

1 **A Spatial Machine Learning Approach to**
2 **Valuing Development and Greenness in Well-**
3 **Being**

4

5 **Abstract**

6 The positive impact of greenness, represented by normalized difference vegetation
7 index (NDVI), in residential environments on human well-being has been documented and
8 recognized. As a widely used proxy, the nighttime light (NTL) indicates the regional socio-
9 economic status and development level. Higher development levels and economic status,
10 coming with high-level NTL, are related to more opportunities and higher income, but also
11 associated with crowded environments and light pollution. However, the relationships
12 between human well-being and greenness and development level remain inconclusive.
13 Here, we demonstrate the complex nexuses between subjective well-being (SWB) and
14 NDVI and NTL by employing the random forest method. We have optimized the partial
15 dependence analysis and created a novel approach to estimate the local marginal effects of
16 NDVI, NTL, and income. Overall, NDVI is positively associated with SWB, while NTL
17 is negatively linked. Those relationships spatially vary. On average, a one-percentage-point
18 absolute increase in local environment NDVI is associated with a 379.01 USD increase in
19 annual income, while a $1-nW/(cm^2 \cdot sr)$ increase in local NTL is equivalent to a 1434.59
20 USD decline in annual income. These values represent average marginal trade-offs derived
21 from locally non-linear relationships. According to our results, greenness and development

22 improvement should consider the local environments rather than simply formulating one-
23 size-fits-all policies or strategies. To conclude, retaining a moderate development intensity
24 and greening based on the local status are necessary ways to achieve a sustainable society
25 and improve human well-being.

26

27 **Keywords:**

28 Human Well-Being; NDVI; NTL; Random Forest; Geographical Connection

29

30 **Introduction**

31 Greenness and the natural environment are positively associated with human well-
32 being (Li & Managi, 2021b; MacKerron & Mourato, 2013; Pinto et al., 2022). Greenness
33 leads to more recreational and physical activities (Hunter et al., 2015; Li et al., 2025), air
34 pollution reduction (Eitelberg et al., 2016; Li & Managi, 2021a), and relief of stress (Astell-
35 Burt & Feng, 2019; Astell-Burt et al., 2014; Barton & Pretty, 2010). However, it remains
36 unclear whether continuous and uniform increases in greenness always lead to
37 improvements in human well-being. While it appears that more greenery typically
38 correlates with greater well-being (Kley & Dovbischuk, 2024; Sharifi et al., 2021), regions
39 with excessive greenness may lack sufficient land for infrastructure development and
40 economic activities, which could negatively impact human well-being (Piao et al., 2024).
41 For instance, due to the crowded living conditions in Japan, particularly in metropolitan
42 areas, like Tokyo, people also desire more urban land, as indicated by previous studies (Li
43 & Managi, 2021b; Piao et al., 2024). In other words, the effect of greenness on human well-
44 being is not consistent in different places, that is, spatial heterogeneity exists (Li & Managi,
45 2021b; Sharifi et al., 2021). Additionally, recent evidence shows that only large urban
46 green spaces could improve well-being (Sharifi et al., 2021). Certain cities emphasize the
47 importance of balancing well-developed urban land with the preservation and utilization
48 of natural landscapes (Kley & Dovbischuk, 2024; Seresinhe et al., 2019). Moderate-
49 intensity development and population density reduction may offer a solution to the conflict
50 between green space and construction (Zhong & Li, 2024). Moderate-intensity
51 development can support the preservation of adequate natural land, and it allows people to

52 experience a more harmonious and nature-integrated urban lifestyle (Hartig & Kahn, 2016;
53 Krischke et al., 2025; Li & Managi, 2024).

54 Nighttime light (NTL) has commonly served as an indicator of socio-economic
55 development, considering geographical information (Chen & Nordhaus, 2011; Zheng et al.,
56 2022). Previous studies show that highly developed regions and places with large amounts
57 of human activities generally have more illumination (Chen & Nordhaus, 2011; Zhao et al.,
58 2022). Conversely, high-intensity development is associated with various environmental
59 challenges, including light pollution (Falchi et al., 2011; Falchi et al., 2019; Li & Managi,
60 2023b), traffic noise (Begou et al., 2020; Prrera et al., 2010), and deprivation of green
61 views (Kley & Dovbischuk, 2024). For example, previous studies have shown that green
62 window views are positively associated with residential satisfaction, whereas high-density
63 buildings tend to obstruct these views and diminish satisfaction (Kley & Dovbischuk,
64 2024). These environmental stressors have been shown to adversely affect public health,
65 contributing to issues such as sleep disruption, psychological stress, and reduced happiness
66 (Falchi et al., 2011; Kley & Dovbischuk, 2024; Li & Managi, 2024; Prrera et al., 2010).
67 Moreover, brighter light at night is detected by both satellites and individuals who are
68 sleeping or attempting to sleep, which reduces living comfort significantly (Li & Managi,
69 2023b). In highly developed areas, lighting often exceeds functional needs, creating excess
70 illumination that disrupts wildlife and poses risks to human health (Falchi et al., 2011).
71 However, in areas lacking sufficient luminosity, safety and development may be
72 compromised, adversely affecting human well-being (Suk & Walter, 2019). In this study,
73 NTL is not treated as a direct measure of light pollution, but rather as a widely used proxy
74 for development intensity and human activity. Excessive NTL, often a sign of

75 overdevelopment, may harm well-being (Falchi et al., 2019; Li & Managi, 2023b), while
76 a lack of infrastructure can hinder safety and economic activity (Gaston et al., 2015). Just
77 as greenness must be balanced with land use, NTL represents the need to balance
78 development intensity, which supports the idea that sustainable well-being requires
79 context-sensitive integration of both natural and built environments. The connections
80 between human well-being, greenness, and development are expected to be intricate and
81 multi-dimensional, but relatively accurate knowledge is helpful to achieve a sustainable
82 society.

83 Evidence has shown that SWB could objectively represent human well-being
84 (Oswald & Wu, 2010). Three distinct approaches, each emphasizing different aspects of
85 well-being, are broadly recognized and widely used for assessing SWB (Steptoe et al.,
86 2015): life evaluation (Diener et al., 2018; MacKerron, 2012; Oswald & Wu, 2010),
87 hedonic well-being (Kahneman et al., 2004), and eudemonic well-being. The life
88 evaluation method is commonly used in previous studies on environmental issues to detect
89 people's well-being based on their overall thoughts of the life quality (Diener et al., 2018;
90 Ryff, 2014). Hedonic well-being focuses on emotions, and eudemonic well-being
91 concentrates on the sense of life. Hedonic well-being is more suitable for real-time analysis
92 because emotions vary more apparently (Mackerron & Mourato, 2009). The links between
93 eudemonic well-being and living environments are relatively weak (Diener et al., 2018;
94 Mouratidis, 2018; White et al., 2020). In this study, overall happiness covering both life
95 satisfaction and hedonic well-being is considered, which is 5 degrees, from negative (1) to
96 positive (5).

97 Earlier studies often assume a globally linear association between environmental
98 factors and well-being (Ghosh et al., 2013; Li & Managi, 2021b; Pinto et al., 2022), which
99 limits their ability to capture context-specific variations. However, real-world relationships
100 are often non-linear and spatially heterogeneous (Sharifi et al., 2021; Zhong & Li, 2024).
101 For example, insufficient lighting reduces feelings of safety (Suk & Walter, 2019), while
102 excessive lighting leads to discomfort (Falchi et al., 2019; Li & Managi, 2023b). Linear
103 models may fail to capture localized non-linear patterns or spatial heterogeneity, and
104 instead produce average effects that mask important variations across contexts (Li &
105 Managi, 2023b; Yang et al., 2024). In ecological contexts, linear models also tend to show
106 poor model fit (Ambrey & Fleming, 2014; Krekel et al., 2016; Pinto et al., 2022; Tsurumi
107 et al., 2018). To address this, we use a random forest model that does not require linearity
108 assumptions, allowing us to uncover more flexible and locally varying patterns in the data.
109 First, the random forest algorithm by ensembling a batch of decision trees does not require
110 any linearity assumption (Breiman, 2001). Second, in random forests, each decision tree is
111 independent, which allows us to do geographical local analysis, compared with boosting
112 models, including light gradient boosting model (LGBM) (Ke et al., 2017), extreme
113 gradient boosting (XGBoost) (Chen & Guestrin, 2016), and CatBoost (Prokhorenkova et
114 al., 2018). It should be noted that Geographically Weighted Regression (GWR) or GWR-
115 family algorithms, such as multiscale GWR and geographically weighted panel regression,
116 are another potential solution, which could capture global nonlinear relationships through
117 local linear relationships (Li & Managi, 2022; Lu et al., 2014). However, GWR and its
118 family typically assume a linear functional form and require careful tuning of spatial
119 bandwidth parameters (Gollini et al., 2015; Li & Managi, 2022), which may limit their

120 ability to capture complex, nonlinear, and high-dimensional relationships (Yang et al.,
121 2024). However, because the idea of GWR is effective, the local connection in this study
122 draws on its local data partitioning idea. In essence, our approach integrates principles from
123 both machine learning and GWR, combining the flexibility of non-parametric modeling
124 with spatially localized interpretation. Here, we utilize over 380,000 observations to
125 explore the links between human well-being and environmental factors, including
126 greenness and development. To make the results understandable, we combine geographical
127 division by the random forest model and build a geographically local explanation. This
128 innovative approach increases the practicality and interpretability of the random forest
129 algorithm in geoscience. The computation workflow is summarized in **Figure 1**, which
130 includes three modules, namely, data preparation, machine learning and local analysis, and
131 interpretation and valuation.

132

133 Materials and Methods

134 *Materials*

135 *Survey Details*

136 Between 2015 and 2017, our research team conducts three nationwide surveys in
137 Japan, targeting a randomly selected sample of 300,000 participants, proportionally
138 representing the country's demographic and geographical distributions. The surveys are
139 executed by Nikkei, Inc., and the questionnaire is designed by a professional team, which
140 includes hundreds of questions covering multiple aspects, including economy,
141 environment, and demographic and socio-economic factors, such as mental health, self-

142 reported health, age, gender, occupation, educational level, income, and more. Nikkei Inc.
143 is a highly reputable and established media and research organization in Japan, which is
144 experienced in conducting nationwide surveys and public opinion polls, often in
145 collaboration with academic and governmental institutions. The investigation by a
146 professional team ensures randomness and reliability. We gather over 450,000 validated
147 responses through three survey waves, which is the “Data Collection” step in **Figure 1**.
148 Including sensitive topics like education level, personal income, and employment status in
149 our internet-based survey potentially reduces the chances of receiving dishonest answers
150 from participants (Chapman et al., 2019; Li & Managi, 2021c). Under the specific laws
151 and with local government authorization, our surveys are conducted within a fixed four-
152 week period each November. While laws and privacy regulations prevent the direct
153 collection of detailed addresses from respondents, we are allowed to gather their residential
154 postal codes. These respondents are from 50,852 postal zones. For our analysis, we utilize
155 the geometric centers of postal zones, as provided by the Google Geocoding API. In
156 defining variables for the model, which is the “Variable Definition” step in **Figure 1**, two
157 key considerations are balanced. On the one hand, including a wider range of variables can
158 improve predictive accuracy by capturing more relevant factors. On the other hand, adding
159 more variables increases the likelihood of missing data, which can reduce the size and
160 quality of the usable dataset. Therefore, we choose variables that are both theoretically
161 meaningful and consistently available to ensure reliable model performance without
162 sacrificing data completeness. After excluding entries with missing data, primarily due to
163 participants selecting “refuse to answer” or “unknown”, our dataset retains over 380,000
164 records, which is the “Data Cleaning” step in **Figure 1**. The final dataset includes complete

165 responses for all variables used in modeling. In the three years, there are 202,331, 109,930,
166 and 70,912 respondents with completed data rows, respectively, where the same
167 respondents are tracked over time.

