sciendo

Journal of Environmental Geography 13 (1–2), 1–12.

DOI: 10.2478/jengeo-2020-0001
ISSN 2060-467X

SPATIAL DEVELOPMENT OF THE 1507–1510 PLAGUE IN POLAND AND ITS CONSEQUENCES. KALISZ DISTRICT CASE STUDY

Tomasz Związek*

Department of Geoecology and Climatology, Stanisław Leszczycki Institute of Geography and Spatial Organisation, Polish Academy of Sciences, Twarda 51/55, 00-818 Warsaw, Poland
*Corresponding author, e-mail: tzwiazek@twarda.pan.pl

Research article, received 27 December 2019, accepted 28 February 2020

Abstract

The article discusses the causes and effects of the plague which is said to have spread over many Polish towns in 1507. The focus is on its possible causes, related to the occurrence of droughts and floods in Central Europe in the late 15th and early 16th century. Available sources from the late mediaeval period have also been analysed for the recorded perceptions of the extreme climatic and weather conditions. Special attention has been paid to the issues of intensity and spatial distribution of the effects of the plague on the example of one district. The analysis covered a variety of issues such as settlement changes, prices of basic goods or even some pollen data. The main results of the study indicate that the climatic extremes at the turn of the 16th century exerted a long-term impact on the society and economy of the region. They also contributed to the abandonment of settlements on rural sites.

Keywords: historical geography, plague, Late Medieval Poland, spatial distribution, droughts, floods, weather extremes

INTRODUCTION

According to the present state of research, the lands of Poland were not affected by the Black Death in the 14th century (Guzowski et al., 2016), however various parts of the country were regularly struck by plagues, which appeared and subsided at quite steady intervals (Karpiński, 2000; Walawender, 1957). According to scholars, their outbreaks are connected to the occurrences of natural disasters, such as droughts, floods, hail, flooding, frosty or warm winters, which contributed to crop failures, rise of food prices and, consequently, famine, which weakened the overall condition of the society (Karpiński, 2000). A good example to illustrate this point was the great plague of Cracow, which was probably a long-term consequence of the drought of 1540 and which in the period between July and November 1543 took the lives of at least 10 thousand persons (which constituted roughly 1/3 of the city's population) (Follprecht and Noga, 2014; Sowina, 2016). Similarly, in Western Europe, numerous periodic plagues appeared at steady intervals and were also likely related to various weather extremes or climatic changes. One particularly difficult time, abundant in plagues, was the period of armed conflicts in the 17th and 18th centuries, which was also the coldest period of the Little Ice Age (Maunder Minimum) (Alfani, 2013; Alfani and Percoco, 2019; Curtis, 2016; Owens et al., 2017). Records contain accounts of economic and demographic consequences of the plagues (Campbell, 2016; Cummins et al., 2016), but attention should also be paid to their socio-cultural dimension (Martín-Vide and Barriendos Vallvé, 1995). Thus, the investigation of the causes and effects of the plagues should be considered one of the fundamental topics in the study of the history of former societies as well as the history of climate and environment. However, recent works have shown that the pre-industrial society (mainly agrarian) was able to deal with crises quite efficiently (Büntgen et al., 2011). They became problematic only when a sudden accumulation of extreme events (natural disasters, famine, wars, etc.) took place, which disturbed every-day balance of the society (Haldon et al., 2018).

Study of extremes and their potential consequences – plagues – are, therefore, complex and interdisciplinary by nature, touching upon various research fields. The basic ones include studies on the problem of disease spread, social reactions to plague, temporary practices of depopulating villages, towns, etc. (Lagerås, 2016). Such studies refer both to the spatial and to the environmental aspects, all of which were as important as the economic elements. The fundamental question which thus arises concerns the relationships between individual crisis events, both climate- and disease-driven.

The main aim of this article is to discuss three major research problems: 1) the extension of the plague in the context of the climatic and weather extremes which contributed to it; 2) the changes to the spatial distribution of the plague on the basis of a case study of the district of Kalisz; 3) the duration of the plague and its impact on the economy of the said area. Answers to the formulated research questions will be given based on the analysis of some of the available materials from the southern part of the Greater Poland region (Fig. 1). For this area, some of the oldest preserved tax registers are available – for the years 1507–1510 – offering lots of data concerning the spread of the disease (Związek, 2013). As regards the extension of the plague, I am assuming that its level can be calculated using a couple of indices: 1) the number of deceased and isolated

people; 2) the number of fugitives from urban and/or rural sites; 3) the degree of abandonment of fields and villages; and 4) the fluctuations of the economic indicators.

The plague of 1507 – written evidence

Chronicles feature numerous pieces of evidence confirming that in 1507, the Kingdom of Poland was struck by a plague which affected many towns („*per multa oppida Poloniae pestis erat*") (Walawender, 1932). The plague reached as far as Vilnius (in the Great Duchy of Lithuania); it was also recorded in far Rus' – in Novgorod – and in Silesia, where it appeared in 1507 and slowly captured subsequent parts of the region together with the towns located there. It struck Wrocław (Breslau) and its vicinity, claiming lives of many people between September 1507 and the beginning of January 1508. Examples based on chronicle mentions can be easily multiplied, but they cannot represent the exact severity of the occurring disease (apart from the frequently exaggerated narratives). Neither are we able to specify its nature. It could have been the pestilence or typhoid – historical sources are very imprecise in this respect (Karpiński, 2000). The severity of the disease given in the narrative sources as well as in sources related to taxes seems to suggest that this was a contagious disease, the germs of which were transmitted between humans and/or animals. One of the chronicles of Novgorod recorded in 1508 that in this year in the city over 5000 men and women died („*w lěto 7016, pomre ljudi mužska połu i ženska tma i 5000 dusz' i 400 bez czetyrech czełowěk*", (Walawender, 1932). The number should not be taken literally but rather as a proof that this was indeed a great plague, exceeding the comprehension abilities of the contemporaries. In this respect, chronicle sources are by no means able to offer us a precise picture of the epidemic.

Plague as a possible consequence of climatic and natural hazards

According to Christian Pfister's preliminary model of the economic impact of climatic events (Pfister, 1988), it is very likely that the outbreak of the plague at the beginning of the 16th century in Poland could have been, at least partly, a long-term effect of weather events in the preceding years. In short, they had created a climatic pressure (e.g. great hunger) which, in subsequent steps, could have led to another crisis, such as the plague (Ell, 1985; Kiss, 2020). The issue is, however, more complicated, and it is also connected with the land use changes over time as well as the scale and range of human impact on the environment during the period of the development of the so-called German Law in Central Europe. Recently archaeologists have pointed out that intensified land use, forest clearance as well as increase in crop production and breeding strongly affected the daily life of the Medieval and Early Modern societies (Schreg, 2011, 2019). In this model, intensive anthropopressure led to microclimatic changes, where floods, droughts or heavy rains were becoming more noticeable in daily life. For instance, extremes such as floods could have caused high soil erosion and bring about the long-term effect of food production drop, rendering the society more vulnerable to diseases. High water could have also caused serious pollution that was dangerous for people and animals, ultimately leading to development of typhoid fever, cholera, etc (WHO).

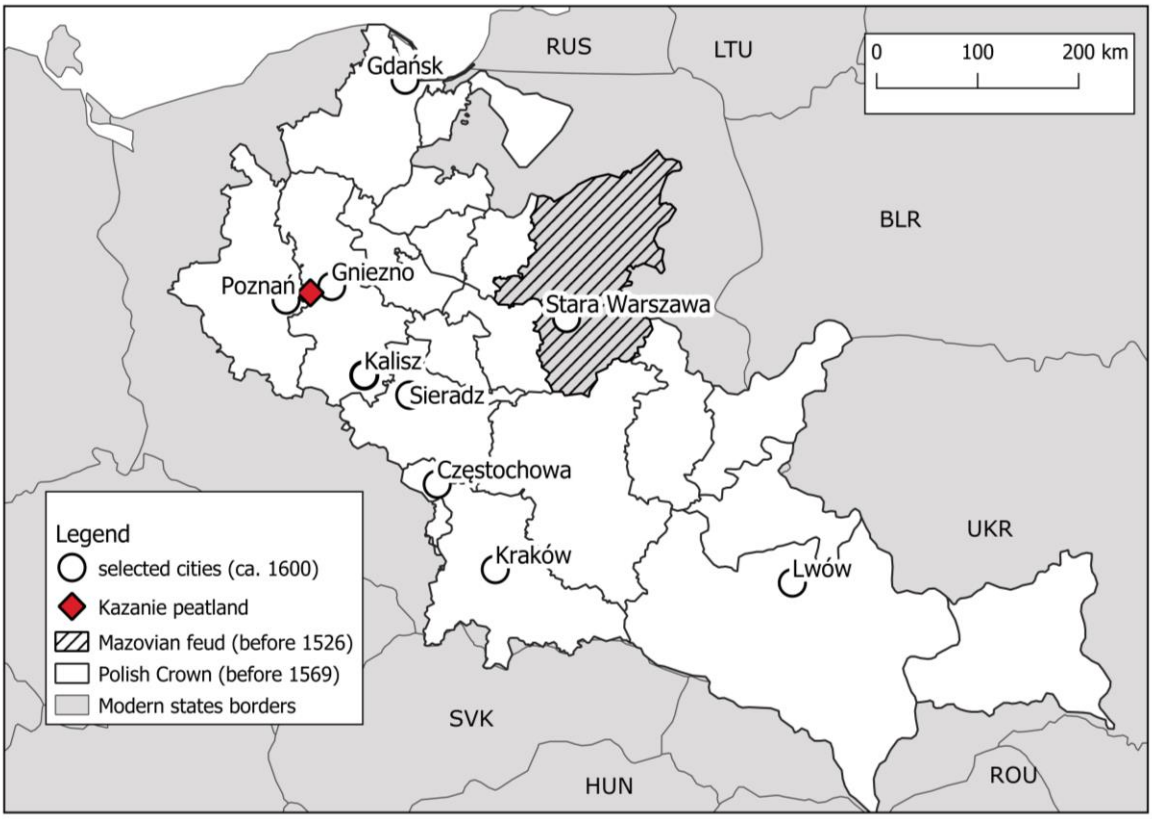

Fig. 1 Study area in the context of the past and contemporary Poland with modern neighbouring states. Basemap source from NaturalEarthData.com

Hungarian and Czech context

Recently Kiss (2017) has shown that some source materials on Hungary definitely point to the occurrence of a drought in 1502 and 1503. She has also indicated that the dry year 1502 had contributed to the aggravation of the consequences of the drought in the following year, which is especially visible in the Old World Drought Atlas (OWDA) model (Cook et al., 2015). This was a period of particularly destructive hail storms and crop failures – affecting cereals, vine and honey alike. In this case, as demonstrated by Kiss (2017), bees '*may act as indicator of weather-, and probably (spring-summer, or earlier) drought-related problems*'. On the other hand, data from the Eger diocese from 1506 have shown that many agricultural fields and vineyards were destroyed within its borders; the income from the tithes was significantly smaller; there was a shortage of fish. It has also been proven that subjects in general suffered from poverty. Thus, the calamities which affected Hungary at that time were complex, and their occurrence can be traced thanks to many types of written records as well as environmental indices (proxies).

On the territory of the Czech Crown, the beginnings of the 16th century were not recorded as an extremely dry period (Brázdil, Dobrovolný et al., 2013). Narrative sources mentioned very cold winters at the turn of almost every year and frequent spring floods. Like in the rest of Central Europe, in 1503 a 30-days drought and low level of crops were recorded. The subsequent years (1505, 1506 and 1507) were also described as periods of drought Brázdil, Kotyza et al., 2013). According to the decadal frequencies of droughts in the Czech territory (Brázdil, Dobrovolný et al., 2013), we can describe the beginning of the 16th century as a period of sinusoidal weather changes with frosty and long winters, frequent floods in springtime and repeatable drought period in summers.

SOURCES

The main sources which have been used in this study can be divided into two basic groups: written records and climatic/environmental archives (tree rings and pollen data). The written sources consist of narrative data (chronicles, memories), which can also be treated as literary texts; this means that they can be used in the context of climatological research only with difficulty due to the personal reception of the phenomena by the authors. The main collection of narrative texts used in this study came from the book by Walawender (1932). The main idea behind his research was to offer as comprehensive view on the climate history and changes as possible based on the contemporary set of published sources. The biggest advantage of this book is that the second part of each volume has been supplemented by the author with extensions from various source editions. Thanks to this, Walawender's works help one reach the basic material and compare his climatic interpretations with the original records written in old-Polish, old-German, Latin and old-Russian.

Another source of materials were the tax registers from the 16th century, which were obtained from the Central Archive of Old Records in Warsaw (AGAD). The registers from the districts of Konin and Kalisz were included in the archive unit no. ASK I 12. The tax registers contain information about the main settlement structure of the late Jagiellonian period of the Polish Crown (Słoń, 2011). They also provide a lot of valuable economic and demographic data on the rural sites and towns of the Early Modern state (Gieysztorowa, Żaboklicka, 1955; Boroda, 2007; Boroda, Guzowski, 2016). Nevertheless, the basic sources were the oldest, and unpublished, tax registers (1507–1510) from southern Greater Poland.

The environmental archives which were used in this study consist of the tree rings data model from the 'Old World Drought Atlas' (ODWA) (Cook et al., 2015) and preliminary pollen data from one of the peatlands in central Greater Poland (Czerwiński et al., 2019). The pollen data had been collected for one of historical and environmental projects in Poland, and at the time when I was doing my research, only one set of proxies (Kazanie peatland) had been elaborated and, therefore, available for me to use. The main difficulty with pollen data concerns the chronology of such materials. In the Kazanie peatland, the resolution of the core appeared in 10-year intervals, hence the analysis made it possible to look only at the beginning and the end of the plague period. Secondly, in general, the data from peatlands are strictly limited to the nearest study area, up to 5 or 10 km away from the reservoir (Birks, Birks, 1980; Theuerkauf, Couwenberg, 2017). For this reason, the final results are more local than in the case of analysis from lake cores. However, the pollen database from Polish lakes and peatlands, which is worthy of application in historical research (with proper ^{14}C chronology), is expected to grow (cf. with Izdebski et al., 2016)

METHODOLOGY

The main methodological idea behind this research was to understand, both in spatial and quantitative terms, the mutual relations between climatic or environmental events (in this case mostly droughts and floods as well as their short- and long-term effects) and the history of the past societies. In this approach, I intend not only to examine past events, but also to pursue some recent climate history research proposals (Camenisch, 2015) which deal with narrative sources as primary materials for the investigation of past climate changes. My main assumption has been to describe and analyse one fairly well-known example of a crisis using the methods of spatial and economic analysis and to compare the preliminary results with the accessible proxies from recent projects published in Poland. In this case, the final result would not be very comprehensive, mostly because of the lack of source materials and proxies at this stage of research. The general idea is also in line with the direction of landscape change studies connected with climate research (among others, the analysis of the process of abandonment of settlements).

RESULTS

OWDA – climatic event and tree-rings data

A comparison of the weather conditions in Hungary, Czechia and in Poland at that time (OWDA) indicates clearly that in the years 1501–1504, there could have been a long-lasting drought (Cook et al., 2015). This scenario

can be supported by weather observations from Cracow (Limanówka, 2001), which show that at that time, the weather was perceived as rather chilly (in the spring) or warm (in the summer). None of the spring days in the years 1503–1507 was considered extraordinarily warm, while days of heat in the summer constituted a tiny fraction of all the observations: in 1503 – 5%, in 1504 – 3%, in 1506 – 11% and in 1507 – 16%. The significant accumulation of dry years in the late 15th and early 16th century (Fig. 3**)** is also visible in the last studies using tree-rings data from Poland (Przybylak et al., 2019). Looking at the narrative sources in this period, there is much more data pertaining to floods mainly sudden ones occurring on the main rivers of the Kingdom (the Warta, the Vistula) (Walawender, 1932). It should be noted, however, that the information about the dry periods, even though theoretically more acute in Poland than in Hungary, did not translate to the reception of these events in the written records in Poland, in which there is a major gap between before 1500 and after 1505 (Fig. 3). In this period, the number of chronicle mentions was bigger for floods (Fig. 2) than for droughts, so one can assume that the perception of the Late Medieval Polish society could have been more sensitive to all water-connected weather events (such as floods, heavy rains, hail storms). Despite the hard weather conditions, the last decade of the 15th century could not be treated in Poland as a breakthrough (cf. with Camenisch, 2018). It is likely that especially the period of very intensive floods and droughts and related with them other weather extremes of approx. 1500–1507 can be treated as the starting point of the later plagues (Fig. 4) in the first decade of the 16th century (Kiss, 2019, 2020; Yue, Lee, 2020).

The weather extreme from the sight of the tax records

The long-term drought, confirmed in all Central Europe (Kiss, 2017, 2019; Kiss, Nikolić, 2015; Brázdil, Dobrovolný et al., 2013) together with numerous instances of flooding caused by sudden rains, had most likely resulted in the outbreak of the epidemic in the Kingdom of Poland and Great Duchy of Lithuania in 1505 (Walawender, 1932). Starting from this year, the plagues repeatedly struck the Polish territories in the subsequent

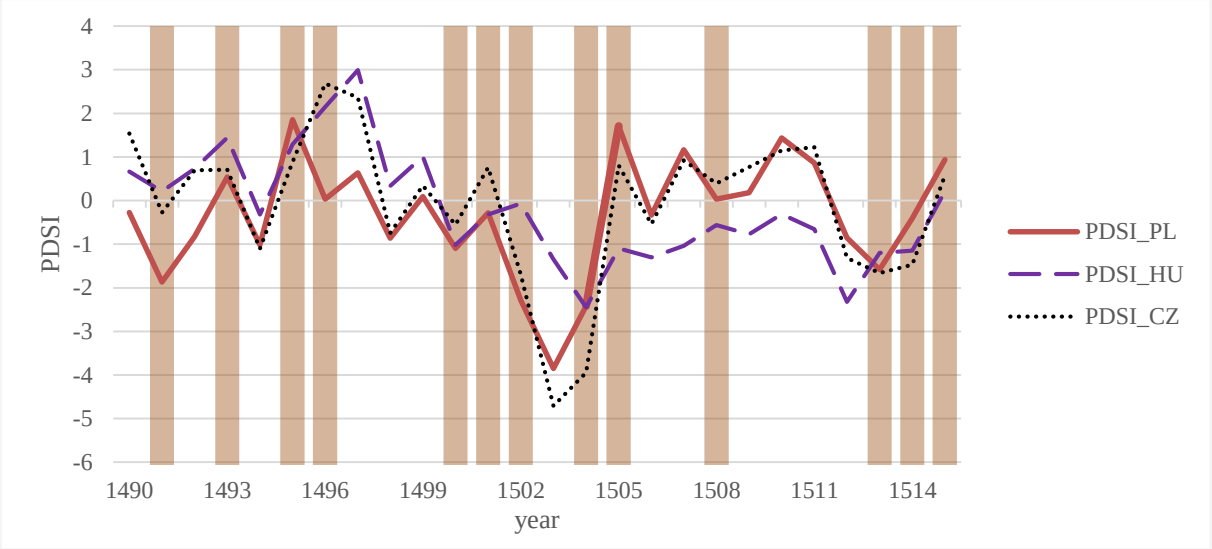

Fig. 2 Potential relationship between data from written records (vertical bars) on floods in Poland and OWDA. Sources (Cook et al., 2015) and (Walawender, 1932)

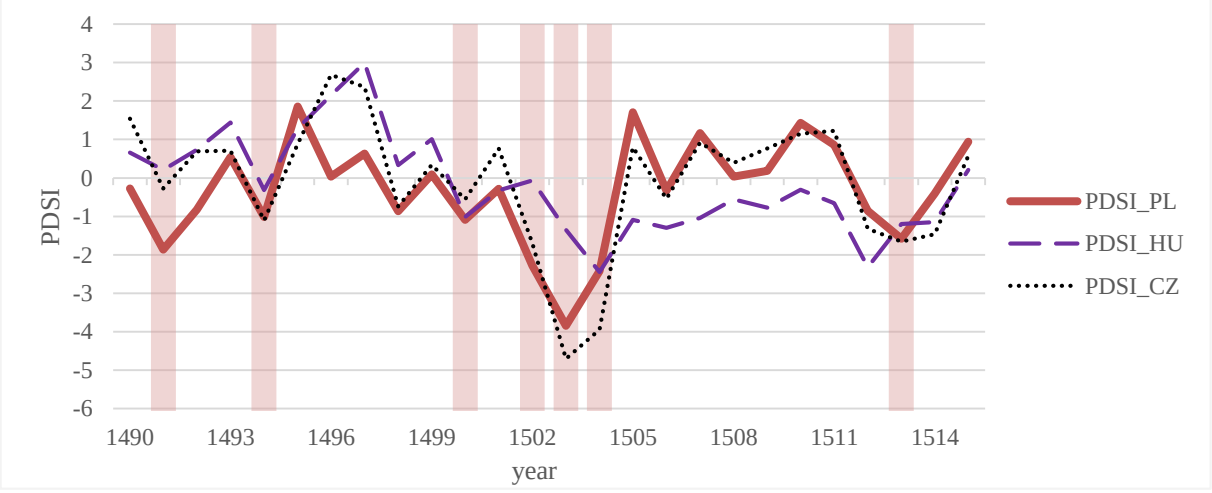

Fig. 3 Potential relationship between data from written records on droughts (vertical bars) in Poland and OWDA. Sources as in Fig. 2

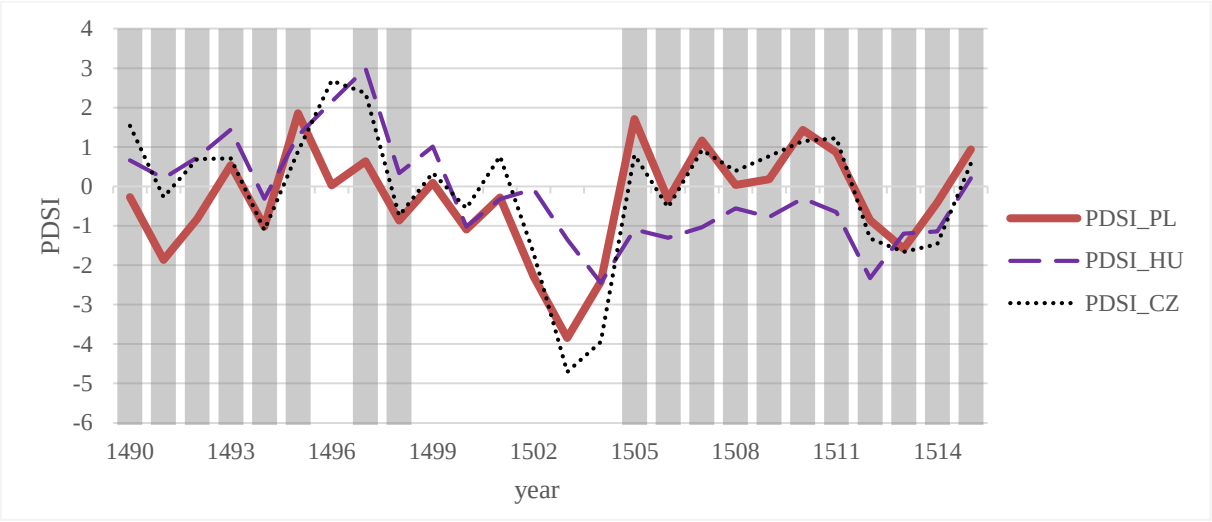

Fig. 4 Information from written records about diseases in Poland (vertical bars) with data on droughts or wet periods from OWDA. Sources as in Fig. 2

decades of the 16th century. When analysing the data from the district of Konin, one can clearly see that information about the disease was recorded in tax-related sources in three main ways: 1) direct mention about the plague striking villages (e.g. "*medius mansus per pestem desertus est*"); 2) mentions of peasants fleeing in connection with the disease ("*duo mansi … per profugos deserti sunt*"); 3) information about peasants dying ("*tribus mansis deserti sunt a tres annis per mortem kmethonis*"). Thus, we are dealing with consequences of a persisting disease given in sources *post factum*. Such data do not make it possible to determine specifically when the disease broke out. It is only clear that it happened in the summer-autumn months of the year.

The analysis of the yearly distribution of data from the period 1507–1510 (Fig. 5) from the district of Konin shows that the disease was most acute in the early autumn (September) of 1507 (Gochna, Związek, 2019). It was then that the greatest percentage of flights and deaths due to the disease was recorded, however it was not until one year later that the greatest part of arable fields became unused. In 1509–1510, the plague was clearly subsiding to practically disappear in the autumn of 1510. Observations based on the data from the Konin district suggest that in this case, we are dealing with consequence-laden diseases (in all likelihood, contagious), which in the view of the contemporary people were so lethal that they forced peasants to flee from their farms. Escape was generally quite a common means of dealing with plague; it was commonly practiced by the contemporary societies (Karpiński, 2000; Sowina, 2020).

In the context of the phenomena in question, however, the manner in which the contemporary tax offices of the state recorded such information seems to be of more importance. In the 15th and 16th century Poland, every village was obliged to pay mandatory, extraordinary taxes, which were set as lump sums. In general, taxes were paid on the basis of the cultivated land, forest clearance, water- and windmills. The amount of the main tax burden in relation to the income was not large (Guzowski, 2005), however failure to pay it was punished with imposition of a penalty upon the entire village (as a community of residents) which in the late 15*th* and early 16*th* century amounted nearly 60 times as much as the initial fee per one łan (Latin: *mansus*) of cultivated land (Gochna, Związek, 2019). For this reason, in the preserved tax registers, tax clerks paid a lot of attention to record the reasons for failure to pay the tax by individual villages. In tax registers from different districts, such information was recorded in various ways.

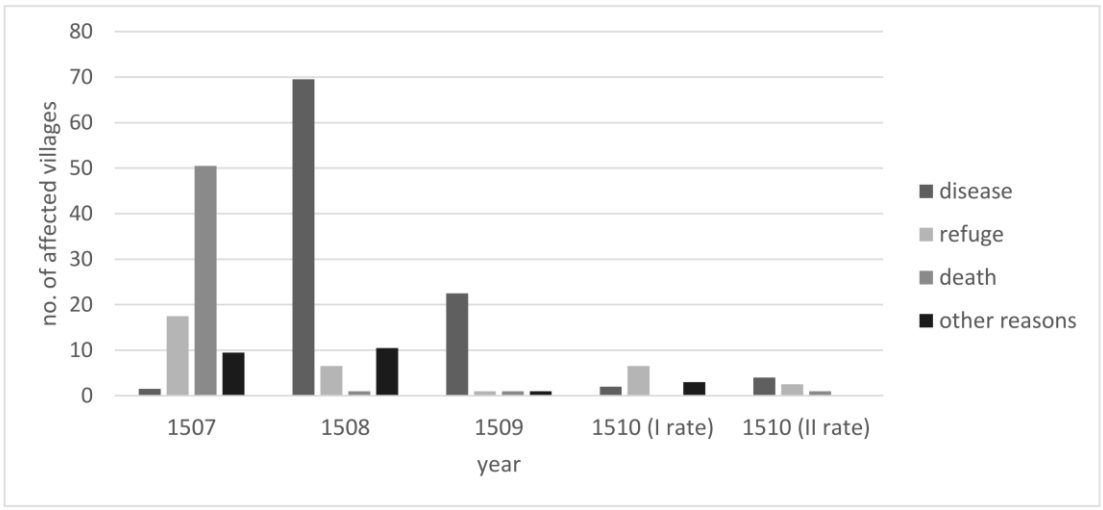

Fig. 5 Information about the disease in 1507–1510 in the Konin district with causes of field desertion. Source: AGAD, ASK I 12

While in the Konin district, each entry in the tax register was accompanied by an appropriate comment, in its southern neighbour – Kalisz district – such notes are not to be found. It is rather puzzling that as far as the registers from Kalisz are concerned, the source sheets often include blank fields (Fig. 6). An analysis of all such fields from the registers from 1507–1510 (Table 1) suggests that they should be interpreted as a direct testimony to the plague spreading over the district. The villages struck by the plague did not pay taxes because the peasants ran away (or possibly died) (*non habet kmethones*) or the villages were simply deserted (*desertum*). A quantitative analysis has made it possible to note that the plague in the vicinity of Kalisz – unlike in the Konin district – was a constant threat for 3 years, from 1508 to 1510, with practically identical impact on the settlement structures (approx. 30% of all villages remained empty).

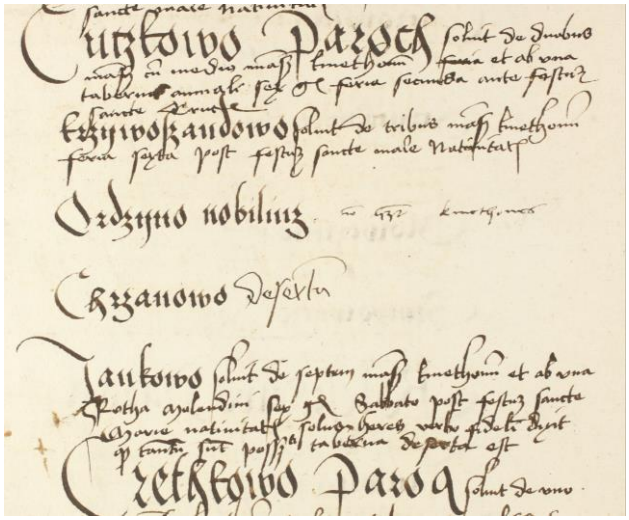

Fig. 6 Fragment of a tax register page from 1507 with two empty records. Source: AGAD, ASK I 12, f. 11v

Plague and the roads network

Tax registers also make it possible to study the spatial development of the plague. Using the available data concerning the settlement in the 16th century Poland (Polish Territories of the Crown in the 16th century. Spatial Database, 2015), one can trace the distribution of the villages affected by the plague in the period in question (Table 1, Fig. 7). Due to its transit character, the district of Kalisz communicated the areas of Lesser and Central Poland (provinces of Sieradz and Łęczyca) with western Greater Poland and German duchies; it also played an important role in the North-South dimension (from Pomerania to Silesia) (Związek, 2017). Wiping the plague out from the outbreak sites – especially given the high mobility at that time – was a very difficult task. The disease affected most acutely the central and eastern parts of the district. The persistence of the disease in this area should also be connected to the existence of a relatively dense and old settlement network, which originated in the early Middle Ages (Dunin-Wąsowicz, 1960). When analysing the map of the development of the plague in the district, it makes one wonder why the plague did not affect so much the areas located to the north-west of Kalisz. The only reasonable explanation of this phenomenon is the proximity of the Prosna – a river splitting the district into two equal parts. It runs through the district from the South-East, via Kalisz, to North-West, supplying the Warta near Pyzdry, and then heading for Poznań. This means that a large portion of the contemporary transit (of items and people) must have been completed using the Prosna, which must have helped mitigate the effects of the plague in this area (Fig. 7).

Economic impact

Background – Crops prices in Cracow

In many Western researches, climatic events are explained using data on fluctuations of crops prices (Campbell, 2016; Pribyl, 2019). However, there is no collection of prices of basic agricultural products available for Greater Poland for historical periods. The best and most comprehensive sets have been elaborated for Cracow (Pelc, 1935), Warsaw, Lublin and Gdańsk (cf. with Boroda, 2019). Due to its state-of-the-art character and the retained source base (e.g. price series for Warsaw start only from 1526 or 1540–1545), only prices from Cracow – as the most important city in the southern part of the Crown – have been used here (Fig. 8). Prices from Lublin have been preserved only starting from the second half of the 16th century, while the market of Gdańsk was connected more with the economic situation in Amsterdam and other Hanseatic cities than with the rest of the Polish state. The dataset for Cracow is not complete and includes many gaps but the prices of rye can suggest that some kind of fluctuation occurred between 1499 and 1507. It is possible that the price fluctuation was caused by climatic pressure, however without further examples

Table 1 Quantitative comparison of data concerning entries in tax registers from the Kalisz district from 1507–1510. Source: AGAD, ASK I 12

Register [year] w/ potential rate number	No. of entries	No. of all villages	No. of deserted villages (*desertum*)	No. of all empty entries	Percentage of empty entries
	A	B	C	E	$^{E}/_{A}$x100%
1507	601	434	25	66	11%
1508	518	431	30	129	25%
1508 (rents)	543	422	23	164	30%
1509	499	430	28	136	27%
1510 (I)	460	419	11	132	28%
1510 (II)	471	424	13	144	30.57%

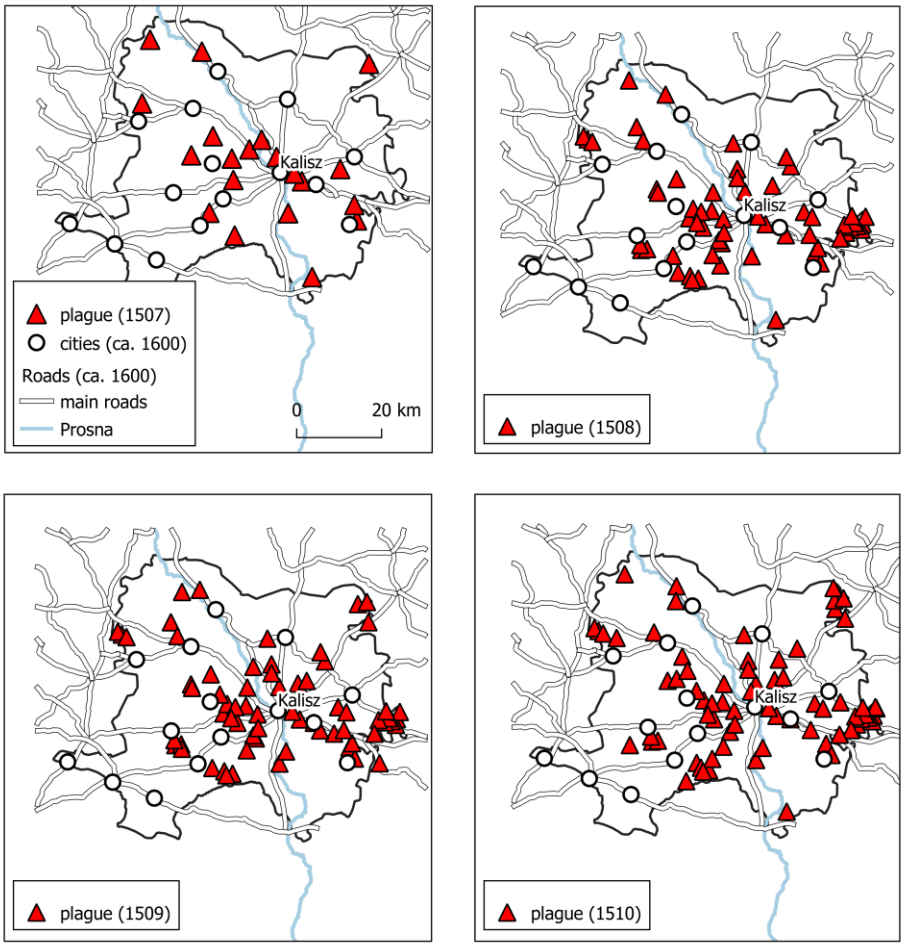

Fig. 7 The spatial development of the plague in Kalisz district villages (1507–1510) according the main roads network. Source ASK I 12

from other cities in the region (Czechia, Hungary, Slovakia) and other studies on prices in Poland in the later centuries, this observation cannot be treated as certain.

