The Spatial and Temporal Distribution of Urban Vegetation Phenology by Local Climate Zone and Urbanization Level

A Capstone Project Paper

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Master of Professional Studies

by

Ruihan Liu

Advisor: David G. Rossiter

2025.05.13

The Spatial and Temporal Distribution of Urban Vegetation Phenology by Local Climate Zone and Urbanization Level

Abstract

This report is for a capstone project for Cornell MPS, concentration Geospatial Applications, to illustrate mastery of a complex geospatial analysis. It used a wide variety of techniques and tools. The project investigated an important topic in the context of the urban environment under climate variability: the spatial-temporal distribution of vegetation phenology in Beijing, focusing on the impacts of local climate zones (LCZ) and urbanization levels. Using MODIS MCD12Q2 data, we analyzed the start of season (SOS), end of season (EOS), and length of season (LOS) across different urbanization gradients and LCZs. Results reveal that urbanized regions exhibit earlier SOS and longer LOS, while natural areas have a later EOS and shorter LOS. LCZs with lower density buildings tend to have a more extended growing season. These findings underscore the significance of urbanization and LCZ in shaping vegetation phenology, providing valuable insights into climate change adaptation and urban ecological management.

1. Background and Reason

Amid the dual challenges of global warming and rapid urbanization, increasing attention has been paid to changes in vegetation phenology within urban environments (Ren et al., 2018). Vegetation phenology, a sensitive indicator of climate change, reflects the timing of key biological events such as green-up, maturity, and senescence. Understanding its spatial and temporal variations is crucial for evaluating ecological responses to anthropogenic impacts and for informing urban sustainability strategies .

Previous studies have primarily focused on large-scale regions, such as national or continental areas, with fewer investigations targeting the fine-scale dynamics of phenology within heterogeneous urban landscapes (S. Li et al., 2023). Moreover, existing research often centers on megacities or urban agglomerations (Zhu et al., 2024), lacking detailed differentiation among smaller urban units that vary in their levels of urbanization and land surface characteristics (Zhang et al., 2022).

This study aimed to fill that gap by examining the spatial-temporal patterns of vegetation phenology within a single metropolitan region, using a combination of phenological metrics and urban form classifications. Specifically, this research leveraged remote sensing datasets—such as MODIS phenology products (MCD12Q2), GAIA impervious surface data, and the Local Climate Zones (LCZ) classification—to quantify variations in Start of Season (SOS), End of Season (EOS), and Length of Season (LOS) across urban, suburban, and rural gradients. Additional datasets, including building height and surface data from GHSL, helped further refine the analysis of urban structure.

By investigating how phenological changes correlate with different urbanization intensities, impervious surface coverage, and LCZ types, this study offers a nuanced understanding of the interactions between urban form and vegetation dynamics (Zhou et al., 2022). It seeks to determine whether higher levels of urbanization necessarily lead to longer or altered growing seasons, and whether certain combinations of urban features (e.g., high building density and low vegetation cover) are more strongly associated with phenological shifts.

Ultimately, this research contributes to the growing body of literature emphasizing the importance of small-scale, high-resolution phenological monitoring (Hu et al., 2024). It offers a novel methodological framework for exploring how urban morphology modulates ecological processes, helping to inform climate-resilient urban planning and green infrastructure development (Zhao et al., 2022).

2. Materials and methods

2.1 Data

2.1.1 Vegetation Phenology parameters

The vegetation phenology parameters are included in the MCD12Q2 MODIS dataset. This dataset records multiple phenological indicators, including greenness, greenup midpoint, maturity, peak greenness, senescence, greendown midpoint and dormancy, with 500m resolution, available since 2000. In this research, we changed the raw data (days since 1970-01-01) into day of year (DOY) for each year. We defined the date when EVI2 first crossed 15% of the segment EVI2 amplitude as the start pf season(SOS), defined the date when EVI2 last crossed 15% of the segment EVI2 amplitude as the end of the season(EOS), and defined the length of the season(LOS) as the difference of the EOS and SOS.

2.1.2 Local Climate Zone

The Local Climate Zone (LCZ) classification system categorizes cities and their surrounding areas based on surface characteristics such as building density, vegetation cover, and surface materials. It is primarily used in urban climate research, including analyses of the urban heat island (UHI) effect and the environmental impacts of urbanization. In this study, we used the LCZ classification proposed by Stewart and Oke (Stewart & Oke, 2012). This is mainly applied to urban climate research, such as the analysis of urban heat island effect and the impact of urbanization on the environment. The LCZ data was from Demuzere's research(Demuzere et al., 2022), which can be download directly, with 100m resolution. During the process of making this data, Sentinel-2 and Landsat were used to get the surface characteristics, MODIS was used to get land surface temperature, Topographic data was from OpenStreetMap. So the LCZ data combined remote sensing data and geographic information data by using machine learning or image processing methods. And it has 17 categories, including 10 built-up types(LCZ1 - LCZ10) and 7 natural types(LCZA - LCZG). The LCZ data represents the year 2018 nominally, but its application is not limited to 2018 studies, with 100m resolution.