168

169 *Human Well-Being*

170 SWB is an individual's evaluation of their lives based on their own judgment and
171 experience (Diener, 1984; Diener et al., 2018). The objective connection between SWB
172 and living conditions has been well documented in the literature (Oswald & Wu, 2010).
173 This study uses happiness, which is widely employed and investigated in SWB-related
174 studies (Diener, 1984; Diener et al., 2018; MacKerron, 2012). To measure people's
175 happiness, we ask the respondents to answer the following question: "overall, how happy
176 are you with your life?" and the potential answers for happiness-related questions are from
177 "very happy (5)" to "very unhappy (1)". It is important to emphasize that happiness scores
178 are qualitative and ordinal in nature, rather than quantitative. In practice, SWB scales,
179 including life satisfaction and happiness, often behave similarly to interval-level variables.
180 Previous studies (Li & Managi, 2021b, 2021c, 2024; MacKerron & Mourato, 2013) have
181 treated SWB as continuous in applied econometric and machine learning literature,
182 especially when the variable has more than three ordered categories. Traditional
183 econometric methods, such as ordered logit and probit models, are widely used to analyze
184 well-being measures due to their latent variable framework and interpretable structure
185 (Astell-Burt et al., 2014; MacKerron, 2012). In multi-valued models, the need to estimate
186 and interpret multiple probability functions can make the results more complex to present
187 and understand (Li & Managi, 2025). For this reason, reclassifying multiple-value output

188 into binary output is another solution often used by previous studies (Astell-Burt et al.,
189 2014; Triguero-Mas et al., 2015). Furthermore, in high-dimensional and spatially
190 heterogeneous contexts, these models can become difficult to specify and compare across
191 space. Machine learning algorithms do not require the outcome variable to be strictly
192 continuous, but artificial neural networks and boosting algorithms also need to establish a
193 softmax function to output multiple classifications, which could not effectively solve the
194 difficulty in the explanations. Bagging methods, including random forest, could avoid the
195 multi-link-function issue since random forest is to aggregate the outputs of each tree. In
196 random forest algorithms, the primary difference between regression and classification
197 tasks lies in the aggregation method: regression uses averaging to combine the predictions
198 of individual trees, whereas classification uses majority voting to determine the final output
199 (Breiman, 2001). The random forest performs well in approximating complex, non-linear
200 relationships, even when the outcome is ordinal (Li & Managi, 2024; Yang & Zou, 2025).
201 Li and Managi (2024) show that random forest outperforms ordered logistic regression,
202 OLS, XGBoost, LGBM, support vector machines, and neural networks in predicting the
203 ordinal GHQ12 mental health scores. Considering the interpretability and flexibility, our
204 analyses are regression tasks. **Figure 2.a** depicts the SWB statistical distribution. The
205 spatial distribution of zonal average happiness is illustrated in **Supplementary Materials**
206 **Figure S1**. The zonal average represents the mean of all respondents' answers within a
207 given postal zone.

208

209 *Remote Sensing Data*

210 To examine the impact of greenness on human well-being, we use the 16-day Level
211 3 Normalized Difference Vegetation Index (NDVI) dataset. This dataset, produced by
212 NASA, provides a 500-meter spatial resolution. This dataset has been widely utilized in
213 previous environmental and demographic research (Lamchin et al., 2018; Li & Managi,
214 2023a; Wang et al., 2020). The NDVI acts as a graphical index reflecting the presence of
215 live green vegetation within a given pixel, with values ranging from -100%, indicating non-
216 vegetated surfaces, such as snow, or ice, to 100%, signifying a high abundance of live green
217 vegetation (Didan et al., 2015). Generally, an NDVI value of 0 typically indicates the
218 presence of bare surfaces with little to no vegetation, such as bare land, rocks, or some
219 urban surfaces. The NDVI data are derived from NASA's MODIS Terra and Aqua
220 satellites, specifically from the products version 6.1 MOD13A1
221 (<https://lpdaac.usgs.gov/products/mod13a1v061/>) and MYD13A1
222 (<https://lpdaac.usgs.gov/products/myd13a1v061/>), respectively. Our study concentrates on
223 the months of October to December each year, aligning with the timing of our survey
224 during each study period. NASA's MOD13A1 and MYD13A1 products offer a 16-day
225 temporal resolution, providing data for the 289th, 305th, 321st, 337th, and 353rd days of
226 the year. To streamline the analysis, we combine the 16-day observations from both the
227 Terra and Aqua satellites into a unified raster for each year. Using the geometric centers of
228 postal zones as reference points, we construct 5-km radius buffers around them. The
229 average NDVI values within these buffers are calculated to represent the natural greenness
230 level of each postal zone. **Figure 2.b** illustrates the NDVI statistical distribution. The
231 spatial distribution of zonal average NDVI is illustrated in **Supplementary Materials**

232 **Figure S2.** It should be noted that in our study, no area had negative NDVI values, so the
233 statistical range we show is from 0 to 100%.

234 Human well-being is influenced by local economic conditions, human activities,
235 and infrastructure quality (Chen & Nordhaus, 2011, 2019; Oswald & Wu, 2010; Piao et al.,
236 2024). NTL data, obtained from the Suomi National Polar-orbiting Partnership Visible
237 Infrared Imaging Radiometer Suite (NPP-VIIRS), are widely utilized to measure human
238 activity and economic conditions (Chen et al., 2021; McCallum et al., 2022). Artificial
239 electric lighting is prevalent in most buildings and infrastructures, making it a reliable
240 indicator for such analyses (Zheng et al., 2022). The Earth Observation Group provides
241 monthly NTL data at a spatial resolution of 1 kilometer. Each year, we compute the average
242 of the NTL rasters for October, November, and December, and subsequently derive the
243 mean NTL values within 5-kilometer radius buffers. In other words, each NTL value
244 represents the average radiance within a 5-kilometer radius buffer, which is an area
245 covering approximately 80 km² around each postal zone centroid. The NTL measurements
246 are expressed in nanowatts per square centimeter steradian ($nW/cm^2 \cdot sr$), which reflect
247 the intensity of light emitted or reflected from the Earth's surface as detected by satellite
248 sensors. While this unit does not translate directly to what the human eye perceives, such
249 as lumens, it is well-established in remote sensing and urban studies as a proxy for
250 nighttime human activity and built-up development. **Figure 2.c** illustrates the NTL
251 statistical distribution. The spatial distribution of zonal average NTL is illustrated in
252 **Supplementary Materials Figure S3.**

253 While NDVI data have a spatial resolution of 500 m and NTL data are at 1 km, both
254 are aggregated to the same spatial unit, 5 km radius buffers centered on postal code

255 centroids. This spatial averaging minimizes the resolution mismatch and ensures
256 comparability across variables. Given the larger aggregation area and the smoothing effect
257 of averaging, we do not expect the resolution difference to introduce substantial bias into
258 our analysis.

259

260 *Demographic and Socio-economic Features*

261 Several demographic and socio-economic factors are accounted for, consistent with
262 prior research (Li & Managi, 2021b, 2021c). As our study examines the relationship
263 between human well-being and living environments, we place significant emphasis on
264 respondents' perceptions of their surroundings, including their sense of safety, satisfaction
265 with living conditions, and feelings of community connection (Li & Managi, 2025). The
266 frequency of experiencing high and low levels of stress, the ability to relax, and self-
267 reported health are indicators used to assess health status in our study (Li & Managi, 2024).
268 In our analysis, we also consider gender, age, annual income, educational background, and
269 employment status (Li & Managi, 2023c). It is worth mentioning that the annual income
270 reported in the survey reflects an income range rather than an exact amount. Individual
271 income includes 14 options. The first 13 are specific income ranges, and the last option is
272 to refuse to answer. If respondents choose the last option, the data is not included in the
273 analysis. The specific thirteen options are: 1, no more than 2 million JPY; 2, 2 – 3 million
274 JPY; 3, 3 – 4 million JPY; 4, 4 – 5 million JPY; 5, 5 – 6 million JPY; 6, 6 – 7 million JPY;
275 7, 7 – 8 million JPY; 8, 8 – 9 million JPY; 9, 9 – 10 million JPY; 10, 10 – 15 million JPY;
276 11, 15 – 20 million JPY; 12, 20 – 30 million JPY; and 13, more than 30 million JPY. The
277 first twelve options are the midpoints of these ranges: 2 million JPY, 2.5 million JPY, 3.5

278 million JPY, 4.5 million JPY, 5.5 million JPY, 6.5 million JPY, 7.5 million JPY, 8.5 million
279 JPY, 9.5 million JPY, 12.5 million JPY, and 17.5 million JPY. The thirteenth option is set
280 as 30 million JPY. **Figure 2.d** illustrates the income statistical distribution. The spatial
281 distribution of zonal average income is demonstrated in **Supplementary Materials Figure**
282 **S4**. It should be noted that all questions must be answered and if the data contains any
283 missing data it will be deleted.

284 In total, 29 features, specifically “Individual Income”, “NDVI”, “NTL”, “Year”,
285 “Latitude”, “Longitude”, “Female Dummy”, “Age”, “Frequency of High-level Stress”,
286 “Frequency of Low-level Stress”, “Easy to Relax”, “Sense of Goodness for Living”, “Safe
287 Feeling of Living Environments”, “Community Attachment”, “Income Level”, “Self-
288 reported Health”, “Student Dummy”, “Worker Dummy”, “Company Owner Dummy”,
289 “Government Officer Dummy”, “Self-employed Dummy”, “Professional Job Dummy”,
290 “Housewife Dummy”, “Retired Dummy”, “Unemployed Dummy”, “College without
291 Diploma”, “Bachelor Dummy”, “Master Dummy”, and “PhD Dummy”, are utilized to
292 construct the models, with descriptive statistics for these features provided in
293 **Supplementary Materials Table S1**. The specific survey questions and the range of
294 responses are summarized in **Supplementary Materials Table S2**.

295

296 **Methods**

297 *Random Forest*

298 Random forest algorithm is employed because it can flexibly model complex, non-
299 linear relationships without requiring strong parametric assumptions. Unlike traditional

300 linear models, random forests accommodate spatial heterogeneity and variable interactions,
301 making them well-suited for analyzing geographically diverse data. Specifically, longitude
302 and latitude can be directly incorporated as input variables in random forests and can
303 interact flexibly with other features (Li & Managi, 2023b, 2024; Su et al., 2015). In contrast,
304 handling the spatial coordinates in this way in linear regression models is more problematic
305 because changes in the geographical environment are usually not linear (Li & Managi,
306 2023b). A summary of decision trees and random forest structures is listed in
307 **Supplementary Materials**. In the “Hyperparameter Search” step in **Figure 1**, the
308 verification is based on 10-fold cross-validation, and the metric is the average validation
309 R^2 in each fold of cross-validation. In 10-fold cross-validation, the dataset is randomly
310 divided into ten equal sub-datasets; the model is trained on sub-datasets parts and tested on
311 the remaining one; and then the training-test process is repeated ten times by using different
312 training and test datasets. The final performance metric, namely average validation R^2 , is
313 the average of these validation results, providing a robust estimate of the model’s
314 generalization ability. This makes average validation R^2 an appropriate metric, as it reflects
315 the model’s predictive accuracy on unseen data while minimizing the risk of overfitting to
316 any particular subset. Since we test 500 hyperparameter combinations, a total of 5000
317 independent models are trained and tested during the hyperparameter search phase. The
318 hyperparameter searching process is conducted on a machine with an NVIDIA 4090 GPU
319 and Python 3.9.12. Scikit-learn does not support GPU acceleration, so we use NVIDIA-
320 developed “rapids” to conduct the process. Even with a GPU, the search process costs
321 roughly 200 hours to obtain the best model. To ensure efficiency, we randomly perform
322 three folds out of 10 folds. The search process could be written as follows:

$$Hyperparameter_{best} = \theta(\mathbf{X}, \mathbf{y}, rf, hyperparameters, m) \quad (1)$$

323 where $Hyperparameter_{best}$ represents the hyperparameter setting to achieve the best
 324 performance in the cross-validation, \mathbf{X} and \mathbf{y} are independent and dependent variables, rf
 325 is the basic random forest algorithm, *hyperparameters* are the potential combinations
 326 based on the potential hyperparameter values, and m is the number of iterations.

327

328 *Local Partial Dependence Analysis*

329 Partial dependence analysis (PDA) is widely employed to explain the machine
 330 learning model (Molnar, 2020). The predictions of decision trees are based on a series of
 331 judgments, and random forests are composed of a series of different decision trees.
 332 Therefore, a simple and intuitive explanation is basically impossible. PDA treats the
 333 random forest model as a black box and compares the difference in the predictions based
 334 on slight adjustments in the feature of interest (Friedman, 2001). Since the results of PDA
 335 are always illustrated by plots and the major programming languages in statistics, such as
 336 Python and R, mainly name the relevant packages as “partial dependence plot”, the partial
 337 dependence plot (PDP) is relatively widely known. The PDA’s process is to predict all
 338 instances by setting a feature of interest to a certain value, average the predictions, and then
 339 gradually change the value of the feature. The range of this set value is equal to the value
 340 domain of the feature in the dataset, and the values are generally quantiles of the domain.
 341 The calculation could be expressed as follows:

$$PD_{fq} = \frac{1}{n} \sum_n rf_{best}(\tilde{\mathbf{X}}), \mathbf{X}^f = q \quad (2)$$

342 where PD_{fq} is the average prediction of the total dataset with setting the feature f equal to
343 q , n is the number of instances in the total dataset, rf_{best} is the well-trained random forest
344 model based on the best hyperparameter combination, $\tilde{\mathbf{X}}$ is the total dataset after
345 adjustments, and $\mathbf{X}^f = q$ represents setting all instances' feature f equal to q . The training
346 process of the model rf_{best} is the “Modelling” step in **Figure 1**.

