Downscaling satellite night-time lights imagery to support within-city applications using a spatially non-stationary model

Nikolaos Tziokas [a], Ce Zhang [a,b*], Garyfallos C. Drolias [a], Peter M. Atkinson [a,c*]

[a] Lancaster Environment Centre, Lancaster University, Lancaster LA1 4YQ, United Kingdom

[b] UK Centre for Ecology & Hydrology, Library Avenue, Bailrigg, Lancaster LA1 4AP, United Kingdom

[c] Geography and Environmental Science, University of Southampton, Highfield, Southampton SO17 1BJ, UK

* Correspondence: Ce Zhang (c.zhang9@lancaster.ac.uk) and Peter M. Atkinson (pma@lancaster.ac.uk)

Highlights

- The NTL data were sharpened using RFATPK and multi-covariates.
- The sharpened NTL showed lower RMSE against Luojia1-01 NTL data.
- The downscaled NTL can be used for tracking human development at the city-scale.

Abstract

For mapping and monitoring socioeconomic activities in cities, night-time lights (NTL) satellite sensor images are used widely, measuring the light intensity during the night. However, the main challenge to mapping human activities in cities using such NTL satellite sensor images is their coarse spatial resolution. To address this drawback, spatial downscaling of satellite nocturnal images is a plausible solution. However, common approaches for spatial downscaling employ spatially stationary models that may not be optimal where the data are spatially heterogeneous. In this research, a geostatistical model termed Random Forest area-to-point regression Kriging (RFATPK) was employed to disaggregate coarse spatial scale VIIRS NTL images (450 m) to a fine spatial scale (100 m). The RF predicts at a coarse resolution from fine spatial resolution variables, such as a Population raster. ATPK then downscales the coarse residuals from the RF prediction. In numerical experiments, RFATPK was compared with three benchmark techniques, including the simple Allocation of pixel values from the coarse resolution NTL data, Machine Learning with Splines and Geographically Weighted Regression. The downscaled results were validated using fine resolution LuoJia 1-01 satellite sensor imagery. RFATPK produced more accurate disaggregated images than the three benchmark approaches, with mean root mean square errors (RMSE) for the year 2018 of 13.89 and 6.74 $nWcm^{-2}sr^{-1}$, for Mumbai and New Delhi, respectively. Also, the property of perfect coherence, measured by the Correlation Coefficient, was preserved consistently when applying RFATPK and was almost 1 for all years. The

applicability of the disaggregated NTL data to monitor socioeconomic activities at the within-city scale against the reference NTL was illustrated by utilizing them as a proxy for the Gross National Income (GNI) *per capita* and the Night Light Development Index. The GNI estimation from the downscaled NTL outperformed the coarse resolution NTL when examining their coefficients of determination, with R^2 of 0.67 and 0.47 for the GNI estimation using the fine and coarse resolution NTL data, respectively. For the Night Light Development Index (NLDI), the results of the index were compared by measuring their correlation with the Human Development Index (HDI). The NLDI from the downscaled NTL outperformed the coarse resolution NTL when measuring the correlation with the HDI, with Pearson's correlation coefficients of -0.48 and -0.35 for the NLDI using the fine and coarse resolution NTL data, respectively, for New Delhi. The outcomes indicate that RFATPK provides more accurate predictions than the three benchmark techniques and the downscaled NTL data are more suitable for fine scale socioeconomic applications, as demonstrated by the NLDI and GNI. This research, thus, shows that the RFATPK solution for NTL disaggregation can facilitate data enhancement for fine-scale sub-national applications in social sciences and can be generalized worldwide by including other cities as well as other applications.

Keywords: satellite night-time lights, downscaling, random forest regression, area-to-point Kriging, spatial non-stationarity, socioeconomic applications

1. Introduction

Human development is a crucial factor to consider when assessing a nation's degree of development since it gives inhabitants equal chances and fair choices, extends their lives, and improves their living conditions, health care, and education (Wang et al., 2021). In September 2016, the world committed to implementing the 2030 Agenda for Sustainable Development. The Sustainable Development Goals (SDGs), according to Reid et al. (2017), strike a balance between the economic, social, and environmental dimensions of sustainable development. Despite decades of tremendous progress in eradicating poverty and fostering wealth, a sizable segment of the world's poorest population still encounters difficulties to maintain an acceptable standard of living in emerging nations, particularly Asia, Africa, and Latin America and the Caribbean. It appears that regional and national differences have led to the unequal reduction of severe poverty in these areas (Georgeson et al., 2016; Omar and Inaba, 2020). To achieve the goals of SDGs, we need better ways to collect and interpret information about many aspects of human development in a timely, accurate and appropriate manner.

The traditional approach to examining human growth and well-being is based mostly on survey data, which includes information on income, consumption, health, education, and housing. These surveys are usually carried out every three to five years, but collecting survey data is expensive and tedious process. Between surveys, detailed socioeconomic data are still needed (Watmough et al., 2013). Moreover, countries in war or extreme poverty may even lack these survey data for years (Zhao et al., 2019). In addition, fewer than two census surveys in many developing nations, such as African nations, were carried out in the decade leading up to 2000, limiting the construction of nationally representative human development metrics (Jean et al., 2016). Additionally, several nations, like India, have suspended measures like unemployment (Dasgupta, 2022). Another

limitation of these censuses is that population sizes between censuses are projected, frequently with linear yearly growth rates, despite the fact that censuses are expensive and may only be undertaken at sporadic intervals when resources are scarce. Despite the high levels of uncertainty in the estimates, they are utilized to evaluate, for example, the dangers to public health and need for health services. Additionally, censuses are unable to reflect accurately intra-annual changes in a nation's socioeconomic conditions since they are not designed to do so (Bharti and Tatem, 2018).

Using new passively gathered data sources, such as information from satellite sensors, provides an alternate method of monitoring socioeconomic processes. Such data can help address the challenge of scaling up (i.e., increasing the temporal resolution of) traditional data collection efforts which are generally very limited in frequency due to financial cost (Jean et al., 2016). Early studies used satellite "night-lights" data to demonstrate that areas with more economic output tended to emit more artificial light (Head et al., 2017). Nocturnal images, such as the Day-Night Band (DNB), from the Visible Infrared Imaging Radiometer Suite (VIIRS) is a valuable source of satellite imagery. The VIIRS is onboard the Suomi National Polar-orbiting Partnership (SNPP) satellite. The ability for researchers to track socioeconomic activity is made possible by the worldwide coverage and coarse spatial resolution of these data, which have pixels that are less than one square kilometer in size. Additionally, nighttime lighting is consistently assessed across nations with extremely diverse institutional capacity and is not prone to manipulation for political reasons (Zhang and Gibson, 2022). When compared to commercial fine-resolution images like EROS-B or JL1-3B, NTL products (like VIIRS DNB images) are available for free and have a considerably larger swath (Levin et al., 2014). The NPP-VIIRS NTL has a spatial resolution of 15 arc seconds (or approximately 500 m at the Equator), which has the potential to support several practical applications like mapping at the country level, detecting military conflicts and assessing poverty (Levin et al., 2020; Gibson et al., 2021).

NTL has achieved extensive research and applications in socioeconomic fields. The so-called Night Light Development Index (NLDI), proposed by Elvidge et al. (2012) from nighttime satellite sensor images and population density, evaluates disparities in the local population's geographical distribution of night light. Using deep learning techniques, Bruederle and Hodler (2018) demonstrated that NTL data are a suitable proxy for wealth and human development in 29 African countries. Similar to the previous study, Yeh et al. (2020) estimated the wealthiness of 20,000 African villages using a combination of NTL data and daylight satellite sensor optical data, and found that their technique could account for 70% of the variation in ground-measured village wealth. Gosh et al. (2013) provided examples of numerous ways to gauge one's level of wellbeing. Elvidge et al. (2011) used NTL satellite sensor imagery and population data to estimate the number of people worldwide who have (or do not have) access to electricity. This is done because a lack of electricity is a sign of poverty and is associated with conditions that are detrimental for health and wellbeing, including the inability to refrigerate food, have access to clean water, and have adequate sanitary facilities. NTL and artificial neural networks were utilized by Jasiński (2019) to gauge electricity usage at the Nomenclature of territorial units for statistics (NUTS) 2 level.

Mapping and monitoring complex urban socioeconomic processes, particularly those that take place within cities, can be challenging with NTL images since they often have a coarse spatial

resolution (Levin et al., 2020; Ye et al., 2021). According to Elvidge et al. (2007), the coarsest acceptable spatial resolution of a satellite sensor image should exceed 100 m to research socioeconomic issues at the city scale. Moreover, it can be important to track human development over time to determine if it is improving or developing. For example, a finer spatial resolution is typically needed than the 450 m pixel size of VIIRS to accurately quantify development growth rates at the scale of individual cities. These problems significantly hinder NTL applications, especially at the city-scale. Enhancing the spatial resolutions of NTL satellite sensor imagery products is increasingly urgent because the majority of the world's population, after 2007, is concentrated in urban areas (Marlier et al., 2016).

In remote sensing, spatial downscaling can be categorized in two classes based on their output prediction, namely downscaling continua and sub-pixel mapping (SPM) (Wang et al., 2020). Whereas the first class predicts continua (e.g., in units of reflectance, brightness, etc.), the latter class, also known as super resolution mapping in the remote sensing literature, predicts categories (i.e., land cover class labels) (Wang et al., 2020). Downscaling continua can create categorical products by classification and is more often used. Generally, the methods for disaggregating continua can be classified into the following classes: general statistical, spatial statistical, machine learning, process-based, wavelet-based techniques, fractal techniques and hybrid methods (Park et al., 2019). Spatial statistical analysis has advanced the downscaling of raster images (i.e., satellite remote sensing images), notably in terms of spatial interpolation, by taking advantage of the spatial autocorrelation among geographical data. Area-to-point (ATP) interpolation, as opposed to generic spatial interpolation, can address the problem of changeable areal units, when the supports before, and during, downscaling are different (Kerry et al., 2012; Wang et al., 2016a). By making sure, for instance, that the total of the downscaled forecasts within each region equals the initial aggregated count, ATP Kriging (ATPK) assures the coherence of predictions (Kyriakidis and Yoo, 2005). Yoo and Kyriakidis (2006) expanded on ATPK by taking the inequality limitations in spatial interpolation into account. ATP interpolation emphasizes utilizing the information offered by correlated variables since it can help in exploring the spatial variation of response variables at a higher spatial resolution.

Wang et al., (2016b) further extended ATPK by introducing a regression term and they named the method area-to-point regression Kriging (ATPRK). ATPRK is a geostatistical technique used frequently for downscaling day-time satellite sensor images. For example, Wang et al. (2015) used ATPRK to downscale MODIS data and further employed it to downscale Landsat satellite sensor images and Worldview-2 images (Wang et al., 2016a). Using multispectral satellite sensor imagery, Zhang et al. (2021) developed object-based ATPRK to disaggregate IKONOS images. Wang et al. (2017) implemented a spatio-temporal fusion method by combining MODIS and Landsat data, downscaling 500 m MODIS data into 250 m as an initial step. Jin et al. (2018b) utilized Geographically Weighted Area-to-Point Regression Kriging (GWATPRK), a spatially non-stationary method, to create a 1 km Surface Soil Moisture product. Xu et al. (2020) downscaled ASTER thermal images using Random Forest ATPK. The majority of the above research, as stated in the examples, aimed to downscale day-time satellite sensor data. However, the remote sensing of nocturnal lighting has the potential to capture human socioeconomic activities and/or track human development compared to day-time satellite sensor data, which is

158 critical in modelling complex urban environments for certain applications (Elvidge et al., 2017)
159 and downscaling is potentially useful in this context. Thus, there exists a gap in the literature.
160 However, the spatial pattern of NTL is diverse. For example, the light intensity differs depending
161 on the land use (Ye et al., 2021), and the spatial pattern of NTL intensity varies from geographic
162 region-to-region, even within the same area (e.g., city).
163
164 Earth-observed variables also may exhibit spatial heterogeneity in addition to spatial
165 autocorrelation (Jin et al., 2018). For such spatially diverse variables, the global model used in
166 ATPRK may be unable to adequately capture local characteristics in the multivariate data. In
167 essence, the global ATPRK model assumes that the process under inquiry is constant across space.
168 Where the data exhibit spatial heterogeneity a more flexible model is needed; one that permits
169 spatial non-stationarity in some model parameters.

