CZECH UNIVERSITY OF LIFE SCIENCES PRAGUE

Faculty of Environmental Sciences
Department of Applied Geoinformatics and Spatial Planning

Spatial Data Quality in Digital Visibility Models

DOCTORAL THESIS

2019

Author:
Ondřej Lagner

Supervisor:
doc. Ing. Petra Šímová, Ph.D.

Reviewers:
Dr. Marta Szostak
University of Agriculture in Krakow

Dr. Vincent Lecours
University of Florida

doc. Ing. Rudolf Urban, Ph.D
Czech Technical University in Prague

Declaration

I have written the thesis by myself, and have not used sources or means without their declaration in the text. Any thoughts by others or literal quotations are clearly referenced. The thesis was not published elsewhere except for individual studies that had already been published in scientific journals.

Acknowledgement

I would like to express my thanks to everyone who helped me and contributed in any way towards this thesis. Thank Tomáš Klouček, Honza Komárek, Jirka Prošek, Michal Fogl, Petr Klápště, David Moravec, Vojta Barták and Katka Gdulová who not only participated in the research studies but also created a friendly and inspiring environment in the halls. Surprisingly, I also feel that I should thank Vítek Moudrý for his incessant criticizing that I should finally start working.

Thank my "new" colleagues from the City Hall of Prague 22 – it is true that only a few of you know about me writing this thesis so far but despite that, thank you for not bombarding me with questions about where have I been when not sitting on my chair.

As an apology for my failure to participate in the track through the beautiful Kokořínsko area, I promised that I would thank Tomáš Kulík, Janča Kořínková, Míša Kafková and Vendík Měšeč. Here are my thanks – and the next time, I promise that I will not forget (or at least will find a better excuse).

I always found it amusing that in every Bachelor or Masters thesis, everyone thanks their supervisors for their valuable advice and comments. Hence, I would like to thank my supervisor, Petra Šímová, for her valuable advice and comments. But seriously – without you, Petra, this thesis would never come to exist. Or would possibly be written some time in the indefinite future. Thank you that you have pushed me in your friendly but uncompromising way through the doctoral studies including writing of this thesis. Thank you for all your support and for the new possibilities you opened for me. I owe you – and let's save the rest of my thanks for Karlov.

My great thanks and respect belong to Katka who suffered through this thesis possibly more than I did and who, despite it all, still loves me. I promise you that I will make it all up to you. Perhaps as soon as next summer. Considerable thanks go to Petra, Jirka and the entire family. Without you and your support, I would not be where I am and it is quite likely that I would not be at all.

Special thanks go to Jarek Janošek who has a lion's share on the fact that the thesis is written in English. On the other hand, my grandmas will probably not be as grateful as they will not find it an easy reading thanks to you.

And to prevent being overcome with emotion, I also need to pragmatically give my thanks to the grants support that made sure that I had no time to be bored over the last years but that also immensely facilitated the creation of this thesis.

Abstract

Digital terrain models and applied visibility analyses are widely used tools of geoinformatics. They have an irreplaceable position in landscape evaluation as well as in evaluation of individual interventions. Similarly, they can facilitate assessment of projects in the fields of landscape ecology, animal ecology, urban planning, architecture and civil engineering, urbanization, planning of tourist paths, lookout towers, etc. The present day modelling of terrain and visibility however faces a problem of existence of a large amount of input geodata of varying quality.

The thesis consists of a set of three commented studies focused on selection and processing of data suitable for visibility analyses. The studies aim to identify spatial data that would lead to the most accurate results of visibility analysis. The accuracy of spatial data is affected by multiple factors such as the method of data acquisition, spatial scale and data processing. All these variables affect, along with the terrain characteristics and the degree of forestation, the accuracy of the visibility results. The three presented studies imply that the greatest effect on the accuracy of the result of visibility analysis can be attributed to the accuracy and spatial scale of the input geodata. This influence is so great that it by far exceeds the effect of other variables such as the degree of forestation, terrain complexity, the number of obstacles to visibility, buildings, etc.

Data acquired through airborne laser scanning are the most suitable input data for visibility analysis. Use of fine scale vector data however may serve as an alternative where LiDAR data are not available. Coarsening the scale however leads to overestimating the visible areas and to the higher false positive rate, i.e., the error where the model predicts more pixels to be visible than really are, which is most likely caused by "flattening" the surface and reducing the number of obstacles when using coarser data. Another factor having a great effect on the accuracy of the digital models is the degree of pre-processing of input geodata. If using raw data that are specifically processed with the aim of the analysis in mind, the results are significantly better than when using ready-made data that are pre-processed with a general purpose algorithm as such data may suffer from processing flaws unknown to the researcher or with loss of detail information caused by an unsuitable processing method.

Abstrakt

Analýzy viditelnosti využívající digitální model terénu jsou velmi rozšířený nástroj na poli geoinformatiky. Mají nezastupitelné místo v hodnocení krajiny jako celku, i při hodnocení jednotlivých krajinných zásahů. Stejně tak jsou uplatňovány při hodnocení projektů v krajinné ekologii, ekologii živočichů, v územním plánování, v architektuře, stavebnictví, při plánování turistických cest, rozhleden apod. Současné modelování terénu a následně analýz viditelnosti čelí problému velkého množství vstupních geodat, která mají rozdílnou kvalitu.

Disertační práci tvoří komentovaný soubor tří studií, zaměřených na výběr a zpracování vhodných dat pro analýzy viditelnosti. Tématem studií je nalezení vhodných prostorových dat, na základě kterých bude provedená analýza viditelnosti nejpřesnější. Přesnost prostorových dat ovlivňuje několik faktorů, mezi které patří způsob pořízení dat, prostorové měřítko a zpracování dat. Veškeré tyto proměnné, rozšířené o charakteristiky terénu a míru zalesnění, se promítají do výsledné přesnosti analýzy viditelnosti. Ze tří předložených studií vyplývá, že nejvýznamnější vliv na výsledek analýzy má přesnost a měřítko vstupních geodat. Tento vliv je natolik značný, že upozadí vliv terénu, vliv vegetace i vliv dalších objektů, které se nacházejí na povrchu Země.

Nejvhodnější data pro analýzy viditelnosti jsou data pořízená leteckým laserovým skenováním. Jako alternativu lze použít vektorová data velkých měřítek. S klesajícím měřítkem dochází k nadhodnocování viditelných míst a k vyšší četnosti false positive, kdy model predikuje více viditelných míst, než jich je ve skutečnosti. Nezanedbatelný vliv na přesnost digitálních modelů terénu má míra zpracování geodat. Prostorová data je nezbytné vybírat s vědomím účelu, ke kterému budou sloužit. Nejvhodnější je využití co nejméně zpracovaných surových dat, u kterých má uživatel jistotu, že neobsahují zásadnější zpracovatelskou chybu. Surová data také mohou obsahovat vice informací, které však byly ztraceny během nevhodně zvolené metody zpracování.

Table of Contents

CHAPTER I
Introduction

I remember that when I was small, I liked looking at maps. I kept looking at them and imagining what must it look like for real. It might have ben one of the reasons why I pursued my Bachelor's degree in Geography and Cartography. When I progressed to the Masters studies, I was asked by Petra Šímová if I would like to work on my Master's Thesis under her supervision and was offered a topic of visibility analyses, I had no clue that this topic would stay with me for another seven years.

The principle of visibility analyses is to model the area visible from a particular point. Over time, this method has been refined almost to perfection and there are many sophisticated methods, all of which are however still based on the same principle. Visibility analyses find their place in many areas of science – from ecology and landscape ecology, through biology and archaeology up to narrow fields of information technology. To create a visibility model, however, there is a crucial prerequisite – spatial data. Its accuracy is a crucial factor for calculations of visibility analyses and to a large degree determines the reliability of the results. There are however many types of spatial data. We can classify them by scale, positional accuracy, by type of processing they were subjected to, by the type of model, etc. Such an unprecedented availability of geodata however often leads to a non-critical approach to the selection of data for a particular analysis. It is not uncommon to see data of global scale used for a local extent analysis just because they were freely available or, contrary, very expensive data were used for a study for which free data would be sufficient.

Most visibility analyses use digital elevation models or digital surface model in a raster format as input data. As mentioned above, the result of the analysis thus depends on the accuracy of the raster, i.e., the used spatial data and processing methods. There are many approaches to creating a digital elevation raster and many studies and papers describing these have been published. Studies focusing on comparison and definition of suitable geodata for application in landscape ecology, namely in bird ecology and biotope modelling, are substantially more scarce. The presented thesis discusses the effect of the input data on visibility analyses, aiming to identify spatial data suitable for particular analyses both from the perspective of processing demands and acquisition costs. The aim of the visibility analyses constituting a substantial part of this thesis is not to find a most suitable location for placement of a lookout tower or evaluate an effect of a proposed construction on the landscape. It can be actually said that visibility analyses are used here as a tool for evaluation of accuracy of digital terrain or surface models. The presented experiments do not change the way of calculating the visibility model – they compare models prepared in the same way based on different input (spatial) data.

Three papers published in scientific journals, which are closely related to the topic of visibility analyses, represent the pillars of this thesis. The first two studies directly analyse the effect of input data on resulting visibility analyses. By comparing results obtained from various spatial data, they are trying to find optimal geodata that would be best suited for the purpose and the method of the analysis. The third paper is aimed at another crucial step of visibility analyses, namely at comparing the suitability of various the suitability of various data collection methods for creating a digital relief or surface model with maximum accuracy.

CHAPTER II
Aims of the thesis

The dissertation thesis aims to present and discuss visibility analyses in particular from the perspective of the effect of input data on the results of the analyses. The opposite is also true - as the input data are typically digital terrain or surface models that are inherently built from spatial data, this also comes hand in hand with evaluation of the accuracy and quality of the input data that were used for creating the digital models. The aims of the thesis were in particular to:

A) Evaluate the suitability of various types of spatial data for analyses in landscape planning.

Visual impact of wind turbines is a topic frequently discussed in association with visibility analyses. The effect of input data on the result of the analysis is however mostly neglected. Thus, we aimed at the evaluation of the reliability of results when using digital surface models at various scales.

> (i) Identify data most suitable for creating digital terrain models
>
> Digital terrain model can be in principle created from any spatial data containing information about elevation. Not every dataset however possesses a sufficient accuracy for creating a model suitable for the analysis in question. One of the aims of this thesis is therefore to find spatial data most suitable for visibility analysis.
>
> (ii) Can vector data compete with LiDAR?
>
> The gradual increase in usage of LiDAR data for creating digital terrain models has pushed the use of vector data into the background. Does the use of vector data still make any sense, or shall it be completely replaced with LiDAR data?

B) Evaluate the effect of the terrain and forestation on the accuracy of visibility analyses

The main goal is to evaluate the effect of different spatial data entering the visibility analysis on its results. Areas of interest are chosen to capture important types of features in the landscape, i. e. with emphasis on capturing a full representation of geomorphologic character of the landscape. The parameters were, in addition, complemented by assigning various heights to the forest canopy to evaluate the effect of the height of the tree stands on visibility analyses.

C) Evaluate the effect of input data processing on the accuracy of resulting digital terrain models.

Despite the fact that geodata are the principal factor defining the accuracy of the performed analysis, little attention is paid to the quality of data and data processing methods in many studies. The presented study compared three available forms of the same LiDAR data – raw data, generalized point cloud and ready-made DTM and DSM. We used our own field measurements to complement those datasets by validation (true) data that were used for accuracy assessment. The aim of this study was to find out how the data processing methods affect the resulting terrain models.

D) How much does the spatial scale affect the accuracy of data for visibility analysis?

The degree of generalization as well as the accuracy of data collection generally affect the results of analyses. We aim to find whether it is necessary for the spatial scale of the data to correspond to the scale of the analysis.

CHAPTER III
Theoretical Background

Digital terrain models – definition and terminology

Digital terrain models as we know them today have developed since their introduction in 1958 (Miller and LaFlamme), in particular from the perspective of their accuracy and level of detail. There have also been some development in the terminology of the models depending on what the particular model depicts. Hence, I believe it useful to clarified the terminology as used in this dissertation right here, at the beginning.

Miller a LaFlame (1958) defined the digital terrain model as a statistical representation of the continuous surface of the ground by a large number of selected points with known xyz coordinates in an arbitrary coordinate field. Storing the DTM data in a way allowing it to be read by computers makes it available for computer analysis of a wide variety of terrain problems as well as for the evaluation of an unlimited number of independent solutions to each type of problem.

A general definition of data model was given almost twenty years later by Tsichritzis and Lochovsky (1977) who defined it as "a set of guidelines for the representation of the logical organization of the data in a database consisting of named logical units of data and the relationships between them". Goodchild (1992) however added that with few (if any) exceptions, the world which is represented in a spatial database is not composed of logical units, and thus must be abstracted, generalized of approximated in the process of creating a database. Podobnikar et al. (2009) define digital terrain model as a continuous surface that besides the grid with values of elevation (known as a digital elevation model — DEM), also consists of other elements that describe the topographic surface, such as slope or skeleton.

The DTM was therefore originally defined as a digital (numerical) representation of the terrain. Since Miller and Laflamme (1958) coined the original term, other alternatives have been brought into use. These include aforementioned digital elevation models (DEMs), digital height models (DHMs), digital ground models (DGMs), as well as digital terrain elevation models (DTEMs). These terms originated from various countries. The term DEM was widely used in America; DHM came from Germany; DGM was used in the United Kingdom; and DTEM was introduced and used by USGS and DMA (Defense Mapping Agency) (Petrie and Kennie, 1987). In practice, these terms (DTM, DEM, DHM, and DTEM) are often assumed to be synonymous and indeed this is often the case. Sometimes, however, they indeed refer to different products. That is, there may be slight differences between these terms. Li (2004) has made a comparative analysis of these differences as follows:

1. Ground: "the solid surface of the Earth"; "a solid base or foundation"; "a surface of the Earth"; "bottom of the sea"; etc.

2. Height: "measurement from base to top"; "elevation above the ground or recognized level, especially that of the sea"; "distance upwards"; etc.

3. Elevation: "height above a given level, especially that of sea"; "height above the horizon"; etc.

4. Terrain: "tract of land considered with regard to its natural features, etc."; "an extent of ground, region, territory"; etc.

From these definitions, some of the differences between DGM, DHM, DEM, and DTM begin to manifest themselves. So, a DGM more or less has the meaning of "a digital model of a solid surface." In contrast to the use of ground, the terms height and elevation emphasize the "measurement from a datum to the top" of an object. They do not necessarily refer to the altitude of the terrain surface, but in practice, this is the aspect that is emphasized in the use of these terms. The meaning of "terrain" is more complex and embracing. It may contain the concept of "height" (or "elevation"), but also attempts to include other geographical elements and natural features. Therefore, the term DTM tends to have a wider meaning than DHM or DEM and will attempt to incorporate specific terrain features such as rivers, ridge lines, break lines, etc. into the model (Li, 1990).

In this thesis, the digital terrain model (DMT, DTM) is defined as a raster matrix with elevations. Digital terrain model is also perceived in this thesis as a term superordinate to the digital surface model (DSM) and digital relief model/digital elevation model (DMR, DEM). DSM is a terrain model representing a real world including all real objects on the surface of the Earth (including e.g. canopy). DMR and DEM, on the other hand, depict only the bare ground without buildings or vegetation.

The evolution of digital terrain models

The aim of this short introduction is to outline the history and development in the field of digital view of the Earth surface. The attempts to depict the surface have been with the mankind from the beginnings and the efforts to depict the surface accurately and to model the uppermost geological layers will probably always be here.

People have always strived to use all available means to find out representations of various terrain elements. Various pictures of the landscape originating in ancient ages may be considered the oldest representations (Maune et al., 2001). Such pictures show general information about terrain characteristics such as the shape and colour of the terrain, however the metrics and accuracy of such pictures were obviously extremely poor (Moore et al., 1991). As already mentioned, the first digital terrain model as we know them was created more than 50 years ago (Miller and LaFlamme, 1958). In the first decades, the focus was on the reliability of the models. The techniques commonly used for evaluating the quality and accuracy of the models were based on statistical comparison of the digital terrain models with small reference areas (Podobnikar, 2009). Till the end of 1990s, data for digital models originated usually from airborne photography and photogrammetry – the acquired stereoimages were manually processed. Another frequently used technique was represented by vectorisation of isohypses from existing maps (Li et al., 2004; Podobnikar, 2009).

Since 1990s, the quality of available DTMs has increased considerably, in particular due to several important factors (Podobnikar, 2009). The first significant breakthrough was represented by development of new data sources, in particular those from satellites and airplanes (Maune et al., 2001). For coarse scales, the utilization of satellite radar interferometry increased while for finer mapping, airborne laser scanning/LiDAR has become a valuable and relatively widely available source (Kraus and Pfeifer, 1998). A separate chapter is represented by installation of sensors on unmanned systems (unmanned aerial vehicles, UAVs). Thanks to the structure from motion (SfM) method, it is possible to create relatively accurate digital surface models from photographs acquired using consumer grade cameras. More sophisticated UAV platforms can even carry a LiDAR sensor. Models created from UAV data are more accurate than those from aforementioned platforms (Glowienka et al., 2017; Tulldahl and Larsson, 2014). In addition, a multispectral or even hyperspectral camera can be mounted as a sensor on a UAV, thus providing additional information (Kaneko et al., 2014; Mårtensson and Reshetyuk, 2017; Pavelka et al., 2015).

Another significant factor contributing towards the improvement in quality of the DTMs is a better availability of various data sources. Besides airborne imaging, radar imaging and isohypses, other datasets containing information about elevations and/or singularities became available. Such datasets include e.g. the anchor points of the geodetic network, boundary points of the cadastre, building databases, spot elevations and other additional information associated with building of high-rise buildings, line constructions, motorways, etc. (Mark A. Maloy and Dean, 2001).

A major leap in landscape modelling came with the increasing availability of LiDAR data. Many countries have released their national point clouds representing a continuous model of the landscape for scientific or, even better, public use. In Europe, such countries include e.g. Poland and Finland (Glowienka et al., 2017; Valbuena et al., 2016). In the Czech Republic, the data release is very slow. Although a LiDAR-derived digital relief model as well as digital surface model are available for almost the entire area of the Czech Republic, the data are only available for sale. Users can purchase individual 2.5 km x 2.5 km mapping sheets; where the area of interest is large, many mapping sheets are needed, which makes the data relatively expensive in the end. Alternatively, users can obtain a raster model of both terrain and surface free of charge; unfortunately, the method of processing the original data and of creating the model is not clearly described, which carries doubts about usability of the models for particular analyses. Obtaining the original raw data is almost impossible for an ordinary user; the situation is somewhat better with the vector data that are available for purchase.

The last but not least factor influencing the data quality is bringing the digital models closer to the public through various mapping geoportals, mapping applications and satellite navigations. We can name e.g. Google Earth, Google Maps and Microsoft Virtual Earth (Podobnikar, 2009). Real digital models are also utilized in the computer gaming industry (Abraham, 2018).

Introduction to visibility analyses

Digital terrain models serve as an essential input for various tasks in environmental and ecological modelling. In the field of applied ecology, modelling of areas that are visible or invisible from an observer point (visibility analyses) are a typical utilization of DTMs. Visibility analyses are a common method of evaluating the impact of buildings on the landscape (Lee and Stucky, 1998; Nijhuis and Van Der Hoeven, 2018), designing landscape elements (Chamberlain and Meitner, 2013), tagging landscape photographs (Brabyn and Mark, 2011), source for historical settlement studies (Sevenant and Antrop, 2007), placement of coastal aquaculture sites (Falconer et al., 2013), or military structures (Smith and Cochrane, 2011). In principle, two extremes are sought for by visibility analyses – either search for spots with maximum visibility (e.g. for placing an object that positively affects the landscape beauty) or with minimum visibility (to find a spot for an object ort construction that is detrimental for landscape beauty).

Nowadays, most geoinformation software is capable of performing such analyses. The viewshed tool is available in geographic programs or platforms such as ESRI, Grass GIS, SAGA GIS, Quantum GIS, etc. (Aben et al., 2018; Sang et al., 2016). A digital elevation model is typically used as an input raster, with a cell value usually corresponding to the elevation of the site. The user must also define the observer location (one or more observer points) and the target area for which the evaluation should be performed. The algorithm may take into account supplementary variables as well (Aben et al., 2018). The basic ones include setting of the height of the observer and of the target object. If, for example, the height of a wind turbine is calculated, the height of the cell where it is expected to be constructed is increased by the expected height of the turbine. It is also possible to limit the azimuth of observation, distance and area or the vertical angle for which the algorithm is expected to perform calculation (Fisher, 1993).

Simple algorithms of visibility analysis are based on comparing elevations on the line connecting the observer's location and the target cell or point (so-called line of sight). The elevation of all cells on the line is evaluated and if elevation of any pixel between the observer's cell and target cell is above the connecting line, the target cell is invisible from the observer's point. If visibility analysis is calculated for the entire area, this calculation is performed individually for every pixel in the area (Kim et al., 2004a; Sang et al., 2016). More complex visibility analyses combine calculations of multiple lines of sight. Such an example may be represented by viewshed analyses, i.e., calculations of areas that can be observed from the observer point. An analogical function is an extended viewshed tool titled "observer point". This tool adds another layer of complexity – it allows several observer points and provides information about the individual observer points or combinations of points from where certain parts of raster can be seen (Caldwell and Mineter, 2003). Even more complex tools are represented by multiple viewshed/cumulative viewshed analysis that calculate mutual visibilities for all pixels in the area. The number of resulting visibility rasters is the same as the number of cells in the area of interest. All those binary rasters are then summed, forming a cumulative raster depicting the most and least exposed cells. (Danese et al., 2009; Fisher, 1996; Wheatley, 1995).

Input data for viewshed analysis

As soon as 1997, Dean (1997) referred that visibility analyses performed on the basis of a DEM/DMR only, i.e, ignoring the vegetation cover, cannot provide correct results in areas with vegetation. It is in accordance with results of Maloy and Dean (2001) who performed visibility analyses using a basic vector map of USA at a scale of 1:24,000 and showed a 56.7 % agreement with reality. Hence, to reach useful results of visibility analyses, it is necessary to use DSMs that model objects on the terrain accurately.

When a ready-to-use DSMs (e.g. LiDAR-based DSMs provided in some countries) are not available, input data for visibility analyses can be prepared from vector datasets. To create a DSM from a vector map, heights of objects such as vegetation and buildings must be added to the DEM interpolated from contour lines. Object heights are not standardized, however, and the assigned heights thus depend on an expert estimate of the user (Klouček et al., 2015; Wallentin et al., 2008). The accuracy of visibility modelling therefore depends on the accuracy of the input surface model combining the accuracy of DTM with the accuracy of the object heights estimations (Lake et al., 2000a; Sander and Manson, 2007a).

Canopy height models (CHMs) are a special type of the digital surface models prepared as a difference between the DTM and DSM after excluding non-vegetation objects such as buildings (where those are not excluded, the correct term is normalized digital surface model, nDSM). CHMs ha been primarily intended for imaging tree crowns (Lim et al., 2003; Lisein et al., 2013); a typical shape of the deciduous tree resembles a canopy, which gave name to this type of model. Canopy height models have an important position in forestry where they can be used for estimation of the wood matter or carbon deposits (Corona and Fattorini, 2010; Steinmann et al., 2013), in agriculture (Næsset, 2002) or in ecosystem modelling. Moreover, they can also represent valuable input for visibility analyses in non-urbanized areas, providing a more realistic model of vegetation then a simple expert estimation of the vegetation height(s).