347 PDA can explain and analyze models quickly and efficiently, but its shortcomings
348 are also significant. First, it is not reasonable to set a feature of all instances to a specific
349 value (Molnar, 2020). For example, under this study’s background, in some rural areas, the
350 NDVI value in the respondents’ residential areas could be 80%, but this value is impossible
351 to be achieved in the urban core of Tokyo. Here, strictly controlling the value range is a
352 solution. Additionally, another key approach is to divide local data sets according to
353 geographical location instead of performing calculations on the entire set. We innovatively
354 propose a geographical clustering method based on random forest segmentation to establish
355 local datasets. The reason for this local division is to ensure the variation of geographically
356 related variables, namely NTL and NDVI. If we divide the local dataset by postal zone, the
357 NDVI and NTL data will only exhibit temporal variation, but ignore spatial variation.
358 Independent variables include the geographical coordinates of the respondents, so most
359 decision trees have a large number of judgments based on geographical coordinates. In
360 other words, a decision tree divides geographical space into many segments. Theoretically,
361 this geographical division process is reasonable, as instances with similar characteristics
362 are typically more likely to be grouped into the same branch during tree-based partitioning.
363 Furthermore, in a random forest, each tree is trained independently on a bootstrapped
364 subset of the data, and splits are made on different combinations of features, often including

365 spatial variables, namely latitude and longitude. Due to randomness in the bootstrapping
 366 process, the geographical segments divided by each tree are different. It should be noted
 367 here that the independence of the trees generated by random forest is relative to the
 368 boosting algorithm because the training data of the next tree in the boosting algorithm is
 369 based on the residual of the previous tree (Chen & Guestrin, 2016). We require that trees
 370 above average are accepted within a certain geographical range. In this way, we obtain all
 371 segments divided by each tree, obtain the boundaries of an instance from each tree, and
 372 then take the medians of the boundaries as the geographical extent of the instance's
 373 neighborhood. This approach is primarily motivated by the requirements of robustness,
 374 ensuring that the majority of trees agree on such a division. Respondents within this range
 375 will be treated as local data for the instance's local PDA, which is the step "Local Dataset
 376 Partitioning" step in **Figure 1**. The local PDA in **Figure 1** could be illustrated as follows:

$$LPD_{ifs} = \frac{1}{n_i} \sum_{n_i} r f_{best}(\widetilde{\mathbf{X}}_i), \quad \widetilde{\mathbf{X}}_i^f = \mathbf{X}_i^f + s \quad (3)$$

377 where LPD_{ifs} represents the local partial dependence value for the instance i after that all
 378 neighboring instances in the instance i 's local dataset adjusted their feature f by adding s ,
 379 n_i is the number of instances in the local dataset of the instance i , $\widetilde{\mathbf{X}}_i$ is the local dataset
 380 after adjustments, the adjustment, $\widetilde{\mathbf{X}}_i^f = \mathbf{X}_i^f + s$, represents setting all neighboring
 381 instances' feature f to increasing s , $\widetilde{\mathbf{X}}_i^f$ represents the adjusted values of the instances'
 382 feature f , and \mathbf{X}_i^f represents the values before adjustments. s is a list of pre-defined values
 383 ranging from 0 to 2 with a 0.05 stride.

384 In terms of the research object of this study, the feature f includes three potential
 385 variables: NDVI, NTL, and income. The pre-defined values of s are considered carefully

386 based on the characteristics of these three variables. In our current design, each respondent
387 receives 41 independent estimates based on 41 potential s values to enhance the robustness
388 of the geographically localized analysis. Some postal zones' NTL values are below 1
389 nW/(cm²·sr). Negative NTL values are not reasonable in the real world. To mitigate this
390 issue, we limit the range of feature changes to increments between 0 and 2 units. For the
391 current NDVI, NTL, and income data distributions, a 2-unit variation is equivalent to a
392 change of 0.15, 0.13, and 0.51 standard deviations, respectively. The practical meaning of
393 the 2-unit increase in each variable could be explained as follows: for NDVI, a 2-unit
394 increase can mean a two-percentage-point increase in the green area while keeping the
395 vegetation composition unchanged, or a higher density of vegetation while keeping the
396 area unchanged, or a combination of the two; for NTL, it means increased nighttime
397 lighting in the area; and for income, it represents a 2-million-JPY increase in personal
398 income. Furthermore, the boundary or extreme values are carefully considered. Increasing
399 NDVI and NTL does not affect the predicted SWB of observations that are already at the
400 maximum value of the feature. Random forest is based on binary calculation, and
401 increasing the feature values to exceed the original maximum value will not cause branch
402 changes, so the final output prediction value will not change. Combined with **Figure 2**,
403 these extreme values are rare. In addition, in local PDA calculations, they do not act alone,
404 but only as a small part of the geographically local data set. Specifically, the maximum
405 NDVI appears in remote areas. Due to the setting of random forest hyperparameters, these
406 observations must form a local geographical dataset with other postal zones. The maximum
407 value of NTL appears in the most developed center of Tokyo, where postal zones are denser
408 due to dense population and space. Local datasets there usually have the most complex

409 spatial variation. For income, its situation is similar to NTL, with the highest income
410 usually sparsely distributed in the city center. Combined with surrounding observations,
411 the local datasets could ensure the robustness of the estimated effects.

412 We connect the change values s and the values of Local PDA through the
413 geographically local linear regression, which is the “Local Effect Estimation” step in
414 **Figure 1**. This step is a local linear interpretation of the global nonlinearity. This
415 connection is intended to estimate the marginal effect of changes in a variable of interest
416 on human well-being. The connection could be expressed as follows:

$$LPD_{ifs} = \alpha_{if}s + \beta_{if} + e_{if} \quad (4)$$

417 where α_{if} and β_{if} are the slope and the intercept of the local connection of an increase in
418 the feature f in i 's neighbor zone. α_{if} could be regarded as the local marginal effect. For
419 example, if feature f is NDVI, α_{if} could be explained as the change in human well-being
420 with a one-percentage-point increase in NDVI value in i 's neighbor zone. It should be
421 stated that the changes in NDVI in this article are all changes in absolute values, such as
422 an increase from 30% to 31%. For each different i , their relationships are not exactly the
423 same due to the changes in the neighborhood space.

424

425 *Monetary Values of Features*

426 To enhance the comparability and understanding of how particular features
427 influence SWB, we convert their effects into monetary values by applying the marginal
428 substitution rate (MSR) between the feature of interest and income (Krekel et al., 2016; Li
429 & Managi, 2021b, 2021c), which is the “Monetary Value Estimation” step in **Figure 1**. In

430 other words, any SWB change stemming from a one-unit increase in a given feature is
431 balanced by an equivalent adjustment in income. Because local PDA value reflects the
432 contribution of a certain feature change, a 1-unit increase will alter SWB accordingly. If
433 we aim to keep SWB constant, then the required income shift to offset that change is
434 considered the feature's monetary value. It should be mentioned that the monetary value
435 estimations are derived under the assumption that SWB is continuous, which is inherited
436 from local marginal effect estimation. Formally, we estimate this monetary value of the
437 feature as follows:

$$MSR_{if} = \frac{\alpha_{if}}{\alpha_{iINC}} \quad (5)$$

438 where MSR_{if} is the MSR of the feature f in the zone i , and α_{iINC} is the local marginal
439 effect of income on SWB in zone i , and α_{if} is local marginal effect of feature f on SWB
440 in zone i . The MSR here could be regarded as the marginal willingness to pay (MWTP)
441 (Li & Managi, 2021b, 2021c).

442 The MSR calculation is based on locally linear approximations of the partial
443 dependence curves. However, we pay more attention to the investigation of local
444 relationships. Specifically, our method estimates marginal effects within geographically
445 defined local datasets, where the assumption of linearity is applied over small ranges. The
446 MSR values are thus derived from local tangents, not global slopes, ensuring that the
447 income-environment tradeoff reflects localized and smoothed interpretations of non-linear
448 relationships. This way reduces the risk of oversimplifying complex tradeoffs and allows
449 monetary valuation to retain sensitivity to spatial variation. All local MSRs' standard errors
450 are estimated through the delta method based on local marginal effects of NDVI, NTL, and
451 income, and their standard errors estimated by local connections. Non-zero MSR values

452 are required to have absolute magnitudes at least 1.96 times greater than their
453 corresponding standard errors, ensuring statistical significance at the 95% level.

454

455 **Results**

456 ***Hyperparameter Combination and Model Performance***

457 Based on the comparison among 500 random combinations, the best model's
458 hyperparameters are 3,900 trees, 8 for the minimum number of instances in nodes, 0.5 for
459 the ratio of total instances for training each tree, 0.4 for the ratio of total features used in
460 each sub-dataset, and 16 for the maximum depth of a tree, respectively. The specific
461 explanation is as follows: our random forest has 3,900 decision trees; if a node has less
462 than 8 instances, it would not be further divided no matter how deep the current tree it is;
463 the instance number of the bootstrapped sub-dataset for a single decision tree is equal to
464 40% of the total training dataset; and 50% features are used for single sub-dataset; and the
465 max depth of the each tree is 16. The validation R^2 of three random folds is 38.40%, 37.84%,
466 and 37.93%, respectively. The average validation R^2 is 38.05%, and the standard deviation
467 is 0.25%. The model's performance is stable in the test. The training R^2 of the three
468 corresponding folds is 51.51%, 51.57%, and 51.89%, respectively, and the mean and
469 standard deviation of the training process are 51.56% and 0.03%, respectively. Overfitting
470 exists but is not serious, since the validation performance is close to the training
471 performance. With an average validation performance of 38.05%, our research model
472 outperforms most happiness research models (Krekel et al., 2016; Li & Managi, 2021c;
473 MacKerron & Mourato, 2013). Specifically, in Krekel et al. (2016)'s, Li and Managi

474 (2021c)'s, and MacKerron and Mourato (2013)'s studies, their accuracy metrics are 5.75%,
475 36.0%, and 13.5%, respectively, which are R^2 . The R^2 used by previous researchers comes
476 directly from the fitting process, which is the training R^2 in this study. In other words, the
477 generalization ability of these models is unknown, that is, it remains unclear whether the
478 model performs well on individuals who meet the sampling conditions but were not
479 included in the sample.

480

481 ***Local PDA Results***

482 **Table 1** summarizes the local marginal effect coefficients of three variables,
483 namely NDVI, NTL, and income, on happiness. We summarize the data at two levels:
484 individual and location levels. Each location has several individuals' responses. When
485 dividing each local dataset, for each individual, their local datasets are the same. Therefore,
486 for each individual at the same location, their local marginal effect coefficients should be
487 the same. In general, NDVI and income growth can bring a certain amount of happiness,
488 while NTL growth will reduce happiness, although these relationships vary locally. We
489 classify these marginal effects into three categories: negative, not monotonic, and positive
490 effects. Specifically, when the effect coefficients are larger than 0, they would be classified
491 as positive effects; when the effect coefficients are smaller than 0, they would be classified
492 as negative effects; if the R^2 is smaller than 50%, the coefficients would be set as 0, and
493 they are classified as not monotonic effects. A less-than-50% local R^2 means that the linear
494 correlation between the change of the variable of interest and the change of happiness
495 caused by it is generally less than 70%. Therefore, we regard that this connection cannot
496 be described as a monotonic linear relationship, or at least this relationship is unstable. It

497 should be noted that all local marginal effect coefficients with a more-than-50% local R²
498 are statistically significant at the 95% level. As NDVI increases, happiness tend to increase
499 in many areas, in approximately 47.05% of all surveyed postal zones. As NTL increases,
500 70.18% of locations are negatively impacted. As income increases, people in 97.64% of
501 postal zones tend to feel happier.

502 **Figure 3** presents the spatial distribution of NDVI local marginal effect coefficients
503 across Japan, capturing the localized relationships between greenness and happiness. Using
504 a color gradient that transitions from blue, representing negative coefficients, to red,
505 representing positive coefficients, the map highlights significant spatial variations. To
506 demonstrate the densely populated areas clearly, six major metropolitan areas in Japan are
507 zoomed in, which are Sapporo, Sendai, Tokyo, Nagoya, Osaka, and Fukuoka from North
508 to South. While red-dominant regions indicate areas where increased greenness strongly
509 correlates with improved SWB, blue regions represent areas where greenness shows
510 negative contributions. This spatial heterogeneity reflects the complex interplay between
511 greenness and socio-economic factors. We find that the local marginal effect coefficient
512 has little correlation with the absolute value of the corresponding greenery state, which is
513 only -0.173, suggesting a weak linear correlation between the existing greenness level and
514 its local marginal effect. Here, the correlation is the Pearson correlation coefficient between
515 the local marginal effect coefficients of NDVI and the average NDVI values within each
516 corresponding 5 km buffer zone. In other words, the marginal effect coefficient is not
517 highly correlated with the corresponding NDVI from a statistically global scale, supporting
518 the need for localized analysis over uniform policy recommendations. From the spatial
519 distribution point of view, the positive relationship between greenery and happiness usually

520 appears in the urban areas, while in the suburban or rural areas, the relationship is negative.
521 **Figure 4** maps the local marginal effects of NTL on happiness, with blue areas indicating
522 negative effects and red showing positive ones. Negative effects dominate, particularly in
523 urban regions, highlighting issues such as crowded environments and light pollution.
524 However, in some rural areas, NTL contributes positively, likely improving the safety and
525 development. **Figure 5** illustrates the spatial distribution of local marginal effects of
526 improved, showing the relationship between income and well-being. Positive coefficients,
527 marked in red, dominate in most locations, indicating that higher income contributes
528 significantly to subjective well-being. This distribution underscores the general importance
529 of income in shaping well-being, and the local marginal effects spatially vary. **Figure 6**
530 demonstrates the distribution of the local marginal effect coefficients of NDVI, NTL, and
531 income on happiness.

