseline year and use it to estimate population during and after the conflict. We validate the estimates with data on displaced and returning population in Mosul.

Literature review

Remote sensing, such as land cover and nighttime light (NTL) imagery, have previously been used effectively to map and model fine-scale populations in the absence of demographic and socioeconomic census or surveys (Yao et al., 2017; Zeng et al., 2011).

Previous studies have extensively utilized nighttime light for various applications, including mapping built-up areas and urban development (Bruederle & Hodler, 2018; Xie et al., 2019), estimating gross domestic product (Shi et al., 2014), monitoring electricity consumption (Kiran Chand et al., 2009), tracking disasters and power outages (Miller et al., 2018; Miguel O. Román et al., 2019), mapping and estimating populations (Zeng et al., 2011), and assessing global CO_2 emissions (Ou et al., 2015). However, limited efforts have been made to evaluate the impact of armed conflicts on population changes using nighttime light. In one of the few attempts to estimate displaced population using nighttime light data, (Witmer & O'Loughlin, 2011) examined the potential of DMSP/OLS nighttime light products to analyze the effects of war in the Caucasus region of Russia and Georgia. They found a strong correlation between the decline in nightlight and population displacement, with a regression R^2 value of 0.7 for 31 cities in two regions. In a more recent attempt, (Li & Li, 2014) investigated the relationship between the loss of nighttime light and internally displaced populations in Syria at the provincial level over a two-year period, revealing an R^2 value of 0.52 using DMSP/OLS nighttime light data.

However, these methods lack the ability to differentiate night light surface reflectance between residential areas and other land use types, including infrastructure, utilities, institutional buildings, and industrial structures. This distinction is crucial to eliminate variations in nightlights that are unrelated to population movement. The loss of light in non-residential neighborhoods can mistakenly be interpreted as a loss of population, or vice versa. (Briggs et al., 2007) used CORINE land cover classes to distinguish nightlight in different land use when mapping small-area population across the European Union with R^2 typically in the order of 0.8–0.9. While land use improved the performance of the population modeling using nightlight, there are still uncertainties in the relationship between population and nightlight. Previous research showed strong relationship between light emission and affluence as well as economic development in large-area aggregates (Elvidge et al., 1997a, b; Sutton et al., 2001). Other factors such as urban configuration, industrial zones, transport infrastructure, energy, and lighting policies and lighting technology may also add uncertainties to the relationship between population and light emission in small-area population estimates (Briggs et al., 2007).

To account for imbalance in population distribution, diversity between rural and urban areas, and other socioeconomic variables, (Zeng et al., 2011) divided their study area in China into several homogeneous population zones using k-means clustering based on population and nighttime light and then modeled population using the Briggs approach. While this approach may consider the spatial non-stationarity, it still needs a separate process for clustering and requires an input number of classes for k-means clustering, which is not preferred. Alternatively, we test the residuals of the regression models for spatial autocorrelation and use local clusters of high and low residuals to add spatial proxies as fixed effects to regression models. Also, the authors use the number of lit and unlit pixels and nightlight emission from DMSP/OLS within each land use category as independent variables in a regression analysis with population count as dependent variables. However, this approach ignores the functional relationship between nightlight and land use categories. For a nighttime light pixel with multiple land use categories, it is unclear how the nightlight emission is distributed. Thus, similar to population, one needs some weighting strategy to distribute nightlight among land use categories. To handle this issue, we take land use as an auxiliary layer and conduct regression analysis to map population and nightlight to different land use categories.

Some of the issues mentioned above are come from limitations in nightlight and land use data. DMSP was operational between 1992-2013. The data is limited in terms of spatial and radiometric resolution and suffers from several sources of error, such as saturation, satellite degradation, and lack of onboard satellite calibration. These limitations make it challenging to use such data for time series analysis (Q. Zhang et al., 2013). Available since 2012, the Day/Night Band (DNB) sensors of the Visible Infrared Imaging Radiometer Suite (VIIRS) provide much higher quality nighttime light data, removing many of the issues with DMSP/OLS, including onboard calibration (M O Román et al., 2018).

Fine-resolution land cover products are rarely available, especially in developing countries and even in developed countries, due to lack of public access or when the purpose is longitudinal analysis. However, with the growing availability of VHSR imagery and the recent developments in pattern recognition methods such as deep convolutional neural networks (CNNs), this is less of a restriction (Pesaresi et al., 2013; Zhu et al., 2017).

Land use and land cover (LULC) classification methods using VHSR imagery can be categorized into four classes: pixel, moving window, object, and scene-based classification (C. Zhang et al., 2018). Pixel level classification, which relies only on spectral characteristics, is helpful to classify land cover. However, it is insufficient to classify land uses that are composed of multiple land covers, such as those typical in urban settings (Zhao et al., 2016). Moving window methods such as grey-level co-occurrence matrix (GLCM) create low-level spatial features that incorporate context or texture information into classification (Niemeyer et al., 2014). However, a regular window shape method still lacks semantic information because objects are irregularly shaped in the real world (Herold et al., 2003). Besides, a separate process is needed to extract these contextual feature descriptors.

Object-based image analysis (OBIA), which adds between-object information such as connectivity, contiguity, distances, and direction amongst adjacent objects, is alternatively used to tackle this problem (C. Zhang et al., 2018). While OBIA methods are powerful, they require a step to recognize objects and need extra ancillary data to define between-object information. Also, to recognize abstract classes such as residential or industrial neighborhoods, a separate post-process is required. Finally, scene-based classification adds semantic information to small land-use scenes (patches) by assigning labels according to their content (Chen et al., 2016). Deep convolutional neural networks such as CNNs are used to classify these scenes. CNN and similar deep neural networks can automatically extract high-level spatial features from the image, making them suitable for the classification of abstract land use classes such as low- and high-residential or industrial. We take advantage of a scene-based deep neural network classification to distinguish between residential neighborhoods and other land use categories (industrial, institutional, etc.).

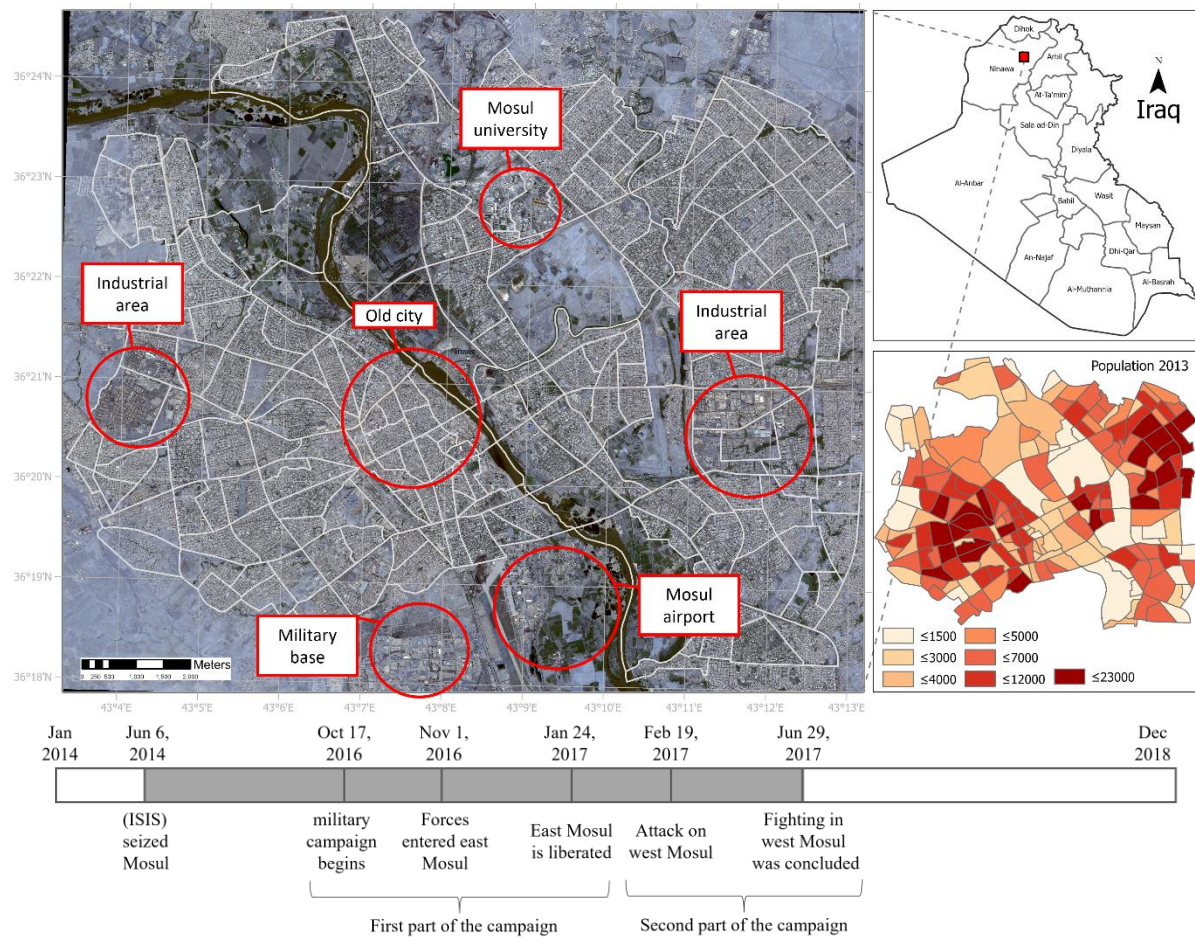Materials and Methods

Study area

Figure 1. The city of Mosul, northern Iraq. The bottom right map shows the distribution of population derived from census population the year before the ISIS captures the city.

We chose a 13km×15km region covering most of the urban area in Mosul in northern Iraq as the study area (figure 1). Mosul is located in Nineveh Governorate, the second most populous first-order administrative district in Iraq (UN Habitat, 2016). The Islamic State of Iraq and Syria (ISIS) seized the city in June 2014. Many of the city's religious and ethnic minorities, such as Kurds, Christians, Turkomans, and Yazidis, immediately displaced, mainly inside the country (UN-Habitat, 2016). However, Rural migrants and IDPs from other parts of Nineveh, Salaheddin and Anbar Governorates, Syrian refugees, and ISIL fighters and their families increased the inflow of population during 2015 and 2016 (UN-Habitat, 2016), especially in the north part of the city (Lafta et al., 2018). Once controlling the city, ISIL enforced strict population control measures. Some reports suggest residents had to provide title to their properties as collateral to temporary leave the city for two weeks (Robinson, et al., 2017). The first part of the military campaign against ISIL began October 2016 focusing mostly on east Mosul. Military forces entered east Mosul on November 1, and declared liberated on January 24, 2017 (Lafta et al., 2018). On February 19, 2017, the campaign started attacks on west Mosul using heavy artillery equipment and air-strikes, which lead to wide damage especially to the old city neighborhoods which was completely obliterated (Lafta et al., 2018). The government of Iraq took control of the city on June 29, 2017, and people began to return from other locations to the city.

We chose 2013 as the reference year to model the urban population spatial distribution using remote sensing imagery to represent the stable state population (baseline population). Then, we used the remotely sensed data of 2014-2018 to study the spatiotemporal change in population during and one year after the end of the conflict. We validate predictions by a longitudinal database of displacement and return for the same period.

