Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 August 2020 doi:10.20944/preprints202008.0499.v1

Article

Application of NDVI in Environmental Justice, Health and Inequality Studies – Potential and Limitations in Urban Environments

Carsten Juergens[1],* & Matthias Fabian Meyer-Heß[1]

[1] Geomatics Group, Institute of Geography, Faculty of Geosciences, Ruhr-University Bochum, D-44870 Bochum, Germany; carsten.juergens@rub.de, matthias.meyer@rub.de

* Correspondence: carsten.juergens@rub.de;

Abstract: This paper discusses the potential and limitations of the Normalized Difference Vegetation Index (NDVI) in environmental justice, health and inequality studies in urban areas. Very often the NDVI is correlated with socioeconomic and/or sociodemographic data to demonstrate the inequality in environmental settings that themselves influence individual health and questions of environmental justice. This paper addresses the limits of the NDVI for such applications and as well its potential, if applied properly. The overall goal is to make people of disciplines other than those that are geo-related aware of the characteristics, limits and potentials of satellite image-based information layers such as NDVI.

Keywords: Greenspace, NDVI, environmental justice, greenness, Sentinel, satellite, urban green, health equity;

---

1. Introduction

Earth observation is a discipline that monitors the earth and the condition of the earth´s surfaces for more than 50 years now. Special earth observation satellites are able to acquire timely data in frequent intervals from the earth surface [1, 2, 3]. One of the most prominent land surfaces is the vegetation cover that has a large extension on the land surfaces. In urban contexts vegetation surfaces are very important for the well-being and health of the urban population. Due to the material composition of urban areas it is helpful to use very high resolution satellite images with decimeter ground resolution. According to [4] "High resolution data are a valuable source for urban and suburban areas and can deliver information in high geometric and semantic quality for various cities and urban agglomerations around the world. Due to accelerating urban sprawl and increasing urban population more and more topics arise where remote sensing is able to support planning and other public duties." The benefit of earth observation data is the ability to keep pace with the development and to keep track of the changes and additions within urban areas in terms of relevant land cover types (e.g. vegetation).

To be able to monitor the vegetation and to judge the condition of photosynthetic active plants one developed vegetation indices. Based on the spectral characteristics of vegetation a comparison between the reflectance in the red (R) and near infrared (nIR) parts of the electromagnetic spectrum is calculated. The selection of these wavelengths results from the absorption and reflection characteristics of vegetation. Due to absorption processes in the visible light, especially in the red part of the electromagnetic spectrum, associated to chlorophyll content of the leaves, one can observe low reflectance values for healthy vegetation. In contrast to that, in the near infrared part of the electromagnetic spectrum one can observe a very strong reflection which corresponds to multiple reflections in the inner cell structure of leaves due to the water content in the cells (see Figure 1).

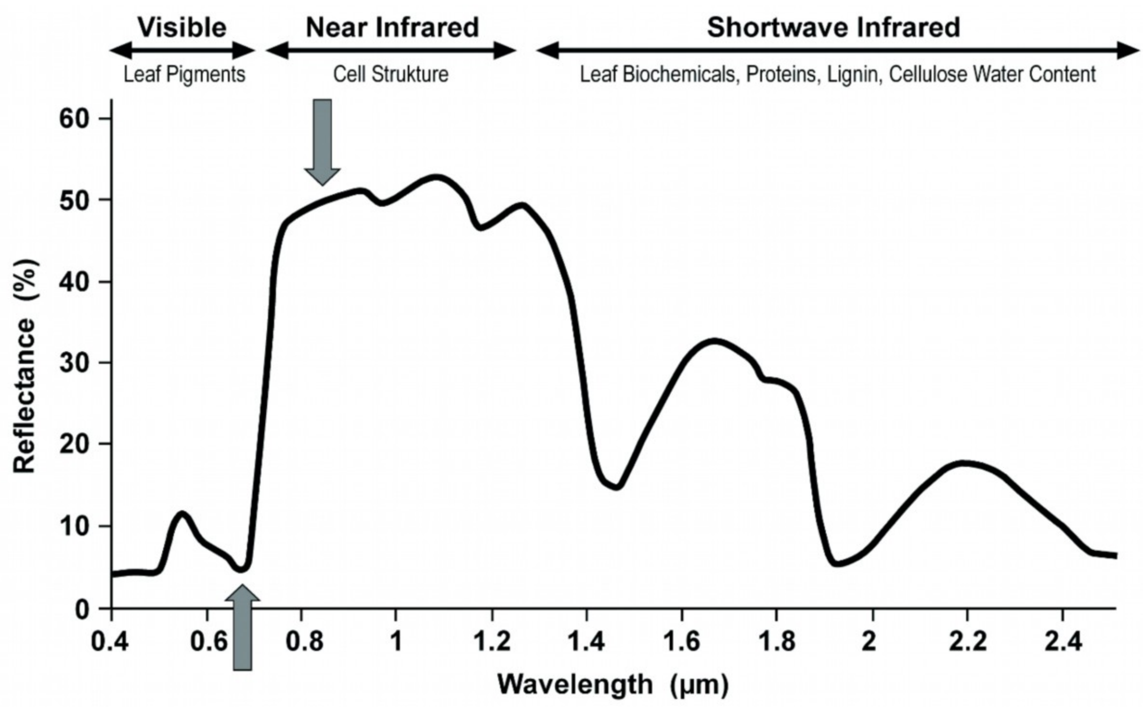

Figure 1. Reflectance curve of photosynthetic active vegetation (modified after [5])

Based on these characteristics [6] calculated the simple ratio vegetation index (RVI) by dividing the red (R) by the near infrared (nIR):

$$RVI = R / nIR$$

"The RVI is widely used for green biomass estimations and monitoring, specifically, at high density vegetation coverage, since this index is very sensitive to vegetation and has a good correlation with plant biomass. However, when the vegetation cover is sparse (less than 50% cover), RVI is sensitive to atmospheric effects, and their representation of biomass is weak." [7].

The most widely used Normalized Difference Vegetation Index (NDVI) was proposed by [8] and is defined as:

$$NDVI = (nIR – R) / (nIR + R)$$

Due to the normalization in the formula the NDVI values appear in the range of -1 to +1. Values below 0 are not related to healthy green vegetation, rather to water, bare soil or abiotic urban surfaces like roofs and road materials. The more the NDVI value tends to +1, the more it is related to vegetation cover and its vigour.

Due to specific needs many other vegetation indices have been developed. For environments with sparse vegetation covers [9] introduced the Soil-Adjusted Vegetation Index (SAVI) which was improved later by the Optimized Soil-Adjusted Vegetation Index (OSAVI) developed by [10]. Many vegetation indices are addressing specific needs, like the modified Normalized Difference Vegetation Index (mNDVI) [11] which is used to estimate frost damages in agriculture based on Landsat data.

Despite all the more sophisticated vegetation indices and the more specific problem-oriented vegetation indices, the NDVI is probably the most used vegetation index today, due to its simple formula and ease of use. NDVI "is often used in research related to regional and global vegetation assessments and was shown to be related not only to canopy structure and LAI but also to canopy photosynthesis" [7] (p. 3). It allows quantitative evaluations and comparisons of different vegetation covers as well as the analysis of vigor and growth dynamics [7]. Consequentially, vegetation indices in general, and the NDVI in particular, are a widely accepted and applied means to assess and monitor spatio-temporal vegetation changes.

Nowadays the NDVI is widely used in environmental justice, health and inequality studies in urban and sub-urban contexts. The application of vegetation indices like NDVI implies that there is vegetation in the investigated urban environment. Vegetation in cities mainly consist of trees, bushes, agricultural fields and meadows/pastures that comprise recreational (e.g. parks, forest), natural (e.g. forests) and agricultural (e.g. fields and pastures) land uses. In general, vegetation is thought to improve our well-being, our health and our quality of life. According to that, one could believe that a high degree of vegetation correlates positively with better quality of life and better individual health. This rather broad perspective does not account for different quality of vegetation, related to the height and its visibility/individual perception or the time period the vegetation appears green (e.g. fields are a rather temporary land use and meadows/pastures are cut and change their appearance). In addition to that, one neglects other influencing environmental factors that could stress individuals and their perception and health situation, like environmental pollution, noise etc.

Vegetation cover in most regions is associated with seasonality aspects which stem from the different seasons during a year. For regional or continental studies one can identify the start of the growing season for large regions by calculating the NDVI [12]. In addition to that during one year one can observe the regional differences in the NDVI values which correspond to vegetation dynamics.

For investigations at larger scales, e.g. for cities one has to use high resolution images to be able to identify as much vegetation details as possible. This is necessary due to the fact that one finds a high number of different surface materials in the city. The smaller the image pixels the higher is the chance to get pure vegetation pixels. If the geometric resolution of the sensor is rather coarse, then one will get many so-called mixed pixels consisting of different materials in one pixel. Then it is almost impossible to isolate the vegetation information. [13] for instance use high resolution satellite images for an urban vegetation phenology analysis in the city of Nanjing, China. Urban vegetation serves a multitude of urban ecosystem functions [14]. "As a main characteristic which is the expression of the seasonal cycles of plant processes and their connections to climate change (temperature and precipitation), vegetation phenology is increasingly significant for a variety of scientific applications nowadays. The timing of phenological events can be used to document and evaluate the effects of climate change on both individual plant species and vegetation communities [15]. To study the features of urban vegetation phenology can better understand the ecological status of the city, the occurrence time of urban vegetation phenology can reflect the response of urban vegetation ecosystem to urban temperature change and precipitation" [13] (p. 43).

Besides phenology driven studies, other authors use urban green indicators for environmental justice/inequality research. [16] for instance use the Spatial Urban Health Equity Indicator Framework (SUHEI) (see also [17]) to relate urban green to other factors in the city with social context, such as air pollution or noise, to estimate the health inequalities for different neighborhoods in the city of Dortmund, Germany. Unfortunately, they did not use the urban green area information mapped to the exact location but calculated the percentage of green area (each area > 1ha) in an administrative unit. Instead, satellite derived NDVI´s could deliver up-to-date information on the current quality of green spaces and help to adapt the SUHEI-results to the true situation and location, e.g. in dry weeks or months of a year, where the green lawn is no longer green.

In contrast to [16], [18] stated that the social sciences increasingly recognize the meaning of georeferenced and geo-spatial data, including remote sensing imagery. He also asserts that more and more social scientists are able to link their data with remote sensing data. A quick literature survey uncovers good and also poor examples. For instance, [19] unfortunately use NDVI data calculated from a Landsat satellite image of 2003 together with Urban Atlas data from 2012. Since high resolution satellite images like Landsat are free of charge and available on a routine basis (revisit time 16 days), one is wondering why they did not use a scene closer to the timeliness of the Urban Atlas data. Besides the gap of nine years, the acquisition date in mid-April seems quite early compared to the vegetation dynamics in their study area. Due to these frame conditions the findings related to green spaces, well-being, health and socio-economic status could be affected from the old base data set and the decision on the season. [20] investigated spatiotemporal contextual uncertainties with MODIS satellite data for the Netherlands and one of his results is: “To mitigate contextual uncertainties, it is advised to integrate temporally well-aligned green space data” (no page numbering). Also, [21] use MODIS satellite data in Taiwan to investigate a linkage between greenness and mortality for a time series of data. They found out that NDVI and mortality causes are negatively correlated.

Another ill-conceived example of NDVI-integration is from [22], who investigate the potential of satellite image-based information for planning authorities to improve the inhabitants´ quality of life. In their publication they do not give any information on the image acquisition dates and the results related to the vegetation`s influence on the urban climate is rather generalized. In this form, the presented results are almost useless, since there is no information on the height of vegetation types, length of green period (e.g. for trees, meadows or fields) or other seasonal effects. And of course, according to [23], [24], [25] or [26] these parameters affect the local climate.

Besides those case studies one can imagine that a closer collaboration between social science and geomatics experts could improve the understanding of socio-spatial phenomena. [27] (p. 262) indicates the potential to “socialize the pixels”. This could assist to analyze socio-spatial indicators together with earth observation image data. [28] created the new term “socio-geomatics” to underline the scientific potential which can be gained by the common use of socio-demographic and socio-economic data together with earth observation and other geo data-related to environmental justice questions. Undoubtedly the interdisciplinary approach will help to come to new insights. However, it is essential, that all used data sets are used properly. As shown in a few examples above, and pointed out by [29], one has to be literate to adequately select and use earth observation and other geo data.