532

533 ***Monetary Values***

534 **Figure 7** illustrates the monetary value of a one-percentage-point absolute increase
535 in NDVI on happiness, converted into 2015 US dollars. All local non-zero MWTP
536 estimates are statistically significant based on the statistical tests using their corresponding
537 delta-method standard errors. We report all monetary values in USD to enhance
538 comparability with international studies and ensure clarity for researchers across different
539 countries. The distributions of NDVI's monetary values align with the geographical
540 average effects of NDVI, as the income effects remain consistent in direction with minimal
541 numerical variations. The average monetary value of NDVI on happiness across all
542 observed locations is 379.01 USD/percentage point, and the monetary value distribution is

543 shown in **Figure 6.d**. Urban areas, particularly around metropolitan regions such as Tokyo
544 and Osaka, exhibit higher monetary values, showing strong positive impacts of the
545 increased NDVI on happiness. In contrast, some rural and mountainous regions, especially
546 in northern Hokkaido and parts of Kyushu, show lower or even negative values, where
547 increased greenness may not align with local priorities or economic activities. **Figure 8**
548 demonstrates the monetary value of a $1\text{-nW}/(\text{cm}^2 \cdot \text{sr})$ increase in NTL on happiness, and all
549 local non-zero MWTP estimates are statistically significant based on the statistical tests.
550 The average monetary values of the increased NTL are $-1434.59 \text{ USD}/(\text{nW}/(\text{cm}^2 \cdot \text{sr}))$ for
551 happiness, and the value distribution is illustrated by **Figure 6.e**. Urban areas with intense
552 NTL, such as Tokyo, Osaka, and Nagoya, generally show more negative impacts.
553 Meanwhile, some rural and less developed regions, particularly in the southern islands,
554 show occasional positive values, indicating the potential benefits of development in
555 improving well-being in these areas. This regional heterogeneity underscores the varying
556 economic and social dynamics influencing well-being across Japan.

557 While the monetary values demonstrated in **Figures 7** and **8** offer intuitive
558 interpretations, they are not global averages derived from a linear model. Instead, they
559 reflect locally estimated MSR based on geographically tailored partial dependence curves.
560 Each estimate accounts for localized, potentially non-linear responses to changes in NDVI,
561 representing local environmental greenness level, and NTL, indicating development
562 intensity. The estimated monetary value is calculated from the slope of the tangent line of
563 the nonlinear relationship. To help readers better understand the non-linearity of these
564 effects, we visualize the distribution of local marginal coefficients and present maps
565 showing both the direction and intensity of relationships, including **Figures 3 - 5** and

566 **Supplementary Materials Figures S1 – S4.** These plots and computation processes
567 clarify that, although average monetary values can be calculated, the underlying functional
568 form is spatially variable and data-driven.

569

570 **Discussion**

571 This study is the first study combining the random forest model, local PDA, and
572 local marginal effects of the variable of interest. We directly take the observations' location
573 information, including longitude and latitude, into the random forest model, which is an
574 important attempt to apply machine learning to a geographical topic. The “random”
575 division of geographical extent in the random forest provides a novel approach to defining
576 the neighbor zones, different from the traditional way based on distance or sharing
577 boundary (Beenstock & Felsenstein, 2019; Fotheringham et al., 2002). The “random”
578 division could guarantee that the people with similar features living close to each other are
579 put into the same group, namely, the end leaf in random forest terminology. We optimize
580 the PDA method by considering the local environments to build geographically local
581 explanations. Compared with a global linear connection, the geographically local link
582 avoids assuming a strong monotonic relationship globally and accepts that the relationship
583 could spatially vary. Finally, our method demonstrates the spatial variation of the local
584 marginal effects and MWTPs.

585 Our key findings reveal that the relationship between SWB and greenness is
586 generally positive, both on average and across most locations. In contrast, higher levels of
587 NTL are associated with a negative impact on subjective happiness. There is ample

588 evidence from previous studies to support the positive association between greenness and
589 human well-being (Li & Managi, 2021b; MacKerron & Mourato, 2013; Tsurumi et al.,
590 2018). People do benefit from a rational level of greenness. Specifically, the amount of
591 green space should be able to meet people's needs for relaxation in life, but not so much
592 lead to a shortage of construction land. On the one hand, greenness could provide
593 recreational activities (MacKerron & Mourato, 2013), air pollution reduction (Chameides
594 et al., 1988; Eitelberg et al., 2016; Mendoza-Ponce et al., 2018), and the creation of
595 aesthetic, artistic, and scientific values for human beings (Felipe-Lucia et al., 2018;
596 Seresinhe et al., 2015). On the other hand, when housing or infrastructure needs are
597 difficult to meet, green space has a relatively low priority, which is consistent with
598 Maslow's hierarchy of needs (Maslow, 1943). For example, in some metropolitan areas,
599 including Tokyo and Nagoya, the increased greenness is negatively associated with the
600 SWB indicators. Intuitively, people in metropolitan areas with less greenness should desire
601 greenness. The possible explanation is that convenient transportation in metropolitan areas
602 increases the accessibility of the natural environment, which fulfills the need for greenness.
603 While increased greenness may not directly diminish human well-being, a notable
604 insufficiency in land allocated for construction and development causes negative impacts.
605 Previous study indicates that the people in Tokyo desire more urban land due to crowded
606 living environments, and increasing greenness in remote areas also cannot have positive
607 effects on SWB (Li & Managi, 2021b). Furthermore, another recent study proposes that
608 only large parks in cities could significantly improve human well-being, so an impactful
609 change in greenness is hard to achieve (Sharifi et al., 2021). In a way, the direct impacts of
610 greenness on human well-being are often overstated, while its indirect effects remain

underexplored or overlooked. Our findings demonstrate that the relationship between greenness and well-being is spatially heterogeneous and context-dependent. The context could include regional culture, geographical conditions, climate environment, and development level, consistent with previous studies (Li et al., 2025; Sharifi et al., 2021). These insights suggest that uniform greening policies may not always be effective and underscore the importance of geographically tailored planning (Sharifi et al., 2021; Zhong & Li, 2024). Although the local marginal effects of greenness are modest in monetary terms, the effects quantify individuals' evaluation of the environmental changes, which supports balanced decisions about land use (Li & Managi, 2021b, 2023b, 2024). This evidence contributes to the broader discourse on sustainable urban development by illustrating that the optimal level of greenness should be aligned with local socio-economic conditions and needs. Rather than simply increasing green areas, effective urban sustainability efforts should seek to integrate nature in ways that maximize well-being while preserving space for essential infrastructure and housing. Thus, sustainable urban development should strive to strike a context-sensitive balance between natural and built environments, a principle aligned with Sustainable Development Goals 3, "Good Health and Well-being" and 11 "Sustainable Cities and Communities". For the purposes of this study, if the local marginal effect of NDVI is positive within a postal area's buffer, then increasing the mean NDVI by up to two percentage points is likely to yield benefits. Here, an increase with two percentage points is considered a moderate range. Conversely, if the local marginal effect of NDVI is negative, increasing greenness in that area may be inadvisable.

NTL serves as a key indicator of economic status and human activities (Chen & Nordhaus, 2011, 2019; Chen et al., 2017; Li & Managi, 2023a; Zhao et al., 2017). A strong

positive correlation between luminosity data and regional GDP could be detected (Chen & Nordhaus, 2011). Brighter areas tend to have higher levels of development and better economic status. In highly developed regions, there are more opportunities and relatively higher incomes, contributing to greater human well-being (Jebb et al., 2018; Li & Managi, 2023c). Our analysis suggests that NTL offers no additional benefit to human well-being in well-developed areas, but even has negative effects, as indicated by consistently negative local marginal effect coefficients in highly illuminated regions in **Figure 4**. In areas with below-median NTL, 38.53% of locations show non-negative NTL effects on well-being, compared to only 21.08% in above-median NTL areas. Positive or neutral effects of NTL on well-being are more common in darker areas than in brightly lit ones. Adequate NTL also provides residents with a sense of safety during the night (Gaston et al., 2015). For this reason, in rural areas without sufficient development, the increased NTL could be a positive factor. However, discussions about the negative effects of NTL on health are often overlooked. Light pollution is linked to adverse health outcomes (Falchi et al., 2011; Falchi et al., 2019; Li & Managi, 2023b). Intense human activities contribute to increased noise levels, particularly from nighttime road traffic. This noise is linked to primary sleep disturbances, cardiovascular diseases, and other health issues (Begou et al., 2020; Pirrera et al., 2010; Welch et al., 2023). Given these factors, the consequences of increased NTL, or overdeveloped environments, are expected to be detrimental. Interestingly, in metropolitan areas, the negative health impacts of NTL are often minimal, potentially counterbalanced by the benefits of developed environments that provide convenience and opportunities. To sum up, in poorly developed areas, more NTL is connected to the improved safety feeling and more infrastructure, which leads to improvements in SWB. In

657 relatively developed areas, the crowded environment with high-level NTL might disturb
658 residents' daily lives, and the development level of the community cannot compensate for
659 the harmful impacts, so the NTL effects are significantly negative. In metropolitan areas,
660 the impacts of NTL are close to zero because people are apt to accept both the negative and
661 positive effects of high-level NTL. Hence, development planning should consider a balance
662 between the development level and its effects on human well-being, fostering the creation
663 of a sustainable society. In areas where the local marginal effect of NTL is found to be
664 positive, a moderate increase, up to $2 \text{ nW}/(\text{cm}^2 \cdot \text{sr})$, in NTL, reflecting higher development
665 intensity, may contribute positively to well-being. However, where the local marginal
666 effect is negative, further development could have adverse impacts and should be
667 approached with caution.

668 Our findings align with international studies showing the importance of urban
669 greenness for subjective well-being. For example, Mackerron and Mourato (2009) indicate
670 that natural environments significantly boosted momentary happiness in the UK, and
671 Ambrey and Fleming (2014) report positive associations between green space and life
672 satisfaction in Australia. However, the overall positive effect may mask local spatial
673 heterogeneity. Where greenness often shows uniformly positive effects, our geographically
674 localized results suggest that the impact of NDVI in Japan varies widely depending on
675 current environments, which is consistent with a case study in Australia with spatial
676 consideration (Sharifi et al., 2021). Similarly, while NTL has been used as a proxy for
677 development globally (Zheng et al., 2022), our study highlights its potential negative
678 effects on well-being at high intensities, particularly in dense urban areas, in line with many
679 international studies. For example, Li et al. (2025) pointed out that sky view and open space

ratio affect the incidence of ischemic heart disease in Wuhan, China. Moreover, in two German cities, dwellers with sufficient green window views have higher human well-being (Kley & Dovbischuk, 2024). High development intensity and high-density buildings inevitably affect these indicators. Our results underscore the importance of spatial and cultural context in interpreting environmental impacts on well-being.

Our research also has certain breakthroughs in methods. This study introduces a novel approach to defining local neighborhoods by leveraging the spatial segmentation of random forests. Unlike traditional methods that rely on fixed-distance buffers or administrative units (Li & Managi, 2021b, 2021c, 2023d), our approach derives each instance's neighborhood based on the overlapping spatial groupings produced by independently trained decision trees. We extract the geographical boundaries defined for the focal instance in each of these trees. The aggregation step ensures that the local zone reflects a consensus among multiple spatial partitions. This method is innovative in two respects. First, it adapts the ensemble model structure to define localized analysis zones in a way that incorporates both model performance and spatial granularity. Second, it allows each instance to have a unique, data-driven neighborhood, which better aligns with the complex and non-uniform spatial relationships. In essence, the random forest becomes not only a predictive model but also a segmentation engine for constructing meaningful spatial context for interpretation. Statistically, random forest and PDA technology are supported by solid theories (Breiman, 2001; Friedman, 2001). Additionally, we show that local connectivity is necessary. The local marginal effect coefficients of a variable of interest are not linearly related to the variable values. This means that directly building a global-scale analysis can only yield a result that is easier to understand, but at the expense of accuracy

703 and plausibility. The relatively strong linear requirement on the local corresponding
704 tangent line of each local marginal effect ensures that the local slopes are statistically
705 significant. The local MSR approach provides a practical framework to translate
706 environmental effects into income-equivalent terms. While it involves an implicit linear
707 tradeoff, we mitigate oversimplification by computing MSR locally and deriving it from
708 non-linear, data-driven partial dependence curves. This avoids imposing a rigid functional
709 form and supports more concrete and spatially grounded policy recommendations.
710 Moreover, statistically, we used the delta method to calculate the significance of all points
711 to ensure the reliability of the MSR values of local points.

712 Several limitations and challenges warrant attention. First, sampling and residential
713 location estimation based on centroids of postal zones might worsen spatial distribution,
714 especially in densely populated metropolitan areas. Second, although this study utilizes a
715 short panel dataset, it does not account for time-fixed effects within individuals. Few
716 studies explore the necessity and nature of data transformations for random forest panel
717 regressions. Additionally, the study pools data across years rather than conducting separate
718 annual analyses. While this approach increases the sample size and geographical coverage
719 for local dataset construction, it limits our ability to explore temporal dynamics or changes
720 over time. The lack of long-term tracking also prevents us from assessing individual-level
721 changes across years, which remains an important direction for future research. Moreover,
722 incorporating additional variables could enhance the analysis. Although our study already
723 includes more variables than past research, important factors like marital status and general
724 health questionnaire scores remain unexamined. The well-being indicators used in this
725 study have limited variability. Expanding the range of SWB assessment options or

726 employing multidimensional evaluation could improve analytical precision. Treating SWB
727 as a continuous variable simplifies analysis but may misrepresent the psychological
728 distance between response levels, which is another limitation. Future research should also
729 refine the random forest model or other powerful machine learning models to better capture
730 time-invariant individual characteristics. Establishing causal links between well-being,
731 greenery, and development is another priority.