As mentioned above, the term canopy height model is very close to another term – normalized digital surface model (nDSM). Both are created as the difference between the digital surface model and digital terrain model; nDSM however contains all objects present on the terrain including anthropogenic structures while canopy height model in the strict sense only shows the vegetation. As laser beams can penetrate under the tree crowns and branches, both DSM and DTM may be derived by point filtering from LiDAR data, which is one of advantages of LiDAR over photogrammetry. Various reflections originate at various levels of the vegetation cover, down to the level of the terrain, which allows the production of DSM, DTM and, in effect, CHM from the same data acquisition mission (Hyyppä et al., 2008; Lim et al., 2003). Also, it is possible to obtain information on the vertical structure of the vegetation under the level of the treetops in this way. When planning a LiDAR mission aimed at deriving a CHM, the choice of the season in which the mission is performed is very important. Such missions are often planned for the leaf-off period. Where this is true, the processed CHM does not represent a full canopy, rather only fragments of tree branches (Kim et al., 2009; Wasser et al., 2013a). On the other hand, a mission in the leaf-off period necessarily provides better information about the terrain as there are significantly more reflections from the ground than during the leaf-on season (Moudrý et al., 2019).

Some studies attempted to use data from airborne or satellite imagery through photogrammetry (Bohlin et al., 2012; Huang et al., 2009; Mora et al., 2013; Nurminen et al., 2013; Véga and St-Onge, 2008), which can also provide a DSM suitable for visibility analysis. In comparison with LiDAR, it is however not possible to use the information about the terrain under the canopy (White et al., 2013). Besides, merging individual images into a large complex imagery is quite difficult due to the amount of vegetation characteristics, repeated texture of vegetation, complicated work with ground control points, etc. (Baltsavias et al., 2008; Eisenbeiß et al., 2009; White et al., 2013). Photogrammetry and satellite imagery is therefore not suitable for creating CHM on itself. However, as many countries provide accurate DTMs derived from airborne LiDAR data

(Bohlin et al., 2012), it is possible to combine airborne or satellite photogrammetry with such DTMs, which results in creating an accurate and up-to-date model of vegetation. Such a combination therefore offers high spatial and temporal resolution. Studies published in the last years also demonstrated the use of UAVs for creating CHM (Dandois and Ellis, 2010; Moudrý et al., 2019). UAVs can capture a significantly smaller area during one mission than a manned airplane, they however provide spatial resolution in the order of centimetres. Another advantage is the possibility of frequently repeated measurements, which provides the information about the vegetation growth over time in the area of interest.

Quality of spatial data

The quality of spatial data used for modelling usually plays a crucial role not just in digital terrain modelling but in landscape ecology in general. Many studies that focused on environmental characteristics and/or relationships among individual components of the environment discussed the effect of the quality of input geodata on resulting models (Klouček et al., 2015; Kumi-boateng and Yakubu, 2010; Li et al., 2012; Meek et al., 2013; Moudrý and Šímová, 2012; Sharma, 2009). The use of suitable data is as important in visibility analyses as in other fields of landscape ecology. Often, unfortunately, case studies utilize data that are readily available and appear suitable at the first sight. Worse, verification of the results is then often omitted or neglected and the poor fit with reality is therefore not even detected. This is discussed e.g. by Lecours (2016) who points out the importance of the selection of suitable data for ecological studies. If the selection of data and variables suitable for an analysis is left to a subjective opinion, the resulting accuracy of the analysis may suffer. In his earlier work, Lecour (2015b) also pointed out another frequent problem associated with the relationship between theoretical spatial resolution and positional accuracy. Researchers often settle with a sufficient spatial resolution of the input data, disregarding the positional accuracy, which can also lead to suboptimal results.

In the ecological literature, three types of scales are usually distinguished – spatial, temporal and thematic (Lecours et al., 2015a). In various contexts, the term spatial scale was defined in several ways. Typically, it concerns spatial characteristics of the object or process including the level of detail and geographical extent

(Lechner et al., 2012). Similarly, temporal scale gives us information about the temporal level of detail, e.g. seconds, days, seasons, etc. Thematic scale, also called level of organization, organizational scale, or ecological organization, is associated with the level at which objects of study are described, for instance taxonomic resolution (Larsen and Rahbek, 2005).

Speaking of the ever increasing utilization of LiDAR-based DSMs, it is necessary to also mention the processing of raw data. No matter how good the raw spatial data are, improper data processing can easily wreck much of the information the raw data contains. The quality of the resulting model is therefore not solely dependent on the density of the original raw point clouds. The computational algorithm (Khosravipour et al., 2016; Xiaoye Liu, 2008) as well as the selected interpolation method (Anderson et al., 2005; Guo et al., 2010) also play an important role. The maximum raw point density therefore does not necessarily warrant the maximum accuracy of the result (Anderson et al., 2006; Jakubowski et al., 2013; Ruiz et al., 2014). In other words, the use of different methods of raw data processing can lead to different quality of models. The difficulties associated with the raw data processing along with easy availability of ready-made LiDAR-based raster DSMs however often steer the researchers in the environmental sciences towards using such "easy" data. However, as the ready-made data products (especially those created at the nationwide extent) are not being prepared in view of particular research needs or of particular analyses, the suitability of such models for studies and analyses at detail scales can be disputable (Mondino et al., 2016) and bulk use of such data can in effect lead to errors in decision making in the field of environmental management.

As mentioned above, very few papers have dealt with the effect of the spatial precision of input geodata on the reliability of results from visibility analyses, even though some authors had previously noted a potential effect (Fisher, 1992; Huss and Pumar, 1997) and input geodata's influence on the results of spatial analyses has been demonstrated many times in other fields. In the case of visibility modelling, the accuracy of the resulting model, whether based on a basic viewshed algorithm or its more advanced variants, potentially depends on the precision of the input digital surface model (DSM), which combines the accuracy of a digital terrain model with the correct heights and spatial determination of objects within the model, particularly of vegetation and structures. Examples of rare studies dealing with input data precision were presented by (Lake et al., (2000b) and Sander and Manson (2007b) who focused on modelling structures representing vertical obstacles to visibility.

Visibility modelling in applied ecology

Besides the use of visibility analyses for historic studies (Carter et al., 2019; Ogburn, 2006; Sevenant and Antrop, 2007); archaeological surveys and exploration (e.g., Paliou, 2011); landscape planning (De Montis and Caschili, 2012); search for sites suitable for buildings that could impact the landscape character (Fernandez-Jimenez et al., 2015) or ski areas (Geneletti, 2008); placing military structures (Smith and Cochrane, 2011); tagging landscape photographs (Brabyn and Mark, 2011); or analysing effects of animal species introduction (Kizuka et al., 2014), visibility analyses have been lately applied to the sea surface (for example in a search of sources of "visual pollution" in the Baltic sea). Anthropogenic activity such as offshore wind farms development, shipping activity, resource extraction platforms or marine aquaculture can have adverse impacts on the visual quality of coastal landscapes (Depellegrin, 2016). Other studies have lately focused on the visual impact of the seashore on social aspects. Qiang et al. (2019)published a study, which showed that view of blue spaces (e.g. ocean, lake, and river) have positive effects on human health and mental well-being. The primary research originated from the use of the digital terrain model (in this case that of seashore and sea level) and the use of visibility analyses in GIS. Similar methods were applied by colleagues of Quiang who in their study Poudyal et al. (2010), investigated the effect of view of a forest on the market price of real estates serving as observer points.

Jiang et al. (2014) also used results of visibility modelling as one of environmental variables when modelling distribution of the Siberian tiger (*Panthera tigris altaica)* and showed it to be a significant component of the resulting distribution model. Alonso et al. (2012) used visibility analysis as a significant factor for selection of nesting locations of the great bustard *(Otis tarda)*. The aim of visibility analyses in such applications is not to investigate whether or not the individuals can see each other but rather to act as a surrogate of the risk of predation or availability of the prey for the predator. A study focused on this topic was published by Olsoy et al. (2015) who evaluated the visibility of potential predators from the perspective of their prey, and with the prey options to take cover from the perspective of predator viewpoints. While the previously mentioned works analysed visibility from animals' perspectives, Kizuka et al. (2014) utilized for their prediction of species occurrence visibility modelled for a human observer. They predicted distribution of two invasive introduced fish species, namely the bluegill (*Lepomis macrochirus*) and largemouth black bass (*Lepomis macrochirus*) in water bodies in a Japanese agricultural landscape. As the connectivity of those water bodies was low, they estimated the influence of human introduction on the presence of those fish in individual reservoirs. They found that the visibility of the reservoir from a road was a better predictor of the presence of those fish than distance from the road or population density in the area.

A paper by Aben et al. (2018) literally calls for increased use of so-called viewshed ecology, i.e, the method of observing the world by animal eyes. If correct parameters and variables are known, we can use visibility analyses for this purpose. An ideal approach would be however represented by scene analysis rather than visibility analysis, i.e., by simulating the orientation in the terrain not on the basis of an elevation raster but rather on the basis of a complete 3D model of a landscape (Murgoitio et al., 2014a). A fictional observer would thus in effect look rather on a photograph-like image than a raster image. This has also been discussed in the work by Sang et al. (2016). At present, we are however limited by hardware capabilities. The data are becoming more accurate, available, with higher density, however the standard computational methods and algorithms are not sufficient to process the huge quantity of such data. Thus, it would be necessary to change the algorithms as well as hardware for such analyses (Xia et al., 2011; Zhao et al., 2013).

Datasets usable as input data for visibility analyses in Czechia

Mapping has a long history in Czechia, reaching as far as the beginning of the second millennium. The oldest surviving map of Czechia is Klaudyan's map from 1518 (Mikšovský and Zimová, 2006). Of course, the map accuracy and practical usability of maps has since increased immensely. This chapter describes datasets that are at present available in the Czech Republic and that are utilized in the individual studies constituting parts of the thesis. Those datasets include geodata acquired both by topographic mapping and by airborne laser scanning.

ZABAGED

ZABAGED (an acronym from Základní báze geografických dat České republiky – Basic dataset of geodata of the Czech Republic) is a digital geographical model of Czechia. The planimetric part of ZABAGED comprises at present 125 types of geographic objects including settlements, communications, distribution networks, water bodies, administrative areas, protected/nature conservation areas, vegetation and surface, terrain and selected survey control points. Objects are represented by a two-dimensional vector component and a descriptive component containing qualitative and quantitative information about the objects. The dataset scale is 1:10 000 (Šíma, 2016)

National map on a 1:5,000 scale

The national map on a 1:5,000 scale (SM 5) is a principal national mapping product of a fine scale. It depicts the entire area of Czechia in a continuous collection of mapping sheets – the entire area is recorded on 16,301 mapping sheets of 2x2.5 km. The dimensions and marking of the mapping sheets are derived from the mapping sheets of the National map on a 1:50,000 scale by dividing them into 100 parts. The layout of the mapping sheets of the National map 1:50,000 is, unlike that of base maps at medium scales, parallel with the axes of the S-JTSK coordinate system (the National map 1:50,000 is, unlike the Base map of the Czech Republic, not published). SM 5 contains planimetry, altimetry and map lettering. It is the finest scale national mapping product containing altimetry. The principal source of planimetric data are cadastral maps, of altimetric data the Base map of the Czech Republic at a 1:10,000 scale. The map lettering comes from both the cadastral maps and a database of geographic names of the Czech Republic Geonames. SM 5 was available only as a fully analogue map prior to 2001, so-called National map (derived) (abbreviated SMO-5) for the entire Czechia. In 2001 to 2007, a vector version, SM-5, was created for approximately 30 % of the Czech Republic and complemented with raster files acquired by scanning of the original SMO-5 mapping sheets for the remaining 70% of the Czech Republic. In 2008-2009, an innovated SM 5 was prepared, including a change of technology, aiming at creating a product called "Vector data of new form SM 5" (and its derived raster version) depending on the gradual digitalization of cadastral maps. The vector version of the new product provides the users with feature type signification, e.g. for spatial planning purposes (ČÚZK, 2019).

DMÚ 25

In the Military Geographic and Hydrometeorologic Office in Dobruska (MGHO), a Digital Model of the Territory (DMÚ 25) is being prepared since 1991 as a complex of data and methods for acquisition, processing and updates of digital information about the territory, the principal part of which is the information about geographical objects acquired from digitalization of 85 sheets of amended topography map TM100 (the basic TM25 at a 1:25,000 scale was created and is updated based on direct mapping, i.e., field collection and verification of data with maximum utilization of airborne imaging; other topography maps up to the 1:1,000,000 scale were derived from it). When creating individual topographic maps of the TM series, the principal properties such as readability, clarity, or map key must be taken into account so when creating a coarser scale maps, the content is simplified – generalized – according to the needs of the particular map. Data are organized in seven logical layers: waters, communications, pipelines/power lines/communication lines, vegetation and land cover, settlements, industrial and other topographic objects, boundaries and fences, terrain relief. Those logical layers are further divided into 20 data layers. The level of the data accuracy and generalization correspond to the 1:25,000 scale (VTÚ, 2019).

ArcČR 500

ArcČR 500 is a digital vector geographic database of Czechia at 1:500,000 scale. It follows up on similar databases created by the Esri Company. The database contains clearly arranged geographic information about the Czech Republic. The data allow a broad spectrum of spatial analyses and visualization as well as linking to statistical data. The geographical information in ArcČR 500 are divided according to the thematic categories including basic geographical elements, layouts of national/state mapping products and administrative structure. ArcČR 500 was created in cooperation of ARCDATA PRAHA, s.r.o., and the State Administration of Land Surveying and Cadastre. It was developed on the basis of maps and databases provided by the State Administration of Land Surveying and Cadastre. The raster digital relief model originates from SRTM data (ArcData, 2019).

LiDAR datasets

The use of laser as a tool for remote sensing has been around for approximately 40 years. In 1960s and 1970s, many experiments demonstrated the use of reflected laser beam in mapping not just the Earth surface but even for study of the surfaces of Moon, Mars and Mercury (Smith et al., 1997; Sun, 2018). Due to the ever more reliable sensors and ever improving resolution of LiDAR systems, LiDAR became one of the most prominent tools in remote sensing and mapping (Ackermann, 1999). This compact and relatively affordable technology has been procured by numerous institutions and users (Klemas, 2015; Tulldahl and Larsson, 2014).

The analysis performed by the State Administration of Land Survey and Cadastre of the Czech Republic in 2006 – 2008 led to a conclusion that the existing altimetric models of the Czech Republic originating from the digitalization of topographic maps are not sufficient for the purposes of the state administration. For this reason, the State Administration of Land Survey and Cadastre prepared a project of creating a new altimetry model of the Czech Republic (ZÚ, 2008). In 2009, an agreement between ministries was signed detailing the cooperation on new altimetric mapping of the Czech Republic that was to take place in 2009-2015. Airborne laser scanning (ALS) was chosen as the most suitable method for this purpose (CUZK, 2009). The resulting products of this project include the Digital Relief Model of the 4th Generation (DMR 4G), Digital Relief Model of the 5th Generation (DMR 5G) and a Digital Surface Model of the 1st Generation (DMS 1G). For the purposes of ALS, the Czech Republic was divided into three areas; the first scanned area was the Central Band in 2010 while in 2011 it was the West Band and in 2013, with a one year delay, the project was completed by scanning the East Band. The used reference coordinate system was UTM / WGS 84, Band 33. The ALS data were acquired using the Litemapper 6800 system consisting of an airborne laser scanner RIEGL LMS Q-680, a recording device, an onboard GNSS system and an inertial measurement unit (IMU) by IGI company (Dušánek, 2014).

To create a digital relief model, the point cloud had to be filtered to identify ground points and above ground objects. The automatic filtering utilized the SCOP++ software with a robust filtering algorithm developed by the Technical University of Vienna. This algorithm works iteratively – it selects lowest points in cells of a regular grid and interpolates a digital terrain model based on such points. In the next step, weights are assigned to individual points – points that are too far from (high above) the first iteration of the terrain are assigned a zero weight and do not enter further analysis. The automatic filtering thus divides the point cloud into ground and above ground points; it is however not perfect and manual inspection and editing of the results of the automatic filtering was also performed.

The processing resulted in three aforementioned models. DMR 4G was the first version created more or less automatically. The reason for releasing this model was to create the first version of the altimetry model required for creating

Orthophoto CR where the DMR 4G constituted the principal data for orthogonalization of airborne survey images. The input data for the DMR 4G were the automatically filtered LiDAR data mentioned above. From data that were identified by the robust filtering method as terrain points, the lowest point in a regular 5x5 m raster was selected. As the point cloud still contained errors caused by incorrect automatic classification, those were manually edited. Such an irregular network of points then served as a basis for interpolating an altimetry model by linear prediction in a 5x5 m raster. The interpolation was performed separately for the geodetic reference systems WGS 84/UTM and SJTSK. DMR 4G is a model provided as a regular grid at a 5x5 m resolution with the median error of 0.30m in the open terrain and 1.00 m in the terrain covered with vegetation (Bělka et al. 2010).

The other digital relief model, DMR 5G, was prepared from data that underwent detailed manual inspection and editing. The final DMR 5G was derived by subsequent three-step model "flattening". In the first step, the lowest point in every 1 x 1 m cell was selected, which reduced the number of points in spots where several blocs were connecting and remainders of the noise from the overlap of scanning bands were removed. In the second step, unwanted local unevenness (e.g. ploughed-up fields) was removed by adjusting the original point elevations by up to ±5 cm, which also flattened the surface. In the third step, the point cloud was diluted while maintaining the median elevation error (Bělka et al. 2012).

The last model of this series is the digital surface model of the 1st generation, DMS 1G (it is a first generation model as such a model was not available for the area of the Czech Republic before). DMS 1G was derived by automatic algorithms. It is in principle DMR 5G supplemented with the above ground objects. In the built-up areas, points that were identified as buildings by automatic filtering were added where they fit the building contours from the cadastre. As far as vegetation was concerned, only reflections that were above ground with a minimum area of 25 m^2 were included in the model (Dušánek et al. 2016).

Summary

Digital terrain models and applied visibility analyses are widely used tools of geoinformatics. They have an irreplaceable position in landscape evaluation as whole as well as in evaluation of individual interventions. Similarly, they can facilitate evaluation of projects in the fields of landscape ecology, animal ecology, urban planning, architecture and civil engineering, urbanization, planning of tourist paths, lookout towers, etc. The present day modelling of terrain and visibility however faces an opposite problem than it has just a few decades ago – too much available input geodata that are of varying quality. When choosing input data for a particular analysis, it is therefore necessary to take the aims of the analysis into account and to consider the model accuracy required for the particular purpose. It is however not uncommon by far that we can see unsuitable selection of input geodata for various research or management purposes, which in turn leads to various misinterpretations and mistakes in the application of the data. The review above aimed at providing a brief summary of the current state of this scientific field and to bring a complex view on data sources, digital terrain models and visibility analyses.

CHAPTER IV
Study 1

Title:
How does data accuracy influence the reliability of digital viewshed models? A case study with wind turbines

Authors:

Tomáš Klouček, **Ondřej Lagner**, Petra Šímová

Author's contribution: 10%

Investigation and field work; Consultation and co-writing of the original draft

Journal:

Applied Geography, IF (2015) **2.565**; Geography ranking 10/77 – **Q1** (2015)

Full citation:

Klouček, T., Lagner, O., & Šímová, P. (2015). How does data accuracy influence the reliability of digital viewshed models? A case study with wind turbines. Applied Geography, 64, 46-54.

Cited:

11 times cited on WOS (2019 August)

Abstract

Viewshed analysis is a GIS tool commonly used in a number of research and practical spatial analyses. Input data and their spatial uncertainty are important aspects affecting analysis reliability. Given that inappropriately selected input geodata can produce imprecise visibility models and as a result cause incorrect spatial decisions, quantifying the effect of this uncertainty on resulting visibility models is important for the models' subsequent use. The objective of our study was to evaluate the suitability of digital surface models with varying levels of detail (a LiDAR-based model and models based upon vector data at differing scales) for simple (binary) viewshed analysis of wind turbines (three wind parks each containing 3–6 turbines). Visibility models based upon this input data were compared with actual visibility from 150 control points at random locations. The study results confirmed the prediction that the viewshed model based on more precise input data corresponded more closely to reality. Moreover, our study is the first to demonstrate that only the number of false positives–(where the model predicts that an object is visible while in reality it is not) depended on input data precision, while input data did not affect the false negatives. In addition, all vector-based models had far more false positives than false negatives, while the opposite was true for the LiDAR-based model.

When we considered the same number of modeled and actually visible wind turbines as a model's matching of reality, there were matches at 83.6–93.7% of control points (95% confidence interval)-for the LiDAR-based model. For models based upon vector maps of various scales, the intervals were 68.4–82.2% (1:10,000), 59.1–74.2% (1:25,000), and 48.1–63.9% (1:500,000). We recorded false positives in 6 cases with the LiDAR-based model and 26, 39, and 59 cases, respectively, for vector-based models.

Keywords

Accuracy; Digital surface model; LiDAR; Uncertainty; Viewshed; Visibility

Introduction

Viewshed analysis is a GIS tool in standard use for more than two decades (e.g., Fisher, 1992; Nagy, 1994; Sansoni, 1996) to perform numerous scientific and practical tasks. Such analyses enable detection of surfaces that are or are not visible from one or more observation locations, and, inasmuch as visibility is symmetrical, identification of surfaces from which certain objects on the Earth's surface are visible. The wide range of possibilities for its use include, for example, planning telecommunications tower placement (De Floriani et al., 1994); constructing military structures (Smith and Cochrane, 2011), observation towers, and tourist routes (Chamberlain & Meitner, 2013; Lu, Zhang, Lv, & Fan, 2008); selecting sites for new photovoltaic power plants (Fernandez-Jimenez et al., 2015); applications in archaeological research (e.g., Paliou, 2011) and landscape planning (De Montis and Caschili, 2012); and tagging landscape photographs in combination with volunteer geographic information (Brabyn and Mark, 2011). Throughout the time that visibility analyses have been used, their limitations and inaccuracies have been discussed. Fisher (1992) noted two mistaken assumptions in visibility analysis: first, that the input digital elevation model is accurate, and second, that viewsheds constitute a Boolean phenomenon. This author (Fisher, 1996, 1995, 1994, 1992; Peter F. Fisher, 1993) as well as a number of later studies (e.g., Chamberlain & Meitner, 2013; Fernandez-Jimenez et al., 2015; Ogburn, 2006) dealt with the possibilities and algorithms of fuzzy viewshed modeling and visual magnitude, the result of which is a raster giving the probability of visibility or degree of visibility, respectively, and not merely binary visible/nonvisible values. Such algorithms enable incorporation of the studied object's distance from the observer, the observation's solid angle, perspective, and so forth. Recent studies have suggested further procedures for individualizing the viewshed that take into account such aspects as solid angle, defined by Domingo-Santos et al. (2011) as the surface area of the observer's retina covered by a given object, and the vertical dimension of terrain, which combines the slope of the visible surface, difference in elevation between the observer and the visible terrain, and relative aspect of the terrain in relation to the observer into a new Vertical Visibility Index (Nutsford et al., 2015). One field with particularly apparent efforts to bring visibility analyses closer to reality and human perception is that of assessing the visual impact of wind turbines (WTs). Here, in addition to the aforementioned improvements in GIS algorithms and creation of specialized software (Manchado et al., 2013), we encounter a number of other evaluation techniques. These include verbal questionnaires, photo-based questionnaires, questionnaires based on computer simulation, and questionnaires completed while viewing actual landscapes (for a review of these methods' use, see Molnarova et al., 2012). Research designed in this way (e.g., Betakova et al., 2015; Bishop & Miller, 2007) conveys information on distances from the observer at which WTs have the greatest visual impact, frequently in combination with such other parameters as the number of WTs, rotor movement, and the landscape's scenic beauty. These results provide a solid foundation for planning studies focused on GIS viewshed analysis quality and selecting specific parameters (landscape character, number of WTs in the study area, viewshed distance, and so on).

