Greenness and Nighttime Light Positively Affect Human Well-being within Certain Ranges - An Empirical Machine Learning Analysis

Abstract

The positive effects of greenness in living environments on human well-being are known. As a widely used proxy, the nighttime light (NTL) indicates the regional socio-economic status and development level. Higher development levels and economic status are related to more opportunity and higher income, ultimately leading to greater human well-being. However, whether simple increases in greenness and NTL always produce positive results remains inconclusive. Here, we demonstrate the complex relationships between human well-being and greenness and NTL by employing the random forest method. The accuracy of this model is 81.83%, exceeding most previous studies. According to the analysis results, the recommended ranges of greenness and NTL in living environments are 10.91% - 32.99% and 0 – 17.92 $nW/cm^2 \cdot sr$, respectively. Moreover, the current average monetary values of greenness and NTL are 3351.96 $USD/\%$ and 658.11 $USD/(nW/cm^2 \cdot sr)$, respectively. The residential areas are far away from the abundant natural resources, which makes the main population desire more greenness in their living environments. Furthermore, high urban development density, represented by NTL, has caused adverse effects on human well-being in metropolitan areas. Therefore, retaining a moderate development intensity is an effective way to achieve a sustainable society and improve human well-being.

Keywords:

Human Well-Being; NDVI; NTL; Environmental Valuation; Random Forest

Introduction

Greenness is positively associated with human well-being (Li and Managi, 2021b; MacKerron and Mourato, 2013; White et al., 2013). Greenness leads to more recreational and physical activities (Hunter et al., 2015), air pollution reduction (Eitelberg et al., 2016; Li and Managi, 2021a), relief of stress (Astell-Burt and Feng, 2019; Astell-Burt et al., 2014; Barton and Pretty, 2010), among others. However, whether simple increases in greenness would always give rise to improving human well-being remains inconclusive. In only terms of the effects of greenness, seemingly, human well-being would be greater with more greenness. Yet, the places with too much greenness do not have adequate land use for infrastructure construction and developing economics, which would hurt human well-being in a way. Previous studies show that people in Japan desire more urban land, especially in metropolitan areas, due to crowded living environments (Li and Managi, 2021b). Some cities express the need for both urban land and natural land use. The contradiction between green space and construction might be solved through moderate-intensity development and population density reduction. Moderate-intensity development could leave a sufficient amount of natural land use, making people live in cities naturally (Hartig and Kahn, 2016). Population density reduction guarantees that most people could have enough opportunities to access necessary infrastructure.

Additionally, high development intensity is related to busy human activities, which causes a series of environmental issues, such as light pollution (Falchi et al., 2011; Falchi et al., 2019), traffic noise (Begou et al., 2020; Pirrera et al., 2010), among others. These environmental issues lead to several detrimental effects on public health (Falchi et al., 2011; Pirrera et al., 2010). The highly developed regions have more human activities and, in turn, have more illumination. Thus, nighttime light (NTL) has long been widely used as a proxy for socio-economic development with geographical information in environmental studies (Chen and Nordhaus, 2011; Yeh et al., 2020). The brighter NTL is not only observed by the satellites but also sensed by sleeping or trying to sleep people. In fact, lighting levels in the highly developed regions are much higher than the needs for their tasks, and these residual lights ultimately affect wildlife and human health (Falchi et al., 2011). Without enough luminosity, the regions might be unsafe and poorly developed, which slashes human well-being (Suk and Walter, 2019). Yet, the surplus of luminosity would induce health issues (Falchi et al., 2011). Therefore, similar to greenness in the living environments, the NTL also needs to be within a moderate range, not too dark and not too bright. However, the rational ranges of greenness and NTL remain elusive.

Previous studies mainly assume the relationships between human well-being and greenness and NTL are linear, even though considering geographical locations or the time-fixed effects within individuals (Ghosh et al., 2013; Li and Managi, 2021b; White et al., 2013). Linear assumptions have a fatal shortcoming that the effects do not vary with the current status. Due to scarcity value, the effects of increases in greenness and NTL might be larger when the current value is low. When the greenness or NTL value is too high, their effects on human well-being might be lower or even become negative. In addition, based on linear assumptions, the accuracy of models is relatively low, generally no greater than 30%, e.g., (Akpinar et al., 2016; Ambrey and Fleming, 2014; Krekel et al., 2016). Hence, for highly accurate analyses, we should not directly assume the relationship between human well-being and features of interest before the analyses. The machine learning methods always build models of data without too many pre-assumptions. The random forest, one of the machine learning methods, does not need to assume the relationship, owing to its non-parametric algorithm (Breiman, 2001). To grasp the relationships between human well-being and greenness and NTL, we employ over 300,000 observations and apply the random forest method to acquire the high-fit model. The random forest is based on boosting technology. Simply speaking, the final results are predicted by a bundle of weak learners, decision trees. In random forests, every decision tree is also non-parametric, making it difficult to explain. To sum up, random forests are typically model-agnostic, so we must use tools to make their results understandable. Accumulated local effects (ALE) describe the average effects of features of interest on the prediction of a machine learning model (Apley and Zhu, 2020). In the light of the relationships illustrated by the ALEs, our research provides the recommended ranges of greenness and NTL value in the living environments to improve human well-being and achieve sustainable cities.

Materials and Methods

Materials

Survey

From 2015 to 2017, our team conducts three waves of nationwide surveys in Japan. The survey covers almost 300,000 and obtains over 450,000 validated questionnaires. To avoid interviewer bias and complete each wave within one month, we perform the surveys via the Internet. Since our survey includes several sensitive questions, such as education background, personal income, employment, among others, the Internet-based approach could reduce the probability of fake answers to some degree (Chapman et al., 2019; Li and

Managi, 2021c). The respondents are randomly sampled according to the population distribution. We obtain permission from the Japanese government to survey around November each year. According to the related laws and regulations, we must complete the survey during the appointed period, which is always four weeks around November. The laws and privacy protection rules make directly acquiring the detailed respondents' addresses difficult and illegal. However, asking the respondents' residential postal code is permitted, so we require the postal code in the survey. We use geometric centers of postal zones from Google Geocoding API as the approximate addresses of the respondents in the analysis. The respondents are from 54,144 postal zones. This survey covers many healthy, demographic and socio-economic characteristics, such as mental health status, self-reported health, age, gender, job, education background, income, among many others. After removing the observation with missing data, 372,685 records are kept. Although the random forest method could fill in missing data (Breiman, 2001), the time cost of computation is huge and unacceptable. Thus, we drop these missing data.

Human Well-Being

We take an overall LS evaluation as the indicator of human well-being. LS is one of the typical indexes of subjective well-being (SWB) (Diener et al., 2018). SWB can demonstrate human well-being with objective justification (Diener et al., 2018; Oswald and Wu, 2010). Furthermore, previous studies also use other indicators, including mental health, self-reported health, morbidity, among others, to represent human well-being (Diener et al., 2018; Park et al., 2021). However, compared with those indicators, SWB is straightforward. Generally, LS, happiness, and Cantril's Ladder could be regarded as SWB indicators individually, but LS is the most widely used (Diener et al., 2018). We apply a single question for evaluating individual LS: “overall, how satisfied are you with your life?”. Then, the respondents should select an answer from “very satisfied (5)” to “very unsatisfied (1)”. Here, it must be underscored that the LS score is qualitative and ordinal rather than quantitative. Yet, this analysis is not a classification task because the output scores are not independent nor categorical. **Figure 1.a** illustrates the statistical distribution of the LS evaluation score. Over 180,000 respondents deem that they are satisfied with their lives, i.e., the LS score is 4, which significantly exceeds other selections. If we perform the classification random forest on this data set, the classification accuracy for the people with lower or higher scores, mainly 1 and 5, would be extremely low, even with under-sampling or over-sampling technologies. In this case, regression random forest is a relatively better strategy, and we, therefore, assume the LS is continuous.

