1.01 Izvirni znanstveni članek

Consequences of Data and Information Uncertainties for Spatial Modelling of the Cultural Landscape

Posledice negotovih podatkov in informacij pri prostorskem modeliranju kulturne krajine

© Tomaž Podobnikar
UNG, Fakulteta za znanosti o okolju
ZRC SAZU, Inštitut za za antropološke in prostorske študije
tp@zrc-sazu.si

Abstract: The quality of data and information depends on many elements that may considerably affect to the results of subsequent spatial modelling and interpretations in archaeology. This complex needs careful treatments that enable recognition the nature of the entirely studied problem. Three applied problems to the uncertainty analyses and modelling of cultural landscapes in landscape archaeology are exposed: (1) influence of the sea level changes on the interpretation of various ancient human activities, (2) analyses of the optimal paths algorithms and reliability of the data sources for interpretation of the ancient paths between settlements, and (3) analyses of the visibility from the significant points and visual communication interpretation. The case study areas are from Croatia (Central Dalmatia), Slovenia (Dolenjska) and Mexico (Yucatan, Campeche).

Keywords: uncertainty, quality of data, model and interpretation, paths, visibility, simulations

Izvleček: Kakovost podatkov in informacij je odvisna od velikega števila elementov, ki lahko pomembno vplivajo na rezultate nadaljnjega prostorskega modeliranja in interpretacije v arheologiji. Vse to potrebuje skrbno obravnavo, kar omogoča celovito spoznavanje narave obravnavanega problema. V okviru arheologije krajine smo opisali tri probleme negotovosti analiz in modeliranja kulturne krajine: (1) vpliv spreminjanja višine morja pri interpretaciji preteklih človekovih dejavnosti, (2) analize algoritmov za iskanje optimalnih poti in zanesljivosti virov podatkov za interpretacijo starodavnih poti med naselbinami in (3) analize vidnosti z značilnih točk ter interpretacija vizualne komunikacije. Izbrana so bila naslednja študijska območja: Hrvaška (Srednja Dalmacija), Slovenija (Dolenjska) in Mehika (Jukatan, Campeche).

Ključne besede: negotovost, kakovost podatkov, modela in interpretacije, poti, vidnost, simulacije

Introduction

The quality of spatially oriented studies depends on many elements, which cannot be easily and uniformly described. It is also impossible to provide a systematic taxonometry for them. The data or model accuracy may affect subsequent processing steps strongly. We can consider the errors as a critical element for determination something as "appropriate for usage", according to our tasks. It means that the error should not be considered just as potential inconvenience. We can distinguish two basic groups of errors (Burrough, McDonnell 1998, 241). The first considers data sources (inherent to the object/entity that is measured or observed) and the second considers a processing/operational error (error/uncertainty propagation considering spatial analyses and modelling). Another view distinguishes between error in the system and human error (factor). Errors in a system can be considered as latent design errors. We should be aware that a low quality of the source data causes a low quality of the final product. The human error includes cognitive bias, which can be minimised by making systems more forgiving or error-tolerant. The human error is also related with interpretation of final results of the modelling.

The majority of error sources can be found, described or evaluated by statistical methods or by other means (Podobnikar 2009). The measurement errors are split to random, systematic and gross. The first group of errors is considered as random variation and would be diminished with higher quality measurements/observations. The second group may be eliminated only when the accurate data are presented. The last group is often considered as a mistake as it is caused by a fault. Although the random measurement errors of spatial data are mainly considered to be normally distributed, their character often contains admixtures of poorly explained properties and other uncertainties that came from difficulties in separation between particular types of errors. Many standard methods for measuring and further modelling errors techniques are known, e.g. known as standard deviation, root means square error, mean error, rank, …, benchmarking, flowcharting, control charts,…, inspection, etc. Error distributions that can be explained with stochastic rules can be even modelled (simulated), for example by Monte Carlo simulations.

Analyses and modelling of the cultural landscapes could be a high level uncertain process if we don't understand the nature of data well, and especially if there is not enough metadata available. Different kind of spatial analyses require various aspects of knowledge of the data sets and their properties.

This paper provides an empirical explanation of the effect of data uncertainty that influence particular modelling tasks (visibility, path modelling, catchment, analysis, etc.). Three selected approaches to uncertainty analyses and modelling of cultural landscapes in landscape

