Over the last five years several European countries have experienced a rapid development of photovoltaic power plants. In the Czech Republic more than 20,000 new photovoltaic power plants have been built with the total nominal power capacity of 2100 MWp. The spatial distribution of these installations is very uneven. Built-up areas are dominated by a large number of small installations while large installations (over 1 MWp) are usually located in agricultural areas. To express a spatial distribution of these installations we have used a synthetic approach combining the spatial density of installations with the power plant capacities. The resulting map in the scale of 1:750,000 shows a continuous distribution of the installed power capacity.

Keywords: photovoltaic power plants; kernel density; GIS; Czech Republic

1. Introduction

Over the last five years many countries in the world experienced a rapid development of photovoltaic applications (EPIA, 2013). This rapid development is associated with a governmental support for green energy and decreasing costs of photovoltaic power plants. For example, between the years of 2009–2011 more than 20,000 new photovoltaic (PV) power plants with the total nominal capacity of 2100 MWp were built in the Czech Republic. Understanding the location and spatial distribution of these installations is important given their recent growth and potential implications associated with power grid management and electricity production/consumption in various parts of the country. Despite the requirement of each PV power plant to be approved by the Czech Energy Regulatory Office ERÚ, a review of the published research showed that no maps exist either for the Czech Republic or any other countries in regards to their distribution and spatial location. The ERÚ list of these power plants contains various data including a nominal power capacity and physical address of the power plant. Initial results indicate that the spatial distribution of these power plants is very uneven. Built-up areas are dominated by a large number of small power plants while large power plants (over 1 MWp) are usually located in agricultural areas. This unevenness creates the map of PV power plants showing the creation and allocation nominal power capacity might be difficult, especially at small scales. The goal of this paper is to present a synthetic map at a scale of 1:750,000 showing a continuous density of PV power plant locations coupled with the installed nominal power capacities. The
Main Map theme is guided with supplementary maps of a smaller scale (1:2,500,000) to assist better understanding of the mapped phenomenon and linkage with other related themes. The map was conceived as being a printed wall map and it was prepared for use by both lay people and experts.

2. Methods and data

The focal map to which this article relates records the spatial distribution of PV power plants and their capacity to produce electricity. The method used for its portrayal is based on application of point density estimation with a spatial kernel weighted by the installed power capacity. While the main map frame shows the continuous distribution of the installed PV capacity, supplementary maps aid to compare this theme with: (i) solar resource map represented by global solar irradiation at optimal PV panel inclination, (ii) point locations of the PV installations displayed as graduated symbols, and (iii) a detail view of a smaller area demonstrating the appropriateness of the continuous representation of a point field.

2.1. Preparation of the topographic background

The scale of the map in the main frame is 1:750,000, and the supplementary maps placed along the sides of the main map frame are at a scale of 1:2,500,000. The topographic background of the main map is based on the ArcČR 500 geodatabase freely available from www.arcdata.cz. The ArcČR 500 is a digital vector geographic database of the Czech Republic prepared at the scale of 1:500,000. The source data used for topographic background required a modification for their larger scale in comparison with the target scales. In particular, the linear features and outline of the polygon features was modified using the Bend Simplify algorithm implemented in ArcGIS 10.1 with a simplification tolerance of 1.9 km and 6.25 km, respectively. The goal was to retain as much of the shape of the feature as possible, eliminating the maximum number of coordinates. Polygon features which were too small for adequate visualisation (below 15 ha and 150 ha, respectively) were eliminated. A similar approach was applied by Zuniga, Pueyo, and Calvo (2012).

2.2. Main map frame

The main map frame at a scale of 1:750,000 displays a continuous representation of PV power plant capacities derived from a spatial density of the PV power plant locations and installed power capacities in the year of 2013. The original data were provided as a MS Excel data sheet by the Czech Energy Regulatory Office ERÚ. The original database contained two types of attributes about licence holders and installations.

The raw data originally comprised a list of registered legal subjects who were given a permission to install PV panels either on roofs or on the ground and connect to the electricity grid. Frequently, several subjects were registered for the same address which means they joined together to install a large group of PV panels which is in fact a single location in real landscape. Therefore, such subjects were filtered and their claimed installed capacities were merged together to get the sum of the installed power at a unique address.