Based on the number of articles published, we can state that the study of those phenomena affecting visibility in terms of humans' subjective perception and improvements to GIS viewshed algorithms constitute a rather frequent topic of research (e.g., Bishop & Miller, 2007; De Montis & Caschili, 2012; Domingo-Santos et al., 2011; Germino et al., 2001; Kim, Rana, & Wise, 2004; Manchado et al., 2013). Very few papers, however, have dealt with the effect of the spatial precision of input geodata on the reliability of results from visibility analyses, even though some authors had previously noted a potential effect (Fisher, 1992; Huss and Pumar, 1997) and input geodata's influence on the results of spatial analyses has been demonstrated many times in other fields. For example, a potentially analogous situation can be seen in ecology, where geodata's spatial uncertainty is an established concept and its effect on analytical results is a known fact (for review see, e.g., Barry & Elith, 2006; Moudrý & Šímová, 2012). In the case of visibility modeling, the accuracy of the resulting model, whether based on a basic viewshed algorithm or its more advanced variants, potentially depends on the precision of the input digital surface model (DSM), which combines the accuracy of a digital terrain model with the correct elevation and spatial determination of objects within the model, particularly vegetation and structures. Examples of rare studies dealing with input data precision were presented by Lake et al. (2000) and Sander & Manson (2007), who focused upon modeling structures as vertical obstacles to visibility. In order to create a DSM, data at differing spatial scales are generally used and are based upon both remote sensing and ground mapping. Probably the most precise inputs are LiDAR-based surface models (see Castro et al. 2015; Lake et al., 2000; Murgoitio et al. 2014). At the same time, LiDAR data is also the most expensive as well as the most demanding in terms of processing the original point cloud into a raster or triangulated surface (a triangulated irregular network). Moreover, it frequently is unavailable for a given study location. A question thus arises as to the degree to which LiDAR-based surfaces can be replaced within visibility analysis by surfaces created through such approaches as using contour lines with elevation values and objects with expertly assigned height, such as polygons of forests boundaries and structure footprints, as well as a question as to the effect that the scale of the data used has on the reliability of visibility analysis.

Evaluating whether a viewshed model has identified visibility in accordance with reality, and therefore whether the tested algorithm and/or input data used are appropriate for the modeling purpose, requires comparison with a control model, a control simulation, or a control dataset. Various methods are used for model verification, including to compare the visible area with a reference visibility model (Lake et al., 2000a; Sander and Manson, 2007b), photographic documentation, or 3D visualization (Germino et al., 2001; Mark A. Maloy and Dean, 2001). A rarely used approach is direct comparison of modeled visibility with actual visibility in the field using visual control from predefined locations, as seen in the work of (Meek et al., 2013). Even though other authors have used direct determination of visibility in the field (Lang et al., 2014), they did not use it to compare a model with reality, but rather as the primary method to determine visibility. This was due presumably to their having insufficiently accurate input data for the purpose of their study. Mark A. Maloy and Dean (2001) used viewpoints to obtain comparison photographs and not for direct visibility control.

The visual impact of WTs is a frequently discussed topic in connection with visibility analyses. Given that, to the best of our knowledge, the effect of input data on the reliability of such analyses has not been resolved, we directed our attention to this issue. The objectives of our study were to evaluate the suitability of DSMs with various levels of detail (a LiDAR-based DSM and DSMs based on vector data at differing scales) for simple binary visibility analyses of WTs at three wind parks and to quantify the extent to which visibility models based on these inputs matched reality. We focused on both the overall extent to which the visibility models matched actual visibility in the field and the structure of those errors occurring, i.e., the occurrence of false positives (where the model predicts that an object is visible while in reality it is not) and false negatives (where the model predicts that an object is not visible while in reality it is). We hypothesized that (i) the extent to which a digital visibility model matched reality would depend on the detail of input DSMs, with more-detailed DSMs recording higher match rates, and (ii) the probability of false positives and false negatives would not be equal in visibility models based upon surface models differing in precision.

Matherial and Methods

Study area and input data

The study analyzed the visibility of WTs in the north of the Czech Republic (50°56' N, 15°08' E). The study area covering 300 km2 is characterized by a wide range of elevations (200–1,120 m a.s.l.) and closely related substantial heterogeneity of land cover. Homogenous spruce monocultures predominate at higher elevations and the percentage of forest stands diminishes with decreasing elevation in favor of agriculturally cultivated areas. The selected area therefore combines several landscape types which differ in terms of their conditions for visibility analysis. Within this broader study area, the evaluation focused on visibility in the surroundings of WTs at three wind parks, defined as a buffer with a radius of 5 km around each WT (see Fig. 1). Wind parks with more than one WT (3–6 per park) and the evaluation distance were selected in view of the findings by Betakova et al. (2015).

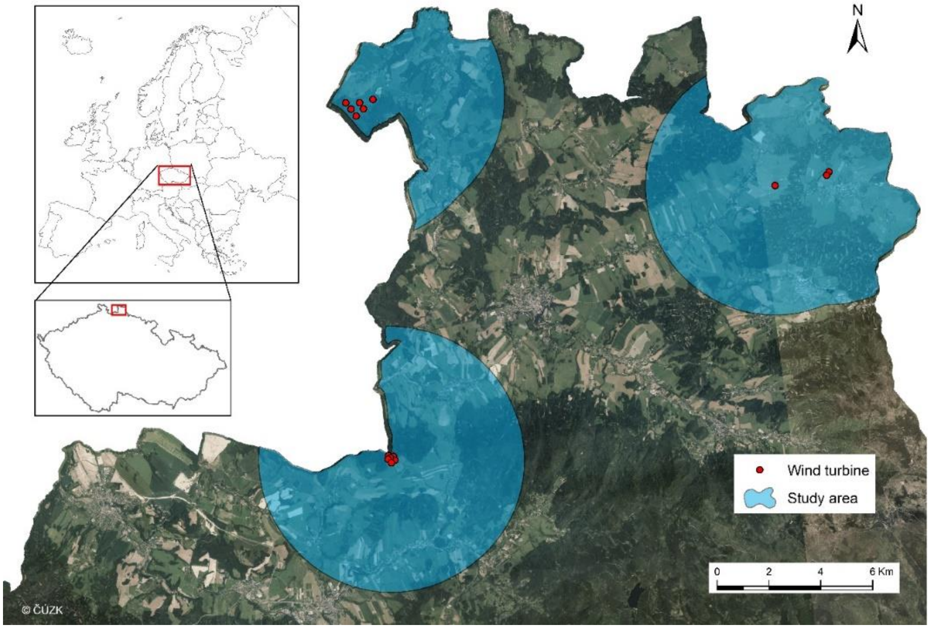

Figure 1. Study area: 5 km buffers around WTs, north Czech Republic.

The study area also has available geodata that differ not only by the method of their acquisition but also in scale and accuracy. Thus, they represent a cross section of products available in the Czech Republic potentially usable for visibility analysis (see Table 1 for details). We used four input datasets, which can be divided into two categories: LiDAR-based and vector-based. The most modern and most accurate is the 1st Generation LiDAR-based DSM of the Czech Republic (LSM), acquired progressively across the entire Czech Republic using airborne laser scanning. Vector-based datasets, within which we include standard, commonly used vector topographic maps, were represented within the study at various scales. Small-scale datasets are represented by a basic national map at a scale of 1:10,000 (Map10). Medium-scale datasets are represented by a vector topographical map called Digital Model of Area 1:25,000 (Map25). Large-scale datasets are represented by a vector geodatabase of the Czech Republic at a scale of 1:500,000 (Map500). In all vector-based datasets, elevation is displayed by contour lines and topography by polygons representing the footprints of individual objects on the ground.

Table 1. Description of input datasets

	Acronym within study	Czech acronym	Scale	Year of last update	Elevation accuracy	Planimetric accuracy	Data description
LiDAR-based dataset	LSM	DMP 1G	Density of elevation point cloud is >1 point/m2	2010	0.4–0.7 m	0.4–0.7 m	Digital surface model represented by elevation point cloud from data acquired by aerial LiDAR covering part of the Czech Republic
Vector-based datasets	Map10	ZABAGED	1:10,000	2011	0.7–5 m	0.5–1 m	Small-scale vector database covering the entire Czech Republic
	Map25	DMU 25	1:25,000	1998	5–10 m	0.5–20 m	Medium-scale vector database covering the entire Czech Republic
	Map500	ArcCR 500	1:500,000	2014	25–50 m	up to 200 m	Large-scale vector database covering the entire Czech Republic

GIS data processing

All GIS analyses were conducted using ArcGIS 10.2 software (ESRI, CA, USA). Based on input geodata, we created four DSMs as inputs for visibility analyses. For vector-based datasets, the DSMs were always calculated as a sum of rasters comprising the terrain (a digital terrain model [DTM]) and objects on the terrain (a digital object model [DOM]). For details, see Table 2. DTMs were calculated by interpolating contour lines using the Topo to Raster method. To create DOMs, we added the estimated elevation of objects on the ground to individual polygons representing said objects and rasterized the polygons. Inasmuch as the study area contains only rural structures mainly comprising houses, we selected the height of 8 m for structures. We assigned the height of 20 m to forest stands as an estimate of the dominant height of forest stands in the area based on data from forest management. Where other woody vegetation types, such as young forests and orchards, were distinguished in the datasets' attributes, we assigned them the height of 5 m. The DSM from the LSM was created by resampling the triangulated irregular network supplied by the State Administration of Land Surveying and Cadastre into a regular raster. All DSMs were created at 5 m resolution.

Table 2. Creation of four digital surface models (DSMs) from input datasets.

DSM		DTM – source elevation data		DOM – source planimetric data
LiDAR-based surface model	=	elevation point cloud		elevation point cloud
Vector-based surface model, 1:10,000	=	MAP10 (contour lines)	+	Map10: forest (20 m), orchard (5 m), built-up area (8 m)
Vector-based surface model, 1:25,000	=	MAP25 (contour lines)	+	Map25: forest (20 m), orchard (5 m), built-up area (8 m)
Vector-based surface model, 1:500,000	=	MAP500 (contour lines)	+	MAP500: forest (20 m), built-up area (8 m)

We assigned the height of hubs (center of blade rotation) to points representing individual WTs according to the wind parks' technical documentation at 40–95 m (OFFSETA parameter) and observer height at 1.8 m (OFFSETB parameter). The Observer Points tool was employed for all visibility analyses. This tool identifies how many and which analyzed objects are visible from each raster location, and so the resulting rasters' pixel values include the number of WTs visible from a given location. Each wind park was analyzed independently and the resulting layers were clipped by the 5 km buffer zones. As it is reasonable to assume that observers in the forest or among structures cannot see anything, we set the value for all forests and built-up zones to zero. The final visibility analysis output is four digital visibility models: a) a LiDAR-based visibility model, b) a vector-based visibility model at 1:10,000, c) a vector-based visibility model at 1:25,000, and d) a vector-based visibility model at 1:500,000 (see Fig. 2 for an example).

Field data collection

The aim of the field data collection and subsequent analysis was to evaluate and compare how digital visibility models matched actual visibility in the field. Prior to the fieldwork, we designated 50 random control points for each wind park (i.e., 150 in total). To avoid spatial autocorrelation of visibility conditions, the minimum distance between control points was set at 200 m. Due to minimal visibility from forest and built-up zones, points were generated only in open areas. At the random control points, the visibility of WT hubs was examined by human eye. Field data were collected in April 2014 under constant meteorological conditions. The weather was clear to partly cloudy, temperatures ranged around 15°C, and wind speeds were under 5 m/s. A portable GPS receiver (Oregon 450t, Garmin) was used to navigate to the coordinates of individual points.

Statistical analysis and evaluation of visibility models' reliability

To evaluate the accuracy of individual digital visibility models, we used as input data the values acquired by comparing visibility modeled at each control point and visibility determined at those points by field examination. We predicted that the LSM-based model would best correspond to reality, followed (in order) by the models based on MAP10, MAP25, and MAP500. We worked in two ways with the research hypothesis that the rate-modeled visibility's match of reality would depend on input DSM precision. First (Section 3.1.), we focused on differences in the number of WTs visible at each control point in the model and in reality. We tested whether these differences between datasets were significant always for two "neighboring" datasets that is to say for the LSM-based model with the MAP10-based model, the MAP10-based model with the MAP25-based model, and the MAP25-based model with the MAP500-based model. As the distribution of the tested values apparently differed from the normal distribution, we used the nonparametric Wilcoxon one-tail paired test.

To compensate for multiple comparisons, we adjusted the significance level for the three tests from $p < 0.05$ to $p < 0.0167$ using Bonferroni correction.

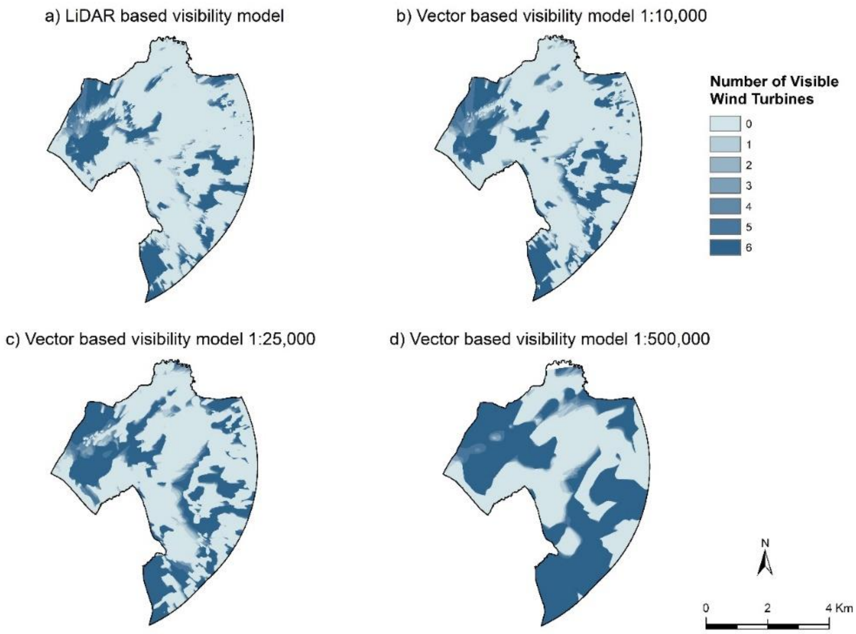

Figure 2. Example of digital visibility models created from different DSMs (5 km buffer): (a) LiDAR-based visibility model, (b) vector-based visibility model at 1:10,000, (c) vectorbased visibility model at 1:25,000, (d) vector-based visibility model at 1:500,000

The second way of comparing visibility models simulated situations when the absolute difference between modeled counts and actually visible counts of objects is not important for a landscape planning task and it is only important whether or not the model agrees with reality in a given way. These binary true and false values were defined in two ways. The first (Analysis 2a, Section 3.2.) worked with absolute accuracy, i.e., for the value to be true the same number of WTs must be visible from the control point as were given by the model. In the second case (Analysis 2b, Section 3.3.), the true value was defined less strictly, simulating such cases as when the visibility of even one WT would be considered as decreasing landscape beauty. For the value to be true, it was therefore enough for the model to determine that some (one or more) WTs were visible from the given location and for some actually to be seen or, alternatively, for no WTs to be visible in both the model and reality. In both cases, we used a test for homogeneity with a binomial distribution to test whether the probability of success (i.e., achieving a value of true) was identical for visibility models based on various datasets. We compared each set with all others using Holm's *p*-value adjustment method to compensate for multiple comparisons. All statistical analyses were done using R software (R Development Core Team, 2015).

For both definitions of the model's matching reality, we evaluated the character of errors, which is to say whether the studied datasets resulted in more false positives (a model overestimating visibility) or false negatives (a model underestimating visibility). For evaluation in accordance with Analysis 2a, we took into account the numbers of visible WTs. Cases where the model predicted more WTs than were seen in reality were considered false positives, and vice versa. According to error definition 2b, false positives occurred when the model predicted that at least one WT would be visible when none were visible in reality. Differences among datasets in terms of the representation of false positives were tested identically as were the total number of errors (test for homogeneity with a binomial distribution, Holm's *p*-value adjustment method).

Results

Difference in the number of visible WTs

A comparison as to the number of visible WTs by which a model based on a given dataset differed from reality confirmed the prediction that visibility models created based on more-detailed input data correspond more closely to actual visibility. The model based on the LSM provided more reliable results than did the model acquired based on vector data at the most detailed scale tested (i.e., MAP10, at a very strong significance level [$p < 0.0001$]). Pairs of models based on vector data at neighboring scales (i.e., MAP10 vs. MAP25 and MAP25 vs. MAP500) can be differentiated at a level of significance an order of magnitude weaker, although still very strong ($p < 0.001$). According to this comparison, the most precise visibility model was the one created from the LSM. Vector-based models' reliability was in accordance with the scale of the input data and the tested datasets yielded significantly different results.

Model matches reality only when the number of visible WTs is the same in the model and in reality

As seen in Table 3, the number of true values (i.e., the number of control points at which the model agreed with reality according to definition 2a) diminished with decreasing precision of input data. In the case of the LSM-based model, there was disagreement at 11.3% of control points, while the model based on MAP500 disagreed in almost half of cases (44%). A similar trend results from mutual comparison of the reliability of models based on individual datasets using the test for homogeneity with a binomial distribution. The LSM-based model displayed significantly better results than did all vector-based models (the significance of the difference strengthened with decreasing vector dataset precision, see Table 4) and MAP10 was significantly better than was MAP500. For neighboring vector datasets, however, it cannot be said that the MAP10-based model provided results significantly different from those of the MAP25-based model; similarly, the reliability of the MAP25-based model did not differ significantly from that of the MAP500-based model.

Table 3. Relative reliability of visibility models as the number and percentage of cases where the model matched reality. Match definition 2a (the number of visible WTs must agree), number of control points n = 150.

Digital Visibility Model	True	False	Relative accuracy (%)	95% confidence interval
LiDAR-based visibility model	133	17	88.7	83.6–93.7
Vector-based visibility model, 1:10,000	113	37	75.3	68.4–82.2
Vector-based visibility model, 1:25,000	100	50	66.7	59.1–74.2
Vector-based visibility model, 1:500,000	84	66	56.0	48.1–63.9

Table 4. Mutual comparison of reliability of visibility models according to match definition 2a (p-value of the test for homogeneity with a binomial distribution, Holm's p-value adjustment). Significant values are in bold.

	LSM	MAP10	MAP25
MAP10	0.0129		
MAP25	< 0.0001	0.1507	
MAP500	< 0.00000001	0.0027	0.1507

Model matches reality if at least one WT is visible in the model and in reality or none are visible in the model and in reality

The numbers of control points at which the model agreed with reality according to definition 2b are given in Table 5. The number of true values recorded follows the same trend as in the previous case, as a more-detailed input data scale corresponded to increased matching between modeled visibility and actual visibility. The LSM-based model failed to match reality at only 3.3% of control points, whereas the MAP500-based model disagreed in almost one-third of cases (28%). As seen in Table 6, mutual comparison of model reliability shows that some model pairs were not significantly different. For match definition 2b, where the model agrees with reality in more cases than it does for the stricter match definition 2a, the LSM-based model was not significantly more accurate than was the MAP10-based model. The MAP10-based model, however, yielded better results than did the models created using MAP25 and MAP500. We can therefore say that for this very loose definition of matching between reality and model it is apparently possible to replace LiDAR-based data with vector data at a similar scale (1:10,000). There nevertheless was still a clear trend that visibility analysis using more-detailed data provided more reliable results.

Table 5. Relative reliability of visibility models as the number and percentage of cases where the model matched reality. Match definition 2b, number of control points n = 150.

Digital viewshed model	True	False	Relative accuracy (%)	95% confidence interval
LiDAR-based visibility model	145	5	96.7	93.8–99.5
Vector-based visibility model, 1:10,000	138	12	92.0	87.7–96.3
Vector-based visibility model, 1:25,000	121	29	80.7	74.3–87.0
Vector-based visibility model, 1:500,000	108	42	72.0	64.8–79.2

Table 6. Comparison of reliability of visibility models according to match definition 2b (p-value of the test for homogeneity with a binomial distribution, Holm's p-value adjustment). Significant values are in bold.

	LSM	MAP10	MAP25
LSM			
MAP10	0.2062		
MAP25	< 0.001	0.0215	
MAP500	< 0.00000001	< 0.00001	0.2062

Character of errors

For both cases of model error definition (false values according to 2a and 2b), it is apparent at first sight that the number of false negatives recorded did not depend on input data precision (see Table 7). Regardless of how we defined true and false values for this study, the datasets used differed in the extent to which they overestimated visibility (false positives), with visibility overestimated more by models based on less-detailed data. All vector-based models, in addition, had more false positives than false negatives, while the opposite was true for the LSM-based model. The occurrence of false positives in individual models mostly differed significantly between neighboring models, although, as in the evaluation of the total number of true and false values, there were cases where differences in the reliability of models versus neighboring datasets were not significant (see Table 8 for *p*-values).

Table 7. Structure (number of cases) of false positives and false negatives in models based on individual datasets. 2a false positive: the model predicts more WTs to be visible than are in reality. 2b false positive: the model predicts at least one WT to be visible while in reality none are visible.

Digital viewshed model	False negative (2a)	False negative (2b)	False positive (2a)	False positive (2b)
LiDAR-based visibility model	11	5	6	0
Vector-based visibility model, 1:10,000	11	2	26	10
Vector-based visibility model, 1:25,000	11	5	39	24
Vector-based visibility model, 1:500,000	7	4	59	38

Table 8. Mutual comparison as to occurrence of false positives in models based on individual datasets. 2a false positive: the model predicts more WTs to be visible than are in reality. 2b false positive: the model predicts at least one WT to be visible while in reality none are visible. p-values, test for homogeneity with a binomial distribution, Holm's p-value adjustment. Significant values are in bold.

	LSM	MAP10	MAP25
2a false positive			
MAP10	0.011		
MAP25	< 0.000001	0.036	
MAP500	< 0.0000000001	< 0.00001	0.064
2b false positive			
MAP10	0.001		
MAP25	< 0.0000001	0.093	
MAP500	< 0.0000000001	0.00017	0.039

Discussion

The results indicate that the reliability of visibility models depended on the scale (level of detail) of input data. This trend was particularly clear when we calculated how the number of objects modeled as visible differed from the number actually visible (Section 3.1.). In controlling at 150 random points, the visibility models created based on the tested datasets differed with very strong significance.

Therefore, if the purpose is to carry out a GIS viewshed analysis in such a manner as to minimize the difference between the number of objects visible in the model and in reality, then it is possible unequivocally to recommend using the most precise input data possible. Visual impact of WTs provides a good example of when large differences in modeled and actual numbers could be important, because, as demonstrated by Betakova et al. (2015), human perceptions of WTs depend on the number of objects seen. In the cases of some evaluation purposes for which GIS viewsheds are modeled, however, it may be more important to achieve a different type of match between model and reality. In our study, we worked with a scenario wherein the purpose of the analysis was not to minimize the difference in numbers, but rather to achieve the best possible match between the number of visible WTs in the model and in reality (2a), which is to say for the model to predict the correct number of visible objects. In contrast, the second scenario (2b) simulated a situation wherein matching numbers would not matter and the visibility of a single tall structure from the given location would be unacceptable (e.g., for a WT) or sufficient (e.g., for a radio mast). In both scenarios, therefore, match (true) and disagreement (false) between the model and reality were defined as binary. Such evaluation is more forgiving (in the case of Analysis 2b versus 2a) of model imprecision. In certain cases, therefore, the difference between neighboring datasets was not significant (e.g., based on the evaluation used in Analysis 2b, a LiDAR-based model can be replaced with a small-scale vector-based model without losing precision). However, the results still clearly indicate a trend that a more precise input surface model leads to a more reliable visibility model. Moreover, it is possible that significant differences between neighboring datasets would have been achieved by increasing the number of control points (i.e., by boosting the test's power). Therefore, the percentage of cases in which the model matched reality may be more interesting than is the significance of the differences. Tables 3 and 5 indicate that this relative accuracy depended on both the dataset used and the specific definition of a match between the model and reality, although the trend was identical for both match definitions used. When selecting input data for GIS viewshed analysis, therefore, it is necessary to take into account not only the scale but also the purpose of the analysis and the relative accuracy that suffices for the given purpose in landscape planning or other field. This study together with papers by other authors (e.g., Berry et al., 2005; Mark A. Maloy and Dean, 2001) can provide guidance as to the degree of accuracy that can be achieved in a given case. Nevertheless, specific accuracy values can, of course, differ under the effect of such factors as the configuration of the area of interest.