2. Materials and Methods

This section presents the geospatial data used and the respective geospatial Analysis. The study area (Figure 2) is a part of the city of Dortmund in North Rhine-Westphalia in Germany. In this study area one finds all relevant land-cover/land-use types of the region.

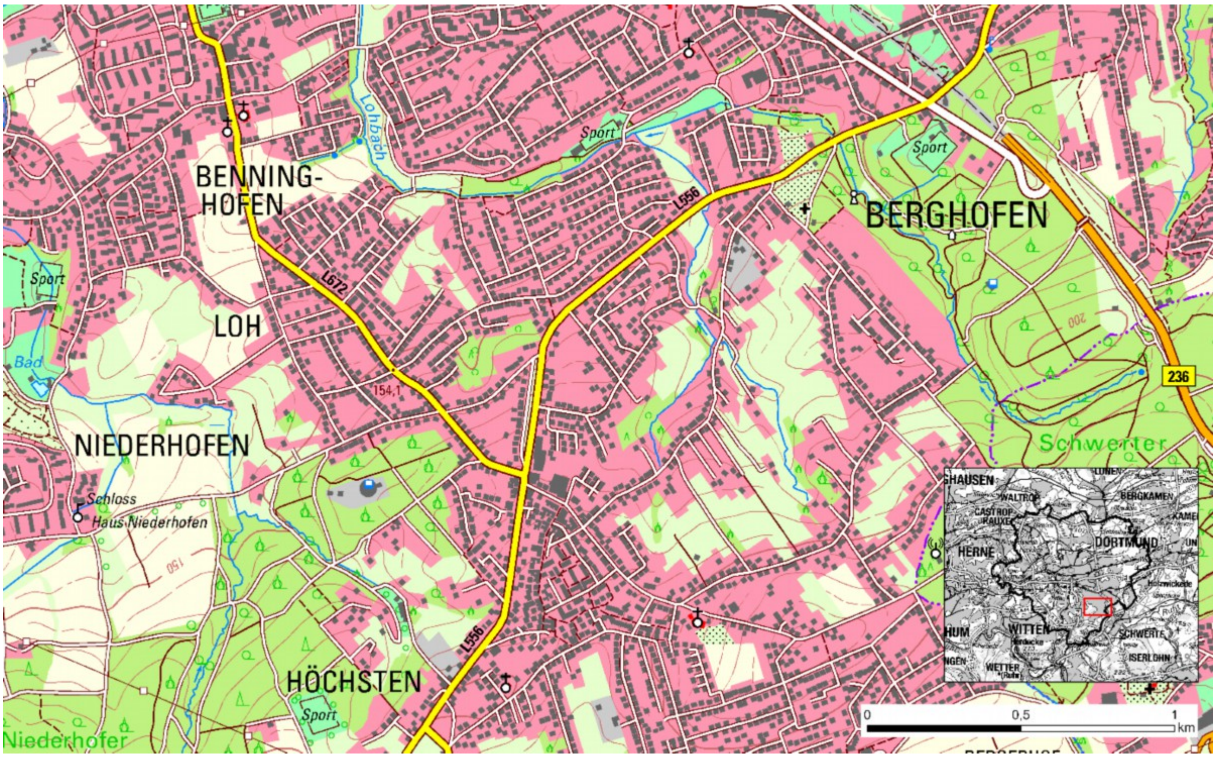

Figure 2. Study area (data source: [**30**], [31])

The land-cover/land-use in the study area was mapped in 2017 by the Regionalverband Ruhr (RVR) [32]. It is a very detailed map product that needed to be simplified for the purpose of this study, to comply with the vegetation cover. Originally the land-cover/land-use map consists of 151 individual categories, from which 57 were meaningful for this study. Those were reduced to 7 main classes of this investigation (Table 1).

For different administrative reasons, the land-cover/land-use map does not cover the whole study area. Therefore, those following maps that are based on this data set have a different extent than those of other geospatial data used in this study.

Table 1. Land-cover/land-use class reduction for this study (base data from [32])

Original land-cover/land-use category (code_akt)	Main land-cover/land-use category
10, 20, 30, 40, 51, 52, 54, 72, 75, 83, 84, 85, 87, 91, 93, 140, 151, 152, 171, 174, 211, 221, 281, 381	Sealed
370	Field
361, 362	Grassland
291, 292, 293, 382	Open Space
400, 410, 420, 431, 432, 441	Forest
233, 271, 273, 282, 305, 321, 451, 471, 472	Other Vegetation
53, 223, 301, 302, 303, 306, 308, 309, 331, 383, 452	Other

Further we used freely available cloud-free high resolution Sentinel-2 satellite images [33] (see Table 2) to investigate the study area during the vegetation period 2019.

Table 2. Sentinel-2 satellite images [33] used for the vegetation period April 1 to August 31, 2019

Acquisition date	Used Bands	Pixel size
2019-04-20	2,3,4,8	10m x 10m
2019-05-15	2,3,4,8	10
2019-06-29	2,3,4,8	10
2019-07-24	2,3,4,8	10
2019-08-26	2,3,4,8	10

Due to phenology the reflectance characteristics of vegetation covers vary during the vegetation period between trees (forests, parks), bushes, meadows and agricultural fields with different crops. Other surfaces like roads or buildings are more or less invariant with time. this is illustrated for a few typical locations in Figure 3.

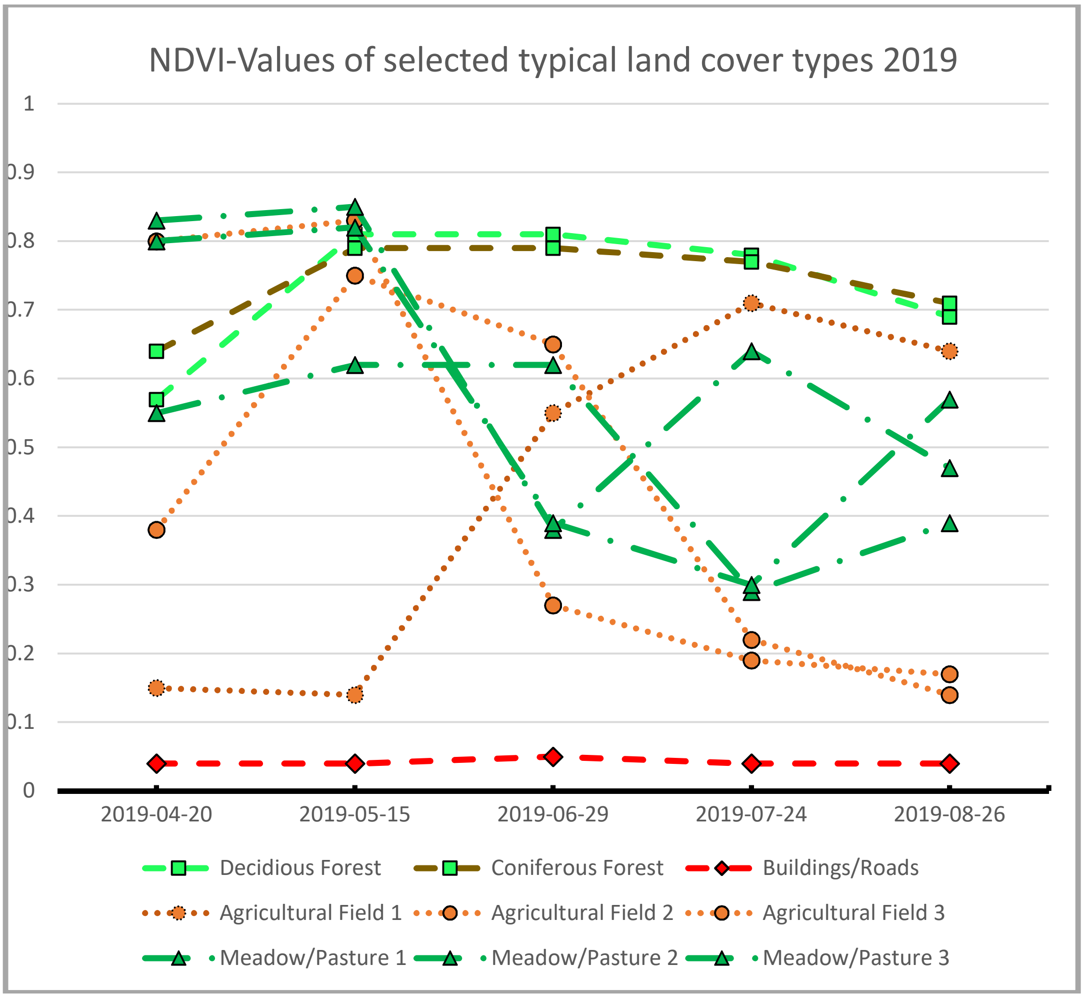

Figure 3. Seasonal NDVI values of selected land cover types from Sentinel-2 images

The five satellite images from which the information in Figure A was extracted are displayed in Figure 4 in natural colors and in Figure 5 in false colours. The images clearly reveal seasonal differences between the different land cover/land use categories. The false colours better pronounce vegetation in red colors. The seasonal effects are also evident in the NDVI-images of the respective dates (Figure 6).

 doi:10.20944/preprints202008.0499.v1

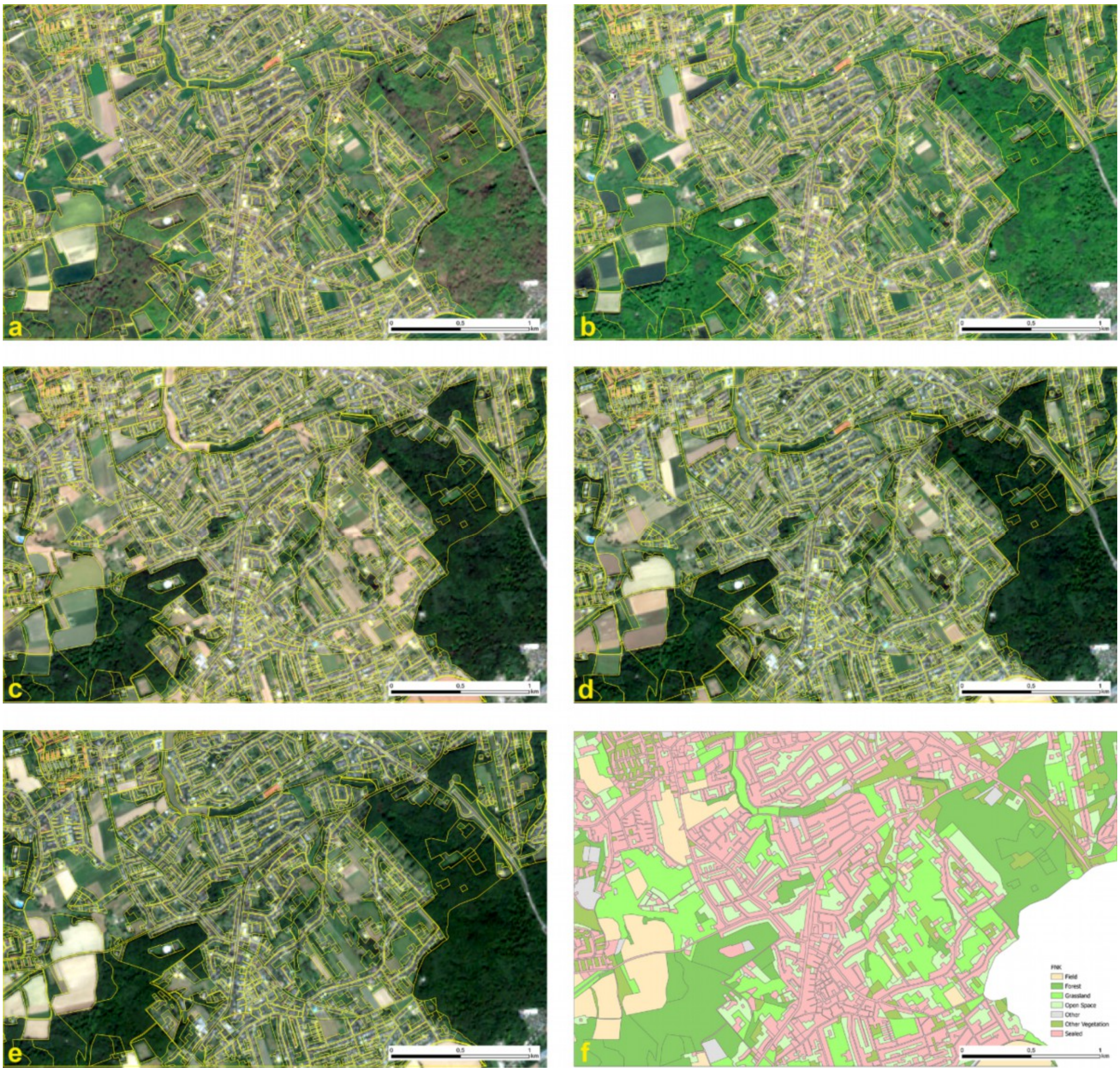

Figure 4. Study area as seen from Sentinel-2 satellite [33] in natural colors (R,G,B=4,3,2) with land use polygons (yellow) on 2019-04-20 (a), 2019-05-15 (b), 2019-06-29 (c), 2019-07-24 (d), 2019-08-26 (e), and as a land use map (data source [32]) (f)