732

733 Conclusion

734 In conclusion, this study highlights the spatially variable impacts of greenness,
735 represented by NDVI, and development, denoted by NTL, on subjective happiness across
736 Japan, providing evidence of their complex interplay with socio-economic and
737 environmental factors. While greenness generally contributes positively to well-being,
738 particularly in urban regions, excessive greenness in rural areas may conflict with
739 developmental priorities. Conversely, NTL's effects are predominantly negative due to the
740 impacts of high-intensity development, especially in urban areas, but moderate increases
741 in rural regions can enhance safety and usability. The MWTP analysis further quantifies
742 these impacts, showing economic values associated with increased NDVI, namely 379.01
743 USD/percentage point, and substantial costs linked to excessive NTL, specifically -
744 1434.59 USD/(nW/(cm²·sr)). The findings emphasize the importance of geographically
745 tailored policies to balance environmental and developmental goals, optimizing human
746 well-being while promoting sustainable urban and rural development.

747

748 **Ethics Statement and Informed Consent**

749 We claim that the ethics review committee for Kyushu University, Japan approved
750 all experimental protocols for the survey, and all methods are carried out according to the
751 relevant guidelines and regulations. All survey methods were carried out following relevant
752 guidelines and regulations. At the beginning of the survey, respondents are informed about
753 the survey's aim and their rights to voluntarily participate. Informed consent was obtained
754 from all participants or, if participants are under 18, from a parent and/or legal guardian
755 before responding to the questionnaire.

756

757 **Data Availability**

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

761

762 **Acknowledgment**

763 This research was supported by the following funding agencies: JSPS KAKENHI
764 (Grant No. JP20H00648), and JST MiraiProgram (Grant No. JPMJMI22I4). This research
765 is a working paper for the Research Institute of Economy, Trade and Industry (RIETI).

766

767 **Table:**

Table 1: The Local Marginal Effect Coefficients of Variables of Interest Summary

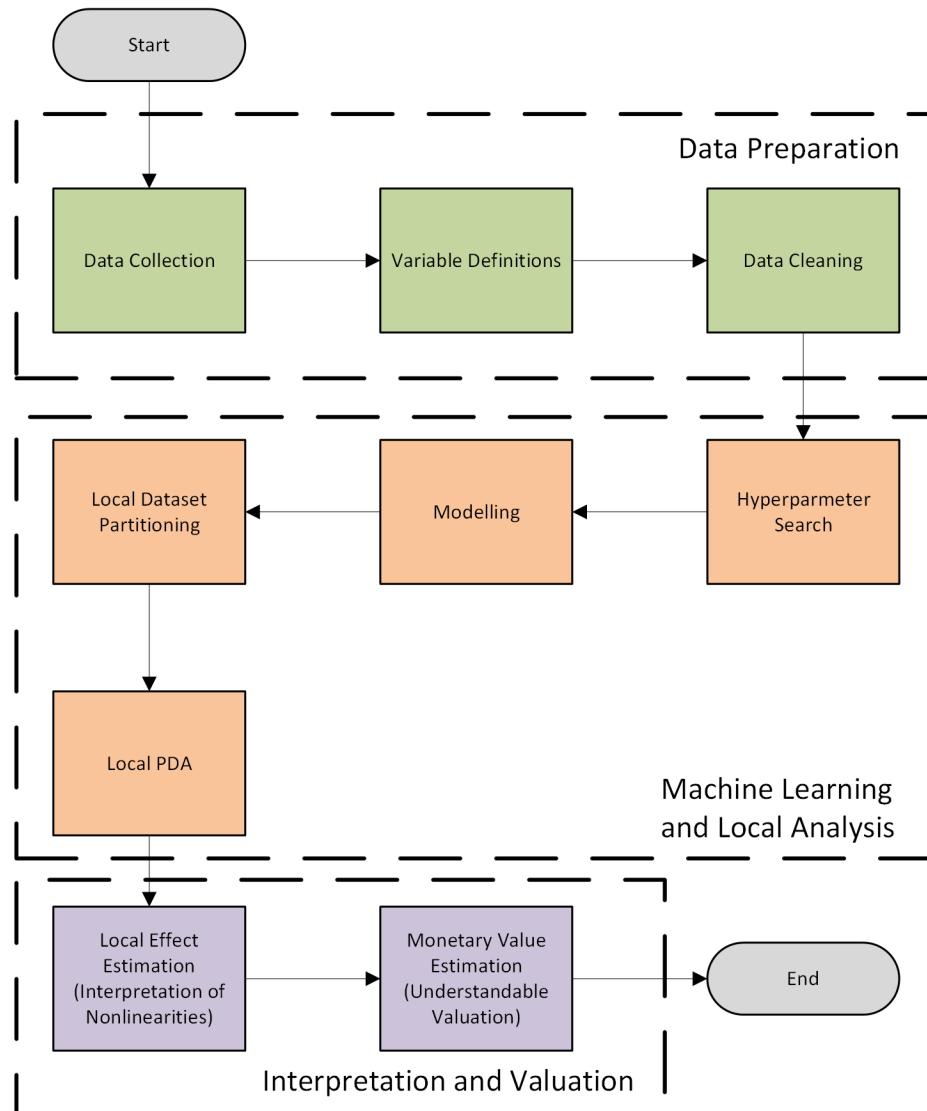
Variable	Level	Count			Percentage			Average Value	
		Negative	Not monotonic	Positive	Total	Negative	Not monotonic	Positive	
NDVI	Individual	94712	117901	170560	383173	24.72%	30.77%	44.51%	0.0003
NDVI	Location	12115	14811	23926	50852	23.82%	29.13%	47.05%	0.0004
NTL	Individual	317427	50580	15166	383173	82.84%	13.20%	3.96%	-0.0017
NTL	Location	35688	10830	4334	50852	70.18%	21.30%	8.52%	-0.0020
Income	Individual	3023	6028	374122	383173	0.79%	1.57%	97.64%	0.0139
Income	Location	1567	1911	47374	50852	3.08%	3.76%	93.16%	0.0148

768

769

770

771 **Figure:**

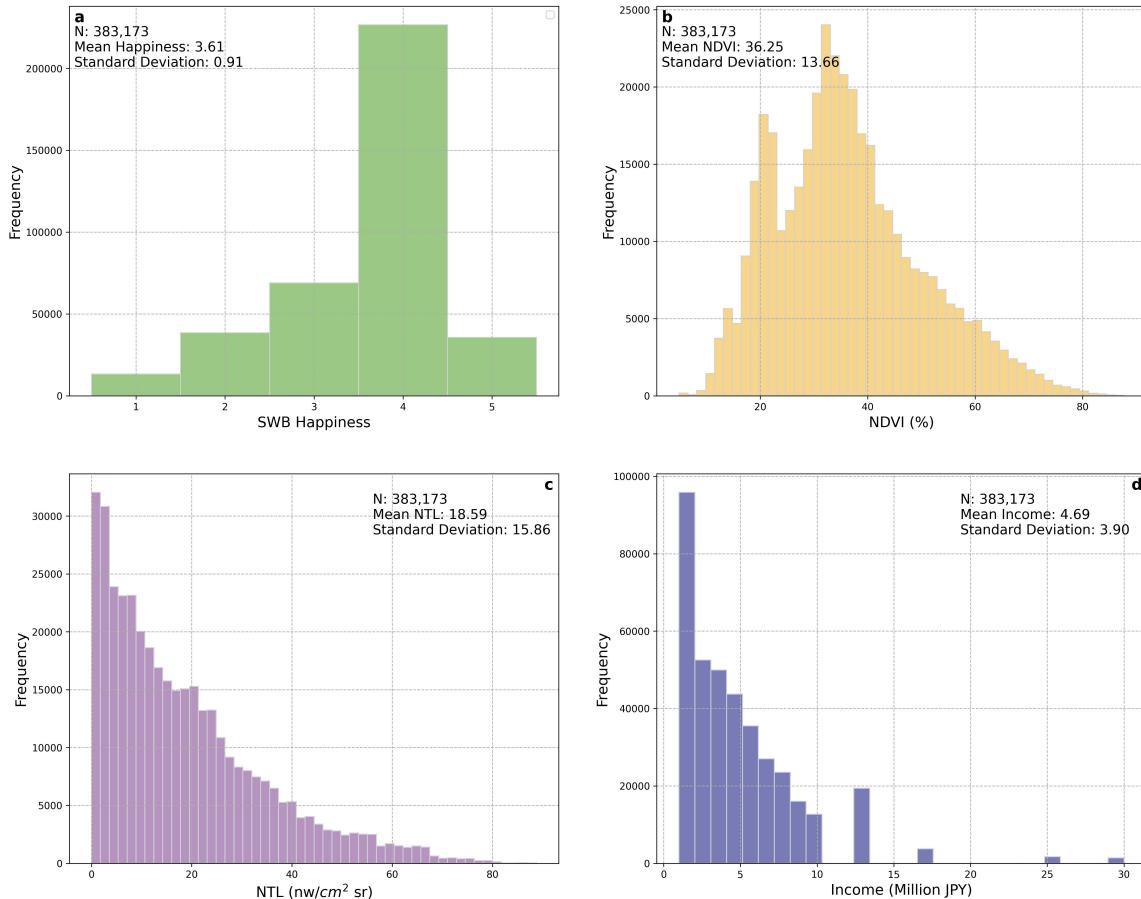


772

773

Figure 1: Conceptual Workflow of the Analytical Pipeline

774



775

776

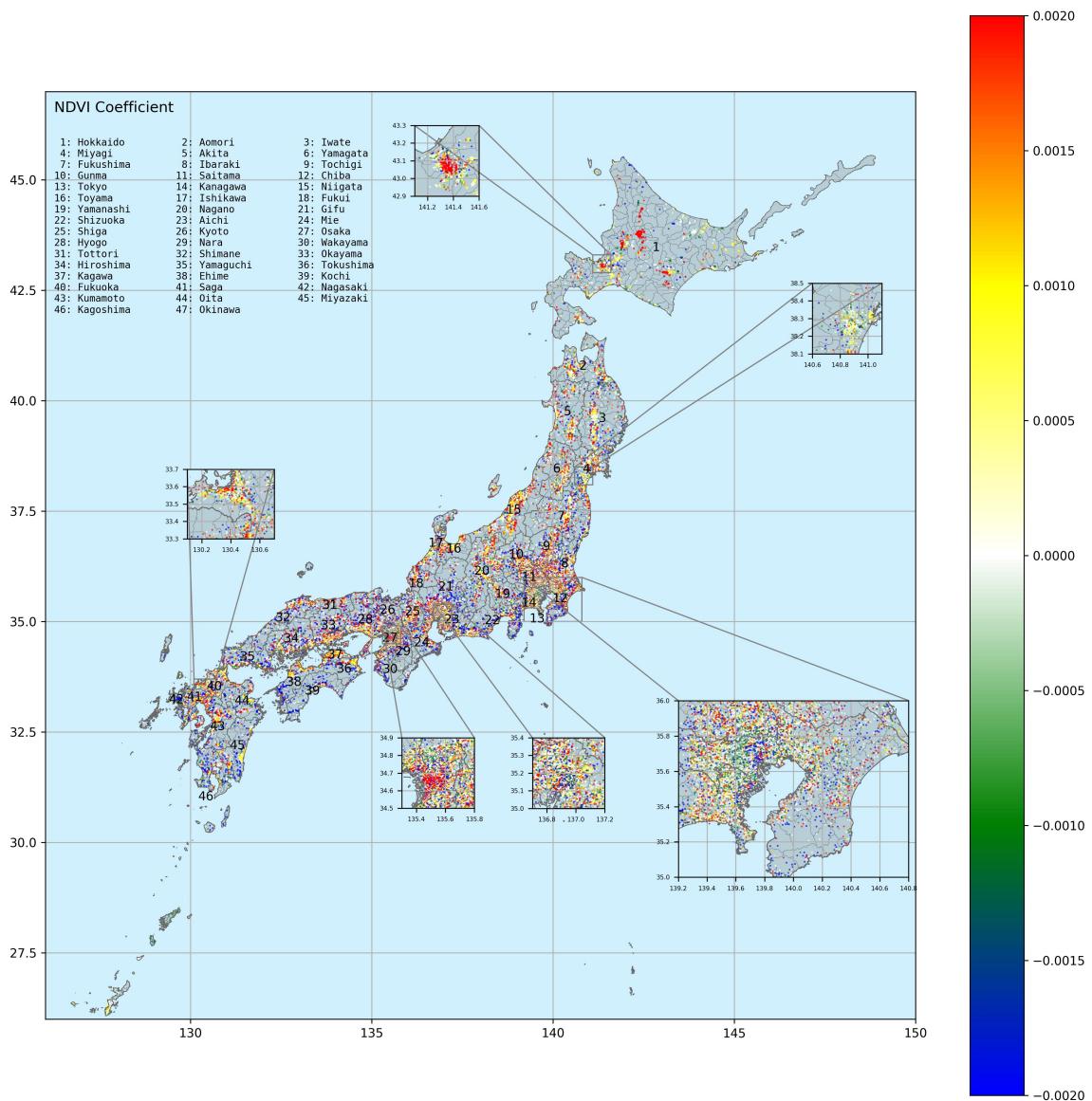
Figure 2: The Statistical Distributions of Critical Variables