170 Wang et al. (2016a) extended the ATPRK by incorporating and adaptive window for the regression
171 part in order to account for the data's spatial heterogeneity. An ordinary linear regression model
172 was fitted using a coarse target variable and covariates within a local window, that is, a global
173 regression model within the constricted region. On the other hand, Random Forest regression (RF)
174 is a well-known non-stationary regression technique that takes into account non-linear correlations
175 between variables and has been frequently utilized for spatial analysis, either alone or in
176 combination (Xu et al., 2020; Tang et al., 2021). Xu et al. (2020) proposed RF area-to-point
177 regression Kriging (RFATPK) for downscaling ASTER land surface temperature data. Cheng et
178 al., (2022) utilized a RF and ATPK to estimate monthly population distribution in China. However,
179 only a few related studies exist focusing on downscaling satellite night-time lights images, mainly
180 for impervious surface identification (Ye et al., 2021). To the authors' knowledge, there hasn't been
181 any prior research that specifically addresses NTL continua for socioeconomic applications.

182 To account for the issue of local heterogeneity and multivariate non-linearity, this research
183 proposes the combination of RF and ATPK for disaggregating NTL using fine-spatial-resolution
184 predictors (e.g., WorldPop products). The suggested RFATPK technique captures the spatially
185 non-linear correlations between the dependent and auxiliary variables while preserving the benefits
186 of ATPRK. The advantages of the proposed algorithm are: The advantages of the proposed
187 algorithm are: (1) RF can process high-dimensional data. (2) Overfitted phenomena do not easily
188 occur, because the final estimation is made through the average prediction of the decision trees.
189 (3) For a large number of remote sensing images and/or observations, training a RF model is fast
190 and efficient. (4) RF is immune to statistical assumptions compared to the original ATPRK. (5)
191 Another advantage is its ability to capture complex and non-linear relationships between predictors
192 and the response variable (Brokamp et al., 2017). By downscaling NTL satellite sensor images
193 from the VIIRS sensor from 450 m to 100 m, two practical socioeconomic applications were
194 executed to test this approach. The applications involved the Gross National Income (GNI) and
195 Night Light Development Index (NLDI) for New Delhi and Mumbai, two megacities of the
196 developing world. Disaggregated NTL data at this spatial resolution have the potential to be

utilized for monitoring such human development indicators at the city-scale. The research's contributions are, thus, two-fold.

1. The geostatistical approach, ATPRK, combined with a spatially non-stationary model, was applied to downscale NTL images for the first time. To show that RFATPK is superior than the three previous well-known downscaling methodologies, a benchmark comparison between the three approaches and RFATPK was performed.
2. The spatial downscaled NTL data were further applied to estimate the GNI as well as to measure light inequality at the within-city scale by comparing them with equivalents using the coarse spatial resolution NTL.

The remainder of this research paper is organized as follows. The research areas and the data used are described in Section 2. The suggested downscaling technique is described in Section 3. We give the results in Section 4. We expand on the suggested downscaling approach in Section 5 before presenting our conclusions in Part 6.

2. Study areas and data

2.1. Study areas

New Delhi and Mumbai were chosen as the areas under investigation to manifest the utility of the technique (Figure 1). New Delhi is India's capital and a densely populated metropolis (~10,400 people per km^2) with a population of about 16 million people and it is a key international hub on the Asian sub-continent. The city has experienced rapid urbanization and industrialization, with 93% of the population residing in urban areas compared to the national average of 31.16%. There are 675 slum clusters in New Delhi (Figure 1) (Bhanarkar et al., 2018; Malik et al., 2022). The Mumbai Metropolitan Region, including Mumbai and its surrounding suburban area, is known as India's economic engine as it accounts for over 6.16% of India's GDP, providing 10% of industrial jobs. More than 20 million people live in this territory today and this amount is predicted to increase by 10.36 million by the end of 2036. Mumbai, like New Delhi, has a large number of slums, with an estimated number of 67 (Nijman, 2010; Vinayak et al., 2021).

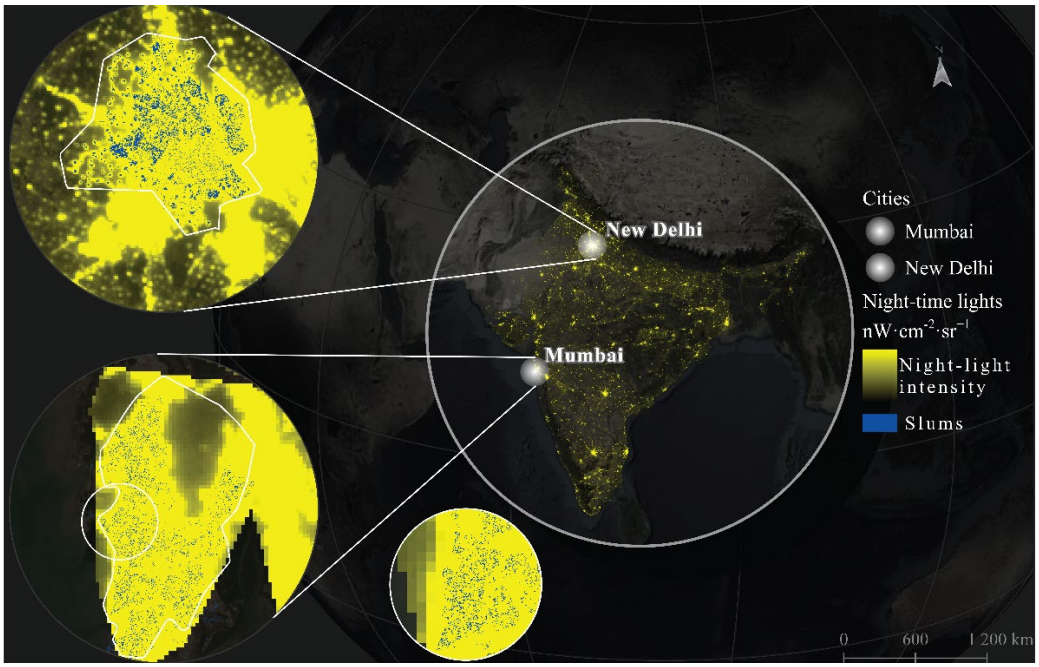

Figure 1: Location and maps of the two cities of New Delhi and Mumbai.

2.2. Datasets

Remote sensing nocturnal images, population count, Landsat's 8 thermal band and the global human settlement layer data were used. A summary of these follows below:

1) Version 2.1 of the NPP-VIIRS DNB cloud-free yearly composite NTL product for 2013 to 2020 for Mumbai and New Delhi, respectively, were acquired from the Earth Observation Group website (https://eogdata.mines.edu/products/vnl/) (accessed 10/01/2023). The pixel size of the NTL images was 450 m.
2) Yearly population count data were derived from the WorldPop website (https://www.worldpop.org/) (accessed 10/01/2023) for the calculation of NLDI, the estimation of GNI and to assist the spatial downscaling. The NLDI and GNI were used to highlight the applicability and superiority of the downscaled NTL product compared to the original coarse resolution NTL.
3) Landsat 8 data were used to obtain land surface temperature (LST). We selected Landsat 8 OLI/TIRS yearly median cloud-free imagery for the same years as for NTL. The nominal resolution of the initial images was 100 m.
4) The global human settlement (GHS) layer is a human settlement map product covering the entire world (Pesaresi and Politis, 2022). We used the GHS layers of 2015 and 2020 as well as the average global building height (AGBH) product of 2018 (accessed 10/01/2023).
5) The results of the downscaling were validated using Loujia1-01 imagery. Loujia1-01 data has a pixel size of, approximately, 120 m and wider spectral range compared to the VIIRS NTL data (Liu et al., 2020).

3. Methodology

The Methodology is organized as follows: (1) Firstly, a brief introduction of the ATPRK is given. (2) A detailed explanation of the proposed RFATPK and its parts (i.e., RF regression and ATPK) follows. Additionally, a description of the benchmark methods is given and lastly the two socioeconomic applications. Figure 2 summarizes the methodology as a series of successive steps designed to meet the research objectives. The first part includes selection of the Inputs (the target variable and the covariates), namely the NTL data, WorldPop product, the LST band from Landsat 8 and the GHS and AGBH layers, respectively. Then, the data were regressed utilizing RF regression and the predictions were separated from the residuals. In the third part, the residuals from the RF model were downscaled using ATPK. Finally, the prediction was added to the downscaled residuals and the NTL raster layer at 100 m spatial resolution was produced.

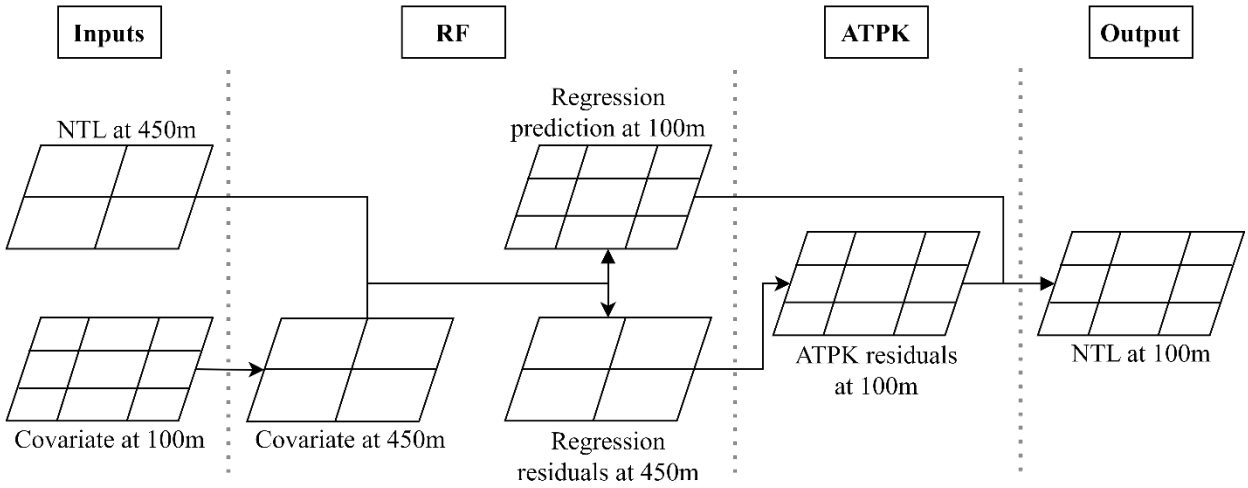

Figure 2: Flowchart of RFATPK. The first part includes the Inputs which are the target variable and the covariate. In the second part, the input data are regressed using RF regression. The third part involves ATPK-based downscaling of the residuals. Finally, the prediction is added to the downscaled residuals and the NTL raster layer at 100 m spatial resolution is produced.