Normalized Difference Vegetation Index (NDVI) and Nighttime Light (NTL)

To examine the impact of greenness on human well-being, we use the 16-day level 3 NDVI data with 500-m spatial resolution produced by the U.S. National Aeronautics and Space Administration (NASA), which is widely used in previous environmental research (Lamchin et al., 2018; Li et al., 2015; Wang et al., 2020). The NDVI is a graphical index to describe whether the observed pixel contains live green vegetation and ranges from -1 (no live green vegetation, -100%) to 1 (rife with live green vegetation, 100%) (Didan et al., 2015). The NDVI data from NASA NASA’s products, MOD13A1 (https://lpdaac.usgs.gov/products/mod13a1v006/) and MYD13A1 (https://lpdaac.usgs.gov/products/myd13a1v006/), are based on MODIS Terra and Aqua

Satellites, respectively. The period of interest is from October to December each year because the survey is conducted around November. The MOD13A1 and MYD13A1's temporal resolution is 16-day. So, NASA provides the data on the 289^{th}, 305^{th}, 321^{st}, 337^{th}, and 353^{rd} in each year. We annually average the 16-day data from Terra and Aqua into one raster. Then, we take the geometric centers of postal zones as the center to build buffers with a 5-km radius. The mean NDVI values of each buffer are extracted, which are considered as the natural level of each postal zone. **Figure 1.b** demonstrates the statistical distribution of the mean NDVI.

Human activities, local economic status, and the goodness of infrastructures are associated with human well-being (Chen and Nordhaus, 2011, 2019). The NTL remote sensing data are widely applied to investigate human activities and economic status (Chen and Nordhaus, 2019; Chen et al., 2021; McCallum et al., 2022) because artificial electric light is equipped in most buildings and infrastructures. The NTL data are extracted from other NASA products, Suomi National Polar-orbiting Partnership Visible Infrared Imaging Radiometer Suite (NPPVIIRS) nighttime light data. The earth observation group provides monthly NTL data with a 1-km resolution. We calculate the average rasters of October's, November's and December's data in each year and then extract the mean NTL values via the buffers with 5-km radius. The unit of NTL is nanowatt per square centimeter steradian ($nW/cm^2 \cdot sr$). The mean NTL data are too scattered, so we use the following equation to reduce the range:

$$\boldsymbol{lnNTL} = \ln(\boldsymbol{NTL} + 1) \tag{1}$$

where $\boldsymbol{NTL}$ is the vector of the mean NTL data of each respondent. **Figure 1.c** shows the statistical distribution of the logarithm of NTL.

Demographic and Socio-economic Characteristics

Several demographic and socio-economic characteristics are controlled, as in previous studies (Krekel et al., 2016; Li and Managi, 2021b, c; MacKerron and Mourato, 2013). Our study is to probe the relationship between human well-being and living environments, so the attitudes toward living environments, including the safe feeling of living environments, the sense of goodness for living, and community attachment, are analyzed. The frequency of high-level and low-level stress and self-reported health represent health status. Gender, age, annual income, education background (one-hot vector), and employment status (one-hot vector) are also employed in the analysis. It must be noted that individual annual income in the survey is an income range rather than an accurate overall income. In total, 24 features are considered in the model (Descriptive statistics of the features shown in **Appendix Table A1**).

Methods

The Regression Decision Tree

A decision tree is the basic element of the random forest method. The decision tree predicts the output variable based on a series of judgments (Breiman, 2001; Liaw and Wiener, 2002). Therefore, the decision tree is an entirely non-parametric approach. When the output variable of the decision tree is categorical, the decision tree is to perform a classification task and called a classification tree. If the output variable is continuous, the decision tree is, in turn, named a regression tree. To analyze ordinal output variables, the previous studies mainly assume that they are continuous and use a regression decision tree (Hothorn et al., 2006). **Figure 2** illustrates a simple example of a regression decision tree with two layers. The algorithm should pass two internodes and do two judgments to complete the prediction in the example tree. In the example, two features, self-reported health and NDVI, are used to predict individual LS. Self-reported health in this study is ordinal, and NDVI is continuous. In the training process, the rules of each judgment and feature range splits are the primary things that the machine needs to "learn". A large amount of data is employed to train the decision tree to minimize the residual sum of squares (RSS). The judgment rules and feature range splits that generate the smallest RSS are the decided rules of the final model. We apply a greedy approach based on the minimization of the RSS to train regression decision trees (Breiman et al., 2017):

$$RSS = \sum_{l \in leaves} \sum_{i \in C_l} (y_i - \bar{y}_{C_l})^2 \tag{3}$$

where l is a leaf, C_l is the cases in leaf l, y_i is the observed value and $\bar{y}_{C_l}$ is the average observed value in leaf l. In this approach, one feature might be split into several ranges rather than only two. For example, the first internode might judge whether the NDVI exceeds 30%, and the second internode might focus on whether the NDVI is more than 50%. The increase in the number of internode of trees would reduce the RSS, but it also leads to over-fitting. Two rules are designed to avoid over-fitting. We set the RSS threshold and the threshold number of remaining cases in the end leaf (Breiman et al., 2017). If the RSS or the number of remaining cases in the end leaf is smaller than the thresholds, the further split stops.

Random Forest

A single regression decision tree is not enough to predict the output variable accurately. In most regression tasks, decision trees are usually deemed as weak learners due to their relatively lower goodness of fit and flexibility. However, they are good at grasping the non-linear relationship compared with the other regression methods. In order to improve the accuracy and keep the advantage of the decision tree, the random forest method is created (Breiman, 2001). Random forest ensembles hundreds of decision trees, lets them individually predict the result, and average all the predicted results. This is typically boosting technology, which aims to use a set of weak learners to create a strong learner (Schapire, 2003). Bootstrapping and bagging are the key technologies to improve the accuracy and reliability of the random forest (Liaw and Wiener, 2002). Bootstrapping is the sampling technology, which is random sampling with replacement. This technology allows the features with almost any statistical distribution to be used in the analysis. To process bootstrapping, the number of trees (N_{tree}) in the random forest should be set. The N_{tree} sub-samples with replacement are extracted from the total sample. The sizes of each sub-sample are 2/3 of the total sample size. Every decision tree is trained by the bootstrapped sub-sample. It should be emphasized that only a fixed number of random features (N_{try}) are used in a single decision tree, rather than all. For regression tasks, N_{try} is one third of the total number of features in the original data set (Breiman, 2001). After training, the random forest can predict the output variable by aggregating the predicted results from each tree. The process that uses the bootstrapped data set and aggregates the predicted results from each tree is terminologically named “bagging”. Because each tree only uses 2/3 data in the bagging, the remained data are called out-of-bagging (OOB) data.

The OOB data are applied to test the accuracy of models. OOB data are employed to calculate the OOB error, which is the residuals of OOB data caused by the trained random forest. This process is similar to cross-validation but has been assembled in the random forest method.

Two parameters in the random forest, N_{tree} and N_{try}, must be pre-defined. The OOB error of the random forest will reduce in the beginning with the number of trees increasing. However, once the tree number exceeds a certain value, the benefit from the increase in the tree number is marginal. According to our experience, the random forest with 1,000 trees could minimize the OOB error to a stable level, so the N_{tree} is set to 1,000. Additionally, previous studies and the method designer recommend the feature number employed in each tree should be 1/3 of all features in the original data set (Breiman, 2001; Breiman et al., 2017; Liaw and Wiener, 2002). Because we have 24 features in the total sample, so N_{try} is set to 8.

Compared with typical linear regression methods widely applied in previous studies, such as ordinary least square (OLS) and ordered logit regression (OLR), the random forest has an obvious advantage that this method does not need to assume the linear relationships between the output variable and features. In fact, the relationships in the real world are normally more complex than linear. For example, the relationship between NDVI and human well-being might not be linear. Without any greenness in the living environments, humans might feel unhappy. In this case, the association is positive. However, if the residential area fills with only greenness, that would be deemed as an entirely undeveloped area. Living in an undeveloped area, humans might also feel uncomfortable. Now, the relationship becomes negative. If we assume a linear relationship, the conclusions might be misleading. Furthermore, the regression methods based on the matrix approach to estimate the coefficients, such as OLS and OLR, must avoid multicollinearity in the analyses. In real-world situations, there are many variables representing completely different data but highly correlated. In the OLS or OLR, we need to drop some variables to guarantee no correlation between independent variables exists. In the random forest, this variable selection process is unnecessary because this model does not require the features to be independent and identically distributed.

Feature Importance

Theoretically, similar to other machine learning methods, we can put as many features into the random forest model as possible to fit if we do not consider time complexity and computer memory limitation. However, unfortunately, we must take time and computer calculation ability into account. Therefore, only essential features could be kept in the analysis. The random forest model estimates the feature importance by computing the increase in root mean squared error (IncRMSE) between before and after permutating a certain feature (Breiman, 2001). A high IncRMSE means that the removed feature is vital to predicting the output variables because the sum of residuals increases significantly with this feature. In our research, we require the IncRMSE of each feature is larger than 10% of the mean of output variable values. Accordingly, 24 features are employed in the analysis. It must be emphasized that the calculation with 24 features should take over 20 hours because we execute the paralleling calculation with 100 cores running together on a high-performance computer. If performing the calculation with one core, the time might be several weeks. Thus, even with a high-performance computer, feature permutation based on feature importance is needed.