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 August 2020 doi:10.20944/preprints202008.0499.v1

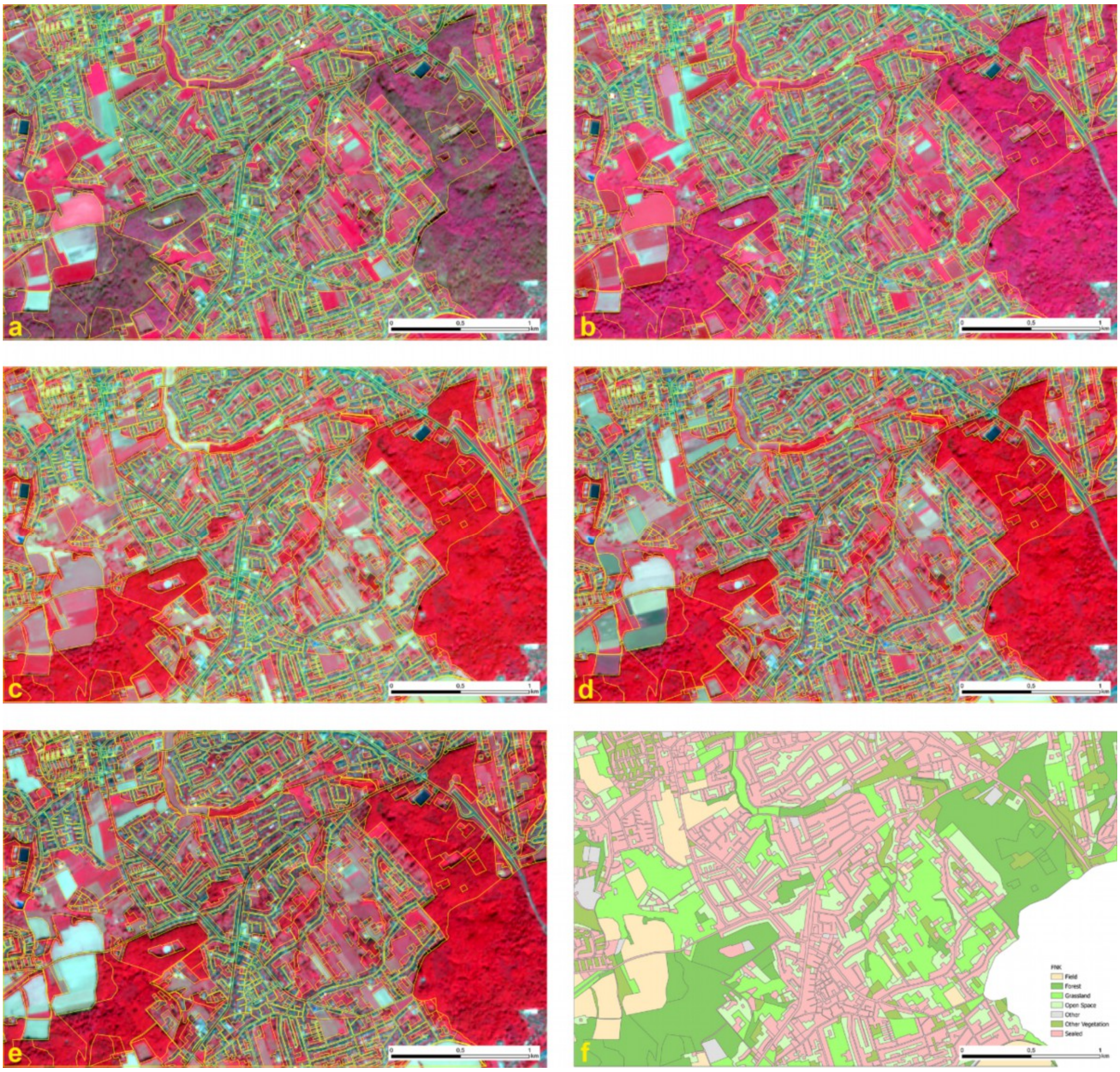

Figure 5. Study area as seen from Sentinel-2 satellite [33] in false colors (R,G,B=8,4,3) with land use polygons (yellow) on 2019-04-20 (a), 2019-05-15 (b), 2019-06-29 (c), 2019-07-24 (d), 2019-08-26 (e), and as a land use map (data source [32]) (f)

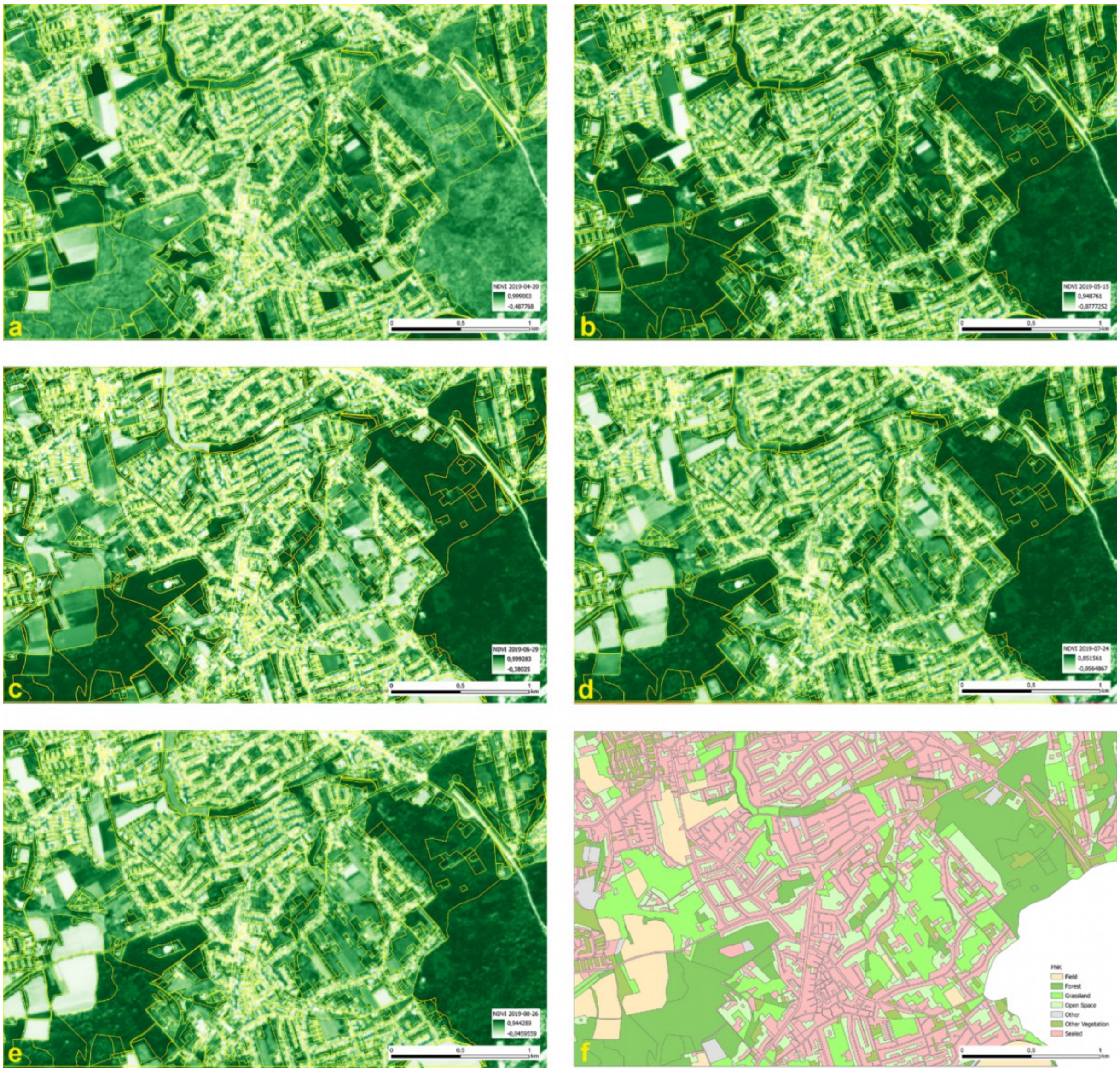

Figure 6. NDVI values (data source [33]) with land use polygons (yellow) for the study area on 2019-04-20 (a), 2019-05-15 (b), 2019-06-29 (c), 2019-07-24 (d), 2019-08-26 (e), and as a land use map (data source [32]) (f)

Besides the calculation of individual NDVI values per image acquisition date, or the calculation of NDVI differences between adjacent image dates, another reasonable approach to calculate mean NDVI values is to do that over time from different images, but always for the same pixel locations. This allows to gain insight into a seasonal average NDVI. Consequently, a mean NDVI across the time period of the five satellite images (temporal mean per pixel) was calculated (Figure 7). In addition to that also the maximum NDVI (temporal maximum per pixel) for the period of observation (April-August 2019) (Figure 8) was extracted.

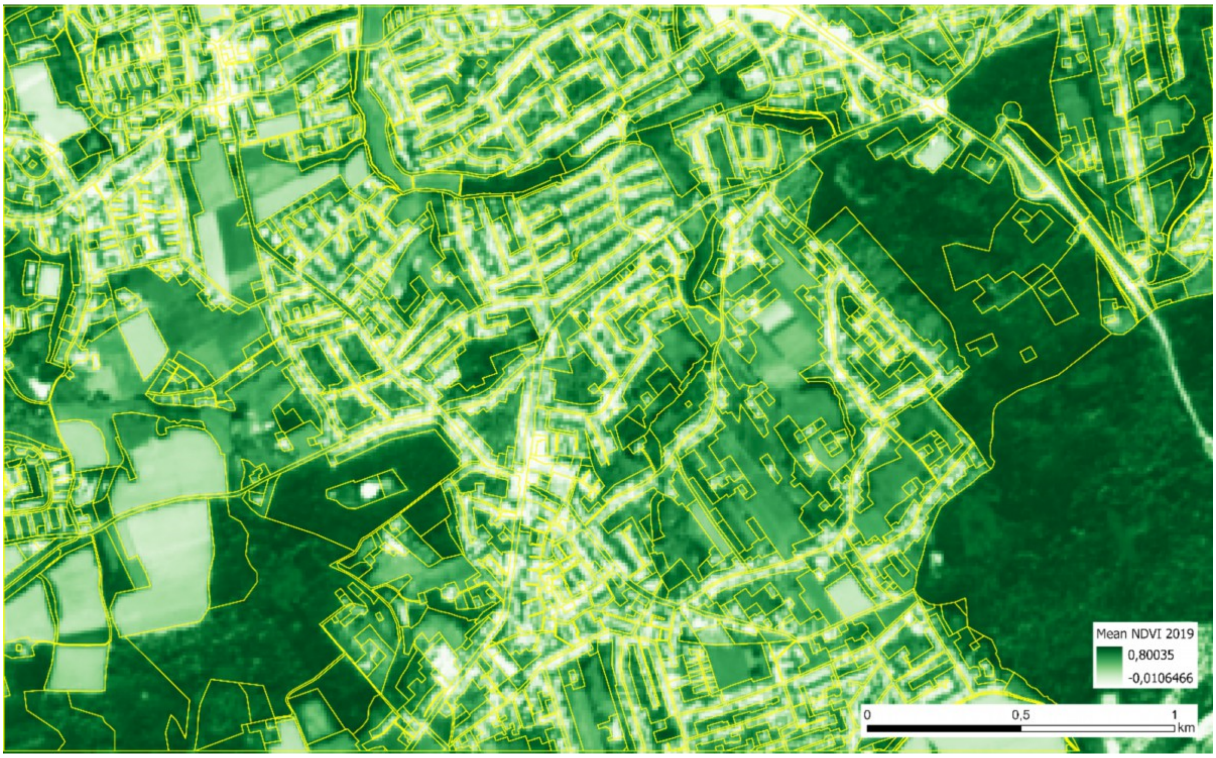

Figure 7. Mean NDVI (data source [33]) for the vegetation period 2019 (data source polygon overlay: [32])