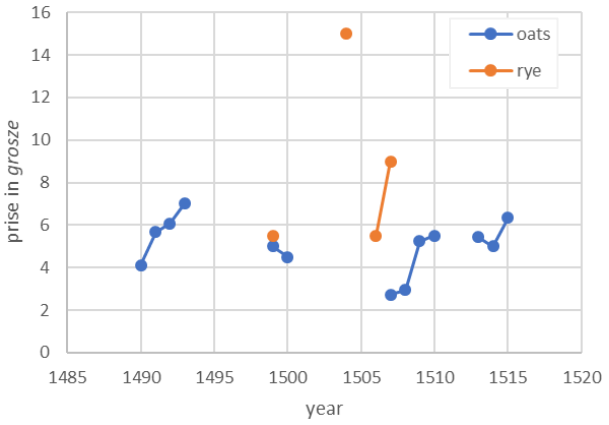

Fig. 8 Recorded prices of oats and rye between 1490 and 1515 in Cracow. Source: (Pelc, 1935).

In towns

The investigation of the impact of the plague upon the economy of the Kalisz district is heavily constrained. There are no comparisons of crop prices available for this area. Changes to the economic structure can be traced using the data included in tax registers or data on potential fluctuations of the prices of rural property (villages, village parts, individual fields, etc.). While the analysis of tax-related data is relatively easy, analysing the prices of property poses problems related to the amount and partial dispersion of the available source material (Gąsiorowski, 1970; Pośpiech, 1989). Such studies also call for a much broader context (territorial and temporal).

The data on the spreading plague can also be studied from the perspective of the economic power of towns in the given district. For years 1507–1509 and 1552, we only have data concerning the liquor tax (Polish: *czopowe*, literally – bung tax) (Table 2), which was a tax paid on the propination – the exclusive right to produce alcoholic beverages in the towns (and sometimes villages) of the former Crown (Boroda, 2016). The data indicate that the plague exerted an unfavourable impact on all the towns of the district, as the income from the liquor tax dropped in 1508 (in relation to the previous year) by approx. 20%. This state of affairs was not long-lasting though, as in the subsequent years the total income from the tax exceeded the amount from before the plague by approx. 6%. On the other hand, when one inspects the impact of the plague on the towns in the context of their overall social, demographic, economic and institutional potential (Słoń, 2017), it turns out that the smaller towns (S) within the district could experience the effects of the plague to a significantly smaller extent than the large (L) and medium (M) towns, such as Kalisz, Pleszew or Stawiszyn. In the times of plague, small towns could successfully compete

with the past production leaders, increasing their own production. The increase in their production could be driven by migrations of craftsmen – temporary escape from one town to another. However, this supposition has not been confirmed in the urban sources for the region. In a longer perspective, it is visible that during the first half of the 16th century, the main urban centres of the district failed to maintain their position. There was a significant economic dispersion in this area, which benefitted small towns, whose economic potential grew in 1552. It is not clear, however, to what extent this effect was caused by the plague of 1507, because there are no sources for the period between the plague in question and the liquor tax register of 1552.

Table 2 Amount of the liquor tax from the towns of the Kalisz district in 1507–1509 and 1552. Source AGAD, ASK I 12

Town / Year	Classification	1507	1508	1509	1552
Kalisz (capital city)	L	13200	10076	12432	6721.5
Pleszew	M	1716	1264	1428	864
Stawiszyn	M	1046	786	3791	2100
Sulmierzyce	S	228	127	168	864
Koźminek	M	226	259	0	768
Zduny	S	222	58	0	192
Iwanowice	S	124	132	126	318
Dobrzyca	S	117	146	66.5	519
Raszków	M	91	596	69	150
Odolanów	S	88	70	74	96
Sobótka	S	48	160	48	96
Kwiatków	S	29	37	13	192
Ostrów Wielkopolski	S	20	70	49	216
TOTAL (in grosze)		17155	13781	18264.5	13096.5

The possibility that the Polish economy of the late 15th and early 16th century functioned in the state of the so-called Late Mediaeval Crisis (Guzowski, 2008) is firmly rejected. Instead, it is argued that this period was one of prosperity, lasting without interruptions from the end of the war against the Teutonic Order (1466) until the so-called Deluge (1655–1660) – a period marked by the catastrophic war against Sweden. As far as this interval is concerned, the vast majority of Polish towns were small or medium-sized (Bogucka, Samsonowicz, 1986). Some towns (reaching 1–2 thousand residents) were clearly agricultural, which was evidenced not only by the small share of crafts and trade, but also by the functioning of a limited economic market in their closest vicinity and existence of many inner gardens (*intra muros*). It appears that this relatively weak urban structure on the Polish territory, disadvantageous in the longer perspective of overall economic development of the country, made it easier to deal with the effects of natural disasters – mainly plagues. It should not come as a surprise that Kalisz, being a medium-sized town and the centre of the region, was most seriously affected by the consequences of the plague in 1507–1510 (Fig. 9). On the other hand, smaller towns benefitted from the situation at this time, offering population running from bigger towns a safe refuge.

In the countryside

The plagues might also cause the significant landscape changes resulting, for example, desertation of arable land or entire villages (Żytkowicz, 1969). In this respect tax registers offer an opportunity to observe the spatial spread of the plague in the countryside areas of the Kalisz district. As far as the general tendencies are concerned, one can see that the size of the agrarian land (Latin: *mansus*, Polish: *łan*) in peasants' [Polish: *kmieć*] and village leaders' [Polish: *sołtys*] farms decreased year on year (Fig. 10). The effects of the plague were becoming more severe, as in the period between 1507 and 1510, approx. 18% of all peasants' farms were completely deserted (Schreg, 2020). The median of the abandonment of the remaining villages in the district was 0.33% of the agrarian land. The crisis affected over half of the villages in the Kalisz district – 186 villages can be considered affected by the disease (52% of all). Their łan area dropped from 4.25 łan to 2.5 (median). In the case of 128 villages (36% of all), no direct correspondence with the plague effects has been noted. It has also been observed that for 40 villages (11% of all), an increase in the number of łans was recorded between 1507 and 1510 (the median of the number of łans from 3 to 4.5). The available data for the Kalisz district do not make it possible to make specific judgements concerning the immediate impact of the disease and its consequences. Apart from following directly from the peasants' death, the phenomena in question could have resulted from their flights as well. The countryside could have also been suffering from overall economic slow-down in this period. It is particularly interesting to note that the unused fields were taken over by neighbouring peasant farms and this practice was of considerable size.

Data from pollen (Kazanie peatland)

In spite of the foregoing, it should be noted that written records alone cannot render the full picture. What can be done about localities for which no tax registers have been preserved? They too were probably struck by the plague; the question is: to what extent? In order to extend the scope of the analysis, some palynological data have been included which can also reflect the economic changes brought about by the plagues with the accompanying transformation of the past landscapes (Mordechai et al., 2019; Słowiński et al., 2019). The peatland is located almost 100 km away from Kalisz, but it can significantly contribute to the data on landscape changes for areas for which we do not have any quantitative written records (in this case Gniezno and Poznań districts). To this end, in order to illustrate the problem and offer a preliminary interpretation, one palynological site was selected, namely the Kazanie peatland (Fig. 1), which reflects environmental changes of a local character (Czerwiński et al., 2019). The chronology of the site was prepared in a 10-year resolution; thus, the paleoecological observations can be made for the period before the climatic extreme year 1505 and after a few-years spread of the plague in the

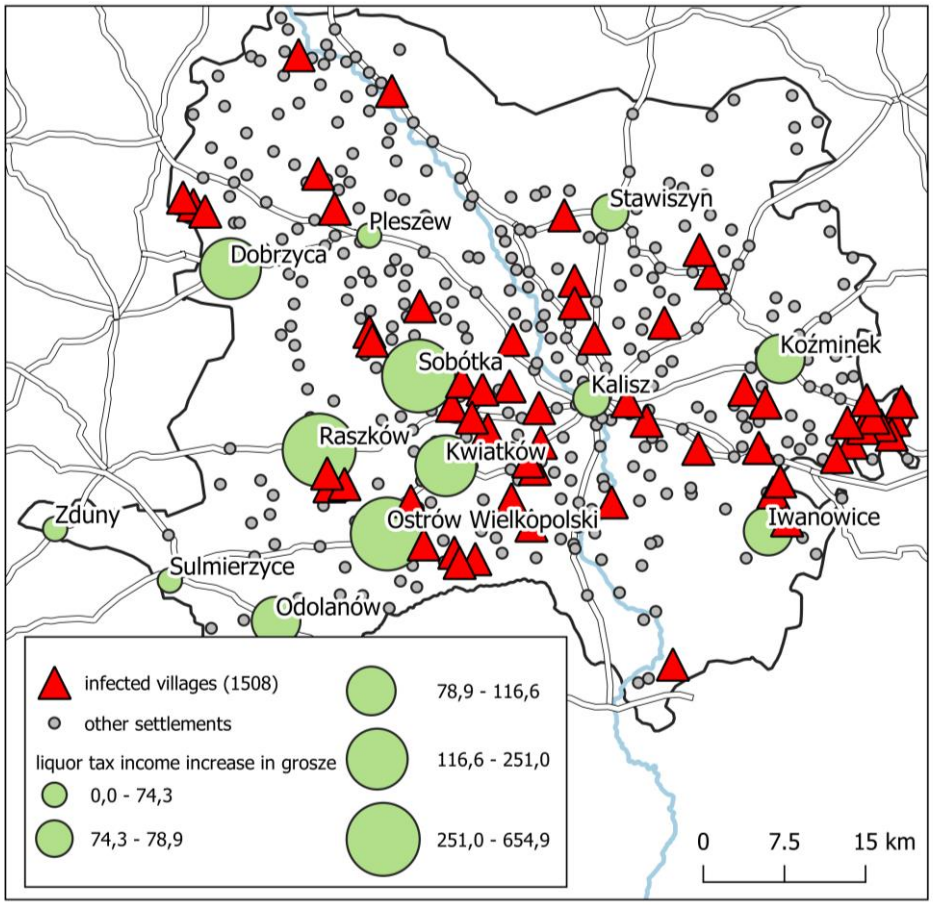

Fig. 9 Increase in income from the liquor tax in 1508 in relation to 1507 – as a percentage. Source AGAD, ASK I 12

Fig. 10 Decrease of the number of arable land in Kalisz district during the plague 1507–1510 with comparison of 1564. Source: AGAD, ASK I 12 and (Kuc et al., 2015).

Polish Kingdom. However, these limited data can offer us the first preliminary look at the environmental and economic impact of the plague on the rural sites in Central Poland. At present, this is the only site in Greater Poland available with reliable absolute chronology and continuous palynological sampling (improving time-resolution of paleoenvironmental interpretation) described by a joint team of historians and paleoecologists.

On the basis of the preliminary materials prepared at the Kazanie site (Fig. 11), one can see (in particular based on the anthropogenic indicators, like cereals and meadow-

pasture plants) that at the beginning of the 16th century, there was a small drop of the pollen share in the analysed samples. It was not rapid, however, reaching for example approx. 1% for cereals in the interval ca. 1501–1511 and nearly 2% in the interval 1501–1521. There is also an almost simultaneous increase of meadow and pasture pollen share, which could indicate the considerable size of the practice of abandoning cultivated fields and transforming them into areas not used in the production of grain crops. One can also note the long-term drop of the percentage value of ruderal plants, which accompany humans' settlement in the given area. Particularly relevant proxies are pollen of cereal and ruderal plants, which declined around 1511. Simultaneously, the level of birch pollen, which is a pioneer species, was rising. In this case, the paleo-proxies should be applied in combination with the written records (abandonment of settlements and fields) in another region. Comparing paleo-data with written evidence, one can assume that the settlement changes in the Kalisz region can be interpreted as a long-term result of the plague, and also, indirectly, as the result of the climate changes during the Little Ice Age (cf. with Schreg, 2011; White 2014).

A detailed analysis of the palynological data from other sites with reliable absolute chronology of sediment would make it possible to offer more accurate observations connected with the reaction of rural communities to all sorts of extreme events – including the plague. The Kazanie fen discussed above offers an example which is limited in terms of range, as it applies at maximum to an area within several kilometres from the site centre. A larger reservoir could even show processes on a regional scale. Secondly, a bigger body of water from which pollen data could be extracted would also be less prone to local environmental transformations (e.g. forest clearing). The observations from the Kazanie peatland are promising when it comes to the use of pollen data analysis in the study of changes brought about by plagues, floods, or droughts in historical-environmental research. But in order to develop a comprehensive narrative about the direct impact and relations between human societies and local environs, it is necessary to build a substantial database of paleo-archives.

CONCLUSION AND RESEARCH PERSPECTIVES

The article has demonstrated that simultaneous application of quantitative and spatial research approaches makes it possible to achieve a deeper understanding of the studied phenomena. Furthermore, the employment of data obtained from nature archives allows one to conduct analyses in a more complex manner. As far as the description of the phenomena in question is concerned, mentions from narrative sources have been proven to be exaggerated. The manner in which they present information is also affected by the problem of the individual perception of the described events – which is difficult to account for. The preliminary analysis of other materials (mainly tax-related and natural) has shown that indeed one can speak of the spread of the plague in the Polish territory at the beginning of the 16th century. Its impact on rural and urban communities was visible in written records (the growing number of deserted villages and fields) and in paleo-data. In particular, the pollen data analysis has shown that settlement abandonment and economy drops in the rural sites can be observed after approx. 1511.

What is definitely needed are further studies of the reactions of pre-industrial societies to various climatic and weather extremes causing plagues. In the subsequent steps, we should focus on the possibility to correlate the data on plagues with the trade in rural property. The same should be carried out for urban areas as well. The analysis of the Polish territories should be extended to cover other provinces – for example Rus' and Mazovia. It would also be advisable to consider the problem of potential economic fluctuations in individual towns within the Polish territory more broadly.

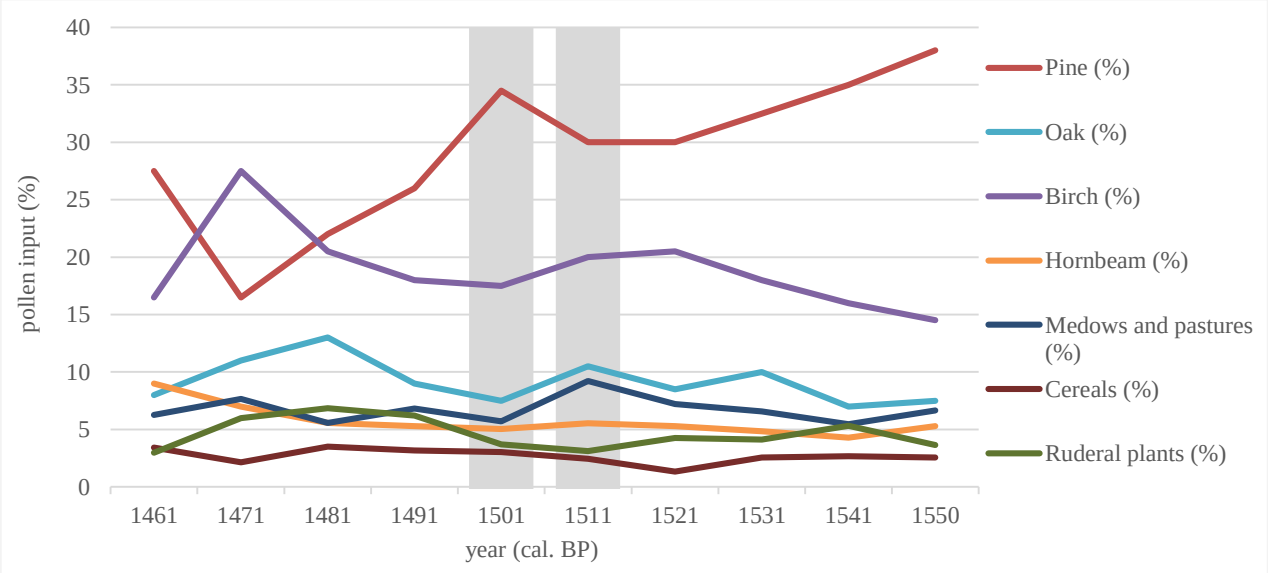

Fig. 11 Simplified pollen data from Kazanie peatland in 10-year resolution. Source (Czerwiński et al., 2019)

In the Polish literature so far, the postulated economic fluctuations have been attributed, among others, to the crisis of feudalism and the resulting general crisis of production in the pre-industrial economy; this seems to call for a reinterpretation. A new approach to the study of factors which shaped the economic changes should consider the mutual relations between the former societies and the changes of environment and climate.

Acknowledgements

I would like to dedicate this article to all the friendly people whom I met during my scholarship at the Central European University in Budapest in the spring of 2019. It is with great pleasure that I recall my meetings especially with Prof. József Laszlovszky, Edit Sárosi, Leslie Carr-Riegel, and János Incze. I would also like to thank Michał Słowiński, Urszula Sowina, Piotr Guzowski, Sambor Czerwiński, Adam Kozak, Agnieszka Mroczkowska, Dominika Łuców, Michał Słomski and Ewa Wółkiewicz for valuable subject matter-related advice. I am also very grateful to Stanisław Butowski for English translation and language corrections of this article. The research was funded using means from the National Programme for the Development of the Humanities no. 1aH15037383, and it is a part of a PhD dissertation called *Man and the environment in the mutual relations. Kalisz District from the end of the 15th century until approx. 1600*.

References

Alfani, G. 2013. Plague in seventeenth-century Europe and the decline of Italy: an epidemiological hypothesis. *European Review of Economic History* 17 (4), 408–430. DOI: 10.1093/ereh/het013

Alfani, G., Percoco, M. 2019. Plague and long-term development: the lasting effects of the 1629–30 epidemic on the Italian cities. *The Economic History Review* 72 (4), 1175–1201. DOI: 10.1111/ehr.12652

Birks, H.J.B. and Birks, H. H. 1980. Quaternary palaeoecology, London, Cambridge University Press.

Bogucka, M., Samsonowicz, H. 1986 Dzieje miast i mieszczaństwa w Polsce przedrozbiorowej, Wrocław.

Boroda, K. 2007. Kmieć, łan czy profit? Co było podstawą poboru łanowego w XV i XVI wieku?'. In: Guzowski, P., Liedke, M. (ed.). Człowiek wobec miar i czasu w przeszłości. Kraków, 152–170.

Boroda, K. 2016. Geografia gospodarcza Królestwa Polskiego w XVI wieku. Instytut Badań nad Dziedzictwem Kulturowym Europy. Białystok.

Boroda, K. 2019. Spatial diversification of prices of rye, wheat, barley, and oats in the Kingdom of Poland in the lustration of royal estates between 1564 and 1565. *Studia Geohistorica* 7, 235–247.

Boroda, K., Guzowski, P. 2016. From King's Finance to Public Finance. Different Strategies of Fighting Financial Crisis in the Kingdom of Poland under Jagiellonian Rule (1386–1572). In Cavaciocchi, S. (ed.) La fiscalità nell'economia Europea secc. XIII-XVIII. Fiscal systems in the European economy from the 13th to the 18th centuries, Firenze, 451–470.

Brázdil, R., Dobrovolný, P., Trnka, M., Kotyza, O., Řezníčková, L., Valášek, H., Zahradníček, P., Štěpánek, P. 2013. Droughts in the Czech Lands, 1090–2012 AD. *Climate and Past* 9 (4), 1987–2002. DOI: 10.5194/cp-9-1985-2013

Brázdil, R., Kotyza, O., Dobrovolný, P., Řezníčková, L. and Valášek, H. 2013. History of weather and climate in the Czech lands. Brno, Masaryk University.

Büntgen, U., Tegel, W., Nicolussi, K., McCormick, M., Frank, D., Trouet, V., Kaplan, J. O., Herzig, F., Heussner, K.-U., Wanner, H., Luterbacher, J., Esper, J. 2011. 2500 years of European climate variability and human susceptibility, *Science* 331 (6017), 578–582. DOI: 10.1126/science.1197175

Camenisch, C. 2015. Endless cold: a seasonal reconstruction of temperature and precipitation in the Burgundian Low Countries during the 15th century based on documentary evidence. *Climate of the Past* 11 (8), 1049–1066. DOI: 10.5194/cp-11-1049-2015

Camenisch, C. 2018. Two Decades of Crisis: Famine and Dearth During the 1480s and 1490s in Western and Central Europe. In Collet, D. and Schuh, M. (ed.) Famines during the 'Little Ice Age' (1300-1800): Socionatural entanglements in premodern societies, Cham, Springer, 69–90.

Campbell, B. 2016. The Great Transition, Cambridge, Cambridge University Press.

Cook, E. R., Seager, R., Kushnir, Y., Briffa, K. R., Büntgen, U., Frank, D., Krusic, P. J., Tegel, W., van der Schrier, G., Andreu-Hayles, L., Baillie, M., Baittinger, C., Bleicher, N., Bonde, N., Brown, D., Carrer, M., Cooper, R., Čufar, K., Dittmar, C., Esper, J., Griggs, C., Gunnarson, B., Günther, B., Gutierrez, E., Haneca, K., Helama, S., Herzig, F., Heussner, K.-U., Hofmann, J., Janda, P., Kontic, R., Köse, N., Kyncl, T., Levanič, T., Linderholm, H., Manning, S., Melvin, T. M., Miles, D., Neuwirth, B., Nicolussi, K., Nola, P., Panayotov, M., Popa, I., Rothe, A., Seftigen, K., Seim, A., Svarva, H., Svoboda, M., Thun, T., Timonen, M., Touchan, R., Trotsiuk, V., Trouet, V., Walder, F., Ważny, T., Wilson, R., Zang, C. 2015. Old World megadroughts and pluvials during the Common Era. *Science advances* 1 (10). DOI: 10.1126/sciadv.1500561

Cummins, N., Kelly, M., Ó Gráda, C. 2016. Living standards and plague in London, 1560-1665. *The Economic History Review* 69 (1), 3–34. DOI: 10.1111/ehr.12098

Curtis, D. R. 2016. Was plague an exclusively urban phenomenon? Plague mortality in the seventeenth-century Low Countries. *Journal of Interdisciplinary History* 47 (2), 139–170. DOI: 10.1162/jinh_a_00975

Czerwiński, S., Guzowski, P., Karpińska-Kołaczek, M., Lamentowicz, M., Gałka, M., Kołaczek, P., Izdebski, A., Poniat, R. 2019. The importance of collaborative historical and palaeoecological research on human impact on the environment: the case of the Kazanie site in Eastern Greater Poland (Eastern Wielkopolska). *Studia Geohistorica* 7, 56–74.

Dunin-Wąsowicz, T. 1960. Kalisz na tle wczesnośredniowiecznej sieci drogowej. In Gieysztor, A. (ed.) Osiemnaście wieków Kalisza. Studia i materiały do dziejów miasta Kalisza i regionu kaliskiego, Kalisz, 69–101.

Ell, S. R. 1985. Iron in Two Seventeenth-Century Plague Epidemics. *Journal of Interdisciplinary History* 15 (3), 445–457. DOI: 10.2307/204140

Follprecht, K. and Noga, Z. 2014. Cracow in 1598. In Słoń, M. (ed.) Historical Atlas of Poland in the 2nd Half of the 16th Century, Frankfurt am Main, Bern, Wien, Peter Lang, 692–765.

Gąsiorowski, A. 1970. Tak zwane prawo wieczności w dawnej Polsce. *Czasopismo Prawno-Historyczne* 22, 31–58.

Gieysztorowa, I., Żaboklicka, A. 1955. Rejestry poborowe Mazowsza z XVI wieku. *Kwartalnik Historii Kultury Materialnej* 3 (2), 343–351.

Gochna, M., Związek, T. 2019. Spatio-temporal aspects of the extraordinary tax collecting system in Greater Poland (1492–1613). *Roczniki Dziejów Społecznych i Gospodarczych* 80, 65–111.

Guzowski, P. 2005. A changing economy: models of peasant budgets in fifteenth- and sixteenthcentury Poland. *Continuity and Change* 20 (1), 9–25. DOI: 10.1017/s0268416004005338

Guzowski, P. 2008. Kryzys gospodarczy późnego średniowiecza czy kryzys historiografii? *Roczniki Dziejów Społecznych i Gospodarczych* 68, 173–193.

Guzowski, P., Kuklo, C., Poniat, R. 2016. O metodach pomiaru natężenia epidemii i zaraz w preindustrialnej Europie w demografii historycznej. In Polek, K., Sroka, Ł. (ed.) Epidemie w dziejach Europy: konsekwencje społeczne, gospodarcze i kulturowe, Wydawnictwo Naukowe Uniwersytetu Pedagogicznego, Kraków, 119–144.

Haldon, J., Mordechai, L., Newfield, T. P., Chase, A. F., Izdebski, A., Guzowski, P., Labuhn, I., Roberts, N. 2018. History meets palaeoscience: consilience and collaboration in studying past societal responses to environmental change. *Proceedings of the National Academy of Sciences* 115 (13), 3210–3218.

Izdebski, A., Koloch, G., Słoczyński, T., Tycner, M. 2016. On the use of palynological data in economic history. New methods and an application to agricultural output in Central Europe, 0-2000 AD.

Explorations in Economic History 59, 17–39. DOI: 10.1016/j.eeh.2015.10.003

Karpiński, A. 2000. W walce z niewidzialnym wrogiem. Epidemie chorób zakaźnych w Rzeczypospolitej w XVI–XVIII wieku i ich następstwa demograficzne, społeczno-ekonomiczne i polityczne, Warszawa, Neriton.

Kiss, A. 2017. Droughts and low water levels in Late Medieval Hungary II: 1361, 1439, 1443-4, 1455, 1473, 1480, 1482(?), 1502-3, 1506: documentary versus three-ring (OWDA) evidence. *Journal of Environmental Geography* 10 (3-4), 43–56. DOI: 10.1515/jengeo-2017-0012

Kiss, A. 2019. A dynamic interplay of weather, biological factors and socio-economic interactions: Late 15th-century - early 16th-century crises in Hungary. In: Kiss, A., Pribyl, K. (ed.) The Dance of Death in Late Medieval and Renaissance Europe: Environmental Stress, Mortality and Social Response, London-New York, 125–145.

Kiss, A. 2020. 'The great (1506–)1507 drought and its consequences in Hungary in a (Central) European context. *Regional Environmental Change* 20 (2), 1–13. DOI: 10.1007/s10113-020-01634-5

Kiss, A., Nikolić, Z. 2015. Droughts, Dry Spells and Low water Levels in Medieval Hungary (and Croatia) I: The Great Droughts of 1362, 1474, 1479, 1494 and 1507. *Journal of Environmental Geography* 8 (1–2), 11–22. DOI: 10.1515/jengeo-2015-0002

Kuc, M., Suproniuk, J., Zachara-Związek, U. 2015. Rejestr poborowy powiatu kaliskiego 1564 r. http://atlasfontium.pl/index.php?article=kalisz_1564.

Lagerås, P. 2016. An interdisciplinary approach. In Lagerås, P. (ed.) Environment, society and the Black Death: an interdisciplinary approach to the late-medieval crisis in Sweden, Oxfrod, Oxbow Books, 2–5.

Limanówka, D. 2001. Rekonstrukcja warunków klimatycznych Krakowa w pierwszej połowie XVI wieku. Instytut Meteorologii i Gospodarki Wodnej. Warszawa.

Martín-Vide, J. and Barriendos Vallvé, M. 1995. The use of rogation ceremony records in climatic reconstruction: a case study from Catalonia (Spain). *Climatic Change* 30 (2), 201–222. DOI: 10.1007/bf01091842

Mordechai, L., Eisenberg, M., Newfield, T. P., Izdebski, A., Kay, J. E., Poinar, H. 2019. The Justinianic Plague: An inconsequential pandemic?. *Proceedings of the National Academy of Sciences of the United States of America*. DOI: 10.1073/pnas.1903797116

Owens, M. J., Lockwood, M., Hawkins, E., Usoskin, I., Jones, G. S., Barnard, L., Schurer, A., Fasullo, J. 2017. The Maunder minimum and the Little Ice Age: an update from recent reconstructions and climate simulations. *Journal of Space Weather and Space Climate* 7, 1–10.

Pelc, J. 1935. Ceny w Krakowie w latach 1369–1600. Kasa im. Rektora J. Malinowskiego. Lwów.

Pfister, C. 1988. Klimatgeschichte der Schweiz 1525-1860. Bern.

Polish Territories of the Crown in the 16th century. Spatial Database. 2015. http://atlasfontium.pl/index.php?article=korona.

Pośpiech, A. 1989. Księgi rezygnacji w kaliskich aktach grodzkich z lat 1580-1655 (Prezentacja źródła, możliwości badawcze. *Rocznik Kaliski* 21, 19–49.

Pribyl, K. 2019. The grain trade, economic distress and social disorder at a time of environmental stress in East Anglia, 1400–ca. 1440. In: Kiss, A., Pribyl, K. (ed.) The Dance of Death in Late Medieval and Renaissance Europe: Environmental Stress, Mortality and Social Response. Routledge. London-New York. 46–65.

Przybylak, R., Oliński, P., Koprowski, M., Filipiak, J., Pospieszyńska, A., Chorążyczewski, W., Puchałka, R., Dąbrowski, H. P. 2019. Droughts in the area of Poland in recent centuries, *Climate of the Past Discussions.* DOI: 10.5194/cp-2019-64, 1–56

Schreg, R. 2011. Feeding the village - reflections on the economy and resilience of the medieval rural economy. In: Klápště, J., Sommer, P. (ed.) Processing, storage, distribution of food in medieval rural environment. Turnhout. Brepols, 301–320.

Schreg, R. 2019. Late Medieval deserted settlements in southern Germany as a consequence of long-term landscape transformations. In Brady, N. and Theune, C. (ed.) Settlement change across Medieval Europe. Old paradigms and new visas. Leiden. Sidestone Press, 161–170.

Schreg, R. 2020. Plague and desertion - a consequence of anthropogenic landscape change? Archaeological studies in Southern Germany. In Bauch, M. and Schenk, G. J. (ed.) The Crisis of the 14th Century. Berlin-Boston, 221–246.

Słoń, M. 2011. Digitale Edition der Ausheberregister aus der Wojewodschaft Kalisch des 16. Jahrhunderts. In Kopiński, K., Flachenecker, H., Tandecki, J. (ed.) Editionswissenschaftliches Kolloquium 2011. Quellen kirchlicher Provenienz: neue Editionsvorhaben und aktuelle EDV-Projekte. Toruń, 396–406.

Słoń, M. 2017. Miasta prywatne w sieci miejskiej Wielkopolski XV–XVI wieku. *Roczniki Dziejów Społecznych i Gospodarczych,* 77, 93–123.

Słowiński, M., Lamentowicz, M., Łuców, D., Barabach, J., Brykała, D., Tyszkowski, S., Pieńczewska, A., Śnieszko, Z., Dietze, E., Jażdżewski, K., Obremska, M., Ott, F., Brauer, A., Marcisz, K. 2019. Paleoecological and historical data as an important tool in ecosystem management. *Journal of Environmental Management* 236, 755–768. DOI: 10.1016/j.jenvman.2019.02.002

Sowina, U. 2016 Kraków wobec zarazy z 1543 roku w świetle rachunków miejskich. In Polek, K. and Sroka, Ł. (ed.) Epidemie w dziejach Europy: konsekwencje społeczne, gospodarcze i kulturowe, Kraków, 169–184.

Sowina, U. 2020. Towary w inwentarzach rzeczy mieszczan krakowskich na przełomie średniowiecza i nowożytności. *Kwartalnik Historii Kultury Materialnej*. in print.