Further steps involved spatial referencing of the locations of the PV power plants within the WGS 1984 coordinate system by the means of the processed tabular data and address information. The online utility BatchGeo (http://batchgeo.com/) was used to find the location of each power plant, i.e. to geolocate each address. The resulting accuracy of geocoding was very high in built-up areas for which street addresses are geolocated in the Google Developers
(2013) (12,941 addresses), whereas only centroid of the respective village was used to locate the power plants in rural areas (4839 addresses). In the latter case, the value of 0.1 MWp was considered as a general limit for which the area coverage of PV panels is large and therefore it is very unlikely the panels could be installed on a roof of building. They should be mainly located in other land cover category than the built-up area thus further away from the centroid. Therefore, it is sensible to verify the exact location of the power plants of 0.1 MWp capacity and higher as their location and magnitude of the capacity could affect the resulting spatial distribution of the resulting density map. Seven hundred fifty-six of such power plants were identified, manually verified and their location was fixed using orthophoto-maps available at www.mapy.cz and using the WMS service of the Czech Office for Surveying, Mapping and Cadastre (2013). We assume the inaccuracy of geocoding the remaining 4083 power plants has a very little influence on the resulting density map image, as these power plants are small and very likely located on the roofs within the built up part of the village, i.e. very close to the centroids.

The resulting GIS data set contained 17,754 unique point locations which were further processed in the ArcGIS 10.1 software. The range of the installed capacities per unique location was between 1 kWp up to 56.961 MWp while the mean value and the median were 9 kWp and 118.3 kWp, respectively. The total sum of the installed PV capacity in the Czech Republic resulted in 2101 MWp. However, the spatial distribution of the PV power plants was very uneven (Figure 1(a)), with more than 12,000 locations clustered in built-up areas, yet, with the total nominal capacity of 363.6 MWp only. Various methods exist for an appropriate

Figure 1. Spatial distribution of PV power plants and their installed capacity in the Czech Republic in 2012 represented as (a) point locations only, (b) proportional symbols sized according to the installed PV power capacity, (c) point density per kilometre squared in a regular grid of 100 m cell size, (d) density of the installed PV power capacity per kilometre squared in a regular grid of 100 m cell size.
cartographic portrayal of spatial distribution of a phenomenon at point locations (Robinson, Morrison, Muehrcke, Kimerling, & Guptill, 2008). Traditional approach to display the magnitude of the phenomenon (e.g. the installed PV power capacity) at a point location uses a method of proportional symbols or graduated symbols (cartodiagram). This is not applicable for a large data set of clustered point data as can be seen in Figure 1(b). Point densities calculated within a regular grid could be applied to display the point distribution, but the magnitude of the mapped phenomenon is not considered in this case (Figure 1(c)). Therefore, we used a synthetic approach to express the spatial distribution of PV power plants together with their installed power capacities per unique location within the territory of the Czech Republic (Figure 1(d)).

The method applied is based on point density estimation with a spatial kernel (Silverman, 1986) which is weighted by the magnitude of the installed PV power capacity. The method is implemented in ArcGIS 10.1 as the Kernel density tool within the Density tool group of Spatial Analyst Tools (Figure 2). The density was calculated for cell centres of a 100 metre regular grid which is appropriate given the scale of the map in the main frame. The cell size mainly impacts on the visual appeal of the kernel density mapping output, with higher resolutions producing maps that avoid the blocky pixelation of outputs generated using larger cell sizes (Chainey, 2013). The kernel density estimator applied in creation of the density raster surface

![Figure 2. Parameter settings of the kernel density estimation in ArcGIS 10.1.](image-url)
is analytically described in Silverman (1986, p. 76, Equation 4.5). In the case of density weighted by an observed value, it can be defined as a bivariate function:

$$f(v) = \frac{1}{n \cdot h^2} \sum_{i=1}^{n} \left\{ K \left[ \frac{1}{h} (v - V_i) \right] \cdot o_i \right\}$$

(1)