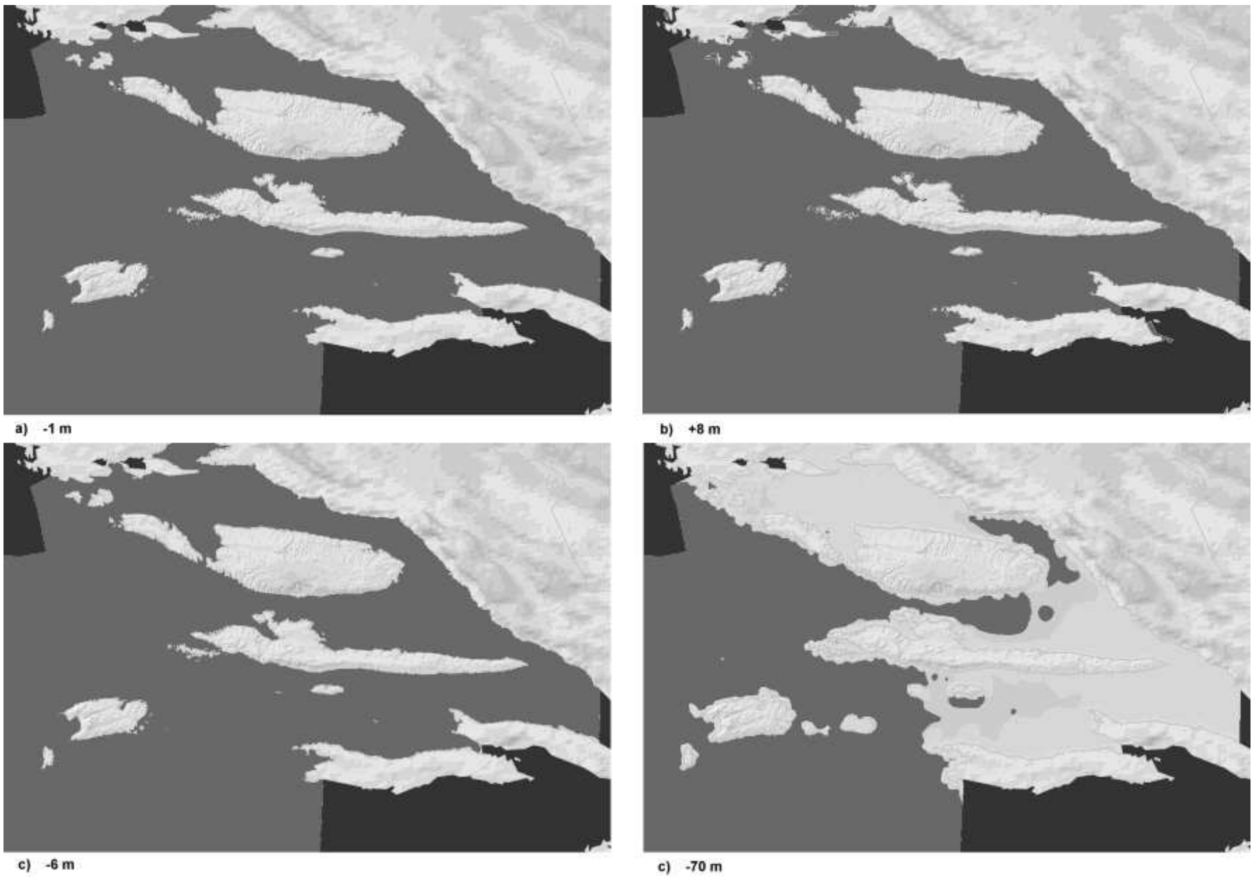

Figure 1. Coastline variation of simulated models of palaeo-landscapes for the Roman period, –1m (a), Bronze Age. +8m (b), Neolithic, –6m (c) and Palaeolithic, –70m (d). The darker area represents the area, in which the sea level was not modelled.
Slika 1. Spreminjanje obalne črte na podlagi modelov paleokrajin za (a) rimsko dobo, -1m, (b) bronasto dobo, +8m, (c) neolitik, -6m in za paleolitik, -70m. Višina morja (obalna črta) ni bila modelirana na potemnjenih območjih.

archaeology are exposed: (1) influence of the sea level changes on the interpretation of various ancient human activities, (2) analyses of the optimal paths algorithms and reliability of the data sources for interpretation of the ancient paths between settlements, and (3) analyses of the visibility from the significant points and visual communication interpretation. The case study areas are from Croatia (Central Dalmatia), Slovenia (Dolenjska) and Mexico (Yucatan, Campeche).

Problem 1: Information on sea level – Central Dalmatia, Croatia

The estimation and simulation of the sea-level regarding present situation for the selected historical periods: Upper Palaeolithic, Neolithic, Bronze Age, and the Roman period, was performed. The main objective of the unpublished paper (Podobnikar et al. 2003) was to discuss sea-level data for the Eastern Adriatic coast, Central Dalmatia, including its uncertainty and deviations, for the last 15,000 years. The paper focused on modelling the average sea-level of the Adriatic basin in order to recognize if and how its dynamics could alter the human activity areas from the Neolithic, through the Bronze and Iron Ages, to the Roman period. The aim was twofold: to observe settlement patterns over their likely past sea-level shape, and to verify regularity of spatial analyses and re-evaluate corresponding hypotheses regarding data sources that correspond to recent environment and water source data.