LCZ	Name
LCZ1	Compact high-rise
LCZ2	Compact mid-rise
LCZ3	Compact low-rise
LCZ4	Open high-rise
LCZ5	Open mid-rise
LCZ6	Open low-rise
LCZ7	Lightweight low-rise
LCZ8	Large low-rise
LCZ9	Sparsely built
LCZ10	Heavy industry

LCZA	Dense trees
LCZB	Scattered trees
LCZC	Bush, scrub
LCZD	Low plants
LCZE	Bare rock or paved
LCZF	Water
LCZG	Snow or ice

2.1.3 DEM

The DEM data source is NASADEM, which can be download from Google Earth Engine, with 30m resolution.

2.1.4 Impervious Surface Area

The GAIA data (X. Li et al., 2020) is available from 1985 to 2018 with 30m resolution, made by Tsinghua University, showing information of global impervious surface area. In each year, 1 represents impervious area, 0 represents natural area.

2.1.5 Building Height and Surface

The built-up height data and built-up surface data are both from Global Human Settlement Layer (GHSL) and represent the actual situation in 2018. The built-up surface data represents the amount of square meters of built-up surface in the cell 10m*10m. The built-up height data represents average height of all building volumes per unit area (building volume/surface area), for example, if there are $10{,}000\text{m}^3$ building volumes (the sum of all building volumes) in a cell distributed on a 100m × 100m grid ($10{,}000\text{m}^2$), then value= $10{,}000\text{m}^3$ / $10{,}000\text{m}^2$ = 1m. The building height's spatial resolution is 100m, building surface's resolution is 10m.(Florczyk et al., 2019)

2.2 Study area

Beijing is located at 40 degrees North latitude, with distinct seasons, most of the vegetation is temperate deciduous broad-leaved forest, with obvious seasonal changes. Moreover, Beijing is rich in topography, diverse in feature types, and wide in LCZ classes. As a rapidly urbanizing mega-city, Beijing has experienced significant urban area expansion in the past few decades. It has research value on both spatial and temporal scales.

In summary, it can be an ideal place to research the relationship between the vegetation phenology and local climate zone or urbanization level.

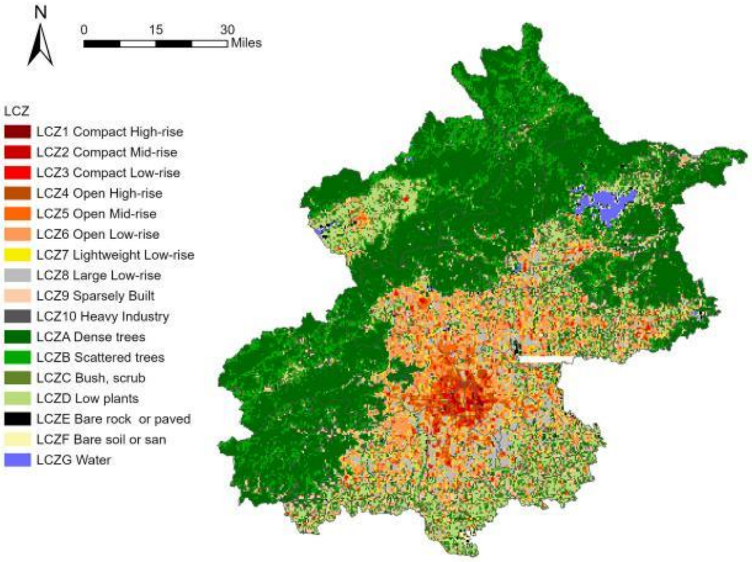

Figure 1. The local climate zone distribution in Beijing, China

But the difference in elevation is something to pay attention to, because the urban areas of Beijing are mainly distributed in plain areas, while the non-urban areas (mainly in the west and north of Beijing) are mostly mountainous and have higher elevations. The difference in altitude will cause the difference in temperature, which will have an impact on LOS, so in the subsequent analysis, in addition to showing spatial distribution of LOS, the change of LOS is more important disentangle the effects of urbanization from secular climate trends.