Theuerkauf, M., Couwenberg, J. 2017. The extended downscaling approach: A new R-tool for pollen-based reconstruction of vegetation patterns. *The Holocene* 27 (8), 1252–1258. DOI: 10.1177/0959683616683256

Walawender, A. 1932. Kronika klęsk elementarnych w Polsce i krajach sąsiednich w latach 1450–1586: Zjawiska meteorologiczne i pomory (z wykresami). Lwów.

Walawender, A. 1957. Obrona przez zarazami w Polsce na przełomie wieków średnich. *Archiwum Historii Medycyny* 20 (1-2), 17–34.

White, S. 2014. The Real Little Ice Age. *The Journal of Interdisciplinary History* 44 (3), 327–352. DOI: 10.1162/jinh_a_00574

WHO Flooding and communicable diseases fact sheet. https://www.who.int/hac/techguidance/ems/flood_cds/en/.

Yue, R. P. H., Lee, H. F. (2020). Drought-induced spatio-temporal synchrony of plague outbreak in Europe. *The Science of the total environment* 698, 134–138. DOI: 10.1016/j.scitotenv.2019.134138

Związek, T. 2013. Najstarszy rejestr poboru nadzwyczajnego i szosu z 1507 roku z terenów powiatu konińskiego. *Rocznik Koniński* 18, 173–201.

Związek, T. 2017. Drogi. In Chłapowski, K. and Słoń, M. (ed.) Wielkopolska w drugiej połowie XVI wieku. Instytut Historii PAN. Warszawa, 268–290.

Żytkowicz, L. 1969. Studia and wydajnością gospodarstwa wiejskiego na Mazowszu w XVII wieku, Warszawa. Państwowe Wydawnictwo Naukowe.

sciendo

Journal of Environmental Geography 13 (1–2), 13–24.

DOI: 10.2478/jengeo-2020-0002
ISSN 2060-467X

THE LAND USE - LAND DEGRADATION NEXUS IN MEDITERRANEAN LANDSCAPES – DRIVERS OF CHANGES AND KEY PROCESSES AT SELECTED NATURA 2000 SITES OF CRETE, GREECE

Burghard C. Meyer*, Fabian Kirsten, Dietmar Sattler, Jürgen Heinrich

Institute of Geography, University of Leipzig, Johannisallee 19a, 04103 Leipzig, Germany
*Corresponding author, e-mail: burghard.meyer@uni-leipzig.de

Research article, received 14 January 2020, accepted 24 March 2020

Abstract

The land use–land degradation nexus in Cretan landscapes in regions with Natura 2000 sites was analyzed by an explorative expert driven study based on literature, field work and photo documentation methods with the aim of determining status, drivers and key processes of change. Drivers of current land use changes have been worked out by (1) general tourism developments and tourism related land uses; (2) irrigated olive yard developments; (3) fenced large-scale goat pastures and (4) large scale greenhouses. Key processes of change have been identified and qualitatively assessed for 5 regions with NATURA 2000 areas based on a non-ranked set of 11 descriptive indicators. The analysis includes the status-description and the importance assessment of land degradation processes in selected NATURA 2000 sites. Threats and pressures taken from the NATURA 2000 documentation and the land use – land degradation nexus and the analysis are a suitable basis for future land management in order to reach land degradation neutrality. The result of our analysis opens a new research field for a better integration of the normally thematically isolated analysis in geography, biology/nature conservation and agricultural policy analysis about the drivers and processes in landscape systems towards a better understanding the trends in land cover change (e.g. vegetation/soil degradation), the trends in productivity or functioning changes caused by land uses and as well for the trends in carbon stock change.

Keywords: integrative land use analysis, landscape change, agriculture, nature conservation, processes of land use change

INTRODUCTION

The Cretan economy and its landscapes as well as most regions in the Mediterranean have been centered on agricultural activities since prehistoric times. Subsistence consisted of cereal cropping in lower regions with fertile soils, whereas in mountainous areas animal husbandry, mostly sheep and goat, was the dominating activity and land use. In some areas, a transhumance system was well developed. The land use history of Crete has caused widespread and still ongoing soil degradation linked with a transformation of the natural vegetation, processes also dating back to ancient times. Locally, the dense Mediterranean evergreen and semi-deciduous forests have already been transformed into Maquis shrublands and Phrygana vegetation in prehistoric times, with an expansion of such areas and further degradation throughout different phases of intense land use (Rother 1992). As a result, nearly the entire island of Crete can be considered and labeled a cultural landscape.

Several sites of this old cultural landscape are protected today as Natura 2000 sites with high qualities for nature protection (EC 1992). Nature conservation, agriculture and tourism are today in a problematic competition for land and water. In order to reach the protection goals of the Natura 2000 guideline, a balanced land management is missing to reduce the negative threats and pressures (Barredo et al., 2016; Dimopoulos et al., 2006; Cuttelod et al., 2008; Underwood et al., 2009). While agriculture was the dominating activity of the Cretan economy until the middle of the 20th century, the touristic sector has become increasingly important in the last fifty years. In the first decades after World War 2, tourists started to visit Crete and created some extra income for local farmers, mostly in the dry summer months (Donatos and Zairis, 1991). However, there was a strong emigration movement, depleting the rural areas of the young generation and workforce. This led to an abandonment of fields and sometimes entire villages, especially in mountainous areas and a reduction in stocking rates and herd numbers (Papanastatis et al. 2004; Watrous, 1982). Besides tourism, agricultural production continues to be an important land use activity. On fertile soils cash crops like olives, citrus and regionally also potatoes have replaced the traditional cereals. Although the touristic infrastructure developments are centered on the northern coast, there is a growing competition for land and water between tourism and agriculture on the entire island. As the demand for irrigation for cultivation as well as tourist arrivals concentrate in the dry summer season, the pressure on water resources is especially pronounced.

With Greece joining the European Union (EU) in 1981, the agricultural and the tourism sector developments started to accelerate. EU Common Agricultural Policy (CAP) subsidies stimulated and incentivized an increase in numbers and stocking rates of grazing animals (Lorent et al. 2009). For example, the number of grazing sheep and goats has increased from 1.56 to 2.51 million in the period

from 1980 to 2010 with stocking rates increasing from around 1 to 2.0-2.3 animal units (AU) per ha (Kosmas et al., 2015). As stocking rates around 1 to 1.5 AU/ha are considered appropriate for sustainable grazing, recent stocking rates result in severe overgrazing and an increasing degradation of soils and vegetation (Ispikoudis et al., 1993; Kairis et al., 2015).

Subsidies from the European Union have also encouraged the intensification of olive cultivation in Crete and across the Mediterranean, leading to rapid landscape changes from a diverse system to monocropping olive yard structures (Allen et al., 2006). In 2017, there were a total number of 27 220 000 olive trees, mostly in compact plantations and under irrigation, on an area of 136 400 ha (Hellenic Statistical Authority 2019). Of these, 41 200 ha with 7 850 000 trees have been in use for more than 50 years, while about 10 000 ha containing 2 000 000 trees are only 10 years old or younger. Most of these are used for oil production. A new form of agricultural land use is increasingly applied by the construction of greenhouses, where mostly legumes and bananas are grown year-around (Karamesouti et al., 2015).

Due to the long history of intense land use, the Mediterranean can be best described as cultural landscapes. It holds valuable assets of culture and nature vital for the cultural land perception and recreation (Farina, 2000). With Natura 2000 (EC, 1992, further referred to as NATURA 2000) the EU developed a coherent network of nature protection areas in its member states facing the goal to halt ongoing biodiversity loss by the protection of habitats and species of flora and fauna on terrestrial and marine ecosystems. So far, only few studies have developed management ideas and practices for a future integration of culture and land use in NATURA 2000 sites by management planning (NK2-Group, no year), or for biosphere reserve developments (Kusova et al., 2008). Even though the Cretan NATURA 2000 Special Areas of Conservation (SAC) network seems to be insufficient to ensure at least a satisfactory representation of the regional plant biodiversity (Dimitrakopoulos et al., 2005), these areas cover about 17 % of the islands terrestrial area. European management guidelines are available for farming (e.g. on the grazing on high value grasslands) in NATURA 2000 (EC, 2014) as well as for the forest's management (EC, 2015). Guidelines for land use practices in NATURA 2000 sites dominated by tourism or recreation - important especially on Mediterranean islands like Crete - have not been developed so far.

The land degradation neutrality goal (defined by the Sustainability Development Goals) has not yet been specified for landscapes comprising NATURA 2000 sites in Crete, since complex explorative studies are missing. Land degradation problems caused by land use, its management and changes, land degradation processes, their determining societal and land drivers and their interlinkages on habitat qualities are generally rarely in the focus of applied research due to their complexity. The aim of the presented explorative study is therefore to work out by literature, photo and expert field analysis important ongoing drivers and processes in landscape systems to determine the status, important to clarify the land use – land degradation nexus in Mediterranean landscapes of Crete.

STUDY AREAS AND METHODS

A number of characteristic Cretan landscapes comprising NATURA 2000 sites were visited during a field campaign in spring 2018 aiming at the clarification and description of critical developments in land use systems including land degradation and depletion of biodiversity (Fig. 1, Fig. 2). The analysis of respective drivers is based on the authors´ field observation and is as well based and supported by literature and data interpretations of the NATURA 2000 public database (EC without year). The selection of study areas is based on a general analysis of different landscape types in Crete, focusing on landscape ecology, geomorphology, soil, vegetation and land use. The study area selection includes varying land uses in high mountains, lower mountain ranges, forests, abandoned and active arable lands, intensively irrigated arable lands, vineyards, irrigated and non-irrigated olive yards, important cultural sites, villages, tourism facilities, beaches, dunes and multiple mixed landscape configurations as Maquis shrublands and Phrygana vegetation as well as abandoned terraced landscapes.

Regions and sites have been selected with the objective to get a major overview of typical current developments within the land use – land degradation nexus, facilitating a better understanding of the status, drivers and active processes potentially causing changes to land and landscape degradation.

Drivers of land use change potentially affecting study areas are analyzed by varying methods focusing on (1) General tourism developments and tourism related land uses; (2) Irrigated olive yard developments; (3) fenced large-scale goat pastures and on (4) Large scale greenhouses. These driver's selection is based on authors expert judgments, on landscape observations, map interpretation and literature.

The main methodological approach was structured observations of recent processes and their activity level influencing land degradation based on photo-documentation and field research. For the traditional Cretan cultural landscapes, the observations are focused on the processes of depopulation, road infrastructure development, forest fires, overgrazing, water resources in quality/quantity, soil erosion, slope modification, disturbances, habitat fragmentation, land cover change and the abandonment of important cultural landscape elements.

Due to the complex geology, relief, related climatological conditions and more or less intense anthropogenic modifications, the island of Crete is characterized by a large variety of natural and man-made habitats at a comparably small surface area of 8265km², mainly given to the intense elevational gradients of the three main mountain ranges (see Fig. 1) and the long history of human presence and activity. The vascular plant flora of Crete comprises 2240 species with a level of endemism of 17% and 6,7% alien or non-native species (Médail, 2017; Turland et al., 1993). With its high level of endemism and high percentage of habitat loss the island is part of the Mediterranean biodiversity hotspot (Médail and Myers, 2005), endangered by continuing habitat destruction and land degradation due to land use

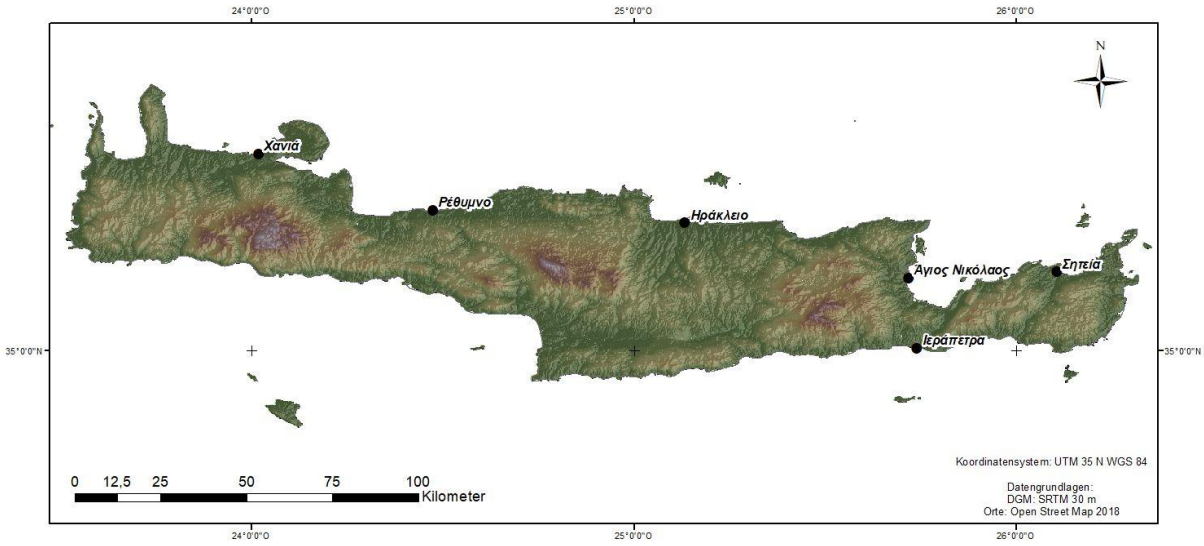

Fig. 1 Topography of Crete (Data Source: SRTM, OpenStreetMap). The map shows the rough and diverse topography of Crete holding high mountains (white, dark brown), multiple hilly areas (brown) and lowland plains and valleys (green)

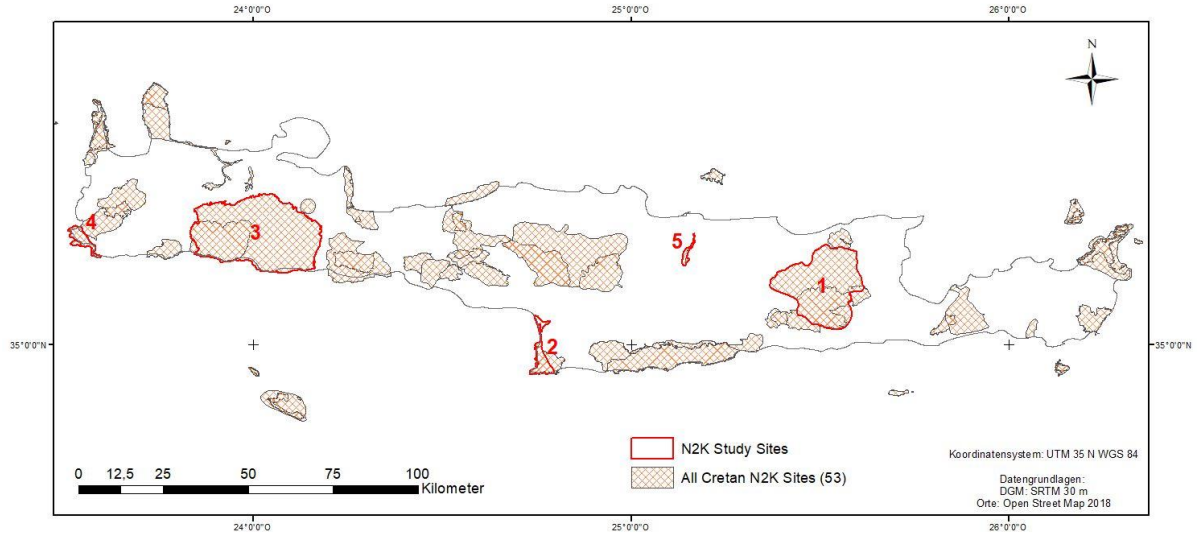

Fig.2 NATURA 2000 (N2K sites) of Crete and analysed N2K study sites (red edging) (1) Lasithi plateau, (2) Ekvoli Geropotamou Mesara, (3) Lefka Ori Kai Parakti, (4) Nisos Elafonisos and Paralia Apo Chrysoskalitissa Mechri Akrotirio Krios and (5) Giouchtas Farangi Agias Eirinis

intensification (olive groves, goat and sheep overgrazing, tourism), forest fires and biological invasions, e.g. *Ailanthus altissima*, *Nicotiana glauca* and *Oxalis pes-capreae* (D´Agata et al., 2009) but as well by abandonment of terraces and land use (e.g. followed by forestation and forest succession).

Crete's habitat diversity has been partially taken into account in our NATURA 2000 area selection process, including areas from the shore up to mountainous regions at 2400 m a.s.l. (Vogiatziakis et al., 2003). For a first overview survey we selected areas comprising NATURA 2000 sites in Mid- and Western Crete ranging from the southwestern coast to the central mountain formations, based on the assumption that all sites are actually affected by more or less strong anthropogenic impacts such as tourism and agricultural land use. Other impacts such as water resources exploitation, land levelling for greenhouse cultures, expansion of livestock breeding (e.g. goats, sheep, bees) and infrastructural development measures have also been taken into account. Basic information about the NATURA 2000 features of these areas can be found in Table 1a,b, that also include threats and pressures recorded at the moment of the designation of the NATURA 2000 sites. The selected areas differ considerably with respect to their geological and geomorphological setting, their traditional land use system and settlement pattern as well as their current significance for agricultural production and tourism.

The site selection therefore addresses combinations of landscapes, habitats and drivers of degradation being typical and representative for Crete. According to Dimitriou (2008) and our own observations, the most significant threats for (natural) habitats in Crete are i) Hydrological alteration due to water overexploitation, artificial drainage and recharge, ii) Overgrazing due to the intense stockbreeding practices, iii) Water pollution and especially eutrophication originating from unsustainable agricultural practices, iv) Solid waste disposal, v) growing and intensified tourism and a lack of legally enforced protection.

Table 1a Threats and pressures of the selected NATURA 2000 study sites in Crete

Name (Code)	**DIKTI: OROPEDIO LASITHIOU (GR4320002)**		**DYTIKA ASTEROUSIA (GR4310004)**		**LEFKA ORI KAI PARAKTIA (GR4340008)**	
Total area (ha)	34007		2651		53364	
Marine area (%)	0,00		0,00		3,35	
N2K Type	B (Habitat directive site)		B (Habitat directive site)		B (Habitat directive site)	
Precipita-tion (mm/y)	900-2200		450-550		500-2800	
Altitude (m a.s.l.)	800-2148		0-400		0-2443	
Threats and Pressures: inside (i), outside (o), both (b) and Rank (High,Medium,Low)	wind energy production	iH	Taking and removal of animals (terrestrial)	bH	grazing	bH
	genetic pollution (animals)	bH	Outdoor sports, leisure and recreational activities	bH	Improved access to site	iH
	grazing	bH/M	Sand and gravel extraction	iM	Hunting and collection of wild animals	oH
	fire (natural)	bH/M	marine constructions	bM	trapping, poisoning, poaching of terrestrial animals	iH
	Cultivation	iM	dispersed habitation	bM	other hunting, fishing or collecting activities	iH
	use of biocides, hormones and chemicals	iM	Professional passive fishing, netting, spear-fishing	bM	attraction park	iH
	Fertilisation	iM	Illegal taking/ removal of marine fauna	bM	roads, motorways, all paved/ tarred roads	iM
	dispersed habitation	iM	camping and caravans, wildlife watching	iM	genetic pollution (plants)	iM
	Hunting and collection of wild animals	oM	human trampling, overuse, vandalism	iM	Erosion	iM
	Taking and removal of animals (terrestrial)	iM	intensive maintenance of public parcs /cleaning of beaches	iM	Sand and gravel extraction	iL
	Taking / Removal of terrestrial plants, general	iM	marine animals death or injury by collision	oM	disposal of household / recreational facility waste	iL
	Erosion	iM	oil spills in the sea	bM	Outdoor sports and leisure activities, recreational activities	iL
	competition	bM	marine macro-pollution	oM		
	Irrigation	iL	Light pollution	iM		
	artificial planting on open ground (non-native trees)	iL	Erosion	iM		
	forestry clearance, clear-cutting, removal of all trees	iL	collapse of terrain, landslide	bM		
	removal of forest undergrowth	iL	use of biocides, hormones and chemicals	bL		
	discontinuous urbanisation	iL	Irrigation	bL		
	disposal of household / recreational facility waste	iL	artificial planting on open ground (non-native trees)	iL		
	walking, horseriding and non-motorised vehicles	iL	car parcs and parking areas	iL		
	motorised vehicles	iL	diffuse pollution to surface waters due to agricultural and forestry activities	bL		
	mountaineering, rock climbing, speleology	iL	inundation (natural processes)	iL		
	fire and fire suppression	bL				
	human induced changes in hydraulic conditions	iL				
	inundation (natural processes)	iL				

Table 1b Threats and pressures of the selected NATURA 2000 study sites in Crete

Name (Code)	**NISOS ELAFONISOS (GR4340002)**		**PARALIA APO CHRYSOSKALITISSA MECHRI AKROTIRIO KRIOS (GR4340015)**		**GIOUCHTAS (GR4310002)**	
Total area (ha)	271		2253		718	
Marine area (%)	83,73		31,18		0,00	
N2K Type	B (Habitat directive site)		B (Habitat directive site)		B (Habitat directive site)	
Precipita-tion (mm/y)	450		450		600-800 mm	
Altitude (m a.s.l.)	0-39		0-160		100-808	
Threats and Pressures: inside (i), outside (o), both (b) and Rank (High,Medium,Low)	Taking / Removal of terrestrial plants	iH	disposal of household / recreational facility waste	iH	Urbanised areas, human habitation	iH
	camping and caravans	iH	Taking / Removal of terrestrial plants	iH	disposal of household / recreational facility waste	iH
	marine constructions	bM	leisure activities with motorized vehicles	iH	grazing	iM
	dispersed habitation	bM	camping and caravans	iH	Animal breeding, stock feeding	oM
	Professional passive net fishing	bM	genetic pollution (animals)	bH	Restructuring agricultural land holding	oM
	leisure spear-fishing	bM	Sand and gravel extraction	iM	artificial planting on open ground (non-native trees)	iM
	Hunting and collection of wild animals	iM	open cast mining	iM	roads, motorways, all paved/ tarred roads	iM
	Accidental capture of terrestrial animals	bM	marine constructions	bM	mountaineering, rock climbing, recreational cave visits	iM
	Illegal taking/ removal of marine fauna	bM	Professional passive net fishing	bM	attraction park	iM
	motorized nautical sports	bM	leisure spear fishing	bM	Vandalism	iM
	recreational visits of terrestrial/marine caves	bM	Hunting and collection of wild animals	iM	missing or wrongly directed conservation measures	iM
	Pollution to surface waters (limnic, terrestrial, marine)	bM	Illegal taking/ removal of marine fauna	bM	fire and fire suppression	iM
	Marine water pollution, oil spills in the sea	bM	motorized nautical sports	bM	cultivation incl. increase of agricultural area	iL
	collapse of terrain, landslide	bM	recreational visits of terrestrial/marine caves	bM	use of biocides, hormones and chemicals	IL
	sea-level changes	bM	Pollution to surface waters (limnic, terrestrial, marine)	bM	open cast mining	oL
			Marine water pollution, oil spills in the sea	bM	paths, tracks, cycling tracks	iL
			collapse of terrain, landslide	bM	discontinuous urbanisation	oL
			sea-level changes	bM	dispersed habitation	iL
			continuous and dispersed urbanisation	bM/L	disposal of inert materials	oL
			grazing	iL	Agricultural structures, buildings in the landscape	oL
					Hunting and collection of wild animals	iL
	Taking / Removal of terrestrial plants	iH	disposal of household / recreational facility waste	iH	Urbanised areas, human habitation	iH
	camping and caravans	iH	Taking / Removal of terrestrial plants	iH	disposal of household / recreational facility waste	iH
	marine constructions	bM	leisure activities with motorized vehicles	iH	grazing	iM
	dispersed habitation	bM	camping and caravans	iH	Animal breeding, stock feeding	oM

Five areas containing NATURA 2000 sites were selected and analyzed to get an overview of important geo-ecological features, land use and main NATURA 2000 area qualities (Fig. 2).

Threats and pressures of the selected NATURA 2000 study sites in Crete as described in Table 1a and Table 1b are taken from the official site descriptions (available by the Natura 2000 Public Database). The same data have been firstly used for the designation of the NATURA 2000 sites. The database as well includes information about the threats and pressures affecting the sites by differentiating them by their activity inside the site (i), outside the site (o), or in both inside and outside (b) and as well by an importance ranking (high, medium, low) to clarify the importance of the treats and pressures in the respective site context.

RESULTS: DRIVERS AND PROCESSES OF LAND USE CHANGE AND LAND DEGRADATION IN REGIONS WITH NATURA 2000 SITES

Drivers of landscapes and NATURA 2000 site change

Four major drivers, all strongly influenced by European policies after the integration of Greece into the European Union in 1981, have been identified as of major importance on land use change and land degradation in NATURA 2000 sites and surrounding areas in Crete. These are interpreted as general determinants of land use intensification. Since the NATURA 2000 site designation (between the years of 1995 and 2004 and today) the main drivers have been found in (a) tourism developments and tourism related land uses, (b) irrigated olive yard developments, (c) fenced large scale goat pasture systems and (d) large scale greenhouse farming strongly modified the land use systems. All four drivers are influenced by regional development policies and measures including water management measures, agricultural measures in production systems development of the European Union as well as by protective settings for NATURA 2000 development. A rural population decrease of 47 % was measured between 1991-1995 at the southern coast of Crete in an investigation area around Omalos and Samaria Gorge (Papanastasis et al. 2004). The same process is as well verified for the Asterousia Mountains by a depopulation of 30 % between 1961 and 2011 (Kizos et al.2014).

The potential general nexus linking the main drivers via key processes to land degradation is shown in Fig.3. The four main drivers of tourism and agricultural developments, irrigated olive yard developments, large scale fenced goat pasturing and greenhouse developments for vegetable production were identified to be the main forces driving the degradation of soil, water, vegetation and landscape components in NATURA 2000 sites and its surroundings.

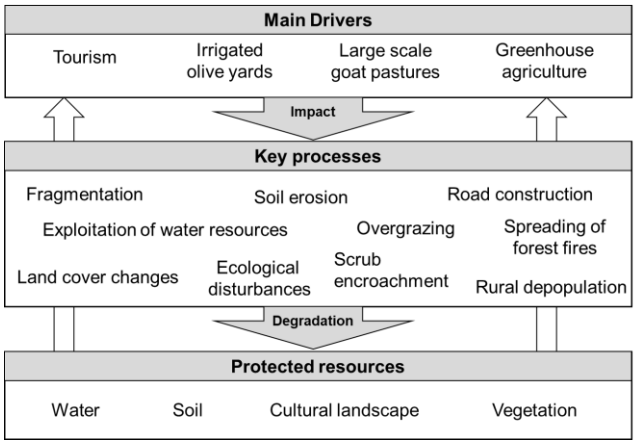

Fig.3 Drivers of change influencing selected processes of change and causing land degradation and degradation of cultural landscape qualities in NATURA 2000 sites and its surroundings of Crete

The four key drivers influencing the NATURA 2000 sites can be described as follows:

(a) General tourism developments and tourism related land uses: Since the 1980s Crete has faced an increased tourism development in number and intensity of beach tourism with large scale touristic estate developments especially on the northern coast. Organized and individual day tourism increased, mainly in the form of ecotourism e.g. by hiking activities. In peripheral communities this driver leads to the abandonment of settlements, depopulation, habitat fragmentation, water scarcity, direct vegetation and ecosystem modifications and ecosystem destruction as well as increased disturbances in NATURA 2000 sites (Kourgialas et al. 2018, Andriotis 2006). See Fig. 4a.

(b) Developments in irrigated olive yards: Large scale irrigated olive yard developments in many Cretan landscapes have been established on arable land formerly used for rain fed agriculture, on former pasture lands as well as on old terrace systems (Karydas et al. 2009, Allen et al. 2006; Price and Nixon 2005, Frederic and Krahtopodlou 2000). Non-irrigated olive cultures have been replaced on a large scale by irrigated olive cultivation. This driver leads to increased water consumption, changes in cultural landscapes and to a strong decrease of the agricultural land use diversity. See Fig. 4b.

(c) Intensive fenced large-scale goat pasture systems: Severe vegetation degradation is caused by management changes from extensive shepherds herding of sheep and goats to large scale fenced areas and the concentration on intensive goat husbandry (with external fodder supply throughout the dry months of the year, when winter and early spring vegetation has run out). See Fig. 4d. The key degradation driver seems to be overstocking and overgrazing of the open landscapes and a problematic grazing management with sheds, feeding facilities and watering places for the animals (Kosmas et al. 2015, Kairis et al. 2015, Lorent et al. 2009). Dispersed extensive grazing of browsing animals is also found in areas of less intensive goat and sheep grazing.

(d) Large scale greenhouse farming developments: The production of vegetables in greenhouses is an important feature of agricultural intensification stimulated by road developments. See Fig. 4c. A strong relief modification by large scale terracing is applied as a baseline for construction; in some places soils and sediments have been excavated from dolines (formerly used for grazing or cereal production) and serve as basic soil substrates in greenhouses (Fig. 4e). Greenhouse farming has a high water consumption and the applied pesticides, herbicides and nutrients bear the risk of negatively influencing public health (Daliakopoulos et al. 2016, Tsatsakis et al. 2008, Anton et al. 2014) and the NATURA 2000 area biodiversity qualities. Furthermore, pollution with macro- and micro-plastics originating from abandoned and degrading greenhouses is affecting both marine and terrestrial ecosystems (Gundogdu and Cevik 2017).

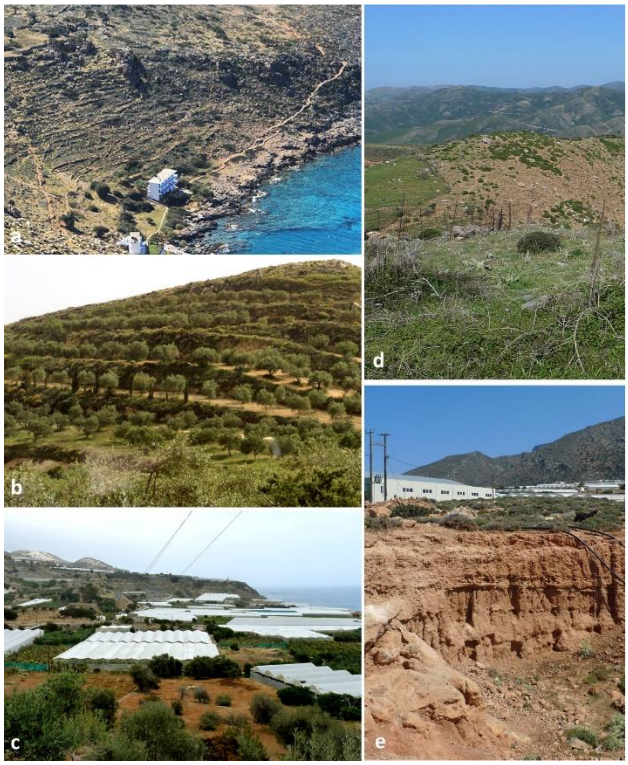

Fig.4 Major drivers affecting Natura 2000 sites – a) Tourism development near Loutro, Western Crete; b) Irrigated olive yards with terracing, Central Crete; c) large scale greenhouse agriculture near Arvi, Southern Crete; d) degraded slopes with intensive fenced large-scale goat pastures near Elos, Western Crete; e) soil excavation for greenhouse substrates near Elafonisi, Western Crete

Important processes observed and identified in the land use change – landscape degradation nexus and described for the selected NATURA 2000 sites are (see Figures 5-7):

(1) depopulation of remote villages and single farmsteads leading to abandoned cultural landscapes and conversion into summer houses (Fig. 6a),

(2) road construction for agricultural, forestry or energy production purposes (Fig. 5d),

(3) frequent large-scale forest fires causing habitat destruction (Fig. 5a),

(4) vegetation degradation on pastures and forest vegetation by overgrazing (Fig. 5b, 5c and 5e),

(5) water resources management measures by intensive water abstraction and modification of water household parameters for irrigation including the desiccation of rivers and water pollution (Fig. 6b and 6c),

(6) soil erosion caused by overgrazing in fenced goat pastures (fig 6d), in olive yards (Fig. 6e) as well as geo-morphological surface change by greenhouse construction (Fig. 4c),

(7) slope modification for the construction of greenhouses (Fig. 4e) or irrigated olive yards (Fig. 4b),

(8) disturbances caused by the increasing individual day tourism (Fig. 7b and 7c),

(9) habitat fragmentation of landscapes by agricultural, forest and road infrastructure developments (Fig. 7a),

(10) land cover changes e.g. from extensive pastures to olive yards, pastures to greenhouses and fenced animal husbandry on areas of formerly shepherds herding of goat and sheep (fig 4 in all photos), as well the replacement of grasslands/cereal fields by forests on terraces (Fig. 7d),

(11) abandonment of characteristic cultural landscape elements including terraces, pathways, stonewalls and former agricultural infrastructure as storage buildings and threshing sites (Fig. 7a).