Many different spatial analyses have been performed and interesting results were obtained for the area of

Central Dalmatia (e.g. Stančič, Kvamme 1999). The analysis and results were promising; nevertheless these spatial analyses were all based on the present coastline. Our further idea was that a paleo-landscape– more particularly the coastline changes–might be an important factor of reliability of the archaeological spatial analysis (Podobnikar et al. 2003). For modelling of the sea-level changes, a digital terrain model (DTM) and a model of the sea bed surface (bathymetrical DTM) were acquired and produced from different data source. These were then integrated into one digital surface model (see the technique in Podobnikar 2005). The global sea-level model indicates 14 m of difference between the Neolithic (–6m) and Bronze Age (+8m on average). According to the nowadays sea-level we can – according to Rohling et al. (1998) – note variations from 1 and up to 8m. With such amplitudes, we can expect substantial landscape modifications, which can alter our perception of territorial dynamics in the past. The result of the sea level change on the coast line is presented in Figure 1.

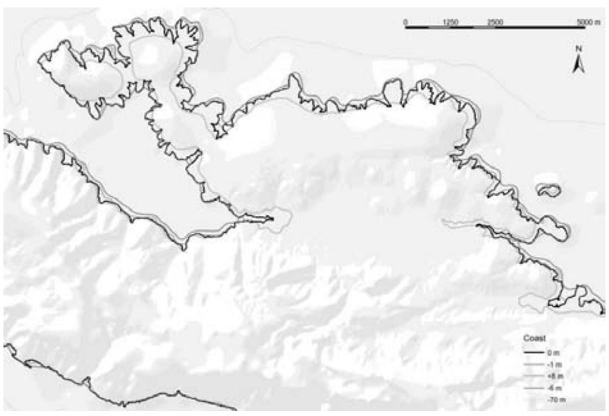

Figure 2. Coast line variation of different periods in the central part of the island of Hvar.
Slika 2. Spreminjanje obalne črte za različna obdobja, za primer osrednjega dela otoka Hvara.

Under these conditions, we can assume that sea-level change has had a significant impact on settlement organisation, especially in the present relatively shallow sea bed areas. This analysis was performed within a geographical information system (GIS). The results verified that the settlement pattern was not noticeably affected by sea-level changes in the last 10,000 years due to the fact that most of the islands had rather steep slopes which resulted in a low coastline variation. However, in the case of the Starigrad plain on the island of Hvar (Figure 2), with a large plain with a low elevation in respect to the sea, a significant impact of sea-level change could be observed. A different fractal dimension of the modelled coast lines that is visually notable (Figure 2) is additionally partly resulted from various qualities of the input data sources for the DTM production that is also a potential source of the uncertainties in the modellin

However, it was discovered that the estimated sea-levels based on the global model were not valid within the project area, before the paper was published. This error in the model negated the neat archaeological theory concerning past human activities in the paleo-landscape in the Central Adriatic Islands for the past 15,000 years, which was constructed from the results (see Podobnikar et al. 2003). The problem was that no archaeological studies were found during further investigations, which supported a sea-level of 8m higher than the present Mediterranean coastline during the Bronze Age. It is confirmed by historians and archaeologists that the recent coastline had effectively been reached by the Late Neolithic in the east Mediterranean. Nevertheless, we should consider minor or occasionally dramatic, local variations (Mitrovica 2003), which are also evident from observation of the remains of certain Roman villas. The main problem in the presented modelling and interpretation failure of the presented study seemed to be that a global model of the sea-level changes was applied, which deviates significantly from the local situation (Adriatic Sea). However, for the more detailed analysis of the ancient coastline, the tectonic movements and dynamics of the sediments depositions of Neretva River should at the very least be analysed using different techniques.

Problem 2: The influence of DTM quality on path simulation – Dolenjska, Slovenia

Within this study the simulation of possible courses of the ancient paths regarding environmental characteristics were performed (Podobnikar et al. 2007). The simulation enables the variation of different parameters, for calculating the path between selected points. The algorithms were applied with spatial analysis in GIS. On the basis of similarities of archaeological remains, we assume strong contacts and exchanges between the settled centres. Our assumptions of the optimal paths modelling were the fol-