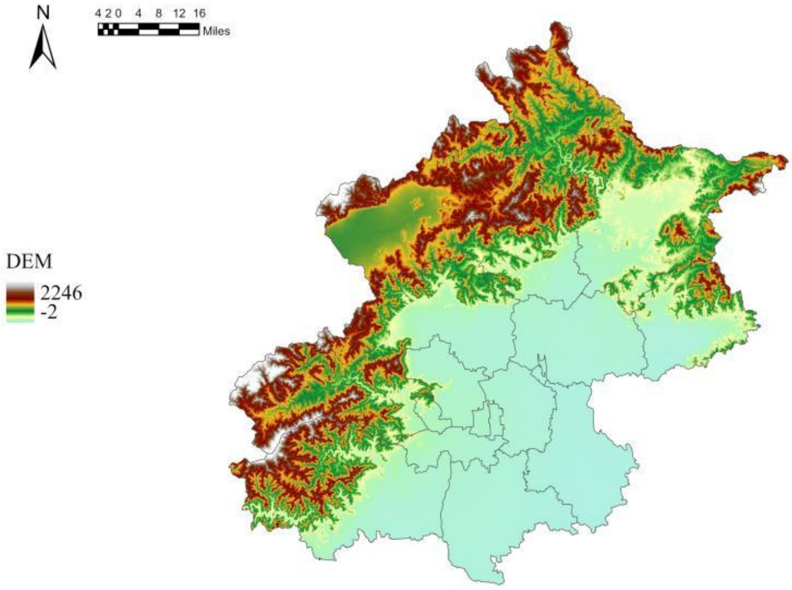

Figure 2. The DEM in Beijing, China

2.3 **Method**

2.3.1 Spatial-Temporal Distribution of Vegetation Phenology

The MCD12Q2 is available in Google Earth Engine(GEE), but for the raw data, it's value is shown as the days since 1907-01-01, so it can be calculated into the day of year(DOY) in each year in GEE. After downloading the data in each year, the raster data was clipped by the Beijing boundary in ArcGIS Pro. To see the spatio-temporal distribution of the phenology indicators, the maps of SOS, EOS and LOS in 2010, 2013 and 2016 were made. To see the temporal distribution, the means of all three vegetation phenology parameters in each year from 2008 to 2018 were calculated and plotted along with their linear regression results. A visual inspection of the trends showed that a linear relation over this time period is justified. For more in-depth analysis, the LOS is the point of my later analysis. This is because it is the most important for the total amount of green vegetation, integrated over the year, and thus is the best indicator of the effect of phenology on the urban environment.

2.3.2 Spatial-Temporal Change of Vegetation Phenology

In order to see the spatial change of LOS from 2008 to 2018, the difference between the value of 2008 and that of 2018 was calculated in the ArcGIS Pro, and shown as a map. To better understand the difference of the LOS between the urban areas and non-urban areas, the GAIA data was used to distinguish the urban and non-urban areas. Using the R environment for statistical computing (R Core Team, 2023) , the region with a value of 1 in the GAIA data, that is, the urban area, and the part with a value of 0 in the GAIA data, that is, the non-urban area, their LOS means for each year were calculated respectively, and linear fitting was performed according to the scatter plot of the means. Again, visual inspection showed that a linear fit is appropriate.

2.3.3 Distribution of LOS on different urbanization level

The resolution of GAIA data is 30m, but the resolution of LOS is 500m. Therefore the GAIA data was resampled from 30m to 500m. During this downsampling, the proportion of impermeable pixels (value 1) in the total number of pixels in each 500m grid was calculated. This proportion represents the proportion of impervious surface area (ISA). ISA can be used an indicator to represent different urbanization levels. Usually, the higher the ISA, the higher the urbanization level. We divided, ISA into 10 intervals at 10% intervals, and the corresponding LOS mean was calculated in each interval. The above operation was carried out for every year, and the mean of all years within the study range was calculated and displayed. In this way, not only can the change of year by year be observed, but also how LOS changes with the change of ISA.

2.3.4 Distribution of LOS within Various LCZs

In order to better observe the difference of LOS in different LCZ, I made the boxplot of LOS in different LCZ. However, the category of LCZ is more complex. Compared with ISA, which only represents the urbanization level as single indicator, LCZ contains more information than only the impervious surface area ratio, especially LCZ1-LCZ9, which usually contains different combinations of building height and building density. In order to better understand the impact of urbanization on LOS, the effect of building height and density on LOS is worth further analysis. Here, because LCZ was generated by machine learning, although it can be applied to studies in other years, it is most accurate in its training year of 2018, so further analysis of 2018 LCZ1-LCZ9 was carried out. For this, binary heat maps of LOS corresponding to different building heights and densities were produced. After resampling, the building height and density data were unified with LOS, and the three kinds of raster data were stacked for statistics, and the mean change of LOS under different building heights and densities was found.