777

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

778

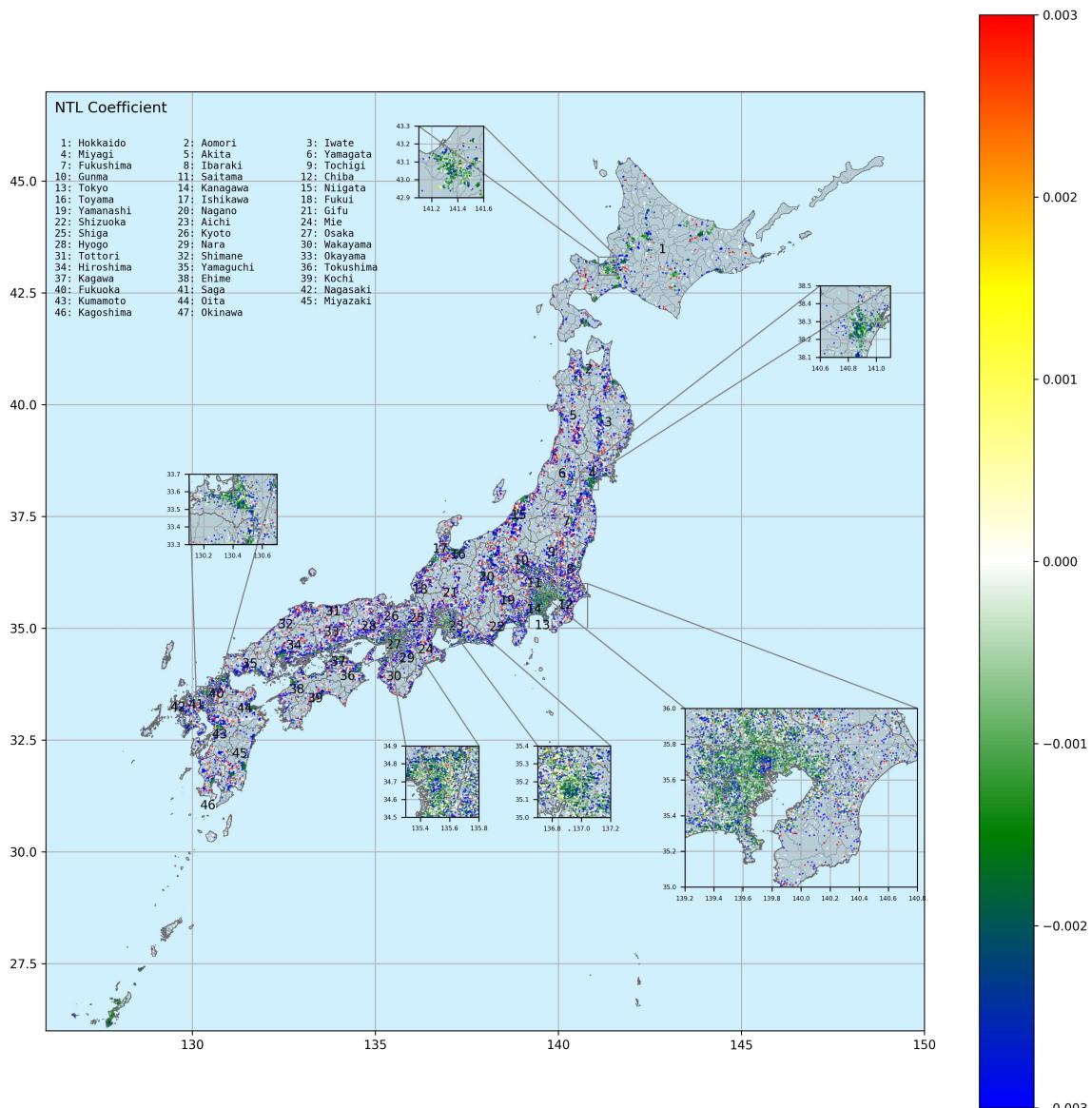
779



780

781 **Figure 3: The Distribution of the Local Marginal Effect Coefficient of NDVI on**
 782 **Happiness**

783

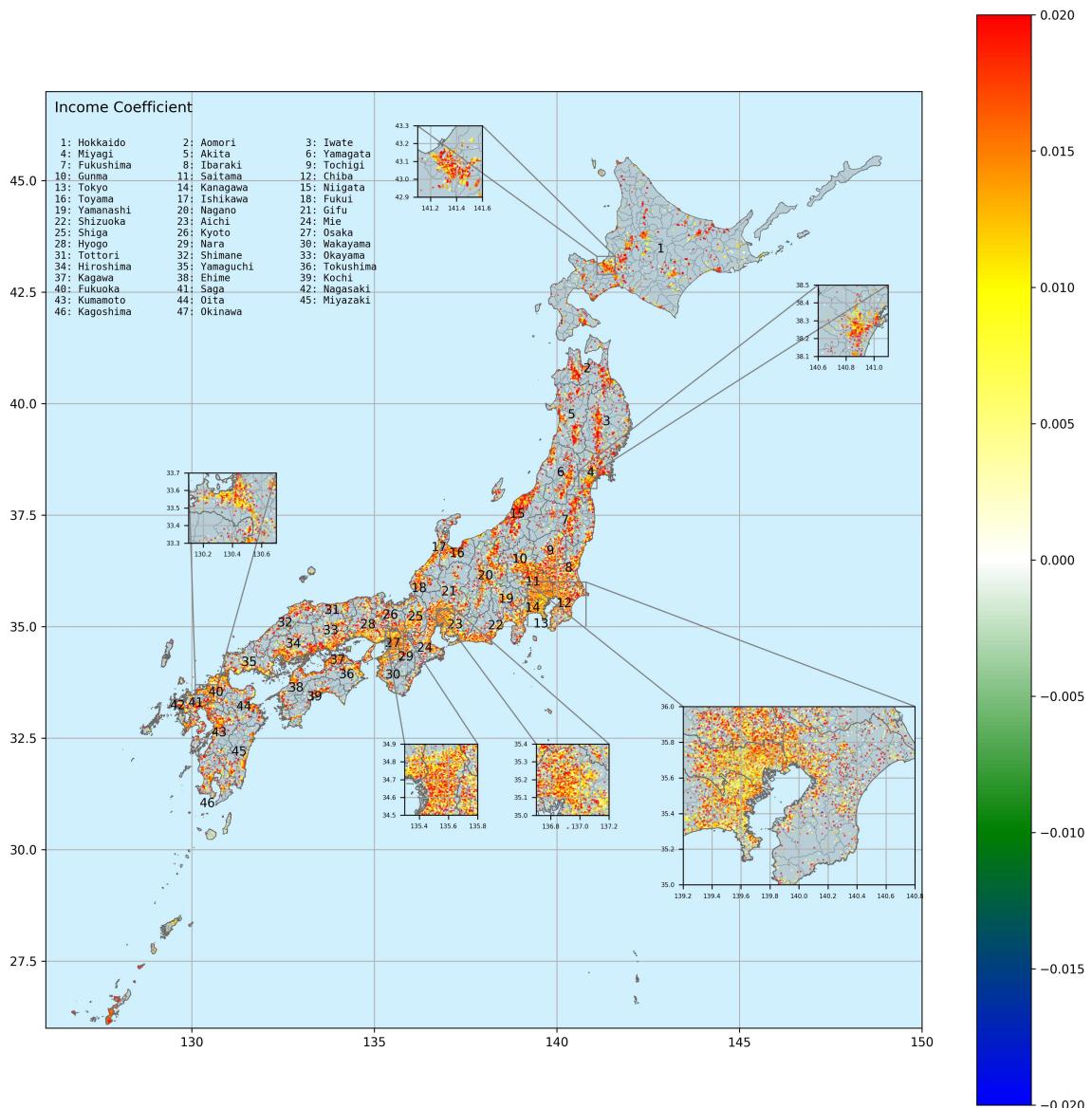


784

785 **Figure 4: The Distribution of Local Marginal Effect Coefficient of NTL on**
Happiness

786

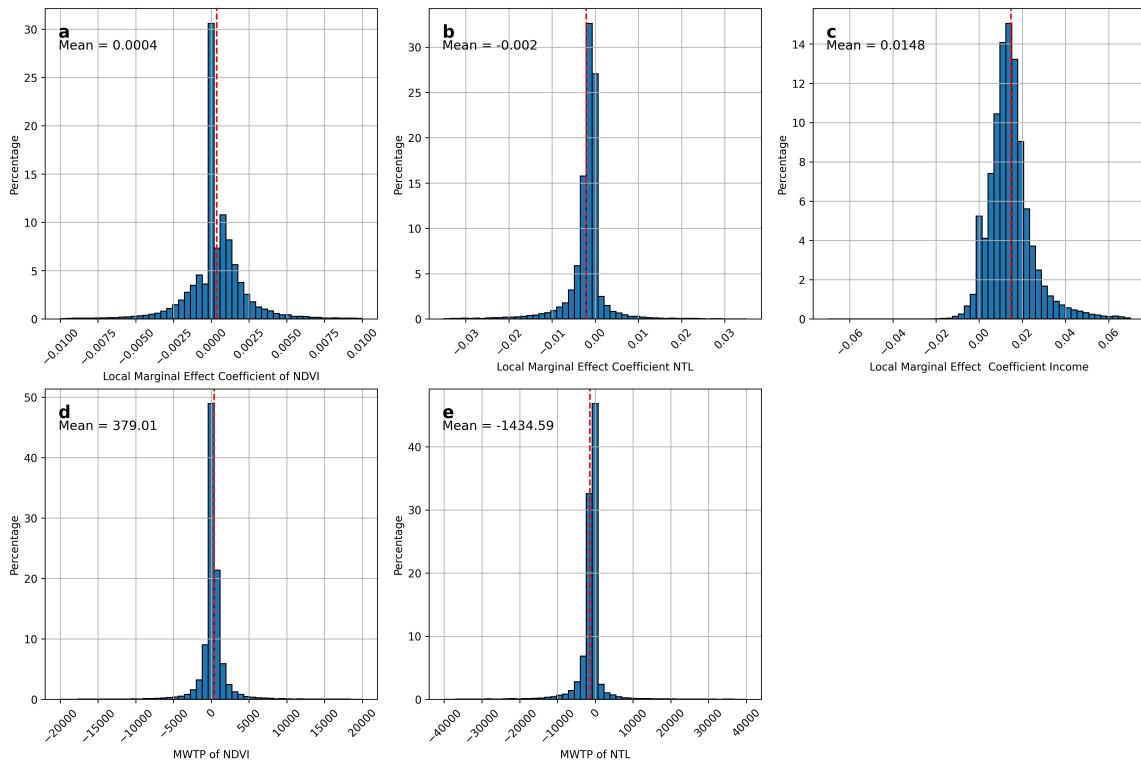
787



788

789 **Figure 5: The Distribution of Local Marginal Effect Coefficient of Income on
790 Happiness**