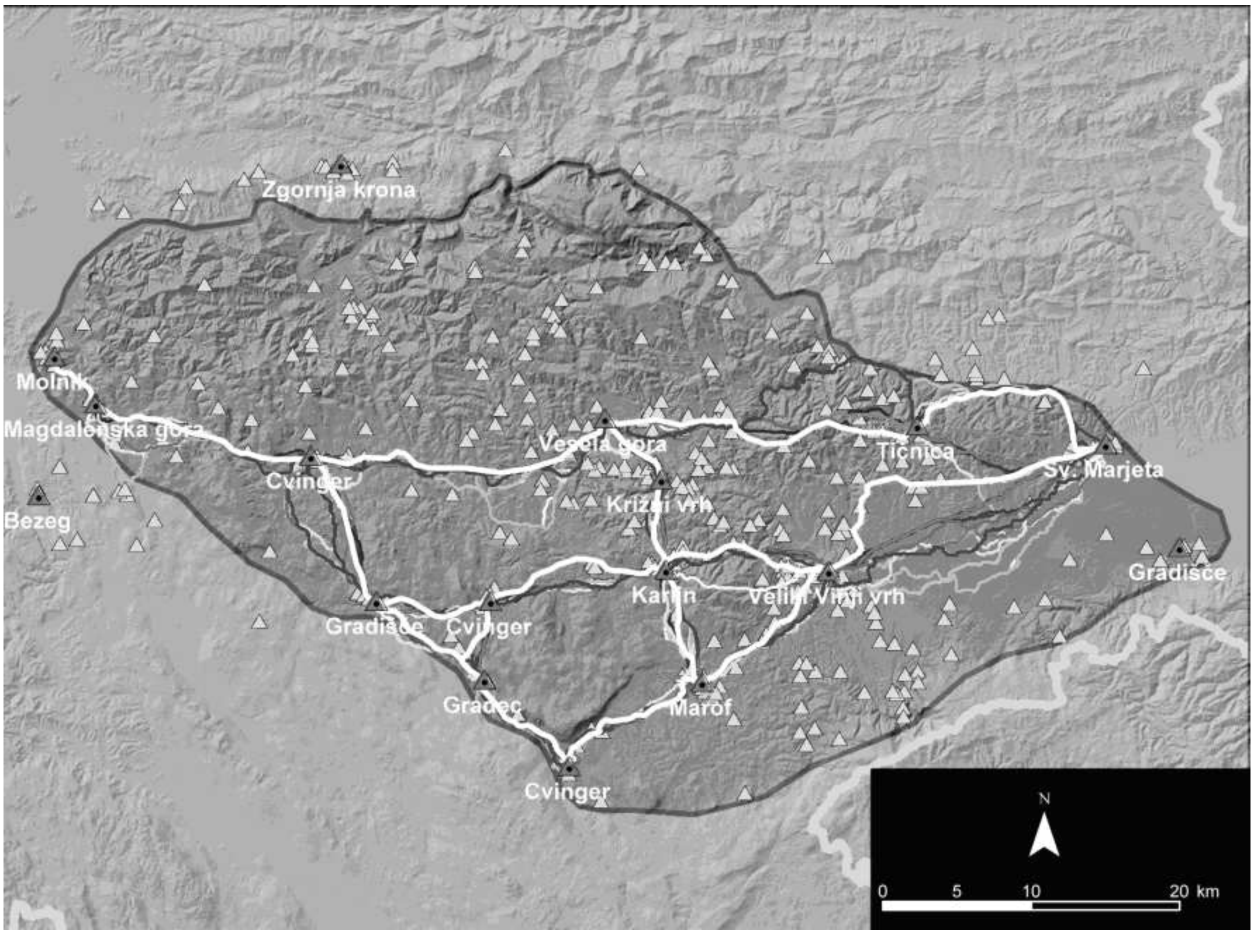

Figure 3. Simulation of ancient paths using nine different cost surface models. On the flat areas, the possible courses are more dispersed. The bold white line illustrates paths that were determined by the subjective knowledge of archaeologists, based on standing monument survey.
Slika 3. Simulacija zgodovinskih poti pri implementaciji devetih različnih modelov stroškovnih ploskev in uporabi DMR-ja. Poteki možnih poti so bolj razpršeni na ravnih območjih. Poudarjena bela črta predstavlja poti, določene na podlagi znanja arheologov.

lowing (Podobnikar et al. 2004, 254): (1) predominantly hilly terrain as the most important factor, (2) climate regime and other environmental conditions such as marshy soil, size of the rivers, vegetation, geology could also be of importance; however, we suggest that these conditions have not been changed drastically to the present day, (3) evidence of boats for river movement and wagons are not available, so we assume the settlers used to walk between the chosen settlements within the landscape, and (4) the impact of cultural and socio-economic factors on the decision making process.

The chosen study area was the central-southern region of Slovenia, Dolenjska (Lower Carniola), which is limited by natural boundaries of the river Krka on the south-west side and the river Sava on the north-east. 14 Iron Age settlements from a total of 38 documented in this area were selected and the presences of 18 paths that run between them were studied. The paths' simulation was performed just based on the DTM which seemed to be the most important factor. All of the algorithms base on least-accumulative-cost distance over a cost surface. In order to evaluate the results nine sets of cost surface parameters within the simulation were selected (Podobnikar et al. 2004, 255). Some results of these simulations can be inspected in Figure 3.

The analysis of the results indicated that it was possible to recognise that simulated paths on the hilly areas are less dispersed and denser than in less hilly areas as such