3. Results

3.1 Spatial-Temporal Distribution of Vegetation Phenology

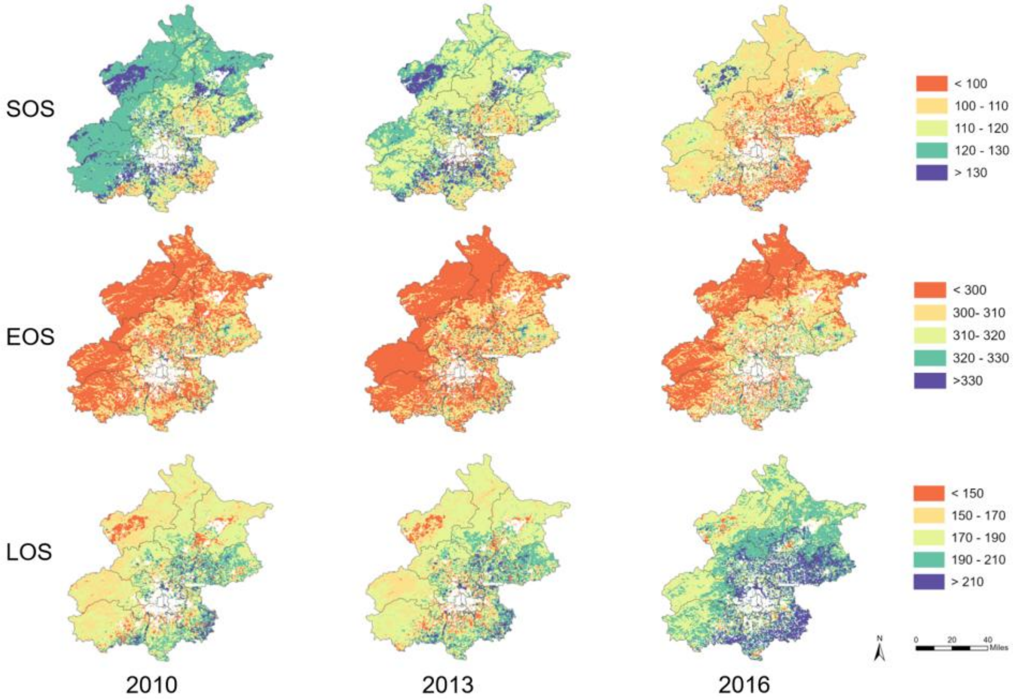

Figure 3. The three phenology parameters for three years

For the SOS, the central and southeastern regions (yellow to orange colors, lower values) show an earlier start of the growing season, strongly influenced by urbanization, suggesting higher temperatures or urban heat island effects. In the northwest, SOS will start later, possibly due to the higher altitude in the mountains and the slower warming. For EOS, in the northern and western regions (green and blue colors, higher values), it ends earlier, possibly due to more abundant natural vegetation and cooler conditions. But the EOS of southeastern regions ends later. SO, for the LOS, the central regions have longer length of season, reflecting the impact of urbanization. But there it’s also possible that the difference is influenced by the different elevations.

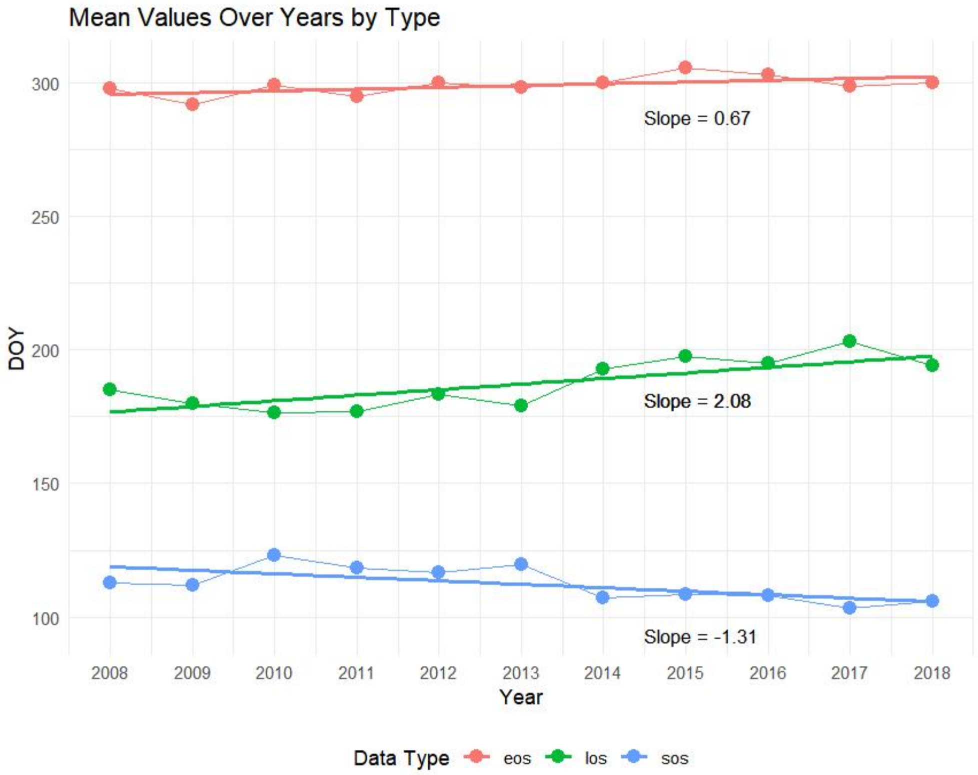

Figure 4. Annual average of three phenology parameters from 2008 to 2018

The graph shows that the trend of SOS, EOS and LOS. They tend to have longer season resulting from earlier starts and later ends. The slope of LOS has the greatest absolute value, and shows a clear trend of about two days per year. These trends highlight the impact of environmental changes, such as urbanization and climate change on vegetation phenology.