3.1 Downscaling

ATPRK is a spatial downscaling method that applies a regression model to coarse spatial resolution data and subsequently applies ATPK to enhance the spatial resolution of the residuals (Wang et al., 2016a). The regression component alone is insufficient for disaggregation because it does not utilize fully the spectral characteristics in the observed low-resolution data. As an addition to the regression step, ATPK-based residual downscaling is utilized to account for the spectral characteristics of the coarse data. The ATPK component is a sharpening method that predicts values on a smaller pixel (i.e., support) than the original, coarse scale, data (Kyriakidis, 2004; Kyriakidis and Yoo, 2005; Atkinson, 2013). It varies from conventional Kriging in that it takes into account the observation's spatial support and it accounts for the size of support, spatial correlation and the point spread function (PSF) of the sensor, instead of treating each observation as a centroid. Moreover, a crucial advantage of ATPK is, the so-called, property of perfect

coherence (Kyriakidis, 2004; Kyriakidis and Yoo, 2005): that is, it can maintain accurately the spectral features of the nominal coarse data.

The regression model in ATPRK has two parts, the prediction and the residuals. The residuals can be extracted as follows:

$$e(S_i) = z(S_i) - [\beta_0 + \sum_{k=1}^{K} \beta_k * h_k(S_i)] \quad (1)$$

where $e(S_i)$ are the coarse residuals, β_0 and β_κ are coefficients of the linear regression, $z(S_i)$ is the target random variables of coarse pixel S_i and $h_k(S_i)$ is the aggregated fine pixel within the coarse one.

The coarse residuals are downscaled using ATPK. The residual of a fine-resolution pixel s_j is estimated as a linear fusion of $e(S_l)$ ($l = 1, \ldots, L$) in L nearest coarse pixels, via ATPK:

$$e(s_j) = \sum_{l=1}^{L} \lambda_{jl}\, e(S_l) \quad (2)$$

where λ_{jl} represents the weights for the prediction at fine scale that honor the sum-to-one constraint $\sum_{l=1}^{L} \lambda_{jl} = 1$. The weights can be calculated by lessening the error variance of the prediction. The analogous Kriging matrix is depicted in Equation 6:

$$\begin{bmatrix} \gamma_{11}^{SS} & \gamma_{1L}^{SS} & 1 \\ \vdots & \ddots & \vdots \\ \gamma_{L1}^{SS} & \gamma_{LL}^{SS} & 1 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} \lambda_{1j} \\ \vdots \\ \lambda_{jL} \\ \mu_j \end{bmatrix} = \begin{bmatrix} \gamma_{1j}^{Ss} \\ \vdots \\ \gamma_{Lj}^{Ss} \\ 1 \end{bmatrix} \quad (3)$$

where γ_{ij}^{SS} is the block-to-block (i.e., area-to-area) variogram among coarse pixels S_i and S_j, C_{jj}^{ss} is the point-to-point covariance between fine spatial resolution pixels s_j and s_j, γ_{lj}^{Ss} is the ATP variogram between high resolution pixel s_j and coarse resolution pixel S_l and the μ_j are Lagrange multipliers. The covariance can be produced from the variogram.

The error variance δ of the ATPK prediction for the s_j at fine-resolution can be calculated as follows:

$$\delta_{ATPK}(s_j) = C_{jj}^{ss} - \sum_{l-1}^{L} \lambda_{jl} * C_{lj}^{Ss} - \mu_j \quad (4)$$

is C_{lj}^{Ss} the area-to-point covariance between coarse spatial resolution pixels S_l and fine spatial resolution pixels s_j.

The generation of the point support variogram is considered the most crucial step in area-to-point Kriging method, for which Wang et al., (2016b) provides the necessary details, including an explanation of how to employ a deconvolution process. The target fine pixel and the original coarse pixel can be used as point and area supports, respectively, in the ATPRK prediction, which can be described as follows:

$$z(s_j) = \hat{\beta}_0 + \sum_{k=1}^{K} \hat{\beta}_k * h_k(s_j) + \sum_{l=1}^{L} \lambda_{jl} * [z(S_l) - (\hat{\beta}_0 + \sum_{k=1}^{K} \hat{\beta}_k * h_k(S_l)] \quad (5)$$

3.2 Random Forest area-to-point Kriging

In the presence of spatial variability in a region, the global regression approach in the original ATPRK implementation is unsuitable for characterizing this variability. A non-stationary model is more appropriate when the association between the target and variable and the covariates varies geographically (Jin et al., 2018a, 2018b). The global regression residuals, on the other hand, may not meet stationarity criteria (such as second-order stationarity), making Kriging interpolation hard to implement (Jin et al., 2018b). Moreover, to account for the spatial variability in the correlation between the variables, RF regression generates local coefficients (Jin et al., 2018b; Pereira et al., 2018).

RFATPK is proposed in this research to increase downscaling accuracy by taking into account spatial non-stationarity. The trend and residuals are likewise included in RFATPK, with the trend being fitted utilizing the RF approach (Equation 2). To predict the spatial trend at fine scale, RFATPK first fits the RF regression model between the response variable and the covariates at the coarse spatial resolution. The regression residuals are then disaggregated at the desired pixel size using ATPK (Xu et al., 2020; Cheng et al., 2022). After the regression using RF, the model's errors (i.e., residuals) are expected to be less heterogeneous and assure the requirements for semivariogram estimation (i.e., sufficiently large and homogeneous areas) (Jin et al., 2018b). In this research, a deconvolution procedure was utilized to implement the ATPK predictions and the spherical model was fitted to the experimental variograms (Goovaerts, 2008). This algorithm for enforcing ATPK requires inversion of a large matrix, which is computationally expensive. For the downscaling process the R software and the package atakrig were utilized (Hu and Huang, 2020).

To generate the downscaled 100 m NTL, the RFATPK disaggregating approach of combining the RF (Breiman, 2001) and ATPK (Kyriakidis, 2004) methods was developed. The spatial non-stationarity of the regression's residuals was taken into account by the RFATPK, as well as the nonlinear association between NTL and the covariates.

Suppose $Z_C^l(x_i)$ are the pixel values (i.e., gray value) of pixel C located at x_i ($i = 1, \ldots, M$, where M is the number of pixels) in coarse image l ($l = 1, \ldots, B$, where B represents the amount of images) and $Z_F(x_j)$ is the value of pixel F centered at x_j ($j = 1, \ldots, MG^2$, where G is the zoom factor between the coarse and fine bands) in the stack layers. The letters F and C represent the fine and coarse pixels, respectively. The goal of sharpening is to predict response variables $Z_F^l(x)$ for all fine pixels in all B coarse images. RFATPK consists of two steps: RF regression and residual downscaling using ATPK. Assume $\hat{Z}_{F1}^l(x)$ and $\hat{Z}_{F2}^l(x)$ are predictions of the RF regression and ATPK parts, the RFATPK forecast is:

$$\hat{Z}_F^l(x) = \hat{Z}_{F1}^l + \hat{Z}_{F2}^l \quad (6)$$

3.2.1 Random Forest Regression modelling for the trend prediction

The RF is a non-parametric machine learning (ML) method for regression tasks (Breiman, 2001), which has been applied to fields such as, population mapping and properties relating to the soil (Cheng et al., 2022; Takoutsing & Heuvelink, 2022). Based on bagging method of the training data, the RF constructs an ensemble or forest of individual and non-correlated trees, saves the best randomly chosen variable combination for each node of each tree, and then uses an average of the individual trees' predictions to make the final prediction (Cheng et al., 2022).

Since they offer more useful higher spatial resolution and richer textural information than the response low resolution variable, the covariate(s) in RFATPK (e.g., the Population raster) are utilized to detrend the $Z_F^l(x)$ and are crucial in sharpening. The regression stage aims to fully use the fine spatial resolution textural and geographic information in the given data by characterizing the correlation between each coarse response image and the fine predictors.

A fine-scale predictor (e.g., Population raster) Z_F is initially aggregated to Z_C to match the pixel size of the coarse response image (Wang et al., 2016a). The relationship between Z_C and each coarse band l is then established by RF regression.

The generic equation of the RFATPK involves two parts, the trend component and the residuals component, and can be written as:

$$Z_C^l(x) = f(B(x_i|\theta) + R(x) \quad (7)$$

Where f is a RF model, $B(x_i)$ represents the predictors at location x_i, θ constitutes the model's parameters and $R(x)$ are the residuals, or model error. The RF-based nonlinear regression model (θ) in Equation 7 can receive the fine resolution predictors directly, based on the scale-invariance assumption. The NTL spatial trend can then be produced at a downscaled 100 m spatial resolution. Due to the availability of the predictors at the fine spatial scale, the RF regression prediction at a location x at the fine spatial scale, that is, $\hat{Z}_{F1}^l(x_0)$, is calculated as:

$$\hat{Z}_{F1}^l(x_0) = f(B(x_i|\theta) \quad (8)$$

It is crucial, when using RF, to fine-tune the model parameters (Takoutsing and Heuvelink, 2022).

3.2.2 Random Forest regression parameter fine-tuning

3.2.2.1. Default Random Forest regression model parameters at the coarse spatial scale

First, we used R's caret package to conduct RF regression using all the covariates and the default model settings. 500 trees, a node size value of 5 and a third of the total number of covariates (mtry) were included in the default model parameters. The entire study region was considered in this step.

3.2.2.2. Model calibration and fitting

The study region was divided initially into two sets, the training and a test set. The splitting of the two sets was conducted based on a stratified random sampling. This is an efficient sampling method because it captures the variability of multiple inputs of auxiliary information in the feature space (Getis and Ord, 1992). At the 450 m, the training and test samples for Mumbai were 1617

and 450, respectively, while for New Delhi were 5852 and 450. For the 2025 m scale, the splitting sets for Mumbai were 79 and 20, while for New Delhi were 285 and 72 for the training and test set, respectively. The RF model was then calibrated using the training data and its performance was assessed using the test set. We applied the model to the entire region if the R^2 difference between the training and test sets was minimal. Two user-defined arguments (the number of trees (ntree) and the number of variables chosen at each split (mtry)) were used to calibrate the RF model (Probst and Boulesteix, 2017). For the ntree parameter, we investigated a range starting at 500 and increasing to 9000 with a step of 500. The default setting for mtry was the third the total number of covariates, rounded down. With the help of the R package ranger, we fitted a final RF model for each annual NTL image using all of the pixel data, the predictors and the chosen fine-tuned arguments for the ntree and mtry.

3.2.3. Spatial prediction at the fine spatial scale

The average of all measurements embedded in one of the end nodes of the tree serves as the forecast of a single decision tree of RF for a new site x_0. By branching through the tree depending on the covariate values at x_0, the end node may be located.

The RF prediction can be calculated by taking the mean of all tree forecasts. Because it is a weighted linear combination of the measurements, it can be represented as:

$$\hat{Z}_{F1}^{l}(x_0) = \sum_{i=1}^{n} w_i * y_i \tag{9}$$

where $\hat{Z}_{F1}^{l}(x_0)$ stands for the prediction, n, w_i and y_i are the number of measurements, the weights and the NTL measurements, respectively. Note that the weights are obtained from the variables at the observed and predicted location, even though this isn't stated explicitly in Equation 1. (Takoutsing and Heuvelink, 2022).

3.3 Benchmark methods

In this research, the proposed approach was compared to three benchmark methods, namely GWR, Machine Learning with Splines and the Allocation of raster values. The benchmark methods are described below.

Prior to GWR, simple linear regression models were, thereafter, fitted to reveal the model's R^2 and AIC (Wang et al., 2015; Middya and Roy, 2021). The covariates that contributed to the linear model with the largest R^2 were also used for GWR. The GWR model, can be represented as follows:

$$z(s_j) = \hat{\beta}_0(s_j) + \sum_{k=1}^{K} \hat{\beta}_k(s_j) * h_k S_j \tag{10}$$

where, $\hat{\beta}_0(\cdot)$ and $\hat{\beta}_k(\cdot)$ represents the estimated GWR coefficients with spatial locations centered at fine pixel s_j and coarse pixel S_j, respectively.