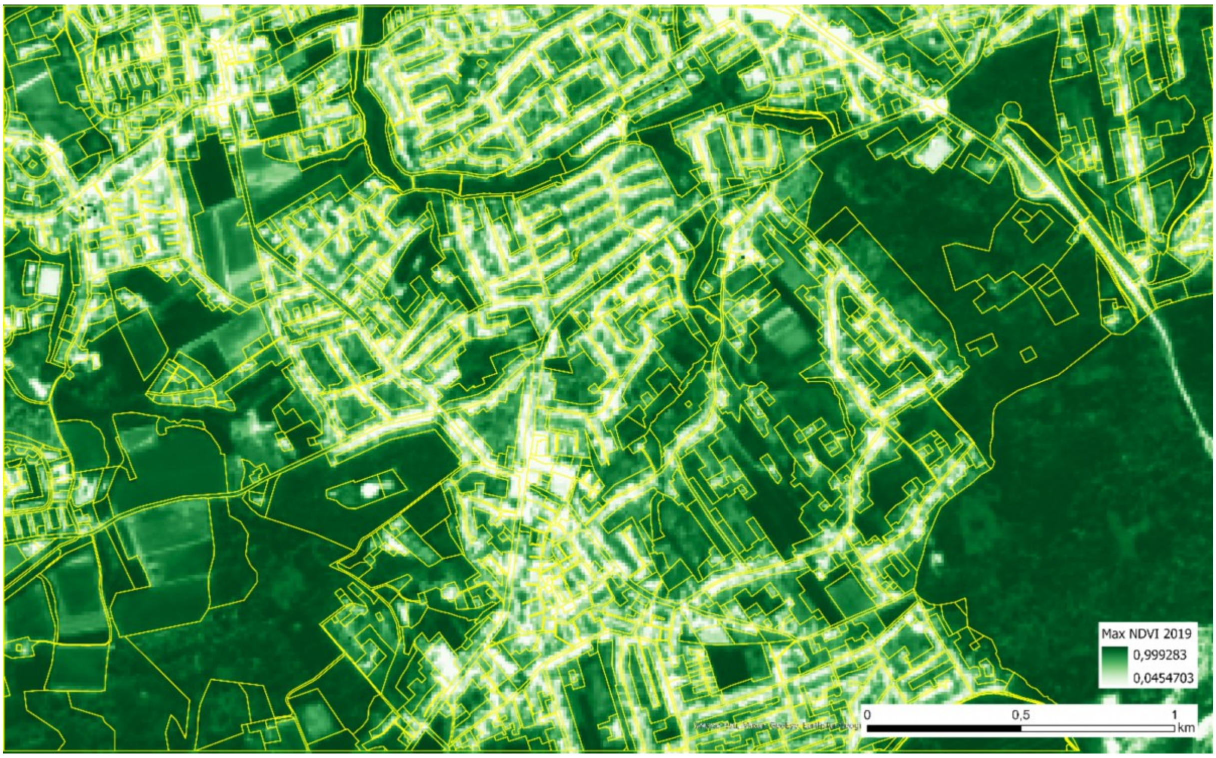

Figure 8. Maximum NDVI (data source [33]) for the vegetation period 2019 (data source polygon overlay: [32])

From the calculated maximum NDVI values one could also extract the month per pixel that corresponds to the respective pixel. This provides the information in which month the greeness is most intense (Figure 9). In a further step one calculated the modus of these pixels with the maximum NDVI-month per polygon of each land-cover/land-use class. This helps to understand the temporal

 doi:10.20944/preprints202008.0499.v1

variability of slightly generalized maximum NDVI values (Figure 10). Having those calculations with maximum NDVI and the modus of the maximum NDVI per polygon, one can identify the maximum intensity of the NDVI (or greeness) per vegetation period and the month with the maximum greeness.

Figure 9. Date of maximum NDVI (data source [33]) per pixel (data source polygon overlay: [32])

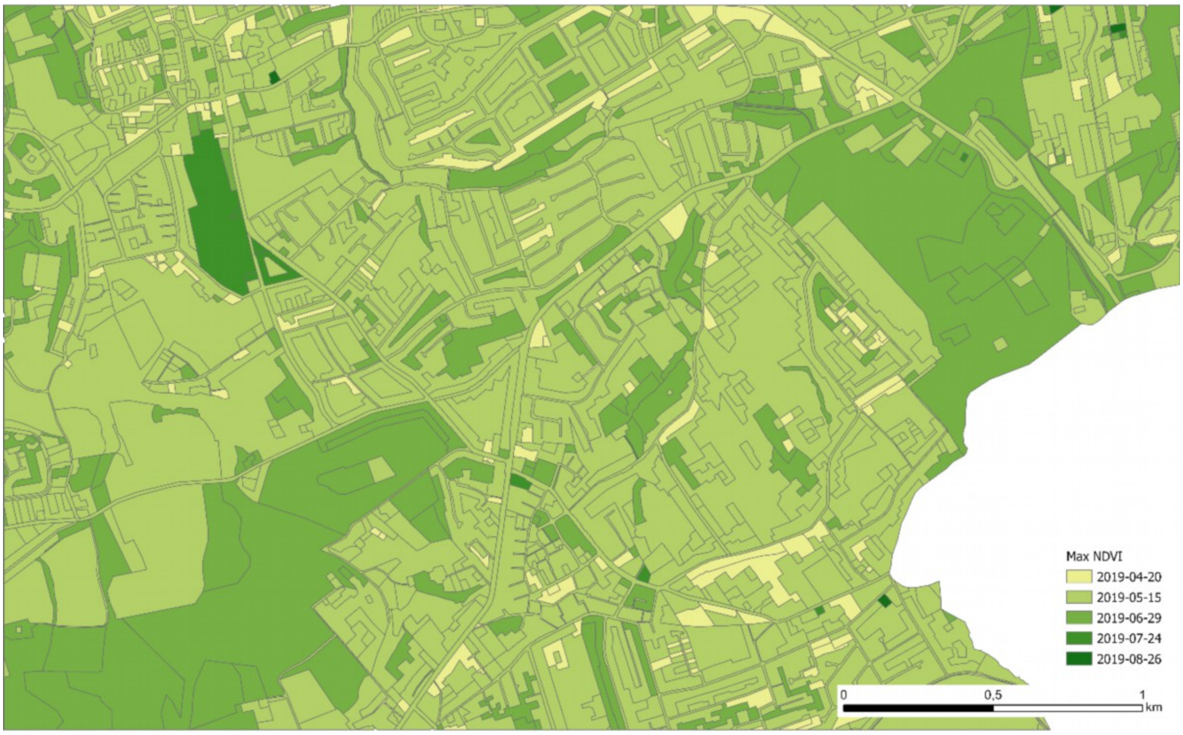

Figure 10. Date of maximum NDVI (data source [33]) per polygon (data source polygon overlay: [32])

Besides the calculation of spatially mean NDVI´s (e.g. for administrative units or fields) another reasonable approach to calculate mean NDVI values is to do that over time from different images, but always for the same location. This allows to gain insight into an annual or seasonal average NDVI. [28] did that for the vegetation period (April-September) for the city of Dortmund in Germany. Due

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 August 2020 doi:10.20944/preprints202008.0499.v1

to his intention to study the mean NDVI for a complete study area, he combined the multi-temporal mean NDVI calculation with the regional mean NDVI calculation. The disadvantage is, that land surfaces with no or little vegetation are included in the calculation and consequently lower the resulting mean NDVI value. Advantageously he could have calculated the mean NDVI on a field or parcel basis. This could give a much better representation of greenness in the urban environment (Figure 11).

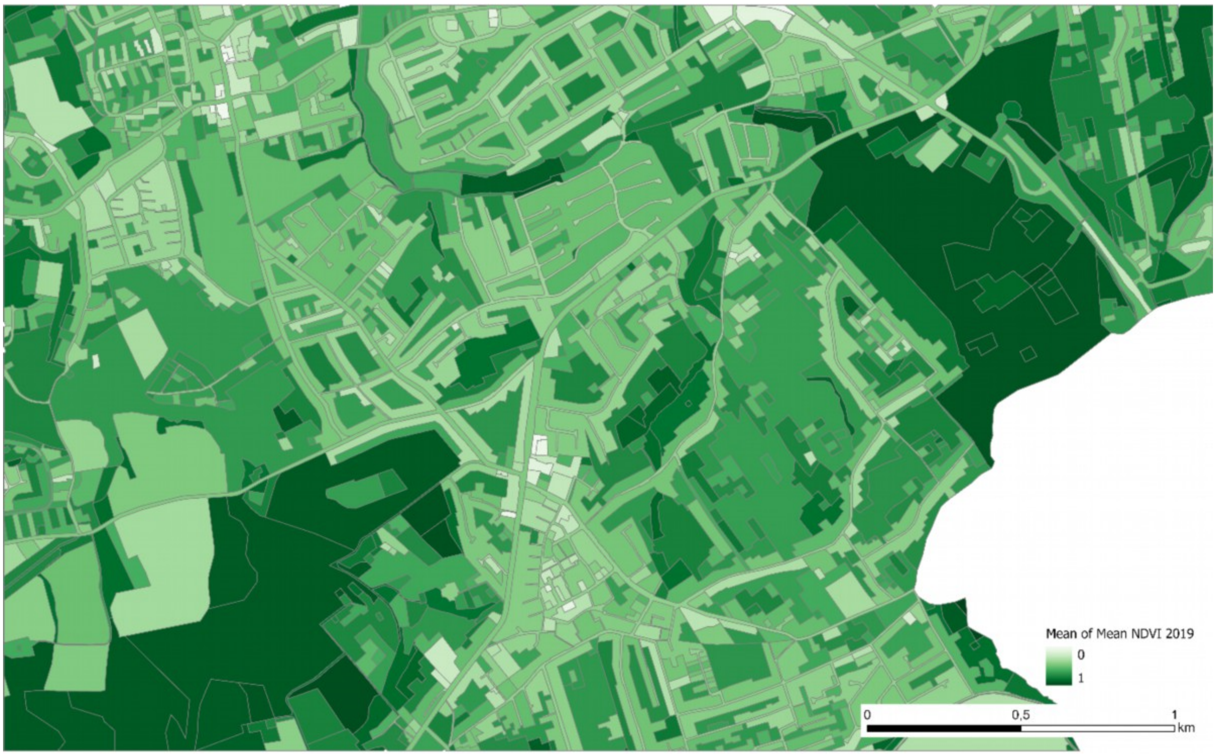

Figure 11. Mean temporal NDVI (data source [33]) calculated as mean NDVI per land cover/land use polygon (data source polygon overlay: [32])

The temporal mean NDVI could be used to generally evaluate the vegetation configuration without consideration of seasonal differences. A high mean value indicates green vegetation for a long period. Lower values could be caused due to less intense green intervals during the vegetation period (e.g. harvested fields). A comparison of the temporal mean NDVI and plant height could give insights into the perception of green within a city. For planning purposes a high degree of vertical green infrastructure could be necessary to reach goals like good quality of life, well-being and health.

For health studies the appearance of urban green could be of great importance. For that kind of investigation one should differentiate different heights of vegetation, due to the fact that human beings perceive vegetation depending on their size. Green meadows probably have another individual perception than a green forest. Since the NDVI does not account for the height of vegetation covers, it would make sense to create a limited number of height classes for studies on urban green and urban health or environmental inequality aspects.