3.2 Spatial-Temporal Change of Vegetation Phenology

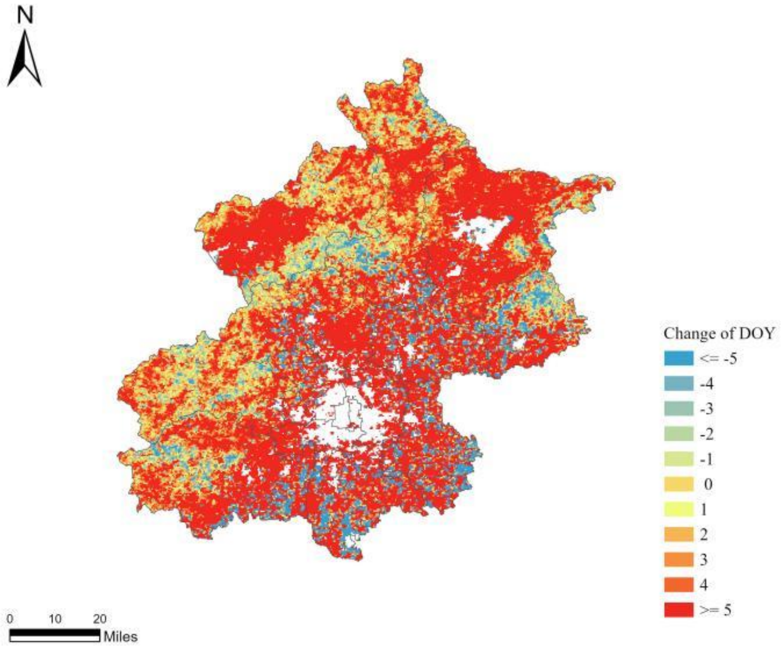

Figure 5. Change of length of season from 2008 to 2018

To determine whether the longer LOS present in urban areas is an effect of urbanization or mainly due to topographic reasons, I calculated the change in LOS from 2018 to 2008. The results shows that except for the western region, LOS remained basically unchanged or had a small increase (about 1-2 days on average), most of the other urban areas had a relatively large increase, and a large range of areas had an average increase of more than 5 days. In short, the growth rate of urban areas is much higher than that of the western mountainous areas, which can indicate that urbanization plays a promoting role in the extension of LOS.

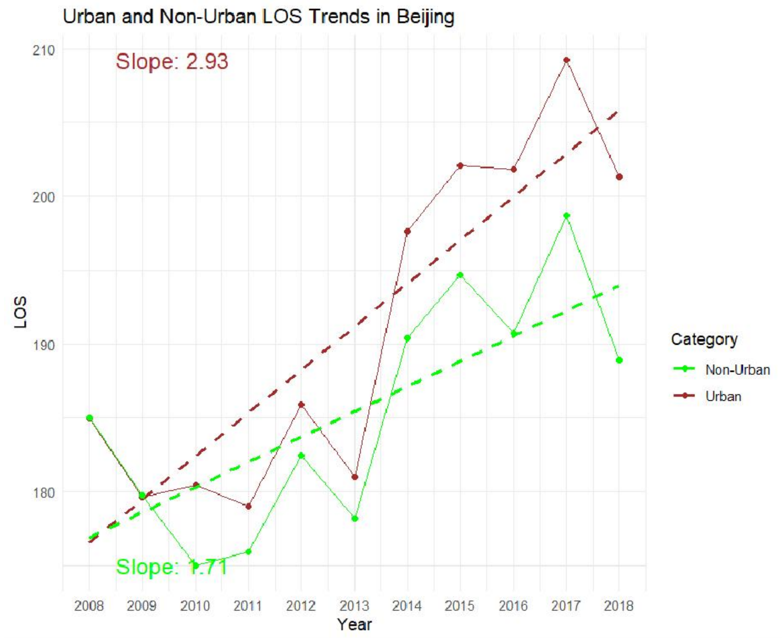

Figure 6. Annual average of LOS in each year for urban area and non-urban area

The results showed that LOS increased year by year in both urban and non-urban areas, but the LOS was longer in urban areas. The linear regression computes a growth rate of the urban areas as 2.93, which is higher than that of non-urban areas (1.71), indicating that this is likely due to the influence of urbanization