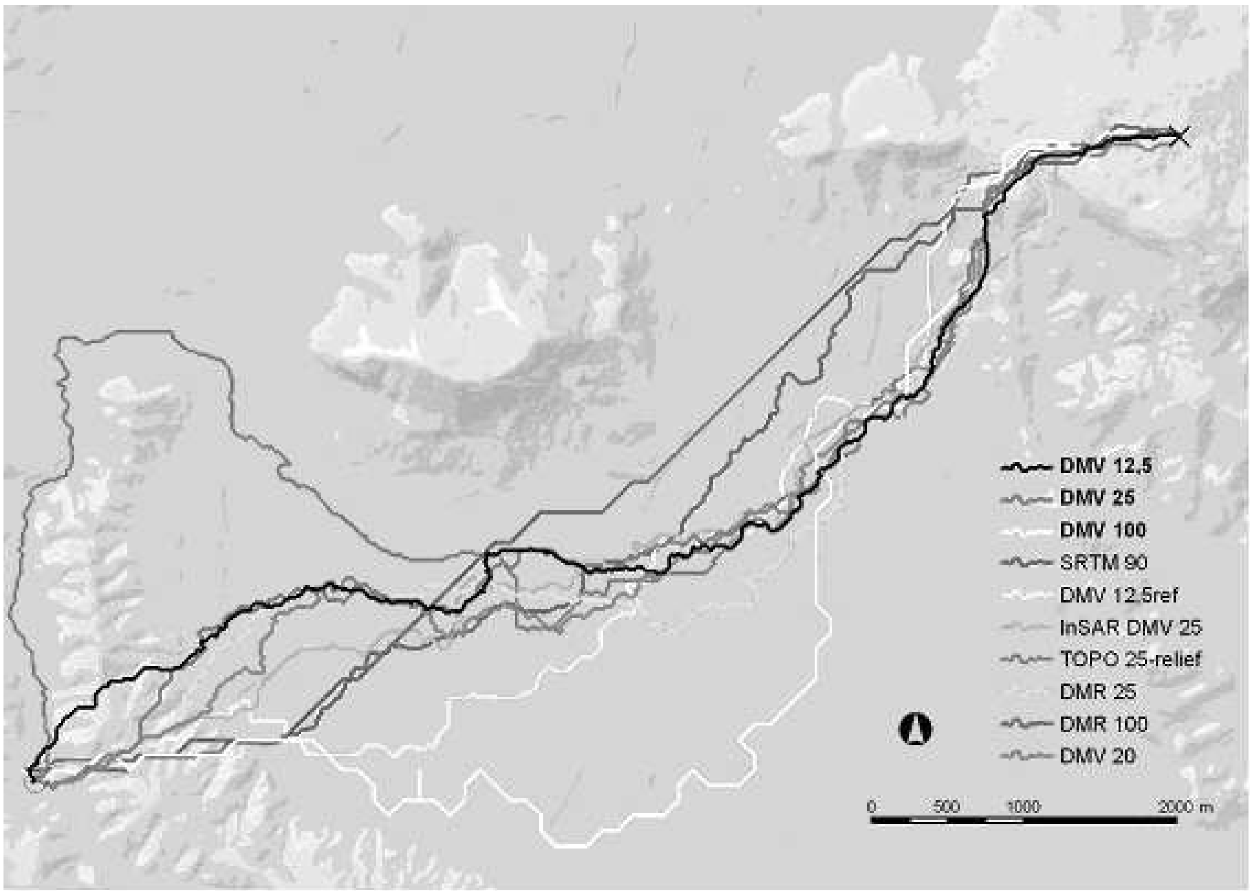

Figure 4. Simulated paths determined from DTMs with different quality using the same path determination algorithm and parameters.
Slika 4. Simulacija poti na podlagi DMR-jev različnih kakovosti pri uporabi istega algoritma in parametrov.

areas offer fewer possibilities for economical transport. In less hilly areas, other factors are more important than terrain. The archaeologist can subjectively interpret which simulated algorithm or which paths are more relevant. On the base of the presented simulation some segments of ancient path in Požarnice were found by archaeologists soon after implementation of the algorithm. However, the excavated material has not yet been completely analysed, therefore they are not yet confident that the segments definitely represent an Early Iron Age path (Tica 2003; Topličanec 2006; Dular, Tecco Hvala 2007, 229–230). This approach shows the importance of parameter variation. The visualisation of all results within one figure allows inspecting possible areas of uncertainties in the simulation results.

However, it has to be considered that for a more reliable study the significance of further factors influencing the run of the paths might be necessary. The importance of the factors fluctuates in respect to the different geomorphology of the area. Within this study we could show with the empirical tests that simulations allow detecting the paths quite reliable. Some of the significant influence factors are: (1) stability of the starting point, depending on the geomorphology; in some cases small changes can lead to a different path, (2) selection of the appropriate algorithm for path calculating, (3) selection appropriate cost surface weights in the algorithm and other constrains (e.g. what kind of path is it, what kind of transport was there), (4) the path from one point of the other is newer the same to the reverse path direction (anisotropic), (5) gross errors on the DTM can play an important role (Podobnikar 2008, 146), (6) the various quality of applied DTMs (considered without of gross errors) and chosen resolution can lead to differently simulated paths (cf. Figure 4).