Partial Dependency Profile

The high accuracy of the random forest model comes with the price that the results are complex and challenging to understand and explain (Friedman, 2001; Greenwell, 2017). PDPs offer a simple solution to global explanations, which demonstrate the relationship between the output variable and features of interest. PDPs are estimated as follows:

$$\hat{f}_{FOI,PDP}(FOI_l) = \frac{1}{n}\sum_{i=1}^{n}\hat{f}(FOI_l, \boldsymbol{OF}_i) \tag{2}$$

where $\hat{f}_{FOI,PDP}(FOI_l)$ is the average predicted outcome given a value FOI_l of a specific feature of interest FOI when all other features ($\boldsymbol{OF}_i$) are the values from the original data set, FOI_l is a given possible value of the feature of interest FOI, $\boldsymbol{OF}_i$ is a vector of feature values except FOI of the respondent i, $\hat{f}$ is the trained random forest model, and n is the data set size, 372,865. All FOI_l belong to an arithmetic progression (AAP):

$$FOI_l \in AAP \ (l \in [1, m]) \tag{3}$$

where m is the user-defined resolution of FOI. FOI_1 is the minimum of FOI's observed values, and FOI_m is the maximum of FOI's observed values. The common difference of AAP is $(FOI_m - FOI_1)/(m-1)$. Simply speaking, PDPs set the FOI value of the original data set to FOI_l, then use the new data set and trained model to predict the output variable. However, the prediction output is still a sequence that might not perfectly fit a continuous function. Therefore, further processes are needed. Additionally, PDPs are global predictions because all observations are employed. In this way, the PDPs work only when no feature correlations exist. It must be underscored that even though the random forest model does not care about the correlation between the features, the explanation method requires no correlation. If the features of interest correlate with other features, their PDPs cannot be trusted. **Equation (2)** assumes that $\boldsymbol{OF_i}$ is fixed. If significantly high-level correlations exist, AAP is unreasonable. Our study is a case in point. The NDVI and logarithm of NTL are significantly and highly correlated because their Pearson's correlation coefficient is -0.811 (p-value < 0.1%). PDPs are not the ideal tool to explain our results.

Accumulated Local Effects

ALEs estimate the effects, similar to PDPs, but using a local data set (Molnar, 2020; Molnar et al., 2020). When ALEs estimate the effect of FOI_j, only the data's FOI values close to the defined FOI_j rather than all data set are included in the process (Apley and Zhu, 2020). The selected data are the local data sets. The local data sets are with lower probability of having irrational other feature values. Therefore, ALEs are more trustable than PDPs in this situation. To compute ALEs of a specific feature of interest (FOI), we first divide the range of FOI into K intervals. We use K-nearest neighbour method to obtain the K intervals. The FOI_j belongs to the set $N(FOI_j)$, which is $(FOI_{k-1}, FOI_k]$. Here, k belongs to $\{1,2, \ldots, K\}$. In practice, the K value is always defined first. Then, based on k-nearest neighbour method, the intervals are divided (Apley and Zhu, 2020). Based on the division, ALE is calculated as the following two steps. The first step is to estimate the uncentered effect of a feature of interest:

$$\hat{g}_{FOI,ALE}(FOI_j) = \sum_{k=1}^{k(FOI_j)} \frac{1}{n(k)} \sum_{[i:FOI_i \in N(FOI_j)]} [\hat{f}(FOI_k, \boldsymbol{OF}_i) - \hat{f}(FOI_{k-1}, \boldsymbol{OF}_i)] \tag{4}$$

where $\hat{g}_{FOI,ALE}(FOI_j)$ is the uncentered effect of the specific feature value FOI_j, $k(FOI_j)$ is the order of the interval that FOI_j belongs to, $n(k)$ is the observation number of the interval k, $\boldsymbol{OF}_i$ is a vector of feature values except FOI of the respondent i in the local data set, and $\hat{f}$ is the trained random forest model.

According to **Equation (4)**, the terminology, ALE, can be clearly clarified (Apley and Zhu, 2020; Molnar, 2020). In the first part of the right side of **Equation (4)**, $\sum_{k=1}^{k(FOI_j)}$ explains the word "Accumulated". For example, if FOI_j belongs to the 3rd interval, $\hat{f}_{FOI,ALE}(FOI_j)$ equals the accumulative values of average uncentered local effects from the first three intervals. The condition of data selection, $[i: FOI_i \in N(FOI_j)]$, limit the local data set, which explains "Local". Only the observed values similar to defined value FOI_j could be put into calculation. The last part, $[\hat{f}(FOI_k, \boldsymbol{OF}_i) - \hat{f}(FOI_{k-1}, \boldsymbol{OF}_i)]$, computes the effects of the FOI value change on the output variable, which explains "Effects".

To make ALEs of each interval comparable, we centralize the $\hat{g}_{FOI,ALE}(FOI_j)$, as follows:

$$\hat{f}_{FOI,ALE}(FOI_j) = \hat{g}_{FOI,ALE}(FOI_j) - \frac{1}{n}\sum_{k=1}^{K} n(k)\, \hat{g}_{FOI,ALE}(FOI_k) \tag{5}$$

where $\hat{f}_{FOI,ALE}(FOI_j)$ is the estimated ALE value of FOI_j. However, in most complex analyses, the function between the feature of interest and its ALE is not continuous. In other words, the function cannot be derived everywhere. Thus, the explanation ability of ALE is still limited in this way.

Pseudo Accumulated Local Effects Function

The function between the feature of interest and its ALE is not continuous, and the feature of interest in the calculation is a sequence. We must use a surrogate function, pseudo accumulated local effects function (PALEF), to demonstrate the complex mathematical relationship between the feature of interest and its ALE. Assuming that the relationships are high-order splines is an effective way. The results from **Equation (5)** are used to fit the PALEF. To keep the results understandable, we require either the R^2 of PALEF to be higher than 99% or the order of PALEF to be not greater than 20. The PALEF is estimated as follows:

$$ALE_{FOI}(k) = \beta_0 + \beta_1 FOI_k + \beta_2 FOI_k^2 + \cdots + \beta_m FOI_k^m \quad (6)$$

where $ALE_{FOI}(k)$ is the estimated ALE of the kth interval of the feature FOI, FOI_k is the pre-defined division boundary value. The m value would keep increasing from 1 until the R^2 of PALEF exceeds 99%, or the m value equals 20. If the order is too high, the PALEF might become shaky, making the model difficult to explain.

Monetary Values of Features

To make the ALE of the feature value changes on human well-being understandable, we estimate their monetary values. The basic pathway of our method is that the ALEs of the features of interest should be offset by the ALE of income change, assuming that nothing else changes:

$$ALE_{FOI}(a) - ALE_{FOI}(b) = ALE_{Income}(c) - ALE_{Income}(d) \quad (7)$$

where ALE_{FOI} is the PALEF of a certain the feature of interest, ALE_{Income} is the PALEX of income, a and b are the feature of interest values after and before the change, and c and d are the income values after and before the change. If we limit the feature of interest value change and income change to zero, **Equation (7)** would be rewritten as follows:

$$\theta(a,c) = \frac{ALE'_{FOI}(a)}{ALE'_{Income}(c)} \tag{8}$$

where $ALE'_{FOI}(a)$ is the derivative of ALE'_{FOI} at a, $ALE'_{Income}(c)$ is the derivative of ALE'_{Income} at c, and $\theta(a,c)$ is the monetary value function of the feature of interest. Apparently, the monetary value of a specific feature of interest in this study is non-stationary and affected by the current situation.

Results

Our random forest builds 1,000 decision trees, and each tree employs eight randomly selected features. The accuracy of the random forest is 81.83%, whereas the regression model, OLS, is only 26.83%. Other indicators, including mean square error (RMSE), mean square error (MSE), and mean absolute error (MEA), also indicate that the random forest is better than OLS. The random forest's RMSE, MSE, and MEA are 0.399, 0.159, and 0.248, respectively, while OLS's RMSE, MSE, and MEA are 0.801, 0.641, and 0.631, respectively. In terms of accuracy, the random forest is significantly better than the linear regression.