Fig.5 Examples of processes with an impact on land degradation – a) Slopes affected by forest fires near Kapetaniana, Southern Crete; b) Tree damaged by goat grazing near Anopolis, Western Crete; c) Destruction of vegetation in fenced goat and sheep browsing area, Central Crete; d) Country road construction for goat pasture management near Anopolis, Western Crete; e) Construction of sheds, watering and fodder places for animals, near Anopolis, Western Crete

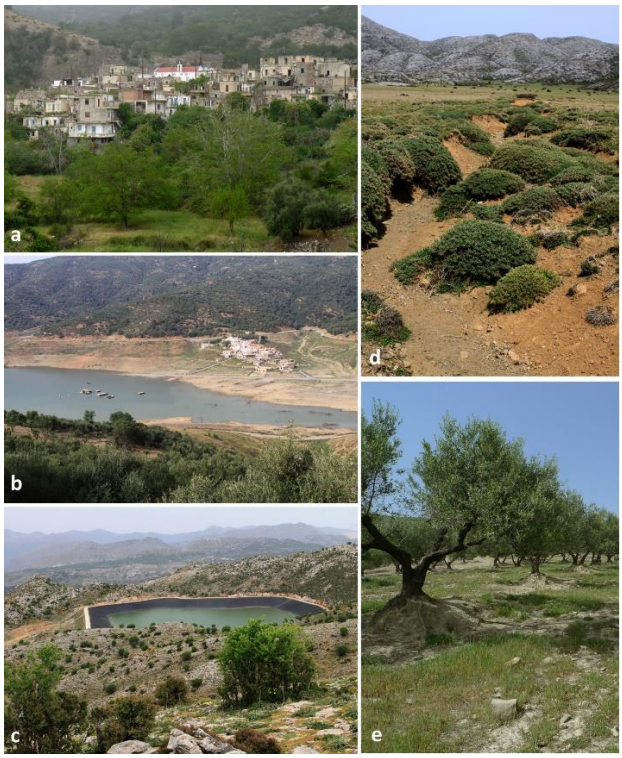

Fig.6 Examples of processes with an impact on land degradation – a) Depopulated village (Kalami, Southern Crete); b) Drowned village (Sfendili) in a large-scale water storage facility in Northern Crete; c) Water pond near Anogia, Central Crete; d) Soil erosion as a result of goat trampling and overgrazing in Ida mountains, Central Crete; e) Olive plantation affected by soil erosion in marly substrates near Agia Varvara, Central Crete

Fig.7 Examples of processes with an impact on land degradation – a) Habitat fragmentation by construction of infrastructure near Livadia, Western Crete; b) Disturbed Juniper thicket at Kedrodassos beach, Western Crete; c) Parking lots in dune formations for day tourism at Elafonisi beach, Western Crete; d) Forest plantations on abandoned traditional terraces near Agios Ioannis, Western Crete; e) Abandoned traditional threshing site (Aloni) near Elafonisi, Western Crete

These processes influence the status in the regions analysed in a complex way and are contradictory to the NATURA 2000 objectives named in the documents submitted to the European Commission during the NATURA 2000 designation process in many cases. The management of these processes should be further detailed as a subject for research based on mapping, process identification and process analysis, process modelling and assessments usable for the regional local and flora/fauna development planning in NATURA 2000 management, which is still lacking at all sites investigated in this study.

Status and intensity of land degradation processes in selected NATURA 2000 sites

The analysis of processes influencing land degradation in the five investigated areas is based on literature analysis and field observations during the field work campaign. Two summarizing tables are study results by substantiating (i) the drivers affecting the NATURA 2000 sites studied (Table 2) and (ii) the intensity of the observed processes with influence on land degradation and the transformation of the traditional cultural landscapes of Crete (Table 3).

The analysis of factors driving the change and the degradation in the selected NATURA 2000 sites (Table 2) highlights the complexity of landscape intensification processes. Direct and indirect forces and processes are often obscure and hard to distinguish. Hence, such drivers are rarely considered and in the focus of NATURA 2000 management activities. Table 2 therefore gives the drivers investigated as most important in the studied areas.

The activity of key processes with influence on land degradation and the degradation of traditional cultural landscapes (see Table 3) is assessed by the authors' expert judgments and based on field observations. The analysis reveals a complex mosaic of drivers and related processes in the NATURA 2000 sites. It is based on a wide range of overlaying and interrelated processes (Fig.3). All NATURA 2000 sites and their surroundings face more or less intensive forms of land degradation. Nevertheless, processes with impact on the Table 1 habitat types are also in interdependences with the processes in the categories of "land degradation" (soil, water and vegetation) and "landscape degradation" (cultural landscape). Such complex landscape nexus is not yet clarified and should be subject of further research.

NATURA 2000 area threats and pressures and the land use – land degradation nexus

The land use – land degradation nexus analysis based on the threats and pressures taken from the official NATURA 2000 site characteristic descriptions in Table 1a,b (last published update 12/2016) and the drivers and processes of land degradation in Crete analysis of our research results in contradictory outcomes. In the following these threats and pressures are shortly interpreted regarding the drivers and processes. While the NATURA 2000 documentation denotes a set of single threats and pressures per site, our approach tries to address the interlinkages (nexus) of such factors in the context of land degradation.

Table 2 Drivers affecting the selected NATURA 2000 sites and its surroundings on Crete: (1) Lasithi plateau, (2) Ekvoli Geropotamou Mesara, (3) Lefka Ori Kai Parakti, (4) Nisos Elafonisos and Paralia Apo Chrysoskalitissa Mechri Akrotirio Krios, (5) Giouchtas Farangi Agias Eirinis

Drivers	**NATURA 2000 areas and surroundings**				
	1	**2**	**3**	**4**	**5**
General tourism developments and tourism related land uses	intense day tourism	medium tourism pressure by summer and day tourism	Intense concentration of day tourism activities along Samaria Gorge, some ecotourism	high pressure by day tourism	low tourism pressures by sports
Irrigated olive yard agricultural developments,	no direct influence; indirect influence via water uptake	strong influence in the wider surroundings	no influence	multiple new developments	strong influence in the surroundings
Intensive fenced large-scale goat pasture agricultural systems	high influence in the surrounding mountains	no influence; some extensive grazing of sheep and goat	high influence in most areas	high influence in the surrounding mountains and on abandoned cultural land	few influence, some extensive grazing of sheep and goat
Large scale greenhouse farming developments	no direct influence; indirect via water uptake	strong influence in the wider surroundings	some developments on coastal sites	multiple new developments	few developments in the surroundings

Table 3 Activity of processes with influence on land degradation and the degradation of traditional cultural landscapes at the selected NATURA 2000 sites and its surroundings: (1) Lasithi plateau, (2) Ekvoli Geropotamou Mesara, (3) Lefka Ori Kai Parakti, (4) Nisos Elafonisos and Paralia Apo Chrysoskalitissa Mechri Akrotirio Krios, (5) Giouchtas Farangi Agias Eirinis

Actual processes with impact on land degradation	**NATURA 2000 areas and surroundings**				
	1	**2**	**3**	**4**	**5**
Depopulation	ongoing	not visible	strong	not visible	not visible
Road infrastructure development	few	strong	few	strong	not visible
Forest fires	in the surrounding	not visible	strong	strong	some
Overgrazing	strong in the surroundings	not clearly visible	very strong	very strong	strong
Water resources in quality /quantity	extraction of water from the plain	strong use of water for irrigation	not clearly visible	intensive	intensive in the surroundings
Soil erosion	water erosion strong in the surroundings	wind erosion active	active	active wind and water erosion	some activities, intensive in the surroundings
Slope modification	ponds developments, overall minor	multiple modification for divers uses	minor modification	large scale slope modification and excavation of dolines for olive and greenhouses	large scale slope modification for wine and olive
Disturbances	minor by local tourism	medium by day and beach tourism	high along Samaria gorge; low in most of the area	very high by day beach tourism	few by sport and recreation
Habitat fragmentation	minor	high caused by new infrastructure activities	medium by activation of old road infrastructure	high by road and agricultural activities	minor; strong in the surroundings
Land cover change	conversion of grasslands to degraded forest; forest succession	diverse local changes, multiple abandoned sites partly military sites	enlargement of forest areas;	change to irrigated olive yards and greenhouses; multiple abandoned sites, some tourism infrastructure, abandoned water infrastructure	forest succession on former terraces
Abandonment of important cultural landscape elements	windmills, divers water pump infrastructure, flooded pastures	old olive orchards, extensive arable land (cereal production) pasture of dunes	terraces, historical paths network, stone-walls around arable former arable fields, old olive orchards, threshing sites	threshing sites, extensive doline land grazing, extensive arable land (cereal production)	pastures, orchards, terraces on lower slopes

The threats and pressures information in the documentation mentioned above is qualitatively described and ranked and attributed with importance levels (high/medium/low) in Table 1a,b. The currentness of the information is not clarified as the threats and pressures may relate to the moment of first reporting for the foundation of the Cretan NATURA 2000 sites. The documentation states important and significant information on highly ranked (important) pressures and threats such as grazing, erosion, fire, fertilization or taking and removal of terrestrial plants. This information is usable as general information about impacts on the habitat types. The land use intensity changes related to agriculture, tourism and infrastructure have been identified as important drivers of degradation of various habitat types in Crete, comparable to and observed in other Mediterranean landscapes (Dimopoulos et al. 2006; Cuttelod et al. 2008; Underwood et al. 2009). Nevertheless, a comparison of status and intensity of land degradation in selected NATURA 2000 sites and of the respective effects of different land uses as described above is difficult as long as detailed model analysis and more statistical data are lacking.

The comparison of major threats stated in the NATURA 2000 documentation and the authors' field observations explained in sections above by the drivers effecting NATURA 2000 sites and its surroundings on Crete reveals some open fields of problem analysis for a better understanding of the land use – land degradation nexus compared with the NATURA 2000 documentation:

(1) The general touristic developments and tourism related land use activities are often named as threats and pressures in the NATURA 2000 documentation without a management goal formulation or related protective measures as subject of the still missing management planning;

(2) The irrigated olive yard agricultural developments are not named as a threat or pressure;

(3) The large-scale fenced goat pastures are named a as threat and pressure by overgrazing and animal breeding, agricultural structures, new agricultural buildings (e.g. sheds in the landscape) in most of the Natura 2000 sites;

(4) Large-scale greenhouse farming as well as the road infrastructure developments as a problem of habitat fragmentation are not a subject of threats and pressures documentation for NATURA 2000 sites.

Landscape degradation processes in the Mediterranean landscapes

The list of current processes determined by land uses with impact on land degradation with importance on landscape, vegetation, soil and water systems is interpreted by the authors as a missing link or knowledge gap for a better understanding of the regeneration potential of Mediterranean landscapes. Land degradation is defined by the United Nations Convention to Combat Desertification (UNCCD 1994) convention as "a reduction or loss, in arid, semi-arid and dry sub-humid areas, of the biological or economic productivity and complexity of rain fed cropland, irrigated cropland, or range, pasture, forest and woodlands resulting from land uses or from a process or combination of processes, including processes arising from human activities and habitation patterns, such as: soil erosion caused by wind and/or water; deterioration of the physical, chemical and biological or economic properties of soil; and long-term loss of natural vegetation". When focussing on the degradation of cultural landscapes and by widening the definition from land degradation analysis towards the analysis of the landscape degradation "as an irreversible or non-resilient system change to a landscape that affects the landscape system components (i.e., their geo-factors, land use and inter-linkages) and the natural and cultural capacities of the landscape (productive, ecological and social structure, processes and landscape functions" (Meyer et al. 2017), the focus of such an analysis in the Mediterranean should be widened into the landscape context - e.g. on the economically very important tourism and agricultural intensification developments in the Cretan context – into a new land use and landscape pattern and landscape configuration perspective.

For both land degradation and landscape degradation analysis of the soil, water and vegetation degradation is required since the authors observations on revegetation, e.g. after fire events in degraded limestone landscapes, shows a fast renewal of the former, mostly already modified status of the Phrygana vegetation. A new view on the degradation problem in the Mediterranean is needed since the changes in the intensity of land uses are not known in a landscape systems context. The land use / landscape changes influencing the NATURA 2000 sites are determined by general and external factors as changes in agricultural or regional policies, management techniques, land use measures and practices interpreted in our study as land use intensity changes. A degradation of vegetation types listed in Habitat directive (Council Directive 92/43/EEC) was found during the field analysis. A system degradation of water management facilities is observed by abandoned irrigation systems and strongly modified surface waters of the karst water systems and multiple other water related infrastructure measures. Soil degradation is obviously found in areas of land and slope surface change by terrace abandonment, active and former erosion gullies, and geomorphologic levelling for greenhouses and by soil abstraction. Landscape degradation is interpreted by the conversion of abandoned or extensive cultural landscape features to a more industrial and less diverse land use pattern, especially by olive yard plantation. Problematic intensification practices of bee keeping and honey production were observed in the analysed NATURA 2000 sites by the introduction of scale insects on pine trees providing a continuous nectar source throughout the entire year. This practice results in local dieback of entire pine tree populations and potentially increases forest fire risk.

CONCLUSION

The land use – land degradation nexus in Mediterranean landscapes was analysed by status, drivers and key pressures using the example of selected areas comprising NATURA 2000 sites. The changing land use intensity e.g. in the fenced large-scale goat pasture systems may

promote further and strong vegetation degradation and is of importance regarding the degradation of soil and water resources. The large-scale greenhouse developments and the irrigated olive yard developments are primarily driving degradation pressures in the landscape context but also lead to degradation of water, soil and vegetation resources. The tourism developments and tourism related land uses including road developments have a significant impact on landscape degradation and may have influence the depletion of vegetation, water and soil resources.

Further analysis is needed for a more complex identification of land degradation within the land use intensity context. The critical limits/levels or thresholds of the observed land degradation are not yet fully known and the respective interpretation of landscape degradation needs to be revised as well. Detailed research is needed for a better understanding of the land use – land degradation nexus in Mediterranean landscapes especially for karst landscapes since degradation processes are less known for soils developed on calcareous rocks, as well as their potential of natural and managed revegetation is not yet scientifically clarified. The Land Degradation Neutrality goal of the UN has to be investigated, measured and analysed by 3 key indicators: (i) by the trends in land cover, (ii) by the trends in land productivity or functioning of the land and (iii) by trends in above and below ground carbon stocks (IUCN 2015). Based on the results of the presented study we suggest that future research on land degradation should lay a stronger emphasis on trends in land use and land cover developments as drivers of change and as a basis for the analysis of the temporal changes in the vegetation, water, soil and cultural landscape systems. At the same time, it is important to understand how the multiple related processes/pressures and threats identified in this study influence land productivity and carbon stocks in the context of climate change.

The aim of the presented explorative study has been to work out by literature, photo and expert field analysis important drivers and ongoing processes in landscape systems to clarify the land use – land degradation nexus in Mediterranean landscapes of Crete. Next steps should be developed towards a better and integrative analysis and description of the sensitivity of the landscapes and its patches and by the determination of thresholds to clarifying the processes changes. Methods used in the expert approach in our study by the example of Crete should be further specified by diverse data sets for a better process description and by model developments applicable for policy and planning advice.

Acknowledgement

We thank Susann Heinrich for the fruitful discussion and the two anonymous reviewers for the helpful comments.

References

Allen, H. d., Randall, R. E., Amable, G. S., and Devereux, B. J. (2006). The impact of changing olive cultivation practices on the ground flora of olive groves in the Messara and Psiloritis regions, Crete, Greece. *Land Degradation Development*, *17*(3), 249–273.

Andriotis, K. (2006). Researching the development gap between the Hinterland and the Coast - Evidence from the Island of Crete. *Tourism Management*, *27*(4), 629–639.

Antón, A., Torrellas, M., Núñez, M., Sevigné, E., Amores, M. J., Muñoz, P., Montero, J. I. (2014). Improvement of agricultural life cycle assessment studies through spatial differentiation and new impact categories: case study on greenhouse tomato production. *Environmental Science & Technology*, *48*(16), 9454–9462.

Barredo, J. I., Caudullo, G., Dosio, A. (2016). Mediterranean habitat loss under future climate conditions: Assessing impacts on the Natura 2000 protected area network. *Applied Geography*, *75*, 83–92.

Cuttelod, A., García, N., Abdul Malak, D., Temple, H., Katariya, V. (2009). The Mediterranean: a biodiversity hotspot under threat. In J.-C. Vié, C. Hilton-Taylor, S. N. Stuart (Eds.), *The 2008 Review of The IUCN Red List of Threatened Species* (pp. 1–13). Gland u.a.: IUCN.

Dal Cin D´Agata, C., Skoula, M., Brundu, G. (2009). A preliminary inventory of the alien flora of Crete (Greece). *Bocconea*. (23), 301–308.

Daliakopoulos, I. N., Pappa, P., Grillakis, M. G., Varouchakis, E. A., Tsanis, I. K. (2016). Modeling Soil Salinity in Greenhouse Cultivations Under a Changing Climate With SALTMED. *Soil Science*, *181*(6), 241–251.

Dimitrakopoulos, P. G., Memtsas, D., Troumbis, A. Y. (2004). Questioning the effectiveness of the Natura 2000 Special Areas of Conservation strategy: the case of Crete. *Global Ecology and Biogeography*, *13*(3), 199–207.

Dimitrou, E. (2008). Conservation plan for the Mediterranean Temporary Ponds habitat in W. Crete. In E. Dimitrou A. Diapoulis (Eds.), *Actions for the conservation of Mediterranean Temporary Ponds of Crete. Final Report:* Institute of Inland Waters Hellenic Centre for Marine Research.

Dimopoulos, P., Bergmeier, E., Fischer, P. (2006). Natura 2000 Habitat Types of Greece Evaluated in the Light of Distribution, threat and Responsibility. *Biology and Environment*, *106B* (3 European Vegetation in the 21st Century), 175–187.

Donatos, G., Zairis, P. (1991). Seasonality of foreign tourism in the Greek island of Crete. *Annals of Tourism Research*, *18*(3), 515–519.

European Commission. Natura 2000 Public Database. Retrieved from http://ec.europa.eu/environment/nature/natura2000/access_data/index_en.htm

Council Directive 92/43/EEC of 21 May 1992 on the conservation of natural habitats and of wild fauna and flora, European Commission 1992.

European Commission. (2014). *Farming for Natura 2000: Guidance on how to support Natura 2000 farming systems to achieve conservation objectives, based on Member States good practice.*

European Commission. (2015). *Natura 2000 and Forests Part I-II*. Technical Report - 2015 – 088.

European Economic Interest Group. *Scoping study on the review of links and complementary between Natura 2000 and cultural sites*. Retrieved from http://ec.europa.eu/environment/nature/natura2000/financing/docs/Scoping%20study%20N2000%20and%20culture.pdf

Farina, A. (2000). The Cultural Landscape as a Model for the Integration of Ecology and Economics. *BioScience*, *50*(4), 313.

Gündoğdu, S., Çevik, C. (2017). Micro- and mesoplastics in Northeast Levantine coast of Turkey: The preliminary results from surface samples. *Marine Pollution Bulletin*, *118*(1-2), 341–347.

Hellenic Statistical Authority. (2019). Crops Surveys / 2017. Retrieved from http://www.statistics.gr/en/statistics/-/publication/SPG63/-

Ispikoudis, I., Lyrintzis, G., Kyriakakis, S. (1993). Impact of human activities on Mediterranean landscapes in western Crete. *Landscape and Urban Planning*, *24*(1-4), 259–271.

IUCN. (2015). *Land Degradation Neutrality: Implications and opportunities for conservation*. Technical Brief 2nd Edition. Nairobi.

Kairis, O., Karavitis, C., Salvati, L., Kounalaki, A., Kosmas, K. (2015). Exploring the Impact of Overgrazing on Soil Erosion and Land Degradation in a Dry Mediterranean Agro-Forest Landscape (Crete, Greece). *Arid Land Research and Management*, *29*(3), 360–374.

Karamesouti, M., v. Detsis, Kounalaki, A., Vasiliou, P., Salvati, L., Kosmas, C. (2015). Land-use and land degradation processes affecting soil resources: Evidence from a traditional Mediterranean cropland (Greece). *CATENA*, *132*, 45–55.

Karydas, C. G., Sekuloska, T., Silleos, G. N. (2009). Quantification and site-specification of the support practice factor when mapping soil erosion risk associated with olive plantations in the Mediterranean island of Crete. *Environmental Monitoring and Assessment*, *149*(1-4), 19–28.

Kizos, T., Detsis, V., Iosifides, T., Metaxakis, M. (2014). Social Capital and Social-Ecological Resilience in the Asteroussia Mountains, Southern Crete, Greece. *Ecology and Society*, *19*(1).

Kosmas, C., Detsis, V., Karamesouti, M., Kounalaki, K., Vassiliou, P., Salvati, L. (2015). Exploring Long-Term Impact of Grazing Management on Land Degradation in the Socio-Ecological System of Asteroussia Mountains, Greece. *Land*, *4*(3), 541–559.

Kosmas, C., Karamesouti, M., Kounalaki, K., Detsis, V., Vassiliou, P., Salvati, L. (2016). Land degradation and long-term changes in agro-pastoral systems: An empirical analysis of ecological resilience in Asteroussia - Crete (Greece). *CATENA*, *147*, 196–204.

Kourgialas, N. N., Karatzas, G. P., Dokou, Z., Kokorogiannis, A. (2018). Groundwater footprint methodology as policy tool for balancing water needs (agriculture & tourism) in water scarce islands - The case of Crete, Greece. *The Science of the Total Environment*, *615*, 381–389.

Kušová, D., Těšitel, J., Matějka, K., & Bartoš, M. (2008). Biosphere reserves—An attempt to form sustainable landscapes. *Landscape and Urban Planning*, *84*(1), 38–51.

Lorent, H., Sonnenschein, R., Tsiourlis, G. M., Hostert, P., Lambin, E. (2009). Livestock Subsidies and Rangeland Degradation in Central Crete. *Ecology and Society*, *14*(2).

Mazzoleni, S., Di Pasquale, G., Mulligan, M., Di Martino, P., Rego, F. (Eds.). (2004). *Recent dynamics of the Mediterranean vegetation and landscape* (Reprinted.). Chichester: Wiley.

Médail, F. (2017). The specific vulnerability of plant biodiversity and vegetation on Mediterranean islands in the face of global change. *Regional Environmental Change*, *17*(6), 1775–1790.

Médail, F., Myers, N. (2005). Mediterranean Basin. In R. A. Mittermeier, G. P. Robles, M. Hoffmann, J. Pilgrim, T. Brooks, C. G. Mittermeier, J. Lamoreux G.A.B. Da Fonseca (Eds.), *Hotspots revisited: Earth's biologically richest and most endangered terrestrial ecoregions* (2nd ed., pp. 144–147). Washington.

Meyer, B. C., Mezősi, G., Kovács, F. (2017). Landscape degradation at different spatial scales caused by aridification. *Moravian Geographical Reports*, *25*(4), 271–281.

Papanastasis, V. P., Ispikoudis, I., Arianoutsou, M., Kakouros, P., Kazaklis, A. (2004). Land-use changes and landscape dynamics in Western Crete. In S. Mazzoleni, G. Di Pasquale, M. Mulligan, P. Di Martino, F. Rego (Eds.), *Recent dynamics of the Mediterranean vegetation and landscape* (pp. 81–93). Chichester: Wiley.

Price, S., Nixon, L. (2005). Ancient Greek Agricultural Terraces: Evidence from Texts and Archaeological Survey. *American Journal of Archaeology*, *109*(4), 665–694.

Rother, K. (1993). *Der Mittelmeerraum: Ein geographischer Überblick; mit 6 Tabellen*. *Teubner Studienbücher der Geographie*. Stuttgart: Teubner.

Tsatsakis, A. M., Tzatzarakis, M. N., Tutudaki, M. (2008). Pesticide levels in head hair samples of Cretan population as an indicator of present and past exposure. *Forensic Science International*, *176*(1), 67–71.

Turland, N. J., Chilton, L., Press, J. R. (1993). *Flora of the Cretan Area: Annotated Checklist Atlas*. London: HMSO.

UNCCD. (1994). United Nations convention to combat desertification in those countries experiencing serious drought and/or desertification, particularly in Africa. Retrieved from https://treaties.un.org/doc/Treaties/1996/12/19961226%2001-46%20PM/Ch_XXVII_10p.pdf

Underwood, E. C., Viers, J. H., Klausmeyer, K. R., Cox, R. L., Shaw, M. R. (2009). Threats and biodiversity in the Mediterranean biome. *Diversity and Distributions*, *15*(2), 188–197.

Watrous, L. V. (1982). *Lasithi: A history of settlement on a highland plain in Crete*. *Hesperia Suppl: Vol. 18*. Princeton, NJ: American School of Class. Studies at Athens.

sciendo

Journal of Environmental Geography 13 (1–2), 25–30.

DOI: 10.2478/jengeo-2020-0003
ISSN 2060-467X

IMPACT OF LOW-DOSE MUNICIPAL SEWAGE SLUDGE COMPOST TREATMENTS ON THE NUTRIENT AND THE HEAVY METAL CONTENTS IN A CHERNOZEM TOPSOIL NEAR ÚJKÍGYÓS, HUNGARY: A 5-YEAR COMPARISON

Zsuzsanna Ladányi[1]*, Katalin Csányi[1], Andrea Farsang[1], Katalin Perei[2], Attila Bodor[2], Adrienn Kézér[1], Károly Barta[1], Izabella Babcsányi[1]

[1]Department of Physical Geography and Geoinformatics, University of Szeged, Egyetem u. 2-6, H-6720 Szeged, Hungary
[2]Department of Biotechnology, University of Szeged, Közép fasor 52, H-6726 Szeged, Hungary
*Corresponding author, e-mail: ladanyi@geo.u-szeged.hu

Research article, received 20 February 2020, accepted 15 April 2020

Abstract

Agriculture is one of the major fields, where sewage sludge can be used. Its high nutrient content can contribute to the improvement of important soil properties, such as nutrient content, water balance and soil structure. However, sewage sludge may contain hazardous components, such as pathogens and pollutants. Therefore, it is important to monitor the effects of its field application. In this paper, we assessed the impacts of two low-dose (2.5 m^3/ha) municipal sewage sludge compost applications (in 2013 and in 2017) in a 5.6 ha arable land in southeast Hungary (near Újkígyós), located in the Hungarian Great Plain. The nutrient and the heavy metal contents in the upper soil layer (0-30 cm) of the studied Chernozem soils were compared between two sampling campaigns in 2013 (before the compost applications) and in 2018 (after the compost applications). Basic soil properties (pH, salinity, humus content, carbonate content, Arany yarn number) complemented with nutrient content (K_2O, P_2O_5, NO_2+ NO_3) and heavy metal content (Cd, Co, Cr, Cu, Ni, Pb and Zn) analyses were performed. The results show that no significant change can be noticed in the baseline parameters over the 5-year period. The slight increase in the P_2O_5, NO_2+ NO_3 content is closely related to the beneficial effects of the sewage sludge deposition. The soil-bound heavy metal load did not increase significantly as a result of the compost treatments, only nickel showed a slight increase in the topsoil. In all cases the heavy metal concentrations did not reach the contamination thresholds set by Hungarian standards. The results provided positive evidences proving that low dose municipal sewage sludge compost disposal on agricultural land is safe, and can be considered as a sustainable soil amendment for agriculture in compliance with legal requirements.

Keywords: sewage compost, treatment, heavy metal, nutrient content, agricultural application

INTRODUCTION

Sewage sludge, as a by-product of the wastewater treatment, is rich in organic and inorganic plant nutrients, therefore it can be used as an alternative to mineral fertilizers in agricultural fields (Singh 2008). Sewage sludge disposal in agriculture is an increasingly popular way of reusing sewage sludge, as it enables the recycling of valuable components, such as organic matter (Csányi et al., 2018), N, P and other nutrients (Kádár, 2013; Babcsányi et al., 2019). Besides amending the soil with plant nutrients and organic matter, sewage sludge application can introduce into the soil some hazardous pollutants, such as heavy metals (Moreno et al., 1997).

According to the 9th Technical assessment on the implementation of the Urban Waste Water Treatment Directive of the European Union (TA UWWTD, 2017), the overall reported production of sludge in EU28 in 2014 was just over 8.7 Mio T/year, from which 44% was re-used for "Soil and agriculture". According to the National Sewage Sludge Management Strategy of Hungary (2014-2023) 38% of the total amount of municipal sludge was used in agriculture as sludge, compost or compost product in Hungary in 2013 (while 46%, 5%, 2%, 9% were used for land reclamation, energy purposes, landfilling and other purposes, respectively). Land application of sewage sludge is regulated only by the concentration of heavy metals as specified in the Council Directive 86/278/EEC (CEC, 1986). Hungary, as most of the EU countries, adopted even more stringent limits for sludge use in agriculture by setting lower limits for heavy metals in comparison with the European Directive 86/278/EEC, furthermore, introduced limit values for other elements as well (Hudcová et al., 2019)

For the disposal of sewage sludge, it is essential to assess the composition and its impact on the soil as it can change the macro- and micro-nutrient content, organic matter content, adsorption capacity, soil structure and water management (Kádár, 2013). In addition, sewage sludge can be a potential source of hazard due to its heavy metal content and microbial composition (Moreno et al. 1997) (Vermes, 1997; Kocsis, 2011). Increased heavy metal content can cause a change in the soil biota (Uri et al., 2005), and can ultimately endanger human health by entering the food chain. Although nutrients (eg. N, P) are essential for plant growth, when applied excessively, however, they may accumulate in the topsoil, can be leached and transported off-site by drainage systems or can be transported by erosion posing risk to (sub)surface

waters (Hudcová et al., 2019). Therefore, it is important to monitor the long-term effects of the land-application of sewage sludge.

This study investigates the impacts of two applications of municipal sewage sludge compost (in 2013 and in 2017) on Chernozem soils based on comparing soil parameters between 2013 and 2018. For this purpose, potential changes in the basic soil physical and chemical features were assessed (pH, salt content, humus content, carbonate content, Arany yarn number, nutrient and heavy metal content). The specific objective is to evaluate if the low dose application of municipal sewage sludge compost causes a significant increase in the soil-bound heavy metal concentrations and how the nutrient content is affected.

STUDY AREA AND METHODS

The study area is located near Újkígyós, a settlement in the Trans-Tisza region of the Hungarian Great Plain (Fig. 1). The studied plot is an 5.6 ha arable land with highly fertile Chernozem soils, where municipal sewage sludge compost was spread twice in 2013 and in 2017 (2.5 m^3/ha). The parameters of the deposited sludge compost are presented in Table 1. Prior to the first sludge compost application in 2013, a soil protection plan was compiled for the study area based on a detailed soil survey (according to the Hun. Decree 40/2008. (II. 26.), the results of which provide the basis for the comparison with data from 2018. In 2013, composite samples were collected from the upper soil (0–30 cm) from 5 sampling areas of 2500 m^2 each within the study area.

The arable land was re-surveyed in March 2018, when 6 composite samples were taken from the upper soil (0-30 cm). During the designation of the sampling areas in 2018, a special attention was paid to include areas affected by sewage sludge compost disposal (1,2, 3,4 plots in Fig.1.) and also unaffected ones (6,7plots in Fig.1.) (Kézér, 2018; Pálffy et al., 2018).

Prior to measurements, samples were dried in the laboratory at 40°C in an oven and powdered (< 2 mm) after removing larger organic debris and foreign material. The pH (H_2O) was determined using a WTW inoLab 720p pH-meter, the total water soluble salt content was determined by measuring the conductivity (OK-104 conductivity meter) of the saturated soil paste, the carbonate content ($CaCO_3$) was determined using the Scheibler type calcimeter according to the Hungarian standard procedure (MSZ-08-0206-2: 1978). To measure texture of the samples, Arany yarn test was carried out based on the MSZ-08-0205-1978 Hun. standard. Humus content was measured by colorimetry using Helios γ type spectrophotometer based on the MSZ 21470-52:1983 standard. The macronutrients P_2O_5 and K_2O were extracted using ammonium-lactate, while the nitrogen forms (NO_2^- + NO_3^- -N) were extracted with KCl-solution according to standard procedures (MSZ20135:1999), measured by FIA spectrometer. Prior to total metal analyses, soil samples were oven-dired again at 105°C for 24 hours. Powdered soil samples (0.5 g) were digested in aqua regia (hydrochloric acid : nitric acid = 3 : 1) in closed vessels in a microwave oven (Anton ar Multiwave 3000) The metal concentrations in the digested samples were measured by an inductively coupled plasma atomic/optical emission spectrometer (Perkin Elmer ICP-OES Optima 7000 DV) (according to the standard: MSZ 21470-50:2006).