Where $h$ is the kernel bandwidth (“Search radius” in ArcGIS 10.1) and $n$ is the number of observed two dimensional locations $V_i$ with coordinates $[x_i, y_i]$ of the $i$-th PV power plant $V_1, \ldots, V_n$. The observed value at each location $V_i$, i.e. the installed PV power plant capacity, is denoted by the vector of $o_i$. The two dimensional kernel function $K(v)$ which is fitted over each point $V_i$ satisfies the condition:

$$\int_{R^2} K(v) dv = 1$$

(2)

Specifically for the production for the main map, $K(v)$ was a quartic kernel weighting function defined as:

$$K_2(v) \begin{cases} 
3 \pi^{-1} \{1 - [(v - V_i)h^{-1}]^2\}^2 & \text{if } (v - V_i)h^{-1} < 1 \\
0 & \text{otherwise}
\end{cases}$$

(3)

With this kernel, if the distance of the $i$-th PV power plant $V$ from the centre of the considered cell is larger than the bandwidth $h$, then the $i$-th PV power plant $V$ does not contribute to the summation. The value of $o_i$ (“Population field” value in ArcGIS 10.1) determines the number of times to count the point $V_i$ in the summation of kernel weights $K(v)$ in Equation (1). The output value of $f(v)$ represents the estimated density of the installed PV capacity in megawattpeak (MWp) per kilometre squared at the location $v[x, y]$ denoting each cell centre of the resulting raster.

The kernel bandwidth $h$ (“Search radius” in ArcGIS 10.1) was initially calculated as 9712 metres by the rule-of-thumb as described in ESRI (2012). A more robust approach was also used which was based on a cross-validation procedure implemented in the spatstat package (Baddeley & Turner, 2005) as the bw.smothppp function. This bandwidth optimisation procedure is designed for weighted kernel density estimation (Nadaraya, 1964; Watson, 1964) which is the case of the analysed PV power plant capacity. The resulting optimal bandwidth was 7627 metres. The choice of bandwidth influences the interpretation of the density and it should be selected also with respect to the target scale of the analysis (Chainey, 2013). The shorter calculated bandwidth produced a more marked circular (‘bull’s eye’) pattern therefore we finally opted for the larger bandwidth which was rounded up to 10,000 metres. This value was considered a more appropriate value for the given scale of map display and spatial distribution of the installed PV capacity to limit (i) the presence of the ‘bull’s eye’ pattern and (ii) to limit smoothing of the resulting density pattern.

The density values in the resulting raster map displayed in the main map frame was classified into nine intervals using the quantile method. The calculated quantile values were rounded to two decimal places.
2.3. Supplementary maps

The main mapped phenomenon is better appreciated with additional maps placed along the main map frame. Two supplementary maps show the territory of the Czech Republic at smaller scale of 1:2,500,000 and display the following themes:

1) Global solar irradiation. The theme is displayed as a regular grid of 1 arcmin cell size showing average values of yearly sum of global solar irradiation striking a planar surface unit inclined at 25 degrees. This inclination angle is calculated as an optimal angle for installation of PV panels in for the territory of the Czech Republic by SolarGIS database v1.9 (Šúr & Cebecauer, 2010) for the time period (1999–2012). It is a trade-off between the optimal and acceptable loss caused by shadow casting of PV panels ordered in rows one behind each other. The raster values are in kilowatthours per metre squared per year. The data layer was supplied by the GeoModel Solar company free of charge.

2) Photovoltaic power plant capacities as graduated symbols. The theme is displayed as black circles sized in three levels. The intervals were selected to differentiate small PV power plants installed mainly on roofs (below 0.3 MWp; 16,862 locations), medium installations mounted on large roofs or on the ground (0.3–1 MWp; 407 locations), and large PV plants (over 1 MWp; 485 locations) installed mainly on the ground which induced change of land cover.

The main map is accompanied with an inset showing the photovoltaic power plant capacities represented by a continuous surface from the main map as well as graduated symbols. This map is a detail of a smaller area near the city of Brno within which a single theme is displayed with two cartographic methods. The aim of placing the inset is to demonstrate the advantages of a continuous representation over the traditional graduated symbols method in case where the point data are densely distributed in space. Despite the larger scale of the inset, some symbols overlap each other and therefore the continuous representation of the PV installed capacity is more appropriate for portrayal of the main mapped theme.