Figure 3 illustrates the IncRMSE after permutating a certain feature. The frequency of high-level stress affects the LS the most. The IncRMSE of this feature is 0.745. Simply speaking, if we do not include the feature, the frequency of high-level stress, in the analysis, the RMSE will increase 0.745. The IncRMSEs of the logarithm of NTL and NDVI are 0.623 and 0.615, respectively. If the logarithm of NTL or NDVI is not put into the model, the RMSE of our model will increase 156.21% or 154.00%. In fact, the IncRMSEs of the logarithm of NTL and NDVI are similar to age's IncRMSE that is 0.632. Therefore, to analyze human well-being, the living environment features, including NTL and NDVI, must be taken into account.

Figure 4 displays the PDPs of NDVI, the logarithm of NTL, and income. Since the features, NDVI and logarithm of NTL, are highly related, **Figure 4.a** and **b** cannot be trusted. However, the feature, income, is not highly correlated with other features. Thus, we could check the validity of ALEs by comparing the ALE and PDP of income. **Figure 5** shows the ALEs of the three variables of interest. **Figures 4.a** and **5.a** show that the ALE and PDP of NDVI are significantly different. According to **Figure 5.a**, people whose environments have 10.91% - 32.99% greenness (NDVI $\in$ $[10.91\%, 32.99\%]$) are more likely to have higher LS. The effects of an increase in NDVI on human well-being are apparently positive before the NDVI value reaches 16.97%. After the greenness in the living environments exceeds 16.97%, the effects of the increase in NDVI on human well-being are mainly negative. People desire greenness, but too much greenness lefts less space for other necessary infrastructure construction. Without convenient living environments, human well-being would not be high. In this way, the explicit relationship between NDVI and human well-being is non-linear, though, in nature, people might be willing to stay with more greenness. **Figure 5.b** demonstrates the ALE of the logarithm of NTL. People's LS is higher than the average level when their environments have 0 – 17.92 $nW/cm^2 \cdot sr$ NTL (the logarithm of NTL $\in [0, 2.94]$). When the logarithm of NTL is less than 1.03, the effects of an increase in NTL on human well-being are positive. However, when it further increases, the effects become negative. Here, we emphasize that the used NTL is the average value of the 5-km buffer. The higher average NTL means the buffer is relatively busier and has more human activities even at night. Noise and light pollution caused by human activities might adversely affect human well-being. Hence, the level of development should be controlled within a reasonable range. According to the comparison between ALE and PDP of the income, the estimations of the effect based on local and global data sets are similar. Because income is not highly correlated with other features in the analysis, the ALE and PDP of the income can replace each other conditionally to some degree. As shown in **Figure 5.c**, in the analyzed range, the impact of an increase in income on human well-being is always positive. Once the individual income exceeds 4.5 million Japanese Yen (JPY) per year, people tend to be more satisfied with their lives than average level.

Figure 6 illustrates the PALEF of NDVI, the logarithm of NTL, and income. We use a 20-order spline to fit the relationship between NDVI and its ALE. The accuracy of the NDVI PALEF is 98.15%. The PALEF of the logarithm of NTL is a 2-order spline, and its accuracy is 99.27%. Furthermore, the 6-order is employed to fit the relationship between income and its ALE. Its accuracy is 99.53%. The average monetary value of a 1%-NDVI increase in living environments is 0.382 (0.379 – 0.387) million JPY, and the average monetary value of a 1-$nW/cm^2 \cdot sr$ increase in NTL is 0.075 (0.072 – 0.077) million JPY. Based on these average monetary values of NDVI and NTL, greenness and NTL are currently desired in Japan. **Figure 7** demonstrates the gridded average monetary value of NDVI. The grid size is $0.25° \times 0.25°$, roughly $30\text{km} \times 30\text{km}$. It must be noted that the grids where the observation number is no greater than 30 have been ignored on the map. According to **Figure 7**, people in most areas of Japan are desired more greenness. This situation is marked in the metropolitan areas, such as Tokyo (number 13), Osaka (number 27), and Nagoya in Aichi (number 23). Intriguingly, the gridded average values in Hokkaido (number 1), Aomori (number 2), Akita (number 5) and Yamagata (number 6) are also higher than the average level. In fact, these areas are the main agricultural zones in Japan. Due to their high latitude, the cropland becomes bare in the winter. This causes the scarcity of greenness in these places during the survey periods. **Figure 8** displays the gridded average monetary values of NTL. In metropolitan areas, including the Great Tokyo region (surrounding number 13), Osaka (number 27), Nagoya in Aichi (number 23), and Sapporo in Hokkaido (number 1), the negative values appear in pieces, which indicates that the NTL there has burdened human well-being. Except the highly developed areas, the negative values also arise sporadically. These scattered negative values are located in the center cities of the rural areas. The respondents cluster in these cities with similar environments, so their attitudes toward the NTL are almost identical.

Discussion

Our main findings are that the relationships between human well-being and greenness as well as NTL are non-linear and non-monotonous. Increases in greenness and NTL in the living environments would improve human well-being only when the current values of NDVI and NTL are low. Once the greenness exceeds about 17% and the NTL is over around 18 $nW/cm^2 \cdot sr$, the increases in greenness and NTL would adversely affect human well-being. Accordingly, the development level and greenness in the living environments need to be moderate. Furthermore, this is the first study that fits ALE to PALEF, which could help researchers further and deeply explain the random forest model. Based on the PALEFs of NDVI, the logarithm of NTL, and income, the monetary values of an increase in NDVI and NTL are estimated which are 0.382 (0.379 – 0.387) million JPY for a 1% increase in average NDVI and 0.075 (0.072 – 0.077) million JPY for a 1-$nW/cm^2 \cdot sr$ increase in NTL in the living environments. In view of the 3-year average exchange rate between JPY and US dollar (113.96 JPY/USD), the monetary values of NDVI and NTL could be rewritten as 3351.96 $USD/\%$ and 658.11 $USD/(nW/cm^2 \cdot sr)$, respectively. It must be underscored that our model's accuracy is relatively high compared with previous studies. The accuracy, RMSE, MSE, and MEA are 81.83%, 0.399, 0.159, and 0.248, respectively, while the accuracy of previous studies is generally no more than 30%.

Numerous previous studies have proved that greenness is positively associated with human well-being. The linear relationship between greenness and human well-being has been widely assumed based on the direct impacts of greenness. However, this assumption has a strong shortcoming. If a person's living environment has only greenness, the human well-being of this person would peak since the impacts of greenness are constant in linear assumption. The reality is far from this assumption. Some studies show that the average human well-being in rural areas is even lower than in urban areas. Although more greenness might not directly reduce human well-being, the land use for construction and development is significantly insufficient. In the rural areas, less opportunity and lower salaries would slash human well-being, while in the urban areas, crowded living environments also lead to poor results (Li and Managi, 2021b). In a sense, the direct impacts of greenness on human well-being are overemphasized, whereas its indirect effects are not well-researched or even simply ignored. In this study, the relationship between greenness and human well-being is complex. Moreover, we recommend that the greenness in living environments should be moderate and belong to a specific range [10.91%, 32.99%]. Within this range, the greenness and land use for construction and development might be balanced. However, greenness in Japan still seems to be inadequate because the most gridded average monetary values of NDVI are positive, even though the forest rate in Japan is 67.26%. In other words, the average impact of an increase in greenness is positive, and people desire more greenness in their living environments. In this case, the Japanese government should consider how to make the greenness closer to people's residential areas. Almost 91.7% urbanization rate indicates that most Japanese population are inaccessible to their abundant natural environments. Therefore, efficient and effective usage of the natural capital is a sustainable pathway to further improving human well-being in Japan.

NTL is an essential indicator representing economic status and human activities (Chen and Nordhaus, 2011, 2019; Zhao et al., 2017). The correlation between luminosity data and the regional GDP is significantly positive (Chen and Nordhaus, 2011). The places are brighter where the development level and economic status are better. The highly developed areas have more opportunities and relatively higher incomes, so people there are more likely to have greater human well-being. However, there is a turning point at which higher income no longer leads to greater human well-being (Jebb et al., 2018). Furthermore, previous studies mostly assume that the relationship between well-being and income is logarithmic (Bertram and Rehdanz, 2015; Kopmann and Rehdanz, 2013; Layard et al., 2008; Li and Managi, 2021c). Based on the logarithmic assumption, when the current income is high, the effects of an increase in income on human well-being are marginal. Moreover, the adverse impacts of NTL on human health are rarely mentioned. The light pollution causes poor health (Falchi et al., 2011; Falchi et al., 2019). Busy human acticities are associated with more noise, especially nocturnal road traffic noise. Primary sleep disturbances, cardiovascular diseases, among others are partially attributable to the detrimental effects of the noise (Begou et al., 2020; Pirrera et al., 2010). For all these reasons, the impact of an increase in NTL should be negative, when the current NTL value has already been high, whereas the impact is positive, when the NTL is insufficient. Our result demonstrate the characteristic of the relationship between NTL and human well-being. It must be underscored that the PALEF of the logarithm of NTL is inverted U-shape, but the mathmatical relationship between NTL and human well-being are more complex. According to mathematical relationship in the PALEX of the logarithm of NTL, after the NTL exceeds 2.80 $nW/cm^2 \cdot sr$, the impacts of a increase in the NTL will become negative. Therefore, development planning should balance the development level and its impacts on human well-being to achieve a sustainable society.