To determine the vegetation height, one downloaded laser scan data of 2018 [30]. To be able to determine the object height, one extracted a digital terrain model (DTM) from the last pulse signal, and a digital surface model (DSM) from the first pulse signal, to be able to subtract the DTM from the DSM to receive the resulting object height of trees, buildings and other rather vertical objects in the area of investigation. The resulting individual object heights with 1m resolution can be seen in Figure 12 and classified into a few height classes in Figure 13 For consistency reasons, one resampled the 1m raster cells to 10m raster cells to do analysis with the 10m NDVI raster cells (Figure 14).

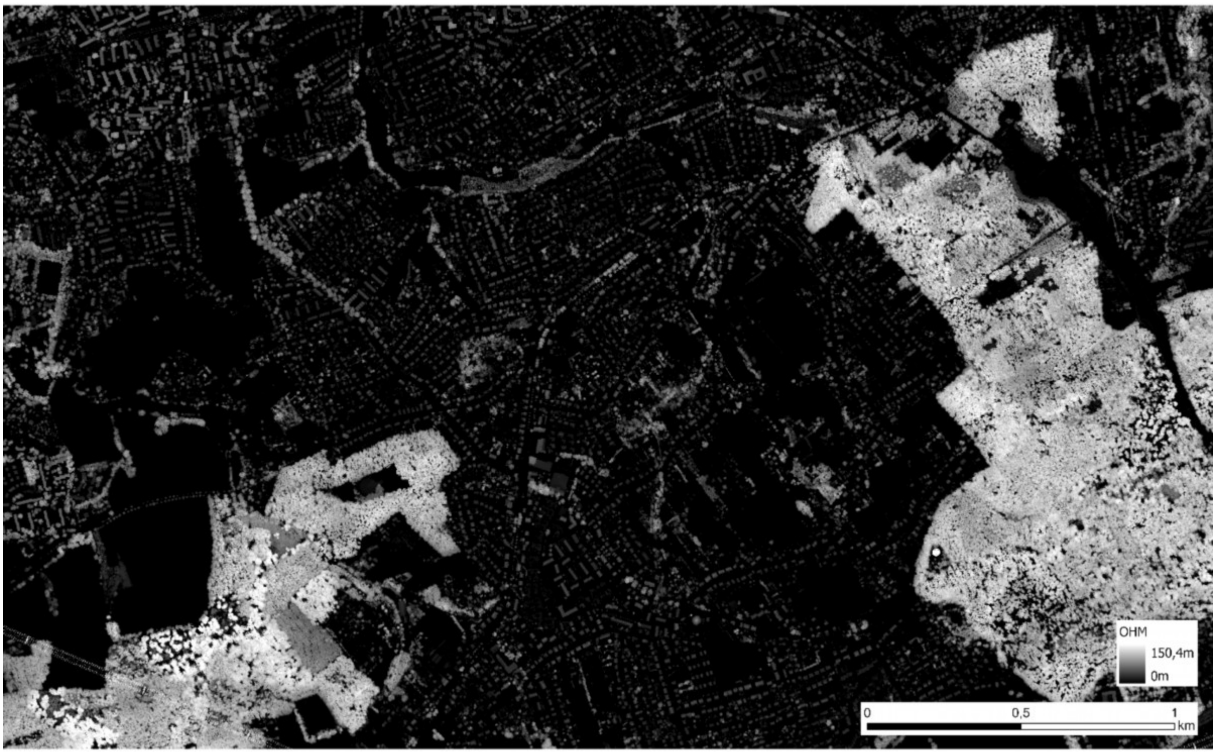

Figure 12. Digital object height model (1m raster cells) of the study area (height data source: [30])

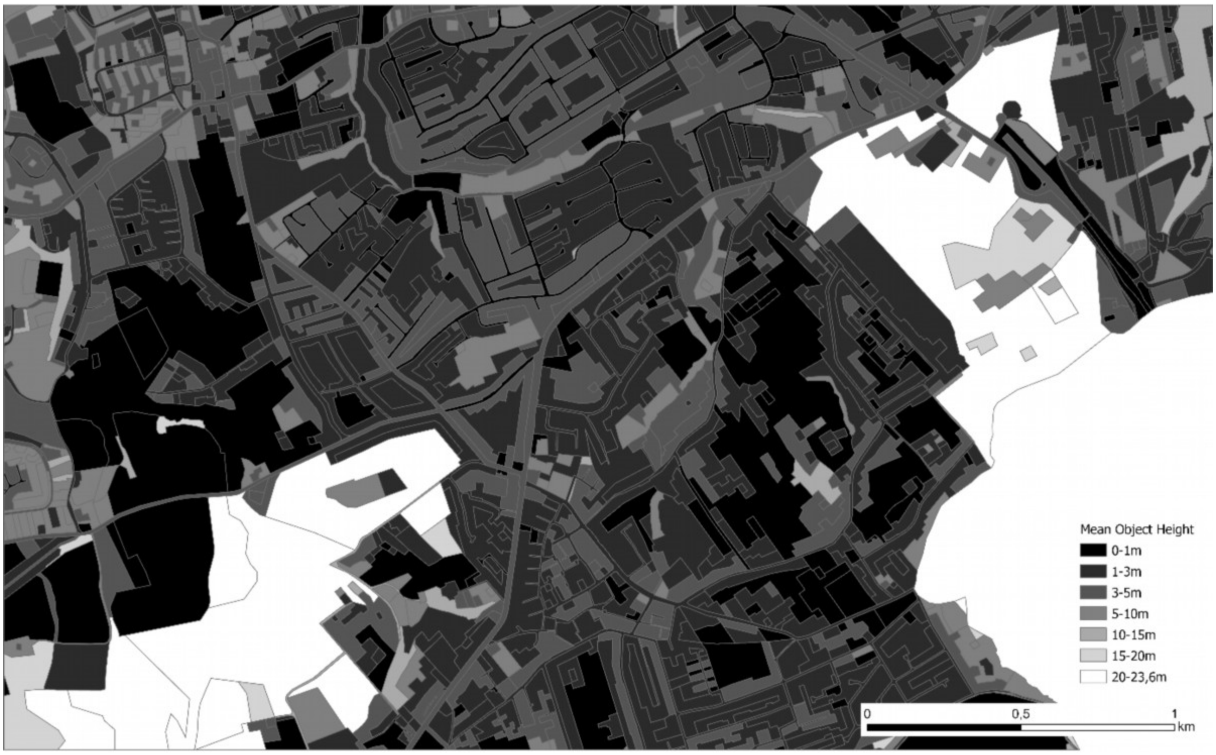

Figure 13. Averaged height classes per land cover/land use parcel (height data source: [30], data source polygon overlay: [32])

 doi:10.20944/preprints202008.0499.v1

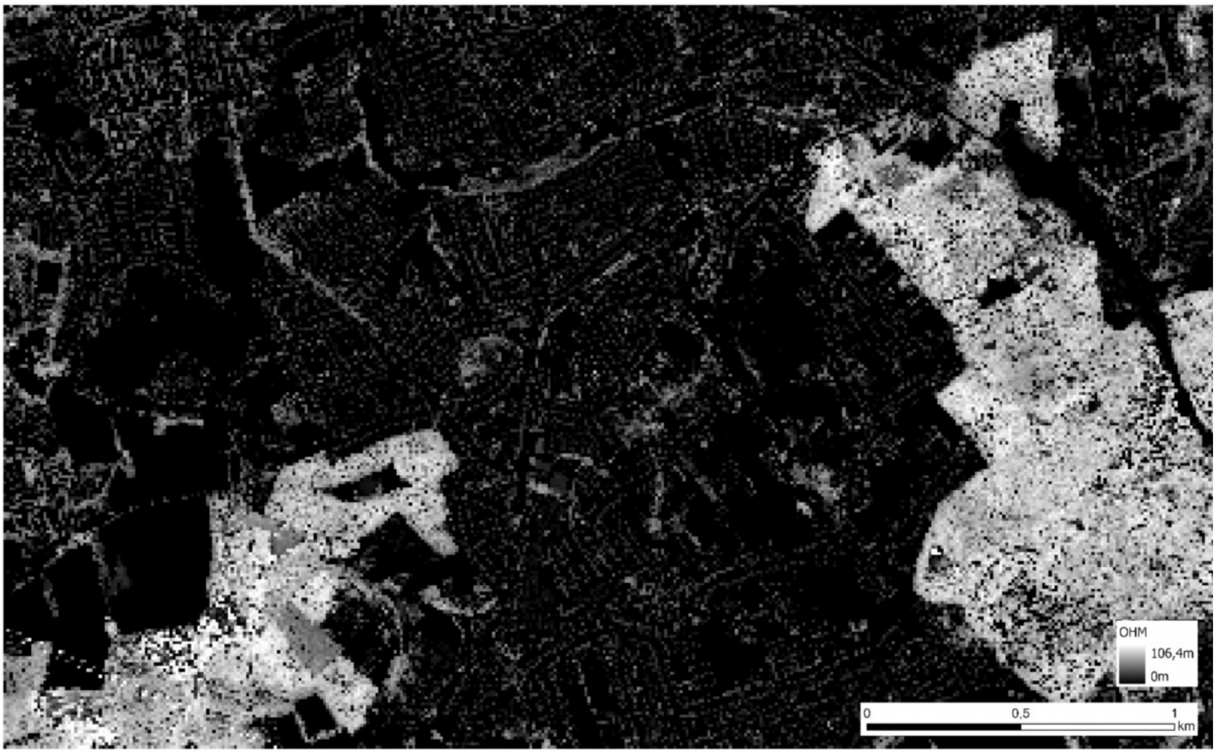

Figure 14. Digital object heights of the study area resampled to 10m resolution (height data source: [30])

The height measurements were then reduced into three meaningful categories: smaller 2m, 2-5m and above 5m height. This height categorization helps to distinguish vegetation categories as is demonstrated in Figure 15.

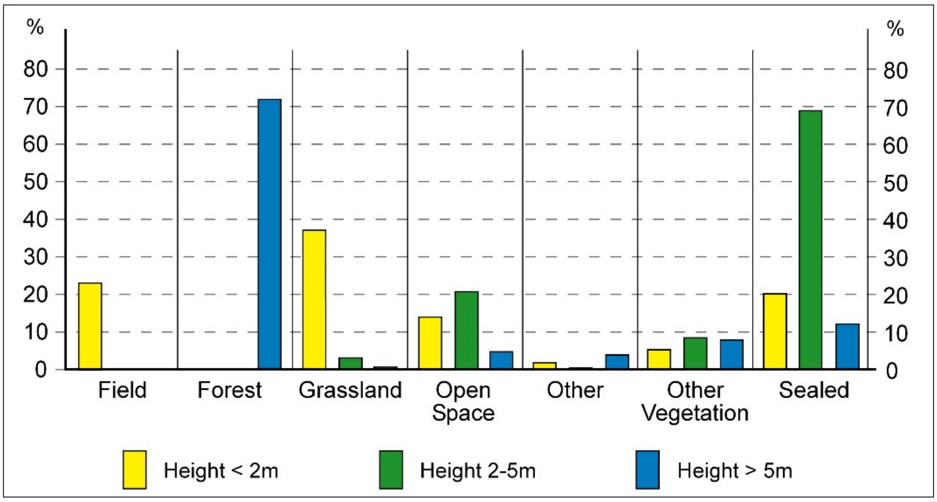

Figure 15. Distribution of land cover/land use categories for three height classes

The major observations regarding the height of the different land cover/land use categories can be summarized as follows. All agricultural fields are in the lowest height category, while the forest areas are completely in the highest height category. About 76% of the sealed surfaces are in the range of 2-5m. Sealed surfaces below 2m are very likely to be streets; higher sealed surfaces than 5m are very likely higher buildings. 89% of the grassland is lower than 2m. The remaining grassland probably is partly covered by trees and bushes. For the remaining classes (Open Space, Other & Other Vegetation) one finds out that their maximum ground coverage ranges between 2-5m (ca. 62%, ca. 59%, ca. 50%). The remaining areas are probably partly covered by trees and bushes as well.