In terms of the distribution and structure of errors of individual models, vector-based models tended to overestimate, to generate false positives (i.e., to predict that an object is visible from more locations than it is in reality). In contrast, the LiDAR-based model predominantly generated false negatives, which is in accordance with the results of (Meek et al., 2013). The occurrence of false negatives (i.e., predicting that a WT is not visible when it is visible in reality) was more or less identical for all models, whereas the number of false positives increased with decreasing input data detail (see Table 7). The dependence of the occurrence of false positives on input precision can be explained by a situation that models based on less-detailed datasets overestimate the extent of the total visible area (see Fig. 2). As suggested by Meek et al. (2013), who reported that a visibility model based on a LiDAR-based DSM originally contained predominantly false negatives but that the opposite situation was true after trees were removed, overestimation of visible area may be caused by inaccurate capture of objects on the ground in coarser-scale data, which causes fewer obstacles to visibility. The same effect may be caused by inaccurate capture of the terrain where the DTM is smother and models only large terrain obstacles. As visible area increases, however, nonvisible area within the study area diminishes and so the independence of the number of false negatives on data accuracy remains surprising in this explanation.

Based on the structure of errors, it can be said that, looking at model accuracy in terms of areas from which a tall structure is not visible, the models display no essential differences. In such analyses, LiDAR-based models can be replaced by vector-based models or a detailed vector-based model by a less-detailed one. Therefore, if for landscape planning purposes we are searching for suitable locations to place a tall structure (e.g., a WT) with the requirement that the structure have the least visibility possible, then it is not a serious mistake to use a large-scale model and place the structure in a location designated as nonvisible. However, we must take into account that using models based on less-detailed data may lead us to overlook potentially suitable locations or not find any suitable locations. In contrast, if a visibility analysis is used that focuses on visible areas based on less precise vector-based models, then visibility is substantially overestimated. This can affect preventive assessment of structure placement in relation to its visibility as well as scenic beauty, where a structure is evaluated as visible from a location with a high aesthetic value and so as having a negative effect even though it would not be visible in reality. Another example of inaccurate modeling of visible areas having an economic impact is the placement of radio masts.

A number of studies mention LiDAR as a theoretically suitable data source for modeling visibility (e.g., Lake et al. 2000, Sander & Manson, 2007) and attention is currently dedicated to quantifying the accuracy of LiDAR-based visibility models (Castr et al., 2015; Murgoitio et al., 2014). In general, we can say that our study confirmed the prediction that LiDAR-based datasets are the most suitable input data for visibility analyses in terms of accuracy and that their accuracy exceeds that of vector-based datasets commonly used in practice. This fact is supported in particular by Analysis 1. If LiDAR data is not available for the study area, it is best to use DSMs created using vector data at the most detailed scale possible. In contrast to vector-based models, the accuracy of LiDAR-based models does not depend primarily on input data scale but rather on the density of the elevation point cloud and the resolution of the DSM created from it (Castro et al., 2015a; Murgoitio et al., 2014b). The provider of the data product used in this study stated a point cloud density greater than 1 point/m^2 and the resolution of the DSM created was 5 m. The precision of our LiDAR-based model (88.7%) matches that of the model created by Berry et al. (2005) with a resolution of 1 m (88.5%). It is probable that if we decreased the pixel size of the DSM we would obtain an even more accurate visibility model, although this would increase computation time demands. The study was conducted in a study area with undulating relief

where elevation ranged between 211 and 723 m a.s.l. It is possible that results could be slightly different in flat areas or in mountains with more dramatic relief. It might logically be assumed that merely slightly undulating landforms will require more detailed data to describe all elevation subtleties, while less detailed datasets could be sufficient for visibility in the mountains. To the best of our knowledge, however, such assumption has not yet been definitively proven and its testing would require systematically selecting a set of sample study areas varying in elevation range from flat land to mountains.

Conclusion

The results of our study confirmed the prediction that the reliability of GIS visibility analyses depends on the input data's level of detail. This dependence was demonstrated through the example of assessing the visibility of tall structures, specifically WTs. Considering the difference between the number of WTs visible from random control points as predicted by GIS visibility models and the number that are visible in reality, the most suitable data input is unequivocally a LiDAR-based DSM. The suitability of visibility models for which the input was surface models created using vector data (contour lines, woody vegetation, and buildings) can be ranked according to input data scale. A similar trend can be observed in the case of the binary evaluation of match and disagreement between modeled visibility and reality, although in certain cases the differences between individual datasets were not unequivocal and depended on how the model's match with reality was specifically defined.

In terms of the reliability of visibility models, none of the input datasets tested differed in the number of recorded false negatives (i.e., cases where the model underestimated WT visibility as compared to reality). Differences consisted in the numbers of false positives (i.e., overestimation of modeled visibility as compared to reality). For both definitions of true and false values, the LiDAR-based model provided the best results. All models based on vector data significantly overestimated visibility compared to the LiDAR-based model and this overestimation was greater for data from less-detailed scales.

In conclusion, we can state that (i) as predicted, more-detailed input data led to more reliable visibility analysis results; (ii) the vector-based models used had more false positives, while the LiDAR-based model had more false negatives; (iii) only the number of false positives depended on input data precision, while the occurrence of false negatives was similar for all datasets used; and (iv) the trends determined are therefore valid also for various definitions of the model's matching of reality. Our conclusions are valid for analyses at a detailed evaluation scale.

Acknowledgements

We acknowledge the anonymous referees for their constructive comments. We owe special thanks to Gale A. Kirking of English Editorial Services for his useful linguistic advice. We thank Jan Komárek for his help with the field work. This study was funded by the Czech University of Life Sciences Prague (grants nos. 20154208 and 20154256).

CHAPTER V
Study 2

Title:
Impact of input data (in)accuracy on overestimation of visible area in digital viewshed models

Authors:

Ondřej Lagner, Tomáš Klouček, Petra Šímová

Author's contribution: 45%

Conceptualization of the study, consultation and co - writing of the original draft, data acquisition, data processing and validation, review and editing

Journal:

PeerJ, **IF** (2017) **2.118**; Multidisciplinary Sciences ranking 19/64 – **Q2** (2017)

Full citation:

Lagner, O., Klouček, T., & Šímová, P. (2018). Impact of input data (in) accuracy on overestimation of visible area in digital viewshed models. PeerJ, 6, e4835.

Cited:

1 times cited on WOS (2019 August)

Abstract

Viewshed analysis is a GIS tool in standard use for more than two decades to perform numerous scientific and practical tasks. The reliability of the resulting viewshed model depends on the computational algorithm and the quality of the input digital surface model (DSM). Although many studies have dealt with improving viewshed algorithms, only a few studies have focused on the effect of the spatial accuracy of input data. Here, we compare simple binary viewshed models based on DSMs having varying levels of detail with viewshed models created using LiDAR DSM. The compared DSMs were calculated as the sums of digital terrain models (DTMs) and layers of forests and buildings with expertly assigned heights. Both elevation data and the visibility obstacle layers were prepared using digital vector maps differing in scale (1:5,000, 1:25,000, and 1:500,000) as well as using a combination of a LiDAR DTM with objects vectorized on an orthophotomap. All analyses were performed for 104 sample locations of 5 km^2, covering areas from lowlands to mountains and including farmlands as well as afforested landscapes. We worked with two observer point heights, the first (1.8 m) simulating observation by a person standing on the ground and the second (80 m) as observation from high structures such as wind turbines, and with five estimates of forest heights (15, 20, 25, 30, and 35 m). At all height estimations, all of the vector-based DSMs used resulted in overestimations of visible areas considerably greater than those from the LiDAR DSM. In comparison to the effect from input data scale, the effect from object height estimation was shown to be secondary.

Keywords

LiDAR, Spatial uncertainty, Digital surface model, Viewshed, Data quality

Introduction

Defining the visibility of objects in the landscape has been important for historical studies (e.g., Ogburn, 2006; Sevenant and Antrop, 2007) and has found application also in a number of areas of current interest, such as seeking locations to place objects potentially harming scenic beauty like photovoltaic power plants (Fernandez-Jimenez et al., 2015), coastal aquaculture sites (Falconer et al., 2013), and ski areas (Geneletti, 2008); placing military structures (Smith and Cochrane, 2011); tagging landscape photographs (Brabyn and Mark, 2011); analyzing the effects of introducing animal species (Kizuka et al., 2014); and modeling predation risk in animal ecology (Alonso et al., 2012; Olsoy et al., 2015). The basic algorithm implemented in most GIS software produces a binary detection of areas that are visible or nonvisible from a point of observation or identifies areas from which a given object in the landscape is or is not visible. Combining several such binary viewsheds created from multiple observation points or from all cells in the raster of the study area creates a cumulative viewshed describing the visual exposure of the study area. As a number of factors may play roles in visibility modeling and using only a binary attribute (0 or 1) constitutes a drastic simplification (Fisher, 1992), other algorithms have been under development for a number of years, such as fuzzy viewshed and visual magnitude (Brent C. Chamberlain and Meitner, 2013; Fernandez-Jimenez et al., 2015; Fisher, 1996, 1995, 1994, 1993, 1992; Ogburn, 2006), as well as visibility indices such as the Vertical Visibility Index (Nutsford et al., 2015), which enrich the model with further parameters and so are used to bring it closer to reality. Due to their simplicity and implementation in common GIS software, however, binary and cumulative viewsheds are still used in a number of studies (Alonso et al., 2012; Falconer et al., 2013; Olsoy et al., 2015; Rosa, 2011; Schirpke et al., 2013).

In addition to the computational algorithm, the reliability of the resulting visibility model also depends on the quality of the input digital surface model (DSM) (Klouček et al., 2015; Lake et al., 2000b; Sander and Manson, 2007a), and Fisher (1992) previously noted that it would be an error to assume the input DSM to be accurate. Although many studies have dealt with improving algorithms, only a few studies have focused on the effect the spatial accuracy of input data has on the reliability of results from visibility analyses, even though, as can been seen in older visibility studies (Fisher, 1992; Huss and Pumar, 1997) and spatial uncertainty research in other areas (for review see Barry and Elith, 2006; Moudrý and Šímová, 2012), it is highly probable that decreased data quality correlates with decreased quality of results. DSMs for visibility analyses are mostly created as combinations of digital terrain models (DTMs) depicting the bare earth surface plus layers of objects on that surface, particularly structures and vegetation. As such layers rarely contain the attribute object height, the height for creating the DSM is estimated based on knowledge of the area or such sources as published works on vegetation in the location, as was done by Schirpke et al. (2013). The accuracy of this estimate represents an additional potential source of DSM inaccuracy beyond the scale of elevation and planimetric data. In extreme cases, objects are entirely omitted from the surface and visibility is modelled based only upon a DTM, even though Dean (1997b) has already demonstrated the logical expectation that using DSM results in higher-quality visibility models.

Examples of rare studies dealing with input data precision have been presented by Lake et al. (2000) and Sander & Manson (2007), who focused upon modeling structures as vertical obstacles to visibility. Some authors have focused on modelling vegetation for visibility analysis, but they did not evaluate the effect of such models' precision on the precision of the visibility model (e.g., Domingo-Santos et al., 2011; Liu et al., 2010). Problems with implementing vegetation and structures into DSMs do not arise when using LiDAR-based surface models, which already contain objects on the surface and are considered by many authors to be currently the most accurate data input for visibility analyses (see Castro et al., 2015b; Lake et al., 2000b; Murgoitio et al., 2013). Using the example of wind turbine visibility and comparing modelled visibility with actual visibility in the field, Klouček et al. (2015) demonstrated that use of a LiDAR-based DSM can result in a ca 90% match rate with reality while use of DSMs based on vector layers of various scales resulted in only 50–80% match rates. Unfortunately, LiDAR-based DSMs cannot yet be used for all study areas due to their high prices, because of the difficulty in processing a point cloud into a raster DSM, and not least for the reason that LiDAR data is not yet available for a number of areas. For these reasons, it is necessary to know how visibility models based on other data differ from LiDAR-based models and whether these differences depend on the quality of the vegetation height estimation and other variables.

While focusing on the immediate vicinity of observer points, the aim of our study was to evaluate the effects that input data accuracy, terrain configuration, number of visual obstacles such as forests and buildings, and the quality of expertly assigned obstacle height (particularly of forests) have on the results of simple binary viewshed analysis.

Matherial and Methods

Sampling locations

We analyzed visibility at 104 sampling locations in the Czech Republic. One location corresponded to a single page of a national map at a scale of 1:5,000 (i.e., a rectangle of 2.5 × 2 km). Selecting locations in this manner provided sufficient areas for visibility analyses at a detailed scale while still enabling acquisition of input data for a sufficient number of locations. The locations (map pages) were selected by stratified random sampling from that section of the Czech Republic which had available DTMs as well as DSMs created from airborne laser scanning data. This section forms a north–south band in the center of the country (Fig. 1) covering elevations ranging from lowlands to mountains (141 to 928 m a.s.l.) and various landscape types from agricultural to forest. Random sampling of locations was stratified so that it would include as equally as possible combinations of variously forested areas (three categories according to the proportion of forest at the location: 0–9%, 10–24%, and 25–60%) and various terrain configurations (three categories according to elevation differences in the area expressed as the standard deviation of elevation in the location: <10, 11–30, >30 m). Another condition was excluding selection of adjacent map pages.

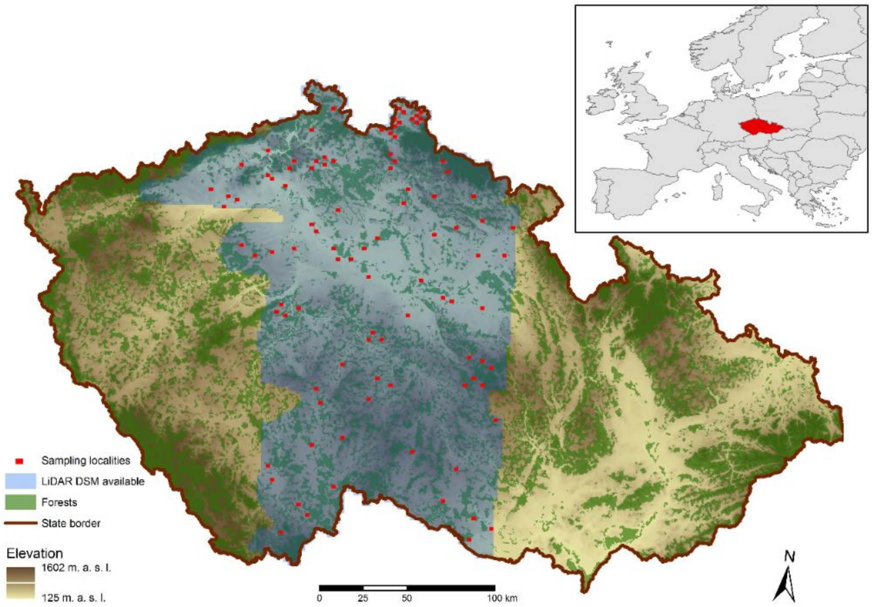

Figure 1. Sampling locations.

Input data and GIS processing

All GIS analyses were conducted using ArcGIS 10.2 software (ESRI, CA, USA). For all viewshed analyses, we used five input DSMs varying in scale and accuracy (see Table 1 for overview). The most accurate was the 1st Generation LiDAR-based DSM of the Czech Republic (hereinafter LiDAR). It was also the only dataset that was available directly as a DSM for the sampling locations. The remaining DSMs were created as sums of rasters comprising the terrain (DTMs) and objects on the terrain (digital object models [DOMs]).

Table 1. Description of input datasets.

Acronym within study	Czech acronym	Scale	Year of last update	Elevation accuracy	Planimetric accuracy	Contour interval	Data description
LiDAR	DMP 1G	Density of elevation point cloud is 1–2 points/m^2	2009–2013	0.4–0.7 m	0.4–0.7 m	No contour	Digital surface model represented by elevation point cloud from data acquired by aerial LiDAR covering part of the Czech Republic
LidOrth	DMR 5G	Density of elevation point cloud is 1–2 points/m^2	2009–2013	0.18–0.3 m	Only elevation dataset	No contour	Digital terrain model represented by elevation point cloud from data acquired by aerial LiDAR covering part of the Czech Republic
	Orthophotomap	Pixel resolution = 0.5 m	2013	Only planimetric dataset	0.25–0.5 m		Orthophotomap covering the entire Czech Republic
Map5	SM 5	1:5,000	2001–2014	0.7–5 m	0.5–1 m	1, 2, or 5 m depending on the character of the terrain	Large-scale vector database covering part of the Czech Republic
Map25	DMU 25	1:25,000	1998	5–10 m	0.5–20 m	5 m	Medium-scale vector database covering the entire Czech Republic
Map500	ArcCR 500	1:500,000	2014	25–50 m	up to 200 m	50 m	Small-scale vector database covering the entire Czech Republic

Working in this manner, one of the inputs combined a LiDAR-based DTM with a vectorization of forests and built-up areas on the actual orthophotomap (hereinafter LidOrth). The remaining DSMs were based on vector topographic maps at scales of 1:5,000 (hereinafter MAP5), 1:25,000 (MAP25), and 1:500,000 (MAP500) (see Table 2 for overview). In all these datasets, elevation was depicted by contour lines and topography by polygons representing the footprints of individual objects on the ground. DTMs were calculated by interpolating contour lines using the topo to raster method. To create DOM rasters, we added estimated values of heights to polygons of visual obstacles and rasterized the layers. Inasmuch as forests were the most important visual obstacles within the locations, we tested five values of forest height (15, 20, 25, 30, and 35 m) to evaluate the effect of the DOMs' height estimates on viewshed results. These values represent a range of mature forest types under various ecological conditions in the Czech Republic. Other woody vegetation types, such as young forests and orchards, were assigned the height of 5 m. We assigned the height of 8 m to buildings and built-up areas as an estimate of the average height of rural structures within the locations.

Table 2. Creation of five digital surface models (DSMs) from input datasets.

DSM		DTM – source elevation data		DOM – source planimetric data
LiDAR	=	elevation point cloud	=	elevation point cloud
LidOrth	=	elevation point cloud	+	vectorization on actual orthophotomap: forest (15–35 m), orchard (5 m), built-up area (8 m)
MAP5	=	MAP5 (contour lines)	+	Map5: forest (15–35 m), orchard (5 m), built-up area (8 m)
MAP25	=	MAP25 (contour lines)	+	Map25: forest (15–35 m), orchard (5 m), built-up area (8 m)
MAP500	=	MAP500 (contour lines)	+	MAP500: forest (15–35 m), built-up area (8 m)

The ArcGIS Viewshed tool, which creates simple binary layers distinguishing between visible and nonvisible areas, was employed for GIS visibility analyses and the process was automatized using a Python script. Within each sampling location, we generated one random point as the location of an observer. We processed a set of viewshed analyses with all of the DSMs and with two heights assigned to the observer point as the OFFSETA parameter within the Viewshed tool. The height of 1.8 m simulated observation of the landscape by a person standing on the ground (*ground* variant). The second variant used the height of 80 m, which can be interpreted as visibility from an observation tower or as visibility from a tall structure such as a wind turbine in the landscape (approximately, disregarding height of the observer; *tower* variant). In this way, we created 2 x 21 viewshed models for each sampling location.

Statistical analysis

We used an R (R Core Team, 2015) script for the nonparametric Friedman's ANOVA with repeated measures design and post-hoc test (available from http://www.r-statistics.com/2010/02/post-hoc-analysis-for-friedmans-test-r-code/) to analyze potential differences among visibilities modeled with different forest heights. The response variable was the amount of visible area as a percent of the location obtained from the viewshed models for each dataset and the forests heights were designed as repeated measures at the same location. The identical procedure was used to analyze differences among visibilities obtained from individual datasets. In this case, the response variable was the percent of visible area modeled with the forest height of 25 m and the datasets were taken as repeated measures at the location. Similarly, we used this design and the Friedman test to test the significance of spatial differences among modelled visibilities. In accordance with previous studies (Castro et al., 2015; Klouček et al., 2015; Lake et al., 2000; Murgoitio et al., 2014), we considered the model based on the LiDAR DSM to be the most accurate (as best matching reality). Hence, the response variable was calculated as the spatial difference between LiDAR visibility and visibility modeled with an individual dataset.

Results

As can be seen in Table 3, using a tower as the observation point or observed object leads, as expected, to larger viewsheds modelled based on each dataset in comparison to the area visible to a ground-level observer, although the trend of differences among datasets is similar for both observer point heights.

Table 3: Sizes of visible area as a percent of the sampling location standard deviation modelled on basis of individual datasets for observation from ground level (ground) and from a height of 80m (tower).

DSM	Level	Forest height 15 m	20 m	25 m	30 m	35 m
LiDAR	*Ground*	------------------------------------- 6.76 ± 6.88 -------------------------------------				
	tower	----------------------------------- 51.40 ± 16.89 -----------------------------------				
LidOrth	*ground*	12.04 ± 10.07	11.75 ± 9.90	11.50 ± 9.77	11.32 ± 9.68	11.29 ± 9.64
	tower	71.46 ± 15.64	69.80 ± 15.80	68.16 ± 16.03	66.56 ± 16.31	65.01 ± 16.66
MAP5	*ground*	16.35 ± 12.85	16.01 ± 12.76	15.72 ± 12.73	15.49 ± 12.71	15.34 ± 12.66
	tower	73.59 ± 15.57	71.92 ± 15.74	70.27 ± 16.02	68.65 ± 16.34	67.06 ± 16.71
MAP25	*ground*	16.18 ± 13.92	15.48 ± 13.71	14.97 ± 13.63	14.56 ± 13.53	14.24 ± 13.41
	tower	72.36 ± 15.94	70.12 ± 16.26	67.95 ± 16.74	65.89 ± 17.25	64.00 ± 17.56
MAP500	*ground*	37.33 ± 22.61	36.74 ± 22.53	36.20 ± 22.47	35.74 ± 22.43	35.36 ± 22.41
	tower	86.37 ± 14.60	85.29 ± 14.97	84.22 ± 15.49	83.18 ± 16.13	82.25 ± 16.65

The smallest average size of visible area in sampling locations came from using the LiDAR DSM, while using all of the remaining datasets led to considerable overestimations in the resulting viewshed (see Fig. 2 for an illustration). For both observer point heights, the viewshed size given by the LiDAR-based model (on average 6.76% of the location when observing from the ground and 51.40% from 80 m) clearly differed from the sizes calculated using the other datasets.

