Using remote sensing data in population density estimation

Evgeniy G. Shvetsov^{1,2}, Nadezhda M. Tchebakova³ and Elena I. Parfenova⁴

¹V.N. Sukachev Institute of Forest, Federal Research Center "Krasnoyarsk Science Center SB RAS", Krasnoyarsk, Russia ²Siberian Federal University, Krasnoyarsk, Russia

Abstract

In recent decades, remote sensing methods have often been used to estimate population density, especially using data on nighttime illumination. Information about the spatial distribution of the population is important for understanding the dynamics of cities and analyzing various socio-economic, environmental and political factors. In this work, we have formed layers of the nighttime light index, surface temperature and vegetation index according to the SNPP/VIIRS satellite system for the territory of the central and southern regions of the Krasnoyarsk krai. Using these data, we have calculated VTLPI (vegetation temperature light population index) for the year 2013. The obtained values of the VTLPI calculated for a number of settlements of the Krasnovarsk krai were compared with the results of the population census conducted in 2010. In total, we used census data for 40 settlements. Analysis of the data showed that the relationship between the value of the VTLPI index and the population density in the Krasnoyarsk krai can be adequately fitted ($R^2 = 0.65$) using a linear function. In this case, the value of the root-meansquare error was 345, and the relative error was 0.09. Using the obtained model equation and the spatial distribution of the VTLPI index using GIS tools, the distribution of the population over the study area was estimated with a spatial resolution of 500 meters. According to the obtained model and the VTLPI index, the average urban population density in the study area exceeded 500 people/km². Comparison of the obtained data on the total population in the study area showed that the estimate based on the VTLPI index is about 21% higher than the actual census data.

Keywords

Population density, VIIRS, nighttime lights, VTLPI, Krasnoyarsk krai.

1. Introduction

Detailed information on the spatial distribution of the population is important for understanding the dynamics of cities and analyzing various socio-economic, environmental and political factors [1, 2]. The most readily available source of population data is usually census data. This method, however, is expensive and time consuming, and therefore, population censuses are rarely conducted [3].

In recent decades, remote sensing methods have often been used to estimate population density, especially using nighttime light (NTL) data. In addition, in recent years, to improve the accuracy of the population density estimates, nighttime lights data have been used in conjunction with other data sources (vegetation indices, data on types of vegetation cover) [4]. We have used several sets of remote sensing data to assess the spatial distribution of the population

SDM-2021: All-Russian conference, August 24-27, 2021, Novosibirsk, Russia

eugeneshvetsov11@yandex.ru (E. G. Shvetsov)

^{© 0 2021} Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings (CEUR-WS.org)

density in the southern and central regions of the Krasnoyarsk krai. The day night band (DNB) of the VIIRS radiometer located on the Suomi NPP satellite was used as a source of nighttime lights data. In addition, data on the Earth's surface temperature and the NDVI vegetation index were used to improve the accuracy of the population density estimates. These datasets were also obtained using measurements made by the SNPP/VIIRS radiometer. The VTLPI (vegetation temperature light population index) index [5] was used as a composite index to estimate the population density.

The main purpose of this work was the estimation of the population density in the central and southern regions of the Krasnoyarsk krai using remote sensing data. This included the following objectives: (1) creation of the composite index, including data on nighttime lights in the study area, surface temperature and vegetation index; (2) estimation of population density using a model derived from the 2010 census.

2. Calculation of the Vegetation Temperature Light Population Index from satellite data

We have generated the GIS layers containing nighttime lights; surface temperature and vegetation index using data from the SNPP/VIIRS satellite system.

The study area included the central and southern regions of the Krasnoyarsk krai, covering an area from 52° N to 58° N and from 88° E to 99° E. (Figure 1). Its total area was about 250 thousand km², including three forest zones: taiga, forest-steppe, and the South Siberian mountain zone [6]. According to the vegetation map [7], the area of forested lands is about 50%.