Also the combination of spatial data sets in a bivariate choropleth map ([34], [28]) is a means to gain additional information from the used data sets. [28] applied three NDVI classes for the characterization of green urban infrastructure: <0.3; 0.3-0.6 and >0.6. The higher the value, the better is the amount of green infrastructure and its condition. This classification scheme with three categories was adopted here according to its practicability and the easy readability of the resulting map. Figure 16 shows the combination of mean object height and mean spatial and temporal NDVI (each per polygon) to evaluate if the vegetation height is of importance for the well-being and health aspects of the nearby population.

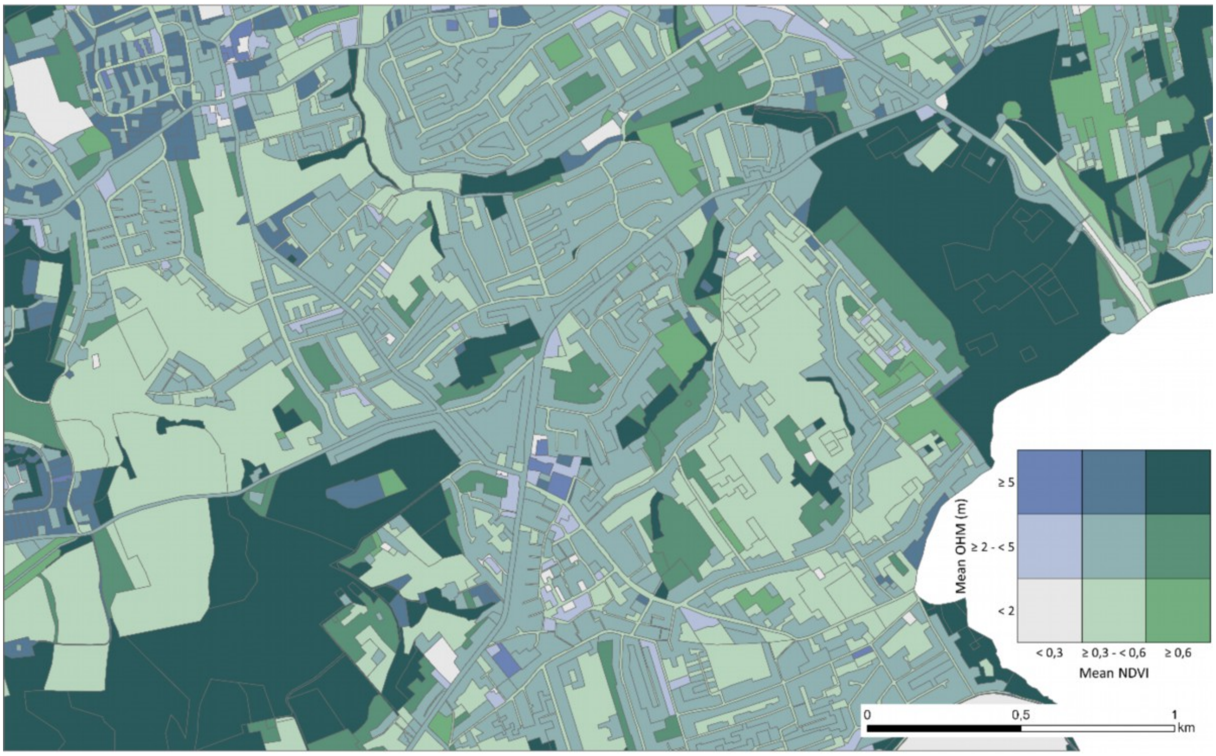

Figure 16. Bivariate choropleth map showing the mean spatial and temporal NDVI (data source [33]) per land cover/land use polygon against the mean object height per land cover/land use polygon

One can identify forested areas with the darkest map color due to their high height and due to their high NDVI values. On the opposite flat areas with low NDVI values represent sealed surfaces and fields. Many residential areas have intermediate NDVI values and a height between 2 and 5m (central raster cell of the legend). This could mean that residents have houses with similar height as the surrounding green area (which is represented by another polygon). So one could assume that if residents look out of their windows they mostly have some green vegetation in their view. This situation probably is advantageous for the residents´ perception of green vegetation and related health aspects. In the upper left corner one can identify a few high-rise buildings (>5m) and a NDVI class surrounding those buildings with vegetation heights less than 5m. This means that residents in the upper stories most likely do not see the vegetation from their windows. As the perception of green vegetation has positive effects on human health, the analysis of building heights and vegetation heights could be of value to identify areas with positive effects and other areas with deficits.

In addition to the bivariate choropleth map one was interested to learn how the land cover/land use categories correspond to the nine object height/NDVI-classes. For this analysis the bivariate choropleth map was calculated on a raster cell basis instead of parcel polygons (Figure 17). A simple frequency analysis of each of the nine classes reveals the class composition. To be able to address the individual combinations of object height and NDVI the nine fields were labelled as follows in Figure 18.

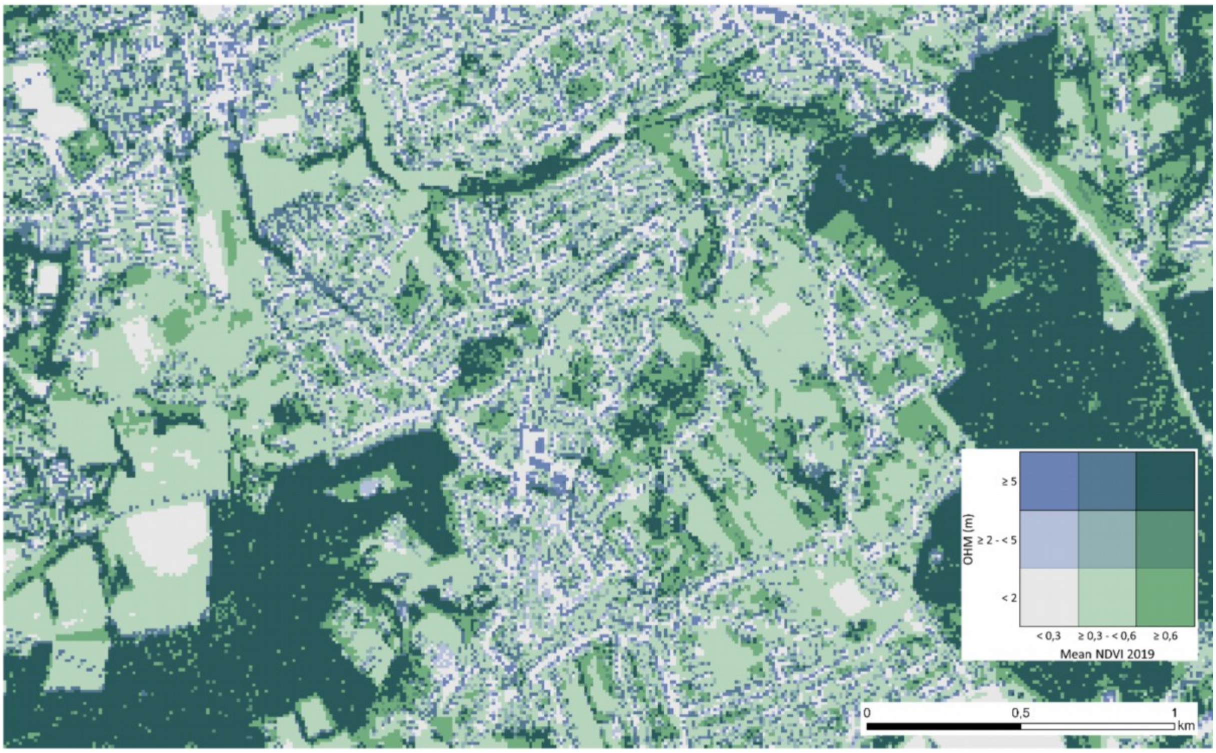

Figure 17. Bivariate choropleth map showing the mean spatial and temporal NDVI (data source [33]) per raster cell against the object height per raster cell

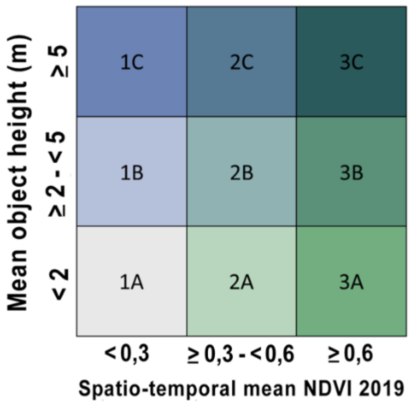

Figure 18. Labels of the legend of the bivariate choropleth map of Figure 16 and Figure 17

The frequency analysis of individual raster cells shows that the land cover/land use categories are distributed as follows (compare Figure 19):

Field 1A is composed mainly by sealed flat surfaces (71.07%) like parking areas or streets and agricultural fields (21.88%).

Field 1B is composed predominantly by sealed surfaces (97.26%) like small buildings.

Field 1C is composed predominantly by sealed surfaces (98.73%) like taller buildings.

Field 2A is composed similar like 1A with sealed flat surfaces like parking areas or streets (40.92%), agricultural fields (13.03%) and additional areas of grassland (23.22%), other vegetation (6.15%) and open space (15.18%).

Field 2B is composed mainly by sealed surfaces (73.42%) like small houses or open space (17.96%).

Field 2C is composed mainly by sealed surfaces (68.47%) like taller houses, other vegetation (6.11%) and open space (13.33%).

Field 3A is dominated by grassland (26.09%), open space (32.87%) and other vegetation (16.56%), in total rather flat vegetated surfaces. Also 12.38% sealed surfaces are present in this category. This is probably due to overhanging effects of vegetation (e.g. bushes) (as seen from the satellite) over sealed materials listed in the land cover/land use map.

Field 3B is composed by open space (33.22%), forest (21.22%), other vegetation (16.88%) grassland (26.09%) and sealed surfaces (16.26%). The sealed surfaces are represented in this category probably due to the same overhanging effects of vegetation (e.g. bushes) over sealed materials listed in the land cover/land use map.

Field 3C is composed predominantly by forest (72.23%)

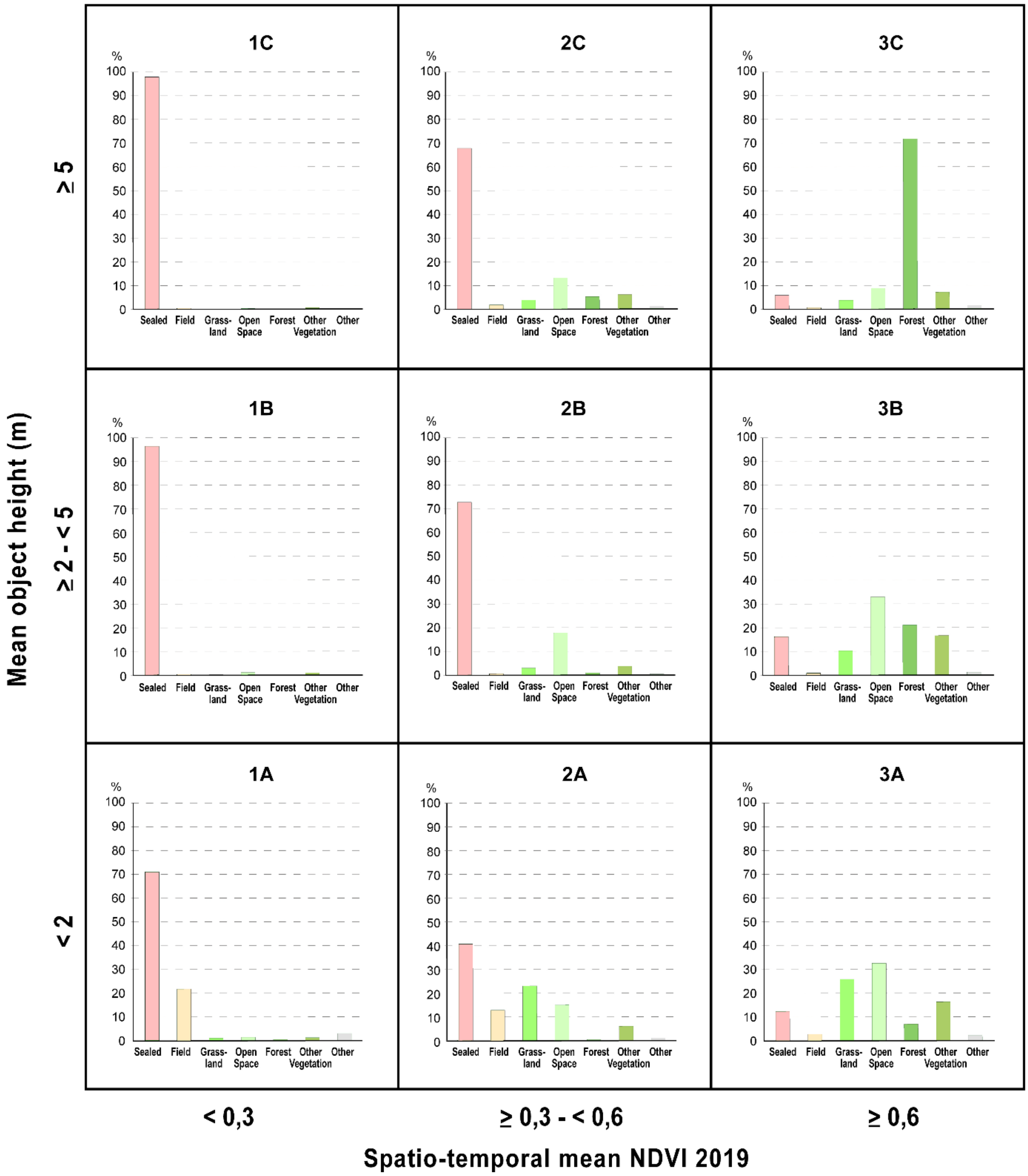

Figure 19. Frequency analysis of the land cover/land use composition of the classes of the bivariate choropleth map