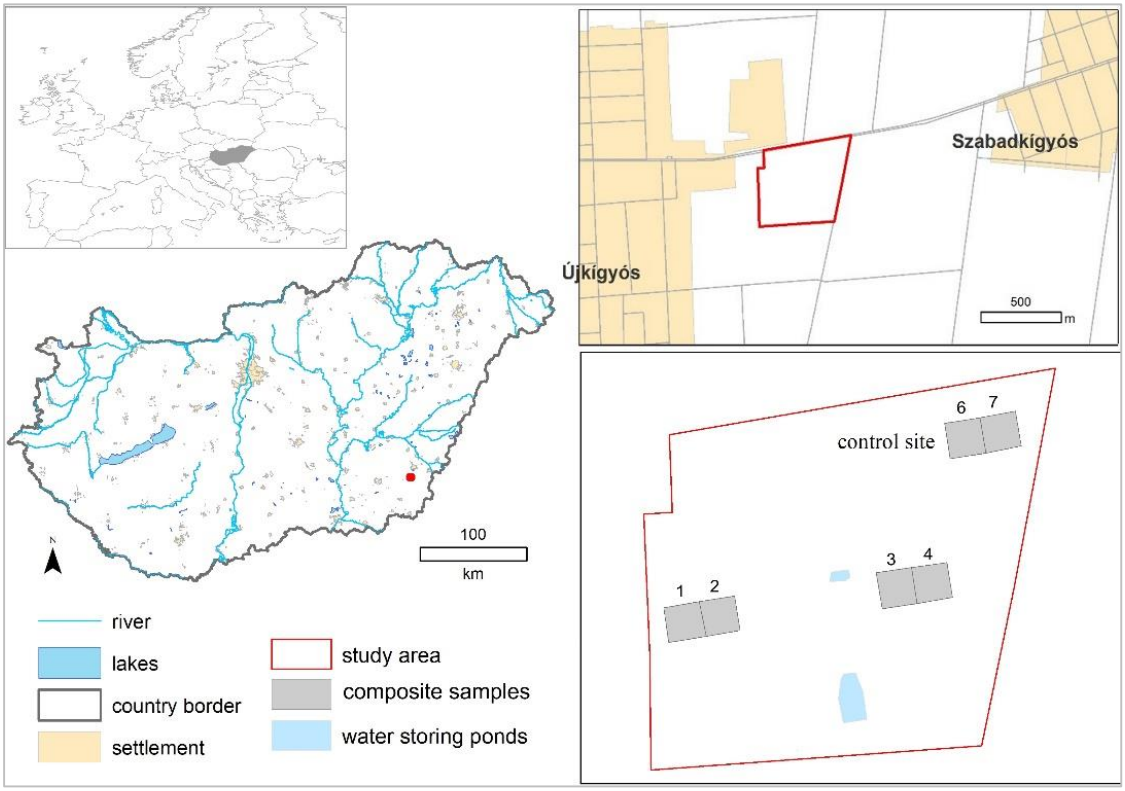

Fig. 1 Location of the study area (Újkígyós, Hungary)

Table 1 The quality of the applied municipal sewage sludge compost in 2013 and in 2017. The displayed Hungarian threshold for sewage sludge compost applied in agricultural land is specified in the Government Decree 40/2008. (II. 26.)

Parameters	Compost quality in 2013	Compost quality in 2017	Hungarian threshold
pH (d.w.)	7.64	8.48	-
Dry matter (mg/kg)	181000	24600	-
Organic matter (m/m%)	13.2	5.7	-
Total nitrogen (mg/kg dry w.)	59100	128700	-
Total phosphorus (P_2O_5 mg/kg dry w.)	19500	16500	-
Total potassium (K_2O mg/kg dry w.)	2560	5860	-
TPH (mg/kg)	<25	2588	1000
As (mg/kg)	67.6	15.5	25
Cd (mg/kg)	0.53	<1	5
Co (mg/kg)	<1	2.31	50
Total Cr (mg/kg)	14.2	17.6	350
Cu (mg/kg)	81.1	103	750
Hg (mg/kg)	0.64	0.46	5
Mo (mg/kg)	6.74	6.35	10
Ni (mg/kg)	12.4	15.1	100
Pb (mg/kg)	18.7	16.3	400
Se (mg/kg)	<1	<1	50
Zn (mg/kg)	444	542	2000

RESULTS AND DISCUSSION

The topsoil in the study area is characterized by slightly alkaline pH (7.0-8.1), low salt content (0.02%), moderate humus content (1.53 – 2.52%), low carbonate content (0.14-3.56 %) and a sandy loam texture (Fig. 2). There was no statistically significant change in these basic soil parameters between 2013 and 2018 ($p \leq 0.05$). As a result of the compost treatment, we would expect the humus content to increase, however a slight decrease can be noticed in the soil humus content in 2018. This observation may be explained by the intensive agricultural production. Another possible explanation may be the low-dose and short-term application of the compost at our study site. A previous field study found that a longer term (50 year-long) annually repeated treatment of arable land with organic amendments may be needed for a significant increase in the soil organic carbon pool (Sleutel et al., 2006).

There is a marked increase in the bioavailable N- and P- contents between 2013 and 2018, whereas no significant change can be observed in the K content (Fig. 3). These observations confirm that the sludge compost treatments supply with valuable plant nutrients agricultural soils and can be a convenient alternative for mineral fertilizers. Moreover, the compost treatment has a longer-term effect compared to mineral fertilizers, because the applied composts supply plants with macronutrients via a slow decomposition process of the organic compost components including the conversion of macronutrients (N, P, S, etc.) into inorganic (and mostly plant available) forms by microorganisms (Diacono and Montemurro, 2011).

Heavy metal load did not increase significantly as a result of the sewage sludge compost applications between 2013 and 2018 (Fig. 4). Among the investigated heavy metals, only nickel showed a slight increase in the topsoil. The heavy metal concentrations are far beyond the contamination thresholds according to the Hungarian standards for soil quality (6/2009. [IV. 14.] KvVM-EüM-FVM). Our results are in line with the outcome of a similar study conducted over a 6-year period that demonstrated that urban sewage sludge compost (mixed with poplar bark) treatments with similar heavy metal contents presented no danger in the short/medium term either to the environment or to crops even when applied in higher doses (80-160 t/ha) compared to the doses in the present field survey (Pinamonti et al., 1997). An incubation experiment also found that the increase in the total metal content following a single sludge addition (48 t/ha) was negligible in an agricultural calcareous silt loam soil using a mixture of sewage sludge (primary aerobic sludge) and cotton waste (Sánchez-Monedero et al., 2004). Similarly, Fang and coworkers (2016) observed no significant increase in the total metal content of a farmland alluvial soil displaying slightly alkaline pH due to a single application of sewage sludge compost at a rate of 48 t/ha. Nevertheless, they found that application intensified the leaching of heavy metals due to the increase in the dissolved organic matter of the soil solution. Depending on the metal content of the applied sewage sludge compost a significant increase in the total metal content of amended soils may be observed, which raises concerns about the long-term applicability of sewage sludge composts (Khadhar et al., 2020).

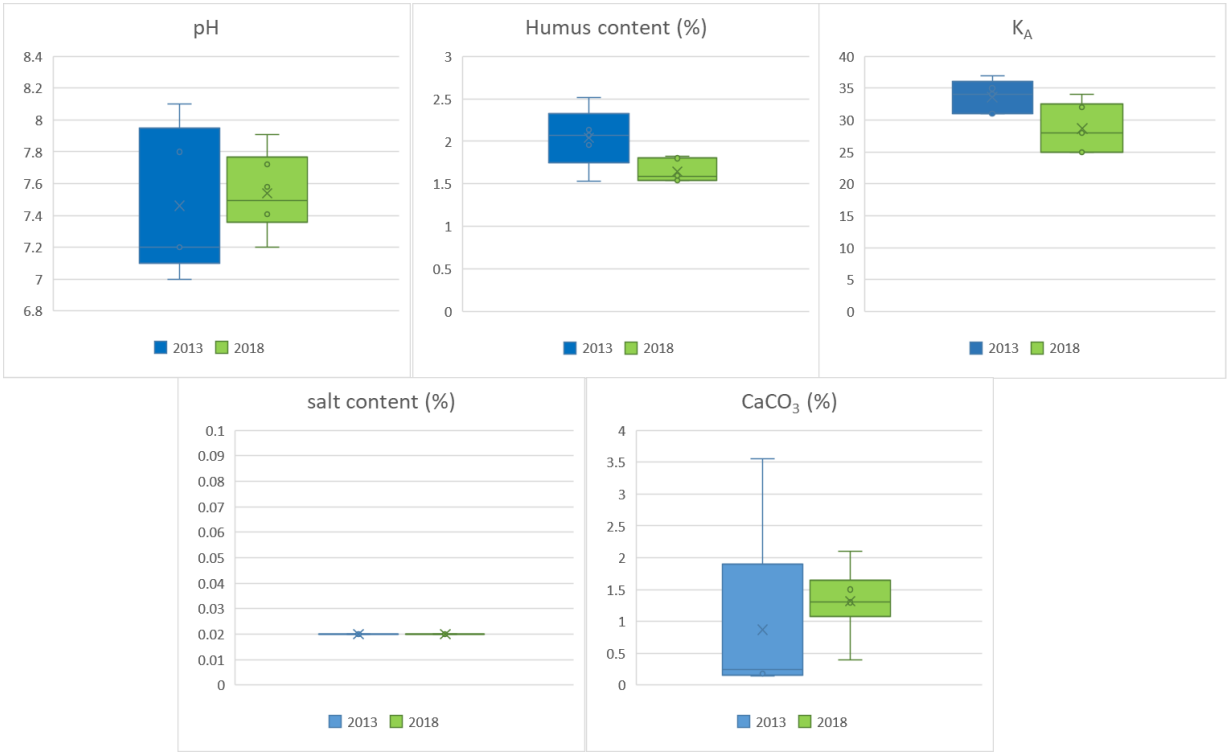

Fig. 2 Basic soil parameters between 2013 and 2018 (before the compost treatments and afterwards) in the topsoils of the study area near Újkígyós (Hungary)

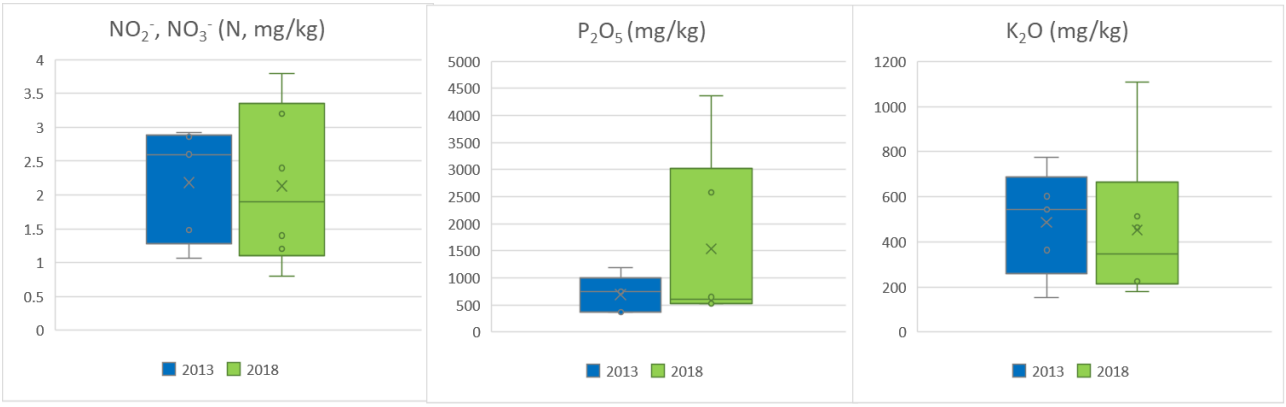

Fig. 3 Changes in the nitrogen, phosphorus and potassium contents between 2013 and 2018 of the topsoil as a result of the two compost treatments

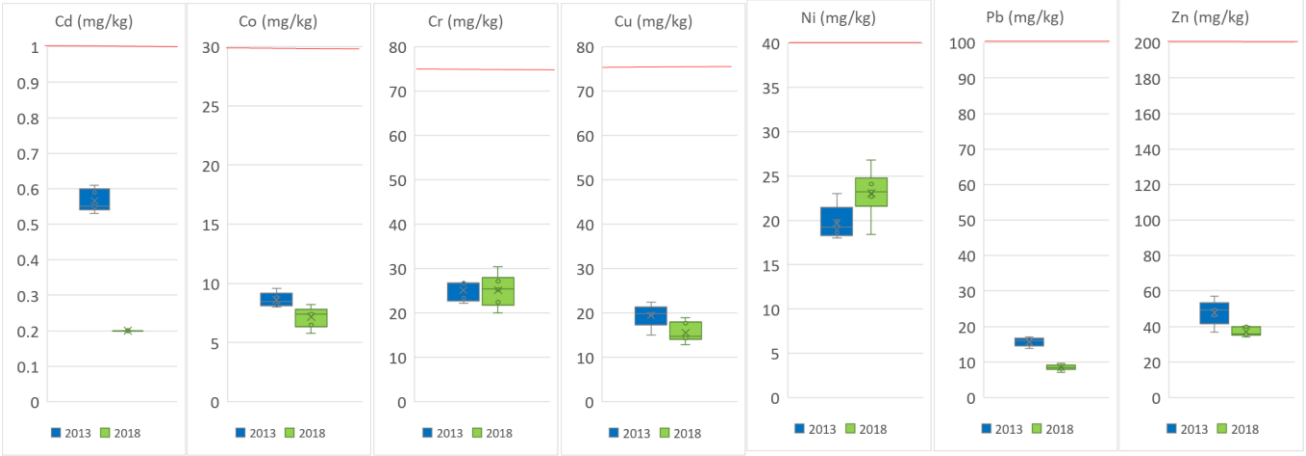

Fig. 4 Evolution of the soil-bound heavy metal contents (Ni, Co, Cr, Cd, Zn, Pb, Cu and As) between 2013 and 2018 at the study site (the red linestands for the contamination thresholds according to the Hungarian standards for soil quality (6/2009. [IV. 14.] KvVM-EüM-FVM)

The beneficial effects of the compost treatment can also be clearly identified if we compare the N-, P- and K-nutrient content of the control and the treated sampling areas in 2018 (Fig. 5). If we separate affected sites according to the timing of the disposal, it can be identified that 1,2 sampling sites (under disposal in 2018) show higher phosphorus content compared to 3,4, where there was no disposal in 2018 (only before). There is no significant difference between 1,2 and 3,4 in case of the nitrate and potassium. Among the heavy metals, copper and zinc show slightly higher concentrations in the soil samples affected by the compost treatment compared to the control soil, which can be explained by their higher concentration in the applied sewage sludge compost as compared to their level in the untreated soils.

CONCLUSIONS

The present study investigated the effects of sewage sludge compost applications on Chernozem soils in Hungary. Our study confirmed that there were no significant changes in either the basic soil parameters or the heavy metal load in the topsoil (0-30 cm) over the studied 5-year-period. In contrast, sludge compost treatment added to the soil an excess of slowly decomposing organic matter, rich in macronutrients (eg. nitrogen, phosphorus). Thus, properly designed sludge spreading on agricultural land in compliance with legal requirements can improve the nutrient and organic matter content of soils in a sustainable way. The accumulation of heavy metals with potential environmental and human health risks could not be identified. According to our results in 2018, the samples affected by sewage sludge compost applications showed an increased nitrogen and phosphorus content compared to the control areas, while the heavy metals zinc and copper showed somewhat higher values in the samples affected by the compost treatments. Our results are in line with previous studies who also did not identify a significant increase in heavy metal concentrations as a result of low doses of sewage sludge disposals (Pinamonti et al., 1997; Sánchez-Monedero et al., 2004). The resulting data also provide evidence that in case the applied sewage sludge composts meet the quality guidelines and requirements for the treatment doses set in the Government Decree 40/2008. (II. 26.), their field application as fertilisers can be considered as a sustainable management practice of secondary raw materials. The study confirmed the safe applicability of low dose municipal sewage sludge composts in a Hungarian study area. Further study sites are under investigation to investigate the impact of regular low-dose municipal sewage disposal within the broader study area.

Acknowledgements

The support of European Union and Hungarian State (grant agreement no. EFOP-3.6.2-162017-00010) is gratefully appreciated.

References

Babcsányi, I., Ladányi, Zs., Perei, K., Bodor, A., Barta, K., Kézér, A., Csányi, K., Pálffy, B., Farsang, A. 2009. The impact of sewage sludge disposal on the bacterial activity, nutrient and heavy metal content of chernozem soils and on the plant productivity, SE Hungary. In: Rákhely, G., Hodúr, C., Lemmer, B., Jákói Z. (ed) II. Sustainable Raw Materials International Project Week And Scientific Conference: Book of Abstracts Szeged, Hungary: University of Szeged, 33–33.

Banerjee, M.R., Burton, D.L., Depoe, S. 1997. Impact of sewage sludge application on soil biological characteristics. *Agriculture, Ecosystems & Environment* 66, 241–249. DOI: 10.1016/S0167-8809(97)00129-1

Csányi, K., Pálffy, B., Ladányi, Zs., Barta, K., Farsang, A. 2018. Tartós szennyvíziszap kihelyezés hatása a talaj szénforgalmára. In: Fazekas, I., Kiss, E., Lázár, I. (ed.) Földrajzi tanulmányok 2018 Debrecen, Magyarország: MTA DAB Földtudományi Szakbizottság, 187–189.

Diacono, M., Montemurro, F. 2011. Long-Term Effects of Organic Amendments on Soil Fertility. In: Lichtfouse, E., Hamelin, M., Navarrete, M., Debaeke, P. (eds.) Sustainable Agriculture Volume 2. Springer Netherlands: Dordrecht.

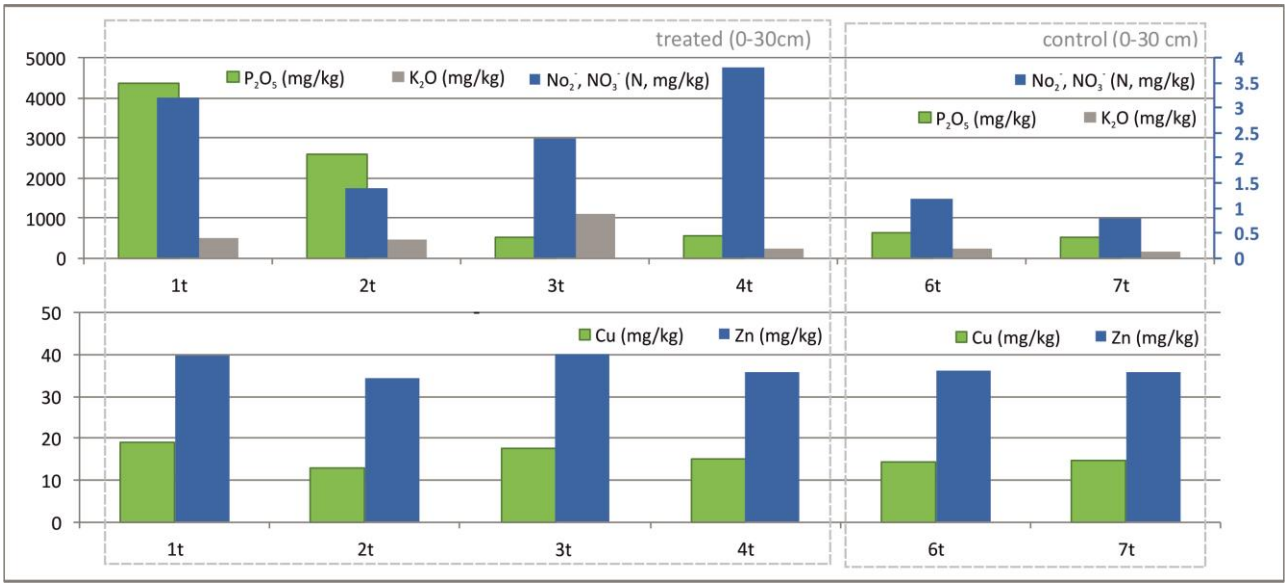

Fig. 5 N-, P- and K-nutrient and heavy metal contents are compared between the compost treated and the control sites in soil samples from the survey in 2018

Fang, W., Wei,Y., Liu, J. 2016. Comparative characterization of sewage sludge compost and soil: Heavy metal leaching characteristics. *Journal of Hazardous Materials* 310, 1–10. DOI: 10.1016/j.jhazmat.2016.02.025

Hudcová, H, Vymazal, J., Rozkošný, M. 2019. Present restrictions of sewage sludge application in agriculture within the European Union. *Soil and Water Research* 14/2, 104–120. DOI: 10.17221/36/2018-SWR

Kádár, I. 2013. Szennyvizek, iszapok, komposztok, szervestrágyák a talajtermékenység szolgálatában. MTA TAKI, Budapest, 347 p.

Kézér, A. 2018. Szennyvízisap-kihelyezés hatása a talaj nehézfémtartalmára és baktérium összetételére. Diplomamunka. Szegedi Tudományegyetem, Természeti Földrajzi és Geoinformatikai Tanszék, 64 p.

Khadhar, S., Sdiri, A., Chekirben, A. Azouzi, R., Charef, A 2020. Integration of sequential extraction, chemical analysis and statistical tools for the availability risk assessment of heavy metals in sludge amended soils. *Environmental Pollution* 263, 114543. DOI: 10.1016/j.envpol.2020.114543

Kocsis, I. (2011). Hígtrágya és szennyvíziszap kezelés. Szent István Egyetem, 93 p.

Moreno, J., Carlos García, L., Hernández, T., Ayuso, M. 1997. Application of composted sewage sludges contaminated with heavy metals to an agricultural soil. *Soil Science and Plant Nutrition* 43, 565–573. DOI: 10.1080/00380768.1997.10414783

National Sewage Sludge Management Strategy of Hungary 2014-2023. Szennyvíziszap kezelési és hasznosítási stratégia (2014-2023). Available at: http://biopsol.hu/files/file/Szennyviziszap_kezelesi_es_hasznositasi_strategia_2018_2023.pdf

Pálffy,B., Farsang, A., Kézér, A., Barta, K., Csányi, K., Ladányi, Zs. 2018. Tartós szennyvíziszap kihelyezés hatása a talaj tápanyag-, toxikuselem-összetételére, valamint a baktériumaktivitásra. Talajtani Vándorgyűlés Program és Absztrakt füzet, 2018. 08. 29 – 09.01. Pécs, 96–97.

Pinamonti, F., Stringari, G., Gasperi, F., Zorzi, G. 1997. The use of compost: its effects on heavy metal levels in soil and plants. *Resources, Conservation and Recycling* 21, 129–143. DOI: 10.1016/S0921-3449(97)00032-3

Sánchez-Monedero, M. A., Mondini, C., de Nobili, M., Leita, L., Roig, A. 2004. Land application of biosolids. Soil response to different stabilization degree of the treated organic matter. Waste Management 24, 325–332.

Singh R.P., Agrawal, M. 2008. Potential Benefits and Risks of Land Application of Sewage Sludge. Waste Manag. 28, 347–358. DOI: 10.1016/j.wasman.2006.12.010

Sleutel, S., De Neve, S., Németh, T., Tóth, T., Hofman, G. 2006. Effect of manure and fertilizer application on the distribution of organic carbon in different soil fractions in long-term field experiments. European Journal of Agronomy 25, 280–288. DOI: 10.1016/j.eja.2006.06.005

TA UWWTD 2017. 9 th Technical assessment on UWWTD implementation https://ec.europa.eu/environment/water/water-urbanwaste/implementation/pdf/9th%20Technical%20assessment%20of%20information%20on%20the%20implementation%20of%20Council%20Directive%2091-271-EEC.pdf

Uri, Zs., Lukácsné Veres, E., Kátai, J., Simon, L. 2005. Különböző módon előkezelt szennyvíziszapok hatása a talaj mikroorganizmusaira és enzimaktivitására. *Agrokémia és Talajtan* 54/3-4, 439–450.

Vermes, L. 1997. Vízgazdálkodás. Mezőgazdasági Szaktudás Kiadó, Budapest, 462 p.

40/2008. (II. 26.) Korm. rendelet a szennyvizek és szennyvíziszapok mezőgazdasági felhasználásának és kezelésének szabályairól

.

sciendo

Journal of Environmental Geography 13 (1–2), 31–42.

DOI: 10.2478/jengeo-2020-0004
ISSN 2060-467X

INVESTIGATING THE DRIVERS OF TOTAL SUSPENDED SEDIMENT REGIME IN THE SENEGAL RIVER BASIN USING LANDSAT 8 SATELLITE IMAGES

Cheikh Faye[1]*, Manuela Grippa[2], Laurent Kergoat[2], Elodie Robert[3]

[1]Department of Geography, U.F.R. Sciences and Technologies, UASZ, Laboratoire de Géomatique et d'Environnement, BP 523 Ziguinchor, Senegal
[2]Geosciences Environnement Toulouse, University of Toulouse III, Route de Narbonne, 31330 Toulouse, France
[3]CNRS, University of Nantes, UMR LETG, Campus du Tertre BP 81227, 44312 Nantes Cedex 3, France
*Corresponding author, e-mail: cheikh.faye@univ-zig.sn

Research article, received 25 February 2020, accepted 27 April 2020

Abstract

Because total suspended sediments (TSS) influence the penetration of light into the water column and are likely to carry pollutants and nutrients, their study is essential for understanding the functioning of African river ecosystems. If the estimation of solid transport is important in the context of hydro-agricultural developments, its quantification often poses a problem. In addition, in situ data for these areas are rare and, as a result, the environmental factors responsible for the variability of TSS can hardly be understood. This work aims to evaluate the spatiotemporal variability of TSS in the surface waters of the Senegal River using satellite data over the 2014-2018 period. The spatio-temporal dynamics of TSS is reconstructed using a relationship established on several West African sites between in situ data from TSS and satellite reflectances from Landsat 8. These data allow analyzing the relationship between TSS and factors such as rainfall and discharge. We found that the TSS peaks in Bakel coincide with the arrival of the first rains and are followed by peaks in discharge with a lag of 2 months. A time lag between TSS and discharge peaks is also observed on its tributaries like the River Falémé. Concerning the spatial variability, TSS generally increase from the river upstream to the downstream and decrease in the Senegal delta after the Diama dam. The analysis of the TSS upstream and downstream of the Manantali dam, in the upstream area, confirms the relatively low sediment deposits in the dam lake.

Keywords: water color, suspended particulate matter, Landsat image, remote sensing, Senegal River

INTRODUCTION

In the context of climate change, a recent trend in increasing daily precipitation extremes has been detected in tropical Africa (Panthou et al., 2014; Taylor et al., 2017). Increased runoff and erosion, as well as increased precipitation, suggest a possible increase in turbidity and suspended sediments (SS) in rivers. Beyond their importance to hydrology, sediments also have an impact on inland water ecology and potentially on human health. Indeed, TSS by forming a refuge and a culture medium conducive to the growth of bacteria generally promote their development while reducing their mortality by protecting them from ultraviolet rays (Palmatee et al., 1993). Thus, in environments with high turbidity, fecal bacteria have high attachment rates to particles (Rochelle-Newall et al., 2015). In addition, Troussellier et al. (2004) noted good bacterial survival after a long residence time in the water (4-5 days) of the Senegal River. Some of these bacteria or viruses cause widespread diseases, such as diarrhoea, which is one of the leading causes of death among children under five in developing countries (Troeger et al., 2017).

Due to the scarcity of in situ monitoring in the Sahel, the development of methods for monitoring turbidity and suspended solids in the Senegal River catchment area is therefore essential. Indeed, the distribution and dynamics of TSS in the Senegal River catchment area (TSS level, seasonal cycles and interannual variability) are not always well known. However, the study of water colour by remote sensing can be used to monitor water turbidity and TSS. Indeed, suspended particles absorb and scatter light affecting the spectral response of surface waters, particularly in the visible and near infrared (NIR), which varies according to concentration, mineralogical composition, organic content and particle size distribution (Reynolds et al. 2010). Although the use of remote sensing in Sub-Saharan Africa is confronted with specificities such as high aerosol content, very high turbidity and TSS values, recent studies have shown that satellite data are effective in monitoring the variability of turbidity and TSS in continental waters in these regions. Kaba et al. (2014) studied the temporal dynamics of an Ethiopian lake using Landsat images. Robert et al. (2016, 2017) used MODIS and Landsat data to document TSS and turbidity in the Bagré reservoir in Burkina Faso (Robert et al., 2016), and in the Gourma region of Mali (Robert et al., 2017).

Many sensors have been developed for various water color applications, including water turbidity assessments, sensors such as AVHRR, SeaWiFS, MODIS, IKONOS, Landsat TM and ETM +. Variations in temporal and spectral resolutions, data availability, calibration issues and temporal coverage are some of the most important factors determining the selection of the best instrument for a particular study (Li and Li, 2004). Although the

application of water color algorithms to estimate water constituents is generally limited by site-specific factors (Hellweger et al., 2004), and the absence of a uniform model of remote sensing to estimate the TSS (Wang et al., 2004), satellite images can be effectively employed to recover water quality parameters. In this context, the objectives of this article are as follows: (i) analyse the seasonal cycle and interannual variability of TSS in the Senegal River basin using Landsat images; (ii) study the links between TSS variability and precipitation and discharges; (iii) explore the variability of TSS in the Senegal River basin, including the Manantali Dam and major tributaries.

STUDY AREA

The Senegal River, some 1700 km long, drains a 300,000 km^2 basin straddling four countries: Guinea, Mali, Senegal and Mauritania (Fig. 1). It ranges from 10°20' to 17° N and 7° to 12°20' W and is made up of several tributaries, the main ones being the Bafing, Bakoye and Falémé rivers that originate in Guinea and form the upper basin (OMVS, 2008) (Fig. 1). The Senegal River, formed by the junction between the Bafing and Bakoye rivers, receives the Kolimbiné and Karokoro rivers on the right and Falémé on the left, 50 km upstream from Bakel. In the southern part of the basin, the density of the surface drainage system reflects the impermeable nature of the land (Rochette, 1974).

The basin is generally divided into three entities: the upper basin, the valley and the delta, which differ greatly in their topographical and climatological conditions. The upper basin extends from the sources of the river (Fouta Djalon) to the confluence of the Senegal River and the Falémé River (downstream of Kayes and upstream of Bakel). It is mainly composed of the Guinean and Malian parts of the river basin and provides almost all of the water inputs (more than 80% of the inputs) of the river to Bakel (OMVS, 2008). The analysis of the hydrological system of the Senegal River basin reveals four hierarchical levels: (1) the natural (uncontrolled) reaches of the Bafing upstream of the dam (Falémé and Bakoye), (2) the reach of the Bafing downstream of the dam (controlled), (3) the reach after the confluence of the Bafing and the Bakoye (at Kayes), (4) the reach after the confluence of the Falémé and the semi-artificialized Bafing-Bakoye subsystem (at Bakel).

From a hydrological point of view, at Bakel, which is often considered the reference station of the Senegal River, the average annual discharge of the river is about 676 m^3/s, corresponding to an annual input of about 21 billion cubic metres. Mean monthly discharges range from extreme values of 3,320 m^3 /s in September to 9 m^3/s in May (Ndiaye, 2003). For the distribution of hydrological inputs from the main rivers in the basin, 50% come from the Bafing (which is the main tributary), 25% from the Falémé, 20%

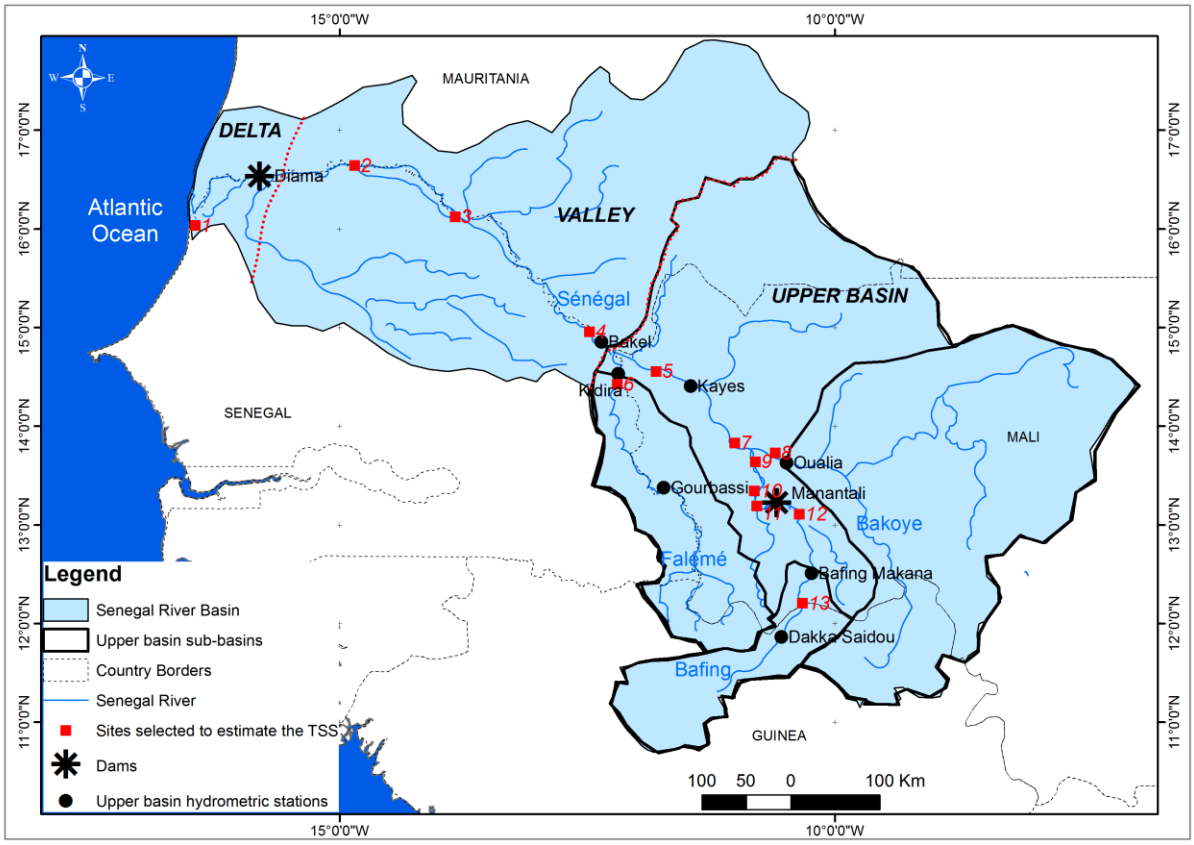

Fig. 1 Location of the Senegal River Basin and sites for estimating TSS. 1: Senegal River site in Saint Louis ; 2: Senegal River site in Podor ; 3: Senegal River site in Matam ; 4: Senegal river site after Faleme ; 5:Faleme River site in Kidira ; 6: Senegal river site before Faleme ; 7: Senegal river site after Bakoye ; 8: Bakoye River site in Oualia ; 9: Senegal river site before Bakoye ; 10: Bafing site after Manantali ; 11: Site downstream of the Manantali Dam : 12: Site upstream of the Manantali Dam ; 13: Bafing site before Manantali

from the Bakoye and the remaining 5% transit from the other tributaries of the Senegal River (OMVS, 2012).