There are several limitations and issues worthy of note. First, although this study employs a short panel data set, the time-fixed effects within indiviudals are not considered. To the best of our knowledge, few studies discuss whether and what data transformations are needed for panel regression random forest. Second, more features should be put into the analysis. Although we have used more features than in previous studies, other features, such as marital status, general health questionnaire score, among others, are still unavailable. Thirdly, the output variable, LS, has too few degrees. An increase in the number of LS evaluation selections could further increase the accuracy of the analysis. Finally, the spatial relationships between observations are not well-detected in our research since the random forest is not so deeply developed as the spatial regression family. In further studies, the random forest model should be optimized to make it able to grasp the time-fixed within individuals. Furthermore, whether geographical contexts could further increase the accuracy of random forests should be probed.

Conclusion

Simple increases in greenness or NTL in living environments without considering the status quo do not always improve human well-being. People are more likely to have greater human well-being when their living environments have 10.91% - 32.99% greenness and 0 – 17.92 $nW/cm^2 \cdot sr$ NTL. In Japan, the current average monetary values of greenness and NTL are 3351.96 $USD/\%$ and 658.11 $USD/(nW/cm^2 \cdot sr)$, respectively. According to these average monetary values, the greenness and NTL are insufficient in most areas in Japan. Japan has no shortage of greenness, but the greenness is far from residential areas, which causes a seeming lack of greenness. Our study illustrates the high-accuracy relationship between human well-being and greenness and NTL to provide more information for governments and the public. This essential information helps to formulate sustainable land-use and development policies to improve human well-being.

Data Availability

The fully reproducible codes are publicly available at https://github.com/MichaelChaoLi-cpu/Greenness_NighttimeLight_WB.git. Data are available from the corresponding author on reasonable request.

Acknowledgment

This research was supported by the following funding agencies: JSPS KAKENHI (Grant No. JP20H00648), the Environment Research and Technology Development Fund of the Environmental Restoration and Conservation Agency of Japan (Grant No. JPMEERF20201001), and also JST SPRING (Grant No. JPMJSP2136).

Figure:

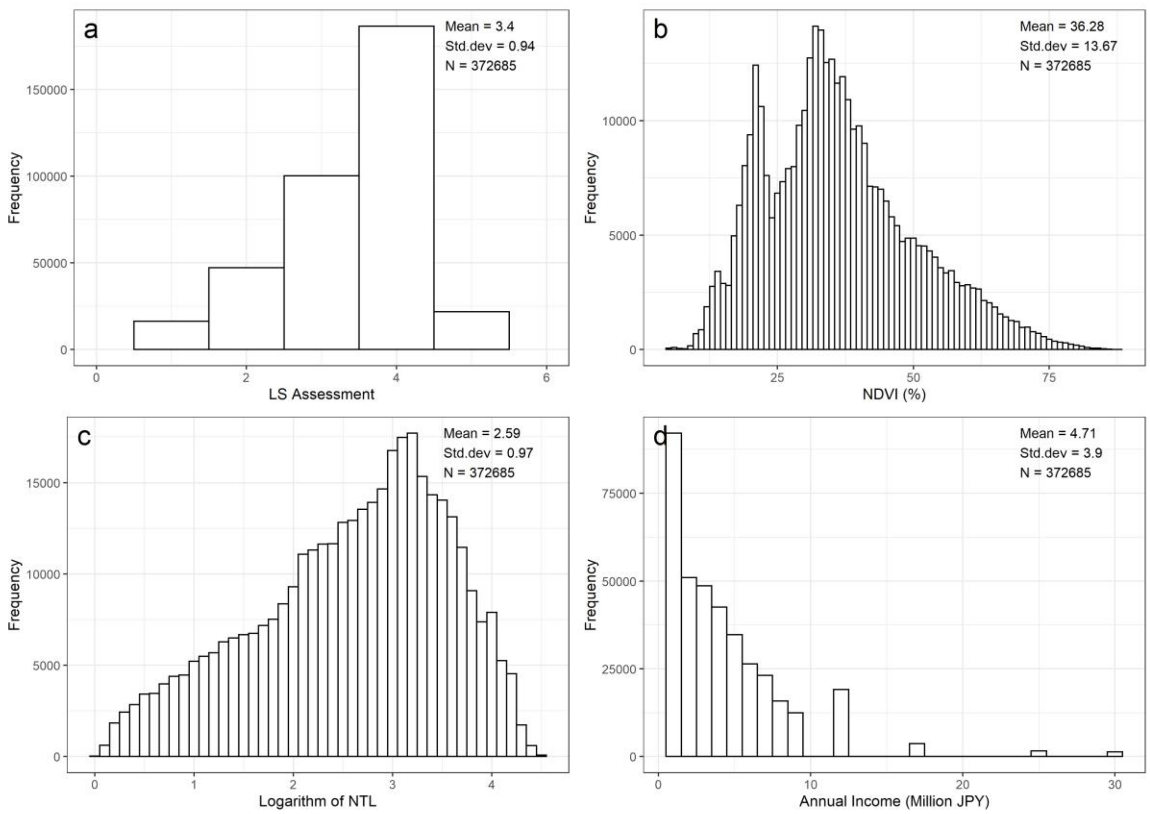

Figure 1: The Statistical Distributions of Critical Variables

(a: LS; b: NDVI; c: NTL; d: Annual Income)