Modeling strategy

The method is structured as follows (figure 2.a). We create a longitudinal data set of land use scenes from VHSR images for Mosul. This data set is used to train a machine learning model and divide the urban area into classes of land use that represent human residential settlements (step one). The strategy to define land use classes (figure 2.b) are explained in the next section. The main goal in step 1 is to distinguish residential neighborhoods from non-residential neighborhoods and from other land covers. Then, the land use layer is coupled with separate spatial regression models to map census population and nightlight into neighborhood level (step two or dasymetric modeling step). We use land use because different land use types impose uneven distribution of both nightlight and population across urban environment. The regression models determine the weights associated with each land use category while mapping population and nightlight. This regression step is also important to address uncertainties stem from land use data, which we discuss in the discussion section in detail. The third and final step estimates the population (as the dependent variable) with night light emissions by land use category (as the independent variables) in a desired aggregated level using a GWR model (step three). This model is used as baseline to estimate population for the next years during and post conflict. That is, steps one and two are repeated to create land use and map nightlight for these years, which are then used to estimate population for each year in step three. The estimates are then validated by a data set of displaced and returnee population.
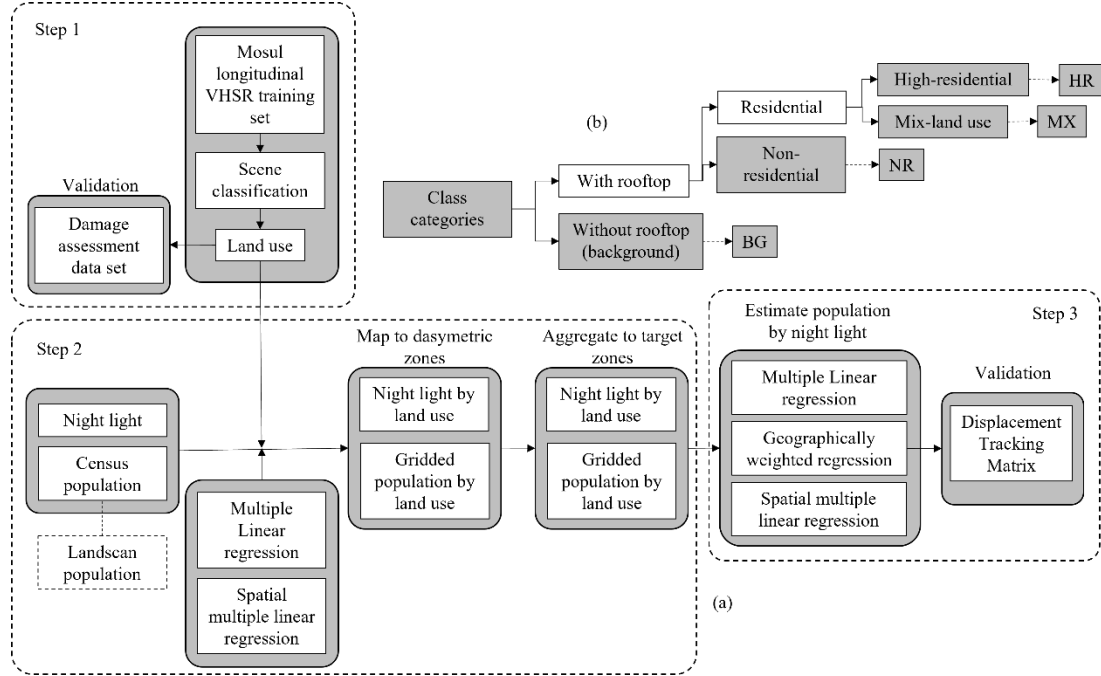
Figure 2. a) Modeling strategy. b) Defining land use classes

Mosul data set

We built the Mosul VHSR scene data set to create land-use products for Mosul, Iraq. The data set consists of more than 20,000 labeled VHSR sample scenes spread over four class categories representing the population density. The scenes represent neighborhoods of 50m×50m in four classes of high-residential (HR), mixed-land use (MX), non-residential (NR), and background (BG). Similar to the census data collection, we target places of residence to define population. People live inside buildings. Thus, initially we aim to distinguish land use that include building rooftops from the land use without building rooftops (figure 2.b). Starting with the latter, all land use without building rooftops is categorized into one single class called background. While it is not idle to merge all these classes into one, it makes the process of creating training data base manageable in terms of manual work.

For the former category, however, we divide it into three separate classes. Not all buildings are places of residence. Typically, people may visit neighborhoods with institutional, industrial, or cultural buildings (or other land cover categories) during the day, but that is not considered their place of residence (except for institutional and non-institutional group quarters, which are not recognized in our study). Thus, we separate land use scenes with residential buildings from those dominated by non-residential buildings (NR). This is where scene-based classification is preferred over pixel-based classification. Pixel-based classification fails to distinguish residential and non-residential pixels with surface reflectance from the same construction material. When integrated with deep learning methods, the scene classification uses neighborhood features to distinguish such categories from each other.

Since a scene can include mixed patterns of residential buildings with other land use and landcover categories, we divide it into high-residential (HR) dominated by residential buildings, and mixed

land-use (MX) where the scene includes residential buildings, but it is dominated by the other land use or land cover types. As a result, these two classes may represent different densities of populations. We assume a scene is HR if it is dominated by residential buildings (more than three rooftops). Scenes with three or fewer residential buildings have a low residential population and are labeled as MX. The functional relationship of HR and MX with population is modeled in the dasymetric mapping step. Figure 3.a shows sample scenes for each class category.

We selected the size of the scenes based on our knowledge of a typical city structure in the Middle East and visual experiments on images of Mosul. A 50m×50m scene of a neighborhood and the number of buildings inside provide information about its land use and likely population density. It may contain many building rooftops, sometimes ten or even more in residential areas.

We created the data set for longitudinal analysis. Five cloud-free VHSR (50 cm spatial resolution) images with appropriate intervals were obtained from multiple sensors based on availability for 2014-2018. The data set includes one GeoEye image (Aug 2014), three Pleiades images (Nov 2015, Aug 2016, Aug 2018), and one World-View2 image (June 2017). We selected the sample scenes in a controlled sampling method to acquire intra-class heterogeneity and inter-class balance. These samples were obtained from the same neighborhoods for all years; that is, each neighborhood has five representative samples collected using multiple sensors at different times (figure 3b). Even if all land covers within a neighborhood remain unchanged over time (i.e., the neighborhood land use label remains the same), conditions such as lighting, satellite angle, sensor characteristics, and atmospheric effects introduce variations in scene imagery. Thus, sample scenes are gathered from images of different years and sensors to make the classification independent of the changes imposed at data collection time.

The labeling process was conducted through visual assessment by a researcher familiar with Middle Eastern city structures. It was then evaluated by the limited land use maps from reports and points of interest data sets. Samples of the same neighborhood were labeled simultaneously in a graphical user interface to minimize human errors. Figure 3c shows the share of each class category in the Mosul data set for a total of 4185 samples per year. Most of the scenes are from the BG and HR categories, while MX and NR categories cover a minor part of the urban area. According to the sample data set, most of the longitudinal changes refer to NR and MX categories.
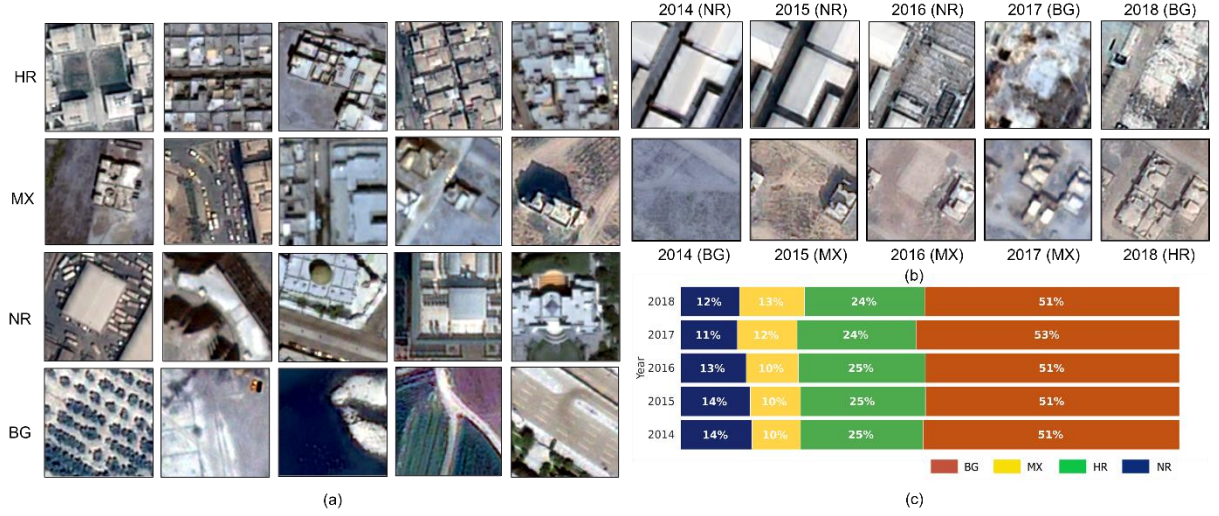
Figure 3. Sample scenes for four classes (a), longitudinal changes for two neighborhoods (b), the data set class balance for different years (c).

Scene classification

We use Xception deep neural network model first introduced by (Chollet, 2017) for this research. Xception has a CNN-based architecture based on the neural network architecture VGG-16 (Simonyan & Zisserman, 2014) and the family of inception networks architectures (Ioffe & Szegedy, 2015; Szegedy et al., 2015, 2016, 2017). The Xception has 36 convolutional layers, which form the feature extraction base of the network. These layers are structured into 14 modules with linear residual connections around them, except for the first and last modules (Chollet, 2017). The main idea behind Xception is to separate the application of filters across the input channels in the convolution process. That is, the convolution occurs in two separate steps: a depthwise convolution followed by a pointwise convolution. In the first step, the filtering is applied on one single channel at a time. Then a one-by-one convolution is used to combine the output channels from the previous step. The depthwise separation of convolution decreases computation time and the number of parameters significantly compared to the standard convolution. We fine tune an Xception network pre-trained on the ImageNet database (more than a million images) by Mosul data set.

Nighttime light

People may still reside in buildings with damaged rooftops during an armed conflict[3]. However, these buildings may not end up in population classes in the process of classification due to damage. Also, land use classes add little information about the current density of population. Thus, we use nighttime light data which can potentially resolve these issues. Artificial lights from damaged buildings are still likely to be captured. The amount of energy recorded by the sensor can be a representative of population density; that is, in a neighborhood with more residential buildings, more radiance is recorded.

---

[3] https://unhabitat.org/the-initial-planning-framework-for-the-reconstruction-of-mosul-0

The Day/Night Band (DNB) sensors of the Visible Infrared Imaging Radiometer Suite (VIIRS) carried by the Suomi-National Polar-orbiting Partnership (S-NPP) and Joint Polar Satellite System (JPSS) platforms provide daily nighttime light emission (M O Román et al., 2018). S-NPP satellites were launched in 2011, and global daily DNB data is available since 2012. Data is accessible from NASA's Black Marble nighttime lights product suite (VNP46) with 500m spatial resolution.[4] Two versions of the product are the daily at-sensor top-of-atmosphere TOA nighttime radiance (VNP46A1) and the daily moonlight and atmosphere corrected NTL (VNP46A2). The quality of the at-sensor TOA nighttime radiance is not suitable for time series analysis due to several sources of error including atmospheric effects, moonlight, and cloud covers. However, these errors are removed in the BRDF corrected product VNP46A2, making them appropriate for this research.

Nevertheless, the VNP46A2 values are not corrected for the anisotropic impact of satellite viewing angle, which may influence the DNB radiance, especially in urban areas (Li et al., 2019). In this research, we create annual median products of nighttime light, so such factors do not significantly influence the annual nighttime light time series analysis. However, especially for daily time series analysis of night light radiance in urban areas, it is necessary to consider the influence of satellite viewing angles.

We collected daily BRDF corrected nightlight values (VNP46A2 product) for the period 2013-2018. Only the uncloudy pixels were selected for each day, and an annual median product was calculated for each year in the period 2013-2018 (figure 4). The nightlight values show a tremendous decline after 2013, especially in 2014, when ISIS seized the city, and 2017, when the international coalition invaded Mosul and evicted ISIS. Values rebounded after the liberation of the city from ISIS control in 2018, although they did not reach levels seen in 2013.
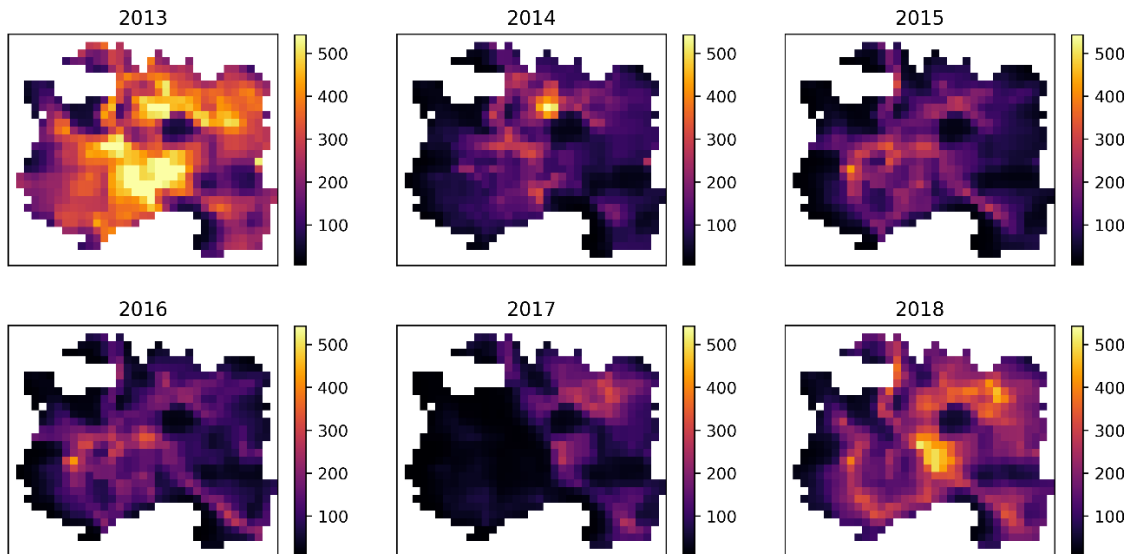

Figure 4. Nightlight emission in Mosul.