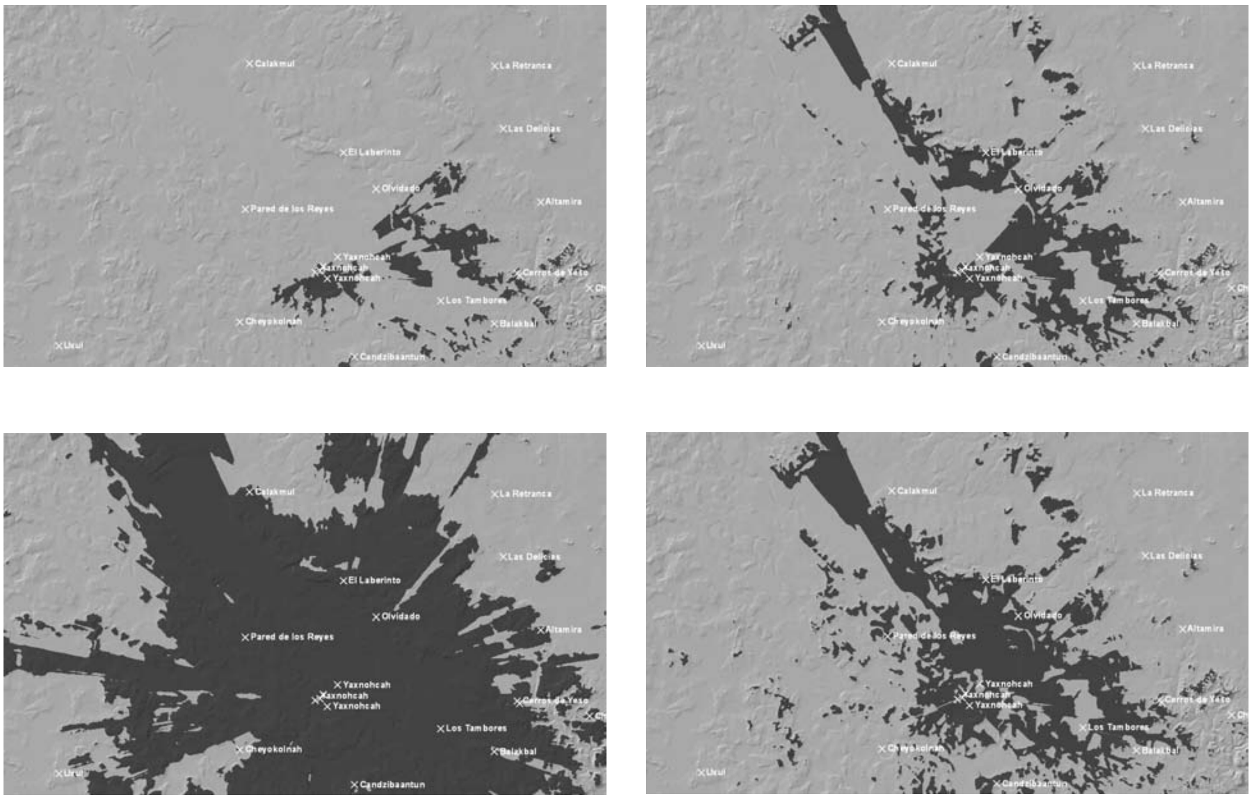

Figure 5. Viewsheds from Yaxnohcah Structure B-3 with different offsets, showing the visibility potential of the site (presented as a rectangle with dimensions of approx. 60 by 40km). Offsets (from/to) are (a) 2/2m, (b) 10/2m, (c) 30/30m and (d) real heights of all selected pyramids (20)m for Yaxnohcah Structure B-3) with offset 2/2m.
Slika 5. Izračuni vidnosti iz Strukture B-3 Yaxnohcaha ob upoštevanju različnih nadvišanj nad terenom, kar kaže na potencial vidnosti z določene točke na območju, predstavljenem kot pravokotnik z merami pribl. 60 krat 40 km. Nadvišanja (od/do), so (a) 2/2 m, (b) 10/2 m, (c) 30/30m in (d) dejanske višine izbranih piramid (20 m za Yaxnohcah Struktura B-3) z dodatnim nadvišanjem 2/2m.

Problem 3: Influence of height uncertainty of pyramids and vegetation cover on visibility studies – Yucatan, Mexico

The visibility studies of the archaeological sites can contribute to the understanding of several questions: Is visibility from archaeological sites connected with the landscape perception? Are natural aspects of site locations significantly different from other (average or randomly selected) areas? How did the visibility and inter-visibility influence the location choices? Are there any significant spatial patterns? Was it possible to communicate between the neighbouring centres or to control others from the tops of the sites? How much of the area was controlled from the sites? Are the answers to those questions clear and significant?

Our primary assumption was that visibility played an important role in settlement patterning in the Maya Lowlands central parts of the Yucatan Peninsula in Mexico (Podobnikar, Oštir 2008, 254). Since the area has so far been practically unexplored, the final results represented a significant contribution to the understanding of the Maya culture in central lowlands of the Yucatan peninsula. Visibility studies have been performed and intervisibility between centres has been determined. Viewsheds from all of the sites, with different offsets of heights for observation and observed (target) points, helped us to understand the terrain characteristics, possible visual communications between different centres, reasons for the chosen centres locations, etc. We have proven that positions of the centres chosen by the Maya lie significantly in areas that are more visible than the average (random) study area.