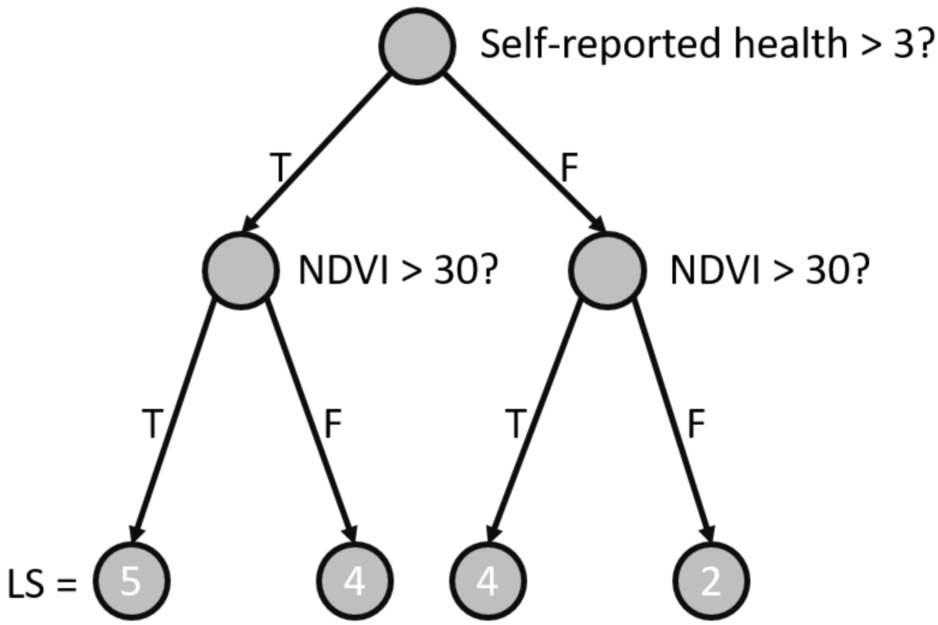

Figure 2: Example of a Regression Decision Tree

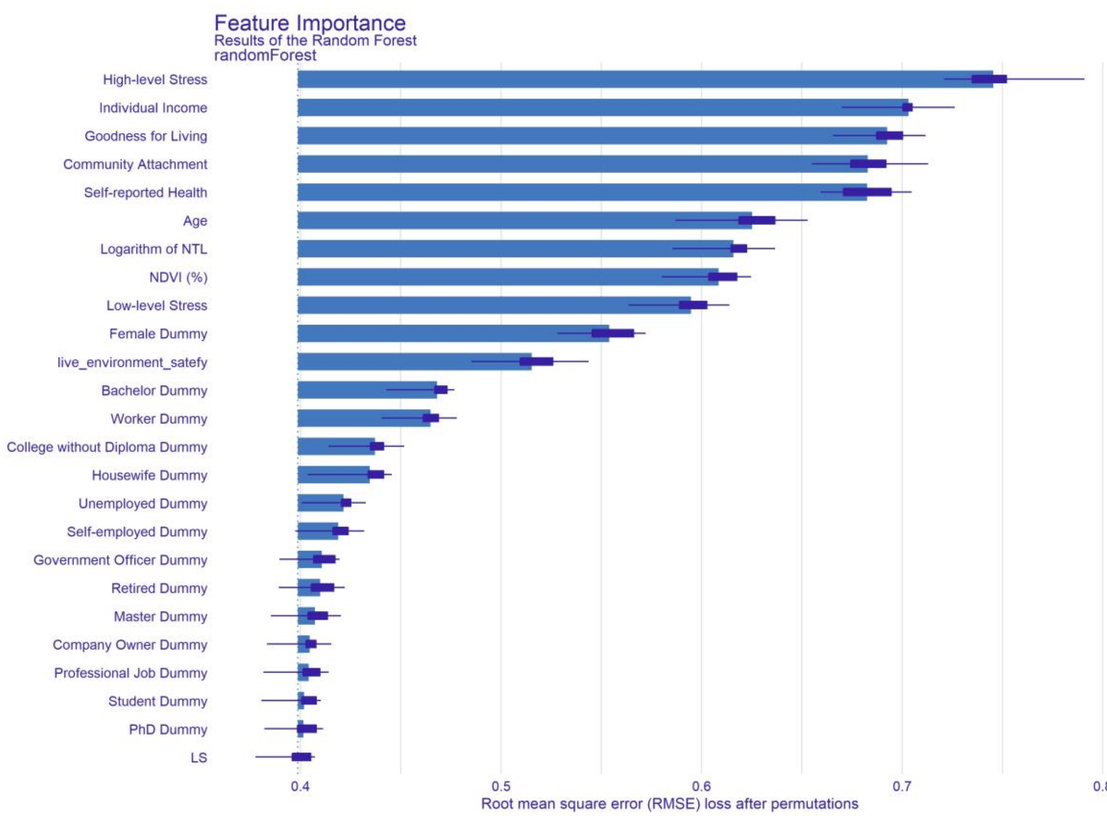

Figure 3: The Importance of Features

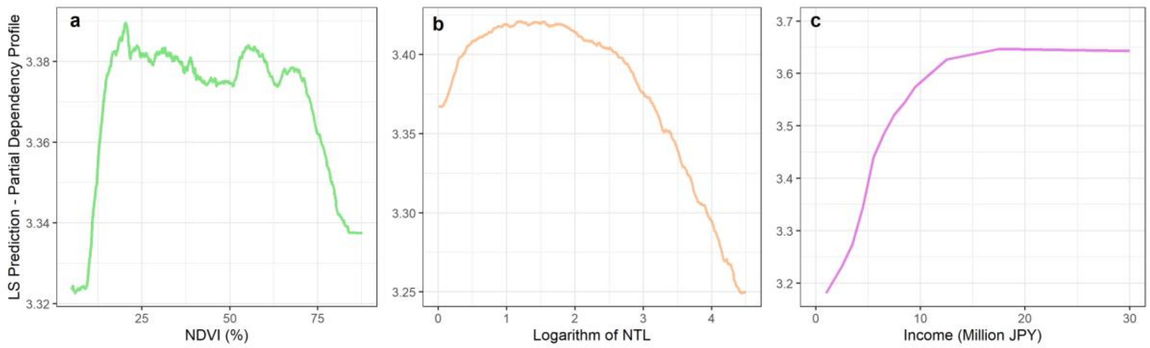

Figure 4: The PDPs of the Features of Interest

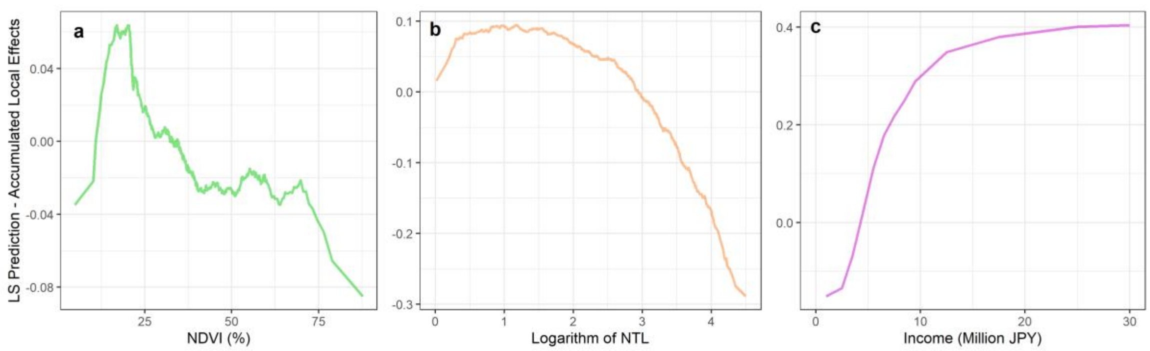

Figure 5: The ALEs of the Features of Interest

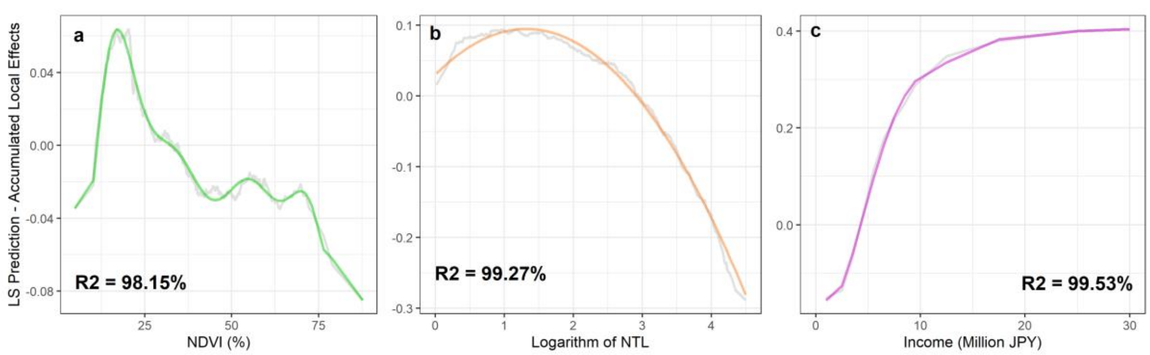

Figure 6: The PALEF of the Features of Interest

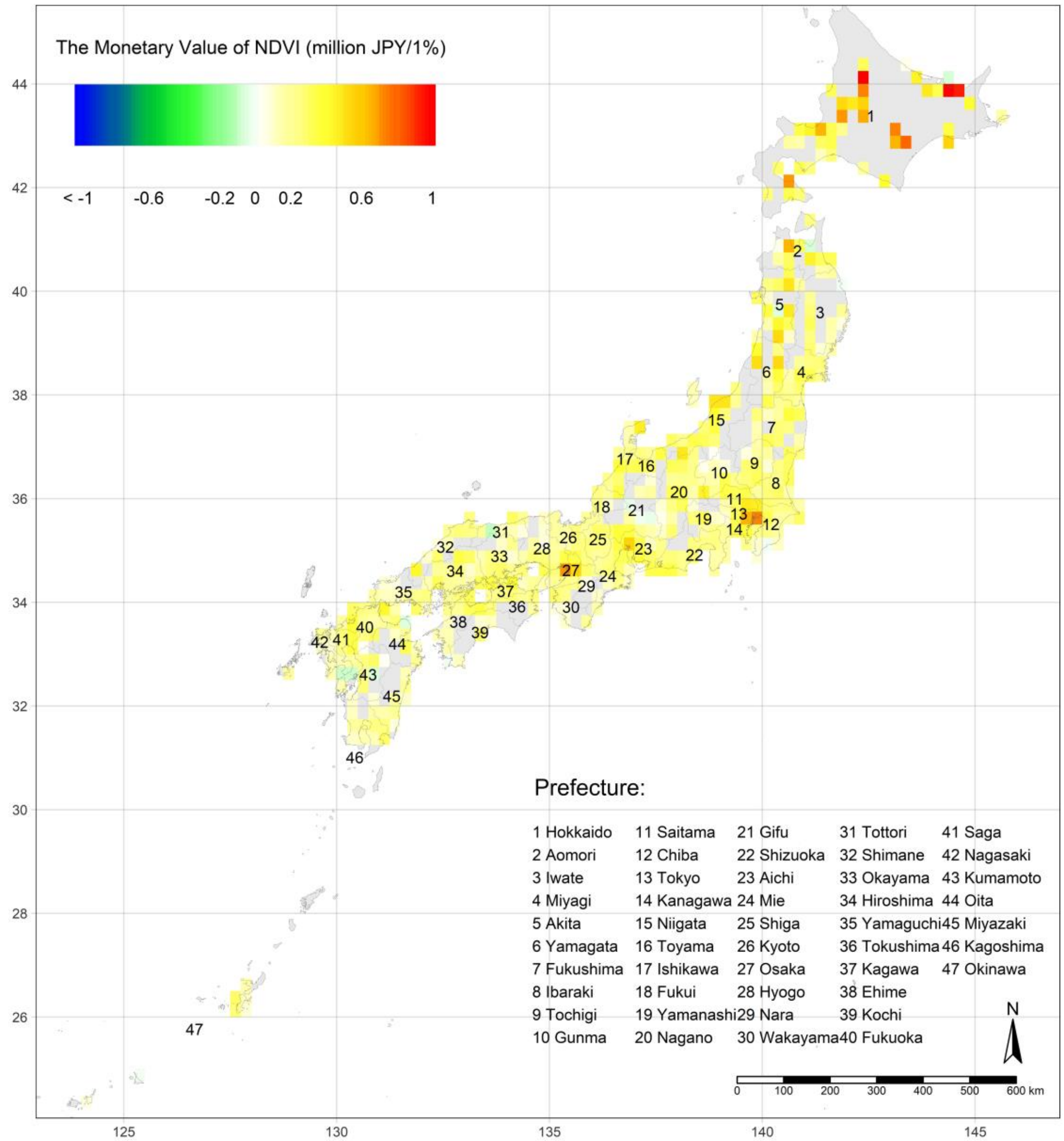

Figure 7: Map of Gridded Average Monetary Value of NDVI

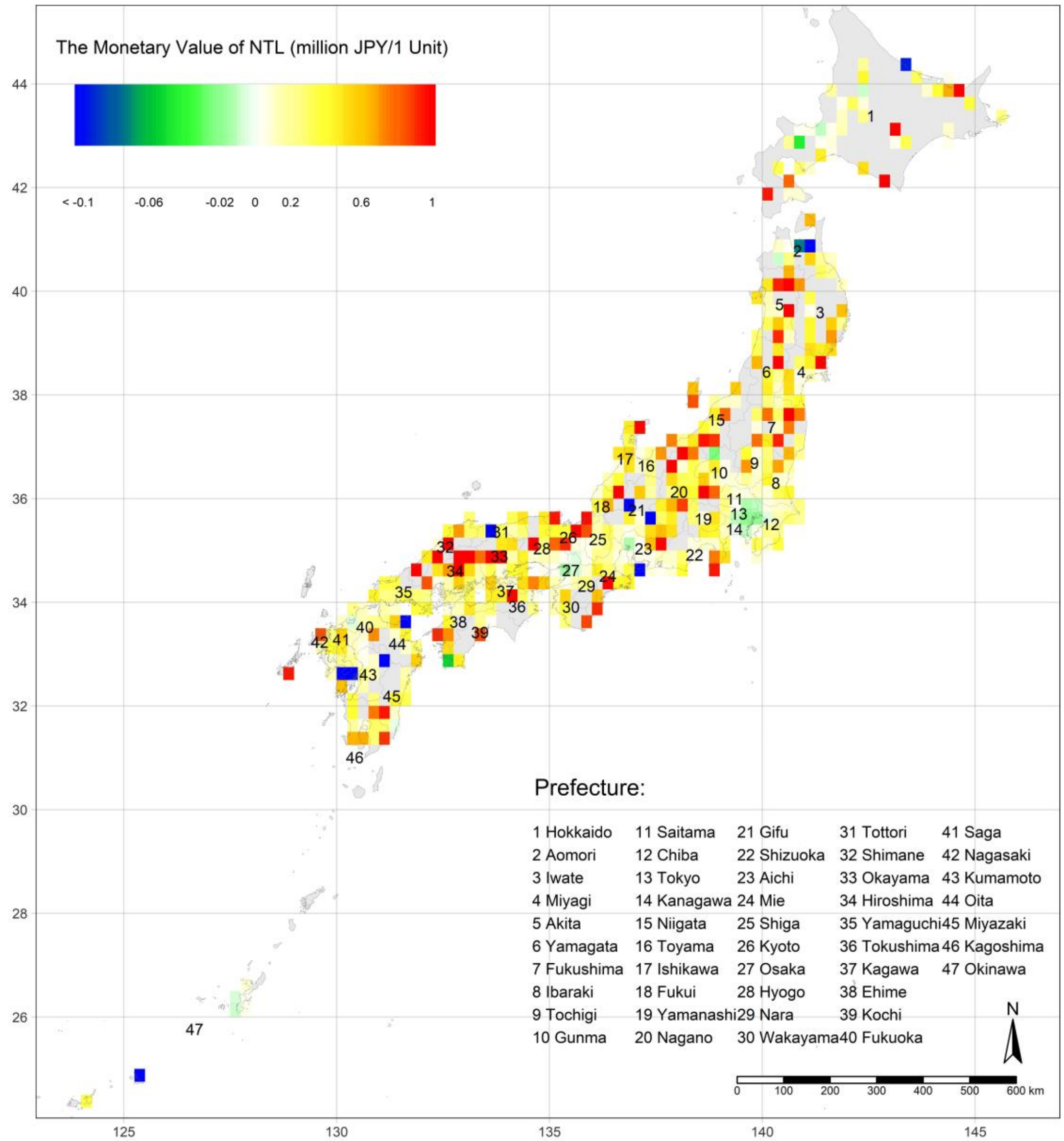

Figure 8: Map of Gridded Average Monetary Value of NTL

Reference:

Akpinar, A., Barbosa-Leiker, C., Brooks, K.R., 2016. Does green space matter? Exploring relationships between green space type and health indicators. Urban Forestry & Urban Greening 20, 407-418.

Ambrey, C., Fleming, C., 2014. Public Greenspace and Life Satisfaction in Urban Australia. Urban Studies 51, 1290-1321.

Apley, D.W., Zhu, J., 2020. Visualizing the effects of predictor variables in black box supervised learning models. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 82, 1059-1086.

Astell-Burt, T., Feng, X., 2019. Association of Urban Green Space With Mental Health and General Health Among Adults in Australia. JAMA Network Open 2, e198209.

Astell-Burt, T., Mitchell, R., Hartig, T., 2014. The association between green space and mental health varies across the lifecourse. A longitudinal study. Journal of Epidemiology and Community Health 68, 578-583.