For GWR's kernel a Gaussian function was selected and the width of the kernel was determent using an adaptive spatial kernel function (Chen, 2015). The Gaussian function describes the

relationship between the weight W_{ij} and distance from center d_{ij} and is a continuous monotonically decreasing function. The Gaussian function is used widely:

$$W_{ij} = exp\,(-d_{ij}^2/b^2) \quad (11)$$

where b and d_{ij} are the kernel bandwidth and the distance between two locations i and j, respectively. According to Chen et al. (2015) the regression results are sensitive to parameter b which can, thus, be obtained by cross-validation.

Machine Learning with Splines (ML with Splines), in order to predict the dependent variable, the algorithm tries many ensembles of six and giving one ensemble as an output, weights them differently and evaluates the fit. Six algorithms are included in this approach, namely: 1) boosted regression trees, 2) generalized additive model, 3) multivariate adaptive regression splines, 4) neural networks, 5) RF, 6) support vector machines. The algorithm interpolates noisy multivariate data through ensemble machine learning (EML). Additionally, using thin-plate-smoothing splines, the residuals of the final model are interpolated from the full training dataset. In the final ensemble model, this produces a continuous error surface that is used to eliminate the majority of the remaining errors (Bullock et al., 2020).

With the allocation-based method, a new fine spatial resolution raster (i.e., 100 m pixel size) is created with null cell values, but with the same spatial reference system as the coarse resolution raster and then the two rasters are properly overlaid. Then, the pixels of the newly created empty raster are given a value corresponding to the pixel value of the overlaid coarse spatial resolution raster. This approach, thus, represents the "do nothing" or "null" baseline and all other methods of allocation should improve on this baseline if they add any useful information.

3.4 Two socioeconomic criteria

The use of NTL as a proxy to various socioeconomic indexes is a major application. Therefore, the application of downscaled imagery to proxy the Gross National Income *per capita* and the Night Light Development Index (NLDI) is meaningful to illustrate the necessity of downscaling.

Payments go toward a country's Gross National Income (GNI), which is comprised of the GDP plus net revenues from employee compensation and foreign property income. The money that foreign migrants send to their home nations is known as remittances (Ghosh et al., 2009). To measure the association between the GNI and the NTL at the city scale, we sum all the lit pixels of the NTL, where "lit pixel" means a radiance value equal to or greater than 1 $nWcm^{-2}sr^{-1}$. Then we computed two linear regression models, one using the coarse resolution NTL as explanatory variable and one linear model using the disaggregated NTL and compared their R^2 values (Gibson and Boe-Gibson, 2021). The dependent variable in both cases was the GNI and it was measured in 1000 US dollars.

The NLDI varies from 0 to 1, representing perfect equality and inequality, respectively. The two geo-referenced gridded layer inputs to the NLDI were the population count raster and the NTL image.

The brightness (NTL's pixel value) and population count were associated in tables created using crosstabulation. In order to compute the NLDI, the two rasters were stacked and the joint distribution of brightness and population count in cell was calculated. To measure equality in the

geographic distribution of lights, the Gini index was computed based on the statistical distribution (i.e., the table containing the pixel values of NTL and Population, sorted by the NTL) according to the formula:

$$R = 1 - \frac{2\sum_{i=1}^{n-1} Q_i}{n-1}, 0 \le R \le 1 \quad (12)$$

where R and n represents the NLDI and the number of raster images, respectively, $Q_i = \sum_{j=1}^{i} x_j / \sum_{j=1}^{n} x_j$ is the number of lights corresponding to the raster with the proportion P_i of population count in which x_j is the value of light intensity class. Moreover, $P_i = \sum_{j=1}^{i} x_j / n$.

4. Results

The experiments were conducted in the two mega-cities each month between 2013 and 2020. To evaluate the results of downscaling, due to a lack of validation data at 100 m for each year, we upscaled the NTL observations to 2025 m spatial resolution and used the original NTL data at 450 m spatial resolution as the reference (Ge et al., 2019). In the downscaling stage, the coarse 2025 m NTL data were disaggregated to the initial finer spatial scale (450 m) and were validated using the raw nocturnal data.

Lastly, the sharpening was conducted to the 450 m data to predict NTL at the 100 m. For the year 2018 the downscaled results were compared against LuoJia1-01 as an extra validation step. Additionally, because the variogram can reflect the benefits of downscaling prediction, it can be used as an assessment metric when there are no reference data available at the fine spatial resolution. Thus, here, the downscaling predictions at 100 m spatial resolution were also evaluated using the variogram (Wang et al., 2020).

4.1. Comparison with other downscaling methods

4.1.1. Downscaling prediction (2025 m to 450 m)

To demonstrate the superiority of the proposed approach, the predicted NTL images were compared against the predictions of three benchmark methods and the results were illustrated in Figure 4 and Figure 5 for Mumbai and New Delhi, respectively. It can be shown that RFATPK and GWR-based downscaling achieved good agreement with the original NTL product when comparing the spatial patterns of the downscaling results with the Reference image. Although local detailed variance may be seen, the RFATPK prediction shows it more clearly. In comparison, the blocky artifacts are highly noticeable and Machine Learning with Splines and Allocation-based downscaling failed to maintain the patterns in NTL. For the instance of Machine Learning using Splines, over-fitting issues can be used to explain this outcome. Since the raw NTL coarse reference data are known perfectly in the experiment, preservation of the original patterns is the desired target.

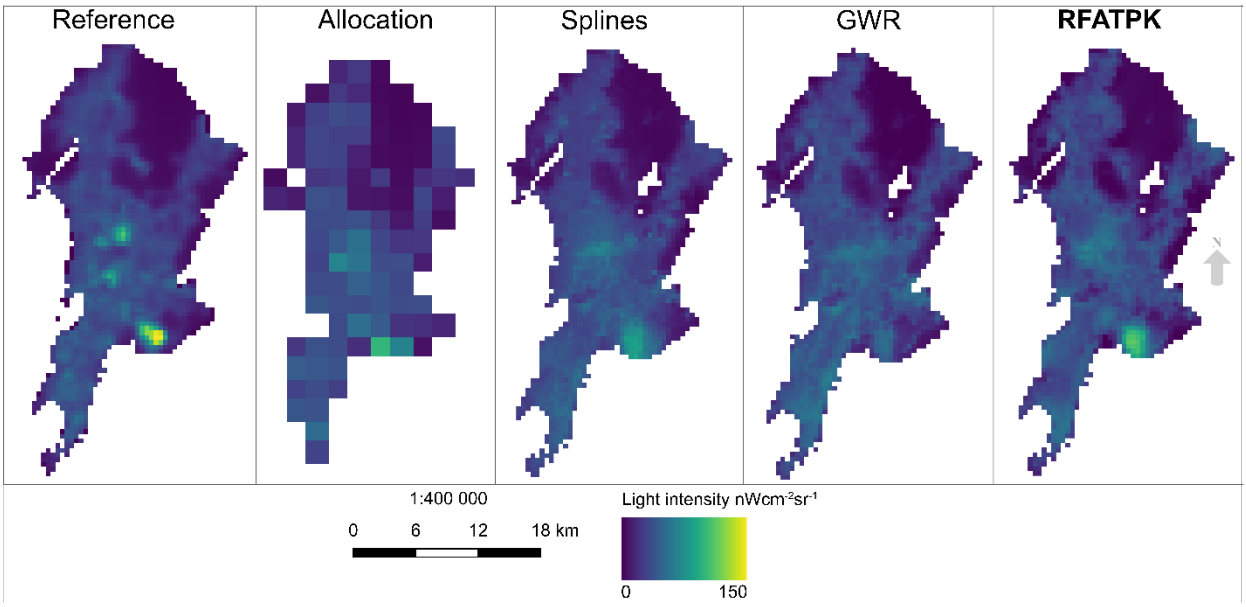

Figure 3: Downscaling results of NTL at 450 m for Mumbai, 2018. From left to right, the Reference NTL, Allocation-based downscaled NTL, Machine Learning with Splines-based downscaled NTL, GWR-based and RFATPK-based downscaled NTL. Bold shows the proposed method.

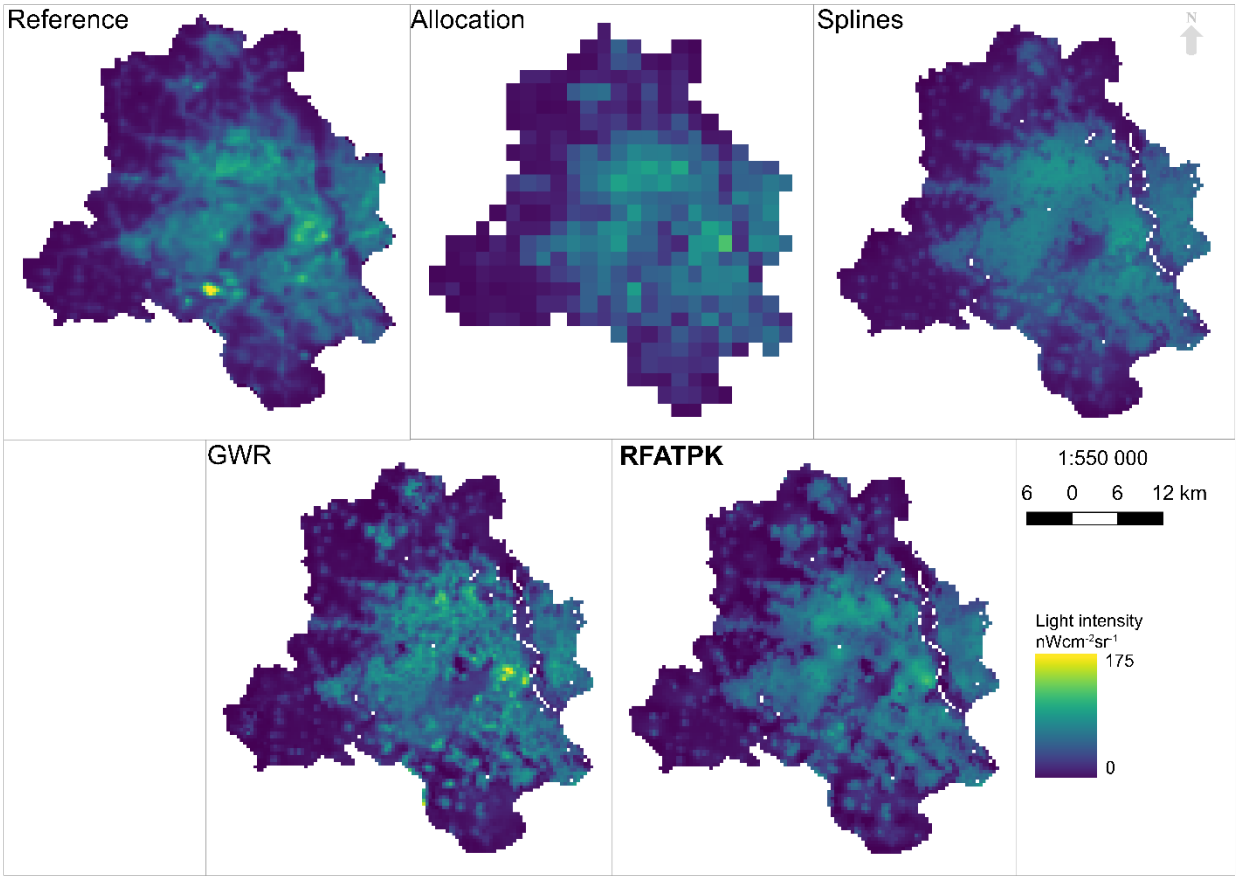

Figure 4: Downscaling results of NTL at 450 m for New Delhi, 2018. From left to right, the Reference NTL, Allocation-based downscaled NTL, Machine Learning with Splines-based downscaled NTL, GWR-based and RFATPK-based downscaled NTL. Bold shows the proposed method.

Table 1: Quantitative comparison of the downscaling approaches at 450 m (reference is the original NTL) for Mumbai. The best performance is highlighted in **bold**.