Various studies in this basin have highlighted climate change and changes in its hydrological regime since the 1970s (Faye et al., 2015; Wilcox et al., 2018). Indeed, faced with a very strong drop in precipitation, which installed the Senegal River basin in a succession of deficit years. the average discharge module at Bakel thus decreased by more than half and went from 1374 m^3/s over the period 1903-1950 to 597 m^3/s over the period 1951-2002; and from an average of 840 m^3 / s over the period 1950-1972 to only 419 m^3/s over the period 1973 (OMVS, 2012). The width of the Senegal river is variable, it is on average 450 m, a width which decreases to reach 400 m only between Guinea and Mali. At the level of the Bakoye, the average width is 250 m. At the level of the valley, the river surrounds a large arm 600 m wide, 12 m deep and a small arm 200 m wide at Saint-Louis.

Faced with the effects of climate change and changes in the hydrological regime, a series of developments (Diama and Manantali dams) were initiated, totally transforming the hydrological dynamics of the Senegal River basin. The Diama dam, located on the Senegal river 27 km upstream of Saint-Louis, was put in water in 1986. Besides its function of blocking the ascent of salt water, it ensures the availability of fresh water while throughout the year for all uses, the development of an irrigable potential of 120,000 ha, the improvement of the filling of lakes (Guiers in Senegal, Rkiz in Mauritania). The Manantali dam, built between 1982 and 1988, is located on the Bafing river, the main tributary of the Senegal river, 90 km upstream of Bafoulabé. At the filling level of 208 meters IGN, its reservoir has a capacity of 11.3 billion m^3 and covers an area of 477 km^2. The Manantali dam regulates the flow of the Senegal river and irrigates a potential of 255,000 ha of land and, in the long term, should make it possible to navigate the river for approximately 800 km from its mouth (OMVS, 2012; Faye, 2018). The Manantali dam was built to guarantee a maximum flow of less than 4,500 m^3/s at Bakel (Thiam, 2016), causing sediment retention at the level of sn lake.

DATA AND METHODS

Satellite data

For this study, we used the surface reflectances of the Landsat 8 OLI satellite available since 2013. These data, corrected by atmospheric effects, are provided by the United States Geological Survey. They were analyzed from Google Earth Engine (and partially downloaded from Earth Explorer). Landsat 8 has several spectral bands in the visible, near infrared and mid-infrared, with a spatial resolution of 30 m and acquires the data with a temporal repetitiveness of 16 days. For this work, we used the blue, green and red bands (2, 3 and 4 at 482 nm, 561.5 nm and 654.5 nm respectively), the near infrared band (5 to 865 nm) and the SWIR bands (6 and 7 at 1613.5 nm and 2202 nm). Data were extracted from 13 sites located along the Senegal River (Fig. 1).

Landsat 8 was chosen for this study because it allows monitoring small rivers such as the tributaries of the Senegal River (spatial resolution of 30 m) using consistent time series since 2014. MODIS satellite data were not retained because their spatial resolution (250 m) does not allow monitoring the dynamics of TSS in small rivers. Sentinel-2 data could provide complementary information to Landsat data, and work in ongoing to evaluate their potential to retrieve TSS over West Africa, and particularly the effect of sunlight, that given the smaller angle of the image acquisition could be more important than for Landsat. Moreover, the first Sentinel-2 images date back to December 2015 and over the study area, at the time of this study, full times series of level 2 products (atmospheric corrected images) were not yet provided.

Cloud images and images with high aerosol levels over the study sites were discarded based on the quality indices provided with the Landsat satellite data (clouds, adjacent clouds, cloud shadows, aerosols, cirrus clouds). TSS were estimated only for open water pixels. Pixels that were not already masked by clouds, aerosols or cloud shadows were classified as open water using a threshold on the MNDWI index (Robert et al. (2017). For this study area a threshold of 0.1 was applied.

Rainfall and discharge data

Rainfall and daily discharge data are also used in this study. The rainfall data come from the database of the National Civil Aviation and Meteorology Agency. These are daily rainfall data from the stations of Kédougou, Podor, Matam and Saint Louis from 2014 to 2018. Hydrological data are daily discharges from the stations of Bakel on the Senegal River and Kidira on the Falémé River from 2014 to 2018. These data come from the database of the Organization for the Development of the Senegal River.

Estimation of TSS

TSS measurements acquired at different sites in West Africa were used to evaluate different satellite indices proposed in the literature, which use visible and infrared bands. Only images acquired within three days of the in situ sampling dates were selected for this analysis. A total of 49 images were used for this study. For the comparison of in situ data and satellite reflectances, the open water pixel closest to the coordinates of the in situ sampling points was chosen to approach as closely as possible the in-situ sampling point.

The in situ water samples used for the estimation of the TSS were collected over the period December 2014 – September 2017 in the Bagré reservoir (Burkina Faso) (Robert et al., 2016), in Lake Agoufou (Gourma region in Mali) (Robert et al., 2017) and from mid-2015 to early 2016 in the Niger River in Niamey (Kennedy Bridge). These sites cover the different types of surface water (lakes, reservoirs and rivers) found in West Africa. At each site, measurements were performed every week during the rainy season and every 15 days during the dry season. For the Bagré site, a field mission was also carried out in July 2015 to study the spatial variability within the lake. In total, 52 in-situ measurements were used for this study: 21 for the Bagre Lake, 25 for the Agoufou Lake and 6 for the Niger River.

For all these sites, the TSS concentration was calculated by filtering the water samples using 0.7 µm glass fibre filters. Before filtration the filters were dried for 1 hour at 105° C and weighed. After filtration, the filters were dried again under the same conditions and weighed. The TSS concentration was calculated as the difference between the weight after and before filtering, divided by the filtered volume. For each sample, two replicas were used and the average TSS value was used for the analysis.

Correlations were evaluated using the R2 (coefficient of determination, equivalent to the Pearson's correlation), with the R software.

RESULTS

Relationship between satellite reflectance and TSS

Landsat 8 reflectance for different combination of visible and near-infrared bands proposed in the literature were plotted towards in-situ TSS and fitted using power relationships (Robert et al., 2016, 2017). Then, the TSS retrieved from Landsat 8 using the relationships obtained for each band combination were compared to in-situ TSS (Table 1).

Table 1 Value of R2, RMSE and bias obtained by linear regression of retrieved TSS towards in-situ TSS

Bands or ratio of bands	R^2	RMSE	Bias
NIR	0.91	389.2	-27.9
R	0.53	925.1	-71.6
NIR + R	0.86	492.3	-15.5
NIR/R	0.76	1562.8	-608.7
NIR/B	0.81	939.7	-262.3
(R-NIR)+(R+NIR)	0.69	2715.8	-1063.2

The best results are obtained for the NIR band (Table 1) with the following power relationship (Fig. 2):

$$Refl\_bande5 = 256.46 \times TSS \text{ ^ } (0.3539),$$

where *Refl_bande5* is the mean reference value of band 5 of the Landsat image. TSS are then estimated by the following equation:

$$TSS = Refl\_bande5 / 256.46 \text{ ^ } 1 / 0.3539$$

TSS retrieved from Landsat 8 are well correlated to in-situ TSS (R^2 = 0.91), with a very low bias (-27.9) and RMSE equal to 389.2 mg/L (Fig. 2b). TSS in the Senegal River basin were therefore estimated by reversing this power relationship, which is assumed to be generally applicable to surface waters up to values of about 2500 mg/l.

Variation of TSS in the upper basin as a function of rainfall and discharge data

Daily data

We compared the TSS, calculated from Landsat 8 satellite data, at the outlet of the upper Senegal River basin with rainfall data from the Kédougou station and hydrological data from the Bakel station (Fig. 3).

The seasonal rainfall pattern shows very heavy rainfall in the rainy season, which can reach more than 100 mm per day with a maximum recorded during the month of August. Total rains are close to zero during the dry season months (Fig. 3). The same is true for discharge in the basin charcaterised by two periods: the high water period centred on the months of July, August, September and October with a maximum in September; the low water period is centred on the rest of the year, with the lowest discharges of the year (Fig. 3) due to the absence of rain. Local precipitation is combined with water inflows from the river upstream of Bakel and its tributaries. The enrichment in suspended solids during the rainy season is due, particularly in the upstream part of the basin, to the strong erosion of the slopes which leads to leaching and particle loading into the river (Mbaye et al., 2016). Surface runoff in the region of

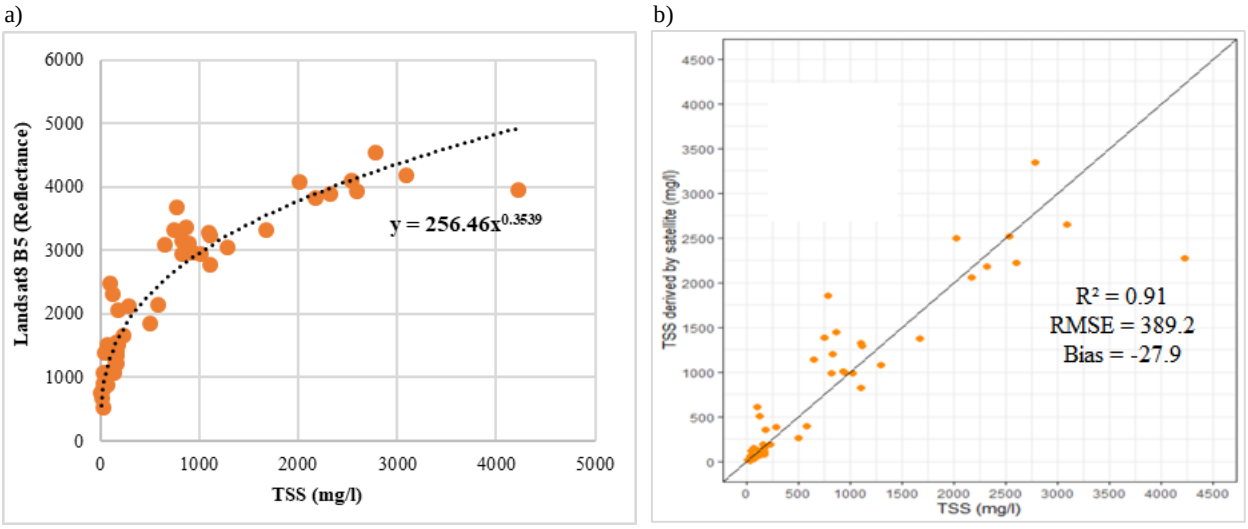

Fig. 2 Relations: a) Landsat reflectance in band 5 versus TSS measured in situ (in mg/l), b) TSS derived from Landsat 8 vs. In-situ TSS (N=52). R^2 indicate the Pearson's correlation. RMSE the Root Mean Square Difference, bias the mean difference. The 1:1 line is also indicated in black

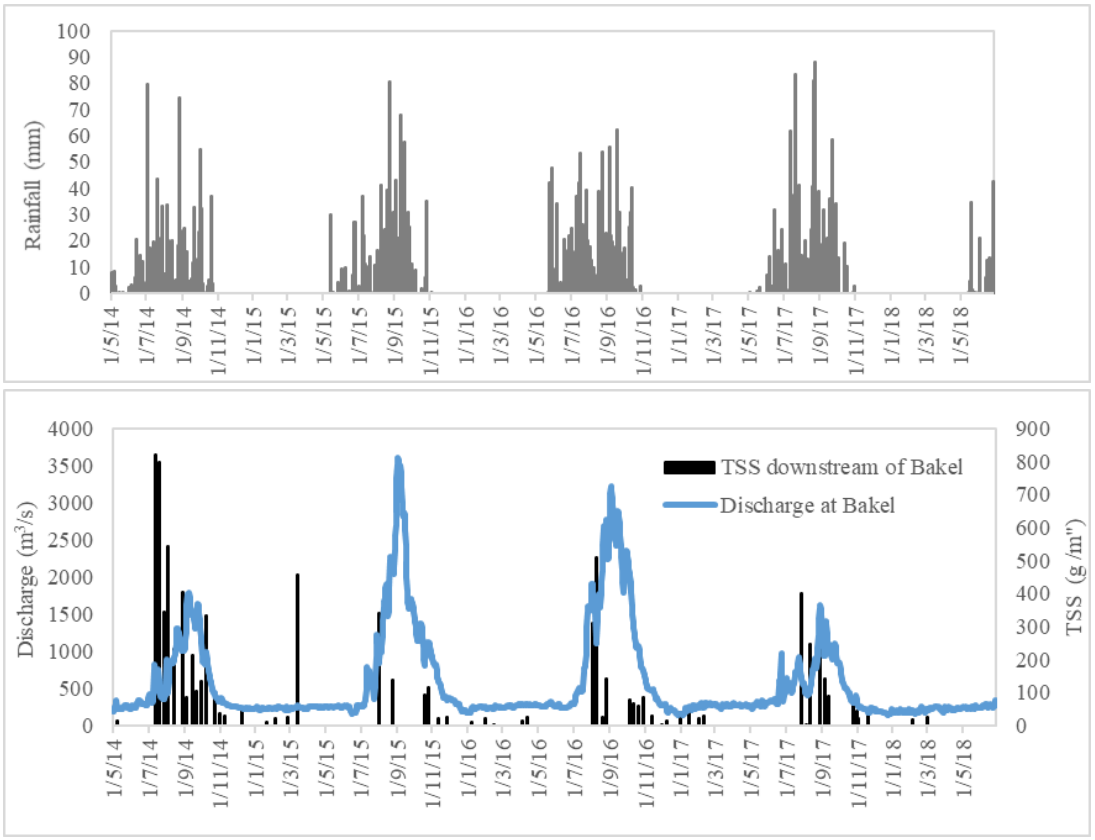

Fig. 3 Relationship between the daily rainfall (Kédougou station), daily concentration of TSS (downstream of Bakel) and the daily discharge at the outlet of the upper Senegal River basin (Bakel station) from 2014 to 2018

these soft sediments have a suspended matter concentration of ~1 g/l and is responsible for 50–80% of the riverine suspended load via slope erosion. In the river bed, this suspended load is diluted by subsurface and groundwater discharge (Kattan et al., 1987). During the dry season, the river becomes very "calm", while its turbidity and discharge remain low.

During the 2013-14 flood period, the main inputs recorded in Bakel came from unregulated tributaries (Falémé: 39.6% and Oualia: 26.8%), with Manantali providing only 10.5% of the 42.2% it received from Bafing upstream. On the other hand, during the low water period of the same year, the regularized tributary of the Bafing River provides almost all the discharge at Bakel, through the support of the low discharge of the dam. Similarly, discharge at Falémé and Bakoye rivers (2.69% and 1.15% respectively) contribute very little to the discharge at Bakel in Senegal River. This reverse operation allows the control of large discharge during the flood period (i.e. an average discharge released of 197 m^3/s) while during the low water period, the gates of the Manantali structure are more open for water releases (i.e. an average discharge released of 262 m^3/s).

Figure 3 shows that the TSS concentration generally increases with the arrival of the first rains and before the increase in the river discharge. However, the TSS experiences significant variations with the highest values between May and September, and which sometimes approach 1000 g / m^3 (820 g / m^3 on July 11, 2014). At this station, while most of the peaks occurred between July (beginning of the rainy season) and August (820 g/m^3 in 2014-15, 342 g/m^3 in 2014-15 and 400 g/m^3 in 2017-18 in July and 508 g/m^3 in 2016-17 in August), some peaks of suspended solids are recorded in September (this is the case on 13 September 2014 with 214 g/m^3), when river discharge is highest and at the time of annual precipitation maxima. A secondary peak, which is not explained by hydrological variations (rain or discharge), is observed in March 2015.

The concentration of TSS at the outlet of the Falémé basin was estimated between 2014 and 2018 and compared to the discharge at Kidira station (Fig.4). For the year 2014, the results remain fairly similar to those noted at the outlet of the upper Senegal River basin, with always the peaks of TSS preceding the peaks in discharge. For the other years, TSS data at the beginning of the rainy season, when cloud cover prevents visibility and therefore detection with optical images, are not in sufficient number to establish a relationship.

Monthly averages

On a monthly basis, the average TSS concentration values over the period 2014-2018 indicate a unimodal evolution (in the form of a bell) with a maximum in July with 590 g/m^3 and a minimum in February with 21.1 g/m^3 in Bakel (Fig. 5). This behaviour alternates between July and August over the different years of the series with a maximum noted in July with 554 g/m^3 in 2014-15, 343 g/m^3 in 2015-16 399 g/m^3 in 2017-18 and in August with 226 g/m^3 in 2016-17. For the monthly minimum of TSS, the value can be very low, falling to 0.09 g/m^3 in January over the 2014-2018 period. Over the period 2014-2018, the Kidira station on the Falémé River shows a maximum in August with 351 g/m^3 and a minimum in December (6.42 g/m^3) (Fig. 5 and Table 2).

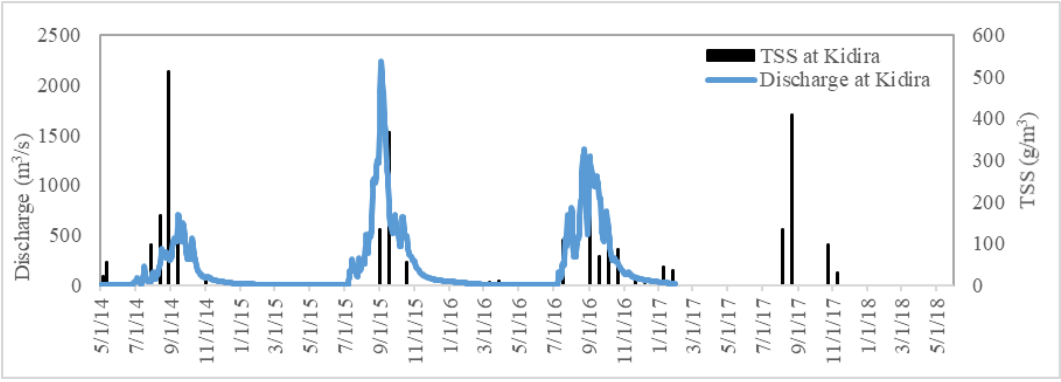

Fig. 4 Relationship between the daily concentration of suspended solids at the outlet of the Falémé basin (upstream of Bakel) and the discharge (Kidira station) from 2014 to 2018

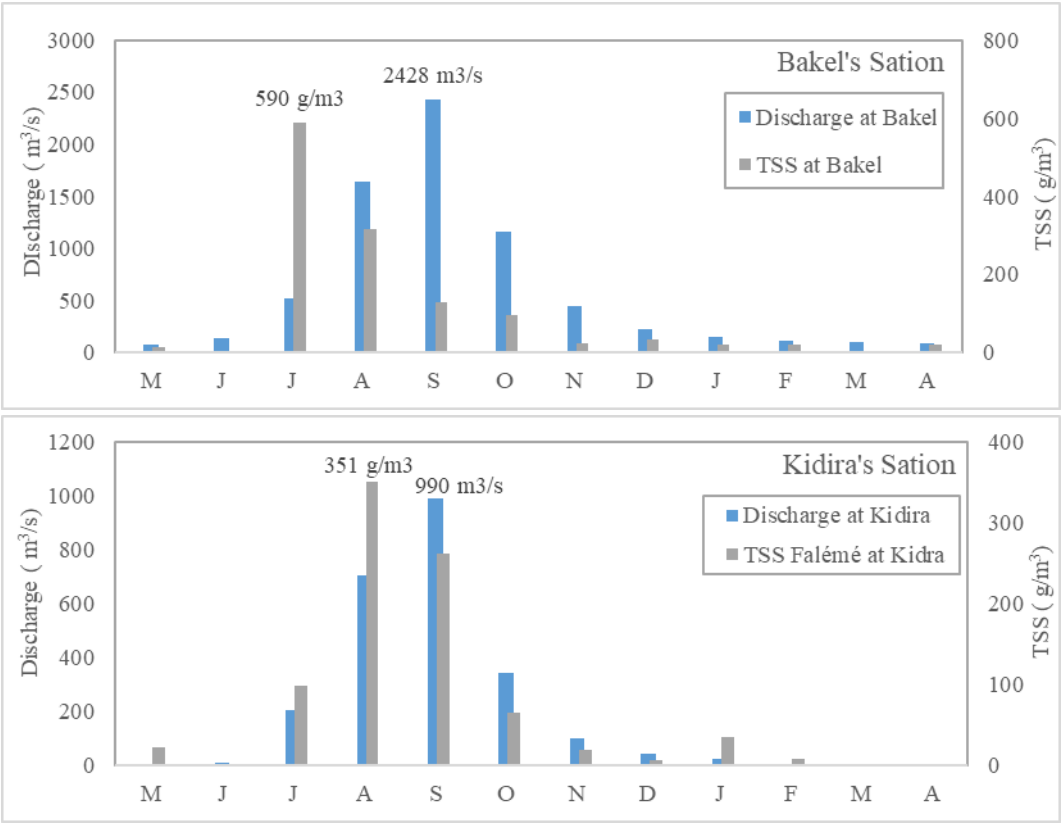

Fig. 5 Relationship between the monthly concentration of suspended solids and the monthly discharge at the outlet of the upper Senegal River basin and at the confluence with the Falémé from 2014 to 2018

The comparison between the discharge and the TSS concentration clearly shows the difference between the two peaks at Bakel and Kidira (Fig. 5). In Bakel, while the peak of TSS is recorded in July with 590 g/m^3 (when the discharge was only 524 m^3/s), the peak of discharge is recorded two months later, in September with a value of 2428 m^3/s (when the TSS was only 128 g/m^3). For the Kidira station on the Falémé, this same lag persists with a peak discharge noted in September (with a discharge of 990 m^3/s corresponding to a discharge of 263 g/m^3 for TSS), one month after that of the TSS with a value of 351 g/m^3 in August (when the discharge was 707 m^3/s).

Evolution of TSS in the valley and delta of the river

Beyond the upper basin, which covers 56% of the basin's surface area and extends from the source to the Bakel hydrological station over a distance of 980 km, the TSS are assessed in the valley, which covers 35% of the basin's surface area and extends over 600 km, and the delta, which extends from Richard-Toll to the mouth of the river over 170 km (Thiam, 2016). To characterize the TSS in the Senegal River valley and delta, the stations of Matam and Podor (for the valley) and Saint Louis (delta) were selected.

Figures 6 and 7 show that the concentration of TSS generally begins to increase with the arrival of the first rains, before the river's discharge increases, and TSS peaks mainly in late July (727 g/m^3 in Matam and 531 g/m^3 in Podor) and August (344 g/m^3 in Saint Louis), often slightly later (Fig. 7) than for some sites previously studied in the upper basin (Bakel for example). At the same time, in the valley and delta of the basin, rain peaks are recorded, as are TSS, in August (137 mm in Matam, 92.1 mm in Podor and

Table 2 Monthly average values of suspended matter concentration in the Senegal River and discharge (at the outlet of the upper Senegal River basin) from 2014 to 2018

TSS and discharges	**M**	**J**	**J**	**A**	**S**	**O**	**N**	**D**	**J**	**F**	**M**	**A**	**Yrs**
Discharge at Bakel	76	143	524	1643	2428	1166	441	226	150	114	97.9	83.1	591
TSS at Bakel	14.1	-	590	316	128	97.0	24.2	35.1	21.6	21.1	-	20.3	127
Discharge at Kidira	0.60	10.2	208	707	990	346	98.8	45.0	23.1	4.04	1.80	0.58	203
TSS Falémé at Kidra	21.8	-	98.3	351	263	65.7	20.1	6.42	34.7	7.54	-	-	96.5
TSS per section	**M**	**J**	**J**	**A**	**S**	**O**	**N**	**D**	**J**	**F**	**M**	**A**	**Yrs**
TSS Senegal in Saint Louis	21.8	-	116	344	181	67.0	27.4	13.3	9.4	9.2	14.7	10.4	74.0
TSS Senegal in Podor	47.3	-	531	457	346	167	95.0	83.5	26.8	14.7	40.1	71.1	171
TSS Senegal in Matam	40.8	-	727	290	132	176	24.8	22.0	27.3	26.6	28.5	22.9	138
TSS Bakoye in Oualia	-	203	108	177	262	135	59.0	25.2	49.5	26.4	16.0	-	106
TSS Senegal after confluence with the Bakoye	4.8	-	29.6	120	122	29.3	12.9	6.2	8.6	3.4	2.0	5.4	31.3
TSS Senegal before confluence with the Bakoye	4.0	-	28.1	47.1	95.7	39.4	10.3	3.4	5.8	4.1	2.0	2.4	22.0
TSS Senegal after confluence with Falémé	-	-	143	298	168	108	34.0	10.7	6.57	18.0	-	-	98.3
TSS Senegal before confluence with Falémé	-	-	47.7	177	239	63.0	7.15	14.7	3.71	4.80	-	-	69.6
Senegal downstream Manantali	-	3.0	32.2	36.3	0.2	53.6	15.4	13.6	3.4	0.6	1.0	-	15.9
Senegal upstream Manantali	-	21.3	12.4	36.5	52.8	17.2	5.9	4.3	2.9	1.7	1.0	-	15.6

81.2 mm in Saint Louis), one month before those of discharge that are recorded in September (502 cm in Matam, 350 cm in Podor and 102 mm in Saint Louis).

Although the TSS show significant variations with the highest values between July and September and a gradual evolution from upstream to downstream, they are, on the other hand, more important in the river valley (the maximum being 811 g/m^3 on 29 July 2018 in Matam and 818 g/m^3 on 25 July 2014 in Podor) than in the delta (where the maximum is only 666 g/m^3 on 12 August 2018 in Saint Louis). The lowest TSS values are recorded during the low water period.

In the valley and the delta which is in the alluvial plain of the basin, it is the heights of water which are available and indicated in place of the discharge as in the other parts of the basin.

Indeed, the TSS values in Saint Louis are much lower than those observed in Podor (Fig. 7 low), despite a fairly similar rainfall dynamic (Fig. 7 high). This is in agreement with Kane (1997) who estimated a solid transport of the Senegal River, although highly variable from one year to another, relatively low during these hydrological cycles, with an average value of 2 million tonnes/year. This may be due to the Diama dam, whose commissioning in 1985 introduced a reduction in solid inputs, and modify the hydrological regime towards the lower estuary, also affecting the sediment balance that contributes to the maintenance of coastal areas and the coast south of the mouth of the Senegal River (Kane, 2005). Thus, the TSS that pass through from the Diama dam are essentially made of fine materials, indicating a river of low competence in its estuary, with clays and silts transported in suspension tending to settle upstream of the dam (Kane, 1997).

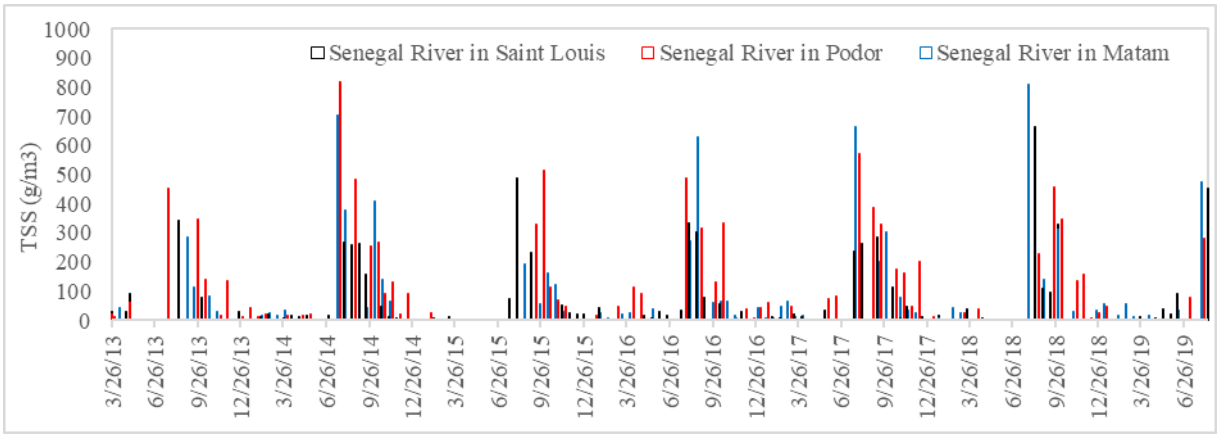

Fig. 6 Daily evolution of the concentration of suspended solids from the Senegal River in the valley (Matam and Podor) and the delta (Saint Louis)

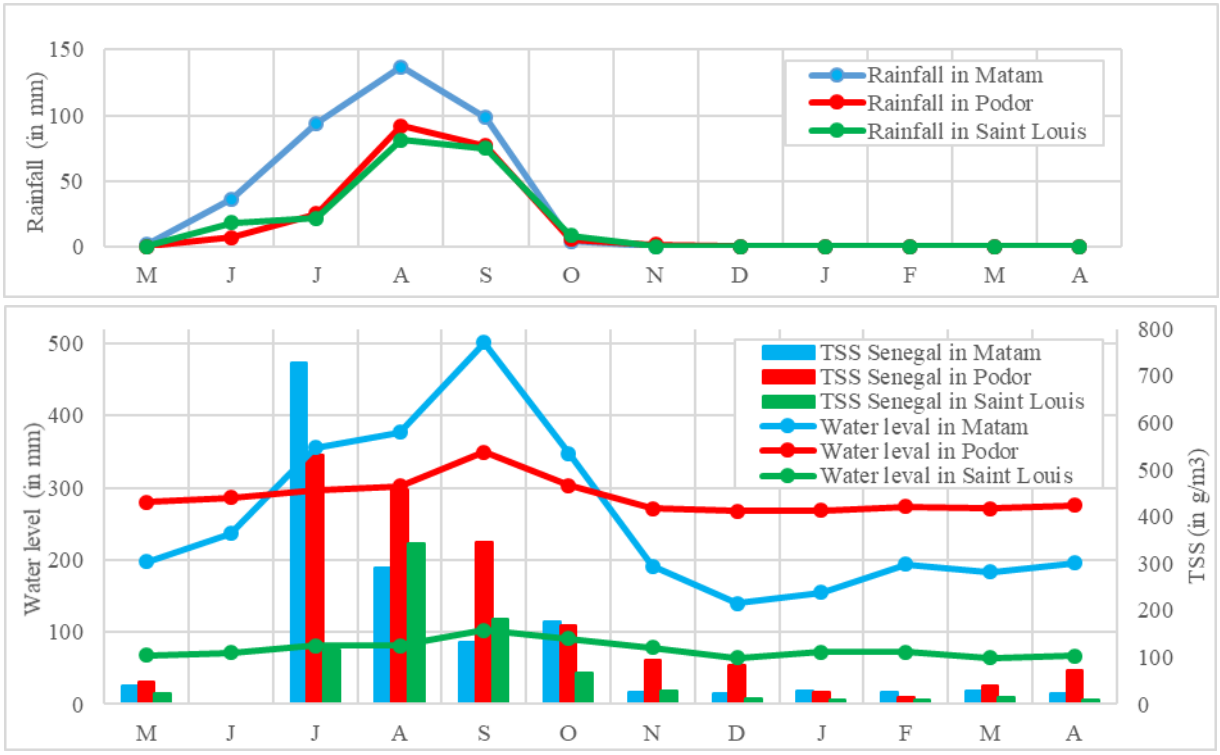

Fig. 7 Monthly evolution of TSS concentration, rainfall and water levels in the valley (Matam and Podor) and delta (Saint Louis) of the Senegal River

The seasonal variability of the concentration of suspended solids in the Senegal River valley and delta shows a peak at the beginning of the rainy season in Matam and Podor that precedes the peak precipitation by one month and the peak water level by two months (Fig. 7). This is in line with the TSS -rain-discharge relationships observed at the Bakel station. In Saint Louis, after the Diama dam, the peak of TSS is observed in August.

Impacts of tributaries on water and sediment supply in the Senegal River

To characterize the impacts of tributaries on the water and sediment supply of the Senegal River, we characterized the TSS of the Senegal River on its upstream and downstream section and each of its tributaries: first on the Senegal River upstream and downstream of the Manantali Dam; then on the Senegal River upstream and downstream of the confluence with the Falémé River; finally on the Senegal River upstream and downstream of the confluence with the Bakoye) (Fig. 8).

On the section of the Senegal River before the confluence with Falémé, the peak of TSS is noted later (in September with 239 g/m^3) compared to the part after the confluence with Falémé (whose peak is noted in August with 298 g/m^3) (Fig. 8). With the exception of September, the concentration of SS on the part after confluence with the Falémé is higher than that of the part before confluence of the Senegal River with the Falémé. On an annual scale, because the concentration of TSS is higher in the Falémé (with 128 g/m^3) than in the Senegal River upstream of Kidira (47.1 g/m^3), it is higher in the downstream part confluence between the Falémé and the Senegal River (with 119 g/m^3) than in the upstream part (Table 3).