Barton, J., Pretty, J., 2010. What is the Best Dose of Nature and Green Exercise for Improving Mental Health? A Multi-Study Analysis. Environmental Science & Technology 44, 3947-3955.

Begou, P., Kassomenos, P., Kelessis, A., 2020. Effects of road traffic noise on the prevalence of cardiovascular diseases: The case of Thessaloniki, Greece. Science of The Total Environment 703, 134477.

Bertram, C., Rehdanz, K., 2015. The role of urban green space for human well-being. Ecological Economics 120, 139-152.

Breiman, L., 2001. Random Forests. Machine Learning 45, 5-32.

Breiman, L., Friedman, J.H., Olshen, R.A., Stone, C.J., 2017. Classification And Regression Trees.

Chapman, A., Fujii, H., Managi, S., 2019. Multinational life satisfaction, perceived inequality and energy affordability. Nature Sustainability 2, 508-514.

Chen, X., Nordhaus, W.D., 2011. Using luminosity data as a proxy for economic statistics. 108, 8589-8594.

Chen, X., Nordhaus, W.D., 2019. VIIRS Nighttime Lights in the Estimation of Cross-Sectional and Time-Series GDP. Remote Sensing 11, 1057.

Chen, Z., Yu, B., Yang, C., Zhou, Y., Yao, S., Qian, X., Wang, C., Wu, B., Wu, J., 2021. An extended time series (2000–2018) of global NPP-VIIRS-like nighttime light data from a cross-sensor calibration. Earth System Science Data 13, 889-906.