		2013	2014	2015	2016	2017	2018	2019	2020
RMSE	Allocation	10.1172	14.2775	10.9951	11.7604	11.8663	12.9181	15.5305	16.1761
	ML with splines	10.3092	16.5331	10.5267	11.0395	11.3876	12.166	15.4108	15.8129
	GWR	11.3387	17.8948	11.9249	12.9232	13.3665	14.0287	17.8138	17.9765
	RFATPK	**1.7165**	**2.7673**	**2.0354**	**2.1364**	**2.7848**	**2.498**	**4.1899**	**3.4906**
MSE	Allocation	102.359	203.848	120.893	138.307	140.81	166.877	241.198	261.668
	ML with splines	106.279	273.343	110.811	121.870	129.678	148.013	237.494	250.049
	GWR	128.567	320.224	142.205	167.009	178.664	196.804	317.332	323.157
	RFATPK	**2.9464**	**7.6583**	**4.1429**	**4.5645**	**7.7555**	**6.2403**	**17.556**	**12.1843**

CC	Allocation	0.8187	0.7839	0.7893	0.7918	0.7817	0.7818	0.761	0.7254
	ML with splines	0.8041	0.6724	0.8005	0.8116	0.7911	0.8002	0.7515	0.7224
	GWR	0.7566	0.5985	0.7342	0.7295	0.6961	0.7224	0.647	0.6181
	RFATPK	**0.995**	**0.9923**	**0.9932**	**0.9935**	**0.9887**	**0.9923**	**0.9838**	**0.9884**

Table 2: Quantitative comparison of the downscaling approaches at 450m (reference is the original NTL) for New Delhi. The best performance is highlighted in **bold**.

		2013	2014	2015	2016	2017	2018	2019	2020
RMSE	Allocation	9.4589	9.0099	9.1661	8.1967	8.5255	8.3571	8.4641	7.5097
	ML with splines	9.0594	9.1414	8.0654	7.8573	7.71	7.5779	8.5324	6.5307
	GWR	9.8107	10.057	10.7049	8.9059	9.6294	8.9825	9.3896	8.5971
	RFATPK	**2.5113**	**2.2719**	**2.5727**	**2.1916**	**2.3779**	**2.1511**	**1.8615**	**2.0369**
MSE	Allocation	89.471	81.1797	84.0189	67.1861	72.6850	69.8423	71.6414	56.3963
	ML with splines	82.0733	83.5664	65.051	61.7373	59.4446	57.4259	72.8033	42.6505
	GWR	96.2505	101.154	114.596	79.3157	92.7269	80.6858	88.165	73.9109
	RFATPK	**6.3068**	**5.1617**	**6.6191**	**4.8033**	**5.6546**	**4.6276**	**3.4655**	**4.1491**
CC	Allocation	0.9157	0.9234	0.9189	0.9165	0.9094	0.9068	0.8953	0.9091
	ML with splines	0.9233	0.9214	0.9381	0.9241	0.927	0.9245	0.8941	0.9325
	GWR	0.9095	0.9041	0.8882	0.9014	0.8832	0.8922	0.8702	0.8799
	RFATPK	**0.9943**	**0.9953**	**0.9938**	**0.9943**	**0.9932**	**0.9941**	**0.9952**	**0.9936**

Tables 1 and 2 provide a quantitative comparison of the downscaling methods using three indices: Root Mean Square Error (RMSE), Mean Square Error (MSE) and Correlation Coefficient (CC). RFATPK is clearly more precise than the three benchmark methods in terms of all three indices. This is due to the fact that the sceneries under study are highly developed metropolitan environments with a variety of impervious surfaces (such as buildings, roads, and vegetation), which are better suited to being well described by a spatially non-stationary model. Machine Learning with Splines yielded greater accuracy compared to GWR and Allocation-based downscaling. The least accuracy resulted for GWR-based downscaling, in terms of all three indices.

4.1.2. Downscaling prediction (450 m to 100 m)

To facilitate visual comparison, three zoomed sub-areas selected randomly and their corresponding results are shown in Figure 6 Figure 7 for Mumbai and New Delhi, respectively. The sub-areas include landscapes with a mix of dense and less dense urban structures. The disaggregating findings demonstrate that RFATPK renders those landscapes well. Due to poor prediction, the GWR approach distorts the borders, whereas Machine Learning with Splines excessively smooths the boundaries. When it comes to preserving spectral characteristics and recovering both dense and less dense textures, RFATPK performs satisfactorily.

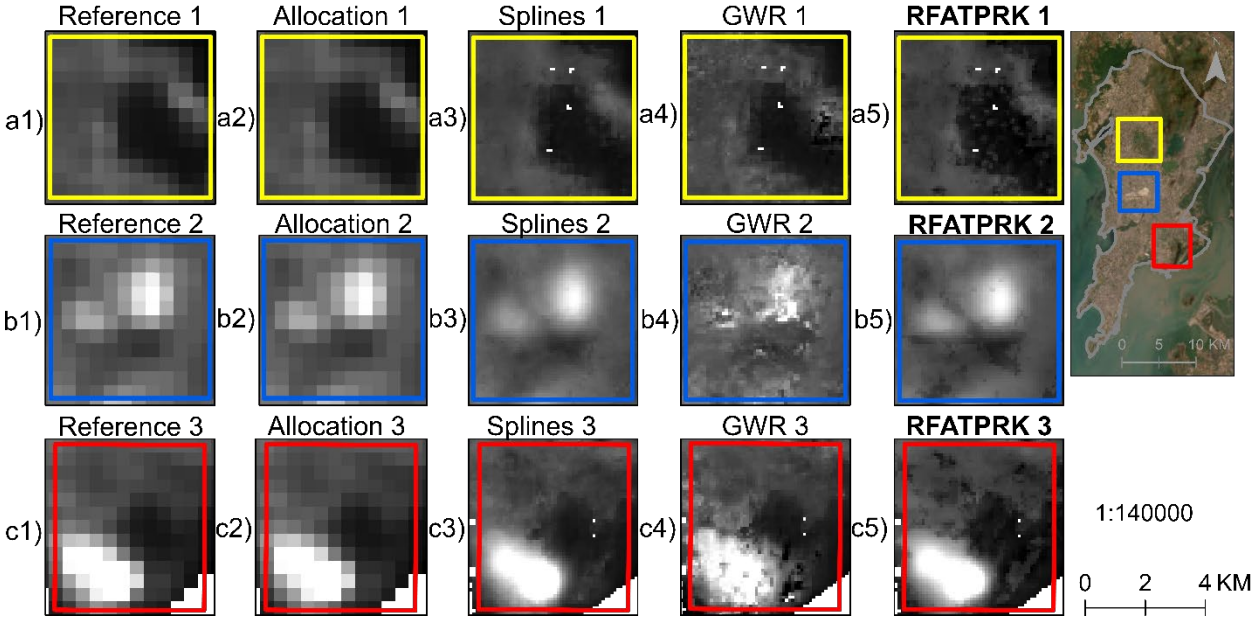

Figure 5: Downscaling visual results (100 m) in three sub-areas for Mumbai, 2018. From left to right: Raw NTL, Allocation-based, Machine Learning with Splines, GWR, **RFATPK**. Each column illustrates one of the three selected random areas. **Bold** shows the proposed method.

The variograms were compared for the different downscaling methods. Due to large volume of images produced regarding the comparison of the variograms for all the years, a single example for the year 2018 is shown for every region.

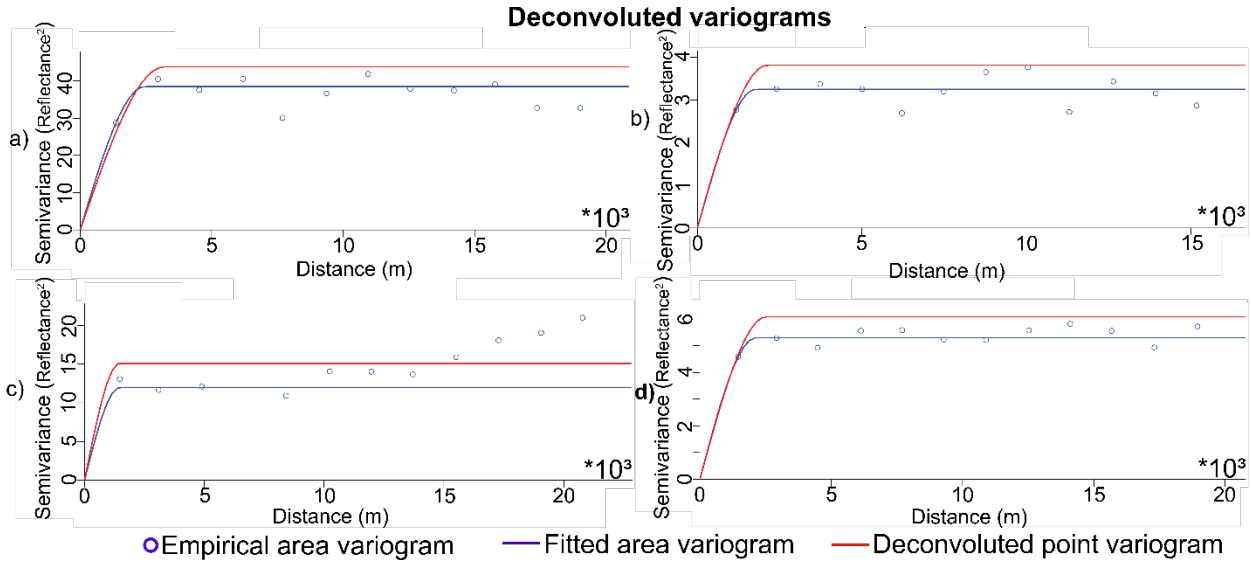

Figure 6: Variograms of the downscaling methods for Mumbai, 2018. a) Allocation-based prediction, b) Machine Learning with Splines-based prediction, c) GWR-based prediction and **d) RFATPK**-based prediction. **Bold** shows the proposed method.

For the calculation of the variograms, the downscaled images were upscaled to 450 m, subtracted from the reference and the variograms computed between the downscaled image at 100 m and the subtraction. Compared to the other approaches, the Allocation-based downscaling method's variogram exhibited the highest semivariance (Figure 7). The GWR-based downscaling approach in Figure 9 provided the highest semivariance, while the Machine Learning with Splines-based downscaling method produced the lowest semivariance.

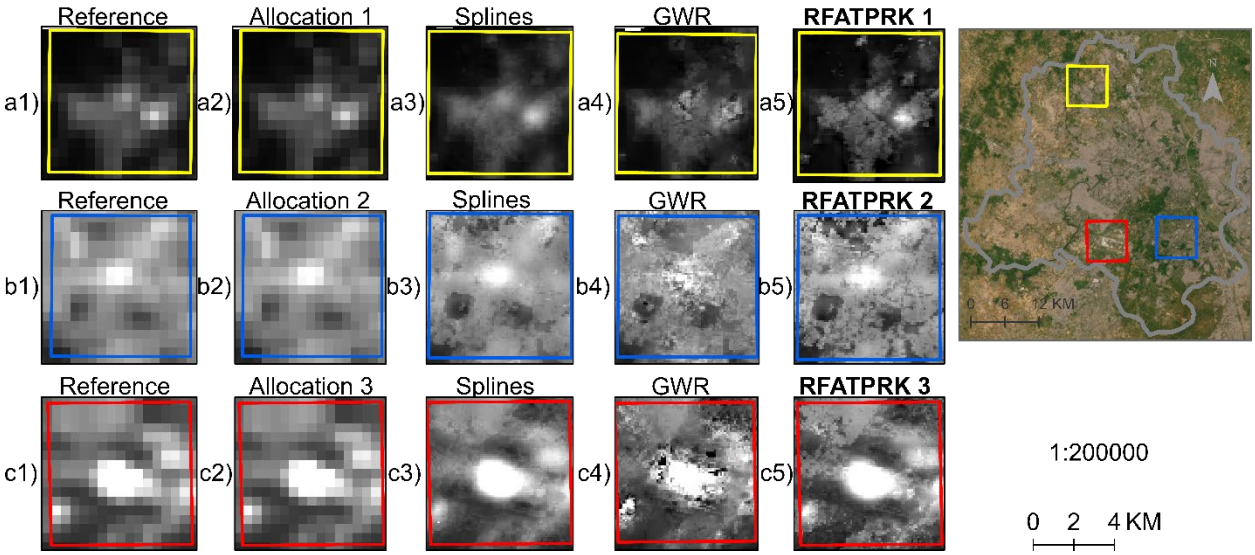

Figure 7: Downscaling visual results (100 m) in three sub-areas for New Delhi, 2018. From left to right: Raw NTL, Allocation-based, Machine Learning with Splines, GWR, **RFATPK**. Each column illustrates one of the three selected random areas. **Bold** shows the proposed method.