[4] https://blackmarble.gsfc.nasa.gov

Population data

Available census population data is very limited for Mosul. The most recent population census conducted by the Iraqi government at the neighborhood level was in 2009 (Iraq National Population Commission,2012). To our knowledge, there is no data available at the neighborhood level for 2013, the stable state for our study. However, estimations of the city's overall population right before the seizure of Mosul by ISIS have been reported by international organizations (UN-Habitat, 2016). Thus, we estimated the most likely Mosul population before the invasion of ISIL in 2013 by leveraging the spatial distribution of census 2009 and longitudinal capacity of LandScan population data, which is known to be the most accurate resource of the world population (Dobson et al., 2000). The population for each census unit i in 2013 was then calculated using the following equation:

$$EstPop_i^{2013} = \left(\frac{P^{2013} - P^{2009}}{LS^{2013} - LS^{2009}} \right) * (LS_i^{2013} - LS_i^{2009}) + P_i^{2009} \qquad (9)$$

Where P^{2009} is the entire population of the study area from census 2009, P^{2013} is the entire population of the study area for 2013 according to reports, and LS^{2009} and LS^{2013} are the entire population from LandScan for 2009 and 2013, respectively. Mapping the population 2009 of Landscan into census level conducted based on the areal weighting method showed positive population value for one zone with zero population count. The number of residential buildings inside the zone was zero confirming the value from the census population. Thus, we adjusted the LandScan 2009 population to match the census population count for this zone. Then we applied equation 9 to estimate the neighborhood level population of Mosul right before the conflict starts. Using the neighborhood level population from census and the temporal power of LandScan, we think this is the most likely population of the study area before the conflict. To avoid confusion with the population's estimations using land use and NTL in the following steps, we refer to the estimated population of 2013 in the table 1 as the census population of 2013 (figure 1 bottom-right). One thing that is noticeable in table 1 is the significant difference between overall population from LandScan and census population of 2009. We further assessed the LandScan data of 2009 for a number of major census zones after mapping. Our assessment showed in many cases the LandScan population count show significant overestimation. For instance, in case of a large zone without any residential building (occupied by amusement park, industrial buildings, and other landcovers) a population of 31,933 were recorded (see supplementary materials).

Table 1. Reference population of study area (2013)

Population source	Zone count	Mean	Std.	Min.	Max.	Overall population
Census (2009)	193	4,660.73	3,681.01	0	19,070	899,521
LandScan (2009)	193	17,284.23	14,049.48	36.858	62,086	3,335,857
LandScan (2013)	193	20,984.08	16,959.53	52.177	75,719	4,049,928
Estimated (2013)	193	5,926.39	4,285.71	0	22,556	1,143,794

Dasymetric mapping

After the land-use product and NTL are ready, we follow the strategy described in figure 2.a to map population and NTL to the dasymetric zones. Dasymetric mapping is not valid unless it meets the following requirements after mapping the population (Briggs et al., 2007). First, all units in the intersection layer (dasymetric layer) must have a non-negative population value. Second, estimations must be non-stationary, meaning change from one pixel to another across time periods. Finally, if the mapped populations are aggregated into the reference census level, the population value must remain unchanged. This last requirement is known as pycnophylactic constraint, which ensures the errors remain intrinsic in the reference census population zones (Tobler, 1979).

We use a variation of the approaches introduced in (Briggs et al., 2007) and (Zeng et al., 2011) to transfer population from census zones and NTL pixels to the dasymetric zones. The following equations represent the regression models for population count P and nighttime light emission E with respect to the area of different land use categories l:

$$P = P_0 + \sum_{l=1}^{L} \beta_l A_l \tag{1}$$

$$E = E_0 + \sum_{l=1}^{L} \alpha_l A_l \tag{2}$$

Where P is the population of census zones, and E is the light emission in the level of nighttime light pixels. P_0 and E_0 are the background population and night light emission, respectively. L is the number of land use categories, A_l is the area associated with land-use class l in a census zone (or night light grid), and β_l and α_l are the regression coefficients. General regression models similar to equations 1 and 2 assume that observations are independent of one another, which is usually not the case for many environmental and socioeconomic variables. Ignoring the violation of independence may lead to inconsistent and biased regression results (Waters, 2012). To account for this possibility, spatial fixed effects are extracted from a spatial relationship matrix derived from residuals analysis with Moran's I local statistics of spatial autocorrelation (Anselin, 1995). Local Moran's I statistics reveal clusters of high and low residuals values, suggesting the relationship among the dependent and independent variables changes across the clusters. The fixed effects are then added to the standard statistical model specification in equations 1 and 2. The coefficients of these models are used to distribute population and NTL to the dasymetric zones. Equations 1 and 2 define the functional relationship between population and land use categories as well as NTL and land use categories in their reference layers (census zones and NTL grid). The following equations show how the population is mapped to the dasymetric zones j using the census population i and weights from equation 1 (Zeng et al., 2011):

$$P_i = P_0 + \sum_{l=1}^{L} \left(\beta_l \sum_{j=1}^{M} A_{jl} \right)$$

$$P_i = P_0 + \sum_{l=1}^{L}\sum_{j=1}^{M} \beta_l A_{jl}$$

$$P_i = P_0 + \sum_{j=1}^{M}\sum_{l=1}^{L} \beta_l A_{jl}$$

$$P_i = \sum_{j=1}^{M}\left(\frac{P_0}{M} + \sum_{l=1}^{L} \beta_l A_{jl}\right)$$

$$P_j = P_0/M + \sum_{l=1}^{L} \beta_l A_{jl} \tag{3}$$

Here P_j is the population in intersection unit j and A_j is the area of dasymetric zone j. M is the number of dasymetric zones inside land use categories l in census unit i. Equation 3 maps population values to the dasymetric zones. However, the dasymetric mapping must have the pycnophylactic (mass preserving) property, meaning the sum of the population must preserve within each census unit i after reaggregation ($\forall\ i: P_i = \sum_{j=1}^{M} P_j$) (Tobler, 1979). This constraint guarantees that the uncertainties remain at the internal level of census zones. Thus, we need to rescale the population within each dasymetric zone using the following equation:

$$P_j{}' = P_j \times \frac{P_i}{\sum_{j=1}^{M} P_j} \tag{4}$$

Where P_i is the population of the census containing P_j, and $P_j{}'$ is the rescaled population of dasymetric zone j. Similar to equations 3 and 4, we distribute night light emission to dasymetric zones. The final equations are as follows:

$$E_j = \frac{E_0}{N} + \sum_{l=1}^{L} \alpha_l A_{jl} \tag{5}$$

$$E_j{}' = E_j \times \frac{E_k}{\sum_{j=1}^{N} E_j} \tag{6}$$

Where E_j is the night light emission at dasymetric zone j. N is the number of dasymetric zones inside land use categories l within each NTL pixel. E_k is the night light emission of the NTL pixel containing E_j, and $E_j{}'$ is the rescaled night light emission of the dasymetric zone j.

Population Estimation

The population and night light emissions of different land-use categories are aggregated into the target level (NTL level). Finally, multiple linear regression with the population as dependent variable and night light emission by land use as independent variables is used to estimate population:

$$P = \gamma_0 + \sum_{l=1}^{L} \gamma_l E_l \qquad (7)$$

Where E_l is the nightlight emission from land-use l; γ_0 and γ_l are the intercept and regression coefficients, respectively. In the presence of spatial non-stationarity, the nature of such models is not fixed over space (Brunsdon et al., 1998). Alternatively, we can use a GWR to estimate population (Fotheringham et al., 2003):

$$P(u) = \gamma_0(u) + \sum_{l=1}^{L} \gamma_l(u)E_l \qquad (8)$$

The notation $\gamma_l(u)$ describes a relationship between population and predictors around location u which is specific to that location. Nearby zones have a higher weight in the calibration than zones that are farther away. The spatial weights can be computed using a desired type of kernel. Since the sample points are regularly spaced, a fixed bandwidth is preferred. We select a bandwidth by minimizing the root mean squared error of the prediction for the generalized geographically weighted regression in a cross-validation process.

Evaluation and validation

The performance of the scene classifier is measured by dividing the Mosul data set into train and test sets and assessing the confusion matrix and kappa value. No data is available to directly assess change in land use. However, we extract changes in pixels that are potentially due to a damage between 2014 and 2017 and compare it with a UN Humanitarian Data Exchange damage assessment, which was conducted on Aug 2017 (after the end of conflict). The data set contains 20,787 sites that are visually assessed for damage in a GeoEye-1 very high-resolution image. In our data set, we define a potential damage as any change in set D = {(HR, BG), (HR, MX), (NR, BG), (MX, BG)}, where the first element of each pair shows the class category of pixel in 2014, and the second element shows the class category in 2017.

We report population estimation performance for the reference year with three measures of root mean square error (RMSE), mean absolute error (MAE), and adjusted R^2. The spatial patterns and total population of returnees for the predictions are then evaluated by a data set of displaced and return from the Displacement Tracking Matrix (DTM) system using spatial cross-correlation analysis. To assess the correlation between predicted and actual maps, we use a modified version of the t-test first proposed by (Clifford et al., 2019) to detect the association between two spatial processes. The test works based on corrections of the sample correlation coefficient between the two spatially correlated processes and requires the estimation of an effective sample size. This factor considers the spatial association of both processes. The test is implemented in R using the SpatialPack package (Cuevas et al., 2020).

Results

Land use

We used an Xception network to classify VHSR images and created land use products. The Mosul data set scenes were resized to meet the input image size requirement of 299×299×3 for the
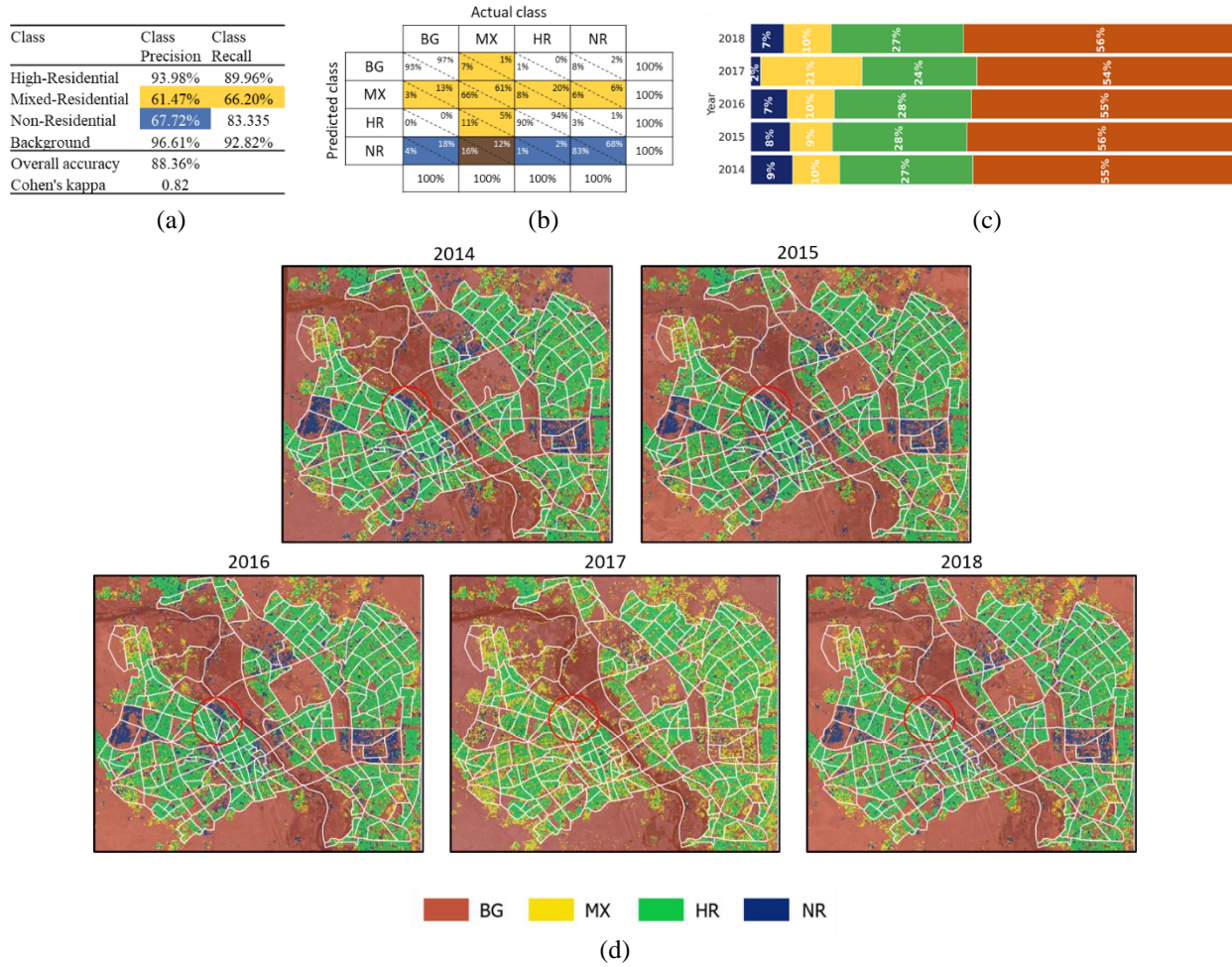
network. Mosul data set were divided into train, validation, and test data sets. 20% of the scenes (equivalent to the sample size for a single year) were left aside for the test. The rest of the scenes were divided between train and validation (80%-20% shares). To leverage high-level features from a pre-trained Xception network on the ImageNet data set, we fine-tuned the Xception network on the training data set by freezing the coefficients of the first eight convolutional layers (entry flow) of the network. We replaced the final fully connected layer (the optional fully-connected layer) and the classification layer to adapt the network to a four-class classification. Finally, the model was trained with an Adam optimizer and an L2 regularizer. It was fine-tuned with an ultimate initial learning rate of 10e-3 and a mini-batch size of 32 for a total of 100 epochs. The classification was implemented in MATLAB and using Tesla V100s GPUs (32GB HBM2 RAM per GPU).