791

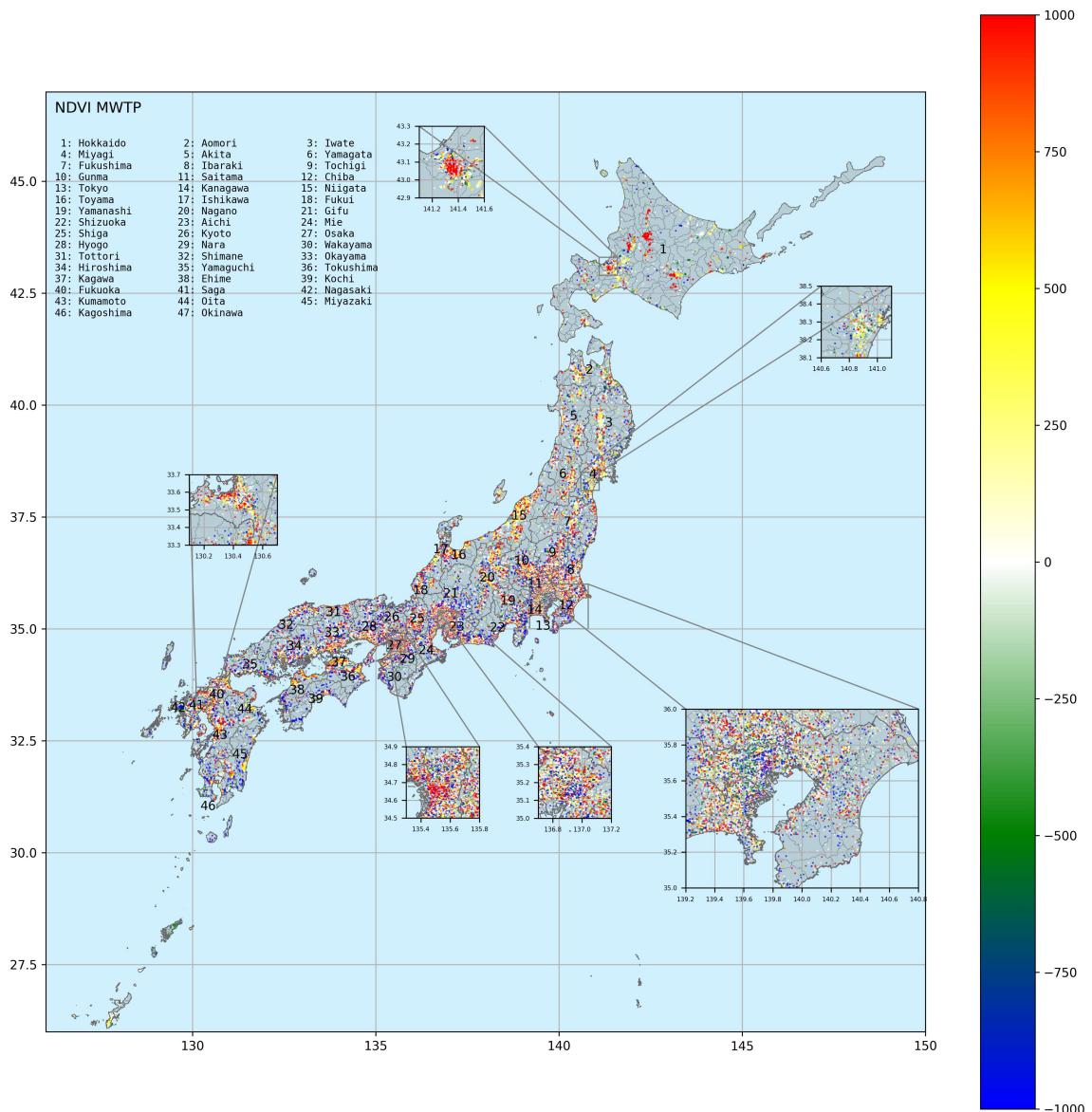


792

793 **Figure 6: The Statistical Distribution of Local Marginal Effect Coefficients of**
794 **Income on Happiness and MWTP**

795

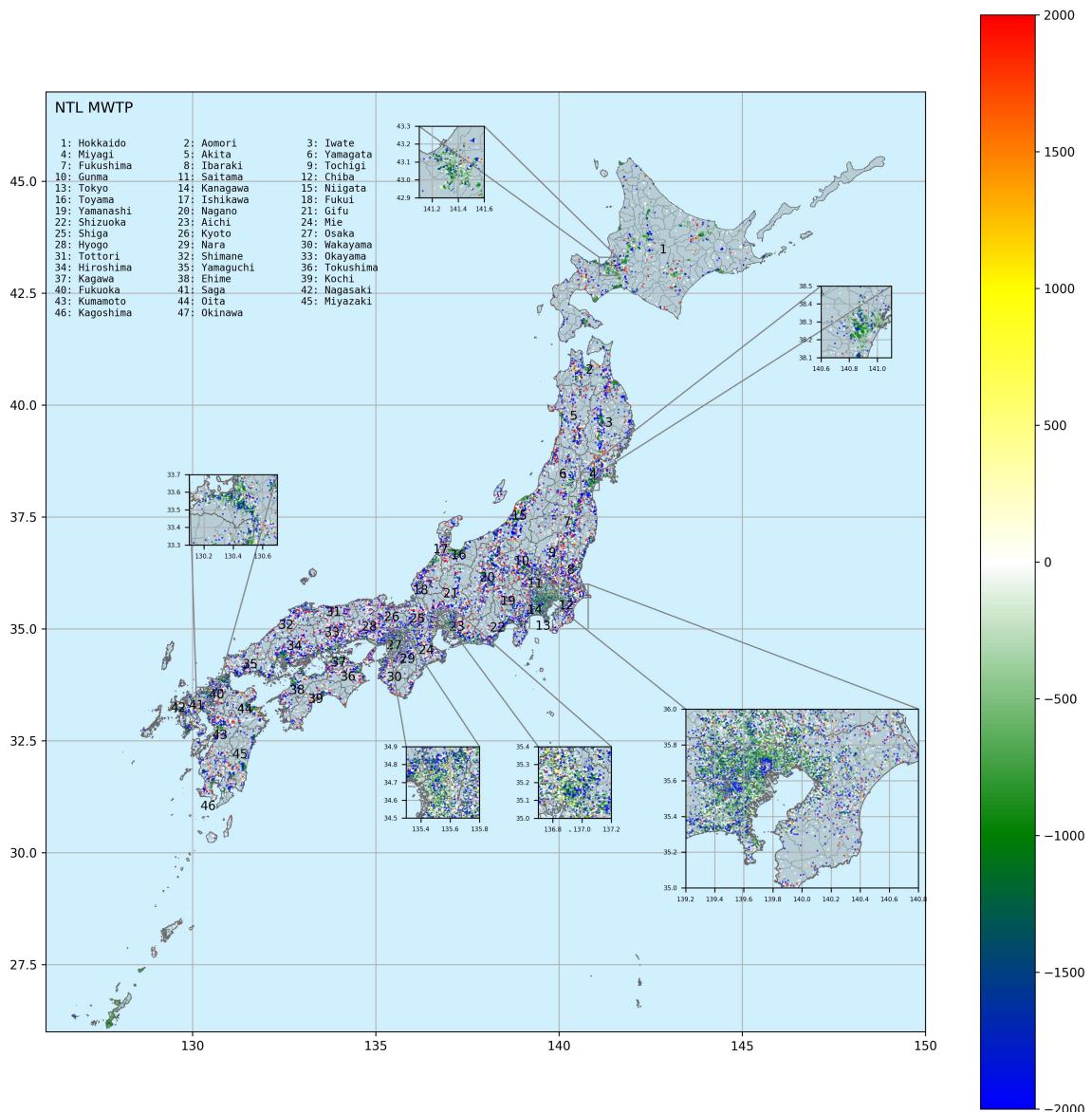
796



797

798 **Figure 7: The Monetary Value of NDVI on Happiness (unit: USD/Percentage Point)**

799



800

801 **Figure 8: The Monetary Value of NTL on Happiness (unit: USD/(nW/cm² · sr))**

802

803

804 **Reference:**

- 805 Ambrey, C., & Fleming, C. (2014). Public Greenspace and Life Satisfaction in Urban
806 Australia [Article]. *Urban Studies*, 51(6), 1290-1321.
807 <https://doi.org/10.1177/0042098013494417>
- 808 Astell-Burt, T., & Feng, X. (2019). Association of Urban Green Space With Mental Health
809 and General Health Among Adults in Australia. *JAMA Network Open*, 2(7),
810 e198209. <https://doi.org/10.1001/jamanetworkopen.2019.8209>
- 811 Astell-Burt, T., Mitchell, R., & Hartig, T. (2014). The association between green space and
812 mental health varies across the lifecourse. A longitudinal study [Article]. *Journal*
813 *of Epidemiology and Community Health*, 68(6), 578-583.
814 <https://doi.org/10.1136/jech-2013-203767>
- 815 Barton, J., & Pretty, J. (2010). What is the Best Dose of Nature and Green Exercise for
816 Improving Mental Health? A Multi-Study Analysis [Article]. *Environmental*
817 *Science & Technology*, 44(10), 3947-3955. <https://doi.org/10.1021/es903183r>
- 818 Beenstock, M., & Felsenstein, D. (2019). *The econometric analysis of non-stationary*
819 *spatial panel data*. Springer.
- 820 Begou, P., Kassomenos, P., & Kelessis, A. (2020). Effects of road traffic noise on the
821 prevalence of cardiovascular diseases: The case of Thessaloniki, Greece. *Science*
822 *of The Total Environment*, 703, 134477.
823 <https://doi.org/https://doi.org/10.1016/j.scitotenv.2019.134477>
- 824 Breiman, L. (2001). Random Forests. *Machine Learning*, 45(1), 5-32.
825 <https://doi.org/10.1023/a:1010933404324>
- 826 Chameides, W. L., Lindsay, R. W., Richardson, J., & Kiang, C. S. (1988). The role of
827 biogenic hydrocarbons in urban photochemical smog: Atlanta as a case study.
828 *Science*, 241(4872), 1473-1475. <https://doi.org/10.1126/science.3420404>
- 829 Chapman, A., Fujii, H., & Managi, S. (2019). Multinational life satisfaction, perceived
830 inequality and energy affordability [Article]. *Nature Sustainability*, 2(6), 508-514.
831 <https://doi.org/10.1038/s41893-019-0303-5>
- 832 Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. 22nd
833 SIGKDD Conference on Knowledge Discovery and Data Mining,
- 834 Chen, X., & Nordhaus, W. D. (2011). Using luminosity data as a proxy for economic
835 statistics. 108(21), 8589-8594. <https://doi.org/doi:10.1073/pnas.1017031108>
- 836 Chen, X., & Nordhaus, W. D. (2019). VIIRS Nighttime Lights in the Estimation of Cross-
837 Sectional and Time-Series GDP. *Remote Sensing*, 11(9), 1057.
838 <https://doi.org/10.3390/rs11091057>
- 839 Chen, Z., Yu, B., Song, W., Liu, H., Wu, Q., Shi, K., & Wu, J. (2017). A New Approach
840 for Detecting Urban Centers and Their Spatial Structure With Nighttime Light
841 Remote Sensing. *IEEE Transactions on Geoscience and Remote Sensing*, 55(11),
842 6305-6319. <https://doi.org/10.1109/TGRS.2017.2725917>
- 843 Chen, Z., Yu, B., Yang, C., Zhou, Y., Yao, S., Qian, X., Wang, C., Wu, B., & Wu, J. (2021).
844 An extended time series (2000–2018) of global NPP-VIIRS-like nighttime light
845 data from a cross-sensor calibration. *Earth System Science Data*, 13(3), 889-906.
846 <https://doi.org/10.5194/essd-13-889-2021>
- 847 Didan, K., Munoz, A. B., Solano, R., Huete, A., & University of Arizona: Vegetation Index
848 Phenology Lab. (2015). *MODIS vegetation index user's guide (MOD13 series)*.

- 849 Diener, E. (1984). Subjective well-being [Review]. *Psychological Bulletin*, 95(3), 542-575.
850 <https://doi.org/10.1037/0033-2909.95.3.542>
- 851 Diener, E., Oishi, S., & Tay, L. (2018). Advances in subjective well-being research
852 [Review]. *Nature Human Behaviour*, 2(4), 253-260.
853 <https://doi.org/10.1038/s41562-018-0307-6>
- 854 Eitelberg, D. A., Van Vliet, J., Doelman, J. C., Stehfest, E., & Verburg, P. H. (2016).
855 Demand for biodiversity protection and carbon storage as drivers of global land
856 change scenarios. 40, 101-111. <https://doi.org/10.1016/j.gloenvcha.2016.06.014>
- 857 Falchi, F., Cinzano, P., Elvidge, C. D., Keith, D. M., & Haim, A. (2011). Limiting the
858 impact of light pollution on human health, environment and stellar visibility.
859 *Journal of Environmental Management*, 92(10), 2714-2722.
860 <https://doi.org/10.1016/j.jenvman.2011.06.029>
- 861 Falchi, F., Furgoni, R., Gallaway, T. A., Rybnikova, N. A., Portnov, B. A., Baugh, K.,
862 Cinzano, P., & Elvidge, C. D. (2019). Light pollution in USA and Europe: The
863 good, the bad and the ugly. *Journal of Environmental Management*, 248, 109227.
864 <https://doi.org/10.1016/j.jenvman.2019.06.128>
- 865 Felipe-Lucia, M. R., Soliveres, S., Penone, C., Manning, P., van der Plas, F., Boch, S.,
866 Prati, D., Ammer, C., Schall, P., Gossner, M. M., Bauhus, J., Buscot, F., Blaser, S.,
867 Bluthgen, N., de Frutos, A., Ehbrecht, M., Frank, K., Goldmann, K., Hansel, F., . . .
868 Allan, E. (2018). Multiple forest attributes underpin the supply of multiple
869 ecosystem services. *Nature Communications*, 9(1), 4839.
870 <https://doi.org/10.1038/s41467-018-07082-4>
- 871 Fotheringham, A., Brunsdon, C., & Charlton, M. (2002). *Geographically Weighted
872 Regression: The Analysis of Spatially Varying Relationships* (Vol. 13). John Wiley
873 & Sons.
- 874 Friedman, J. H. (2001). Greedy Function Approximation: A Gradient Boosting Machine.
875 *The Annals of Statistics*, 29(5), 1189-1232. <http://www.jstor.org/stable/2699986>
- 876 Gaston, K. J., Gaston, S., Bennie, J., & Hopkins, J. (2015). Benefits and costs of artificial
877 nighttime lighting of the environment. *Environmental Reviews*, 23(1), 14-23.
878 <https://doi.org/10.1139/er-2014-0041>
- 879 Ghosh, T., Anderson, S., Elvidge, C., & Sutton, P. (2013). Using Nighttime Satellite
880 Imagery as a Proxy Measure of Human Well-Being. *Sustainability*, 5(12), 4988-
881 5019. <https://doi.org/10.3390/su5124988>
- 882 Gollini, I., Lu, B., Charlton, M., Brunsdon, C., & Harris, P. (2015). GWmodel: An R
883 package for exploring spatial heterogeneity using geographically weighted models.
884 *Journal of Statistical Software*, 63(17).
885 <https://doi.org/https://doi.org/10.18637/jss.v063.i17>
- 886 Hartig, T., & Kahn, P. H. (2016). Living in cities, naturally [Editorial Material]. *Science*,
887 352(6288), 938-940. <https://doi.org/10.1126/science.aaf3759>
- 888 Hunter, R. F., Christian, H., Veitch, J., Astell-Burt, T., Hipp, J. A., & Schipperijn, J. (2015).
889 The impact of interventions to promote physical activity in urban green space: A
890 systematic review and recommendations for future research. *Social Science &
891 Medicine*, 124, 246-256. <https://doi.org/10.1016/j.socscimed.2014.11.051>
- 892 Jebb, A. T., Tay, L., Diener, E., & Oishi, S. (2018). Happiness, income satiation and turning
893 points around the world [Article]. *Nature Human Behaviour*, 2(1), 33-38.
894 <https://doi.org/10.1038/s41562-017-0277-0>