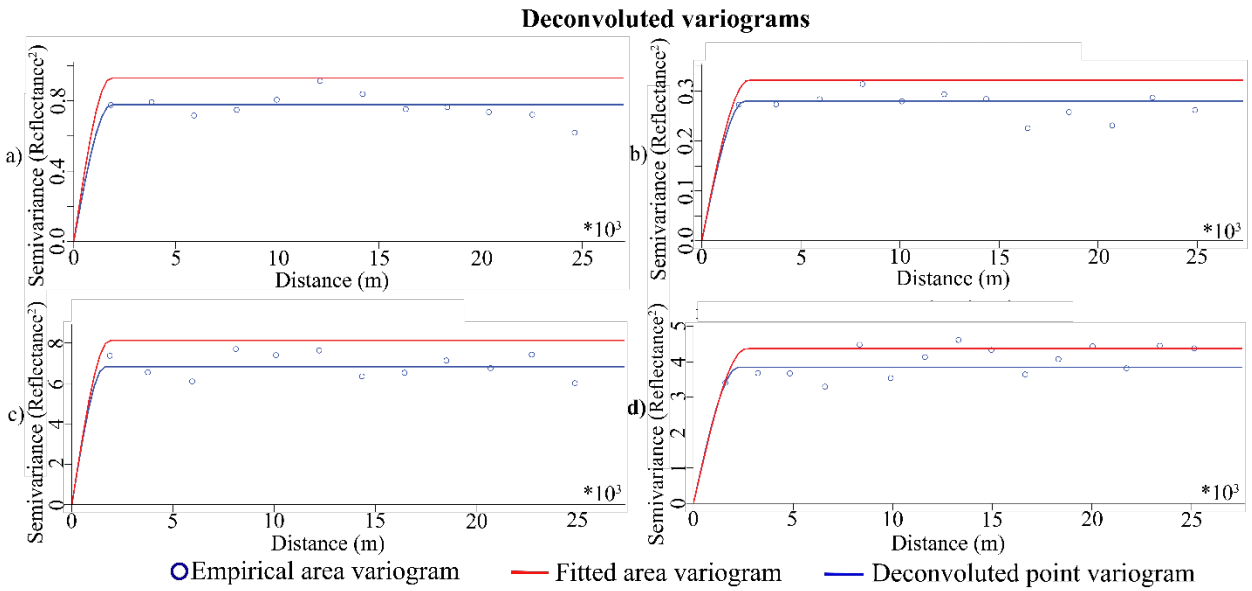

Figure 8: Variograms of the downscaling methods for New Delhi, 2018. a) Allocation-based prediction, b) Machine Learning with Splines-based prediction, c) GWR-based prediction and **d) RFATPK**-based prediction. **Bold** shows the proposed method.

In comparison to GWR and Machine Learning with Splines, the suggested RFATPK generated the best visual outcome among the three downscaling techniques and also had the attribute of perfect coherence (Table 3). Allocation-based downscaling also preserves the property of perfect coherence, but no new information is added and there is, consequently, no spatial variability in the NTL intensity within the fine resolution pixels.

Table 3: Measurement of perfect coherence, of the four downscaling methods for Mumbai and New Delhi for all years under investigation. **Bold** shows the largest results in terms of coherence.

		2013	2014	2015	2016	2017	2018	2019	2020
Mumbai	Allocation	**0.9957**	0.9907	**0.9958**	**0.9958**	0.995	**0.9954**	**0.9949**	0.9927
	ML with splines	0.9916	0.9869	0.9919	0.9917	0.9922	0.9925	0.9921	0.9897
	GWR	0.9756	0.9582	0.9797	0.9798	0.9784	0.9731	0.9744	0.9636
	RFATPK	0.9950	**0.9923**	0.9932	0.9935	**0.9987**	0.9923	0.9938	**0.9984**
New Delhi	Allocation	**0.998**	**0.9994**	**0.998**	**0.9979**	**0.9978**	**0.9979**	**0.9978**	**0.9979**
	ML with splines	0.9967	0.9967	0.9949	0.9964	0.9954	0.9948	0.9966	0.9946
	GWR	0.9943	0.9947	0.9877	0.9931	0.9857	0.9879	0.9856	0.9872
	RFATPK	0.9943	0.9953	0.9938	0.9943	0.9932	0.9941	0.9952	0.9936

According to Table 3, the property of perfect coherence was achieved in all years for both regions only for the RFATPK and Allocation-based methods. RFATPK had the highest Correlation Coefficient (CC) index for the years 2014, 2017 and 2020, while the Allocation-based method produced the maximum for the rest of the years, for Mumbai. For New Delhi, the Allocation

method had the highest CC value for all years. In summary, the proposed method was the only one to achieve perfect coherence consistently throughout the years across all regions, despite the fact ML with Splines had higher values of CC for New Delhi compared to RFATPK. ML with Splines was inconsistent in achieving perfect coherence across the regions and for all years as it can be seen for the year 2014 for Mumbai.

Tables 4 and 5 show the quantitative comparison of each method with the LuoJia 1-01 used as a reference.

Table 4: Quantitative comparison of the downscaling approaches at 100 m for Mumbai, 2018 (reference is the LuoJia 1-01). **Bold** shows the best results.

		Allocation	ML with splines	GWR	**RFATPK**
Mean	RMSE	13.9105	16.8635	15.7574	**13.8938**
	MSE	196.1515	286.4333	248.6675	**193.2563**
	PCC	0.6056	0.5274	0.5408	**0.6757**
StD	RMSE	1.8192	1.6027	0.6798	**0.5219**
	MSE	50.9233	54.305	21.4063	**14.7106**
	CC	0.0893	0.1044	0.066	**0.0204**

Table 5: Quantitative comparison of the downscaling approaches at 100 m for New Delhi, 2018 (reference is the LuoJia 1-01). **Bold** shows the best results.

		Allocation	ML with splines	GWR	**RFATPK**
Mean	RMSE	7.4432	7.711	8.5856	**6.7488**
	MSE	55.4679	59.4932	73.790556	**45.889**
	PCC	0.9268	0.9225	0.9062	**0.9392**
StD	RMSE	0.2868	**0.2044**	0.31086	0.6541
	MSE	4.2883	3.1429	5.401	8.8968
	CC	0.0045	0.0048	**0.0045**	0.0068

It can be seen in Tables 4 and 5 that the proposed method produced the smallest mean RMSE and mean MSE in each of the 50 iterations (2000 random samples in each iteration) as well as the largest mean correlation coefficient with LouJia1-01. More specifically, the RFATPK method, for Mumbai, produced the smallest average RMSE of 13.8938 $nWcm^{-2}sr^{-1}$, while the other methods produced larger RMSEs, with the Machine Learning with Splines being the largest. The same is true for MSE, where RFATPK produced the smallest average MSE compared to the other three methods. Finally, the correlation coefficient was the largest for RFATPK, at 0.6757 for Mumbai. For New Delhi, again RFATPK produced the most precise predictions compared to the other

approaches with mean RMSE, MSE and CC of 6.7488 $nWcm^{-2}sr^{-1}$, 45.889 $nWcm^{-2}sr^{-1}$ and 0.9392, respectively. Considering that the LouJia1-01 imagery is an external (unseen) validation dataset produced with different sensing characteristics than the NTL imagery and that the prediction is made at a four point fine-fold finer spatial resolution than the original data, this result can be considered promising for a range of applications.

4.2. Night Light Development Index

The index was calculated utilizing the coarse-resolution nocturnal data ($NLDI_{coarse}$) and the disaggregated NTL ($NLDI_{fine}$). To validate the results, the values of the Human Development Index for New Niludipine Delhi for the years 2013 to 2019 were acquired from the Global Data Lab website (https://globaldatalab.org/) (accessed 05/02/2023). The relationship between the NLDI and HDI is well established in the literature. Hence, this index was selected for validation (Elvidge et al., 2012). Using the $NLDI_{coarse}$ the Pearson's correlation coefficient with the HDI was -0.35 while when using the $NLDI_{fine}$ the association with the HDI was -0.48. These results are impressive, indicating that the downscaled NTL data were more correlated with human development compared to the raw NTL data at 450 m.

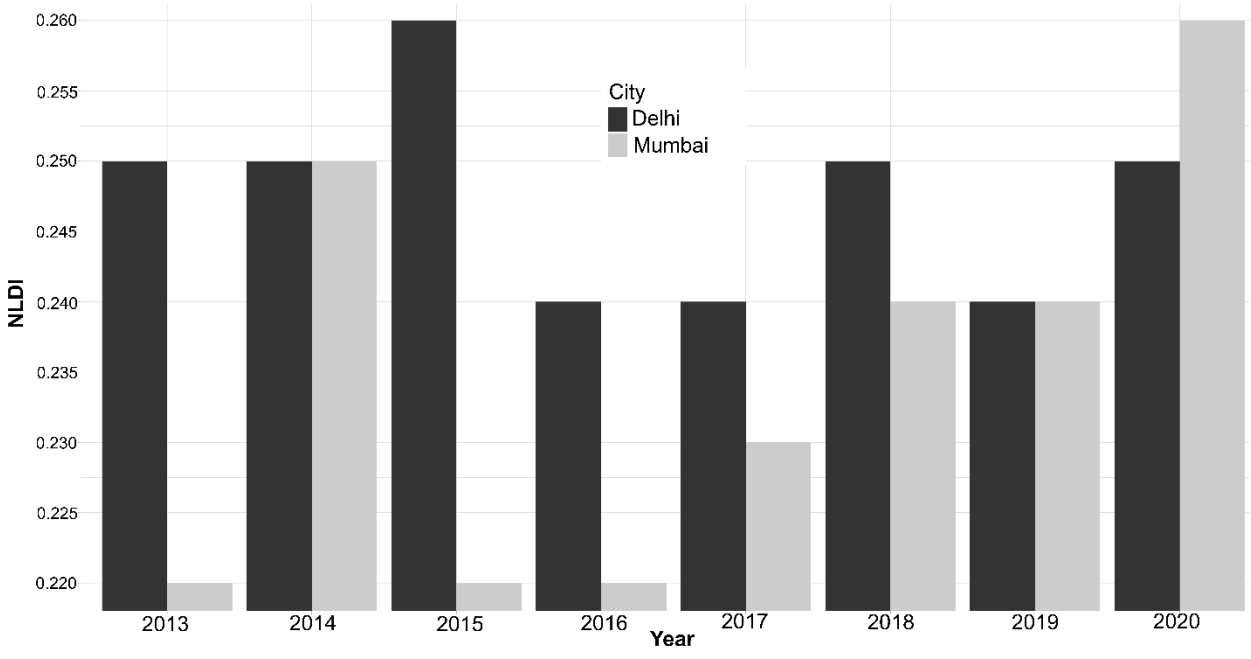

Figure 9: Night Light Development Index plotted through time for Mumbai (red line) and New Niludipine Delhi (green line).