The classification results are available in figure 5a. The overall accuracy equals 88.36%, with a Cohen's Kappa coefficient of 0.82. This accuracy is high compared to most remote sensing classifications, which rarely reach accuracies higher than 80%. The class precision and recall are relatively high for the HR (93.98%, 89.96%) and BG (96.61%, 92.82%) categories making them acceptable for further population modeling. The precision and recall for the MX (67.72%, 66.20%) and NR (61.47%, 83.33%) categories, however, are lower. That means the model performs well for highly populated and unpopulated areas but has trouble distinguishing non-residential and mixed land use neighborhoods.

Further analysis of the classification results is demonstrated in the confusion matrix of figure 5b (percentages are rounded to the nearest whole number). A closer look at the MX prediction errors shows most of the false positives are from the HR class (20%), and similarly, most of the NR class's prediction errors are from the BG class (18%). Although most of the misclassified MX scenes are from NR, 11% of them are still from the HR class. This information indicates that the classifier is still highly capable of distinguishing classes representing population (HR and MX) from classes that are not considered residence places (NR and BG). One may use a binary classification by merging those two sets of class categories. However, we prefer to distinguish MX and HR since these classes have quite distinct levels of population density. Also, when possible, larger number of classes are preferred in remote sensing. Thus, we decided to continue with the current four-class classification. However, the classification can still be improved. Subsequent work could seek to improve classification results by developing logical rules to refine the classification results among different years. Another approach would be to use a combination of convolutional neural networks with recurrent neural networks (especially LSTMs) to further develop a spatiotemporal classification.

Figure 5c illustrates the share of each land use category in the classification of images from different years. Each pixel (scene) represents the 50m×50m neighborhood classified by the trained Xception network. The two blue clusters in the eastern and western side of the city are the two main industrial zones, and the sparse cluster of blue areas contains the academic buildings of the University of Mosul (also see figure 1). We see most of the changes occurred in 2017, which is in line with our understanding of the evolution of the conflict in Mosul. The coalition of Iraqi and international forces used airstrikes, artillery, and ground units to evict ISIS from the city in 2017. This resulted in the destruction or damage of many buildings, especially non-residential structures.

The pattern is evident in Figure 5d and 5e, which show the classification results for the entire study area and for the Old City and its surrounding at the center of Mosul, which experienced the heaviest fighting. The Old City is reported to be obliterated completely during the military attacks (Lafta et al., 2018). While the classifier shows high damage surrounding and to some degree in Old City, it still recognizes many scenes as HR class. This is mainly because damaged buildings are less available in the training data set and they are categorized to the same class (BG or MX) with many other land use and land covers. The main impact of air strikes seems to be to reclassify NR (and to a lesser extent HR) to MX. This is consistent with airstrikes mostly targeting infrastructure. Since we did not have access to the VHSR image of 2013, we used 2014 classified land use for 2013. Our visual assessment with the older images of online database engines such as ArcGIS and Google Earth showed we are not missing too much information (since the conflict starts after 2014, we did not see significant changes in land use). We further assessed the land use classification with a data set of damage assessment from UN Humanitarian Data Exchange (see supplementary materials). It is, however, important to note that the purpose is not to assess damage in this study. Displaced and returnee populations may stay in buildings with damaged rooftops.

Class	Class Precision	Class Recall
High-Residential	93.98%	89.96%
Mixed-Residential	61.47%	66.20%
Non-Residential	67.72%	83.335
Background	96.61%	92.82%
Overall accuracy	88.36%	
Cohen's kappa	0.82	

(a)

(b)

(c)

2014   2015

2016   2017   2018

BG   MX   HR   NR

(d)

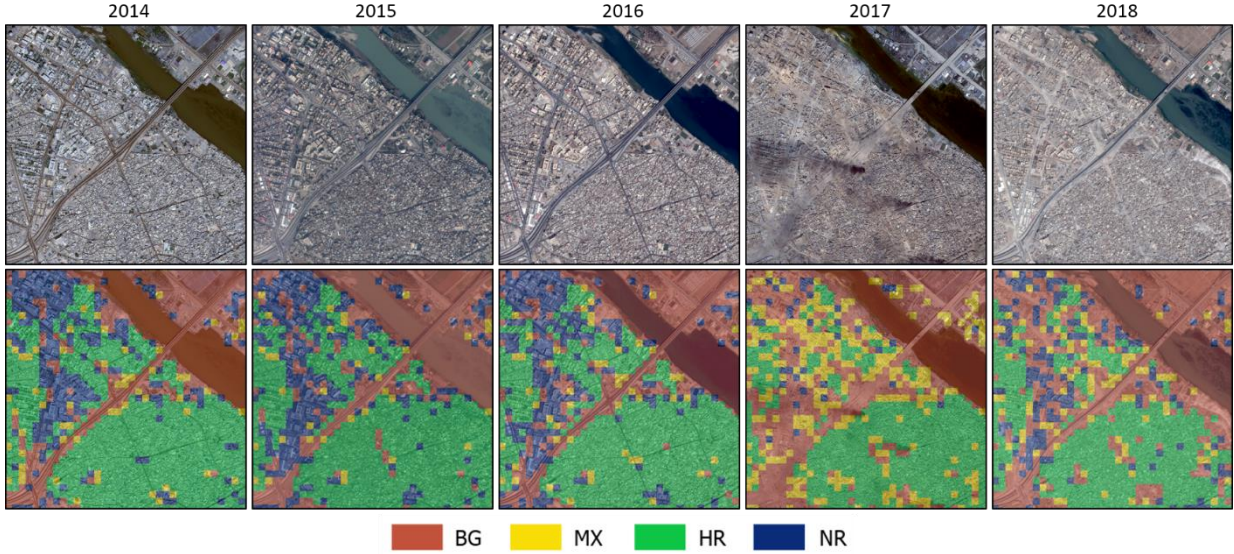| 2014 | 2015 | 2016 | 2017 | 2018 |

BG   MX   HR   NR

(e)
Figure 5. Classification results

Dasymetric modeling of night light emission

We used Equation 2 to map night light emission to the dasymetric zones. Areas of four categories of land use were used as the regression model's independent variables. Keeping all variables in the model will not result in perfect multicollinearity because the area of night light zones varies in border regions. To further check collinearity we used the variance inflation factor (VIF) test to measure the amount of multicollinearity between independent variables. The VIF score remains less than two for all variables in all models, indicating negligible multicollinearity. As the results in Table 2 show, while all signs and magnitudes are meaningful, conventional multiple linear regression performed poorly for all years of analysis based on the adjusted R^2 values. However, the residuals analysis using Global Moran's I index shows a very high positive spatial autocorrelation. To identify the local clusters of high and low values in residuals, we conducted an Anselin Local Moran's I statistic, which identifies statistically significant hot and cold spots. To account for spatial non-stationarity, we added clusters of High-High, High-Low, Low-High, Low-Low, and zones with spatial autocorrelation as fixed effects in equation 2 (figure 6a).

Table 2. Dasymetric mapping of nighttime light using ordinary multiple linear regression

Land use class	2013	2014	2015	2016	2017	2018
Intercept	-30.11	-16.19	-22.32	-17.20	-8.61	-23.70
High-residential Area (HR)	0.0023***	0.0007***	0.0008***	0.0008***	0.0008***	0.0014***
	(0.000)	(6.14e-05)	(5.66e-05)	(5.21e-05)	(5.36e-05)	(7.56e-05)
Mixed-residential Area (MX)	0.0003	-3.328e-05		0.0001	0.0003***	0.0002
	(0.000)	(0.000)		(0.000)	(8.51e-05)	(0.000)
Non-residential Area (NR)	0.0024***	0.0011***	0.0011***	0.0010***	0.0008**	0.0015***
	(0.000)	(0.000)	(0.000)	(9.58e-05)	(0.000)	(0.000)
Background Area (BG)	0.0013***	0.0007***	0.0006***	0.0006***	0.0002***	0.0009***
	(9.19e-05)	(5.41e-05)	(5.16e-05)	(4.73e-05)	(4.38e-05)	(6.72e-05)
Adjusted R^2	0.450	0.25	0.27	0.29	0.28	0.36
Residuals' Global Moran's I	0.71	0.71	0.66	0.64	0.79	0.70

Significance level: ***0.001 **0.01 *0.05

Table 3 summarizes the results for the dasymetric mapping with spatial fixed effects defined based on residual clusters. Overall, the performances increase significantly with the lowest and highest R^2 values of 0.71 and 0.79, respectively. Coefficients are meaningful in terms of signs and magnitudes. All variables are significant ($p < .05$). Most of the light emission is from HR and NR neighborhoods. The amount of light emitted from the BG is greater than MX, which may be due to the existence of different land covers such as road networks and parks that are lit in BG. In the models for 2013, 2017, and 2018, LH, HL, and LH are absent because they are not among the residuals' clusters in ordinary multiple linear regression. The LL variable remains outside of the models to avoid perfect multicollinearity. We see significant decrease in coefficients after 2014 which remains consistent during the period the city is in control of the ISIS. While the coefficients increase in 2018, they do not return to the pre-conflict level.