- 895 Kahneman, D., Krueger, A. B., Schkade, D. A., Schwarz, N., & Stone, A. A. (2004). A
896 Survey Method for Characterizing Daily Life Experience: The Day Reconstruction
897 Method. *Science*, 306(5702), 1776-1780. <http://www.jstor.org/stable/3839780>
- 898 Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., & Liu, T.-Y. (2017).
899 *Lightgbm: A highly efficient gradient boosting decision tree* Advances in neural
900 information processing systems,
- 901 Kley, S., & Dovbischuk, T. (2024). The equigenic potential of green window views for city
902 dwellers' well-being. *Sustainable Cities and Society*, 108, 105511.
903 <https://doi.org/https://doi.org/10.1016/j.scs.2024.105511>
- 904 Krekel, C., Kolbe, J., & Wuestemann, H. (2016). The greener, the happier? The effect of
905 urban land use on residential well-being [Article]. *Ecological Economics*, 121, 117-
906 127. <https://doi.org/10.1016/j.ecolecon.2015.11.005>
- 907 Krischke, J., Beckmann-Wübbelt, A., Glaser, R., Dey, S., & Saha, S. (2025). Relationship
908 Between Urban Tree Diversity and Human Well-being: Implications for Urban
909 Planning. *Sustainable Cities and Society*, 124, 106294.
910 <https://doi.org/https://doi.org/10.1016/j.scs.2025.106294>
- 911 Lamchin, M., Lee, W.-K., Jeon, S. W., Wang, S. W., Lim, C. H., Song, C., & Sung, M.
912 (2018). Long-term trend and correlation between vegetation greenness and climate
913 variables in Asia based on satellite data. *Science of The Total Environment*, 618,
914 1089-1095. <https://doi.org/https://doi.org/10.1016/j.scitotenv.2017.09.145>
- 915 Li, C., & Managi, S. (2021a). Contribution of on-road transportation to PM2.5. *Scientific
916 Reports*, 11(1), 21320. <https://doi.org/10.1038/s41598-021-00862-x>
- 917 Li, C., & Managi, S. (2021b). Land cover matters to human well-being. *Scientific Reports*,
918 11(1). <https://doi.org/10.1038/s41598-021-95351-6>
- 919 Li, C., & Managi, S. (2021c). Spatial Variability of the Relationship between Air Pollution
920 and Well-being. *Sustainable Cities and Society*, 103447.
921 <https://doi.org/10.1016/j.scs.2021.103447>
- 922 Li, C., & Managi, S. (2022). Estimating monthly global ground-level NO₂ concentrations
923 using geographically weighted panel regression. *Remote Sensing of Environment*,
924 280, 113152. <https://doi.org/https://doi.org/10.1016/j.rse.2022.113152>
- 925 Li, C., & Managi, S. (2023a). Gridded Datasets for Japan: Total, Male, and Female
926 Populations from 2001–2020. *Scientific Data*, 10(1).
927 <https://doi.org/10.1038/s41597-023-01989-4>
- 928 Li, C., & Managi, S. (2023b). Inappropriate nighttime light reduces living comfort.
929 *Environmental Pollution*, 334, 122173.
930 <https://doi.org/https://doi.org/10.1016/j.envpol.2023.122173>
- 931 Li, C., & Managi, S. (2023c). Income raises human well-being indefinitely, but age
932 consistently slashes it. *Scientific Reports*, 13(1). <https://doi.org/10.1038/s41598-023-33235-7>
- 934 Li, C., & Managi, S. (2023d). Natural land cover positively correlates with COVID-19
935 health outcomes. *BMC Public Health*, 23(1). <https://doi.org/10.1186/s12889-023-15484-3>
- 937 Li, C., & Managi, S. (2024). Mental health and natural land cover: a global analysis based
938 on random forest with geographical consideration. *Scientific Reports*, 14(1), 2894.
939 <https://doi.org/10.1038/s41598-024-53279-7>

- 940 Li, C., & Managi, S. (2025). Impacts of community attachment and community livability
941 on environmental activity according to XGBoost and SHAP. *Cities*, 156, 105559.
942 <https://doi.org/https://doi.org/10.1016/j.cities.2024.105559>
- 943 Li, X., Liu, Q., Liu, X., & Ren, Y. (2025). Greener is healthier: Unraveling the distinction
944 of health impacts of urban green indicators in Wuhan, China. *Sustainable Cities
945 and Society*, 125, 106318. <https://doi.org/https://doi.org/10.1016/j.scs.2025.106318>
- 946 Lu, B., Charlton, M., Harris, P., & Fotheringham, A. S. (2014). Geographically weighted
947 regression with a non-Euclidean distance metric: a case study using hedonic house
948 price data. *International Journal of Geographical Information Science*, 28(4), 660-
949 681. <https://doi.org/10.1080/13658816.2013.865739>
- 950 MacKerron, G. (2012). Happiness Economics from 35 000 Feet [Review]. *Journal of
951 Economic Surveys*, 26(4), 705-735. [https://doi.org/10.1111/j.1467-6419.2010.00672.x](https://doi.org/10.1111/j.1467-
952 6419.2010.00672.x)
- 953 Mackerron, G., & Mourato, S. (2009). Life satisfaction and air quality in London.
954 *Ecological Economics*, 68(5), 1441-1453.
955 <https://doi.org/10.1016/j.ecolecon.2008.10.004>
- 956 MacKerron, G., & Mourato, S. (2013). Happiness is greater in natural environments
957 [Article]. *Global Environmental Change*, 23(5), 992-1000.
958 <https://doi.org/10.1016/j.gloenvcha.2013.03.010>
- 959 Maslow, A. H. (1943). A theory of human motivation. *Psychological review*, 50(4), 370.
- 960 McCallum, I., Kyba, C. C. M., Bayas, J. C. L., Moltchanova, E., Cooper, M., Cuartero, J.
961 C., Pachauri, S., See, L., Danylo, O., Moorthy, I., Lesiv, M., Baugh, K., Elvidge,
962 C. D., Hofer, M., & Fritz, S. (2022). Estimating global economic well-being with
963 unlit settlements. *Nature Communications*, 13(1). [https://doi.org/10.1038/s41467-022-30099-9](https://doi.org/10.1038/s41467-
964 022-30099-9)
- 965 Mendoza-Ponce, A., Corona-Núñez, R., Kraxner, F., Leduc, S., & Patrizio, P. (2018).
966 Identifying effects of land use cover changes and climate change on terrestrial
967 ecosystems and carbon stocks in Mexico. *Global Environmental Change*, 53, 12-
968 23. <https://doi.org/10.1016/j.gloenvcha.2018.08.004>
- 969 Molnar, C. (2020). *Interpretable machine learning*. Lulu. com.
- 970 Mouratidis, K. (2018). Built environment and social well-being: How does urban form
971 affect social life and personal relationships? *Cities*, 74, 7-20.
972 <https://doi.org/https://doi.org/10.1016/j.cities.2017.10.020>
- 973 Oswald, A. J., & Wu, S. (2010). Objective Confirmation of Subjective Measures of Human
974 Well-Being: Evidence from the USA [Article]. *Science*, 327(5965), 576-579.
975 <https://doi.org/10.1126/science.1180606>
- 976 Piao, X., Li, C., & Managi, S. (2024). Infrastructure distribution in cities and the
977 improvement of the well-being of citizens in Japan: evidence from the quantile
978 regression. *Sustainable and Resilient Infrastructure*, 1-15.
979 <https://doi.org/10.1080/23789689.2024.2403882>
- 980 Pinto, L. V., Inácio, M., Ferreira, C. S. S., Ferreira, A. D., & Pereira, P. (2022). Ecosystem
981 services and well-being dimensions related to urban green spaces – A systematic
982 review. *Sustainable Cities and Society*, 85, 104072.
983 <https://doi.org/https://doi.org/10.1016/j.scs.2022.104072>

- 984 Pirrera, S., De Valck, E., & Cluydts, R. (2010). Nocturnal road traffic noise: A review on
985 its assessment and consequences on sleep and health. *Environment International*,
986 36(5), 492-498. <https://doi.org/https://doi.org/10.1016/j.envint.2010.03.007>
- 987 Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V., & Gulin, A. (2018).
988 *CatBoost: unbiased boosting with categorical features* Advances in neural
989 information processing systems,
- 990 Ryff, C. D. (2014). Psychological Well-Being Revisited: Advances in the Science and
991 Practice of Eudaimonia. *Psychotherapy and Psychosomatics*, 83(1), 10-28.
992 <https://doi.org/10.1159/000353263>
- 993 Seresinhe, C. I., Preis, T., MacKerron, G., & Moat, H. S. (2019). Happiness is Greater in
994 More Scenic Locations [Article]. *Scientific Reports*, 9, 11, Article 4498.
995 <https://doi.org/10.1038/s41598-019-40854-6>
- 996 Seresinhe, C. I., Preis, T., & Moat, H. S. (2015). Quantifying the Impact of Scenic
997 Environments on Health. *Scientific Reports*, 5(1), 16899.
998 <https://doi.org/10.1038/srep16899>
- 999 Sharifi, F., Nygaard, A., & Stone, W. M. (2021). Heterogeneity in the subjective well-
1000 being impact of access to urban green space. *Sustainable Cities and Society*, 74,
1001 103244. <https://doi.org/https://doi.org/10.1016/j.scs.2021.103244>
- 1002 Steptoe, A., Deaton, A., & Stone, A. A. (2015). Subjective wellbeing, health, and ageing.
1003 *The Lancet*, 385(9968), 640-648. [https://doi.org/10.1016/s0140-6736\(13\)61489-0](https://doi.org/10.1016/s0140-6736(13)61489-0)
- 1004 Su, J. G., Jerrett, M., Meng, Y.-Y., Pickett, M., & Ritz, B. (2015). Integrating smart-phone
1005 based momentary location tracking with fixed site air quality monitoring for
1006 personal exposure assessment. *Science of The Total Environment*, 506-507, 518-
1007 526. <https://doi.org/https://doi.org/10.1016/j.scitotenv.2014.11.022>
- 1008 Suk, J. Y., & Walter, R. J. (2019). New nighttime roadway lighting documentation applied
1009 to public safety at night: A case study in San Antonio, Texas. *Sustainable Cities
1010 and Society*, 46, 101459. <https://doi.org/https://doi.org/10.1016/j.scs.2019.101459>
- 1011 Triguero-Mas, M., Dadvand, P., Cirach, M., Martinez, D., Medina, A., Mompart, A.,
1012 Basagana, X., Grazuleviciene, R., & Nieuwenhuijsen, M. J. (2015). Natural outdoor
1013 environments and mental and physical health: Relationships and mechanisms
1014 [Article]. *Environment International*, 77, 35-41.
1015 <https://doi.org/10.1016/j.envint.2015.01.012>
- 1016 Tsurumi, T., Imaugi, A., & Managi, S. (2018). Greenery and Subjective Well-being:
1017 Assessing the Monetary Value of Greenery by Type [Article]. *Ecological
1018 Economics*, 148, 152-169. <https://doi.org/10.1016/j.ecolecon.2018.02.014>
- 1019 Wang, R., Yang, B., Yao, Y., Bloom, M. S., Feng, Z., Yuan, Y., Zhang, J., Liu, P., Wu,
1020 W., Lu, Y., Baranyi, G., Wu, R., Liu, Y., & Dong, G. (2020). Residential greenness,
1021 air pollution and psychological well-being among urban residents in Guangzhou,
1022 China. *Science of The Total Environment*, 711, 134843.
1023 <https://doi.org/10.1016/j.scitotenv.2019.134843>
- 1024 Welch, D., Shepherd, D., Dirks, K. N., & Reddy, R. (2023). Health effects of transport
1025 noise. *Transport Reviews*, 43(6), 1190-1210.
1026 <https://doi.org/10.1080/01441647.2023.2206168>
- 1027 White, M. P., Elliott, L. R., Gascon, M., Roberts, B., & Fleming, L. E. (2020). Blue space,
1028 health and well-being: A narrative overview and synthesis of potential benefits.