Grassland is a typical flat land cover/land use type but one wonders why it is present in categories above 2m height. For this study we joined two originally separately mapped grassland types to keep things simple. Meadows and pastures were one category and the other was meadows and pastures with scattered fruit trees. In Figure 20 one can see the individual trees at 1m resolution. These individual trees integrated in the grassland class are the reason why grassland is represented in height classes above 2m.

Figure 20. Object height at 1m resolution for a grassland parcel. Individual fruit trees are clearly visible (height data source: [30])

Other land cover/land use classes that mostly appear flat like grassland, have the same phenomenon that bushes and trees are integrated in the class. This results in the appearance of these rather flat land cover/land use classes in categories higher than 2m and also in higher NDVI values. Even fields have high NDVI values with corresponding tall height information. This is because often trees are along field borders and so they were integrated in the analysis.

Most residential areas are characterized by category 2B, due to the fact that this is a mixture of small houses and garden vegetation mostly without very high trees. However, sometimes trees obscure the buildings and/or the sealed surface like roads with their branches and leaves overhanging the sealed material (see Figure 21). Some larger properties are characterized by solitary trees and meadows with bushes and fall into 3B.

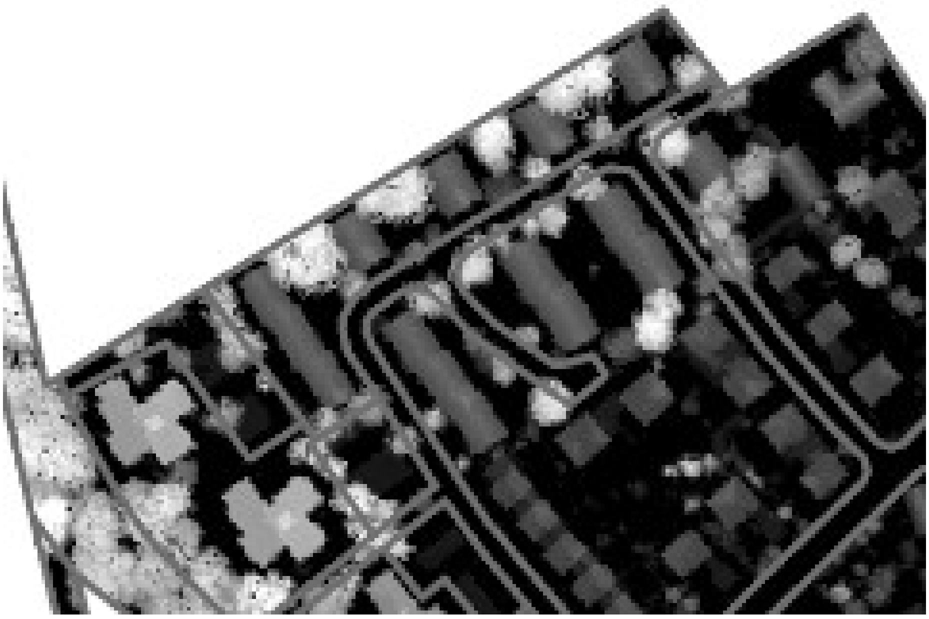

Figure 21. Tall vegetation obscuring sealed surfaces (height data source: [30])

The examples show that the explanation of the height inconsistencies can be performed by visual inspection of the relevant parcels. Scattered tall vegetation is present in almost all land cover/land use categories. The calculated relationships between NDVI and object height are valid and one can conclude that high mean NDVI values in the two-dimensional choropleth map correspond mostly to trees scattered in other categories. For an analysis interested in the value of green vegetation it is a benefit to know that even on rather flat green surfaces one can find tall vertical green structures that increase the value for recreation and health aspects.

3. Results

As described in [35] and [36], urban greenspace has to be considered as an environmental resource having positive health effects. This statement invites for further investigations regarding the green vegetation in urban areas. In this study we focussed on the NDVI as a widely used indicator of urban green vegetation and demonstrated a range of applications.

The results show the importance of NDVI calculations based on high resolution satellite images. Such images are of high timeliness and provide up-to-date information on vegetation surfaces. Due to the high repetition rate of earth observation satellites like Sentinel-2 with five days, one can also consider vegetation monitoring approaches to track vegetation changes within one vegetation period as well as between different years of observation. While Sentinel satellite images are freely available there seems to be no longer the necessity to laboriously extract greenspace information from existing map products with disadvantageous up-to-dateness and possibly questionable land cover/land use class definitions.

Using satellite images requires some geospatial data literacy for proper data selection and analysis. One requirement refers to the selection of satellite images. This implies to consider an image acquisition date with relevance to the intended vegetation analysis. That refers to phenological development stages of vegetation as well as to an acquisition date close to other (geo) data to be used together with the image(s). Fortunately, satellite images are available quite frequently. Optical images can suffer from clouds, but high revisit frequencies offer good chances for cloud-free images on a later date.

Besides the date of image acquisition, technical characteristics of the sensor system are also relevant. Here, one should think about the size of the image pixels. Smaller pixels offer more details and less mixed information. Especially in urban areas one can observe a large variety of surface materials that compose the image. Due to frequent surface material change, one can expect more than one surface material in a single pixel, if the pixel is too coarse. Smaller pixels offer a better chance to get more „pure“ pixels of one surface material. This means that high or very high resolution satellite images are advantageous for most studies.

In the Figures 3, 4, 5, and 6 one can identify seasonality in the individual seasonal images as well as in the NDVI images across the vegetation period. Some land-cover/land-use classes are very stable and others are highly variable. For vegetation especially forests are very stable, while fields and meadows are affected by harvest operations. As long as the crop/grass is not harvested, it appears in most cases as a green surface. Depending on the individual crop calendar, during a season one will find different stages of green vegetation.

Having the effects of seasonality in mind, one is interested to characterize the vegetation with respect to individual perception and health aspects. Having more than one image for NDVI calculation, one could generate a temporal mean NDVI. In case of forests one can observe relatively high NDVI values, which is caused by the stable green appearance of forests. In the case of fields or pastures the harvest leads to points in time with no or less green. The related low NDVI for those moments has a lowering effect on the temporal mean NDVI. In short, the temporal mean helps to judge the intensity and duration of green appearance.

The extraction of the temporal maximum NDVI provides a time stamp that identifies the moment with the highest green appearance. This could help to temporally arrange green vegetation areas for a well-balanced mixture of land covers/land uses for recreational and well-being aspects.

Due to the seasonality aspects, which offer additional analytical options, one can conclude, that the examination of the vegetation or greenspace via NDVI based on only one observation could not be sufficient for questions related to environmental justice, health or inequality.

Another aspect relates to vegetation height, since the perception of high vegetation covers like forest are more beneficial to the individual and the resulting positive health aspects. Depending on the land cover/land use class, varying typical heights could be identified.

The combination of height and NDVI revealed the spatial distribution of potential areas of high recreational and health value, as well as poor areas. Higher green areas are believed to have more positive effects on human health than lower green areas. For instance, areas with high NDVI and tall objects represent most likely forested areas. These represent relatively cool areas during summer months, which results in healthier life conditions. Analyzing the bivariate choropleth map could also assist in optimizing the urban land cover/land use mixture with respect to environmental health aspects. The frequency analysis of the land cover/land use classes related to the nine height/NDVI-complexes revealed which land cover/land use areals could be improved for a beneficial and healthy environment.

4. Discussion

To clearly structure the following observations/findings, we structure the discussion to address limitations as well as potentials of NDVI application in environmental justice, health and inequality studies.

4.1. Limitations

The application of NDVI values in urban environments is limited due to various reasons. First of all, the growing seasons of plants are different from plant species to plant species. This includes times when some areas are not covered with photosynthetic active vegetation. This is very clear for agricultural fields, e.g. there is no biomass after harvest. Also, deciduous trees lose their leaves and look different in winter and fall compared to spring and summer. However, this phenomenon of seasonality is quite normal and it could question other studies, which (directly or indirectly) assume a constant green situation throughout the year.

Another aspect is the degree of vegetation cover. In sparsely vegetated areas image pixels are composed of reflectance coming from vegetation and the soil, due to the fact that remote sensing systems have the vertical view on the earth´s surface. To describe the vegetation coverage one developed the leaf area index (LAI), which describes the amount of green leaves. All values above 1.0 describe plants (like trees) with more than 100% ground covered by vegetation due to the fact that the plant has more than one level of leaves. Values below 1.0 describe sparse vegetation covers with soil and vegetation associated in one pixel.

The pixel size affects the NDVI values as well. The smaller the pixel size, the higher is the chance that 100% of the pixel area is covered by vegetation. Larger pixels might have less vegetation cover and in addition to that also soil cover. So, larger pixels tend to result in mixed pixels, compared to smaller pixels. Such mixed pixels reduce the pure information content since they are composed from more than one land cover type [29]. This means that they do not represent one specific land cover type or class, but a mixture of at least two, with unknown spatial composition. One could try to get VHR images with small pixels to reduce the mixed spectral information. However, mixed pixels are present in any image. One can only try to reduce the area they represent by reducing the pixel size. As a rule of thumb one can say that the smaller the pixel size is, the smaller is the area affected/represented by mixed pixels.

Optical satellite images suffer from clouds. Cloudy situations obscure the view to the earth´s surface. Consequently, one needs cloud-free satellite images to calculate proper NDVI values. For time series investigations this could be problematic since data gaps destroy optimal time series analysis with equal interval image dates. Depending on the type of investigation, one could overcome this problem by calculation maximal NDVI values per season. Of course, then the seasonality information is lost. As a compromise one could track the date of the maximum NDVI value for each pixel to identify the exact date of the maximum NDVI (relative to the available cloud-free images).

To calculate mean spatial NDVI values it is not wise to do that for a certain administrative area. By doing this, one includes every surface material into the calculation. For instance, water and buildings or streets are included in the calculation of a mean NDVI value. This automatically will reduce the mean NDVI due to the fact that water and artificial surfaces have NDVI values close to 0 or below 0. Another problem could be the comparability of mean NDVI values of differently sized administrative areas.

4.2. Potential

The calculation of NDVI in urban areas has a high potential to identify relatively well-equipped greenspace areas with high potential for well-being and a healthy environment and on the opposite relatively poor-equipped green areas with rather low potential for well-being and a healthy environment. For instance, the high repetition rate of optical satellites like Sentinel-2 (five days) allows to detect changes in the NDVI response of the vegetation cover on a weekly basis. This could be the basis for a monitoring approach. After a certain monitoring period one could try to give advice to planners to improve the green situation for instance to have longer periods of visible green areas, due to height considerations. In this context one could also study the length of the green period of individual land cover types to perhaps find a good mixture of land cover types to have a long green period for a specific neighborhood.