3. Conclusions

The presented map of PV power plants in the Czech Republic at the scale of 1:750,000 shows an uneven distribution of the installed power capacities with dominance of several regions (South Moravia, Western Northern and Southern Bohemia). The applied methodology enables readers to easily understand the information on the size and density of the mapped theme which is difficult to display in the medium scale with traditional cartographic methods such as proportional symbols. The maps of smaller scale positioned along the sides of the main map frame provide additional information for better understanding of the spatial distribution of the PV power plants and its relationship with potential global irradiation per year. The continuous representation of the mapped phenomenon provides new possibilities for further analysis and understanding of the phenomenon. For example, the map can be directly compared to areas with the highest electricity consumption or to power grid lines and transmission capacities helping to optimise the usage of the national power grid.

Acknowledgements

This work was supported by the Ministry of Education of the Slovak Republic under Grant APVV 0176-12, VEGA 1/0272/12 and KEГA 002PU-4/2012.
Software
The raw data on PV installations were processed using Microsoft Excel 2010, with geolocation performed using BatchGeo (http://batchgeo.com/). Esri ArcGIS 10.1 was used to generate the thematic layers used for map production. R statistical software (R Core Team, 2012) was used for the selection of the kernel bandwidth. Corel Draw X5 graphic suite was used for final map production.

References
Baddeley, A., & Turner, R. (2005). Spatstat: An R package for analyzing spatial point patterns. Journal of Statistical Software, 12(6), 1–42. URL http://www.jstatsoft.org/v12/i06/

Chainey, S. P. (2013). Examining the influence of cell size and bandwidth size on kernel density estimation crime hotspot maps for predicting spatial patterns of crime. Bulletin of the Geographical Society of Liege, 60, 7–19.

Czech Office for Surveying, Mapping and Cadastre. (2013). Prohlížecí služba WMS – Ortofoto. Český úřad zeměměřický a katastrální [Web viewing service WMS – Orthophoto, format description]. Retrieved from: http://geoportal.cuzk.cz/WMS_ORTOFOTO_PUB/WMService.aspx

EPIA. (2013). Global market outlook for photovoltaics 2013–2017. Brussels: European Photovoltaics Industry Association.

ESRI. (2012). Kernel density (Spatial Analyst). ArcGIS resources help. Desktop: Release 10.1 Redlands, CA: Environmental Systems Research Institute. Retrieved from: http://resources.arcgis.com/en/help/main/10.1/index.html#//009z0000000s000000

Google Developers. (2013). Google maps geocoding API. Retrieved from: https://developers.google.com/maps/documentation/geocoding/

Nadaraya, E. A. (1964). On estimating regression. Theory of Probability and its Applications, 9, 141–142.

R Core Team. (2012). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from: http://www.R-project.org

Robinson, A. H., Morrison, J. L., Muehrcke, P. C., Kimerling, A. J., & Guptill, S. C. (2008). Elements of cartography (6th ed.). Chichester: Wiley.

Silverman, B. W. (1986). Density estimation for statistics and data analysis. New York: Chapman and Hall.

Šúri, M., & Cebeceauer, T. (2010). SolarGIS: New Web-Based Service Offering Solar Radiation Data and PV Simulation Tools for Europe, North Africa and Middle East. Proceedings of the International Conference on Solar Heating, Cooling and Buildings EUROSun 2010, 28 September – 1 October 2010, Graz, Austria. Retrieved from: http://geomodelsolar.eu/_docs/papers/2010/Suri-Cebecauer_SolarGIS_EUROSUN2010.pdf

Watson, G. S. (1964). Smooth regression analysis. Sankhy A, 26, 359–372.

Zuniga, M., Pueyo, A., & Calvo, J.-L. (2012). The Spanish population during the twentieth century and beyond. Journal of Maps, 8(4), 386–391, doi: 10.1080/17445647.2012.744364