Table 3. Dasymetric mapping of nighttime light using the spatial regression model

Land use class	2013	2014	2015	2016	2017	2018
Intercept	-125.68	-63.102	-60.08	-60.04	-42.90	-97.23
High-residential Area (HR)	0.002***	0.0006***	0.0006***	0.0006***	0.0006***	0.0011***
	(6.63e-05)	(3.85e-05)	(3.82e-05)	(3.44e-05)	(2.99e-05)	(4.73e-05)
Mixed-residential Area (MX)	0.0008***	0.0003**	0.0003**	0.0003**	0.0004***	0.0007***
	(0.000)	(0.000)	(0.000)	(9.56e-05)	(4.78e-05)	(0.000)
Non-residential Area (NR)	0.0021***	0.001***	0.0008***	0.0008***	0.0004**	0.0013***
	(0.000)	(6.8e-05)	(6.8e-05)	(6.22e-05)	(0.000)	(0.000)
Background Area (BG)	0.0012***	0.0005***	0.0005***	0.0005***	0.0002***	0.0008***
	(5.98e-05)	(3.61e-05)	(3.4e-05)	(3.17e-05)	(2.38e-05)	(4.27e-05)
High-High residuals (HH)	297.64***	153.23***	156.20***	145.64***	135.92***	224.93***
	(8.504)	(4.916)	(4.445)	(4.270)	(3.181)	(6.127)
High-Low residuals (HL)	115.30***	52.93**	51.90***	53.10***		87.55***
	(28.621)	(19.623)	(3.38e-05)	(11.850)		(16.960)
Low-High residuals (LH)		55.32**	50.50***	49.16**	42.19*	
		(16.239)	(13.584)	(15.513)	(20.845)	
Insignificant residuals (NS)	109.15***	50.29***	52.40***	53.78***	42.14***	88.30***
	(7.101)	(4.028)	(3.619)	(3.383)	(2.616)	(4.884)
Low-Low residuals (LL)	-	-	-	-	-	-
Adjusted R^2	0.79	0.71	0.73	0.72	0.79	0.77

Significance level: ***0.001 **0.01 *0.05

Dasymetric modeling of population

Similar to the previous section, we distributed the census population to the intersection level using the land use categories as independent variables (equation 1). The VIF score remains less than 1.6 for all independent variables, indicating little multicollinearity. Table 4 summarizes the results for the ordinary and spatial multiple linear regression. The adjusted R^2 value for the OLS model is 0.59, with a moderate and significant Global Moran's I index indicating spatial autocorrelation in residuals. Adding clusters of the high and low residuals values to the model as spatial fixed effects increase the R^2 value to 0.68 (figure 6b). Among the land use covariates, all coefficients are positive except for the MX. While small in magnitude and insignificant, positive signs for NR and BG may be due to misclassification or complex interaction between land use categories. For example, schools or places of worship are non-residential but are usually located in census units with high population counts. The same is true for some sub-categories of land use in BG, such as

roads or parks. In the final model, NR and BG have been left aside because they are not significant. This is in agreement with the expectation that people rarely live in BG and NR neighborhoods.

The MX category's coefficient is negative, while it was expected to be positive representing neighborhoods with lower residential populations. One possible reason is the low accuracy in MX classification, mainly because it is less discriminative than the other classes (mixed of many land use categories). Splitting that into more classes may improve the class accuracy of sparse residential areas. The MX has also been left aside in the final model because of the negative sign. The implication is that the population is only mapped to the high-residential areas. Given that most MX areas are in the suburbs, we may under-represent population in these areas and over estimate population in high residential neighborhoods. Among the fixed effects, only HH, HL, and NS are kept in the final model.

Table 4. Dasymetric mapping of population for the baseline year (2014) using ordinary and spatial multiple linear regression

Land use class	Ordinary regression	Spatial regression
Intercept	1651.11	-1405.01
High-residential (HR)	0.0237***	0.0246***
	(0.002)	(0.001)
Mixed-residential (MX)	-0.0221***	-0.0223 ***
	(0.004)	(0.004)
Non-residential (NR)	0.0027	0.0027
	(0.002)	(0.002)
Background (BG)	0.0004	0.0005
	(0.000)	(0.000)
High-High residuals (HH)		5367.88***
		(745.258)
High-Low residuals_(HL)		4382.48**
		(1193.332)
Low-High residuals (LH)		2100.38
		(1029.111)
Insignificant residuals (NS)		2796.40***
		(591.413)
Low-Low residuals (LL)		-
Adjusted R^2	0.59	0.68
Residuals' Global Moran's I	0.36	

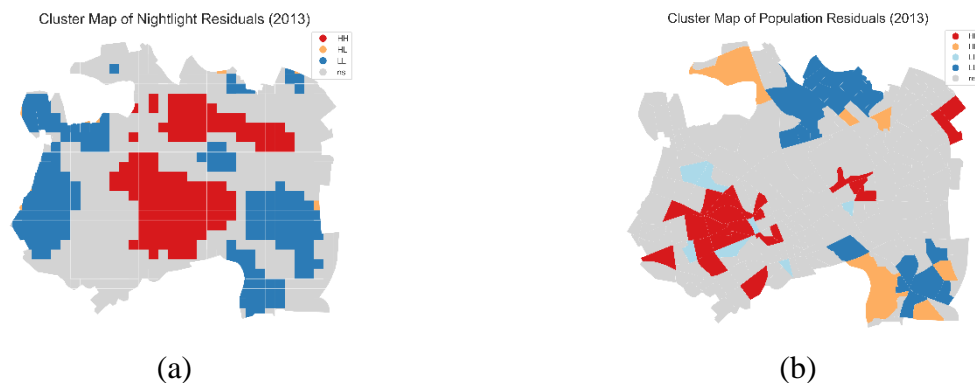Significance level: ***0.001 **0.01 *0.05

Figure 6. Local cluster of residuals for nightlight (a) and population (b)

Population estimation

Model calibration (baseline year)

After both population and nightlight are aggregated by land use categories in the nighttime light level, the population is estimated using equations 7 and 8. Based on the results reported in the previous section, the population is only distributed in HR neighborhoods. That means the only independent variable that remains in model 7 is nighttime light in HR neighborhoods. Thus, we present the results for four models shown in table 5. MLR-E is the multiple linear regression with night light emission by land use as independent variables based on equation 7. The residuals of this model are highly clustered with a significant Global Moran's I index of 0.56. SMLR-E is a model with the same independent variable as in MLR-E but with spatial fixed effects from high and low MLR-E residuals clusters. The model performance improves significantly with an R-squared value of 0.87.

Alternatively, we can use geographically weighted regression based on equation 8 with two different specifications. One is with the area of HR as the independent variable (GWR-A) and the other with the nighttime light value of HR neighborhoods (GWR-E). Fixed bandwidth is considered due to the regular distribution of the samples and calculated based on cross-validations. Both GWR models show significant improvement in comparison to the other models. GWR model can only be used after determining the MLR model's specification; that is why it is used only in the final population estimation stage. GWR-A models the remaining spatial non-stationarity from mapping population to the dasymetric zones layer. Figure 7 demonstrates the four model predictions. We use the GWR models to estimate the population for upcoming years. The GWR-A and GWR-E have negative values for some neighborhoods. However, the sum of negative values for the entire study area remains 0.043% and 0.012% of the total population for GWR-A and GWR-E, respectively.

Table 5. Population estimation models

| Variable	Model			
	MLR-E	SMLR-E	GWR-A	GWR-E
Intercept	192.61***	-669.13***	-11.37***[2]	192.61***[2]
	(44.61)	(61.38)		
Nigh light emission High-residential (NTL_{HR})	13.86***	13.37***	0.0257***[2]	13.86***[2]
	(0.32)	(0.233)		
High-High residuals (HH)		2198.96***		
		(74.647)		
High-Low residuals_(HL)		1206.46***		
		(221.330)		
Low-High residuals (LH)		738.70***		
		(167.061)		
Insignificant residuals (NS)		818.68***		
		(61.314)		
Low-Low residuals (LL)		-		
Bandwidth[1]			316.52m	240.65m
Adjusted R_squared (Quasi R_squared[1])	0.71	0.87	0.979	0.989
MAE	602.64	773.92	149.17	102.88
RMSE	895.97	902.76	239.89	172.54
Residuals' Global Moran's I	0.56			
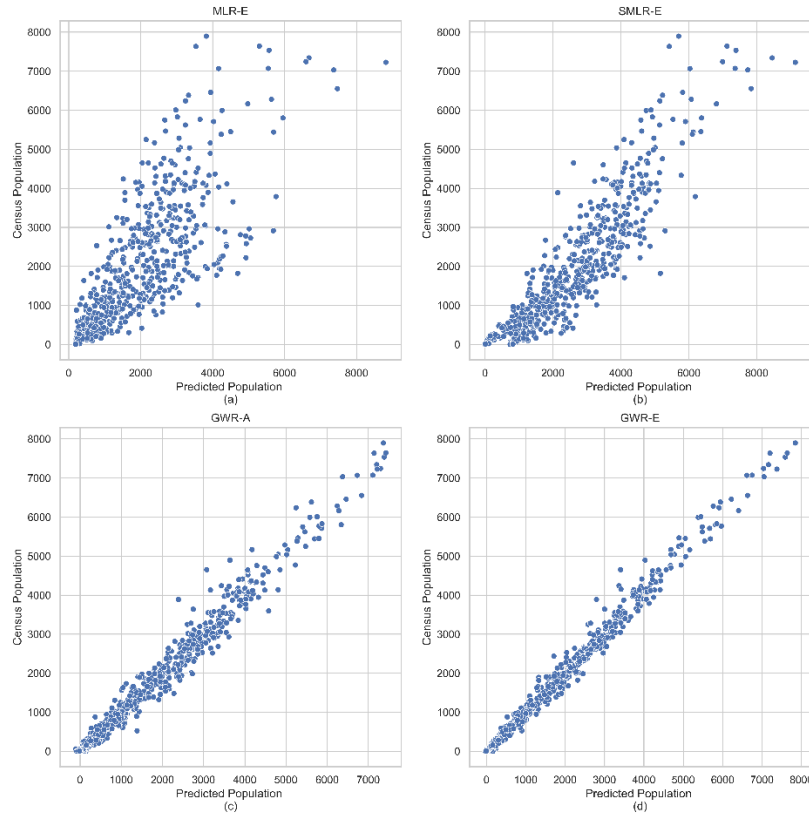
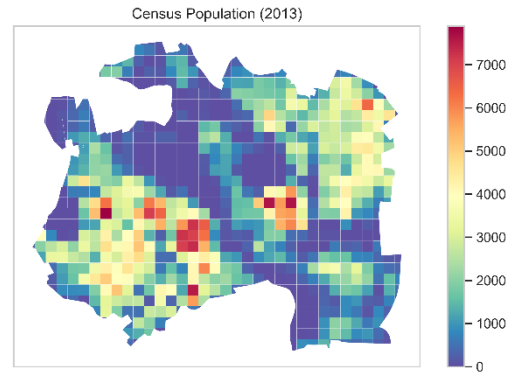Figure 7. Population prediction comparison between models.

Population estimation during and after the conflict

We use the spatial models in section X (mapping ntl) and Y (mapping population) together with the two GWR models' coefficients in section Z to estimate the population for the next five years of the study. Figure 8a shows the estimated total populations for the reference year (2013) and these next five years based on GWR models. The GWR-A model shows little change in the overall population. The GWR-E, however, shows a significant decline in population in 2014 and then a second time in 2017, following a pattern in nighttime light (see figure 5). GWR-E estimates are more in line with the reported displacements, so this observation shows the importance of night light to estimate population. Figures 8b and 8c represent the census population of 2013 and the estimated population of 2013-2018 using the GWR-E model, respectively. The population map of the estimation for 2013 shows the GWR-E model can estimate the population's spatial pattern. Figure 8d shows the change in population compared to the previous year, providing a better sense of population changes within each year. Many of the city's religious and ethnic minorities