Didan, K., Munoz, A.B., Solano, R., Huete, A., University of Arizona: Vegetation Index Phenology Lab, 2015. MODIS vegetation index user’s guide (MOD13 series).

Diener, E., Oishi, S., Tay, L., 2018. Advances in subjective well-being research. Nature Human Behaviour 2, 253-260.

Eitelberg, D.A., Van Vliet, J., Doelman, J.C., Stehfest, E., Verburg, P.H., 2016. Demand for biodiversity protection and carbon storage as drivers of global land change scenarios. 40, 101-111.

Falchi, F., Cinzano, P., Elvidge, C.D., Keith, D.M., Haim, A., 2011. Limiting the impact of light pollution on human health, environment and stellar visibility. Journal of Environmental Management 92, 2714-2722.

Falchi, F., Furgoni, R., Gallaway, T.A., Rybnikova, N.A., Portnov, B.A., Baugh, K., Cinzano, P., Elvidge, C.D., 2019. Light pollution in USA and Europe: The good, the bad and the ugly. Journal of Environmental Management 248, 109227.

Friedman, J.H., 2001. Greedy Function Approximation: A Gradient Boosting Machine. The Annals of Statistics 29, 1189-1232.

Ghosh, T., Anderson, S., Elvidge, C., Sutton, P., 2013. Using Nighttime Satellite Imagery as a Proxy Measure of Human Well-Being. Sustainability 5, 4988-5019.

Greenwell, B.M., 2017. pdp: an R Package for constructing partial dependence plots. R J. 9, 421.

Hartig, T., Kahn, P.H., 2016. Living in cities, naturally. Science 352, 938-940.

Hothorn, T., Hornik, K., Zeileis, A., 2006. Unbiased Recursive Partitioning: A Conditional Inference Framework. Journal of Computational and Graphical Statistics 15, 651-674.

Hunter, R.F., Christian, H., Veitch, J., Astell-Burt, T., Hipp, J.A., Schipperijn, J., 2015. The impact of interventions to promote physical activity in urban green space: A systematic review and recommendations for future research. Social Science & Medicine 124, 246-256.

Jebb, A.T., Tay, L., Diener, E., Oishi, S., 2018. Happiness, income satiation and turning points around the world. Nature Human Behaviour 2, 33-38.

Kopmann, A., Rehdanz, K., 2013. A human well-being approach for assessing the value of natural land areas. Ecological Economics 93, 20-33.

Krekel, C., Kolbe, J., Wuestemann, H., 2016. The greener, the happier? The effect of urban land use on residential well-being. Ecological Economics 121, 117-127.

Lamchin, M., Lee, W.-K., Jeon, S.W., Wang, S.W., Lim, C.H., Song, C., Sung, M., 2018. Long-term trend and correlation between vegetation greenness and climate variables in Asia based on satellite data. Science of The Total Environment 618, 1089-1095.

Layard, R., Mayraz, G., Nickell, S., 2008. The marginal utility of income. Journal of Public Economics 92, 1846-1857.

Li, C., Managi, S., 2021a. Contribution of on-road transportation to PM2.5. Scientific Reports 11, 21320.

Li, C., Managi, S., 2021b. Land cover matters to human well-being. Scientific Reports 11.

Li, C., Managi, S., 2021c. Spatial Variability of the Relationship between Air Pollution and Well-being. Sustainable Cities and Society, 103447.

Li, W., Saphores, J.D.M., Gillespie, T.W., 2015. A comparison of the economic benefits of urban green spaces estimated with NDVI and with high-resolution land cover data. Landscape and Urban Planning 133, 105-117.

Liaw, A., Wiener, M., 2002. Classification and regression by randomForest. R news 2, 18-22.

MacKerron, G., Mourato, S., 2013. Happiness is greater in natural environments. Global Environmental Change 23, 992-1000.

McCallum, I., Kyba, C.C.M., Bayas, J.C.L., Moltchanova, E., Cooper, M., Cuaresma, J.C., Pachauri, S., See, L., Danylo, O., Moorthy, I., Lesiv, M., Baugh, K., Elvidge, C.D., Hofer, M., Fritz, S., 2022. Estimating global economic well-being with unlit settlements. Nature Communications 13.

Molnar, C., 2020. Interpretable machine learning. Lulu. com.

Molnar, C., Casalicchio, G., Bischl, B., 2020. Interpretable Machine Learning – A Brief History, State-of-the-Art and Challenges. Springer International Publishing, pp. 417-431.

Oswald, A.J., Wu, S., 2010. Objective Confirmation of Subjective Measures of Human Well-Being: Evidence from the USA. Science 327, 576-579.

Park, J., Kim, H.-J., Lee, C.-H., Lee, C.H., Lee, H.W., 2021. Impact of long-term exposure to ambient air pollution on the incidence of chronic obstructive pulmonary disease: A systematic review and meta-analysis. Environmental Research 194, 110703.

Pirrera, S., De Valck, E., Cluydts, R., 2010. Nocturnal road traffic noise: A review on its assessment and consequences on sleep and health. Environment International 36, 492-498.

Schapire, R.E., 2003. The Boosting Approach to Machine Learning: An Overview. Springer New York, pp. 149-171.

Suk, J.Y., Walter, R.J., 2019. New nighttime roadway lighting documentation applied to public safety at night: A case study in San Antonio, Texas. Sustainable Cities and Society 46, 101459.

Wang, R., Yang, B., Yao, Y., Bloom, M.S., Feng, Z., Yuan, Y., Zhang, J., Liu, P., Wu, W., Lu, Y., Baranyi, G., Wu, R., Liu, Y., Dong, G., 2020. Residential greenness, air pollution and psychological well-being among urban residents in Guangzhou, China. Science of The Total Environment 711, 134843.

White, M.P., Alcock, I., Wheeler, B.W., Depledge, M.H., 2013. Would You Be Happier Living in a Greener Urban Area? A Fixed-Effects Analysis of Panel Data. Psychological Science 24, 920-928.

Yeh, C., Perez, A., Driscoll, A., Azzari, G., Tang, Z., Lobell, D., Ermon, S., Burke, M., 2020. Using publicly available satellite imagery and deep learning to understand economic well-being in Africa. Nature Communications 11.

Zhao, N.Z., Liu, Y., Cao, G.F., Samson, E.L., Zhang, J.Q., 2017. Forecasting China's GDP at the pixel level using nighttime lights time series and population images. Giscience & Remote Sensing 54, 407-425.

Appendix:

Table S1: Descriptive Statistics of Features

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
LS	372,685	3.404	0.936	1	3	4	5
Safe Feeling of Living Environments	372,685	3.036	0.588	1	3	3	4
Sense of Goodness for Living	372,685	3.851	0.880	1	3	4	5
Community Attachment	372,685	3.348	1.020	1	3	4	5
Frequency of High-level Stress	372,685	3.164	1.165	1	2	4	5
Frequency of Low-level Stress	372,685	3.605	1.059	1	3	4	5
Female Dummy (Gender)	372,685	0.351	0.477	0	0	1	1
Age	372,685	49.069	11.745	17	41	57	101
Self-reported Health	372,685	3.255	1.269	1	2	4	5
Annually Individual Income (Million JPY)	372,685	4.710	3.905	1	2.5	6.5	30
College without Diploma	372,685	0.209	0.407	0	0	0	1
Bachelor Dummy	372,685	0.440	0.496	0	0	1	1
Master Dummy	372,685	0.047	0.212	0	0	0	1
PhD Dummy	372,685	0.014	0.116	0	0	0	1
NDVI (%)	372,685	36.276	13.673	4.861	26.285	44.273	87.709
Logarithm of NTL	372,685	2.588	0.968	0.014	1.951	3.323	4.498
Student Dummy	372,685	0.012	0.107	0	0	0	1
Worker Dummy	372,685	0.597	0.491	0	0	1	1
Company Owner Dummy	372,685	0.025	0.157	0	0	0	1
Government Officer Dummy	372,685	0.053	0.224	0	0	0	1
Self-employed Dummy	372,685	0.071	0.257	0	0	0	1
Professional Job Dummy	372,685	0.027	0.163	0	0	0	1

Housewife Dummy	372,685	0.086	0.280	0	0	0	1
Retired Dummy	372,685	0.066	0.248	0	0	0	1
Unemployed Dummy	372,685	0.045	0.208	0	0	0	1