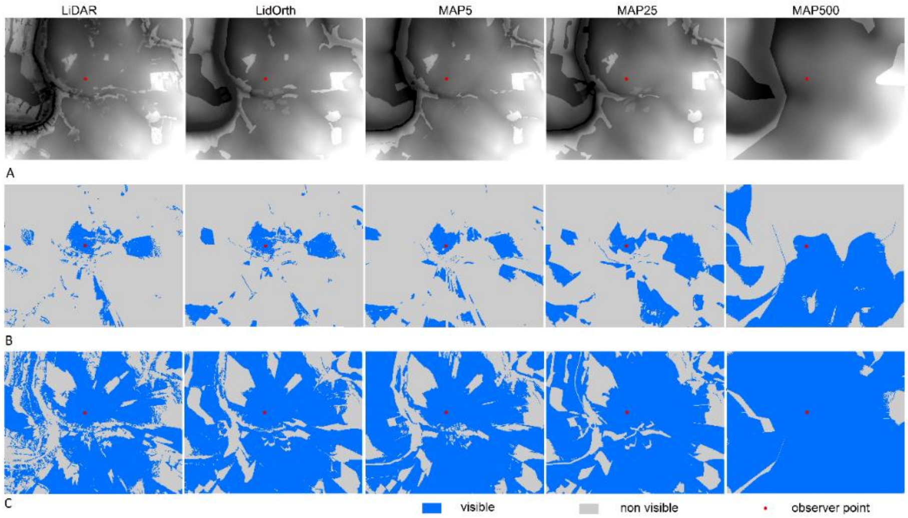

Figure 2. Overestimation of visible area depending on input DSM scale and observer point height_an example of one sampling location (A) Digital surface model. (B) Visibility model_ground variant. (C) Visibility model_tower variant.

For the ground-level variant, the results closest to those of the LiDAR-based model were achieved by the model combining the LiDAR DTM with the vectorized orthophotomap, followed by the models based on MAP25 and MAP5, which had average visibility similar to one another. For the observer point height of 80 m, there were minimal differences among results acquired using LidOrth, MAP5, and MAP25. For both variants, the model based on MAP500 produced the largest viewshed overestimations. For ground-level observation, the LidOrth-based model produced visible areas ca 70% larger than those produced by the LiDAR-based model. The MAP5- and MAP25-based models resulted in visible areas more than twice as large and the MAP500-based model more than five times as large as those produced by the LiDAR-based model. Although the differences in visible area did not come to such large multiples for the tower variant, the visibility modelled based on various datasets differed by more percentage points and the differences therefore concerned a larger proportion of the area. The LidOrth-, MAP5-, and MAP25-based models produced visible areas almost 20 percentage points larger than did the LiDAR-based model, and the MAP500-based model produced visible areas about 30 percentage points larger (Table 3). This simple overview of percentages also corresponds to the results of the Friedman test for models using a forest height of 25 m (Table 4). For both observer point heights, there were no significant differences in visibility modelled based on the MAP5 and MAP25 datasets. In addition, there was no significant difference for the tower variant between the LidOrth-based and MAP25-based models. All remaining differences among datasets were significant, and frequently very highly so (see Table 4).

Table 4. Significance of size differences among visible areas modelled based on individual datasets using a forest height of 25mfor ground and tower variants. Friedman test with repeated measures design and post-hoc test. Significant values are in bold.

	LidOrth		MAP5		MAP25		MAP500	
	ground	*tower*	*ground*	*tower*	*ground*	*tower*	*ground*	*tower*
LiDAR	<0.002	$<1e^{-9}$	$<1e^{-16}$	$<1e^{-16}$	$<1e^{-11}$	$<1e^{-16}$	$<1e^{-16}$	$<1e^{-16}$
LidOrth			$<2e^{-5}$	<0.0002	<0.01	0.051	$<1e^{-16}$	$<1e^{-16}$
MAP5					0.554	0.455	$<1e^{-9}$	$<1e^{-9}$
MAP25							$<5e^{-14}$	$<1e^{-14}$

Spatial differences between visibility based on individual datasets and LiDAR-based visibility were larger than were the differences from numerically subtracting visible areas, although numerical and spatial differences displayed the same trend (the Spearman correlation coefficient for numerical and spatial differences varied between 0.845 and 0.957). In terms of spatial differences, the LidOrth-based model differed in resulting visibility from the LiDAR-based model by 8.05 percentage points in the *ground* variant and by 25.75 percentage points in the *tower* variant. The spatial differences between other models and the LiDAR-based model were significantly greater than was that for the LidOrth-based model (see Table 5 for *p*-values). The differences between the remaining datasets and the LiDAR-based model expressed as percentage points were (*ground*, *tower*): MAP5 12.52, 26.56; MAP25 12.44, 26.62; and MAP500 32.5, 35.29. Similarly as for the analyses focused on total visible area (Table 4), the spatial difference analysis also resulted in no significant differences from the LiDAR-based model for visibilities calculated based on MAP5 and MAP25 (Table 5). Visibility based on MAP500 again very significantly differed from that based on all of the others.

Table 5. Overestimations by individual models when compared to LiDAR results.

Model based on:	LiDAR	LidOrth	MAP5	MAP25	MAP500
Difference Ground (1.8m)	0	8.05	12.52	12.44	32.5
Difference Tower (80m)	0	25.75	26.56	26.62	35.29

In general, it can be concluded from the analysis as to effect of dataset used on resulting visibility that viewshed models calculated using a combination of a LiDAR-based DTM with vectorization on an orthophotomap provide similar results as do models created based on maps at scales 1:5,000 to 1:25,000, although modelled visibility is strongly overestimated in comparison to models based on LiDAR-based DSMs.

Looking at the effect of forest height (Table 3), it is clear that visible area decreases with taller forest height. Other effects of forest height are demonstrated in the LidOrth dataset, representing datasets giving similar visible area values (LidOrth, MAP5, and MAP25), and the MAP500 dataset as the dataset giving the most different results (see Table 6). For the *ground* variant, visible area extent was in most cases significantly different when the forest height was changed by 10 m, while a change of 5 m was sufficient in the *tower* variant. All of the heights produced results significantly different from those of the LiDAR-based model. The significance of all of the differences had a decreasing tendency with coarser scale and tended to be lower for the *ground* variant than for the *tower* variant.

Table 6. Significance of spatial differences among modelled visibilities. The response variable was calculated as the spatial difference between LiDAR visibility and the visibility modeled by an individual dataset. Friedman test with repeated measures design and post-hoc test. Significant values are in bold.

	MAP5		MAP25		MAP500	
	ground	*tower*	*ground*	*tower*	*ground*	*tower*
LidOrth	$<5e^{-8}$	<0.005	$<5e^{-7}$	<0.0001	$<1e^{-16}$	$<1e^{-16}$
MAP5			0.998	0.852	$<5e^{-16}$	$<1e^{-16}$
MAP25					$<1e^{-16}$	$<5e^{-15}$

For the combination of all effects, the difference from the LiDAR-based model was least apparent for the *ground* model with MAP500 as the input dataset and forest height of 35 m ($p = 0.048$). Given the overall overestimation of visibility by all datasets, however, it cannot be stated that the tallest forest height estimate is the most suitable for calculating viewshed. Despite their statistical significance, percentage differences in visible area size caused by changes in forest height were minimal in comparison to those caused by input data accuracy. It can therefore be stated that the effect of data detail on modelled visibility is dominant and that when using surfaces not based on LiDAR object height accuracy has only a secondary effect on the accuracy of the result.

How spatial differences between the visibility modelled with a given dataset and LiDAR-based visibility depended on terrain configuration and number of obstacles cannot be generalized, because individual datasets in combination with the *ground* and *tower* variants produced varying results.

Discussion

This study compares the results of visibility models based on data of various spatial accuracy with models based on a LiDAR-based dataset, the latter of which most closely matches reality according to the field comparison of modelled visibility by Klouček et al. (2015). Based on our results, all of the other models considerably overestimated visibility in comparison to the LiDAR-based model. We had expected that, with the exception of the LidOrth dataset combining a LiDAR-based DTM with objects digitized on an actual orthophotomap, the smallest difference would appear in visibility modelled on basis of the most detailed vector data (i.e., MAP5). Surprisingly, however, MAP5 provided similar results as did MAP25, whether working with numerical or spatial differences in visible areas. In addition to the fact that MAP5 is at a more detailed scale than is MAP25, MAP5 is the only tested dataset that depicts individual buildings and not just outlines of built-up areas. In accordance with the results of Sander and Manson (2007), who stated that generalizing building locations has a significant effect on the resulting viewshed model and that this effect is more important than is that from imprecise building height determination, we predicted that MAP5 would produce a more precise viewshed model. However, Sander and Manson (2007) analyzed visibility in cities and our results indicate that in locations in the countryside, where buildings occur to a lesser extent and are predominantly part of smaller municipalities as in our study, then generalizing buildings does not have a significant effect on visibility modelling results. For all variants evaluated, visibility modelled on the basis of MAP500 differed the most from the other visibility models. This result corresponds to the low reliability of visibility models based on this dataset (48.1–63.9%) found by Klouček et al. (2015). Data generalized to such an extent as is found in maps at a scale of 1:500,000 therefore cannot be used at such a detailed scale (areas of 5 km^2) for modelling visibility, not even when objects on the surface are included from the planimetric layers of such a map.

In relation to the findings of this study and those of Klouček et al. (2015), it can be difficult to understand the results of studies that do not describe in detail the input data used to model visibility. This is a problem for certain applied studies that do not have as their primary objective to study the effect of geodata on the results. For example, Geneletti (2008) modelled the visibility of ski areas in a range of 5 km, which means within the zone of greatest visual effect (e.g., Betakova et al.), based on a DSM, but that author did not state how and from what data the DSM was assembled. Etherington and Alexander (2008) stated the scale of the digital elevation model (1:20,000) and the resolution of the raster (30 m) used for their viewshed model, but it is not clear whether this raster included vegetation. Given that the scale of data used has a dominant effect on visibility results, all future studies should describe the input data so that the applicability of the study results can be evaluated.

According to our results, therefore, it cannot be stated unequivocally that the rate of spatial overestimation by datasets would be, say, higher in flat or mountainous terrain or in areas that are more or less forested. Our work considered obstacles to visibility to be opaque. This is not necessarily the case, however, and particularly not in the case of forest stands. Therefore, searches are underway for techniques to model forests more realistically than as solid polygons with uniform tree height (Domingo-Santos et al., 2011; Liu et al., 2010). Such forest models work with individual trees and thereby take into account both stand density and set crown height, with stands having crown height set higher being more transparent. Our results indicate, however, that at the given evaluation scale (locations of 5 km^2) such labor-intensive modelling of stands is not significant for the results, as the effect of input data scale is dominant. This can be seen in the fact that all of the datasets used produced overestimations in comparison to the LiDAR-based model. Making forest stands transparent would result in a higher percentage of visible area at a given location (i.e., even greater overestimations) and thus increasing the accuracy of obstacle models would paradoxically further add to viewshed model inaccuracy.

The LiDAR-based DSMs used in this study originate from nationwide imaging which did not have as its primary objective to create DSMs in non-built-up areas.

The fact that the imaging took place also outside of the growing season can, together with the low point cloud density, lead to inaccuracy in the DSMs, particularly in places with broadleaf vegetation. It is therefore possible that use of more detailed LiDAR captured during the growing season would reveal even greater spatial overestimation of visibility by all tested datasets.

Conclusion

This comparison of visibilities modelled using the LiDAR-based DSM and DSMs based on vector datasets or on a combination of the LiDAR DTM and an orthophotomap indicates that all of the other models considerably overestimated visibility in comparison to the LiDAR-based model. The overestimation rate was greater in absolute numbers with a higher observer point, although trends in overestimations were identical in models simulating observation from the ground and those simulating observation from a tower. In both cases, it can be stated that none of the other datasets with any set height for obstacles to visibility approached the accuracy of the LiDAR-based visibility model and that the established obstacle height had a minor effect on resulting visibility in comparison to the effect of the dataset.

Acknowledgements

We acknowledge the anonymous referees for their constructive comments. We owe special thanks to Gale A. Kirking of English Editorial Services for his useful linguistic advice.

CHAPTER VI
Study 3

Title:
The significance of using raw data: a case study with canopy height models of shrubs

Authors:

Ondřej Lagner, Tomáš Klouček, Michal Fogl

Author's contribution: 60%

Field work and data acquisition, conceptualization of the study, consultation and writing of the original draft, data acquisition, data processing and validation, review and editing

Journal:

SGEM2019 Conference Proceedings, **IF** (2016) **0, 45**; SJR (2017) 0,211

Full citation:

Lagner, O., Klouček, T., Fogl, M. (2019). The significance of using raw data: a case study with canopy height models of shrubs. 19th International Multidisciplinary Scientific GeoConference SGEM 2019, SGEM2019 Conference Proceedings, ISBN 978-619-7408-79-9 , ISSN 1314-2704, 30 June - 6 July, 2019, Vol. 19, Issue 2.1, 1089-1098.

Abstract

The quality of spatial data plays a crucial role in environmental modelling and management, especially in local scale studies needing a detail mapping of vegetation elements in a mosaic, near-natural landscape. One of the sources of spatial data for such modelling is airborne LiDAR. Although LiDAR-based vegetation and terrain models are often considered accurate, their quality is dependent on the density of the original raw point clouds and the computation algorithm. The aim of this study was to answer a question of how the method of LiDAR raw data processing affects the accuracy of the resulting canopy height models of shrubs in the mosaic landscape consisting of herbaceous plants and shrub formations. We hypothesize that using raw LiDAR data in conjunction with a suitable algorithm, we can obtain a more accurate shrub model than that acquired from the same raw LiDAR data through a general all-purpose processing used for computation of nationwide digital surface models. The comparison of vertical accuracy of individual models with reference field data showed that combining raw LiDAR data with an algorithm suitable for the studied area could lead to creating better shrub vegetation models than those available from the governmental products. Besides, our results also imply that even data with relatively low point cloud density that are not primarily intended for creating digital models of vegetation can yield a good canopy height model eligible for shrub detection if processed in a suitable way.

Keywords

Raw data, LiDAR, Digital Surface Model, Canopy Height Model, LAS data

Introduction

The quality of spatial data used for modelling usually plays a crucial role in studies focused on environmental characteristics and/or relationships among individual components of the environment (Li et al., 2012; Moudrý and Šímová, 2012). The need for detailed spatial data correctly reflecting the reality increases with smaller scale of the research – for example, for local scale modelling aiming to capture small elements in a scattered vegetation landscape. One of the sources of spatial data for environmental modelling in such scale is airborne LiDAR (Wu et al., 2006), the availability of which is growing. LiDAR-based digital surface models (DSMs), i.e., detailed landscape models including buildings and vegetation, are used e.g. in the studies on modelling the visibility (Klouček et al., 2015) or on the solar potential assessment on the local scale, taking into account shadows cast on the roofs by surrounding vegetation (Fogl and Moudrý, 2016). LiDAR-based DSMs capturing the horizontal and vertical structure of vegetation (also called canopy height models, CHMs) have been for a long time used as an important source of data for forest ecology, forestry management, biomass estimation, estimates of carbon uptake or as a source of explanatory variables in animal ecology (Melin et al., 2016).

While many studies have discussed the issue of the quality of CHMs in the forest environment, only a few studies have focused on the use of LiDAR for detection of shrub vegetation in an open landscape. The shrub vegetation however plays an important role in the environment, constituting an important habitat for many animals, and the knowledge of its structure can be crucial for land management. Most studies on this topic published so far concentrated on semi-arid areas. Other studies were focused on accuracy of techniques for estimating shrub height in a similar environment (Glenn et al., 2011) or on determination of the shrub biomass volume (Estornell et al., 2011). More recent studies use LiDAR (in conjunction with spectral remote sensing) for example for mapping of invasive shrub species in the urban environment (Chance et al., 2016) or for mapping of shrub habitats in the arctic tundra (Boelman et al., 2016). Creating CHMs for the open landscape of the temperate zones is however a challenge that has not been sufficiently explored so far.

The aim of this study is to answer a question of how the method of LiDAR raw data processing affects the accuracy of the resulting CHM of shrubs in the mosaic landscape consisting of herbaceous plants and shrub formations. In the area of interest, the real (manually measured) height of shrubs is compared to (i) LiDAR-based models available for the entire area of the Czech Republic created through general all-purpose processing without any special attention to recording shrubs and (ii) a model created from the same input LiDAR raw data designed to capture shrub vegetation at the studied site. We hypothesize that using raw LiDAR data in conjunction with a suitable algorithm, we can obtain a more accurate shrub model than that acquired from the same raw LiDAR data through a general all-purpose processing used for computation of a nationwide CHM.

Matherial and Methods

Study area

Study area (20 km²) is located in the west of the Czech Republic (Central Europe) in Doupovske hory (50°18' N, 13°8' E; Figure 1). It is a part of a military area used for NATO military exercises. Human activities (other than military) in this area are strictly limited to forestry and game management and, to a lesser extent, ecological and forestry research. Within the territory of the Czech Republic, this area is specific in its character, consisting of a mosaic of herbaceous and shrub vegetation with some remnants of forest vegetation. The predominant shrub species are slow-growing shrubs (approx. 20cm per year) such as the blackthorn *(Prunus spinosa)*, hawthorn *(Crataegus oxycantha)* and dog rose *(Rosa canina)*.

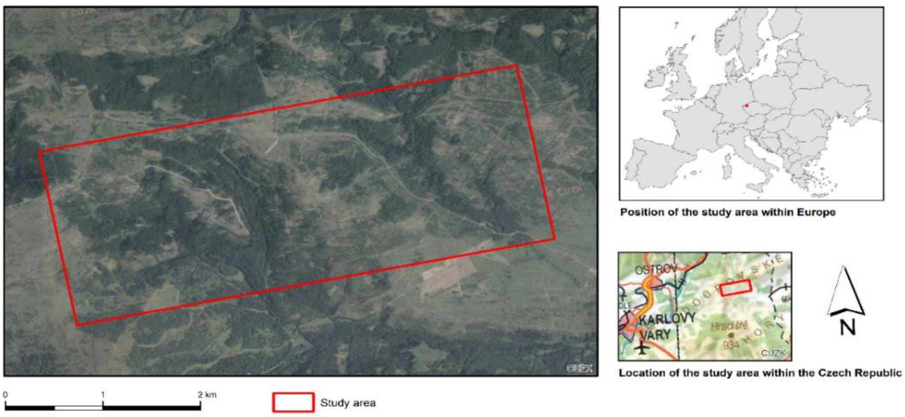

Figure 1. The area of interest (20 km2) is located in the western part of the Czech Republic, in Doupovske hory (Hradiste military area).

Field data collection

The aim of the field data collection was to obtain reference data on the real shrub vegetation height in the area of interest. The measurements took place in August 2015 at 78 shrub locations, which were randomly selected from an orthophoto of the studied area. A portable GPS receiver (Oregon 450t, Garmin) was used to navigate to the selected points. The shrub height was measured using an ultrasound tree height meter (Vertex IV).

LiDAR data

The airborne LiDAR data was acquired for the entire area of the Czech Republic by the State Administration of Land Surveying and Cadastre (CUZK) using LiteMapper 6800 (IGI mbH) system combined with a RIEGL LMS-Q680 scanner within a framework of a national project undertaken between 2009 – 2013. The area of interest was scanned in March 2011. This data exists in three formats: (1) raw LiDAR - original raw LiDAR dataset in LAS format, (2) TIN XYZ - discrete points dataset represented by a triangle irregular network (TIN) with 3D coordinates in XYZ ASCII format, (3) DEM raster - freely available LiDAR derived digital elevation models in raster format (see Table 1 for details). TIN XYZ and DEM rasters are provided free of charge for non-commercial purposes while the raw LiDAR data can only be procured on request for scientific purposes.

Table 1. Description of input datasets.

Acronym used throughout study	Spatial resolution	Data format	Vertical accuracy	Availability	Data description
Raw LiDAR	minimum 1 - 2 points/m^2	LAS	not specified	scientific	Dataset represented by raw elevation point cloud from data acquired for the entire Czech Republic
TIN XYZ	1 - 2 point/m2	XYZ ASCII	terrain accuracy (0.18 - 0.30 m) object accuracy (0.40 - 0.70 m)	commercial/ anybody	Dataset represented by generalized elevation point cloud created from raw airborne LiDAR data (processed by CUZK), covering the entire area of the Czech Republic
DEM raster	2 m	raster	terrain accuracy (0.18 - 0.30 m) object accuracy (0.40 - 0.70 m)	Free for anybody	Datasets represented by digital elevation models created from raw airborne LiDAR data (processed by CUZK), covering the entire area of the Czech Republic

CUZK = State Administration of Land Surveying and Cadastre

LiDAR data processing

From each of the three available datasets (raw LiDAR, TIN XYZ, DEM raster), a canopy height model (CHM) was created: (1) Raw LiDAR data were processed in a specialized LiDAR software LAStools. The points in the cloud were classified as either ground or vegetation and subsequently, the entire point cloud was normalized (the altitude values were replaced with relative elevation above he terrain). After that, the canopy height models were generated using the las2dem tool. For the best results, *spike-free* algorithm [11] integrated into las2dem tool was used to generate the CHM. (2) Data for TIN XYZ were processed in a similar way. To achieve a correct classification of terrain points and vegetation points, two available datasets (digital terrain model and digital surface model) were merged into a single point cloud. The resulting point cloud was processed in the way described for raw LiDAR data, yielding a second CHM. Finally, (3) the last dataset (DEM raster) was processed in ArcGIS (version 10.4) and the CHM was obtained by subtracting DTM raster from DSM (See Figure 2 for the full data processing scheme). In total, three canopy height models were therefore obtained: (a) Raw LiDAR CHM, (b) TIN XYZ CHM, (c) DEM raster CHM, see Figure 3.

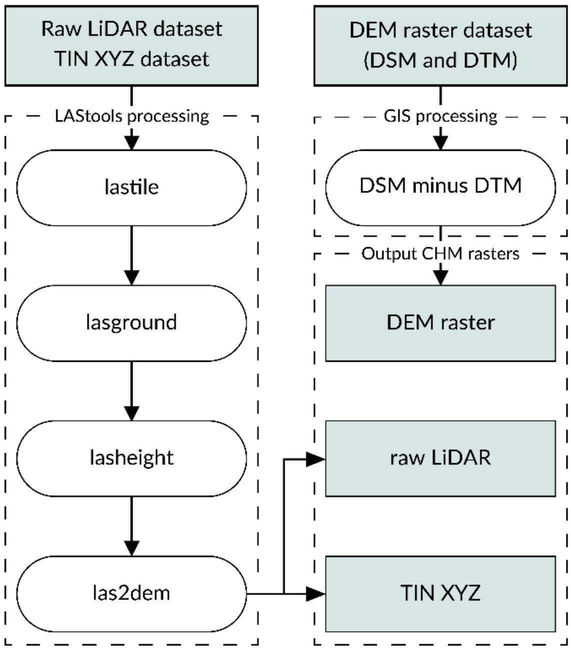

Figure 2. Diagram showing the preparation of individual CHM models.

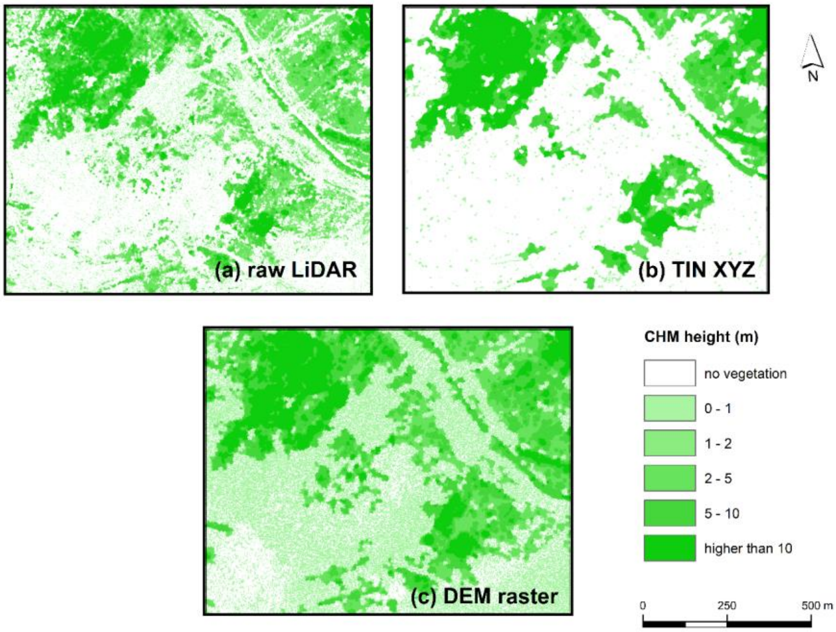

Figure 3. CHM model examples: (a) Raw LiDAR CHM; (b) TIN XYZ CHM and (c) DEM raster.