We have used satellite data obtained using the VIIRS radiometer of the SNPP satellite. As a source of nighttime lights data, we have used a product generated by the National Geophysical Data Center (NOAA, USA) and available at (https://www.ngdc.noaa.gov/eog/viirs/download_ dnb_composites.html). This product contains information on the light intensity from areas of anthropogenic activity, as well as certain other objects or phenomena with a spatial resolution of 15 arc seconds. We used data on nighttime lights for the territory of the Krasnoyarsk krai for 2013, which were averaged over 12 months, so each image pixel represented the mean annual illumination value. After the averaging procedure, the final layer was re-projected into the Albers Equal-Area Conical Projection with a spatial resolution of 500 meters.

To calculate the vegetation indices, we used the VNP13A1 product, containing vegetation indices, such as NDVI, EVI, as well as reflectance values in the visible and near infrared wave ranges. During data preprocessing, we also have excluded areas containing clouds or cloud shadows. Further, a composite image was created using the maximum NDVI values obtained between the beginning of June and the end of August of the year 2013.

Additionally, to calculate the composite VTLPI index, we also used data on the earth's surface temperature. The VNP21A1 product created using the VIIRS radiometer data was used as a source of surface temperature data. Using product quality flags we have excluded pixels containing clouds, cloud shadows, or low quality data. The final temperature layer was calculated as the mean annual temperature. Further, the GIS layers containing the NDVI values and surface temperature were also re-projected into the Albers equal-area conical projection with a spatial resolution of 500 meters using the nearest neighbor method.



Figure 1: GIS layer containing VTLPI spatial distribution for the territory of the central and southern regions of the Krasnoyarsk krai, obtained using VIIRS radiometer data.

To exclude areas occupied by water bodies, a map of land surface types was used [7]. To assess the population density, we used the VTLPI index [5], calculated as follows:

$$VTLPI = rac{\sqrt{DNB_{norm}}}{NDVI_{max}} \cdot LST_{norm},$$

where DNB_{norm} is the normalized DNB value, NDVI_{max} is the maximum NDVI value, and LST_{norm} is the normalized surface temperature value. The normalized DNB_{norm} and LST_{norm} values were calculated as follows:

$$LST_{norm} = \frac{LST - LST_{min}}{LST_{max} - LST_{min}}$$

Here LST_{norm} is the normalized surface temperature, ranging from 0 to 1. LST_{max} and LST_{min} are the maximum and minimum temperature values within the study area, respectively, and LST is the temperature value of the analyzed pixel.

$$\sqrt{\text{DNB}_{\text{norm}}} = \frac{\sqrt{\text{DNB}} - \sqrt{\text{DNB}_{\min}}}{\sqrt{\text{DNB}_{\max}} - \sqrt{\text{DNB}_{\min}}}.$$

Here DNB_{norm} is the normalized value of the nighttime light for the area, ranging from 0 to 1. DNB_{max} and DNB_{min} are the maximum and minimum nighttime light values within the study area, respectively, and DNB is the value of the analyzed pixel.

Unlike [5], we did not correct the VTLPI value for the height of the terrain above sea level, since there were no mountain regions within our study area.

2.1. Modeling population density using remote sensing data

To use the VTLPI to assess the population density of the Krasnoyarsk krai, a model relationship relating these two data sets was obtained. The materials of the year 2010 census were used as a source of data on the actual population density. Census materials are publicly available on the Rosstat website (https://rosstat.gov.ru/free_doc/new_site/perepis2010/croc/perepis_itogi1612. htm).

We compared the averaged VTLPI values calculated for a number of settlements within the Krasnoyarsk krai with the population values obtained from the census data. To calculate the mean VTLPI value, we used geoinformation layer of the settlements boundaries. Using these boundaries, we have extracted fragments from the VTLPI raster layer corresponding to settlements for which several statistical parameters were calculated. In total, we used data for 40 settlements in the central and southern regions of the Krasnoyarsk krai.

According to the census data, the number of people in the analyzed settlements ranged from 756 to 973826. At the same time, the mean VTLPI value for the considered settlements ranged from 0.7 to 17. Figure 2 shows a diagram of the VTLPI values and population density in the analyzed settlements according to the census data. Analysis of the data showed that the relationship between these data series can be fitted using a linear function, such as P = 136.5 * VTLPI - 105, which reasonable well describes the dependence between the population density and the value of the composite VTLPI index ($R^2 = 0.65$). Here P is the population density (people/km²). In this case, the value of the root-mean-square error was 345, and the median relative error was 0.09. The relative error was calculated as the difference between the census data and the model data referred to the census data.