The visibility analysis requires a high quality DTM, information about heights of the buildings – mostly pyramids (53 from 40 major, medium and minor centres) that are used for observations – artificial outlooks and information about vegetation cover. Visibility maps from all 53 structures were calculated. Figure 5 presents the results of these simulations using different parameters for Structure B-3 of Yaxnohcah. Figure 5a shows visibility from the ground with offsets of 2m (2/2m – for observation/target or from/to points) applied to the observation and target point; it reveals that the site's visibility potential was rather poor before the pyramid was built. Figure 5b with offsets 10/2m presents visibility from small pyramids or trees about 10m high; we can see that large areas to the NW have become visible. Picture c shows an extreme case with offsets of 30/30m, when the observer is located in a tree top and observes the tops of other trees. The examples illustrated in Figure b and c are useful for a better understanding of the natural terrain characteristics regarding chosen observation points and data quality. They can also help us to understand the why the founders of the particular sites have chosen their positions and what they might have possibly seen from the trees before they built the settlements. Figure 5d, shows a situation after the pyramid with a height of 20m was built and furthermore assumes that the vegetation was cut down in order to produce viewsheds that are as realistic as possible.

Visibility analysis helps to understand the importance of visual communication. This seemed to be very important for the Maya settlements on Yucatan peninsula. The basic data were the DTM and heights of the most dominant pyramids. The most uncertain variable was vegetation. The visibility was simulated with or without vegetation consideration from the areas of pyramids or their tops. The study shows on one hand a quite high sensitivity to possible errors on DTM (see also Podobnikar 2008) or heights of pyramids for some sites and on the other hand it describes the intervisibility between some pyramids that can not be tested in reality, due to the dense modern vegetation that cover the selected area. Namely it seems that some pyramids are positioned on special positions that allowed visual communication, as it was shown by different analyses. However, it is impossible to validate these results in field. The main problems that increase uncertainties of the model interpretation are therefore potential lack of quality of the DTM, inadequate classification of the pyramids according to their functionality, and uncertain or non-stop changed vegetation coverage.

Conclusion

In the study, three different examples were explained and possible problems of the results were analysed. (1) The analysis of the Bronze Age settlements and their cultural landscape show that just slightly wrong information on a sea level in the past, can produce unpredictable results on spatial analysis and at the same time deny all the reliable looking hypotheses behind. The main problem was therefore the sources of the information and additional potential the unpredictable influences of the DTM quality to the coastline modelling that should be both carefully evaluated for the reliable results of such studies. (2, 3) Analysis of the ancient paths or visual communications shows potentially unexpected results, therefore they need additional verification. In our case the simulation of different possible parameters was performed. The results show that in some cases just slightly different coordinates from the present locations of starting and target points of paths or points of observation can introduce completely different solutions. Additional problem can be uncertain quality of the used DTMs. One of the reliable solutions of such problem might be in (Monte Carlo) simulations of different parameters and finding the most feasible result(s).

Quality of the applied data, either spatial datasets or archaeological evidences is a crucial factor for the reliable results in modelling. The chosen modelling techniques are therefore less important than the quality of the data. Significant and not enough respected value in spatial modelling are availability of the data/information sources and especially their quality. Cost of the data is nowadays much higher than of the hardware or software and even of the management of the whole system.

What can be the solutions of the presented problems? The described examples show the importance of understanding the data that are used for the certain application and algorithms behind the modelling. To enhance the understandability, the data/information and their implementation (procedures) to the models should be carefully studied. Not perceived gross or systematic error or not assessed random error may cause unexpected and even wrong results of the particular study. Decision maker should be then able to do a relevant interpretation of his work. An important problem presented in this study is availability of appropriate quality data/information. It is much bigger than the quality of the numerical algorithms, models or simulations. There are many empirical and numerical tools and techniques for quality assessment, which are presented in the introduction, but one of the first techniques that might be applied is "rule of thumb" that or the best implementation needs a lot of experiences. Further complex techniques includes study of the metadata with the data profiling, including sampling, applying the ETL (extract, transform, and load) functionalities with quality assessment/control techniques or even total quality management approaches.

Posledice negotovih podatkov in informacij pri prostorskem modeliranju kulturne krajine (Povzetek)

Kakovost podatkov in informacij je odvisna od velikega števila elementov, ki lahko pomembno vplivajo na rezultate prostorskega modeliranja in interpretacije v arheologiji. Vsi elementi potrebujejo skrbno obravnavo, ki nam omogoča celovito spoznavanje narave obravnavanega problema. V študiji so bila analizirana izbrana območja kulturnih krajin Hrvaške (Srednja Dalmacija), Slovenije (Dolenjska) in Mehike (Jukatan, Campeche).