3.3 Distribution of LOS on different urbanization level

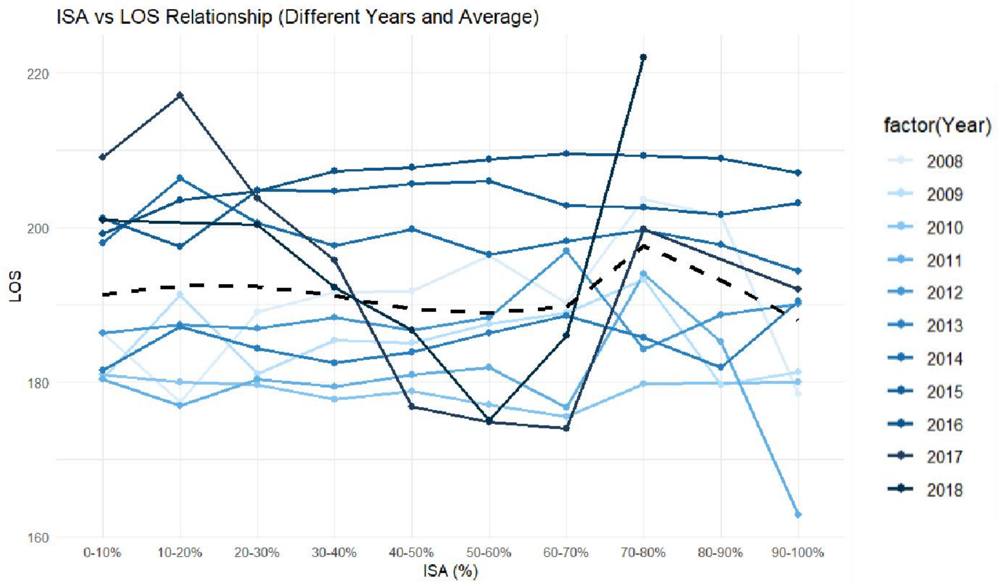

Figure 7. Annual average of LOS in each year for different ISA level

Previous analysis have looked at both urban and non-urban areas, with urban areas showing larger LOS increases. There are also different levels of urbanization in urban areas. In the process of resampling, the impervious area ratio of each pixel was calculated, and the ratio can represent different urbanization level. In 2018, the LOS data can't cover all the ISA data bigger than 80%, so there was data missing in the figure. On the whole, the LOS shows an increasing trend from 2008 to 2018. But it didn't show the same trend as ISA increased. For the area that ISA is smaller than 70%, the LOS remained unchanged. And there is a peak in the area that ISA between 60% and 70%, and then decrease. So the relationship between the ISA or the urbanization level and LOS is not completely positive. In summary, bigger ISA or higher urbanization level don't mean longer LOS.

3.3 Distribution of LOS on Various LCZs

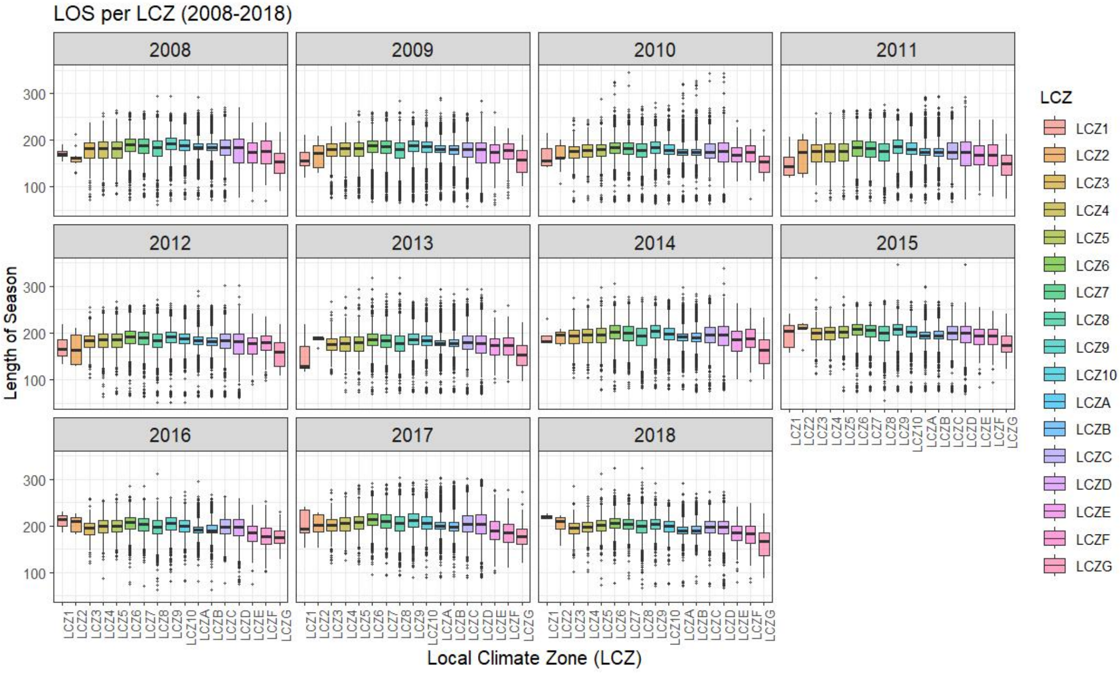

Figure 8. LOS in each year for different LCZ

In order to better understand the relationship between the urbanization level and LOS, I used the local climate zone to analyse the LOS. LCZ1, LCZ2 and LCZ3 represented compact urban areas, among these three LCZs, the LCZ3 usually have longest LOS in many years, and the LCZ3 have the lowest building height. The same trend is also shown among the LCZ4, lcz5 and LCZ6 which represent open urban area, the LCZ6 usually have the longest LOS with the lowest building height. Also, the open urban areas have longer LOS than compact urban areas on the whole. LCZ7 and LCZ8 both represent the low-rise urban, but LCZ7 is more compact compared with LCZ8, and LCZ7 usually have shorter LOS. Based on what we have summarized, low and open urban areas are more likely to have longer LOS.