Figure 2: The relationship between the VTLPI index and population density according to the 2010 census. Each point corresponds to one settlement. The solid line represents the linear fit.



Figure 3: Spatial distribution of population density in the study area.

Using the obtained model equation and the spatial distribution of the VTLPI index, we have estimated the distribution of the population over the territory of the study area with a spatial resolution of 500 m (Figure 3). According to the resulting model and the VTLPI index, the highest population density was observed in cities, which corresponds to the actual census data. The average density of the urban population in the considered territory exceeded 500 people/km².

Comparing our results with the census data for the study area we found that the VTLPI-based estimate is about 21% higher than the actual census data. Thus, according to the model estimate, the total population was 2.9 million, while according to the census data, this was approximately 2.4 million.

There are several factors that can influence the accuracy of estimates. First, the VTLPI index does not allow excluding the influence of industrial objects, which are characterized by rather high values of the nighttime light measurements, temperature, as well as low values of the NDVI index. For example, on the territory of the Achinsk refinery, the population density, according to the model estimate, reached 4–5 thousand people/km², which is obviously associated with significant emissions of thermal and light radiation. Also, errors in determining the boundaries of settlements used to create a sample in the process of the model parameterizing can have an impact on the accuracy of estimates. In addition, the time difference between census data (2010) and the time of satellite data acquisition (2013) is likely to have influenced the results.

3. Conclusions

Using the archive of thematic products, generated using the data of the VIIRS radiometer, we have created the GIS layers of the nighttime lights, the vegetation index NDVI and the land surface temperature for the central and southern regions of the Krasnoyarsk krai for the year 2013. Using these layers, we have calculated the VTLPI index related to population density.

According to the year 2010 census data obtained for 40 settlements of the Krasnoyarsk krai, a training sample was created, which was used to parameterize a linear model relating the VTLPI index and the actual population density.

Using the obtained model equation and the spatial distribution of the VTLPI index, the population density for the study area was estimated with a spatial resolution of 500 m. At the same time, comparison with the census data showed that the estimate based on the VTLPI index is about 21% higher than the actual data.

Acknowledgments

This work was financially supported by the Russian Foundation for Basic Research, the Krasnoyarsk Regional Fund for the Support of Scientific and Scientific and Technical Activities, project No. 19-45-240004-r_a "Predictions of the ecological-economic potential for possible "climatic" migrations in the Angara-Yenisei macroregion in a changing climate of the 21st century".

References

- Ahola T., Virrantaus K., Krisp J.M., Hunter G.J. A spatio-temporal population model to support risk assessment and damage analysis for decision-making // International Journal of Geographical Information Science. 2007. Vol. 21. P. 935–953.
- [2] Aubrecht C., Özceylan D., Steinnocher K., Freire S. Multi-level geospatial modeling of human exposure patterns and vulnerability indicators // Natural Hazards. 2013. Vol. 68. P. 147–163.
- [3] Yang X., Huang Y., Dong P., Jiang D., Liu H. An updating system for the gridded population database of China based on remote sensing, GIS and spatial database technologies // Sensors. 2009. Vol. 9. P. 1128–1140.
- [4] Roy Chowdhury P.K., Maithani S., Dadhwal V.K. Estimation of urban population in Indo-Gangetic Plains using night-time OLS data // International Journal of Remote Sensing. 2012. Vol. 33. Is. 8. P. 2498–2515.
- [5] Luo P., Zhang X., Cheng J., Sun Q. Modeling population density using a new index derived from multi-sensor image data // Remote sensing. 2019. Vol. 11. P. 2620.
- [6] Decree of the Ministry of Natural Resources of Russia "On approval of the list of forest zones of the Russian Federation and the list of forest regions of the Russian Federation" No. 367 dated 18.08.2014 (as amended on 19.02.2019). Moscow, 2014. (In Russ.)
- [7] Bartalev S.A., Belward A., Erchov D.V., Isaev A.S. A new SPOT4-VEGETATION derived land cover map of Northern Eurasia // International Journal of Remote Sensing. 2003. Vol. 24. Is. 9. P. 1977–1982.