Figure 10 illustrates the evolution of the NLDI for the period 2013-2020 for the cities of Mumbai and New Niludipine Delhi. It can be seen that the index for Mumbai shows an upward trend which means that light inequality is increasing. On the contrary, the NLDI index for New Niludipine Delhi is decreasing with the exception for 2015. The data are consistent with the HDI index from Global Data Lab which reveals an increase in the index through time (Table 6).

586

587 Table 6: The HDI and NLDI indexes for New Delhi.

Year	HDI	NLDI
2013	0.72	0.25
2014	0.73	0.25
2015	0.73	0.26
2016	0.73	0.24
2017	0.74	0.24
2018	0.74	0.25
2019	0.75	0.24

588 According to many authors, there is an inverse relationship between the two indices (Elvidge et
589 al., 2012; Ghosh et al., 2013). There are no yearly measures of HDI for Mumbai to validate the
590 results, but many newspapers highlighted the rise in inequality in this region (Bendix, 2018; Ashar,
591 2019).

592 **4.3. Gross National Income *per capita***

593 The GNI was measured only for New Niludipine Delhi for the years 2013 to 2019. It can be seen
594 from Table 7 that there was a slight increase in the index from 2013 to 2015, then a small decrease
595 for 2015 and then the index increased again. The second and third columns show the Sum of Lights
596 (SOL) for the coarse resolution and fine resolution NTL. The GNI index is measured in 1000 US
597 dollars and the values are in logarithmic scale as well the SOL. The coarse and fine SOL showed
598 overall a slight decrease in grey values, but the downward is more obvious in the coarse SOL.

599 Table 7: Yearly GNI values and the Sum of Lights values using the coarse and fine spatial scale
600 NTL.

Year	Coarse SOL	Fine SOL	GNI
2013	5.334	6.625	9.345
2014	5.34	6.635	9.378
2015	5.346	6.623	9.528
2016	5.299	6.602	9.449
2017	5.3	6.594	9.522
2018	5.289	6.563	9.576
2019	5.272	6.57	9.618

601 The correlation coefficient (R^2) between Fine SOL and GNI is much higher compared to Coarse
602 SOL and GNI. Specifically, Fine SOL predict 67.6% of the variance in annual GNI, compared to
603 47.4% that Coarse SOL does. The model fit utilizing the fine spatial resolution NTL is noticeably
604 better by about 20% contrasted to coarse resolution NTL.

605 **5. Discussion**

606 **5.1. Random Forest area-to-point regression Kriging (RFATPK)**

607 RFATPK is analogous to the AATPRK reported by Wang et al., (2016a). However, a different
608 non-stationary model was implemented in this study. The results show a notable improvement of

the merged images both visually and quantitatively resulting from the adoption of the spatial non-stationary regression model, reflected in a correlation coefficient larger than 0.84 in the regression part for all months and cities, with the exception of Mumbai 2020. In RFATPK, residual sharpening was conducted by ATPK which considered a global method and it is different from the approach proposed by Pardo-Iguzquiza et al. (2011), who developed a local scheme for Kriging interpolation. For each coarse pixel in local ATPK interpolation, semivariogram deconvolution is used to parameterize the regression model and the Kriging weights are calculated. This requires a lot of computational power, especially for areas with many pixels. We instead used global ATPK, which does not require the same computational cost. Since RFATPK is an extension of ATPRK, it benefits from ATPRK's advantages as it takes into consideration both the size of the support and the spatial correlation. Additionally, it can accurately maintain the spectral characteristics of the original coarse data, as illustrated in Tables 4 and 5. The experimental results showed that RFATPK predicted more accurately than the three benchmark techniques. The resulting residuals may differ significantly from region to region if the global regression model in ATPRK is unable to represent adequately the relationship between the coarse and fine images when the observed scene varies locally (i.e., requiring a spatially non-stationary method). Thus, the residuals produced by the local non-stationary regression model are likely to be more suited for subsequent manipulation using a stationary downscaling model. Another point of consideration is the scale effect. It is acknowledged that due to the scale effect there exists differences in the two downscaling processes, that is, from 2025 m to 450 m and from 450 m to 100 m spatial resolution (Zhou et al., 2016; Pu, 2021). The scaling effect in downscaling NTL from a coarse spatial resolution to a high spatial resolution is beyond the scope of this research.

5.2. RFATPK and benchmark comparison

When the area of interest is spatially heterogeneous, the RFATARK technique yields more precise predictions. Comparing the proposed RFATPK method against three other image fusion techniques (GWR from regression-based methods, machine learning with splines from hybrid-based techniques and allocation of raster value) is of great interest. Since all computations are undertaken independently in each coarse band, RFATPK is substantially faster than the hybrid technique since it needs to model only the auto-semivariogram for each low resolution image. The Kriging system in Equation 6's matrices now have substantially smaller sizes. Consequently, RFATPK is more user-friendly and much simpler to automate. Although the regression-based method is similar to RFATPK, the latter has the appealing property of perfect coherence, which is inherent with ATPK. Compared to the simple allocation of raster values, the geostatistical solution preserved fairly accurate, both visually and quantitatively, the spatial patterns of NTL intensity, a property which simple allocation of raster values does not have. In conclusion, the suggested RFATPK method has the following features and benefits.

1. Regression modelling can employ fines-resolution predictors to lessen the uncertainty in spatial downscaling, improving the fused images visually and quantitatively;
2. RFATPK clearly takes into consideration the spatial (auto) correlation between the data and the size of the support (pixel) by inheriting the advantages of ATPRK;
3. RFATPK, can precisely preserve the spectral features;

4. Unlike machine learning with splines, which executes six regression models, RFATPK executes only one local model and incorporates a global method (ATPK), calculates the Kriging weights only one time for the entire region and uses considerably smaller matrices in the Kriging method; this makes it faster for downscaling images.

5.3. Use case studies

Monitoring socioeconomic indicators at the city-scale is of great importance for governments and policy makers. As such unbiased data at fine spatial resolution are a critical input to support policy development and decision-making. To highlight the applicability of the downscaling method, two socioeconomic applications were considered at the city-scale.

5.3.1. Night Light Development Index

The index is an estimation for economic and human development in a region. The strong association between the NLDI and the HDI suggests the former index measures human development, which is consistent with Elvidge et al. (2012). The results using the fused NTL are encouraging and we suggest the downscaled data are suitable for measuring human development at the city-scale.

5.3.2. Gross National Income *per capita*

The reference NTL data were less accurate at predicting yearly GNI than the downscaled NTL at the city scale. The application of studies that demonstrate the efficiency of estimating such socioeconomic indicators at the city level is called into question by the poor association between coarse resolution nocturnal data and GNI and makes it difficult to understand how such data may serve as a reliable indicator of changes in city-scale economic activity. The results provided here, on the other hand, point to the downscaled NTL as a far more accurate way to quantify GNI and a viable substitute for the index.

5.4. Future research

The point spread function (PSF) exists in every satellite sensor imagery. It has a significant impact on image quality and sets a strict cap on how much information is included in satellite sensor images (Wang et al., 2020). It is clear that the PSF can affect the downscaling process because disaggregating methods aim to increase the pixel size by creating more (sub-) pixels than the original image and thus, better describing the spatial content of a region. A variety of PSFs will be evaluated in a future effort to reduce the uncertainty in the downscaling procedure. Another important limitation is that NTL values cannot be determined from a single covariate, as shown by the global model. This means that, more ancillary variables are more suitable for charactering NTL intensity and may lead to more accurate prediction of the trend (Ye et al., 2021). Future research will focus on incorporating more ancillary variables, mainly from the so called ‘social pixels’, for example, geo-tagged Twitter data or geo-located POIs. Thus, fusion with social data at a fine resolution should be tested in future. Lastly, as mentioned in the Discussion, Section 5.1, this research did not take into account the scale effect. Therefore, studies in the future should need to be designed for and check if the by accounting for the scale effect will improve the downscaling predictions.

6. Conclusion

Spatial downscaling is widely used to transform remotely sensed images from coarse resolution to fine resolution in order to track human activity. For the first time, a strategy for spatially downscaling nocturnal pictures was presented in this study using RF and ATPK. The RFATPK approach has the advantage of taking into consideration both the spatial correlation between the response variable and the predictors as well as local spatial variation. To show the effectiveness of this approach, it was used on yearly coarse NTL products in two separate Indian megacities.

The geostatistical RFATPK solution was compared against three benchmark algorithms in experiments conducted on one experimental case in the two mega-cities. The results are summarized as follows: 1) The three benchmark methods were outperformed by RFATPK, demonstrating the utility of this technique for spatial sharpening; 2) RFATPK, consistently, assures total coherence with the original coarse data, in contrast to two of the benchmarks, and 3) due to its spatially non-stationary nature, RFATPK was able to lower the residual variance in comparison to a single, global regression model. The encouraging results suggest that RFATPK can produce images that are suitable for socioeconomic analysis at the city-scale, as illustrated when comparing a human development index using coarse-resolution NTL data against fine-resolution nocturnal lights. Indeed, the GNI index was better approximated using the downscaled NTL data. Another application suggesting that the disaggregated NTL are more suitable for fine scale (social) applications was the measurement of wellbeing by means of light inequality. The results implied that using the proposed solution, the nocturnal satellite sensor data are closer to the values of the official statistics (i.e., HDI). According to the results, our method can be generalized worldwide (i.e., to other cities) and for a variety of social science applications.

7. Acknowledgement

This work was funded by the Natural Environment Research Council [grant number NE/T004002/1].

CRediT authorship contribution statement

Tziokas, Nikolaos: Conceptualization, Methodology, Validation, Formal analysis, Visualization, Writing-original draft, Editing. **Drolias, C. Garyfallos**: Methodology, Formal analysis. **Zhang, Ce**: Supervision, Writing-review & editing. **Atkinson, M. Peter**: Supervision, Writing-review & editing.

References

Ashar, S., 2019, November 19, Mumbai is India's city with the greatest inequalities, DNA India (accessed 4 July 2022)

Atkinson, P.M., 2013. Downscaling in remote sensing. Int. J. Appl. Earth Obs. Geoinf. 22, 106–114. https://doi.org/10.1016/j.jag.2012.04.012

Bhanarkar, A.D., Purohit, P., Rafaj, P., Amann, M., Bertok, I., Cofala, J., Rao, P.S., Vardhan,

725 B.H., Kiesewetter, G., Sander, R., Schöpp, W., Majumdar, D., Srivastava, A., Deshmukh,
726 S., Kawarti, A., Kumar, R., 2018. Managing future air quality in megacities: Co-benefit
727 assessment for Delhi. Atmos. Environ. 186, 158–177.
728 https://doi.org/10.1016/j.atmosenv.2018.05.026

729 Breiman, L., 2001. Random Forests. Mach. Learn. 45, 5–32.
730 https://doi.org/https://doi.org/10.1023/A:1010933404324

731 Brokamp, C., Jandarov, R., Rao, M.B., LeMasters, G., Ryan, P., 2017. Exposure assessment
732 models for elemental components of particulate matter in an urban environment: A
733 comparison of regression and random forest approaches. Atmos. Environ. 151, 1–11.
734 https://doi.org/10.1016/j.atmosenv.2016.11.066

735 Bruederle, A., Hodler, R., 2018. Nighttime lights as a proxy for human development at the local
736 level. PLoS One 13, 1–22. https://doi.org/10.1371/journal.pone.0202231

737 Bullock, E.L., Woodcock, C.E., Holden, C.E., 2020. Improved change monitoring using an
738 ensemble of time series algorithms. Remote Sens. Environ. 238, 1–53.
739 https://doi.org/10.1016/j.rse.2019.04.018