Statistical analysis

Thus acquired CHMs were compared to the results measured manually during the field data collection. For each reference point, the maximum height within a 2m buffer space in the CHM was recorded, which ensured the capture of the highest point of the shrub even if the exact position of reference point was not on the top of the shrub. The accuracy of the models was then evaluated through Root Mean Square Error (RMSE). The differences between models were tested by Friedman rank sum test and subsequent paired comparisons using Conover's test for a two-way balanced complete block design. All statistical analyses have been performed in R software (R Development Core Team., 2017).

Results

Friedman rank sum test revealed that all models differed both mutually ($P < 0.001$) and from the real field measurements ($P < 0.001$). A comparison of the vertical accuracy of CHM with the real vegetation heights showed that the root mean square error (RMSE) for Raw LiDAR model was 1.59m. The results for DEM raster (RMSE 2.50m) and TIN XYZ (2.60m) were inferior to the Raw LiDAR dataset.

Table 2. Basic statistical parameters of models: Mean difference between field measurements and model-derived values, and root mean square error in differences between measured and modelled values.

	Raw LiDAR	DEM raster	TIN XYZ
Mean difference (m)	-1.18	-1.95	-2.28
RMSE (m)	1.59	2.50	2.60

The basic descriptive characteristics mentioned above show an obvious trend of underestimating the vegetation height when compared to the real life values (Figure 4). The Raw LiDAR model evaluation was the closest to the real data (mean difference -1.18m) while the TIN XYZ was the furthest from the real data (mean difference -2.28 m), see Table 2. A higher system error was therefore recorded for the models based on derived datasets than the model based on raw data.

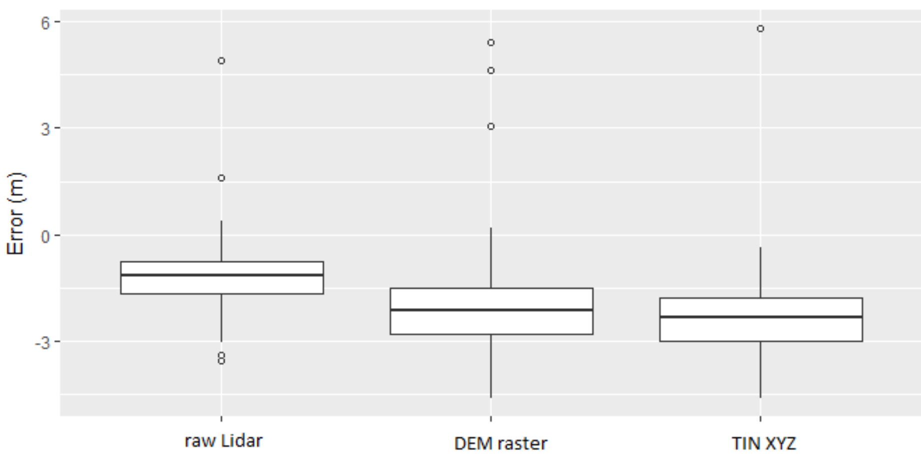

Figure 4. A box plot of differences between results obtained from individual canopy height models and true heights of shrubs (median, quartiles, range without outliers and outliers defined as more than 2/3 times of upper quartile).

This trend was confirmed through a more detailed analysis of the elevation errors. Raw LiDAR model overestimated the shrub height in four cases only, DEM raster in five and TIN XYZ derived model just once. Raw LiDAR model also correctly detected the shrubs (predicted a height greater than zero) for 72 shrubs, DEM raster for 63 and TIN XYZ for 4 shrubs only. However, only 67 % (48/72) of shrubs were detected to be over 0.5m tall, which dropped to 10 % (6/63) in case of DEM derived model and 25 % (1/4) in TIN XYZ, see Table 3.

Table 3. CHM deviations from reality, focused on vertical height overestimation and shrub detection.

	Raw LiDAR	DEM raster	TIN XYZ
Height overestimation	4/78 (5.1%)	5/78 (6.4%)	1/78 (1.2%)
Shrub detection	72/78 (92.3%)	63/78 (80.7%)	4/78 (5.1%)

Discussion

The comparison of the vertical accuracies of individual canopy height models confirms the original hypothesis that models based on raw data are more accurate than those based on pre-processed datasets. The most precise CHM with the lowest RMSE values was the model created from raw LiDAR data. The accuracies of CHMs derived from pre-processed DEM Raster and TIN XYZ data were substantially lower as the original method of processing led to a partial loss of spatial and vertical information, in particular the part capturing low vegetation. Such a loss of information was probably caused by using a general unified method applied to the entire area of the Czech Republic within the framework of a state-funded project, which is, according to our expectations, not the most suitable for such a specific location.

The comparison of CHMs and real life heights of the shrub vegetation clearly shows a trend of underestimating the tree heights in the models. This is undoubtedly associated with the 4 year difference between the LiDAR and field data collection (2011 and 2015). Taking into account the shrub growth, we could be mislead to the conclusion that the model showing the greatest heights in 2015 (closest to the 2015 reality) could be the one most overestimating the 2011 heights, and therefore the least accurate. It is however unlikely to be the case. It has been repeatedly shown that LiDAR-based CHMs have a general tendency to underestimate the real heights (Wasser et al., 2013b), and it is therefore unlikely that the 2011 values would be systematically overestimated. The reason for such tendency to underestimate the heights is logical – the LiDAR beam can miss the tallest part of the shrub and penetrate deeper before finding an obstacle, especially where the density of the raw point cloud is as low as in our study (1 - 2 points/m^2) and where the LiDAR data collection was performed in the leafless period of the year. Under the conditions of the study area, the recorded shrub species count among the slow growing ones, with average annual growth of approximately 20cm. Hence, we can conclude that, despite the several years difference between the aerial recording and field measurement, the model closest to the 2015 reality was at the same time the closest to the 2011 reality and, therefore, that the best CHM was derived from the raw LiDAR data.

The comparison of raw LiDAR and TIN XYZ suggests that unlike the CHM created from derived data, the raw data model utilizes the full potential of all available LiDAR reflections. The processing for all-purpose datasets can lead to both the generalization of the input data and to dilution of the original points below a level required for a study of the vegetation cover. Jakubowski et al. (2013), report such a threshold value to be 1 point per square meter in their study dealing with the issue of point cloud filtering.

The CHM derived from the DEM raster, which was created by subtracting the digital terrain raster from digital surface raster, is burdened with a high amount of noise mimicking low vegetation. This model appears at the first glance to be more accurate than the model created from a filtered cloud point (TIN XYZ) and the two models are statistically significantly different. However, after a closer look at the individual errors, it is apparent that the errors are in the magnitude of centimetres and the registered difference therefore has no real impact on the practical usability of the models; it can be safely stated that both models derived from the pre-processed data are of equal quality. For creation of raw LiDAR and TIN XYZ based CHMs, the *spike-free* algorithm (Khosravipour et al., 2016) was used; this algorithm is one of the few methods described in the literature calculating vegetation models directly from the point cloud instead of its translation into individual vertical raster layers and their summation.

The algorithm therefore allows to utilize the maximum amount of information present in the LiDAR data. The LiDAR data used in the study was not primarily intended for the creation of DSMs but of DTMs. Hence, the LiDAR scanning was performed in winter (March), which, being a leafless period, potentially reduces its value for creating CHMs while, on the other hand, facilitating a better terrain detection thus allowing a more precise DTM calculation. Performing two scans – one in winter and another in summer – would be of course optimal from the accuracy point of view but it would be more time consuming and problematic from the economic point of view. Our results imply that using a suitable processing method can allow us to use even suboptimal raw data for tasks as demanding and sensitive as creating CHMs of shrub vegetation. The raw data from this state funded project is unfortunately not freely available and its procurement is restricted by numerous conditions (scientific purposes only, limited area, etc.). Our results thus indirectly support the conclusions of Turner et al. (2015) that free availability of environmental raw data would constitute a major contribution to the environmental research and management.

Conclusions

Study results show that combining raw LiDAR data with an algorithm suitable for the studied area can lead to creating better shrub vegetation models than those freely available from the governmental products. The uniform processing of such raw data used for creating these national datasets leads to loss of information, especially where low vegetation such as shrubs is concerned. Besides, our results also imply that even data with relatively low point cloud density that are not primarily intended for creating digital models of vegetation can yield a good canopy height model eligible for shrub detection if processed in a suitable way. Such models may somewhat underestimate the real height of the shrub vegetation but the error is sufficiently low. Study results also highlight the importance of using raw geodata, which is applicable (besides environmental science) in multiple fields employing LiDAR data for spatial analyses.

Acknowledgements

We acknowledge the anonymous referees for their constructive comments. We thank Petr Klápště and Jiří Prošek for their help with the field work and Karel Hrach for statistical advice. This study was funded by the Czech University of Life Sciences Prague (grants nos. 20154208 and 42300131242315**6).**

CHAPTER VII
Discussion and Conclusion

Study 1 compared the modelled situation with reality. The field validation was performed on 150 control points randomly generated before the study. Besides proving that more accurate input data lead to more accurate results, we were the first to introduce the terms false positives and false negatives into visibility modelling. False positives mean errors in the model when the observed point can be seen from the control point in the model while this is not true in reality. By analogy, false negatives mean the opposite, i. e. the situation when according to the model, the observed point cannot be seen while in reality, it can. In principle, false positive appears at the first sight to be a less damaging mistake for a user utilizing a visibility analysis e.g. for evaluation of the impact of the construction on the landscape character. The evaluated construction would be in reality visible from a smaller area than what has been modelled and the reality will therefore look "better" than predicted. In reality, however, there is a major downside to this from the investor's point of view. If the investor needs to submit an environmental impact assessment (EIA), the false positive impact damages his cause. If the model predicts the visibility from a larger area than the reality, it may lead to rejection of the project by the administration. It is also unfortunate for the administration that rejected the project on the basis of an incorrect analysis as it can lead to lengthy appeals or even legal proceedings. This type of error was reduced when using more accurate data in our study. The false negative error is however also damaging for visibility analysis. In this type of error, however, we have not detected a significant difference in the error rate when using input datasets with different accuracies.

We can thus conclude that better accuracy of input data led to an overall improvement in accuracy of the visibility analyses; this improvement was however predominantly in the false positive type of error while false negatives were not improved. The probable explanation lies in the fact that more accurate and detailed data capture the objects on the ground that may represent obstacles to visibility better than coarser data. Coarser data have a tendency to "flatten" the surface and, hence, to erroneously remove obstacles present in the real landscape. A similar effect may be observed when using data with lower accuracy of the digital terrain model as such data would only capture more pronounced features than more accurate data. The fact that the false negative error does not depend on the accuracy of the input data is quite interesting from this point of view.

The study 2 was not a case study; as it was performed after the first one, it built on data from the first paper. The model based on LiDAR data proved to be better than any models based on vector DTM with assigned expert heights in this study. While the first study was aimed at validation of the predicted visibility of a particular object (wind turbines), the second one was more focused on the differences among models with various input datasets in various terrain types (differing in ruggedness) including lowlands, hills and mountains. Areas with different representations of forest were selected for analysis – mapping sheets that included only minimum of forests as well as those that were almost completely forested were represented among the study areas. In addition, the tree heights in the areas of study were different, which also affected the results. While in LiDAR data, the tree height is clearly specified by the beam reflection, vector databases do not contain the vegetation height. We have assigned multiple heights to the vegetation and analysed which one would yield results closest to the LiDAR data. In addition, we have calculated the visibility models from two points above the terrain – observer height (human height) and a lookout tower height. The results again confirmed the trend of overestimating the visible area when using less accurate data. We can thus say that LiDAR data have a better information value for visibility modelling than data from topographic mapping. It is necessary to say that vector data of various scales that have however a similar degree of generalization appropriate to the size of the area of interest provide similar results. If the scale is however not sufficient to show details and is strongly generalized, the results of visibility analysis are only indicative.

The Studies 1 and 2 demonstrated that LiDAR data are the most suitable for visibility analyses. In addition, the availability of LiDAR data is growing. Many countries have released their national data for scientific purposes as well as for the use by general public. The provided data can however often be pre-processed in various ways, generalized and such processing may have a negative impact on the results of the analysis. Thus, the Study 3 evaluated the accuracy of CHMs derived from LiDAR data (that can subsequently serve as inputs for visibility analyses) in relation to the processing methods. A canopy height model was chosen as an example. The results imply that the best fit of the model elevations to reference field measured elevations can be achieved when using raw data and processing them in view of the purposes of the study. However, surprisingly, we also found that even partially generalized and processed data (in an XYZ format) that were originally created for coarse scale digital terrain models had a good agreement with reality and could be therefore in effect successfully used for very detailed analyses. The poorest results when compared with reality were obtained from the ready-made final product, which lost a lot of information contained in the raw data.

The accuracy of input data

Results of Studies 1 and 2 demonstrate that the accuracy of resulting models is always dependent on the accuracy of input data. If unsuitable data are used for analysis, the results may be insufficient for the purposes of applied ecology. Ideally, we should use raw, unprocessed data as this is the only way to make sure that the data was not affected by processing flaws. If we work with data pre-processed by another party, so-called spatial data uncertainty enters the equation. This term was mentioned by Moudrý a Šímová (2012) and is associated with the suitability of data for a particular purpose. The data selection must always correspond with the purpose of the analysis (Lecours et al., 2015a, 2015b).

Data uncertainty is closely related to data accuracy. It can be affected by the type and spatial resolution of the used sensor, temporal resolution, spectral resolution, etc. All these variables can become a source of a potential spatial error and if such errors occur, they make the data less suitable or less accurate for the purposes of landscape or visibility analyses. When planning any research, it is thus crucial to consider all possible sources of potential errors and to try to eliminate them (Podobnikar, 2009).

Results of the Studies 1, 2, 3 are in accordance with the above mentioned findings. The dependence of the reliability of visibility analyses on the quality of input data has been demonstrated on the example of visibility assessment of wind turbines in the Study 1 and partially in the Study 2. Based on our results, the best model for visibility analyses is undoubtedly the LiDAR-based digital surface model. This is in accordance with studies by Cramer et al. (2018); Hopkinson et al. (2005) Melin et al. (2016). If LiDAR data are for some reason not available, a digital relief model based on interpolation of isohypses supplemented with surface objects of assigned height can be also used for visibility analyses but the level of accuracy will depend on the level of detail of data collection (cartographic scale of the original vector map). Studies 1 and 2 as well as a study by Mark A Maloy and Dean (2001) confirm that there is a surprisingly low level of difference in the quality of analyses based on scales of 1:5,000 and 1:25,000. On the other hand, the difference of results obtained from those two scales and that of 1:500,000 is enormous (which was however expected considering that the areas of the individual analysed units were in the order of kilometres). If calculating an analysis for the entire area of Europe, even data at a scale of 1:500,000 could be usable but for detailed studies (e.g. for local landscape planning), such data are indicative only.

An interesting fact is the overestimation of the results of visibility analyses when using less accurate data. This is implied by the fact that LiDAR data exactly depict every tree, every single obstacle in visibility while when using vector data, forests and buildings are only roughly modelled from polygons with assigned heights. As the overestimation results predominantly from the false positive error, the model incorrectly shows too many locations from which the evaluated objects are visible. The opposite error, i.e., false negative, is almost equally represented in all datasets. If drawing conclusions for practical visibility assessment, the above findings mean that less accurate data provide "more strict" results – the false negative error (i.e. failure to correctly say that the construction will be visible) is similar regardless of data accuracy and worse data accuracy leads predominantly to the increase of the area from which the model predicts visibility, although in reality, it would not be visible.

Airborne laser scanning is at present probably the most accurate data source for fine scale studies where detailed information about elevation is needed. Airborne LiDAR offers a suitable compromise between requirements on data processing, accuracy and size of the area of interest (Alonso et al., 2012). From the perspective of applied ecology, LiDAR data represent a valuable source due to the broad scale of their possible use. Their advantages include high vertical and horizontal accuracy allowing the acquisition of both the digital surface and digital terrain model from the same mission. It is also capable of providing information about the vertical structure of vegetation, which can be very important e.g. in species distribution modelling (Murgoitio et al., 2014a). LiDAR data can be acquired both at the nationwide level and, if need be, as very fine data for local studies (Koska et al., 2017).

It can be generally said about spatial data originating from a point cloud (most LiDAR data) that the denser the point cloud, the more accurate the result of any analysis using it as input data (Anderson et al., 2006; Xiaoye Liu, 2008). On the other hand, the Study 3 partially disproves that opinion as it showed that even low density LiDAR data can be successfully used for modelling of shrubs on a small scale if proper processing methods and algorithms are utilized.

Raw data utilized in the Study 3 originated from the nationwide scanning with a minimum density of 1-2 points per square meter. Unlike many end users of the products, however, the authors of this study were aware of the low density and the study aimed to determine the success rate of shrub identification. It can be often observed that authors use available data for their research without critically evaluating the suitability of the data.

All the above leads to the same conclusions. Before using any data for analysis, the researcher must critically consider the properties of the data along with the purpose of the analysis, consider the extent of the area of interest and what is the purpose of processing the data. Data accuracy is a relative term and it always depends on the requirements of the particular application. While a certain accuracy can be sufficient for one analysis, it can be insufficient for another purpose and unnecessarily good for yet another.

The effect of the terrain and the surface

Many studies have proved that objects on the surface have a significant effect on the result of the visibility analysis (Lake et al., 2000b; Lecours et al., 2015b; Sander and Manson, 2007a). This was one of the reasons why we have focused on this topic, especially in the Study 2, in more detail. We were surprised to find that the terrain ruggedness and the effect of surface objects has a smaller effect on the accuracy of visibility analyses than the quality of the spatial data. It however again depends on the particular conditions of an individual study. In the study by Sander and Manson (2007), the results of which we have partially built upon, performed their visibility analysis in a highly urbanized area. In our study, however, we performed analysis in a rural landscape where the buildings are much more scarce and smaller. Our studies have demonstrated that in such a landscape, generalization of the heights of buildings does not have a significant effect on the outcomes of visibility analyses.

Similarly, it cannot be simply said that the rate of error would be higher in a flat or mountainous landscape, in forested areas or those without forests. For explanation of this phenomenon, it is necessary to again mention the principal difference between LiDAR and vector data. LiDAR-based surface models provide information about individual trees that may act as obstacles while in the case of polygon vector layers, we perceive any tall vegetation as a homogenous, completely opaque obstacle with strictly defined shapes. In LiDAR data, a partial transparency of such trees is possible. This is one of the reasons while researchers are looking for new methods allowing modelling of individual trees in a way corresponding to reality as much as possible (Domingo-Santos et al., 2011; Liu et al., 2010). In the results of visibility analyses, the opacity of such obstacles when using polygon vector layers leads to the aforementioned overestimation of the visibility as false positives.

The assigned forest height affects the outcomes of visibility analyses. It is not surprising that with increasing forest height, visible areas are growing smaller. It however again depends on the coarseness of data and rate of generalization. When using data of the coarse scale, the forest height has practically not affected the results while when using the finest scale, even the difference of five meters played a role. The best agreement with the LiDAR data was registered where the assigned tree height was 35 metres. As the vector data however generally tend to overestimate the visibility, it cannot be simply said that assigning a high height to the obstacles makes the data more realistic; while the overall visibility is the closest to LiDAR data when assigning the maximum heights to the obstacles, the "distribution of visibility", i.e., the exact locations visible according to LiDAR data and according to vector data with assigned vegetation heights, will likely differ.

The Study 2 originally also evaluated the effect of the distance of the obstacle on the result of visibility analysis. This assessment was not included in the final paper due to its complexity and length of the paper. It was not possible to generalize the relationship between the visibility modelled with a given dataset/LiDAR-based visibility and terrain configuration/number and distance of obstacles because individual datasets produced varying results. The only common features were that no effect of the distance to obstacle on the accuracy was found in any of the generalized linear models and that in the instances where the effect of the number of obstacles or terrain configuration was significant, this explained a relatively low proportion of variability in the response variable. As the line-of-sight principle indicates, the distance of obstacles to visibility from the observer point is essential to the resulting visibility, particularly if this point is low above the terrain (e.g. Nagy, 1994). In our study, the distance to the nearest obstacle did however not have a significant effect on how much other modelled visibilities spatially differed from the LiDAR-based model. This result does not negate the fact that closer obstacles have a greater effect on visibility. The finding merely demonstrates that the distance to the nearest obstacle has no effect on the accuracy of the visibility model. Some previous studies (Kim et al., 2004a) have indicated a possible effect of terrain configuration and number of obstacles on the resulting visibility, which is why we tested whether these variables affect the sensitivity of the visibility model to uncertainties in input data. When evaluating the rate of spatial overestimation of visibility when using various datasets and comparing it to the LiDAR-based model in our study, these variables did not usually have a significant effect. According to our results, therefore, it cannot be stated unequivocally that the rate of spatial overestimation by datasets is higher e.g. in flat or mountainous terrain or in areas that are more or less forested.

It is often impossible to say that a visibility obstacle is completely opaque or completely transparent. This is especially true about forest stands and techniques are being developed for a more realistic representation of forest stands than a simple polygon with a unified tree height as is most often the case at present (citace). Such models of forests consider individual trees and thus take into account both the density of the forest stand and the height where the tree crown begins as the forest is much more transparent below the crowns. Our results however imply that in the scale we used, such a laborious process does not bring any significant benefit for the validity and accuracy of the visibility model as the scale of the input data plays a much greater role. This is supported by a fact that apart from LiDAR, all used digital surface models significantly overestimated the visible areas. If making the forest stands more transparent, we would achieve an even higher percentage of the visible area, i.e. even greater overestimation. In other words, the model inaccuracy would probably get even worse if the obstacles were modelled in a more detailed way.

In all, LiDAR data have provided better overall results than vector data. There is however one more fact that should not be forgotten, namely the season in which the LiDAR data were acquired. While vector data are usually only updated once in a long time, LiDAR data can be acquired more often. A possible problem lies in detection of tree leaves. We can assume that if we compared data from winter LiDAR scanning against data from summer LiDAR scanning, the results would likely be different. The reason is that if data were acquired in the winter season, LiDAR scanning would underestimate the tree heights as the beam enters deep into the tree crown before being reflected. It is therefore necessary to find out if the scanning was performed during leaf-on or leaf-off season. This is also confirmed by the study Wasser et al. (2013).

Despite all statistical evaluations, we can only confirm the original hypothesis that in the case of vector data, the effect of the terrain complexity and assigned vegetation height on the accuracy is negligible and does not have a major impact on the results of the visibility analyses. Contrary, the greatest impact on the visibility analyses can be attributed to the quality of the spatial data and it is necessary to pay great attention to their selection for the particular purpose.

Use of raw data

Raw data acquired in the field in an appropriate way with a suitable sensor possess accuracy that is being reduced with every subsequent processing step. This is true of any data including those acquired by LiDAR. In the Studies 1 and 2, we were searching for the most suitable dataset for maximizing accuracy of visibility analyses. The best dataset was that acquired by airborne LiDAR scanning. In the Czech Republic, raw LiDAR data have been obtained for the entire area, they have however not been released. They are available either commercially (to be procured) or as a processed raster format, which is however generalized to a great degree and it is difficult to acquire information about the processing methods used for creating the product. Verification of suitability of such generalized data for a particular example of ecological analysis was the principal reason for undertaking Study 3.