immediately displaced after ISIS attack in June 2014 (UN-Habitat, 2016). At the same time some literature reported inflows to the north of the city from other towns (Lafta et al., 2018). Rural migrants and IDPs from other parts of Nineveh, Salaheddin and Anbar Governorates, Syrian refugees, and ISIL fighters and their families increased the inflow of population during 2015 (UN-Habitat, 2016). Historically, most of the Sunni population of Mosul lived in western part of the city where we see the increase in population in 2015 (see also ethnic group maps in the supplementary materials). Once controlling the city, ISIL enforced strict population control measures. Some reports suggest residents had to provide title to their properties as collateral to temporary leave the city for two weeks (Robinson, et al., 2017). This may explain lack of significant change in population in 2016. The first part of the military campaign against ISIL began October 2016 focusing mostly on east Mosul. Military forces entered east Mosul on November 1, and declared it liberated on January 24, 2017 (Lafta et al., 2018). The 2017 population estimation shows a decrease in west and increase in eastern part of the city. On February 19, 2017, the campaign started attacks on west Mosul using heavy artillery equipment and air-strikes, which lead to wide damage especially to the old city neighborhoods which was completely obliterated (Lafta et al., 2018). The government of Iraq took control of the city on June 29, 2017, and people began to return from other locations to the city. To further analyze the pattern in 2017, we estimated monthly population change for March and June 2017 (see the bottom two maps in Figure 8.d). The large reduction in the population of the west Mosul may be associated with start of the attacks. The large increase in east Mosul population may be due to displacement inside the city (e.g., from west side to east to avoid military operations), or inflows of returnees from outside of the city after liberation of East Mosul. We later show the first possibility is more likely using the displacement and returnee data set. Estimates show high increase in population for both sides of the city in 2018.
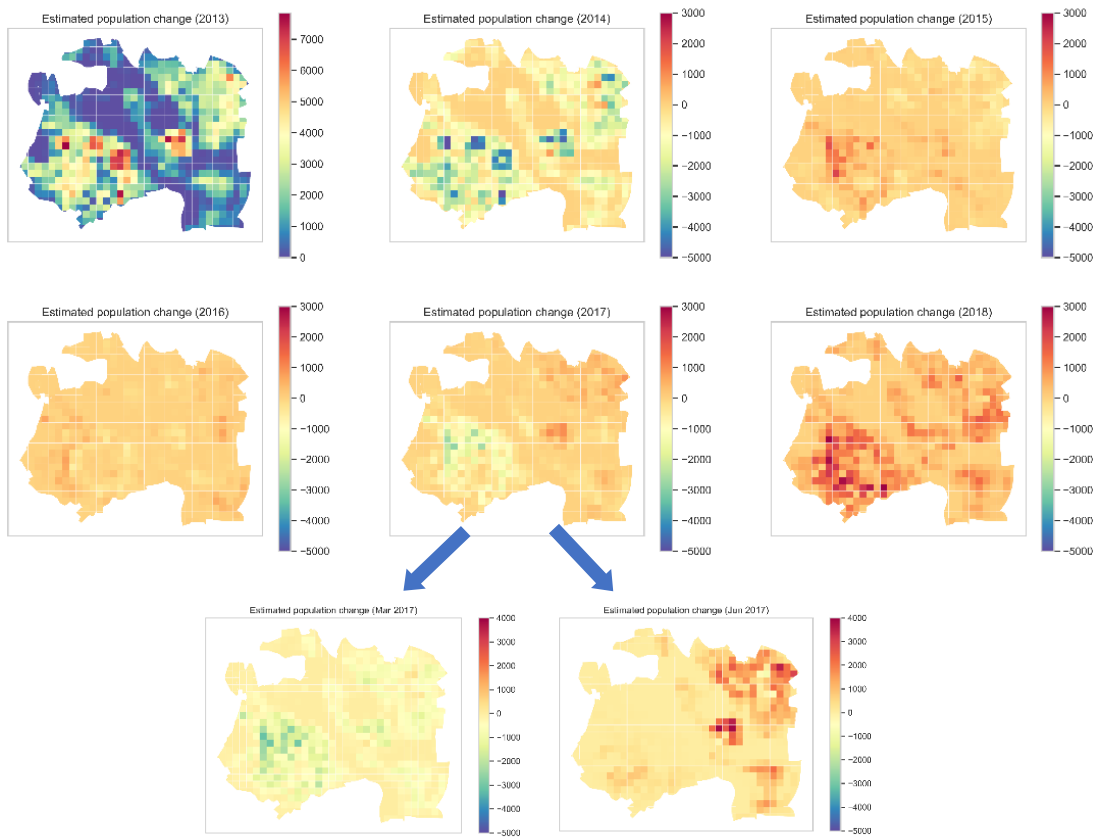

(a)

(b)

22
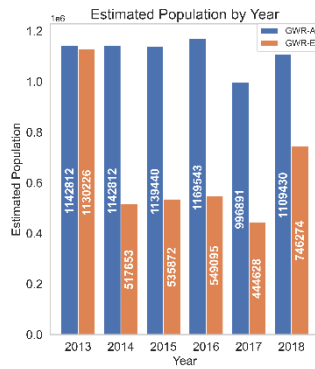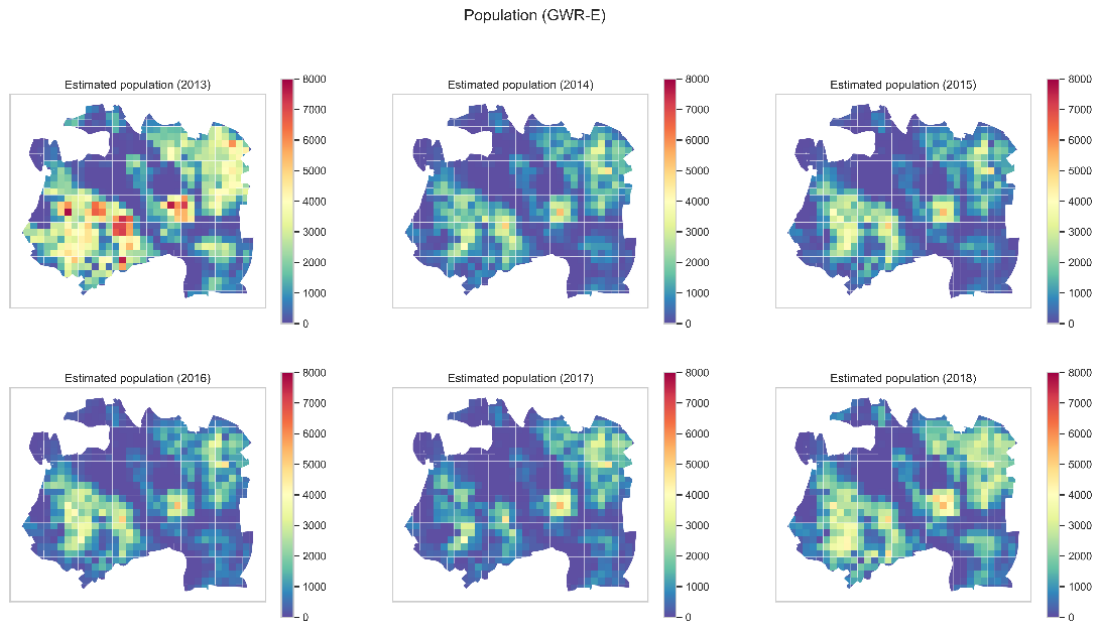
Figure 8. Population prediction results.

Evaluation

Displacement and return data

To evaluate the findings reported above, we compare them to data from the Displacement Tracking Matrix (DTM) that tracks and monitors population mobility. The system regularly and systematically captures, processes, and disseminates information to understand the movements of displaced populations, primarily through surveys[5]. We obtained data of families displaced from or returned to their place of residence in Iraq after December 2013. The data is collected through repeated household surveys. At the start of each data collection cycle, locations are assessed, and a master list of locations is determined. A location is a camp, village, or neighborhood and is updated based on expert assessment. The locations within this list may change over time, but the same neighborhood should theoretically only have one assessment location in a month. Half of the locations are updated every two weeks, which means that every month data on the entire country is available.

A Rapid Action and Response Team (RART) goes to each location in Iraq and engages in direct observation, interviews, and communication with knowledgeable informants, such as local leaders. The RART identifies the number of families displaced to a location and the number of formerly displaced families who have returned to a location as of the time of data collection. The number of displaced families is the new number of displaced families still in a location. The number of returnee families is the new number of returnee families to a location. Both values replace the numbers of the previous month. If a location no longer has displaced families present or all families have returned, it is zeroed out and removed from the master location list. Each assessment is automatically given a credibility score used to determine the reliability of these estimates of population movement.

We collected the returnee and displacement locations for the study area. For Mosul, the displacement and return data are only available after January 2017, most likely due to the risk of collecting survey data. Since the locations may slightly change during data collection, we used an average latitude longitude for each location ID. While locations represent neighborhoods, they have no pre-defined boundaries. Thus, we used Thiessen polygons of the locations to assign an approximate neighborhood based on the relative distance of survey locations, and intersected these with the nightlight level grid (model target level) to transfer the returnee and displaced family populations to the NTL level grids for comparison.

Figure 9 shows the total number of current displaced and returnee families from the DTM on a monthly basis starting January 2017. The temporal pattern is in harmony with the information from reports. June 2017 is the month that the government took complete control of the city. The bar chart shows that over the next six months until December 2017, the displaced population (notice that there are not many returnees yet) increases and starts to decrease starting January 2018 (when the returnee population starts to increase). This pattern confirms the idea of internal displacements potentially between East and West Mosul and vice versa due to events mentioned in previous section. Our estimations of the population dynamics in 2017 are inline with these observed patterns

---

[5] https://dtm.iom.int/

in DTM data. The decline in returnee population in May and June 2018 may be due to inability to conduct survey for some neighborhoods.
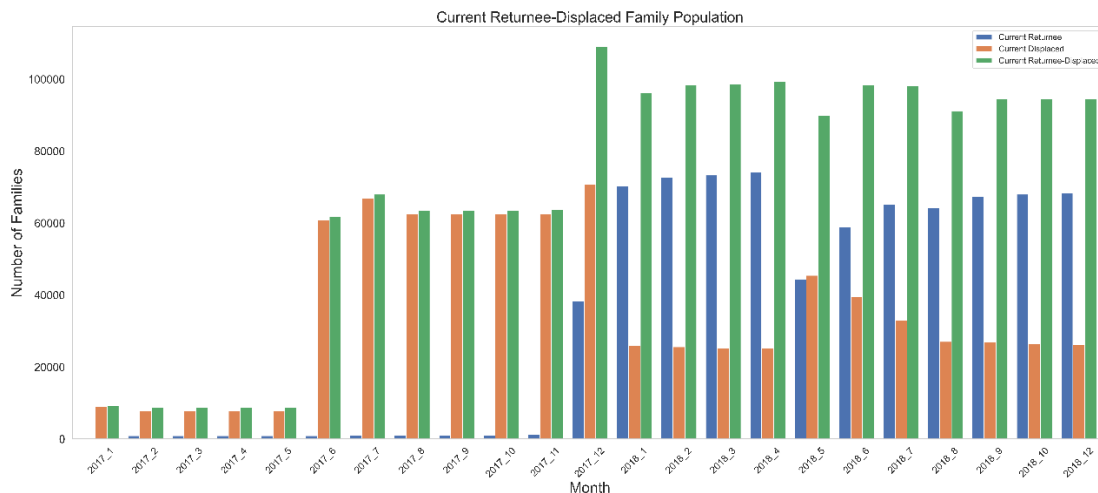

Figure 9. Current returnee-displaced family population

Evaluating prediction results

In many crises that produce large-scale displacement, a spatially and temporally granular field data is unavailable, making it difficult or impossible for researchers to validate their predictions. While validation with DTM data is a unique and important part of our analysis, it is important to highlight elements of the DTM that create some limitations in our comparison. First, displaced families in a neighborhood may be displaced either from outside of the city or from other neighborhoods inside Mosul; in other words, the DTM do not provide information about origin of displaced persons[6]. We also lack data about the population of families displaced from Mosul to locations outside of the city. That means we can neither fully use data of displaced families to evaluate our predictions because the displaced population may be floating internally over time, nor can we evaluate how many people have left the city. Thus, DTM can only confidently represent the number of unique returnee families (and not displaced families) for each month.

Second, it was only after the government took the city's control, and mainly in December 2017, that displaced persons start to return in larger numbers. Thus, 2018 is the only year we can quantitatively evaluate our predictions with the DTM data set. To make this comparison, we calculate the number of returnee families and the number of returnee and displaced families in 2018 with respect to 2017 from DTM. Similarly, we calculate the population change (either displaced or returned) of 2018 from our predictions.

Figure 10 presents a visual assessment of the spatial pattern for the predicted population change (left), real returnee population change (middle), and real returnee and displaced population change (right). Our estimated population change indicates positive population change across the city. The

---

[6] https://www.internal-displacement.org/expert-opinion/for-cities-specify-the-neighbourhood

areas with the largest increase in population in 2018 are in the eastern part of the city, the far southwest, and the eastern bank of the Tigris river. Population increase has been modest in the neighborhoods comprising and surrounding the Old City in western Mosul. This is consistent with available qualitative evidence. Economic activity in eastern Mosul rebounded more quickly after the end of active fighting. The neighborhoods in the far southwest are primarily residential and experienced less wartime damage than the Old City[7]. To compare the predicted and real spatial pattern, we use a spatial cross-correlation method based on the correlation of predicted and real population change. The spatial cross-correlation is based on a variation of t-test to consider the spatial characteristics of predicted returnee population and real returnee population maps. Table 6 shows the sample correlation and p-values for the pair of predicted and real maps of 2018. The correlation is moderate, positive, and statistically significant at a 0.01 confidence level, indicating similarity in predicted and actual spatial patterns.
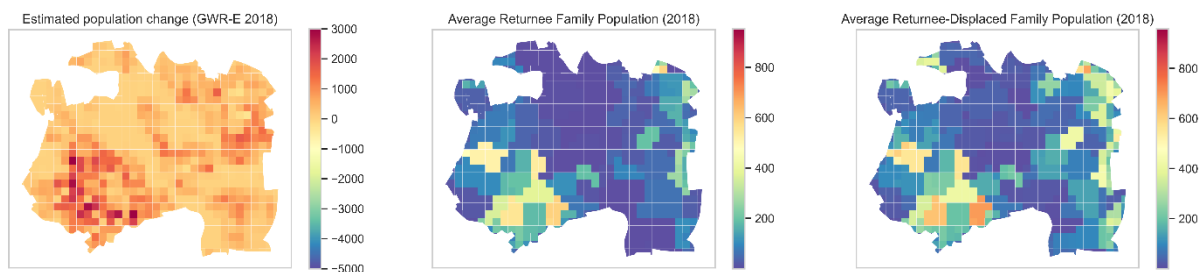

Figure 10. predicted population change (left) and actual returnee (center) and returnee-displaced (right) population change

Table 6. Spatial cross-correlation of predicted and actual maps

	2018	
Population	Sample correlation	P-value
Returnee (average)	0.404	0.0026
Returnee & Displaced (average)	0.390	0.0011

According to the UNHCR field survey, the average family size in Mosul is 6[8]. Using the number of returnee families and the average family size, we estimate the number of returnees in 2018. The increase in GWR-E population (301,722) is similar to the returnee population increase (372,092) and returnee and displaced population increase (305,918) from the DTM data set for 2018. Both the returnee and returnee and displaced populations from the DTM data are reported due to limitations of the dataset.