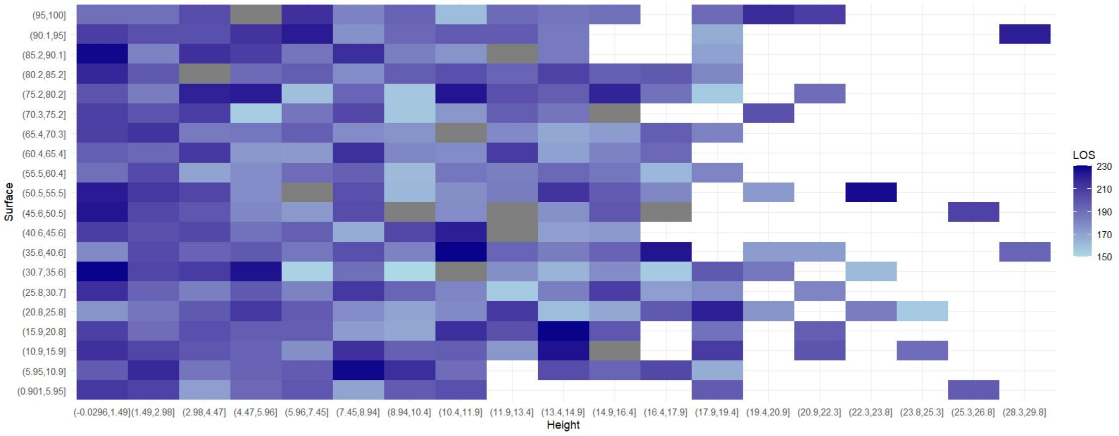

Figure 9. LOS in different building height and density in 2018 in LCZ1-LCZ9

In order to deeply explore the influence mechanism of building height and density on LOS, LCZ1-LCZ9 was selected for research. LCZ partition is the most accurate for the training year 2018, so the data of this year was analyzed in this part. Both the height and the compactness(surface) of building have effect on the LOS, the urban areas with lower building have longer LOS, the areas with less build-up surface have longer LOS, but that’s not significant in the areas with low building height. So there is a combination influence for the LOS. In summary, in the urban area, both the height of the building and the build-up surface can influence LOS, usually, the low and open urban have longer LOS, rather than the regions have highest urbanization level.

4. Discussion

This research aimed to explore the effects of urbanization on vegetation phenology, especially LOS. This research found that compared with non-urban areas, LOS is longer in urban areas, and the growth trend of LOS is faster in the study scope (2008-2018), indicating that urbanization will promote LOS prolongation. However, this does not mean that higher levels of urbanization will always lead to longer LOS. If the urban area is further detailed, the ratio of impervious area and different LCZ zones are observed, it is found that the urbanization level is not completely proportional to LOS, but contains a complex influence mechanism. Built-up areas in LCZ are mainly classified according to the height and density of buildings. Therefore, when further exploring the influence of these two factors on LOS, it is found that they synergistically affect LOS, and there will be longer LOS in low-height and more open areas, rather than the areas with the highest urbanization as we usually assume. It is speculated that this may be due to the fact that in areas with high density and tall buildings, the buildings will block the sunlight, affect the growth of vegetation, or there will be stronger winds between the buildings.

Although the urban area of Beijing is large, MCD12Q2 limits the data accuracy of the study to 500m, and the GAIA data is only updated to 2018, so the time range cannot be extended. Subsequent studies can extract EVI data for self-calculation, so as to obtain higher spatial resolution data for research.

5. Evaluation of geospatial tools

Google Earth Engine (GEE) was primarily used to acquire and initially process remote sensing data sets, including MODIS phenological data (MCD12Q2), NASADEM elevation data, and GAIA impervious surface data. In this research, MODIS phenological data (MCD12Q2) and NASADEM elevation data can be downloaded directly through GEE platform, while GAIA data can be resampled on GEE platform to calculate the impervious area ratio. And because this study only focuses on the scope of Beijing, GEE platform can choose the scope of data download instead of the whole remote sensing image, avoiding the problem of long download time due to large data. In addition, GEE platform has strong computing power, which can efficiently process and quickly download time series data, greatly improve the efficiency of data extraction, and is suitable for long-term time series research.

ArcGIS Pro is the primary platform for spatial visualization and cartographic design. In this research, it is mainly applied to the production of maps to show the distribution of LCZ, DEM, SOS, EOS and LOS in Beijing, as well as the changes of LOS. At the same time, due to the simple and convenient operation, the data in this study was clipped to the administrative scope of Beijing, and the LOS difference calculation from 2008 to 2018 was also carried out on this platform. Compared with the same class of GIS software, ArcGIS Pro is, in my opinion, more beautiful and concise in map design, and the tool operation is more convenient. At the same time, ArcGIS Pro can also carry out data processing, such as the GAIA data resampling carried out in GEE platform, but it is not as convenient as processing in GEE platform. However, the use of ArcGIS Pro software requires permission or purchase, which may limit some users.