The answer is simple. The more general pre-processing is applied on the data, the greater deviation from reality. The comparison of vertical accuracy of individual digital models confirmed the hypothesis that models based on raw unprocessed data are more accurate than those based on pre-processed data. The most accurate model in our study was based on raw LiDAR data. Both the commercially and freely available data were less accurate than raw data as a partial loss of both the spatial and vertical information occurred during processing. This loss was most likely caused by a use of a general processing method that was applied on the entire area of the Czech Republic. It is therefore not surprising that when using such data processed using a general algorithm for analysis of a much smaller area, the processed data are not sufficient. Our findings are in accordance with those by Mondino et al.(2016) who also concluded that pre-processed data may not be suitable for special analyses.

Comparison of raw LiDAR data and derived pre-processed models demonstrates that the model from raw data uses the full potential of all available laser reflections. The use of already generalized data may however lead to dilution of the original number of reflections below the level necessary for the detailed identification of the vegetation cover. Jakubowski et al. (2013) reported 1 point per square meter to be a threshold density for study of vegetation cover; our Study III confirms that result. Our results however show that when using a suitable processing method, even pre-processed input data that were originally not intended for such analysis and have a small density of reflection points can be utilized with relative success. It is however again necessary to mention the effect of the vegetation season of LiDAR data collection.

Further work

I dare say that the visibility analyses are, as far as researchers' interest is concerned, past their prime. In recent years, not many studies have been published that would deal with this topic and go into the principles of the method. Of late, most studies utilising visibility analyses are case studies where the visibility analysis serves as a tool for some discovery, often related to the historical development of the area (e.g. Carter et al., 2019; Paliou, 2011). A tendency to move from the dry land to analysing visual pollution on the surface of seas and oceans is also apparent (Depellegrin, 2016; Qiang et al., 2019; Robert, 2018). Attention is also paid to the development of visibility analyses and suitable data for realistic visibility modelling in the urban environment or to use visibility analyses for creating automatic views and landscape evaluation not only from the bird perspective but also as simulations of real view of the landscape from the perspective of an observer (e.g. Sahraoui et al., 2016; Yamagata et al., 2016). From time to time, a study is published that offers a way to improve and optimize the data processing methods in a faster way, in a way allowing processing of a larger area or allowing the use of more detailed and accurate data (e.g. Xia et al., 2011; Zhao et al., 2013).

Personally, I believe there is still a great potential in combining the visibility analyses and raw LiDAR data where the visibility would be calculated directly from a point cloud rather than from derived models. This would help resolve the problems associated with opaque or (partially) transparent obstacles. Of course, such a method would require overcoming of the problem with the points having no size. The point clouds do not need to originate from LiDAR data, either. The ever increasing use of UAVs corresponds with the ever increasing utilization of the structure from motion algorithm. While preparing orthomosaics from the individual images, identical points are searched for in the images and subsequently used for creating a point cloud. Such a point cloud could also be used for visibility analysis.

Another type of visibility analysis is a calculation of so-called Sky-view factor. Simply said, it is a cumulative visibility analysis providing information for every pixel about the relative size of the sky visible from the particular point. (see for more information e.g. Middel et al., 2018; Štular et al., 2012; Zakšek et al., 2011).This method is especially used in archaeology and in research of historic settlement. The view of the world through airborne images can capture even remnants of the landscape from the past including extinct settlements and other proofs of the human presence. If suitable geodata and methods of detection are chosen (and visibility analyses undoubtedly represent one of such methods), we acquire a valuable tool for reconstruction of the historic landscape.

Conclusion

Visibility analyses represent a powerful and widely used tool (not only) in landscape ecology. While the methods of the analysis has in principle reached their maximum efficiency, the question of selecting the proper spatial data for visibility analyses is far from concluded. This thesis compared various spatial input data to find out the most suitable data for creating digital terrain models and subsequent visibility analyses. At the beginning, the size of the problem was not so apparent but gradually, we began to appreciate the huge amount of available spatial data. Although airborne laser scanning has recently become the principal source of data, the quality of vector datasets is still sufficient for replacing LiDAR data for some types of analyses.

It is not possible to simply say what data are the best. There are no ideal data that would be suitable for any purpose without any or with only a minimum processing. It is true that airborne LiDAR has a lot of advantages – high horizontal and vertical accuracy, the possibility to acquire custom data for any area size, and penetration of the canopy by the laser beam, which results in the possibility of creating a digital terrain model in addition to the digital surface model even in inaccessible areas. The downside is however the relatively complicated processing of such data (unless using a pre-processed digital terrain/surface model). Vector data are usually obtained as a ready-made product that can be subsequently easily interpolated using several simple steps to create a digital terrain model. If using LiDAR data, especially raw LiDAR data, we have to perform a number of sophisticated steps and classifications to obtain a worthwhile model. Therefore, although LiDAR data are more suitable for creating digital terrain models in landscape ecology, vector data can be used with only a minor loss of accuracy if LiDAR data are unavailable. I do not expect cessation of the use of vector data due to increase in utilization of LiDAR data; there are many fields where a line cannot be replaced with a point cloud. New maps will still be created and updated, the records of private property will still be kept by the cadastral bureaus, drivers will still drive with the help of SatNavs... Here, LiDAR can help with mapping but cannot replace vector maps.

The effect of data accuracy on the result of the analysis is so great that it can completely suppress the influence of other factors. One of the aims of my thesis was to find out how much the complexity/ruggedness of the terrain and the degree of forestation influence the accuracy of visibility analysis. We found out that the effect was minimal. I have to admit that this finding was very surprising for me. I guessed that in a complex terrain, visibility analyses will be less accurate and that the same would be true about areas with a high degree of forestation. On the other hand, this finding is in accordance with the proposition that when using accurate data, the analysis captures all significant phenomena in the area and it does not matter how complex the terrain is. Visibility is of course also affected by the distance from the obstacle; our research however showed that as long as the obstacle is recorded in the input data, the distance of the obstacle has no notable effect on the accuracy of the analysis.

Our results obviate that when using any input data, the best results can be achieved when using data with a minimum degree of pre-processing or, ideally, to process the data individually in view of the purpose of its usage. My supervisor likes to say that she does not trust any statistics that she has not done herself and I believe that the same can be true about processing input data. If using pre-processed data, we usually have no information about the method by which the product was created, what methods and algorithms have been used, what data was removed and what the remaining data actually represent. It is often possible to find at least partial answers to those questions from metadata that should be provided together with the product. Such information is however often less than accurate and not completely reliable. In our study, we utilized a ready-made raster digital terrain model, its text version in the x,y,z format and raw LiDAR data. After processing and comparison of the results, we found out that the model created from the raw data contained the most information while the pre-processed ready-made raster model carried the least information. One can ask why the ready-made raster did not contain the information that we were able to extract from the raw data. The reason is quite simple – the LiDAR data and, consequently, the ready-made DTM were product of a nationwide campaign covering the entire Czech Republic and the acquisition and processing was thus not intended for analyses of sites as small as (in our case) several hectares. Another question can be posed – why have we even tested data that were apparently not intended for such detailed analyses and so might have seemed unsuitable? We wanted to test what information can be mined from the same data if we process them ourselves with the purpose of the analysis in mind. And the result is very interesting – although the raw data was on a nationwide scale, we managed to perform an accurate analysis at the level of individual shrubs when carefully processing the raw data – at the level that is not even indicated in the ready-made raster product. A regular user who utilizes the ready-made product therefore does not have any idea how great data he could have at his disposal if the part of information was not lost during processing.

The above mentioned example obviates that even data acquired for a rough scale can carry information sufficient for the studies on a local scale. At the same time, we however have to mention that this was rather an exception than a rule. In general, it can be said that the data mostly correspond with their scale as far as the level of generalization and detail are concerned and that we should always critically evaluate the data and consider its use for a particular analysis. Our studies showed that there was not a major difference between results of visibility analyses at 1:5,000 and 1:25,000 scales while the difference between those two datasets and that with a scale of 1:500,000 was immense. We can even say that the difference was so huge that the latter scale is only suitable for indicative purposes, not for drawing any firm conclusions. It can be also said that vector datasets overestimate the result of visibility analyses in comparison with a LiDAR dataset and this overestimation grows with coarsening the scale. For analyses in the landscape ecology in general and in practice, there are efforts to minimize the overestimation. An interesting phenomenon has arisen when evaluating control points of visibility analysis of a wind farm. With coarsening the scale, the false positives rate (i.e. instances when the model states that the wind farm is visible while in reality, it is not) grows. The number of false negatives (i.e. situations when the farm is visible but the model states it is not) however remains practically unchanged when using geodata of various scales.

Any data have their limits. As mentioned before, it is therefore necessary to critically evaluate the geodata in view of the analysis for which they shall serve.

This is especially true in the landscape ecology. If the circumstances allow, the best solution is to use raw data. This is the only way to make sure that a previous processing has not led to an unintended loss of spatial information. If we have LiDAR data at our disposal, the best solution for creating the digital terrain or surface model that shall serve as an input for the visibility analysis is to use such data. If, however, LiDAR data are not available, we can utilize vector data of a scale corresponding with the extent of the analysis and the area of interest. Usually, it is better to use finer scale data as such data come with a better chance of accurately depicting all terrain phenomena and objects that in the end form a terrain model with the best fit to reality. And, as also mentioned above, the accuracy of the visibility analysis will be closely related with the accuracy of the input digital model.

Although visibility analyses may be on the verge of scientific interest at present, they are far from obsolete. There are many possible applications of the visibility analyses in the landscape ecology, although it is unlikely that we should see any major breakthrough in the near future. Still, the input spatial data will always form the basis for those analyses and it is always necessary to critically consider the use of a particular input dataset. If the scale corresponds with the scale of the analysis, if we know the way in which the data were processed and if we use a proper algorithm for preparing a digital terrain model, we can assume that the resulting analysis will be of sufficient quality. This is not only true in landscape ecology but in all studies using spatial data as their inputs.

REFERENCES

Aben, J., Pellikka, P., Travis, J.M.J., 2018. A call for viewshed ecology: Advancing our understanding of the ecology of information through viewshed analysis. Methods Ecol. Evol. 9, 624–633. https://doi.org/10.1111/2041-210X.12902

Abraham, B., 2018. Video Game Visions of Climate Futures: ARMA 3 and Implications for Games and Persuasion. Games Cult. 13, 71–91. https://doi.org/10.1177/1555412015603844

Ackermann, F., 1999. Airborne laser scanning - Present status and future expectations. ISPRS J. Photogramm. Remote Sens. https://doi.org/10.1016/S0924-2716(99)00009-X

Alonso, J.C., Álvarez-Martínez, J.M., Palacín, C., 2012. Leks in ground-displaying birds: Hotspots or safe places? Behav. Ecol. 23, 491–501.

Anderson, E.S., Thompson, J. a., Austin, R.E., 2005. LIDAR density and linear interpolator effects on elevation estimates. Int. J. Remote Sens. 26, 3889–3900. https://doi.org/10.1080/01431160500181671

Anderson, E.S., Thompson, J.A., Crouse, D.A., Austin, R.E., 2006. Horizontal resolution and data density effects on remotely sensed LIDAR-based DEM. Geoderma 132, 406–415. https://doi.org/10.1016/j.geoderma.2005.06.004

ArcData, 2019. ArcČR 500. Arcdata Praha, Esri official distributor, https://www.arcdata.cz/produkty/geograficka-data/arccr-500

Baltsavias, E., Gruen, A., Eisenbeiss, H., Zhang, L., Waser, L.T., 2008. High-quality image matching and automated generation of 3D tree models, in: International Journal of Remote Sensing. https://doi.org/10.1080/01431160701736513

Barry, S., Elith, J., 2006. Error and uncertainty in habitat models. J. Appl. Ecol. 43, 413–423.

Bělka, L., Dušánek, P., Fiala, R., Gamrát, J., Kafka, O., Peichl, J., Šíma, J., 2010. Technická zpráva k digitálnímu modelu reliefu 4. generace (DMR 4G). Zeměměřičský úřad.

Bělka, L., Dušánek, P., Fiala, R., Gamrát, J., Kafka, O., Peichl, J., Šíma, J., 2012. Technická zpráva k digitálnímu modelu reliéfu 5. generace (DMR 5G). Zeměměřičský ústav.

Berry, R., Fitzell, I., Kidner, D., 2005. The transition to high resolution digital surface models: Improvements in visibility analysis performance, in:

Proceedings of the GIS Research UK Annual Conference 2005, GIS Research UK Annual Conference 2005; Glasgow, UK. pp. 245–251.

Betakova, V., Vojar, J., Sklenicka, P., 2015. Wind turbines location: How many and how far? Appl. Energy 151, 23–31.

Bishop, I.D., Miller, D.R., 2007. Visual assessment of off-shore wind turbines: The influence of distance, contrast, movement and social variables. Renew. Energy 32, 814–831.

Boelman, N.T., Holbrook, J.D., Greaves, H.E., Krause, J.S., Chmura, H.E., Magney, T.S., Perez, J.H., Eitel, J.U.H., Gough, L., Vierling, K.T., Wingfield, J.C., Vierling, L.A., 2016. Airborne laser scanning and spectral remote sensing give a bird's eye perspective on arctic tundra breeding habitat at multiple spatial scales. Remote Sens. Environ. 184, 337–349. https://doi.org/10.1016/j.rse.2016.07.012

Bohlin, J., Wallerman, J., Fransson, J.E.S., 2012. Forest variable estimation using photogrammetric matching of digital aerial images in combination with a high-resolution DEM. Scand. J. For. Res. https://doi.org/10.1080/02827581.2012.686625

Brabyn, L., Mark, D.M., 2011. Using viewsheds, GIS, and a landscape classification to tag landscape photographs. Appl. Geogr. 31, 1115–1122.

Caldwell, D., Mineter, M., 2003. Analysis and visualization of visibility surfaces. Proc. 7th ….

Carter, N., Santini, L., Barnes, A., Opitz, R., White, D., Safi, K., Davenport, B., Brown, C., Witschey, W., 2019. Country Roads: Travel, Visibility, and Late Classic Settlement in the Southern Maya Mountains. J. F. Archaeol. https://doi.org/10.1080/00934690.2019.1571373

Castro, M., García-Espona, A., Iglesias, L., 2015a. Terrain model resolution effect on sight distance on roads. Period. Polytech. Civ. Eng. https://doi.org/10.3311/PPci.7658

Castro, M., García-Espona, A., Iglesias, L., 2015b. Terrain model resolution effect on sight distance on roads. Period. Polytech. Civ. Eng. 59, 165–172. https://doi.org/10.3311/PPci.7658

Chamberlain, Brent C., Meitner, M.J., 2013. A route-based visibility analysis for landscape management. Landsc. Urban Plan. https://doi.org/10.1016/j.landurbplan.2012.12.004

Chamberlain, Brent C, Meitner, M.J., 2013. A route-based visibility analysis for landscape management. Landsc. Urban Plan. 111, 13–24.

Chance, C.M., Coops, N.C., Plowright, A.A., Tooke, T.R., Christen, A., Aven, N.,

2016. Invasive shrub mapping in an urban environment from hyperspectral and LiDAR-derived attributes. Front. Plant Sci. 7, 1–19. https://doi.org/10.3389/fpls.2016.01528

Corona, P., Fattorini, L., 2010. Area-based lidar-assisted estimation of forest standing volume. Can. J. For. Res. https://doi.org/10.1139/x08-122

Cramer, M., Haala, N., Laupheimer, D., Mandlburger, G., Havel, P., 2018. Ultra-high precision UAV-based LiDAR and dense image matching, in: International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives. https://doi.org/10.5194/isprs-archives-XLII-1-115-2018

ČUZK, 2019. Mapové podklady, Český úřad zeměměřičský a katastrální. https://geoportal.cuzk.cz

Dandois, J.P., Ellis, E.C., 2010. Remote sensing of vegetation structure using computer vision. Remote Sens. https://doi.org/10.3390/rs2041157

Danese, M., Nolè, G., Murgante, B., 2009. Visual impact assessment in urban planning. Stud. Comput. Intell. https://doi.org/10.1007/978-3-540-89930-3_8

De Floriani, L., Marzano, P., Puppo, E., 1994. Line-of-sight communication on terrain models. Int. J. Geogr. Inf. Syst. 8, 329–342.

De Montis, A., Caschili, S., 2012. Nuraghes and landscape planning: Coupling viewshed with complex network analysis. Landsc. Urban Plan. 105, 315–324.

Dean, D.J., 1997a. Improving the accuracy of forest viewsheds using triangulated networks and the visual permeability method. Can. J. For. Res. 27, 969–977.

Dean, D.J., 1997b. Finding optimal routes for networks of harvest site access roads using GIS-based techniques. Can. J. For. Res. 27, 11–22. https://doi.org/10.1139/x96-144

Depellegrin, D., 2016. Assessing cumulative visual impacts in coastal areas of the Baltic Sea. Ocean Coast. Manag. 119, 184–198. https://doi.org/10.1016/j.ocecoaman.2015.10.012

Domingo-Santos, J.M., de Villarán, R.F., Rapp-Arrarás, Í., de Provens, E.C.P., 2011. The visual exposure in forest and rural landscapes: An algorithm and a GIS tool. Landsc. Urban Plan. 101, 52–58.

Dušánek, P., 2014. Nové výškopisné mapování české republiky. Zeměměřičský úrad.

Dušánek, P., Hron, V., Hoffman, P., Setnička, M., Šíma, J., 2016. Technická zpráva k digitálnímu modelu povrchu 1.generace (DMP1G). Zeměměřičský ústav.

Eisenbeiß, H., Zurich, E.T.H., Eisenbeiß, H., Zürich, E.T.H., 2009. UAV photogrammetry, Institute of Photogrammetry and Remote Sensing. https://doi.org/doi:10.3929/ethz-a-005939264

Estornell, J., Ruiz, L.A., Velázquez-Martí, B., Fernández-Sarría, A., 2011. Estimation of shrub biomass by airborne LiDAR data in small forest stands. For. Ecol. Manage. 262, 1697–1703. https://doi.org/10.1016/j.foreco.2011.07.026

Etherington, T.R., Alexander, S.M., 2008. Using viewsheds to determine area sampled by ground-based radiotelemetry. J. Wildl. Manage. 72, 1043–1046.

Falconer, L., Hunter, D.C., Telfer, T.C., Ross, L.G., 2013. Visual, seascape and landscape analysis to support coastal aquaculture site selection. Land use policy 34, 1–10.

Fernandez-Jimenez, L.A., Mendoza-Villena, M., Zorzano-Santamaria, P., Garcia-Garrido, E., Lara-Santillan, P., Zorzano-Alba, E., Falces, A., 2015. Site selection for new PV power plants based on their observability. Renew. Energy 78, 7–15.

Fisher, P.F., 1996. Extending the applicability of viewsheds in landscape planning. Photogramm. Eng. {&} Remote Sens. 62, 1297–1302.

Fisher, P.F., 1995. An exploration of probable viewsheds in landscape planning. Environ. Plan. B Plan. Des. 22, 527–546.

Fisher, P.F., 1994. Probable and fuzzy models of the viewshed operation, in: Innovations in GIS. Taylor {&} Francis London, pp. 161–175.

Fisher, Peter F, 1993. Algorithm and implementation uncertainty in viewshed analysis. Int. J. Geogr. Inf. Syst. 7, 331–347.

Fisher, Peter F., 1993. Algorithm and implementation uncertainty in viewshed analysis. Int. J. Geogr. Inf. Syst. https://doi.org/10.1080/02693799308901965

Fisher, P.F., 1992. First experiments in viewshed uncertainty: Simulating fuzzy viewsheds. Photogramm. Eng. {&} Remote Sens. 58, 345–352.

Fogl, M., Moudrý, V., 2016. Influence of vegetation canopies on solar potential in urban environments. Appl. Geogr. 66, 73–80. https://doi.org/10.1016/j.apgeog.2015.11.011

Geneletti, D., 2008. Impact assessment of proposed ski areas: A GIS approach integrating biological, physical and landscape indicators. Environ. Impact Assess. Rev. 28, 116–130.

Germino, M.J., Reiners, W. a., Blasko, B.J., McLeod, D., Bastian, C.T., 2001.

Estimating visual properties of rocky mountain landscapes using GIS. Landsc. Urban Plan. 53, 71–83.

Glenn, N.F., Spaete, L.P., Sankey, T.T., Derryberry, D.R., Hardegree, S.P., Mitchell, J., 2011. Errors in LiDAR-derived shrub height and crown area on sloped terrain. J. Arid Environ. 75, 377–382. https://doi.org/10.1016/j.jaridenv.2010.11.005

Glowienka, E., Michalowska, K., Opalinski, P., Hejmanowska, B., Mikrut, S., Kramarczyk, P., 2017. Use of LIDAR Data in the 3D/4D Analyses of the Krakow Fortress Objects. IOP Conf. Ser. Mater. Sci. Eng. 245. https://doi.org/10.1088/1757-899X/245/4/042080

Goodchild, M.F., 1992. Geographical data modeling. Comput. Geosci. https://doi.org/10.1016/0098-3004(92)90069-4

Guo, Q., Li, W., Yu, H., Alvarez, O., 2010. Effects of topographic variability and Lidar sampling density on several DEM interpolation methods. Photogramm. Eng. Remote Sens. 76, 701–712. https://doi.org/10.14358/PERS.76.6.701

Hopkinson, C., Chasmer, L.E., Sass, G., Creed, I.F., Sitar, M., Kalbfleisch, W., Treitz, P., 2005. Vegetation class dependent errors in lidar ground elevation and canopy height estimates in a boreal wetland environment. Can. J. Remote Sens. 31, 191–206. https://doi.org/10.5589/m05-007

Huang, H., Gong, P., Cheng, X., Clinton, N., Li, Z., 2009. Improving measurement of forest structural parameters by co-registering of high resolution aerial imagery and low density LiDAR data. Sensors. https://doi.org/10.3390/s90301541

Huss, R.E., Pumar, M. a, 1997. Effect of database errors on intervisibility estimation. Photogramm. Eng. {&} Remote Sens. 63, 415–424.

Hyyppä, J., Hyyppä, H., Leckie, D., Gougeon, F., Yu, X., Maltamo, M., 2008. Review of methods of small-footprint airborne laser scanning for extracting forest inventory data in boreal forests, in: International Journal of Remote Sensing. https://doi.org/10.1080/01431160701736489

Jakubowski, M.K., Guo, Q., Kelly, M., 2013. Tradeoffs between lidar pulse density and forest measurement accuracy. Remote Sens. Environ. 130, 245–253. https://doi.org/10.1016/j.rse.2012.11.024

Jiang, G., Sun, H., Lang, J., Yang, L., Li, C., Lyet, A., Long, B., Miquelle, D.G., Zhang, C., Aramilev, S., Ma, J., Zhang, M., 2014. Effects of environmental and anthropogenic drivers on Amur tiger distribution in northeastern China. Ecol. Res. https://doi.org/10.1007/s11284-014-1160-3

Kaneko, Korehisa, Nohara, Seiich, Kaneko, K, Nohara, S, 2014. Review of

Effective Vegetation Mapping Using the UAV (Unmanned Aerial Vehicle) Method. J. Geogr. Inf. Syst. 733–742. https://doi.org/10.4236/jgis.2014.66060

Khosravipour, A., Skidmore, A.K., Isenburg, M., 2016. Generating spike-free digital surface models using LiDAR raw point clouds: A new approach for forestry applications. Int. J. Appl. Earth Obs. Geoinf. 52, 104–114. https://doi.org/10.1016/j.jag.2016.06.005

Kim, S., McGaughey, R.J., Andersen, H.E., Schreuder, G., 2009. Tree species differentiation using intensity data derived from leaf-on and leaf-off airborne laser scanner data. Remote Sens. Environ. https://doi.org/10.1016/j.rse.2009.03.017

Kim, Y.H., Rana, S., Wise, S., 2004a. Exploring multiple viewshed analysis using terrain features and optimisation techniques. Comput. {&} Geosci. 30, 1019–1032.