The exact assessment of the urban vegetation is beneficial for the assessment of any local climate situation. Under the perspective of environmental justice, one could come to the conclusion that much vegetation (and therefore high NDVI values) corresponds with cooler air temperature in summer and results in healthier life conditions. This means the degree of green vegetation is an environmental indicator/parameter that is related to health risks. On the contrary the degree of

5. Conclusions

This study was motivated to investigate the limitations and potential of NDVI with other spatial data for application in the field of environmental justice and inequality related to health and recreation in urban environments. It is clear that inequality of environmental settings influence the individual health situation causing health equity or inequity. It was intended to demonstrate a literate approach to use NDVI information and to point out potential problems or drawbacks. From an urban test site in Dortmund, Germany one can draw many conclusions which are summarized hereafter.

NDVI calculations from remotely sensed earth observation images is an easy task but needs some degree of data literacy. One should be aware of the later use of generated data and be able to judge which image acquisition date is appropriate. In case of time series data for monitoring purposes one should be aware of clouds, that could obscure the ground and affect the NDVI calculation.

For some studies a mean NDVI could be of interest. One should consider two types of mean NDVI calculations. In one case the mean over time for exactly the same location is calculated to have a mean value for e.g. one year at the same location (e.g. land cover type or plant association). In another case one could calculate the mean value for an area (e.g. administrative or statistical unit) and across all land cover types. The resulting mean NDVI would give an idea on how much or how less green is in this area but does not give any spatial differentiation.

The calculation of the maximum NDVI makes sense only for a time series like a vegetation period, to identify the date of the maximum chlorophyll activity.

The NDVI values per pixel help to determine the plant activity and vice versa allows to identify the environmental burden (e.g. heat). In cases with much vegetation and high NDVI values respectively the environmental burden is rather low.

To assist in statements related to environmental burdens, well-being or health issues, maps are helpful to visualize and locate environmental concerns or consternation of the population, e.g. the heat vulnerability due to low quality green vegetation infrastructure. The combination of height and

NDVI revealed the spatial distribution of potential areas of high recreational and health value, as well as poor areas.

Earth observation allows to map all green spaces in an urban area and is not limited to public ground like tree cadastres. For instance, this supports the evaluation of urban micro climatic conditions. A high degree of vegetation (e.g. trees) generates cooling effects of the neighborhood and therefore leads to healthier conditions of life in the particular neighborhood.

Supplementary Materials: The original data sources are available under: https://geo-cloud.geographie.ruhr-uni-bochum.de/index.php/s/eKMxqWGYHzSbc9P

Author Contributions: Conceptualization, C.J.; methodology, C.J. and M.F.M.-H.; geo-processing, M.F.M.-H.; Writing – original draft preparation, C.J.; writing – review & editing, C.J. and M.F.M.-H.; visualization, C.J. and M.F.M.-H. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding

Acknowledgments: This paper is supported by the project no. 2019-1-CZ01-KA203-061374 Spatial and economic science in higher education - addressing the playful potential of simulation games (Spationomy 2.0) funded by the European Union within the Erasmus+ program. The support is highly appreciated.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Albertz, J. Einführung in die Fernerkundung: Grundlagen der Interpretation von Luft- und Satellitenbildern, 5th ed.; Wissenschaftliche Buchgesellschaft: Darmstadt, Germany, 2016
2. Lillesand, T.M.; Kiefer, R.W.; Chipman, J. Remote Sensing and Image Interpretation, 7th ed.; Wiley: New York, NY, USA, 2015.
3. Jürgens, C. Earth Observation Data. In SPATIONOMY—Spatial Exploration of Economic Data and Methods of Interdisciplinary Analytics; Pászto, V., Jürgens, C., Tominc, P., Burian, J., Eds.; Springer: Cham, Switzerland, 2020; pp. 30–38, doi:10.1007/978-3-030-26626-4.
4. Juergens, C.; Crespi, M.; Maktav, D.; Goossens, R.; Jacobsen, K. & Nascetti, A. Foreword to the European journal of remote sensing special issue: urban remote sensing – challenges and solutions. *European Journal of Remote Sensing* **2019**, 52:sup2, 1, DOI: 10.1080/22797254.2019.1643574
5. Roman, A.; Ursu, T. Multispectral satellite imagery and airborne laser scanning techniques for the detection of archaeological vegetation marks. In: Landscape Archeology on the northern frontier of the Roman Empire at Porolissum – an interdisciplinary research project; Opreanu, C.H.; Lăzărescu, V.-A., Eds.; Mega Publishing House, 2016; pp.141-152
6. Jordan , C. F. Derivation of leaf-area index from quality of light on the forest floor. *Ecology* **1969**, Vol. 50, no. 4, pp. 663–666
7. Xue J.; Su, B. Significant Remote Sensing Vegetation Indices: A Review of Developments and Applications. *Journal of Sensors* **2017**, Article ID 1353691, https://doi.org/10.1155/2017/1353691
8. Rouse Jr. J. W.; Haas, R.; Schell, J.; Deering, D. Monitoring vegetation systems in the great plains with ERTS. NASA Special Publication 351, 1974, 309
9. Huete, A. R. A soil-adjusted vegetation index (SAVI). *Remote Sensing of Environment* **1988**, Vol. 25, no. 3, pp. 295–309
10. Baret, F.; Jacquemoud, S.; Hanocq, J. F. The soil line concept in remote sensing. *Remote Sensing Reviews* **1993**, Vol. 7, no. 1, pp. 65–82
11. Jurgens, C. The modified normalized difference vegetation index (mNDVI) a new index to determine frost damages in agriculture based on Landsat TM data. *International Journal of Remote Sensing* **1997**,Vol. 18, Issue 17, pp. 3583-3594, DOI: 10.1080/014311697216810
12. Beck, P.; Karlsen, S.; Skidmore, A.; Lennart, N.; Høgda, K. The onset of the growing season in northwestern Europe, mapped using MODIS NDVI and calibrated using phenological ground observations. 2005. https://www.isprs.org/proceedings/2005/ISRSE/html/papers/261.pdf
13. Feng, L.; Song, G.; Zhu, L.; Yanan, Z.;Di, L. Urban vegetation phenology analysis using high spatio-temporal NDVI time series. *Urban Forestry & Urban Greening* **2017**. 25, pp. 43-57, DOI: 10.1016/j.ufug.2017.05.001

14. Bolund, P.; Hunhammar, S. Ecosystem services in urban areas. *Ecol. Econ.* **1999**. 29, 293–301
15. IPCC. Fourth Assessment Report: Climate Change. 2007. http://www.ipcc.ch/publications_and_data/ar4/syr/en/contents.html
16. Flacke, J.; Schüle, S.A.; Köckler, H.;Bolte, G. Mapping Environmental Inequalities Relevant for Health for Informing Urban Planning Interventions—A Case Study in the City of Dortmund, Germany. *Int. J. Environ. Res. Public Health* **2016**. 13(7), 711, https://doi.org/10.3390/ijerph13070711
17. Bolte, G.; Bunge, C.; Hornberg, C.; Köckler, H. Umweltgerechtigkeit als Ansatz zur Verringerung sozialer Ungleichheiten bei Umwelt und Gesundheit. *Bundesgesundheitsbl* **2018**. 61, 674–683. https://doi.org/10.1007/s00103-018-2739-6
18. Wagner G.G. Über den Tellerrand hinaus – eine Fernerkundungsinfrastruktur für bessere Sozialforschung!. In: Globale Urbanisierung. Taubenböck H.; Wurm M.; Esch T.; Dech S., Eds. Springer Spektrum, Berlin, Heidelberg, 2015. https://doi.org/10.1007/978-3-662-44841-0_33
19. Zepp, H.; Inostroza, L.; Sutcliffe, R.; Ahmed, S.; Moebus, S. Neighbourhood Environmental Contribution and Health. A novel indicator integrating urban form and urban green. *Change and Adaptation in Socio-Ecological Systems*, **2018**, 4(1), 46-51. doi: https://doi.org/10.1515/cass-2018-0005
20. Helbich, M. Spatiotemporal Contextual Uncertainties in Green Space Exposure Measures: Exploring a Time Series of the Normalized Difference Vegetation Indices. *Int. J. Environ. Res. Public Health* **2019**, 16, 852
21. Lee, H.-Y.; Wu, C.-D.; Chang, Y.-T.; Chern, Y.-R.; Lung, S.-C.C.; Su, H.-J.; Pan, W.-C. Association between Surrounding Greenness and Mortality: An Ecological Study in Taiwan. *Int. J. Environ. Res. Public Health* **2020**, *17*, 4525, https://doi.org/10.3390/ijerph17124525
22. Möller, M.; Bartlomiejczak, K. Urbanes Grün und Stadtklima: Analyse von Vegetation in europäischen Großstädten.- *gis.Business3/2020*, p. 21-23
23. Colunga, M. L.; Cambrón-Sandoval, V. H.; Suzán-Azpiri, H.; Aurelio Guevara-Escobar, A. & Luna-Soria, H. The role of urban vegetation in temperature and heat island effects in Querétaro city, Mexico. *Atmósfera* **2015**, Volume 28, Issue 3, p. 205-218, https://doi.org/10.20937/ATM.2015.28.03.05
24. Hart, M.A.; Sailor, D.J. Quantifying the influence of land-use and surface characteristics on spatial variability in the urban heat island. *Theor. Appl. Climatol.* **2008**, 95, 397–406.
25. Duarte, D.; Shinzato, P.; Gusson, C. & Alves, C. The impact of vegetation on urban microclimate to counterbalance built density in a subtropical changing climate. *Urban Climate* **2015**. 14. 10.1016/j.uclim.2015.09.006.
26. Alves, E.D.L. & Lopes, A. The Urban Heat Island Effect and the Role of Vegetation to Address the Negative Impacts of Local Climate Changes in a Small Brazilian City, *Atmosphere* **2017**, 8, 18; doi:10.3390/atmos8020018
27. Jürgens C. Urbane Fernerkundung mit Satellitenbildern:gestern – heute – morgen. In: Globale Urbanisierung. Taubenböck H.; Wurm M.; Esch T.; Dech S., Eds. Springer Spektrum, Berlin, Heidelberg, 2015, https://doi.org/10.1007/978-3-662-44841-0_29
28. Moos, N. Soziogeomatik – Möglichkeiten und Grenzen der Verwendung von Erdbeobachtungsdaten und Geodaten zusammen mit soziodemographischen und sozioökonomischen Daten. Ph.D. Thesis, Ruhr-University Bochum, Germany, July 14, 2020
29. Juergens, C. Digital Data Literacy in an Economic World: Geo-Spatial Data Literacy Aspects. *ISPRS Int. J. Geo-Inf.* **2020**, *9*(6), 373, https://doi.org/10.3390/ijgi9060373
30. Land NRW - WMS NW DTK25 – dl-de/by-2-0 (www.govdata.de/dl-de/by-2-0), available online as WMS: https://www.wms.nrw.de/geobasis/wms_nw_dtk25? (Accessed on 15 August 2020)
    Title of Site. Available online: URL (accessed on Day Month Year)
31. BKG – WMS Digitale Topographische Karte 1:500.000 Graustufendarstellung, available online as WMS: https://sgx.geodatenzentrum.de/wms_dtk500_grau (Accessed on 15 August 2020)
32. RVR - Flächennutzungskartierung Regionalverband Ruhr, Essen, Referat 9, Team Geodaten, Stadtplanwerk, Luftbilder Stand 2017
33. ESA Copernicus Open Access Hub Available at scihub.copernicus.eu/dhus/ (Accessed on 13 April 2020)
34. Götze, W.; van den Berg, N. Techniken des Business Mapping. De Gruyter, Berlin, Boston (= Managementwissen für Studium und Praxis), 2017
35. Flacke; J.; Köckler, H. Spatial urban health equity indicators – a framework-based approach supporting spatial decision making. *WIT Transactions on Ecology and The Environment* **2015**. Vol. 193, DOI 10.2495/SDP150311

 doi:10.20944/preprints202008.0499.v1

36. Krekel, C.; Kolbe, J.; Wüstemann, H. The Greener, The Happier? The Effects of Urban Green and Abandoned Areas on Residential Well-Being. SOEP papers 728, 2015.