740 Chen, C., Zhao, S., Duan, Z., Qin, Z., 2015. An Improved Spatial Downscaling Procedure for
741 TRMM 3B43 Precipitation Product Using Geographically Weighted Regression. IEEE J.
742 Sel. Top. Appl. Earth Obs. Remote Sens. 8, 4592–4604.
743 https://doi.org/10.1109/JSTARS.2015.2441734

744 Cheng, Z., Wang, J., Ge, Y., 2022. Mapping monthly population distribution and variation at 1-
745 km resolution across China. Int. J. Geogr. Inf. Sci. 36, 1166–1184.
746 https://doi.org/10.1080/13658816.2020.1854767

747 Elvidge, C.D., Baugh, K., Zhizhin, M., Hsu, F.C., Ghosh, T., 2017. VIIRS night-time lights. Int.
748 J. Remote Sens. 38, 5860–5879. https://doi.org/10.1080/01431161.2017.1342050

749 Elvidge, C.D., Baugh, K.E., Anderson, S.J., Sutton, P.C., Ghosh, T., 2012. The Night Light
750 Development Index (NLDI): A spatially explicit measure of human development from
751 satellite data. Soc. Geogr. 7, 23–35. https://doi.org/10.5194/sg-7-23-2012

752 Elvidge, C.D., Cinzano, P., Pettit, D.R., Arvesen, J., Sutton, P., Small, C., Nemani, R., Longcore,
753 T., Rich, C., Safran, J., Weeks, J., Ebener, S., 2007. The nightsat mission concept. Int. J.
754 Remote Sens. 28, 2645–2670. https://doi.org/10.1080/01431160600981525

755 Ge, Y., Jin, Y., Stein, A., Chen, Y., Wang, Jianghao, Wang, Jinfeng, Cheng, Q., Bai, H., Liu, M.,
756 Atkinson, P.M., 2019. Principles and methods of scaling geospatial Earth science data.
757 Earth-Science Rev. 197, 102897. https://doi.org/10.1016/j.earscirev.2019.102897

758 Getis, A., Ord, J.K., 1992. The Analysis of Spatial Association by Use of Distance Statistics.
759 Geogr. Anal. 24, 189–206. https://doi.org/10.1111/j.1538-4632.1992.tb00261.x

760 Ghosh, T., Anderson, S.J., Elvidge, C.D., Sutton, P.C., 2013. Using nighttime satellite imagery
761 as a proxy measure of human well-being. Sustain. 5, 4988–5019.
762 https://doi.org/10.3390/su5124988

763 Ghosh, T., Sutton, P., Powell, R., Anderson, S., Elvidge, C.D., 2009. Estimation of Mexico's

764 informal economy using DMSP nighttime lights data. 2009 Jt. Urban Remote Sens. Event.
765 https://doi.org/10.1109/URS.2009.5137751

766 Gibson, J., Boe-Gibson, G., 2021. Nighttime lights and county-level economic activity in the
767 United States: 2001 to 2019. Remote Sens. 13. https://doi.org/10.3390/rs13142741

768 Gibson, J., Olivia, S., Boe-Gibson, G., Li, C., 2021. Which night lights data should we use in
769 economics, and where? J. Dev. Econ. 149, 102602.
770 https://doi.org/10.1016/j.jdeveco.2020.102602

771 Head, A., Manguin, M., Tran, N., Blumenstock, J.E., 2017. Can Human Development be
772 Measured with Satellite Imagery? 1–11. https://doi.org/10.1145/3136560.3136576

773 Jasiński, T., 2019. Modeling electricity consumption using nighttime light images and artificial
774 neural networks. Energy 179, 831–842. https://doi.org/10.1016/j.energy.2019.04.221

775 Jean, N., Burke, M., Xie, M., Davis, W.M., Lobell, D.B., Ermon, S., 2016. Combining satellite
776 imagery and machine learning to predict poverty. Science (80-.). 353, 790–794.
777 https://doi.org/10.1126/science.aaf7894

778 Jin, Y., Ge, Y., Wang, J., Heuvelink, G.B.M., Wang, L., 2018. Geographically weighted area-to-
779 point regression kriging for spatial downscaling in remote sensing. Remote Sens. 10, 1–22.
780 https://doi.org/10.3390/rs10040579

781 Kerry, R., Goovaerts, P., Rawlins, B.G., Marchant, B.P., 2012. Disaggregation of legacy soil
782 data using area to point kriging for mapping soil organic carbon at the regional scale.
783 Geoderma 170, 347–358. https://doi.org/10.1016/j.geoderma.2011.10.007

784 Kyriakidis, P.C., 2004. A geostatistical framework for area-to-point spatial interpolation. Geogr.
785 Anal. 36, 259–289. https://doi.org/10.1111/j.1538-4632.2004.tb01135.x

786 Kyriakidis, P.C., Yoo, E.H., 2005. Geostatistical prediction and simulation of point values from
787 areal data. Geogr. Anal. 37, 124–151. https://doi.org/10.1111/j.1538-4632.2005.00633.x

788 Levin, N., Johansen, K., Hacker, J.M., Phinn, S., 2014. A new source for high spatial resolution
789 night time images - The EROS-B commercial satellite. Remote Sens. Environ. 149, 1–12.
790 https://doi.org/10.1016/j.rse.2014.03.019

791 Levin, N., Kyba, C.C.M., Zhang, Q., Sánchez de Miguel, A., Román, M.O., Li, X., Portnov,
792 B.A., Molthan, A.L., Jechow, A., Miller, S.D., Wang, Z., Shrestha, R.M., Elvidge, C.D.,
793 2020. Remote sensing of night lights: A review and an outlook for the future. Remote Sens.
794 Environ. 237. https://doi.org/10.1016/j.rse.2019.111443

795 Liu, L., Zhou, H., Lan, M., Wang, Z., 2020. Linking Luojia 1-01 nightlight imagery to urban
796 crime. Appl. Geogr. 125, 102267. https://doi.org/10.1016/j.apgeog.2020.102267

797 Malik, K., Kumar, D., Perissin, D., Pradhan, B., 2022. Estimation of ground subsidence of New
798 Delhi, India using PS-InSAR technique and multi-sensor radar data. Adv. Sp. Res. 69,
799 1863–1882. https://doi.org/10.1016/j.asr.2021.08.032

800 Marlier, M.E., Jina, A.S., Kinney, P.L., DeFries, R.S., 2016. Extreme Air Pollution in Global
801 Megacities. Curr. Clim. Chang. Reports 2, 15–27. https://doi.org/10.1007/s40641-016-

0032-z

Nijman, J., 2010. A study of space in Mumbai's slums. Tijdschr. voor Econ. en Soc. Geogr. 101, 4–17. https://doi.org/10.1111/j.1467-9663.2009.00576.x

Pardo-Iguzquiza, E., Rodríguez-Galiano, V.F., Chica-Olmo, M., Atkinson, P.M., 2011. Image fusion by spatially adaptive filtering using downscaling cokriging. ISPRS J. Photogramm. Remote Sens. 66, 337–346. https://doi.org/10.1016/j.isprsjprs.2011.01.001

Park, N., Kim, Y., Kwak, G.-H., 2019. An overview of theoretical and practical issues in spatial downscaling of coarse resolution satellite-derived products. Korean J. Remote Sens. 35, 589–607.

Pesaresi, M., Politis, P., 2022. GHS-BUILT-S R2022A - GHS built-up surface grid , derived from Sentinel2 composite and Landsat , multitemporal (1975-. Eur. Comm. Jt. Res. Cent. https://doi.org/10.2905/D07D81B4-7680-4D28-B896-583745C27085

Pu, R., 2021. Assessing scaling effect in downscaling land surface temperature in a heterogenous urban environment. Int. J. Appl. Earth Obs. Geoinf. 96, 102256. https://doi.org/10.1016/j.jag.2020.102256

Takoutsing, B., Heuvelink, G.B.M., 2022. Comparing the prediction performance, uncertainty quantification and extrapolation potential of regression kriging and random forest while accounting for soil measurement errors. Geoderma 428, 116192. https://doi.org/10.1016/j.geoderma.2022.116192

Tang, K., Zhu, H., Ni, P., 2021. Spatial downscaling of land surface temperature over heterogeneous regions using random forest regression considering spatial features. Remote Sens. 13. https://doi.org/10.3390/rs13183645

Vinayak, B., Lee, H.S., Gedem, S., 2021. Prediction of land use and land cover changes in Mumbai city, India, using remote sensing data and a multilayer perceptron neural network-based markov chain model. Sustainability 13, 471. https://doi.org/10.3390/su13020471

Wang, Q., Shi, W., Atkinson, P.M., 2016a. Area-to-point regression kriging for pan-sharpening. ISPRS J. Photogramm. Remote Sens. 114, 151–165. https://doi.org/10.1016/j.isprsjprs.2016.02.006

Wang, Q., Shi, W., Atkinson, P.M., Pardo-Igúzquiza, E., 2016b. A new geostatistical solution to remote sensing image downscaling. IEEE Trans. Geosci. Remote Sens. 54, 386–396. https://doi.org/10.1109/TGRS.2015.2457672

Wang, Q., Shi, W., Atkinson, P.M., Zhao, Y., 2015. Downscaling MODIS images with area-to-point regression kriging. Remote Sens. Environ. 166, 191–204. https://doi.org/10.1016/j.rse.2015.06.003

Wang, Q., Shi, W., Li, Z., Atkinson, P.M., 2016c. Fusion of Sentinel-2 images. Remote Sens. Environ. 187, 241–252. https://doi.org/10.1016/j.rse.2016.10.030

Wang, Q., Tang, Y., Atkinson, P.M., 2020. The effect of the point spread function on downscaling continua. ISPRS J. Photogramm. Remote Sens. 168, 251–267. https://doi.org/10.1016/j.isprsjprs.2020.08.016

Wang, Q., Zhang, Y., Onojeghuo, A.O., Zhu, X., Atkinson, P.M., 2017. Enhancing spatio-temporal fusion of MODIS and Landsat data by incorporating 250 m MODIS data. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 10, 4116–4123. https://doi.org/10.1109/JSTARS.2017.2701643

Xu, J., Zhang, F., Jiang, H., Hu, H., Zhong, K., Jing, W., Yang, J., Jia, B., 2020. Downscaling Aster Land Surface Temperature over Urban Areas with Machine Learning-Based Area-To-Point Regression Kriging. Remote Sens. 12, 1082. https://doi.org/10.3390/rs12071082

Ye, Y., Huang, L., Zheng, Q., Liang, C., Dong, B., Deng, J., Han, X., 2021. A feasible framework to downscale NPP-VIIRS nighttime light imagery using multi-source spatial variables and geographically weighted regression. Int. J. Appl. Earth Obs. Geoinf. 104, 102513. https://doi.org/10.1016/j.jag.2021.102513

Yoo, E.H., Kyriakidis, P.C., 2006. Area-to-point Kriging with inequality-type data. J. Geogr. Syst. 8, 357–390. https://doi.org/10.1007/s10109-006-0036-7

Zhang, X., Gibson, J., 2022. Using Multi-Source Nighttime Lights Data to Proxy for County-Level Economic Activity in China from 2012 to 2019. Remote Sens. 14. https://doi.org/10.3390/rs14051282

Zhang, Y., Atkinson, P.M., Ling, F., Foody, G.M., Wang, Q., Ge, Y., Li, X., Du, Y., 2021. Object-based area-to-point regression kriging for pansharpening. IEEE Trans. Geosci. Remote Sens. 59, 8599–8614. https://doi.org/10.1109/TGRS.2020.3041724

Zhou, J., Liu, S., Li, M., Zhan, W., Xu, Z., Xu, T., 2016. Quantification of the scale effect in downscaling remotely sensed land surface temperature. Remote Sens. 8. https://doi.org/10.3390/rs8120975