Discission

---

[7] https://www.npr.org/sections/parallels/2017/09/18/551482635/as-east-mosul-comes-back-to-life-west-mosul-remains-in-ruins

[8] https://unhabitat.org/the-initial-planning-framework-for-the-reconstruction-of-mosul-0
https://conflictandhealth.biomedcentral.com/articles/10.1186/s13031-018-0167-8

We showed how using land use and nightlight together with spatial models can improve mapping and prediction of population displacement during conflicts. Particularly, we separated the human settlements from other land use and land cover with fine-grained resolution to make sure population is distributed accurately across urban area. Compared to the previous work, we found the statistical functional relationship between population and nightlight with land use in a separate process. Nightlight may be reflected from many sources in different types of land use and land cover. Thus, to predict the population using nightlight we need to distinguish lights from human residential areas from these other sources. Using a regression model to statistically find the relationship between population and nightlight with land use was also important to reflect uncertainties due to inaccuracy in land use products. The nightlight is also depended on differences due to affluence, urban configuration, as well as energy and lighting policies and technology. By using a geographically weighted regression and spatial residuals we attempted to implicitly account for these variations.

An important finding in our results is the significance of nighttime light in estimating population density in the context of conflict-related population movement. Figure 8.a shows how adding nightlight in GWR-E model influences the total population estimation. At the same time, mapping of the GWR-A and comparing it with population estimations of GWR-E shows the spatial distribution of the changes are similar. These results further support the idea that land use and night light collaborate to model distribution and population density. While nightlight and land use cannot provide direct information about origin and destination of displacements, they can provide some insights especially in the context of conflicts. Some of our results indicate potential displacements between East and West Mosul during the conflict.

There are limitations in the method we used in this study. While we consider the heterogeneity in functional relationship between population and nightlight across space, we assume it remains the same over the course of the time. For example, reduction in nightlight during conflicts is not only due to population loss. Other, sources such as interruptions in supply due to damage or control of one side of a conflict may impact amount of light and cause underestimation of population. There are some reports about energy supply issues in Mosul during 2015 (Robinson, et al., 2017). Thus, when available these factors must be included in the regression models.

Lastly, the field data is very limited and available only for 2017 and 2018. Moreover, such data is based on surveys subject to many operational challenges. For example, the surveyed neighborhoods and people may have changed during the period, and the displacement and return values are approximate (based on answers from families). In addition, the boundaries of the neighborhoods are defined using a pure geometric criterion and displaced and returned populations were disaggregated into the target grids using an aerial weighting. The DTM displaced and returnee data do not provide fine-grained origin information about returnee and displaced population, which limits the validation. As a result, the field data is also subject to uncertainties, whilst this is still the best available data.

Our model underrepresents population in low residential (mostly suburban) areas and overestimate the population in high residential areas due to low accuracy in MX class. Mixed-residential scenes are composed of multiple land uses making it less discriminative from other classes. Leaving this

27

variable together with BG and NR outside of the dasymetric model means the population is only distributed in high-residential areas. Finally, it is worth noting that high-rise buildings and satellite viewing zenith and azimuth angles can create uncertainties in urban areas (Li et al., 2019). Such an effect is less likely to influence the night light values in Mosul since buildings are not high and we use annual estimations.

Notwithstanding these limitations, the present study has been one of the first attempts to thoroughly examine estimation of population mobility and displacement using remotely sensed data at the time of violent conflicts. As suggested in (Baal & Ronkainen, 2017), creating a link between the limited field data surveys and broader systems that include data and modeling strategies such as those presented here may help us develop systems that can be extremely helpful to study population displacement and return at the time of conflicts. Models like the one presented here are useful when the purpose is to track population change in a region. It can potentially provide interesting findings about the dynamics of population when it is scaled to a large region composing multiple urban areas. When the area becomes large or the spatial resolution increases there will be computational challenges with GWR. Parallel versions of GWR have been developed, but more research is needed for models specific to population estimation.
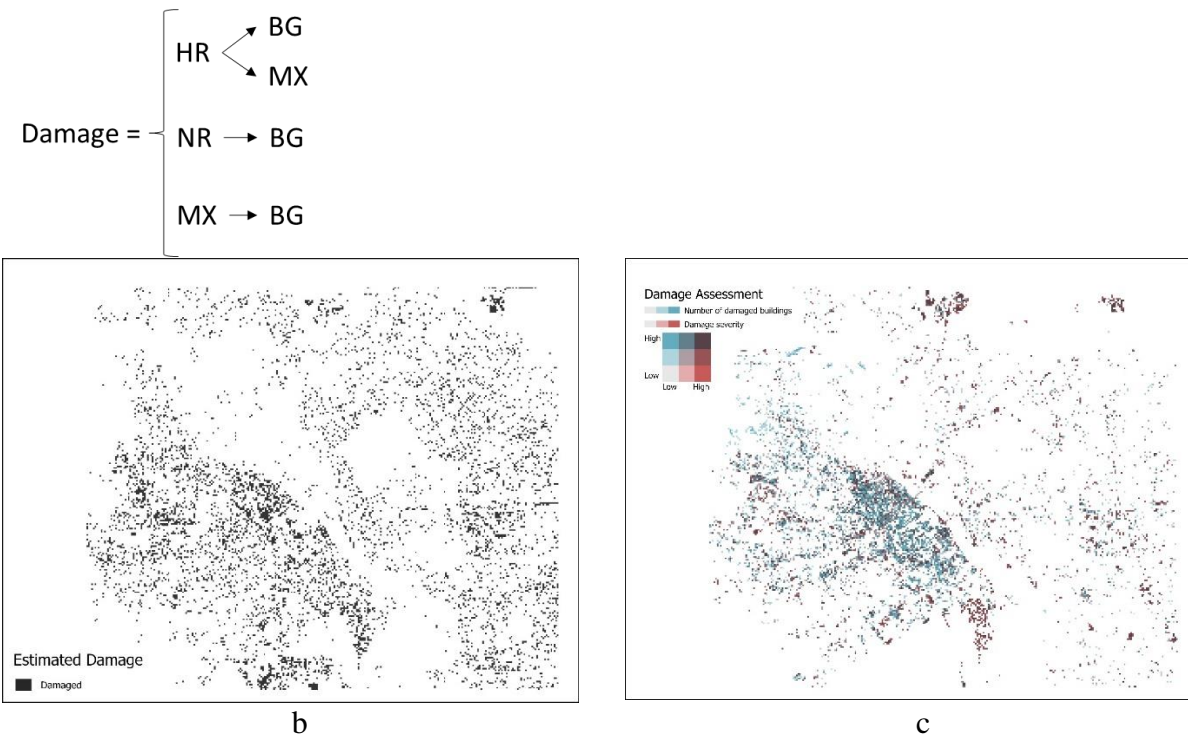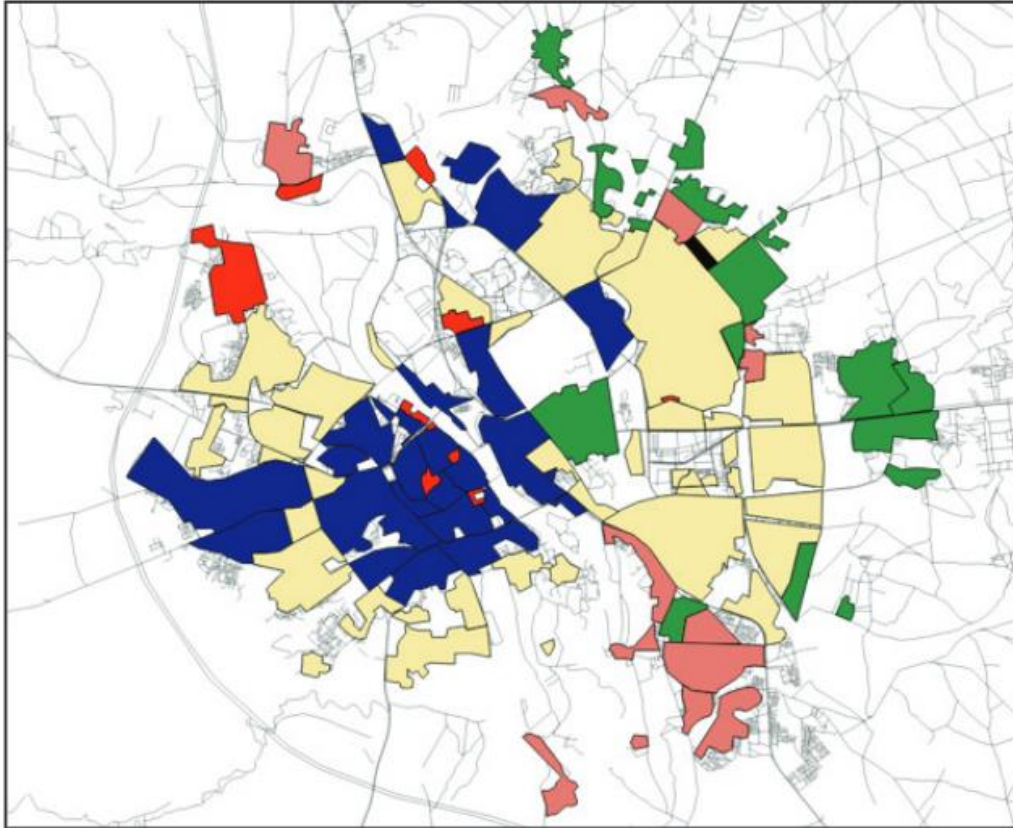
Supplementary materials:

Figure 1. Damage assessment: a) damage definition. b) Pixels that changed from 2014 to 2017 based on damage definition. c) Damage assessment from UN Humanitarian Data Exchange

Satellite damage assessment suggest that many top-floors have been damaged due to airstrikes, which may hinder the return of shopkeeper families to the shops.

Many shopowners have started reconstruction of the structures on their own.

https://unhabitat.org/the-initial-planning-framework-for-the-reconstruction-of-mosul-0

GLOBAL TRENDS FORCED DISPLACEMENT IN 2019

https://www.unhcr.org/60b638e37/unhcr-global-trends-2020

SOURCE: Michael Izady, "Mosul and Environs: Ethnic Composition, 2013," Columbia University Gulf 2000 Project, 2017.
NOTE: Red areas indicate Assyrian Christian neighborhoods, green areas indicate Kurdish neighborhoods, light red areas are Turkmen, black areas are Shia, and purple areas are Sunni. Yellow areas indicate neighborhoods with mixed ethnicity.

https://apps.dtic.mil/sti/citations/AD1085619

References

Anselin, L. (1995). Local Indicators of Spatial Association—LISA. *Geographical Analysis*, *27*(2), 93–115. https://doi.org/10.1111/j.1538-4632.1995.tb00338.x

Briggs, D. J., Gulliver, J., Fecht, D., & Vienneau, D. M. (2007). Dasymetric modelling of small-area population distribution using land cover and light emissions data. *Remote Sensing of Environment*, *108*(4), 451–466. https://doi.org/10.1016/j.rse.2006.11.020

Bruederle, A., & Hodler, R. (2018). Nighttime lights as a proxy for human development at the local level. *PLOS ONE*, *13*(9), e0202231. https://doi.org/10.1371/journal.pone.0202231

Brunsdon, C., Fotheringham, S., & Charlton, M. (1998). *Geographically weighted regressionÐmodelling spatial non-stationarity*. 431–443.