In general, discharge at Bakel station are much higher than those measured at Kidira (Fig. 8). According to OMVS (2012**)**, the annual modules of the main rivers in the basin are as follows: 180 m^3 /s at Manantali on the Bafing; 149 m^3 /s at Oualia on the Bakoye; 134 m^3 /s at Gourbassi on the Falémé; 676 m^3 /s at Bakel on the Senegal River. A tributary such as the Falémé, whose discharge remains natural, has a slightly higher concentration of suspended solids than the Senegal River upstream of Kidira, a fact that could be explained by the role of the Manantali dam in retaining sediments in its lake and thus reducing sediments in the downstream part of the dam. Thus, although Falémé is more heavily loaded with suspended solids, its sediment supply to the Senegal River is mainly marked by its relatively smaller volume of water than that of the Senegal River (Falémé representing only 25% of the discharge flowing into Bakel.

By comparing the TSS of the different tributaries of the Senegal River (Bafing, Bakoye and Falémé), through a characterization of the TSS of the Senegal River on its upstream and downstream section and of each of the tributaries, we can generally observe an increase in the concentration of TSS after the confluence of the tributaries (Table 2 and 3). If we take the section of the confluence with the Bakoye basin, the average concentration of TSS over the period studied before the confluence, which is 18.1 g/m^3, increased slightly to 19.4 g/m^3 after the confluence, an increase of 1.3 g/m^3 (at a rate of 7.2%), indicating a slight excess of the TSS concentration in the Bakoye over the Senegal River. Similarly, for the section of the confluence with the Falémé basin, the average concentration of TSS over the period studied before the confluence, which is 69.6 g/m^3, increased to 98.3 g/m^3 after the confluence, an increase of

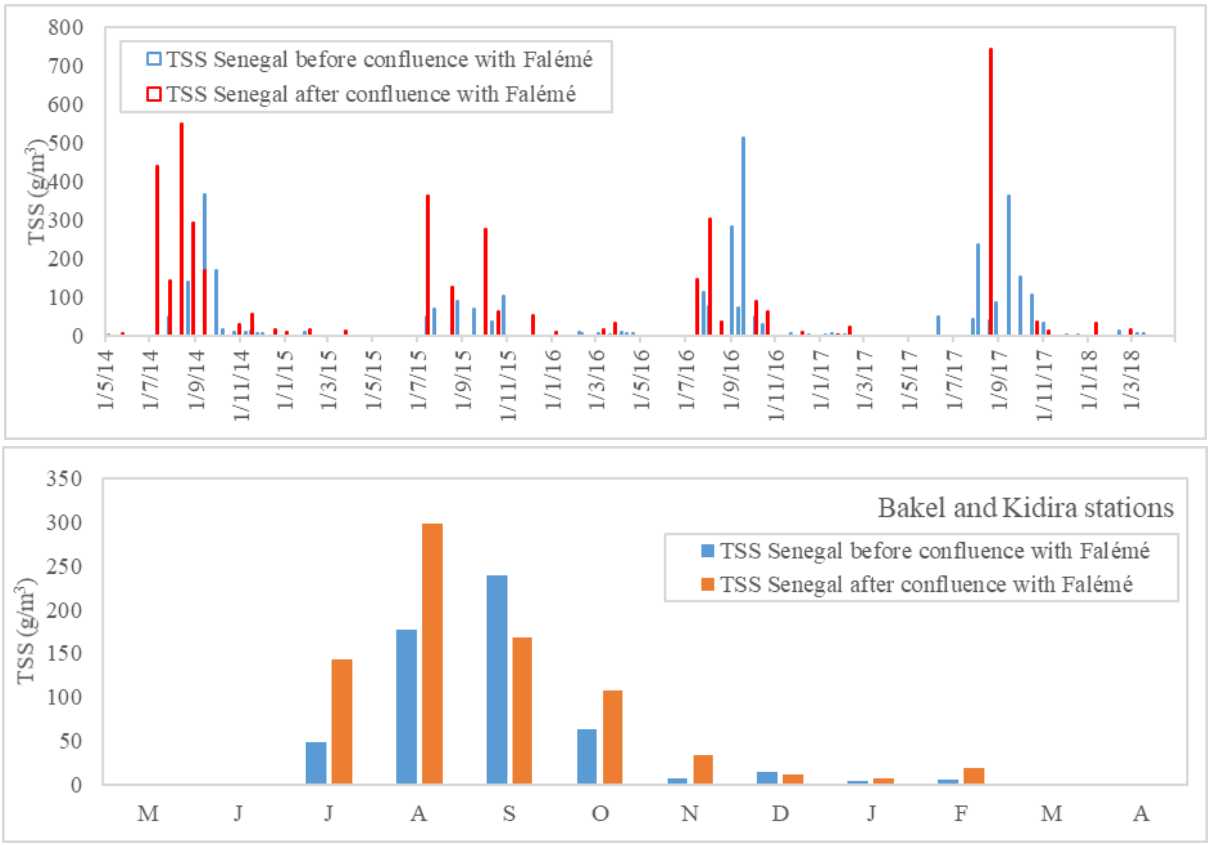

Fig. 8 Daily and monthly evolution of the Senegal River TSS concentration (upstream and downstream of the confluence with Falémé) from 2014 to 2018

Table 3 Evolution of the annual concentration of TSS (in g/m^3) of the Senegal River from upstream to downstream in the basin

Valley and delta	**2014-15**	**2015-16**	**2016-17**	**2017-18**	Average
Senegal in Saint Louis	82.3	65	88.2	104.2	84.9
Senegal in Podor	168.8	144.8	181.8	130.8	156.6
Senegal in Matam	130.7	99	123	142.6	123.8
Senegal in Bakel	237	74	98.1	115	131
Falémé in Kidira	169	98	78.6	167	128
Bakoye in Oualia	90.7	90.6	99.9	71.3	88.1
Senegal - Falémé	**2014-15**	**2015-16**	**2016-17**	**2017-18**	Average
Senegal downstream of Falémé	102	123	70.3	169	119
Senegal upstream of Falémé	68.2	49.0	66.5	57.7	47.1
Senegal - Bakoye	**2014-15**	**2015-16**	**2016-17**	**2017-18**	Average
Senegal downstream of Bakoye	23.0	28.3	10.5	16	19.4
Senegal upstream of Bakoye	33.0	13.9	16.3	9.3	18.1
Senegal - Manantali	**2014-15**	**2015-16**	**2016-17**	**2017-18**	Average
Bafing downstream of Manantali	19.9	48.3	60.6	35.3	41.0
Bafing downstream of Manantali	21.4	21.1	49.6	53.3	36.4
Section at the entrance of the Manantali dam	13.9	6.9	13.7	16.5	15.6
Section at the exit of Manantali	21.4	39.9	16.1	18.0	15.9

28.6 g/m^3 (with a rate of 41.1%), also indicating a very high sediment supply from the Falémé to the Senegal River. In the Bafing basin, the concentration on both sides of Manantali remains lower in the upstream part (36.4 g/m^3) than in the downstream part (41.0 g/m^3). In the valley, the transverse gradient, marked by an increase in TSS from upstream to downstream, remains strong (123.8 g/m^3 in Matam; 156.6 g/m^3 in Podor), while in the delta, with the presence of marine dynamics and the sediment retention role of the Diama dam (Kane, 2005), the concentration of TSS is declining (84.9 g/m^3 in Saint Louis).

Figure 9 shows the monthly evolution of the concentration of suspended solids along the Senegal River. It appears that, in general, the concentration of TSS increases, in relation to the increase in the volume of water flowing, from upstream (the source) to downstream (valley area), even if the seasonal distribution is not the same.

Impacts of the Manantali dam on sediment deposition in the Senegal River

The comparison of the upstream and downstream TSS concentration curves shows a more or less similar evolution over some periods, and different over others (Fig. 10). However, analysis of TSS values in the upstream part of Manantali indicates the relatively low sediment supply of Bafing in the lake of the Manantali dam (the annual average of the series being 13.3 g/m^3). This indicates a relatively limited accumulation of SS in the lake of the Manantali dam.

If we compare the average concentration of TSS in the upstream part (with 13.0 g/m^3) with those in the downstream part (with 20.1 g/m^3), the average concentration of TSS in the downstream part of the dam is sligthly higher than that in the upstream part of the dam (with a difference of 7.1 g/m^3). Only for 4 dates the average concentration of TSS in the upstream part is higher than in the downstream part indicating a potential sediment deposition in the dam lake.

Overall the potential sediment deposition is relatively low. However, some peaks can occur upstream and are accompanied by significant retention. This is why, the potential weak sediment deposits in the lake of the Manantali dam (noted over certain periods where the TSS upstream are higher than those downstream) do not allow us to affirm the absence of excessive sedimentation. This is why, the Manantali dam is for the moment spared from excessive sedimentation which constitutes one of the major problems associated with dams (this sedimentation can block the valves of a dam and cause its failure under certain conditions) (Zhang, 2014).

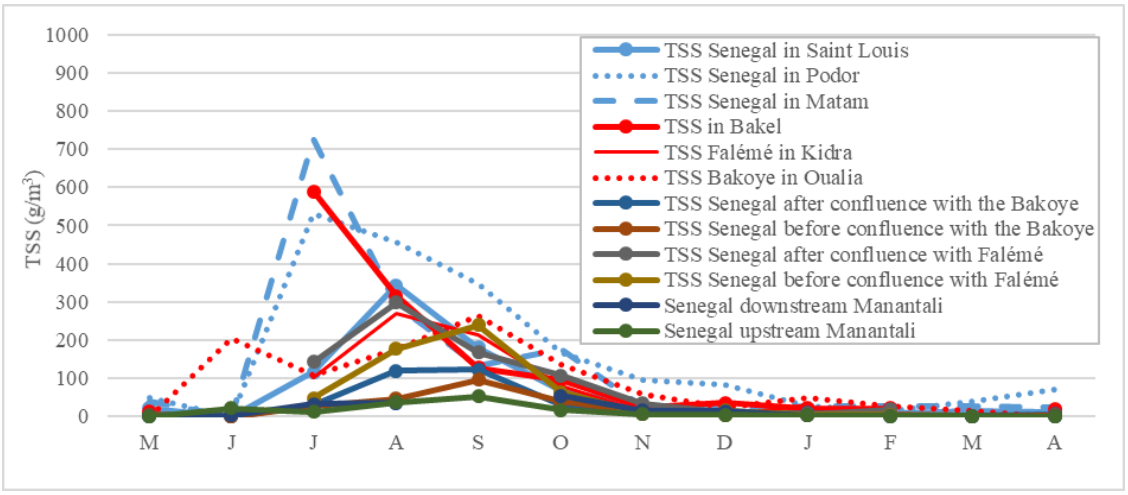

Fig. 9 Monthly evolution of the Senegal River TSS concentration from upstream to downstream in the basin

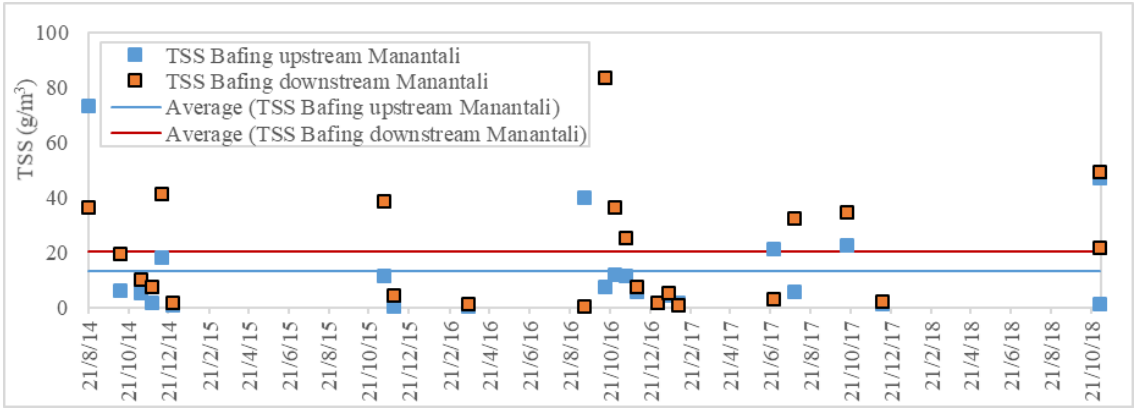

Fig. 10 Relationship between the concentration of TSS on the Bafing section (upstream and downstream of the Manantali dam) from 2014 to 2018

DISCUSSION

The use of the NIR band of Landsat 8 images has proven to be very appropriate for turbidity monitoring in a river characterized by high turbidity values and high spatial and temporal variability. These same Landsat satellite data were selected by Robert et al. (2017) to document TSS and turbidity in the Gouma region of Mali. Our study confirms the potential of sensors and the infrared band for monitoring the dynamics of TSS in West African surface waters.

The surface reflectances of Landsat 8 made it possible to analyse the spatial and temporal variability of turbidity in the Senegal River basin over the period 2014-2018. In terms of seasonal cycle, the increase in TSS concentration is mainly caused by rainfall between the early rainy season and the mid-rainy season. The concentration of TSS therefore generally increases with the arrival of the first rains before the river's discharge increases. However, this does not concern the whole basin (but only the upper basin), because there is sometimes a temporal difference between the upstream (upper basin) and downstream (valley and delta) zone of the basin. In semi-arid environments, runoff induced by the first rains could lead to increased transport because the soil surface has very low vegetation cover (Mbaye et al., 2016). Later in the season, with the development of vegetation in cropland and pasture, erosion and sediment transport are reduced, as well as the concentration of TSS in rivers and lakes or reservoirs (Robert et al., 2016).

At the stations in the upper part of the basin, most of the peaks occurred between July and September (high water period), often before the annual peaks in precipitation and river discharge (as in the lower part of the basin where peaks are noted in late July and early August). These results confirm those of Mbaye et al. (2016) which indicate a concentration of TSS measured at a station in July 2012 (Dembancane; between Bakel and Mattam) of 540 mg/l, well above the concentrations observed in August 2013. They are also in agreement with the TSS concentrations presented in Kattan et al. (1987), which found maximums during the first rainy season in June/July before the peak discharge from August to October. Our work also confirms the results of Kane (2005) who notes a rapid increase in TSS with the arrival of annual flood waters that carry sediment from watershed erosion.

At the interannual scale (over the period 2014-2018), a large variability in the concentration of TSS was observed in the basin on all sections. Although there are significant variations in TSS, they are higher in the river valley than in the delta and upper basin. These results are consistent with those reported by Mbaye et al. (2016) which indicate much higher TSS concentrations during the rainy season with values between 5.4 and 366.4 mg/l, with a general increase in TSS at downstream stations and a slight reduction at estuarine stations. They are also in agreement with those of Kane (2005) who, based on the TSS measurements made, noted a rapid increase in TSS with the arrival of annual flood waters that carry sediment from watershed erosion. His study showed that concentrations remain relatively high (almost 200 mg/l) until the end of September - beginning of October, after which time solid loads decrease very rapidly. According to his study, hydrological cycles show that on a monthly scale, the months of August, September and October account for almost all (95%) of the solid load. It is worth noting the drastic decrease in tonnage from November onwards (around 2% and even less) and the insignificant percentages in December (less than 1%).

According to Zhang (2014), excessive sedimentation is one of the major problems identified and associated with dams and can block its doors, causing the dam to rupture under certain conditions). Fortunately, in the Manantali dam, small peaks of TSS are noted, peaks which can hardly cause the silting up of the dam. Finally, satellite estimates were also used to assess the dynamics of TSS in the Manantali dam. The comparison between an upstream and a downstream site shows first of all the low sediment supply from the upstream section and the relatively low impact of the Manantali dam on sediment retention in its lake. These results therefore confirm the estimates that have been made, which revealed that the Senegal River has brought small proportions of silt into this reservoir on average per year. They are also in line with the observations of SOGEM (Société de Gestion de l'Energie de Manantali) officials who confirm that the current siltation rate in the Manantali reservoir is negligible, which is consistent with the extreme clarity of the water in the reservoir (Faye, 2018). An impact study estimated that 530,000 tonnes of suspended solids, 50% of which were particles less than 0.002 millimetres in diameter, were added to the reservoir each year, estimating a lifespan of 450 years (estimated period of filling in the dead water volume of the reservoir by sedimentary deposits observed at the time) (Faye, 2018). However, TSS peaks may occur, as noted in this study, and lead to a significant potential sediment deposition. In addition, over an extended period of time, sediment will continue to accumulate behind the walls of the dam.

CONCLUSION

Landsat 8 satellite data provided valuable information on the spatial and temporal variability of suspended particulate matter in surface waters of the Senegal River basin. In the Senegal River basin, TSS are characterized by a high daily variability in their concentration, despite a fairly regular seasonal rhythm. TSS remains relatively high for much of the year. Daily rainfall and hydrological data were used to establish a relationship between river discharge and inverted TSS concentration from Landsat 8 satellite data for the site located at the outlet of the upper Senegal River basin. Shifts are noted between the peaks of TSS observed often in July and those of rainfall (August) and discharge (September). Satellite estimates are also used to assess the dynamics of TSS in the Senegal River basin and in the Manantali dam. Comparison of the TSS concentration on either side of the Manantali dam suggests low sediment retention in its reservoir, with possible peaks from time to time. The low concentrations of satellite derived TSS are consistent with the ranges of TSS deposits noted in the Manantali Dam reservoir.

This study has demonstrated the value of including water colour data in environmental studies to expand and complement analyses carried out through *in situ* measurements at broader spatial and temporal scales. In the future, it is therefore necessary to continue the qualification of Landsat-8 products, to carry out a post-processing chain for Landsat 8 images for "water colour" products, to develop regional algorithms for continental waters, to valorise products, to integrate water colour products into hydrological modelling (SWAT), to use water colour products for human health.

Acknowledgement

The authors warmly thank the teams of the GET Laboratory (Géosciences Environnement Toulouse) for their hospitality and the Embassy of France which awarded us the grant of French cooperation which allowed the stay in France.

References

Etcheber, H., Schmidt, S., Sottolichio, A., Maneux, E., Chabaux, G., Escalier, J.M., H., Derriennic, H., Schmeltz, M., Quemener, L., Repecaud, M., Woerther, P., Castaing, P. 2010. Monitoring water quality in estuarine environments: lessons from the MAGEST monitoring programme in the Gironde fluvial-estuarine system. Hydrology and Earth Systems Science Discussions, 7, 9411–9436.

Faye, C. 2018. Weight of transformations and major drifts related to major river water projects in Africa: case of the Manantali dam on the Senegal river basin. *Journal of Research in Forestry, Wildlife & Environment* 10(3), 13–24.

Faye, C., Diop, E. S., Mbaye, I. 2015. Impacts des changements de climat et des aménagements sur les ressources en eau du fleuve Sénégal : caractérisation et évolution des régimes hydrologiques de sous-bassins versants naturels et aménagés. *Belgeo 4*, 1–22. DOI: 10.4000/belgeo.17626

Hellweger, F. L., Schlosser, P., Weissel, J. K. 2004. Use of satellite imagery for water quality studies in New York Harbor. *Estuarine, Coastal and Shelf Science*, 61,437–448. DOI: 10.1016/j.ecss.2004.06.019

Janicot, S., Ali, A., Asencio, N, Berry, G., Bock, O., Bourles, B., Caniaux, G., Chauvin, F., Deme, A., Kergoat, L. et al. 2008. Large-scale overview of the summer monsoon over West and Central Africa during the AMMA field experiment in 2006. *Ann. Geophys*. 26, 2569–2595. DOI: 10.5194/angeo-26-2569-2008

Kaba, E., Philpot, W., Steenhuis, T. 2014. Evaluating suitability of MODIS-Terra images for reproducing historic sediment concentrations in water bodies: Lake Tana, Ethiopia. *Int. J. Appl. Earth Obs. Geoinform*. 26, 286–297. DOI: 10.1016/j.jag.2013.08.001

Kane, A. 1997. L'après-barrages dans la vallée du fleuve Sénégal: Modifications hydrologiques, morphologiques, géochimiques et sédimentologiques. Conséquences sur le milieu et les aménagements hydro-agricoles. Thèse de doctorat d'Etat, Univ. Dakar, Senegal.

Kane, A. 2005. Régulation du Fleuve Sénégal et flux de matières particulaire vers l'estuaire depuis la construction du Barrage de Diama. Sediment Budgets 2 (Proceedings of symposium S1 held during the Seventh IAHS Scientific Assembly at Foz do Iguaçu, Brazil, April 2005). IAHS Publ. 292, 2005.

Kattan, Z., Gac, J.Y., Probst, J.L. 1987. Suspended sediment load and mechanicalerosion in the Senegal basin – estimation of the surface runoff concentrationand relative contributions of channel and slope erosion. *J. Hydrobiol*. 92, 59–76. DOI: 10.1016/0022-1694(87)90089-8

Lafon, V., Robinet, A., Donnay, T., Doxaran, Lubac, B., Maneux, E., Aldo Sottolichio, A., Hagolle, O. 2014. RIVERCOLOR: chaîne de traitement des séries temporelles LANDSAT, SPOT et MODIS dédiée à la cartographie des matières en suspension en zone estuarienne. XIIIèTSS Journées Nationales Génie Côtier – Génie Civil Dunkerque, 2-4 juillet 2014, Editions Paralia CFL, 611–620.

Li, R., Li, J. 2004. Satellite Remote Sensing Technology for Lake Water Clarity Monitoring: An Overview. *Environmental Informatics Archives* 2, 893–901.

Mbaye, M.L., Gaye, A.T., Spitzy, A., Dähnke, K., Afouda, A., Gaye, B., 2016. Seasonal and spatial variation in suspended matter, organic carbon, nitrogen, and nutrient concentrations of the Senegal River in West Africa. *Limnologica* 57, 1–13. DOI: 10.1016/j.limno.2015.12.003

Moore, G.K. 1980. Satellite remote sensing of water turbidity/Sonde de téléTSS ure par satellite de la turbidité de l'eau. *Hydrol. Sci. Bull.* 25, 407–421. DOI: 10.1080/02626668009491950

Ndiaye, E. M. 2003. Le fleuve Sénégal et les barrages de l'OMVS : quels enseignements pour la mise en œuvre du NEPAD ? », *VertigO*, 4 (3) http://journals.openedition.org/vertigo/3883

OMVS, 2008. Plan d'Action Stratégique de Gestion des ProblèTSS Environnementaux Prioritaires du Bassin du Fleuve Sénégal, Projet FEM/Bassin du fleuve Senegal. 133 p.

OMVS, 2012. SENEGAL-HYCOS Document de projet Une composante du Système Mondial d'Observation du Cycle Hydrologique (WHYCOS). I.00585.001- DI-SFA 12-173, 110 p.

Palmateer, G.A., Mc Lean, D.E., Kutas, W.L., Meissner, S.M. 1993. Suspended particulate/ Bacteria interaction in agricultural drains, Ed. By S.S. Rao, 1–40.

Panthou, G., Vischel, T., Lebel, T. 2014. Recent trends in the regime of extreme rainfall in the Central Sahel. *Int. J. Climatol.* 34, 3998–4006. DOI: 10.1002/joc.3984

Reynolds, R.A., Stramski, D., Wright, V.M., Wozniak, S.B. 2010. Measurements and characterization of particle size distributions in coastal waters. *Journal of Geophysical Research* 115, C08024, DOI:10.1029/2009JC005930.

Robert, E., Grippa, M., Kergoat, L., Pinet, S., Gal, L., Cochonneau, G., Martinez, J.-M. 2016. Monitoring water turbidity and surface suspended sediment concentration of the Bagre Reservoir (Burkina Faso) using MODIS and field reflectance data. *Int. J. Appl. Earth Obs. Geoinform* 52, 243–251. DOI: 10.1016/j.jag.2016.06.016

Robert, E., Kergoat, L., Soumaguel, N., Merlet, M., Martinez, J-M. Diawara, M. and Grippa, M. 2017. Analysis of Suspended Particulate Matter and Its Drivers in Sahelian Ponds and Lakes by Remote Sensing (Landsat and MODIS): Gourma Region, Mali. *Remote Sensing* 9, 1272, 1–23. DOI: 10.3390/rs9121272

Rochelle-Newall, E., Nguyen, T.M.H., Le T.P.Q. 2015. Sengtaheuanghoung, O., Riblozi, O. A short review of fecal indicator bacteria in tropical aquatic ecosystem: Knowledge gaps and future directions. *Front. Microbiol*. 6, 1–15. DOI: 10.3389/fmicb.2015.00308

Rochette, C. 1974. *Monographie hydrologique du fleuve Sénégal*. Coll. Mém. ORSTOM, 1442 p.

Sottolichio, A., Castaing, P., Etcheber, H., Maneux, E., Schmeltz, M., Schmidt, S., 2011. Observations of suspended sediment dynamics in a highly turbid macrotidal estuary, derived from continuous monitoring. *Journal of Coastal Research* SI 64, 1579–1583.

Taylor, C., Beluši´c, D., Guichard, F., Parker, D.J., Vischel, T., Bock, O., Harris, P.P., Janicot, S., Klein, C., Panthou, G. 2017. Frequency of extreme Sahelian storms tripled since 1982 in satellite observations. *Nature* 544, 475–478. DOI: 10.1038/nature22069

Thiam, N. A. 2016. Allocation optimale de l'eau dans le bassin versant du fleuve Sénégal. Mémoire de Maîtrise en génie des eaux, Université de Laval, Québec, Canada, 84 p.

Troeger, C., Forouzanfar, M., Rao, P.C., Khalil, I., Brown, A., Reiner, R.C., Fullman, N., Jr., Thompson, R.L., Abajobir, A., Ahmed, M., et al. 2017. Estimates of global, regional, and national morbidity, mortality, and aetiologies of diarrhoeal diseases: A systematic analysis for the Global Burden of Disease Study 2015. Lancet Infect. Dis.

Troussellier, M., Got, P., Bouvy, M., Boup, M., Arfi, R., Lebihan, F., Monfort, P., Corbin, D., Bernard, C. 2004. Water quality and health status of the Senegal River estuary, *Marine Pollution Bulletin* 48, 9–10, 852–862. DOI: 10.1016/j.marpolbul.2003.10.028

Wang, X., Wang, Q., Liu, G. and Li, H., 2005. A study on the Quantitative Remote Sensing Model for the Suspended Sediment Concentration in Coastal Water with ASTER Conference paper, Report no. A290054.

Wilcox, C., Vischel, T., Panthou, G., Bodian, A., Blanchet, J., Descroix, L., ... & Kone, S. 2018. Trends in hydrological extremes in the Senegal and Niger Rivers. *Journal of Hydrology 566*, 531–545. DOI: 10.1016/j.jhydrol.2018.07.063

Zhang, W. 2014. Weighing the Pros and Cons: Transformation of Angle of View for Three Gorges Dam. *Natural Resources* 5, 1048–1056. DOI: 0.4236/nr.2014.516088

.

sciendo

Journal of Environmental Geography 13 (1–2), 43–52.

DOI: 10.2478/jengeo-2020-0005
ISSN 2060-467X

MACHINE LEARNING TECHNIQUES FOR LAND USE/LAND COVER CLASSIFICATION OF MEDIUM RESOLUTION OPTICAL SATELLITE IMAGERY FOCUSING ON TEMPORARY INUNDATED AREAS

Boudewijn van Leeuwen*, Zalán Tobak, Ferenc Kovács

Department of Physical Geography and Geoinformatics, University of Szeged, Egyetem u. 2-6, 6722 Szeged, Hungary
*Corresponding author, e-mail: Corresponding author: leeuwen@geo.u-szeged.hu

Research article, received 1 April 2020, accepted 1 May 2020

Abstract

Classification of multispectral optical satellite data using machine learning techniques to derive land use/land cover thematic data is important for many applications. Comparing the latest algorithms, our research aims to determine the best option to classify land use/land cover with special focus on temporary inundated land in a flat area in the south of Hungary. These inundations disrupt agricultural practices and can cause large financial loss. Sentinel 2 data with a high temporal and medium spatial resolution is classified using open source implementations of a random forest, support vector machine and an artificial neural network. Each classification model is applied to the same data set and the results are compared qualitatively and quantitatively. The accuracy of the results is high for all methods and does not show large overall differences. A quantitative spatial comparison demonstrates that the neural network gives the best results, but that all models are strongly influenced by atmospheric disturbances in the image.

Keywords: Sentinel 2, artificial neural network, random forest, support vector machine, machine learning, classification

INTRODUCTION

Accurate data about land use/land cover (LULC) of our surroundings continues to be important information for many applications like the monitoring and management of natural resources, development strategies, and global change studies. LULC changes due to changes in for example biological diversity, climate and terrestrial ecosystems, but are also drivers of change for these systems (Baamonde et al., 2019; Chatziantoniou et al., 2017). Satellite data classification for mapping of LULC is a common approach. Automatic classification of LULC with high accuracy based on medium resolution optical satellite imagery has been a challenge for decades. In earlier days, spatial and temporal resolution of the input data were limiting factors for accurate monitoring of LULC change. With the appearance of global medium to high resolution multispectral satellite data with a temporal resolution of just several days, in many situations input data for LULC change studies is available in abundance, even in situations where the phenomena change quickly. Advanced algorithms to process and classify large amounts of data can be used to produce accurate thematic maps over large areas and in a timely manner.

Supervised algorithms are a common approach to extract thematic information from multispectral satellite images. This research applied different nonparametric, machine learning algorithms for classification, namely support vector machine (SVM), random forest (RF) and deep artificial neural network (ANN).

Random Forest has been developed rapidly and has been widely used in many fields such as medicine, economics, and geography during the past twenty years. Breiman (2001) proposed Random Forest, which changes the way the classification or regression tree is constructed. It is an ensemble classification method consisting of many decision tree classification models (Jin et al., 2018). The RF algorithm exhibits good robustness compared to other traditional methods in the classification of a remote sensing image, because it requires fewer parameters, minimal manual intervention, and yields high classification accuracy; it can also manage high-dimensional data and obtain classification results rapidly (Ming et al., 2016). SVM employs optimization algorithms to locate the optimal boundaries between classes. Statistically, the optimal boundaries (hyperplanes) should be generalized to unseen samples with least errors among all possible boundaries separating the classes, therefore minimizing the confusion between classes (Huang et al. 2002). SVM were originally designed as a binary linear classifier, which assumes two linearly separable classes to be partitioned. SVM are further extended to deal with non-linear classification by using a non-linear kernel function to replace the inner product of optimal hyperplane. Moreover, SVM have been used for multi-class mapping through reducing the multi-class problem into a set of binary problems so that the basic SVM principles can be still applied (Shi and Yang, 2015). SVM and RF are able to deal with unbalanced data. Therefore, SVM and RF are becoming increasingly popular in image classification studies (Thanh and Kappas, 2018; Gudmann et al., 2019). In the 2000-s, (deep) neural networks started to make their

comeback due to the increased availability of data, optimization of the training algorithms and network architectures, and improvements of hardware, mainly the availability of affordable GPUs (Zhu et al., 2017). Recently, this development was combined with easier access to the algorithms via open source machine learning libraries, like scikit-learn, Pytorch and Keras-Tensorflow (Pedregosa et al., 2011; Paszke et al., 2019; Chollet, 2015; Abadi et al., 2015). Using open source programming language Python, these libraries can be used in combination with scientific data processing libraries like pandas, numpy, matplotlib and gdal.

This research aims to classify LULC based on a Sentinel 2 satellite image with special focus on inland excess water (IEW). This is a type of inundation that occurs in regions with very low relief intensity, where large areas get flooded due to a combination of a surplus of water, and limited infiltration and evaporation, or due to upwelling of groundwater. Due to its geographical characteristics the Great Hungarian Plain is particularly vulnerable to this phenomenon. IEW mostly occurs in agricultural areas where it results in reduced production and financial loss. Long term inundations cause reduced quality of agricultural soil (Szatmári and Van Leeuwen, 2013). Quite some earlier research has been published using traditional and more novel algorithms to extract inland excess water inundations from satellite data. Maximum likelihood classification was applied to identify IEW by Rakonczai et al. (2001) and Van Leeuwen and Tobak (2014). Szántó et al. (2008) applied unsupervised classification with a self-organising map to identify IEW. Mucsi and Henits (2010) applied spectral mixture analysis in a subset of our research area but with different classes. Van Leeuwen et al. (2012) applied a small feed forward multilayer perceptron to detect inland excess water on a mosaic of aerial photographs, spectral angle mapping was applied to hyperspectral data by Csendes and Mucsi (2017), Balázs et al (2018) used Random forest and SVM and received overall accuracies of over 90% while classifying PCA data extracted from Landsat 7 data. Other methods to extract water from medium resolution satellite data have been based on various indices (Lacaux et al., 2007; Feyisa et al., 2014). Szantoi et al. (2015) presents a comparison between maximum likelihood, decision tree and feed forward multilayer perceptron algorithms to classify different types of grass in a wetland area. Our research compares the results of SVM, RF and a deep ANN to detect LULC classes with inland excess water represented by two water classes. The accuracy of the results is statistically compared using overall accuracy and Cohen's Kappa.