Kim, Y.H., Rana, S., Wise, S., 2004b. Exploring multiple viewshed analysis using terrain features and optimisation techniques. Comput. Geosci. https://doi.org/10.1016/j.cageo.2004.07.008

Kizuka, T., Akasaka, M., Kadoya, T., Takamura, N., 2014. Visibility from roads predict the distribution of invasive fishes in agricultural ponds. PLoS One 9, 1–10.

Klemas, V. V., 2015. Coastal and Environmental Remote Sensing from Unmanned Aerial Vehicles: An Overview. J. Coast. Res. https://doi.org/10.2112/jcoastres-d-15-00005.1

Klouček, T., Lagner, O., Šímová, P., 2015. How does data accuracy influence the reliability of digital viewshed models? A case study with wind turbines. Appl. Geogr. 64, 46–54. https://doi.org/10.1016/j.apgeog.2015.09.005

Koska, B., Jirka, V., Urban, R., Křemen, T., Hesslerová, P., Jon, J., Pospíšil, J., Fogl, M., 2017. Suitability, characteristics, and comparison of an airship UAV with lidar for middle size area mapping. Int. J. Remote Sens. https://doi.org/10.1080/01431161.2017.1285086

Kraus, K., Pfeifer, N., 1998. Determination of terrain models in wooded areas with airborne laser scanner data. ISPRS J. Photogramm. Remote Sens. https://doi.org/10.1016/S0924-2716(98)00009-4

Kumi-boateng, B., Yakubu, I., 2010. Assessing the Quality of Spatial Data. Eur. J. Sci. Res. 43, 507–515.

Lake, I.R., Lovett, A. a., Bateman, I.J., Day, B., 2000a. Using GIS and large-scale digital data to implement hedonic pricing studies. Int. J. Geogr. Inf. Sci. 14, 521–541.

Lake, I.R., Lovett, A.A., Bateman, I.J., Day, B., 2000b. Using gis and large-scale digital data to implement hedonic pricing studies. Int. J. Geogr. Inf. Sci. https://doi.org/10.1080/136588100415729

Lang, C., Opaluch, J.J., Sfinarolakis, G., 2014. The windy city: Property value impacts of wind turbines in an urban setting. Energy Econ. 44, 413–421.

Larsen, F.W., Rahbek, C., 2005. The influence of spatial grain size on the suitability of the higher-taxon approach in continental priority-setting. Anim. Conserv. https://doi.org/10.1017/S1367943005002362

Lechner, A.M., Langford, W.T., Bekessy, S.A., Jones, S.D., 2012. Are landscape ecologists addressing uncertainty in their remote sensing data? Landsc. Ecol. https://doi.org/10.1007/s10980-012-9791-7

Lecours, V., Brown, C.J., Devillers, R., Lucieer, V.L., Edinger, E.N., 2016. Comparing selections of environmental variables for ecological studies: A focus on terrain attributes. PLoS One 11, 1–18. https://doi.org/10.1371/journal.pone.0167128

Lecours, V., Devillers, R., Schneider, D.C., Lucieer, V.L., Brown, C.J., Edinger, E.N., 2015a. Spatial scale and geographic context in benthic habitat mapping: Review and future directions. Mar. Ecol. Prog. Ser. 535, 259–284. https://doi.org/10.3354/meps11378

Lecours, V., Simms, A., Devillers, R., Lucieer, V., Edinger, E., 2015b. Finding the Best Combinations of Terrain Attributes and GIS software for Meaningful Terrain Analysis. Geomorphometry Geosci. 133–136.

Lee, J., Stucky, D., 1998. On applying viewshed analysis for determining least-cost paths on digital elevation models. Int. J. Geogr. Inf. Sci. https://doi.org/10.1080/136588198241554

Li, D., Zhang, J., Wu, H., 2012. Spatial data quality and beyond. Int. J. Geogr. Inf. Sci. 26, 1–14. https://doi.org/10.1080/13658816.2012.719625

Li, Z., 1990. Sampling Strategy and Accuracy Assessment for Digital Terrain Modelling. University of Glasgow.

Li, Z., Zhu, Q., Gold, C., 2004a. Digital terrain modeling: Principles and methodology, Digital Terrain Modeling: Principles and Methodology. https://doi.org/10.1201/9780203357132

Li, Z., Zhu, Q., Gold, C., 2004b. Digital terrain modeling: Principles and methodology, Digital Terrain Modeling: Principles and Methodology. https://doi.org/10.1201/9780203357132

Lim, K., Treitz, P., Wulder, M., St-Onge, B., Flood, M., 2003. LiDAR remote sensing of forest structure. Prog. Phys. Geogr. 27, 88–106.

https://doi.org/10.1191/0309133303pp360ra

Lisein, J., Pierrot-Deseilligny, M., Bonnet, S., Lejeune, P., 2013. A photogrammetric workflow for the creation of a forest canopy height model from small unmanned aerial system imagery. Forests. https://doi.org/10.3390/f4040922

Liu, L., Zhang, L.Q., Ma, J.T., Zhang, L., Zhang, X.M., Xiao, Z.Q., Yang, L., 2010. An improved line-of-sight method for visibility analysis in 3D complex landscapes. Sci. China Inf. Sci. 53, 2185–2194.

Lochovsky, F.H., Tsichritzis, D.C., 1977. User performance considerations in DBMS selection, in: Proceedings of the ACM SIGMOD International Conference on Management of Data. https://doi.org/10.1145/509404.509425

Lu, M., Zhang, J.F., Lv, P., Fan, Z.H., 2008. Least visible path analysis in raster terrain. Int. J. Geogr. Inf. Sci.

Maloy, Mark A., Dean, D.J., 2001. An accuracy assessment of various GIS-based viewshed delineation techniques. Photogramm. Eng. Remote Sens. 67, 1293–1298.

Maloy, Mark A, Dean, D.J., 2001. An accuracy assessment of various GIS-based viewshed delineation techniques. Photogramm. Eng. Remote Sensing.

Manchado, C., Otero, C., Gómez-Jáuregui, V., Arias, R., Bruschi, V., Cendrero, A., 2013. Visibility analysis and visibility software for the optimisation of wind farm design. Renew. Energy 60, 388–401.

Mårtensson, S.G., Reshetyuk, Y., 2017. Height uncertainty in digital terrain modelling with unmanned aircraft systems. Surv. Rev. 49, 312–318. https://doi.org/10.1080/00396265.2016.1180754

Maune, D.F., Huff, L.C., and Guenther, G.C., 2001. DEM user applications. Am. Soc. Photogramm. Remote Sens. 367–394.

Meek, S., Goulding, J., Priestnall, G., 2013. The Influence of Digital Surface Model Choice on Visibility-based Mobile Geospatial Applications. Trans. GIS 17, 526–543. https://doi.org/10.1111/tgis.12021

Melin, M., Matala, J., Mehtätalo, L., Pusenius, J., Packalen, P., 2016. Ecological dimensions of airborne laser scanning - Analyzing the role of forest structure in moose habitat use within a year. Remote Sens. Environ. 173, 238–247. https://doi.org/10.1016/j.rse.2015.07.025

Middel, A., Lukasczyk, J., Maciejewski, R., Demuzere, M., Roth, M., 2018. Sky View Factor footprints for urban climate modeling. Urban Clim. https://doi.org/10.1016/j.uclim.2018.05.004

Mikšovský, M., Zimová, R., 2006. Historická mapování českých zemí.

Miller, C.L., LaFlamme, R.A., 1958. The digital terrain model - Theory & Application. Am. Soc. Photogramm.

Molnarova, K., Sklenicka, P., Stiborek, J., Svobodova, K., Salek, M., Brabec, E., 2012. Visual preferences for wind turbines: Location, numbers and respondent characteristics. Appl. Energy 92, 269–278.

Mondino, E.B., Fissore, V., Lessio, A., Motta, R., 2016. Are the new gridded DSM/DTMs of the Piemonte region (Italy) proper for forestry? A fast and simple approach for a posteriori metric assessment. IForest 9, 901–909. https://doi.org/10.3832/ifor1992-009

Moore, I.D., Grayson, R.B., Ladson, A.R., 1991. Digital terrain modelling: A review of hydrological, geomorphological, and biological applications. Hydrol. Process. https://doi.org/10.1002/hyp.3360050103

Mora, B., Wulder, M.A., Hobart, G.W., White, J.C., Bater, C.W., Gougeon, F.A., Varhola, A., Coops, N.C., 2013. Forest inventory stand height estimates from very high spatial resolution satellite imagery calibrated with lidar plots. Int. J. Remote Sens. https://doi.org/10.1080/01431161.2013.779041

Moudrý, V., Šímová, P., 2012. Influence of positional accuracy, sample size and scale on modelling species distributions. Int. J. Geogr. Inf. Sci.

Moudrý, V., Urban, R., Štroner, M., Komárek, J., Brouček, J., Prošek, J., 2019. Comparison of a commercial and home-assembled fixed-wing UAV for terrain mapping of a post-mining site under leaf-off conditions. Int. J. Remote Sens. https://doi.org/10.1080/01431161.2018.1516311

Murgoitio, J., Shrestha, R., Glenn, N., Spaete, L., 2014a. Airborne LiDAR and terrestrial laser scanning derived vegetation obstruction factors for visibility models. Trans. GIS 18, 147–160.

Murgoitio, J., Shrestha, R., Glenn, N., Spaete, L., 2014b. Airborne LiDAR and terrestrial laser scanning derived vegetation obstruction factors for visibility models. Trans. GIS. https://doi.org/10.1111/tgis.12022

Murgoitio, J.J., Shrestha, R., Glenn, N.F., Spaete, L.P., 2013. Improved visibility calculations with tree trunk obstruction modeling from aerial LiDAR. Int. J. Geogr. Inf. Sci. 27, 1865–1883. https://doi.org/10.1080/13658816.2013.767460

Næsset, E., 2002. Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. Remote Sens. Environ. https://doi.org/10.1016/S0034-4257(01)00290-5

Nagy, G., 1994. Terrain visibility. Comput. {&} Graph. 18, 763–773.

Nijhuis, S., Van Der Hoeven, F., 2018. Exploring the skyline of rotterdam and the hague. visibility analysis and its implications for tall building policy. Built Environ. https://doi.org/10.2148/benv.43.4.571

Nurminen, K., Karjalainen, M., Yu, X., Hyyppä, J., Honkavaara, E., 2013. Performance of dense digital surface models based on image matching in the estimation of plot-level forest variables. ISPRS J. Photogramm. Remote Sens. https://doi.org/10.1016/j.isprsjprs.2013.06.005

Nutsford, D., Reitsma, F., Pearson, A.L., Kingham, S., 2015. Personalising the viewshed: Visibility analysis from the human perspective. Appl. Geogr. 62, 1–7.

Ogburn, D.E., 2006. Assessing the level of visibility of cultural objects in past landscapes. J. Archaeol. Sci. 33, 405–413.

Olsoy, P.J., Forbey, J.S., Rachlow, J.L., Nobler, J.D., Glenn, N.F., Shipley, L.A., 2015. Fearscapes: Mapping functional properties of cover for prey with terrestrial LiDAR. Bioscience. https://doi.org/10.1093/biosci/biu189

Paliou, E., 2011. The communicative potential of theran murals in late Bronze Age Akrotiri: Applying viewshed analysis in 3D townscapes. Oxford J. Archaeol. 30, 247–272.

Pavelka, K., Faltynova, M., Matouskova, E., Sedina, J., 2015. Biodiversity, forest condition mapping and forest harvesting monitoring using UAV. ACRS 2015 - 36th Asian Conf. Remote Sens. Foster. Resilient Growth Asia, Proc.

Petrie, G., Kennie, T., 1987. An Introduction to Terrain Modeling: Applications and Terminology, Terrain Modelling in Surveying and Civil Engineering: A Short Course. University of Glasgow.

Podobnikar, T., 2009. Methods for visual quality assessment of a digital terrain model. Sapiens 2.

Poudyal, N.C., Hodges, D.G., Fenderson, J., Tarkington, W., 2010. Realizing the economic value of a forested landscape in a viewshed. South. J. Appl. For. 34, 72–78. https://doi.org/10.1093/sjaf/34.2.72

Qiang, Y., Shen, S., Chen, Q., 2019. Visibility analysis of oceanic blue space using digital elevation models. Landsc. Urban Plan. 181, 92–102. https://doi.org/10.1016/j.landurbplan.2018.09.019

R Core Team, 2015. R: A Language and environment for statistical computing.

Robert, S., 2018. Assessing the visual landscape potential of coastal territories for spatial planning. A case study in the French Mediterranean. Land use policy 72, 138–151. https://doi.org/10.1016/j.landusepol.2017.12.037

Rosa, D. La, 2011. The observed landscape: map of visible landscape values in the province of Enna (Italy). J. Maps 7, 291–303.

Ruiz, L.A., Hermosilla, T., Mauro, F., Godino, M., 2014. Analysis of the influence of plot size and LiDAR density on forest structure attribute estimates. Forests 5, 936–951. https://doi.org/10.3390/f5050936

Sahraoui, Y., Vuidel, G., Foltête, J.-C., Joly, D., 2016. PixScape – un outil logiciel intégré pour l'analyse du paysage visiblePixScape - an integrated software tool for analyzing landscape visibility. Cybergeo. https://doi.org/10.4000/cybergeo.27862

Sander, H.A., Manson, S.M., 2007a. Heights and locations of artificial structures in viewshed calculation: How close is close enough? Landsc. Urban Plan. 82, 257–270.

Sander, H.A., Manson, S.M., 2007b. Heights and locations of artificial structures in viewshed calculation: How close is close enough? Landsc. Urban Plan. https://doi.org/10.1016/j.landurbplan.2007.03.002

Sang, N., Gold, C., Miller, D., 2016. The topological viewshed: embedding topological pointers into digital terrain models to improve GIS capability for visual landscape analysis. Int. J. Digit. Earth 9, 1185–1205. https://doi.org/10.1080/17538947.2016.1192229

Sansoni, C., 1996. Visual analysis: A new probabilistic technique to determine landscape visibility. Comput. Aided Des. 28, 289–299.

Schirpke, U., Tasser, E., Tappeiner, U., 2013. Predicting scenic beauty of mountain regions. Landsc. Urban Plan. 111, 1–12.

Sevenant, M., Antrop, M., 2007. Settlement models, land use and visibility in rural landscapes: Two case studies in Greece. Landsc. Urban Plan. 80, 362–374.

Sharma, A., 2009. Comparing the radiometric behavior of LISS III data of IRS 1D and IRS P6 for landscape ecological application. J. Environ. Informatics 13, 66–74. https://doi.org/10.3808/jei.200900141

Šíma, J., 2016. Obsah Základní báze geografických dat 73–84.

Smith, C., Cochrane, E.E., 2011. How is visibility important for defence? A GIS analysis of sites in the western Fijian Islands. Archaeol. Ocean. 46, 76–84.

Smith, D.E., Zuber, M.T., Neumann, G.A., Lemoine, F.G., 1997. Topography of the Moon from the Clementine lidar. J. Geophys. Res. E Planets. https://doi.org/10.1029/96JE02940

Steinmann, K., Mandallaz, D., Ginzler, C., Lanz, A., 2013. Small area estimations of proportion of forest and timber volume combining Lidar data and stereo aerial images with terrestrial data. Scand. J. For. Res. https://doi.org/10.1080/02827581.2012.754936

Štular, B., Kokalj, Ž., Oštir, K., Nuninger, L., 2012. Visualization of lidar-derived relief models for detection of archaeological features. J. Archaeol. Sci. https://doi.org/10.1016/j.jas.2012.05.029

Sun, X., 2018. Lidar Sensors From Space, in: Comprehensive Remote Sensing. https://doi.org/10.1016/b978-0-12-409548-9.10327-6

Tulldahl, H.M., Larsson, H., 2014. Lidar on small UAV for 3D mapping, in: Electro-Optical Remote Sensing, Photonic Technologies, and Applications VIII; and Military Applications in Hyperspectral Imaging and High Spatial Resolution Sensing II. https://doi.org/10.1117/12.2068448

Turner, W., Rondinini, C., Pettorelli, N., Mora, B., Leidner, A.K., Szantoi, Z., Buchanan, G., Dech, S., Dwyer, J., Herold, M., Koh, L.P., Leimgruber, P., Taubenboeck, H., Wegmann, M., Wikelski, M., Woodcock, C., 2015. Free and open-access satellite data are key to biodiversity conservation. Biol. Conserv. 182, 173–176. https://doi.org/10.1016/j.biocon.2014.11.048

ZÚ., 2008. Projekt tvorby nového výškopisu České republiky. Zeměměřičský úřad.

Valbuena, R., Maltamo, M., Packalen, P., 2016. Classification of multilayered forest development classes from low-density national airborne lidar datasets. Forestry. https://doi.org/10.1093/forestry/cpw010

Véga, C., St-Onge, B., 2008. Height growth reconstruction of a boreal forest canopy over a period of 58 years using a combination of photogrammetric and lidar models. Remote Sens. Environ. https://doi.org/10.1016/j.rse.2007.09.002

VTÚ, 2019. DMÚ 25 - popis dat. Vojenský topografický ústav Dobruška

Wallentin, G., Tappeiner, U., Strobl, J., Tasser, E., 2008. Understanding alpine tree line dynamics: An individual-based model. Ecol. Modell. https://doi.org/10.1016/j.ecolmodel.2008.07.005

Wasser, L., Day, R., Chasmer, L., Taylor, A., 2013a. Influence of Vegetation Structure on Lidar-derived Canopy Height and Fractional Cover in Forested Riparian Buffers During Leaf-Off and Leaf-On Conditions. PLoS One 8. https://doi.org/10.1371/journal.pone.0054776

Wasser, L., Day, R., Chasmer, L., Taylor, A., 2013b. Influence of Vegetation Structure on Lidar-derived Canopy Height and Fractional Cover in Forested Riparian Buffers During Leaf-Off and Leaf-On Conditions. PLoS One. https://doi.org/10.1371/journal.pone.0054776

Wheatley, D., 1995. Cumulative viewshed analysis: a GIS-based method for investigating intervisibility, and its archaeological application, in: Archaeology and Geographic Information Systems: A European Perspective.

White, J.C., Wulder, M.A., Vastaranta, M., Coops, N.C., Pitt, D., Woods, M., 2013. The utility of image-based point clouds for forest inventory: A comparison with airborne laser scanning. Forests. https://doi.org/10.3390/f4030518

Wu, S., Li, J., Huang, G., 2006. Deriving vegetation structure in ecological applications from LiDAR imagery. J. Environ. Informatics 8, 111–114. https://doi.org/10.3808/jei.200600081

Xia, Y., Kuang, L., Li, X., 2011. Accelerating geospatial analysis on GPUs using CUDA. J. Zhejiang Univ. Sci. C 12, 990–999. https://doi.org/10.1631/jzus.c1100051

Xiaoye Liu, 2008. Airborne LiDAR for DEM generation: some critical issues. Prog. Phys. Geogr. 32, 31–49. https://doi.org/10.1177/0309133308089496

Yamagata, Y., Murakami, D., Yoshida, T., Seya, H., Kuroda, S., 2016. Value of urban views in a bay city: Hedonic analysis with the spatial multilevel additive regression (SMAR) model. Landsc. Urban Plan. 151, 89–102. https://doi.org/10.1016/j.landurbplan.2016.02.008

Zakšek, K., Oštir, K., Kokalj, Ž., 2011. Sky-view factor as a relief visualization technique. Remote Sens. https://doi.org/10.3390/rs3020398

Zhao, Y., Padmanabhan, A., Wang, S., 2013. A parallel computing approach to viewshed analysis of large terrain data using graphics processing units. Int. J. Geogr. Inf. Sci. 27, 363–384. https://doi.org/10.1080/13658816.2012.692372

CURRICULUM VITAE

Personal:
Ondřej Lagner

Born: 30th October, 1989, Praha (Czech Republic)

Education:

2014 – present: Ph.D. studies,

Applied and landscape Ecology, Faculty of Environmental science,
Czech University of Life Sciences, Prague, Czech Republic

Thesis: Spatial data quality in digital visibility models

2012 – 2016: Master's degree

Faculty of Science, Charles University in Prague, Prague, Czech Republic

Master thesis: A development of the GIS tool for evaluation of population mobility based on mobile phone operators data

2012 – 2014: Master's degree

Faculty of Environmental science, Czech University of Life Sciences, Prague, Czech Republic

Master thesis: Evaluation of visibility analyses based on different geodata

2009 – 2012: Bachelor's degree

Faculty of Science, Charles University in Prague, Prague, Czech Republic

Bachelor's thesis: Creating 3D model of the White tower in Hradec Kralove from terrestrial laser scanning data

Professional experience:

Since 2018: Alderman in Prague 22 district

Since 2015: Czech University of Life Sciences Prague, Department of Applied Geoinformatics and Spatial Planning (assistant professor of GIS and Cartography, authorized operator of UAV's).

Since 2015: Member of academic senate Faculty of Environmental science, Czech University of Life Sciences, Prague, Czech Republic

Since 2014: Member of the Prague 22 district council

Publications:

Moudrý, V., Beková, A., Lagner, O. 2019. Evaluation of a high resolution UAV imagery model for rooftop solar irradiation estimates. *Remote Sensing Letters*, *10*(11), 1077-1085

Lagner,O., Klouček, T., Fogl. M., 2019. The signifikance of using raw data: A case study with canopy height models of shrubs. SGEM2019 Conference Proceedings, ISBN 978-619-7408-79-9, ISSN 1314-2704, vol. 19, Issue 2.1. 1089-1098 pp.

Klouček, T., Moravec, D., Komárek, J., Lagner, O., & Štych, P. (2018). Selecting appropriate variables for detecting grassland to cropland changes using high resolution satellite data. *PeerJ*, *6*, e5487.

Lagner, O., Klouček, T., & Šímová, P. (2018). Impact of input data (in) accuracy on overestimation of visible area in digital viewshed models. *PeerJ*, *6*, e4835.

Klouček, T., Lagner, O., & Šímová, P. (2015). How does data accuracy influence the reliability of digital viewshed models? A case study with wind turbines. *Applied Geography*, *64*, 46-54.

Grants and project:

National grants

2018 - 2019: Early Detection of Forest Infestation by Bark Beetle (Ips Typographus) Using Unmanned Aerial Vehicles (principal investigator (2018) co-investigator (2019)).

2015 - 2016: Norway grants: The Reduction of Habitat Fragmentation Consequences in Various Types of Landscape in the Czech Republic (co-investigator).

Internal grants

(Founded by Internal Grant Agency of the University/Faculty)

2018 - 2019: Remote Sensing: an Effective Tool for the Study of Spatial Dynamics of Bark Beetles at Krkonoše Mountains National Park (co-investigator).

2017 - 2018: Influence of Remote Sensing Data Resolution in Evaluating Ecological Measures (co-investigator).

2015 - 2017: Usability of Modern Geodata in Ecology and Landscape Ecology (principal investigator).

2015 - 2016: Usability of Digital Surface Models for Selected Tasks in Animal and Landscape Ecology (co-investigator).

2013 - 2013 Influence of Input Geodata on Visibility Analysis of Wind Turbines (co-investigator).

Teaching activities

Since 2014 Lecturer of GIS, Cartography, Remote Sensing using UAVs

Since 2014 Supervisor and reviewer of 2 Bachelor and 4 Master Thesis (GIS applications, Landscape ecology)