For statistical analysis and visualization, the R environment for statistical computing was used. In this research, the comparison of phenological parameters in different spatial units, such as LCZ class and impervious area ratio, and the annual trend analysis were carried out using R. R's flexibility and its wide range of spatial and statistical software packages (e.g., raster, sf, ggplot2) make it particularly suitable for repeatability analysis. Compared to other tools that can analyze data, R produces more beautiful graphics, especially with the ggplot2 graphics package. But computational efficiency when working with large raster files requires careful data management and sometimes pre-processing on other platforms.

6. Recommendation

To better assess the impact of urbanization on vegetation phenology, future studies should strive to eliminate confounding non-urbanization factors, especially topographic differences that can influence temperature and subsequently affect plant phenology. We recommend calculating changes rather than absolute values, in order to control for terrain-induced bias and isolate the effects of urbanization.

There is currently a common misconception that urban heat island (UHI) effects promote plant growth and extend the growing season(Weng, 2009). However, our study reveals that vegetation in high-density, high-rise urban areas often performs poorly, likely due to factors such as reduced sunlight exposure, limited growing space, and degraded soil conditions. To address this issue, the government should pay greater attention to the actual condition of urban vegetation, particularly in densely built-up areas. Urban planning should prioritize the preservation and creation of low-rise, open spaces that are more conducive to healthy plant growth.

References

Demuzere, M., Kittner, J., Martilli, A., Mills, G., Moede, C., Stewart, I. D., et al. (2022). A global map of local climate zones to support earth system modelling and urban-scale environmental science. Earth System Science Data, 14(8), 3835–3873. https://doi.org/10.5194/essd-14-3835-2022

Florczyk, A. J., Corbane, C., Ehrlich, D., Freire, S., Kemper, T., Maffenini, L., et al. (2019). GHSL data package 2019: public release GHS P2019. Luxembourg: Publications Office of the European Union. https://doi.org/10.2760/290498

Hu, M., Li, X., Xu, Y., Huang, Z., Chen, C., Chen, J., & Du, H. (2024). Remote sensing monitoring of the spatiotemporal dynamics of urban forest phenology and its response to climate and urbanization. Urban Climate, 53, 101810. https://doi.org/10.1016/j.uclim.2024.101810

Li, S., Li, Q., Zhang, J., Zhang, S., Wang, X., Yang, S., & Zhang, S. (2023). Study on the Spatial and Temporal Distribution of Urban Vegetation Phenology by Local Climate Zone and Urban–Rural Gradient Approach. Remote Sensing, 15(16), 3957. https://doi.org/10.3390/rs15163957

Li, X., Gong, P., Zhou, Y., Wang, J., Bai, Y., Chen, B., et al. (2020). Mapping global urban boundaries from the global artificial impervious area (GAIA) data. Environmental Research Letters, 15(9), 094044. https://doi.org/10.1088/1748-9326/ab9be3

R Core Team (2023), R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria. Available at: https://www.R-project.org/

Ren, Q., He, C., Huang, Q., & Zhou, Y. (2018). Urbanization Impacts on Vegetation Phenology in China. Remote Sensing, 10(12), 1905. https://doi.org/10.3390/rs10121905

Stewart, I. D., & Oke, T. R. (2012). Local Climate Zones for Urban Temperature Studies. Bulletin of the American Meteorological Society, 93(12), 1879–1900. https://doi.org/10.1175/BAMS-D-11-00019.1

Weng, Q. (2009). Thermal infrared remote sensing for urban climate and environmental studies: Methods, applications, and trends. ISPRS Journal of Photogrammetry and Remote Sensing, 64(4), 335–344. https://doi.org/10.1016/j.isprsjprs.2009.03.007

Zhang, Y., Yin, P., Li, X., Niu, Q., Wang, Y., Cao, W., et al. (2022). The divergent response of vegetation phenology to urbanization: A case study of Beijing city, China. Science of The Total Environment, 803, 150079. https://doi.org/10.1016/j.scitotenv.2021.150079

Zhao, C., Weng, Q., Wang, Y., Hu, Z., & Wu, C. (2022). Use of local climate zones to assess the spatiotemporal variations of urban vegetation phenology in Austin, Texas, USA. GIScience & Remote Sensing, 59(1), 393–409. https://doi.org/10.1080/15481603.2022.2033485

Zhou, L., Ma, L., Johnson, B. A., Yan, Z., Li, F., & Li, M. (2022). Patch-Based Local Climate Zones Mapping and Population Distribution Pattern in Provincial Capital Cities of China. ISPRS International Journal of Geo-Information, 11(8), 420. https://doi.org/10.3390/ijgi11080420

Zhu, E., Fang, D., Chen, L., Qu, Y., & Liu, T. (2024). The Impact of Urbanization on Spatial–Temporal Variation in Vegetation Phenology: A Case Study of the Yangtze River Delta, China. Remote Sensing, 16(5), 914. https://doi.org/10.3390/rs16050914