V okviru arheologije krajine smo razčlenili tri probleme negotovosti rezultatov analiz in modeliranja kulturne krajine. (1) Analizirali smo vpliv spreminjanja višine morja pri interpretaciji preteklih človekovih dejavnosti, kot so bronastodobne naselbine. Prikazali smo, da je rezultat na podlagi analiz prostorskih podatkov lahko nepredvidljiv in popolnoma napačen, četudi je informacija višine morja za obdobje bronaste dobe le malo različna od pravilne. Dodatno vpliva tudi kakovost digitalnega modela reliefa (DMR) na podlagi katerega lahko določamo obalno črto v povezavi z višino morja. Pod današnjo morsko gladino je DMR praviloma nenatančno izdelan, poleg tega je bilo paleo površje drugačno od današnjega. Ob vsem tem se na podlagi napačnih rezultatov lahko še tako verodostojna hipoteza izkaže za neprimerno.

Nadalje smo analizirali (2) algoritme za iskanje optimalnih poti in zanesljivosti virov podatkov za interpretacijo starodavnih poti med naselbinami in (3) vidnost z značilnih točk ter interpretirali vizualne komunikacije. Pri omenjenih nalogah so se izkazali za pomembne izbor parametrov algoritmov, natančnost določitve koordinat začetka in konca poti oz. stojišč (še posebej na razgibanem reliefu) ter kakovost DMR-ja.

Opisani primeri kažejo na pomembnost razumevanja podatkov, ki jih uporabljamo za izbrane aplikacije in algoritme v sklopu modeliranj. Kakovost uporabljenih prostorskih podatkov ali arheoloških dokazov so porok za zanesljivost rezultatov prostorskih analiz.

References

BURROUGH, P. A., R. A. MCDONNELL 1998, *Principles of Geographical Systems Information Systems*: *Spatial Information Systems and Geostatistics*. Oxford University Press.

DULAR, J., S. TECCO HVALA 2007, *South-eastern Slovenia in the early Iron Age: settlement, economy, society = Jugovzhodna Slovenija v starejši železni dobi: poselitev, gospodarstvo, družba*. Opera Instituti archaeologici Sloveniae 12, Ljubljana.

MITROVICA, J. X. 2003, Recent controversies in predicting post-glacial sea-level change. – *Quaternary Science Reviews* 22, 127-133.

PODOBNIKAR, T. 2005, Production of integrated digital terrain model from multiple datasets of different quality. – *International journal of geographical information science* 19/1, 69–89.

PODOBNIKAR, T. 2008, Simulation and representation of the positional errors of boundary and interior regions in maps. In: A. Moore, I. Drecki (eds.), *Geospatial vision: new dimensions in cartography: selected papers from the 4th National Cartographic Conference Geocart*'2008, New Zealand. – Lecture notes in geoinformation and cartography, Berlin, Heidelberg, 141–169.

PODOBNIKAR, T. 2009, Methods for visual quality assessment of a digital terrain model. – *S.A.P.I.EN.S*, Special Issue, 2/20, http://sapiens.revues.org/index738.html.

PODOBNIKAR, T, C. BRIESE, M. DONEUS, N. PFEIFER 2007, Uncertain observations of the Past with a modern outcome. In: W. Börner (ed.). *Archäologie und Computer*, 5.-7. November 2007: Workshop 12 : Kulturelles Erbe und neue Technologien , Wien, 1–9 .

PODOBNIKAR, T., L. NUNINGER, T. VELJANOVSKI, Z. STANČIČ, V. GAFFNEY 2003, *Past human activities concerning palaeo-environment in the Central Adriatic Islands*, 23 pp. (unpublished paper).

PODOBNIKAR, T., K. OŠTIR 2008, Geographical information systems and remote sensing analysis. In: I. Šprajc (ed.). *Reconocimiento arqueologico en el sureste del estado de Campeche*, Mexico : 1996-2005, BAR international series, Paris monographs in American archaeology, 1742, 19. – Archaeopress 243–261.

PODOBNIKAR, T, S. TECCO HVALA, J. DULAR 2004, *Iterative approach to ancient paths modelling in the Iron Age study of the Dolenjska region* (Slovenia). In: K. Fischer Ausserer (ed.), Enter the past: the e-way into the four dimensions of cultural heritage: proceedings of the 30th conference, BAR international series, 1227. – Archaeopress 254–257.

ROHLING, E. J., M. FENTON, F. J. JORISSEN, P. BERTRAND, G. GANSSEN, J. P. CAULET 1998, Magnitudes of sea-level lowstands of the past 500,000 years. – *Nature* 394, 162–165.

STANČIČ, Z., K. KVAMME 1999, Settlement patterns modelling through Boolean overlays of social and environmental variables. In J. Barcelo, I. Briz, A. Vila (eds.), *New technique for old times*. – Archaeopress 16 pp.

TICA, G. 2003, Požarnice pri Družinski vasi. In: B. Djurić (ed.), *Zemlja pod vašimi nogami. Arheologija na avtocestah Slovenije. – Vodnik po najdiščih*, Ljubljana, 229-231.

TOPLIČANEC, M. 2006, Rimska in prazgodovinska cesta pri Požarnicah: rimska cesta Emona-Neviodunum-Siscia. – *Rast* (104) 17/2, 189–190.