Castelluccio, M., Poggi, G., Sansone, C., & Verdoliva, L. (2015). *Land Use Classification in Remote Sensing Images by Convolutional Neural Networks*. 1–11. http://arxiv.org/abs/1508.00092

Chen, C., Zhang, B., Su, H., Li, W., & Wang, L. (2016). Land-use scene classification using multi-scale

completed local binary patterns. *Signal, Image and Video Processing*, *10*(4), 745–752. https://doi.org/10.1007/s11760-015-0804-2

Chollet, F. (2017). Xception: Deep learning with depthwise separable convolutions. *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, *2017-Janua*, 1800–1807. https://doi.org/10.1109/CVPR.2017.195

Clifford, P., Richardson, S., & Hemon, D. (2019). Assessing the Significance of the Correlation between Two Spatial Processes Author (s): Peter Clifford , Sylvia Richardson and Denis Hemon Published by : International Biometric Society Stable URL : https://www.jstor.org/stable/2532039 REFERENCES Linked. *Biometrics*, *45*(1), 123–134.

Cotter, C. (2019). Displacement in Times of Armed Conflict. How International Humanitarian Law Protects in War, and Why It Matters. *Icrc*, 1–72.

Cuevas, F., Mancilla, D., & Osorio, M. F. (2020). *Package ' SpatialPack .'*

Dobson, J. E., Bright, E. A., Coleman, P. R., Durfee, R. C., & Worley, B. A. (2000). LandScan: A global population database for estimating populations at risk. *Photogrammetric Engineering and Remote Sensing*, *66*(7), 849–857. https://doi.org/10.1201/9781482264678-24

Elvidge, C. D., Baugh, K. E., Kihn, E. A., Kroehl, H. W., & Davis, E. R. (1997a). Mapping city lights with nighttime data from the DMSP Operational Linescan System. *Photogrammetric Engineering and Remote Sensing*, *63*(6), 727–734.

Elvidge, C. D., Baugh, K. E., Kihn, E. A., Kroehl, H. W., Davis, E. R., & Davis, C. W. (1997b). Relation between satellite observed visible-near infrared emissions, population, economic activity and electric power consumption. *International Journal of Remote Sensing*, *18*(6), 1373-1379.

Fotheringham, A. S., Brunsdon, C., & Charlton, M. (2003). *Geographically weighted regression: the analysis of spatially varying relationships*. John Wiley & Sons.

Herold, M., Liu, X., & Clarke, K. C. (2003). *03-915.Qxd*. *69*(9), 991–1001.

Ioffe, S., & Szegedy, C. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. *32nd International Conference on Machine Learning, ICML 2015*, *1*, 448–456.

Iraq National Population Commission. (2012). Iraq Population Situation Analysis Report. https://iraq.unfpa.org/sites/default/files/pub-pdf/PSA%20English%202012.pdf. (accessed March 18,2022).

Kiran Chand, T. R., Badarinath, K. V. S., Elvidge, C. D., & Tuttle, B. T. (2009). Spatial characterization of electrical power consumption patterns over India using temporal DMSP-OLS nighttime satellite data. *International Journal of Remote Sensing*, *30*(3), 647–661. https://doi.org/10.1080/01431160802345685

Lafta, R., Cetorelli, V., & Burnham, G. (2018). Living in Mosul during the time of ISIS and the military liberation: results from a 40-cluster household survey. Conflict and health, 12(1), 1-8.

Li, X., Chen, F., & Chen, X. (2013). Satellite-observed nighttime light variation as evidence for global armed conflicts. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *6*(5), 2302–2315. https://doi.org/10.1109/JSTARS.2013.2241021

Li, X., & Li, D. (2014). Can nighttime light images play a role in evaluating the Syrian Crisis?

International Journal of Remote Sensing, *35*(18), 6648–6661.
https://doi.org/10.1080/01431161.2014.971469

Li, X., Liu, S., Jendryke, M., Li, D., Wu, C., Li, X., Liu, S., Jendryke, M., Li, D., & Wu, C. (2018).
Night-Time Light Dynamics during the Iraqi Civil War. *Remote Sensing*, *10*(6), 858.
https://doi.org/10.3390/rs10060858

Li, X., Ma, R., Zhang, Q., Li, D., Liu, S., He, T., & Zhao, L. (2019). Anisotropic characteristic of
artificial light at night – Systematic investigation with VIIRS DNB multi-temporal observations.
Remote Sensing of Environment, *233*(August). https://doi.org/10.1016/j.rse.2019.111357

Mennis, J. (2009). Dasymetric mapping for estimating population in small areas. *Geography Compass*,
3(2), 727–745. https://doi.org/10.1111/j.1749-8198.2009.00220.x

Mennis, J., & Hultgren, T. (2006). Intelligent dasymetric mapping and its application to areal
interpolation. *Cartography and Geographic Information Science*, *33*(3), 179–194.
https://doi.org/10.1559/152304006779077309

Miller, S. D., Straka, W. C., Yue, J., Seaman, C. J., Xu, S., Elvidge, C. D., Hoffmann, L., Azeem, I.,
Miller, S. D., III, W. C. S., Yue, J., Seaman, C. J., Xu, S., Elvidge, C. D., Hoffmann, L., & Azeem,
I. (2018). The Dark Side of Hurricane Matthew: Unique Perspectives from the VIIRS Day/Night
Band. *Bulletin of the American Meteorological Society*, *99*(12), 2561–2574.
https://doi.org/10.1175/BAMS-D-17-0097.1

Niemeyer, J., Rottensteiner, F., & Soergel, U. (2014). Contextual classification of lidar data and building
object detection in urban areas. *ISPRS Journal of Photogrammetry and Remote Sensing*, *87*, 152–
165. https://doi.org/10.1016/j.isprsjprs.2013.11.001

Ou, J., Liu, X., Li, X., Li, M., & Li, W. (2015). Evaluation of NPP-VIIRS nighttime light data for
mapping global fossil fuel combustion CO_2 emissions: A comparison with DMSP-OLS nighttime
light data. *PLoS ONE*, *10*(9). https://doi.org/10.1371/journal.pone.0138310

Penatti, O. A. B., Nogueira, K., & Dos Santos, J. A. (2015). Do deep features generalize from everyday
objects to remote sensing and aerial scenes domains? *IEEE Computer Society Conference on
Computer Vision and Pattern Recognition Workshops*, *2015-Octob*, 44–51.
https://doi.org/10.1109/CVPRW.2015.7301382

Pesaresi, M., Huadong, G., Blaes, X., Ehrlich, D., Ferri, S., Gueguen, L., Halkia, M., Kauffmann, M.,
Kemper, T., Lu, L., Marin-Herrera, M. A., Ouzounis, G. K., Scavazzon, M., Soille, P., Syrris, V., &
Zanchetta, L. (2013). A global human settlement layer from optical HR/VHR RS data: Concept and
first results. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*,
6(5), 2102–2131. https://doi.org/10.1109/JSTARS.2013.2271445

Robinson, E., Egel, D., Johnston, P. B., Mann, S., Rothenberg, A. D., & Stebbins, D. (2017). When the
Islamic State comes to town: The economic impact of Islamic State governance in Iraq and Syria.
RAND CORP ARLINGTON VA ARLINGTON.

Román, M O, Wang, Z., Sun, Q., Kalb, V., Miller, S. D., Molthan, A., Schultz, L., Bell, J., Stokes, E. C.,
Pandey, B., & Seto, K. C. (2018). NASA's Black Marble Nighttime Lights Product Suite. *Remote
Sensing of Environment*, *210*, 113–143. https://doi.org/10.1016/j.rse.2018.03.017

Román, Miguel O., Stokes, E. C., Shrestha, R., Wang, Z., Schultz, L., Sepúlveda Carlo, E. A., Sun, Q.,
Bell, J., Molthan, A., Kalb, V., Ji, C., Seto, K. C., McClain, S. N., & Enenkel, M. (2019). Satellite-
based assessment of electricity restoration efforts in Puerto Rico after Hurricane Maria. *PLoS ONE*,
14(6), 1–22. https://doi.org/10.1371/journal.pone.0218883

Shi, K., Yu, B., Huang, Y., Hu, Y., Yin, B., Chen, Z., Chen, L., & Wu, J. (2014). Evaluating the ability of NPP-VIIRS nighttime light data to estimate the gross domestic product and the electric power consumption of China at multiple scales: A comparison with DMSP-OLS data. *Remote Sensing*, *6*(2), 1705–1724. https://doi.org/10.3390/rs6021705

Simonyan, K., & Zisserman, A. (2014). *Very Deep Convolutional Networks for Large-Scale Image Recognition*. http://arxiv.org/abs/1409.1556

Sutton, P., Roberts, D., Elvidge, C., & Baugh, K. (2001). Census from Heaven: An estimate of the global human population using nighttime satellite imagery. *International Journal of Remote Sensing*, *22*(16), 3061–3076. https://doi.org/10.1080/01431160010007015

Szegedy, C., Ioffe, S., Vanhoucke, V., & Alemi, A. (2017). Inception-v4, inception-resnet and the impact of residual connections on learning. *Proceedings of the AAAI Conference on Artificial Intelligence*, *31*(1).

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., & Rabinovich, A. (2015). Going deeper with convolutions. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 1–9.

Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the inception architecture for computer vision. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2818–2826.

Tobler, W. R. (1979). Smooth pycnophylactic interpolation for geographical regions. *Journal of the American Statistical Association*, *74*(367), 519–530. https://doi.org/10.1080/01621459.1979.10481647

UN Habitat. (2016). City Profile of Mosul, Iraq.http://reliefweb.int/sites/reliefweb.int/files/resources/UN-Habitat_MosulCityProfile_V5.pdf (accessed March 18,2022).

UNHCR. (2020). Trends at a glance: Global trends forced displacement in 2020. *UNHCR The UN Refugee Agency*, 1–84. https://www.unhcr.org/60b638e37/unhcr-global-trends-2020

Waters, N. (2012). A Review of "Handbook of Applied Spatial Analysis: Software Tools, Methods and Applications." In *Annals of the Association of American Geographers* (Vol. 102, Issue 1). https://doi.org/10.1080/00045608.2011.624965

Witmer, F., & O'Loughlin, J. (2011). Detecting the effects of wars in the Caucasus Regions of Russia and Georgia using radiometrically normalized DMSP-OLS nighttime lights imagery. *GIScience and Remote Sensing*, *48*(4), 478–500. https://doi.org/10.2747/1548-1603.48.4.478

Xie, Y., Weng, Q., & Fu, P. (2019). Temporal variations of artificial nighttime lights and their implications for urbanization in the conterminous United States, 2013–2017. *Remote Sensing of Environment*, *225*(November 2018), 160–174. https://doi.org/10.1016/j.rse.2019.03.008

Yao, Y., Liu, X., Li, X., Zhang, J., Liang, Z., Mai, K., & Zhang, Y. (2017). Mapping fine-scale population distributions at the building level by integrating multisource geospatial big data. *International Journal of Geographical Information Science*, *31*(6), 1220–1244. https://doi.org/10.1080/13658816.2017.1290252

Zeng, C., Zhou, Y., Wang, S., Yan, F., & Zhao, Q. (2011). Population spatialization in china based on nighttime imagery and land use data. *International Journal of Remote Sensing*, *32*(24), 9599–9620. https://doi.org/10.1080/01431161.2011.569581

Zhang, C., Sargent, I., Pan, X., Li, H., Gardiner, A., Hare, J., & Atkinson, P. M. (2018). An object-based

convolutional neural network (OCNN) for urban land use classification. *Remote Sensing of Environment*, *216*(April), 57–70. https://doi.org/10.1016/j.rse.2018.06.034

Zhang, S., Zhang, X., Zhang, A., Fu, H., Cheng, J., Huang, H., Sun, G., Zhang, L., & Yao, Y. (2019). Fusion Of Low-And High-Level Features For Uav Hyperspectral Image Classification. *2019 10th Workshop on Hyperspectral Imaging and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, 1–4.

Zhao, B., Zhong, Y., & Zhang, L. (2016). A spectral-structural bag-of-features scene classifier for very high spatial resolution remote sensing imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, *116*, 73–85. https://doi.org/10.1016/j.isprsjprs.2016.03.004

Zhu, X. X., Tuia, D., Mou, L., Xia, G. S., Zhang, L., Xu, F., & Fraundorfer, F. (2017). Deep Learning in Remote Sensing: A Comprehensive Review and List of Resources. *IEEE Geoscience and Remote Sensing Magazine*, *5*(4), 8–36. https://doi.org/10.1109/MGRS.2017.2762307

Zou, Q., Ni, L., Zhang, T., & Wang, Q. (2015). Remote Sensing Scene Classification. *IEEE Transactions on Geoscience and Remote Sensing Letters*, *12*(11), 2321–2325.