DATA AND STUDY AREA

The research is focused on an area in the south of the Great Hungarian Plain that is vulnerable to inland excess water (Fig. 1). On average IEW inundations occur every two to three years in Hungary. The latest IEW period was in 2018, with moderate inundations in February and March. A Sentinel 2B Level 2A image (Drusch et al., 2012) from 18 March 2018 has been selected to test the different algorithms. Sentinel 2 L2A images contain Bottom Of Atmosphere (BOA) reflectance values stored in 100 x 100 km tiles. Bands

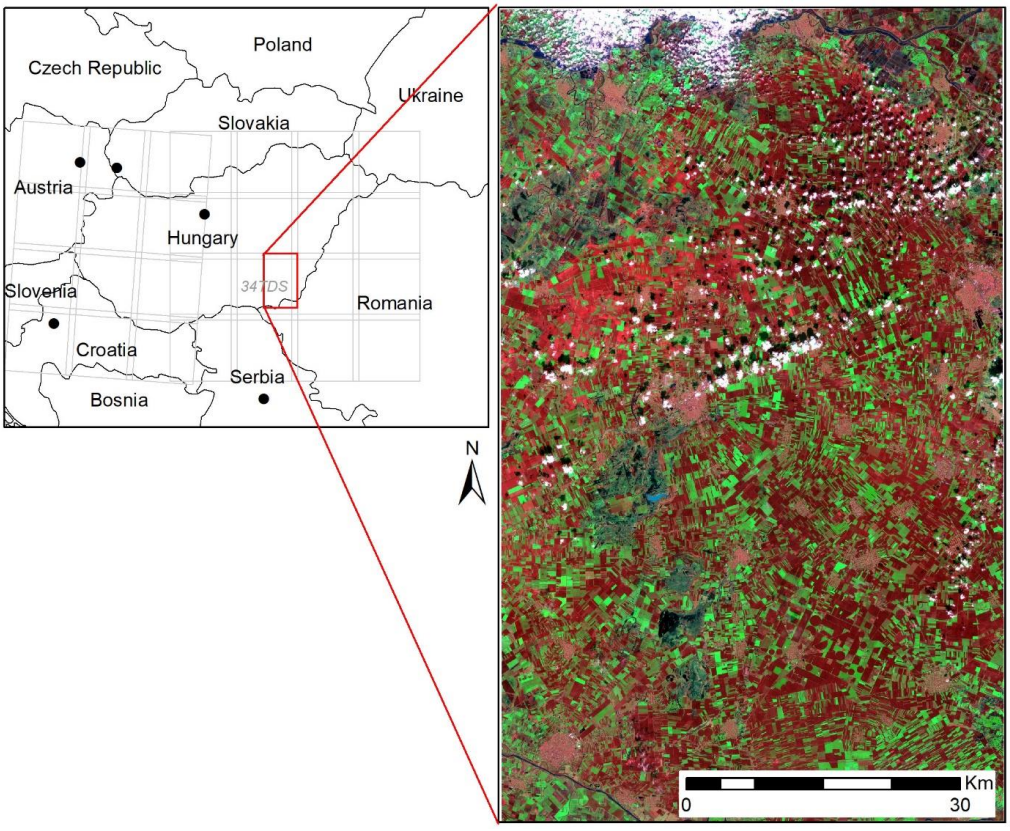

Fig. 1 Sentinel 2 false color composite (RGB843) showing the study area, its location in Hungary and the 34TDS tile in the Sentinel 2 tiling grid

2-8, 11 and 12 with spatial resolutions of 10 and 20 meter have been resampled to 10 meter and stacked into one composite file with 9 bands. Part of the original 34TDS tile did not have reflectance information due to the diagonal orientation of the satellite path compared to the Sentinel 2 tiling grid and was therefore cut off not to disturb the classifications. To exclude artifacts due to mosaicking, only one Sentinel 2 tile was used for the classifications, this way an area of 4900 km^2 remained to be classified. As usual during IEW periods, the image showed clouds and cloud shadows, which have a strong disturbing influence on classifications. The land use in the area is mainly agricultural, but there are several smaller cities and towns. In the north, the area is bounded by the Körös river, in the center and northwest larger natural wetlands and grasslands can be found (Mezősi, 2017). With elevations between 77 m and 105 m (above Baltic mean sea level) the relief intensity in the area is very small.

METHODS

Each experiment to classify the satellite image was designed in the same way (Fig. 2). First, the sample data was split into the three sets. Then, the inputs of the training and validation sets were standardized. In the next step, a model was defined as described in the next session. Several hyperparameters were tested for each model and once the optimal hyperparameters were determined, the model was trained using these hyperparameters. Then, the complete satellite image was imported and converted to a large 9-dimensional numpy array. The arrays were split into smaller subarrays to reduce memory use. The subarrays were sequentially fed to the trained model to predict new outputs. These outputs were concatenated to form an array with the same number of rows and columns as the input satellite image. The output array was then converted to a geoTiff file to be evaluated in a geographic information system (GIS). In the GIS, the test points were used to extract the classes from the model output (prediction) and compared to the actual classes (reference). Finally, a confusion matrix, overall accuracy and Cohen's Kappa were calculated.

Creation of the training, validation and test data set

Supervised classification methods require a large set of samples of input and output data to train the model to recognize the patterns forming the classes in the data set. A second data set is needed to define the hyperparameters and to validate if the model is not overfitting during the training phase. A third data set is used to independently test the predictions made using the trained model. To create the three datasets, polygons were digitized, by visual inspection of the different land cover classes on different RGB composites of the Sentinel 2 image. For each polygon, it was stored which class it represented. Then, the polygons were randomly split into three categories according to a 60/20/20 ratio and finally, they were converted to points (Fig. 3). Each point is a sample from one of eight LULC classes (Table 1).

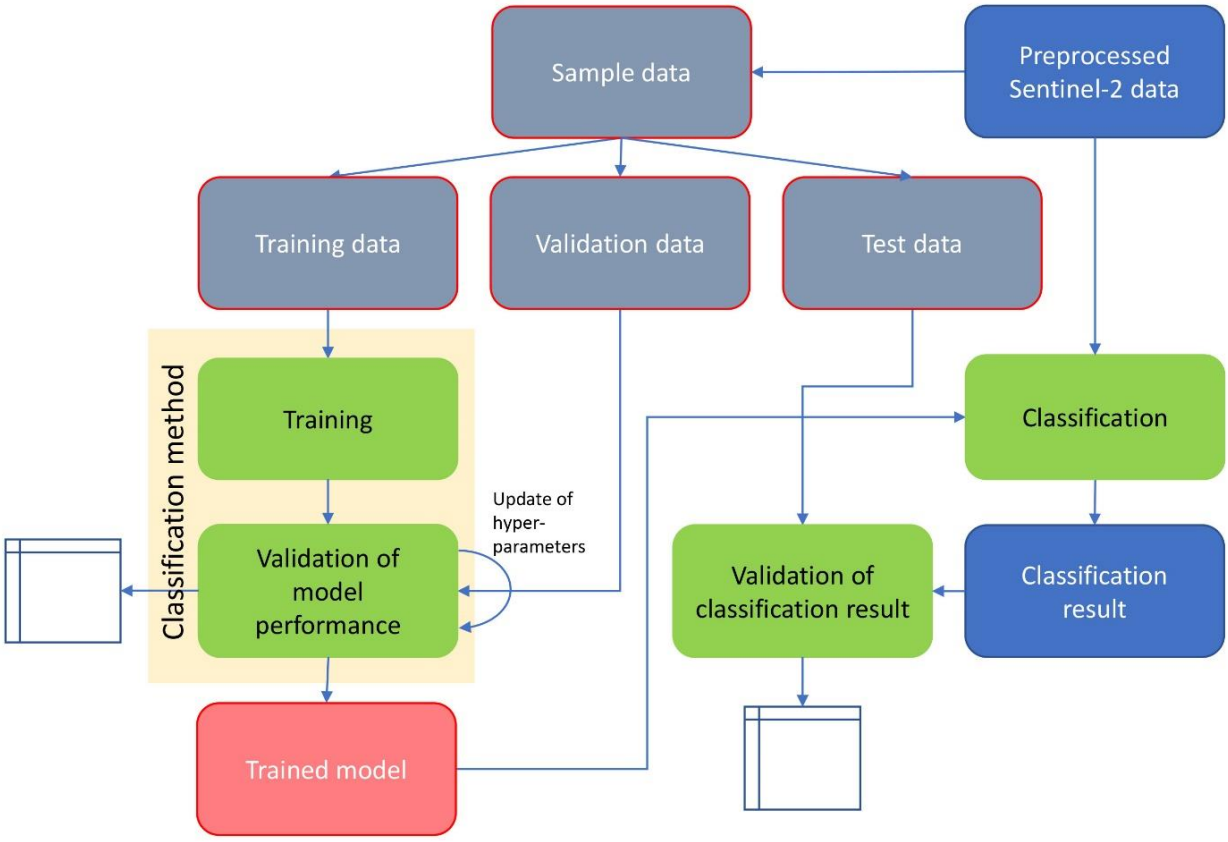

Fig. 2 Classification methodology

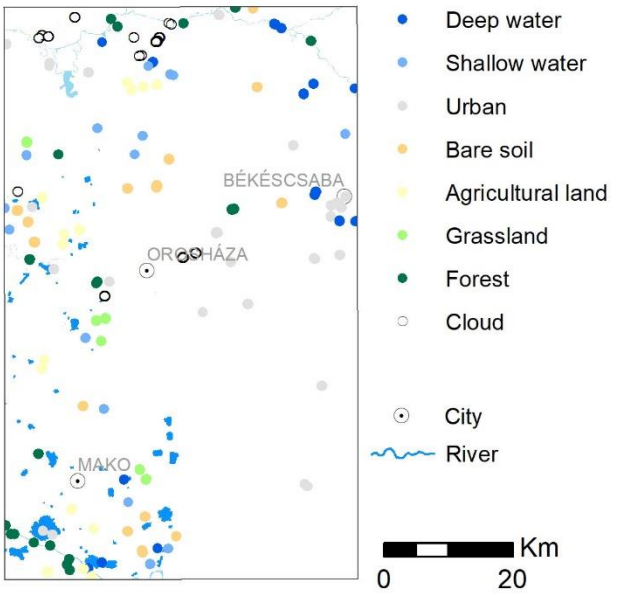

Fig. 3 Distribution of test points for each class in the study area

Table 1 Number of samples per LULC class in the training, validation and test sets

	LULC class	Number of samples		
		Training	Validation	Test
1	Deep water	1825	1040	1068
2	Shallow water	846	694	292
3	Urban	5294	2358	3112
4	Bare soil	6095	2733	2973
5	Agricultural land	2958	1597	1881
6	Grassland	1475	681	847
7	Forest	2805	2157	1668
8	Cloud	3793	2127	1585
	Total	25091	13387	13426

For each of the modelling algorithms, the same training, validation and test set was used. Each set was standardized by removing the mean and scaling to unit variance using Standard Scaler from the Scikit-learn library (Pedregosa et al., 2011).

Modelling

Three different models have been used to determine the best classification result. With the Scikit-learn gridSearch function many different hyperparameter combinations were evaluated and only results calculated using the optimal combinations are presented here. The first classification model is the Random Forest method. It was determined that the best result was reached with 200 trees in the forest and a maximum depth of 20 trees. The second method is the Support Vector Machines algorithm. For the kernel, a linear and an RBF kernel were tested, and the linear kernel showed better results. The regularization parameter C was set to 1.0.

The final model was a sequential deep artificial neural network with two hidden layers with 16 and 12 neurons in each layer built with Keras/Tensorflow (Abadi et al., 2015). This neural network is considered a deep neural network since it has two hidden layers, contrary to shallow neural networks with maximum one hidden layer. After each hidden layer, 20% dropout occurred to prevent overfitting. ReLu activation functions were used for the hidden layers and a softmax function for the output layer. The Adam optimization function with a learning rate of 0.001 was used for training. The ANN was trained with a batch size of 32 and 50 epochs on a Graphical processing unit (GPU). While training the model, 10-fold cross validation was calculated to determine the mean accuracy and variance (Chollet, 2015).

Prediction on the complete image using the trained model

After each model was trained, it was fed with the complete satellite image. Since the image dimension is 7382 x 8921 x 9 (columns x rows x bands) with 32 bits values, it was too large to fit it as a whole to the model. Therefore, after converting the image to a 3-dimensional numpy array, it was split in equal subarrays with a dimension of more or less 1000 x 1000 x 9. Each of the subarrays was then fed to the models and the prediction was calculated. The resulting predictions were concatenated to the original shape of the input numpy array and then the reconstructed array was exported to a TIFF image with the same spatial extent and coordinate system as the input satellite image.

Testing

The classified output image was read into a geographic information system (GIS) and at the locations of the random test points the classes were extracted. These classes were evaluated with the user defined classes at the same locations. Finally, a confusion matrix with overall accuracy and Cohen's Kappa were calculated to determine the independent validation accuracy (Congalton and Green, 2008).

RESULTS AND DISCUSSION

Qualitative assessment of the training and validation samples

The results of the predictions have been compared in a qualitative as well as quantitative matter. The qualitative comparison was performed on the whole image and on smaller areas with interesting features. To understand the distribution of the training and validation data sets, the distributions of the different classes were compared per satellite band (Fig. 4). The cloud class is not shown because in each band it is well separated from the other classes with much higher reflectance values. The training and validation samples have been randomly selected from all samples, which is reflected by the similarities of the patterns shown in both graphs. It can also be seen that the variation between the deep and shallow water classes is large, compared to urban, bare soil and grass land among most bands. Agricultural soil has the largest variation in band 6, 7 and 8. Forest has the largest variation in the bands with the longest wavelengths. Furthermore, it shows

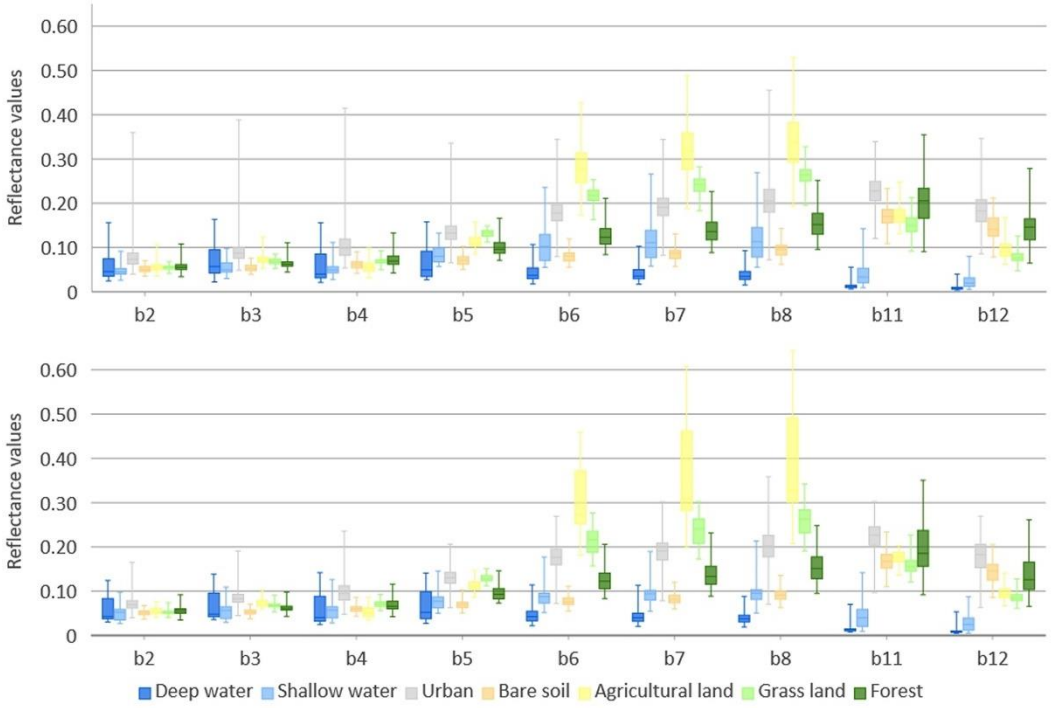

Fig. 4 Distribution of reflectance values for each class per band of the Sentinel 2 image for the training (top) and the validation set (bottom)

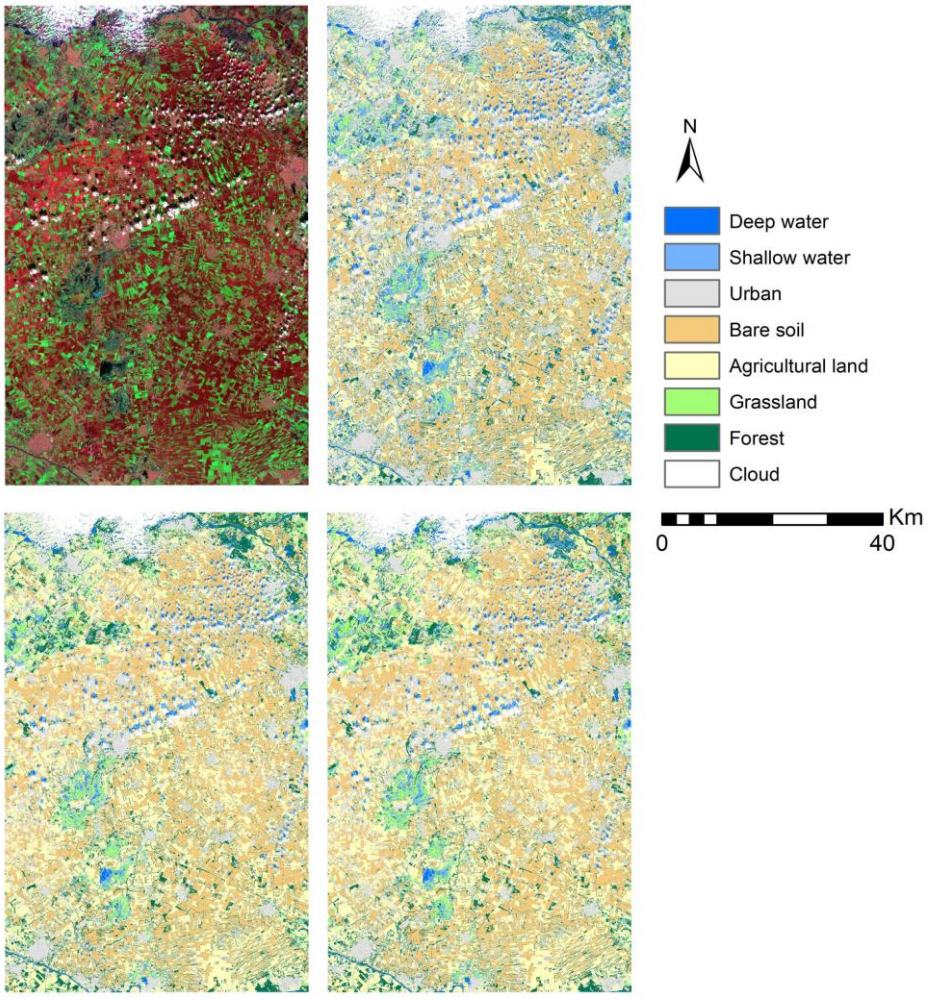

Fig. 5 Classification results for the total area: Sentinel 2 false color composite (RGB843) (upper left), Random forest result (upper right), SVM result (lower left), ANN result (lower right)

that, as expected, the variance between the reflectance values is the lowest in the bands with the shortest, visible wavelengths.

Qualitative comparison of the classification results

Visual inspection of the classification results shows a very similar pattern for all classification methods (Fig. 5). Large areas with continuous clouds in the north and scattered clouds in the center and top part of the images can be recognized in each result. Also, the pattern of large water bodies and urban centers in the area is shown in all three result. The mixture of bare soil and agriculture, typical for March in this area, is dominating the overall LULC pattern in the classifications. Overall, the classification results for SVM and ANN are more similar than for Random forest. This is confirmed by the total number of pixels classified in each class per method (Fig. 6).

Evaluating a smaller area with a large lake, surrounded by a mixture of wetland and grassland clearly shows differences between the Random forest, Support vector machine and Artificial neural network approaches (Fig. 7). The large, shallow lake in the south part of the subset is misclassified as urban by the RF method, SVM partly identifies the lake, while it is properly delineated by the ANN approach. All approaches overestimate the amount of urban in the area, but RF does this more often than the other approaches. Also, in many places, RF identified grassland is as scattered water pixels. The classification of bare soil and agricultural land is similar.

A subarea with more forests is shown in Figure 8. The forests along the river in the center are properly delineated by all approaches, but there are large differences between the amount of water south of the river where large parcels with soil heavily saturated with water can be found. ANN and RF classify these parcels almost exclusively as shallow water, while SVM designated them as forests. Many areas are misclassified as urban in the RF classification.

The third subarea is showing an urban area surrounded by a mixture of agricultural land and bare soils (Fig. 9). The urban area is classified similarly in all three approaches, but the small river and its banks flowing through the small city is only shown in the RF result. In SVM and ANN only the forest on the riverbanks is detected. Overall, the RF method is more sensitive to water than the other methods.

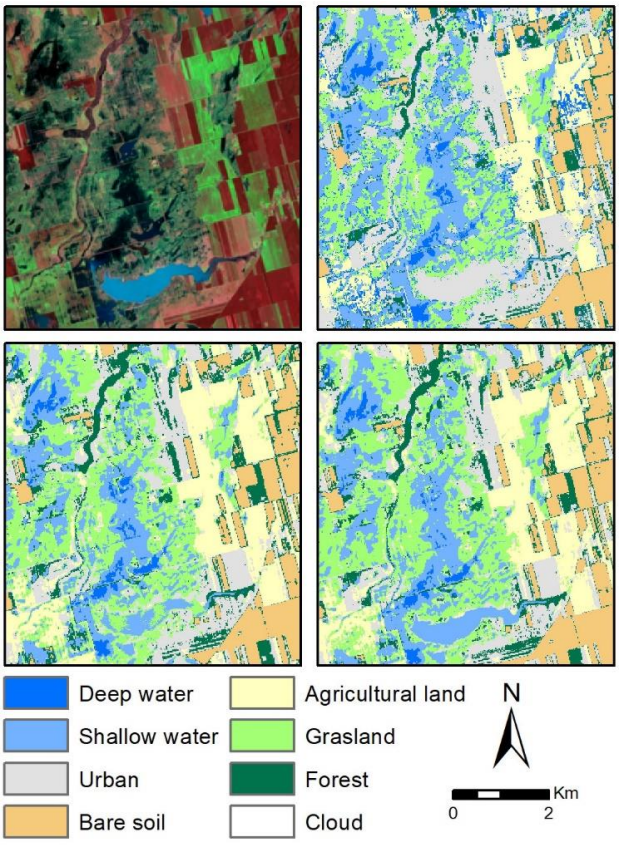

Fig. 7 Subset of the classification results with mainly inland excess water, wetlands and agriculture: Sentinel 2 false color composite (RGB843) (upper left), Random forest result (upper right), SVM result (lower left), ANN result (lower right)

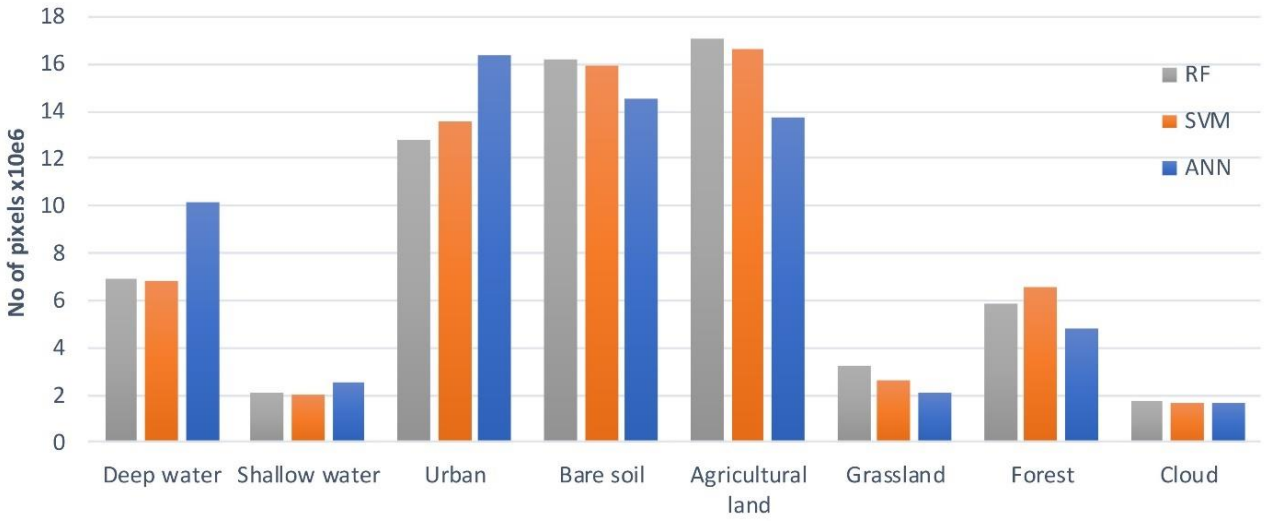

Fig. 6 Number of pixels per class for each classification method

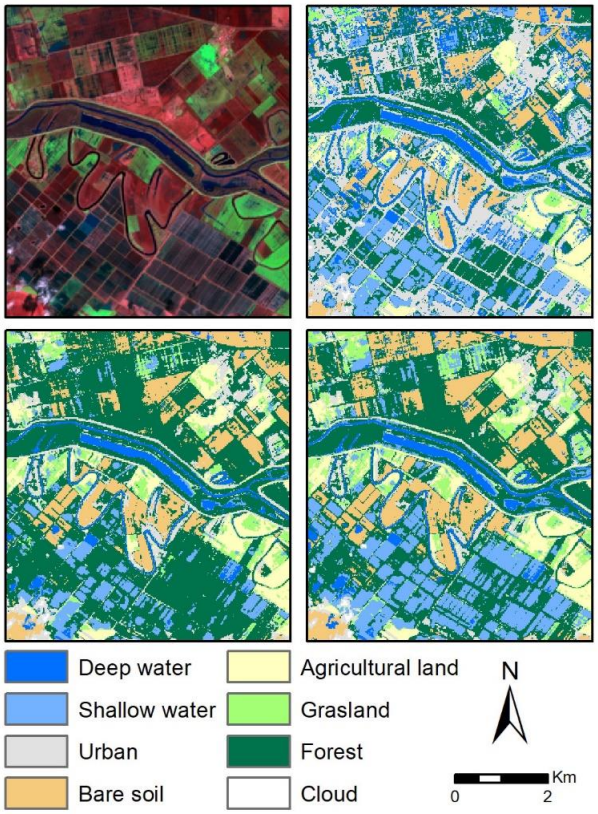

Fig. 8 Subset of the classification results with a mixture of open water, inland excess water and saturated soils: Sentinel 2 false color composite (RGB843) (upper left), Random forest result (upper right), SVM result (lower left), ANN result (lower right).

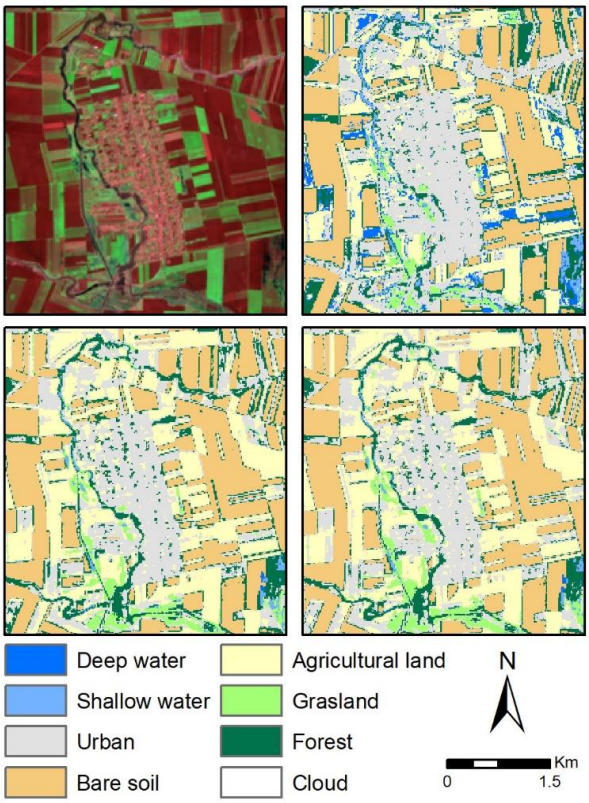

Fig. 9 Subset of the classification results with mainly urban land use and agricultural land and bare soil: Sentinel 2 false color composite (RGB843) (upper left), Random forest result (upper right), SVM result (lower left), ANN result (lower right).

The last subarea shows the effect of clouds and cloud shadows on the classifications (Fig. 10). In all three methods, this causes problems, even though a cloud class was added to the training set. The clouds themselves are classified properly, although at their boundaries, where they are less thick, they cause every method to misclassify them as urban. The cloud shadows cause bigger problems. Without exception, the shadows are misclassified as deep or shallow water due to their darkening effect on the land cover. Often bare soil is misclassified as deep water, while agricultural land is wrongly identified as shallow water.

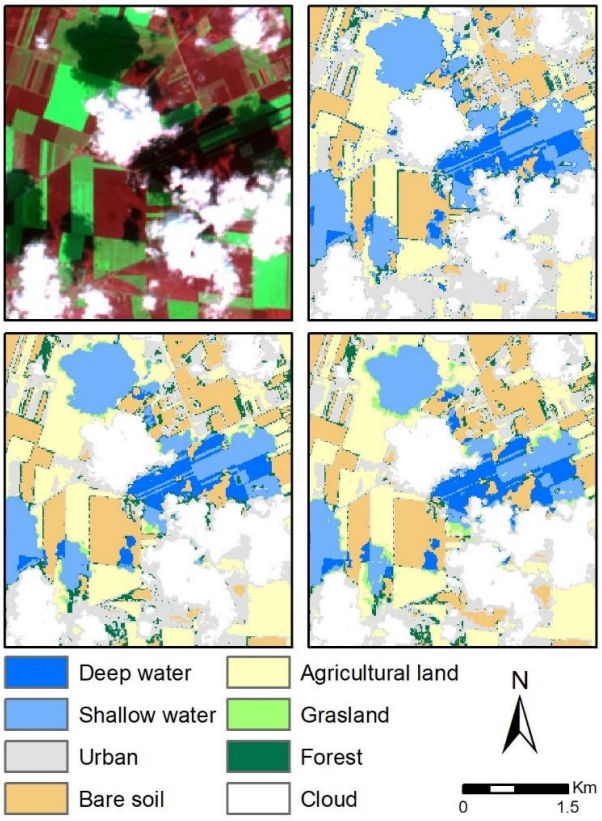

Fig. 10 Subset of the classification results with clouds and cloud shadows: Sentinel 2 false color composite (RGB843) (upper left), Random forest result (upper right), SVM result (lower left), ANN result (lower right)

Quantitative comparison of the classification results

Apart from the visual comparison between the classification results, two quantitative comparisons have been performed to evaluate the methods. The first comparison provides the average accuracy and its variation for each method based on a 10-fold cross validation calculated using the validation set. The second method is based on the test set that was used to calculate the confusion matrix giving the results from the trained models and the expected results. Based on the confusion matrix, the overall accuracy (OA), User's accuracy, producer's accuracy and Cohen's Kappa were calculated (Congalton and Green, 2008).

The Random Forest method had an average accuracy on the cross validation of 0.9275, with a variation of 0.0381. The confusion matrix is shown in Table 2. The urban prediction class contained most misclassified pixels, mainly agricultural land, grassland and forest. To

a smaller extent, these classes were also classified as deep water. The bare soil, grassland and clouds output classes never contain wrongly classified pixels.

Compared to RF, the support vector machine approach has a higher average accuracy of 0.9724 with a lower variation of 0.0184, while the test results with the independent data set are also slightly better (Table 3). Most misclassification occurred in the urban class. These are mainly agricultural land and grassland, and to a lesser extent forest and clouds. Also urban was misclassified as forest and grassland as agricultural land.

Finally, the ANN has a slightly lower average accuracy of 0.9628 with a variation of 0.0258. The test results with the independent data set is very similar (Table 4). The misclassification pattern for ANN is more or less the same as for the SVM method, although some forest was also misclassified as shallow water.

All three methods gain very high accuracy classification results (above 0.9), with RF having the lowest accuracies and SVM having very similar, but slightly better results than ANN. Comparing the training times for the three algorithms, it is clear that training the SVM model (0.567 seconds) is much more efficient than the other models (RF: 18.8 seconds, ANN: 157 seconds).

Although, in general water can usually be detected with high accuracy in multispectral images, in case of the ANN model there was a relatively high error due to misclassification as forest. This might be caused by the forests in shallow water along the Tisza river in the study area. RF and SVM did not show this misclassification. Deep water was classified almost perfectly by SVM and ANN, but RF had more problems with this class, with misclassifications in multiple other classes. All three classifications show relatively large errors for the urban class. The main reason for this might be the mixed pixels in the urban class due to the resolution of Sentinel 2 data. Agricultural land and grassland are mixed by the SVM and ANN classifications which can be explained by thematic misclassification of the training samples.

Misclassifications often happened in areas with shadow from clouds. To overcome this problem, many methods have been developed, ranging from shadow modelling, cloud and shadow masking, to sampling of classes in the shadows (Shahtahmassebi et al., 2013; Foga et al., 2017). The aim of this research was to evaluate the differences between the selected algorithms therefore the cloud/shadow problem has been ignored in the calculations.

The data sets that have been used for training the models is relatively small. In the current revolution of deep learning the larger the data set, the better for deep neural networks. A larger training set may therefore result in higher accuracy for the ANN, but of course this will also result in longer training times. For the current classifications, the accuracy is already well above 0.90, therefore the need for higher accuracy is not apparent for the presented LULC application with relatively few classes. If the number of classes would increase, more training is required and the advantage for ANN would be higher (Rai et al., 2020).

Experiments were executed to automatically extract LULC classes from the Corine Land Cover 2018 database (CLC 2018) and use these as labels for the training data sets. Unfortunately, this method to automatically create a larger training data set was not successful because the spatial and thematic resolution of CLC2018 compared to Sentinel 2 data is too low resulting in many mixed classes within one CLC2018 polygon. Training the models with these labels caused large errors. The application of other land cover data sets with higher resolution, like Copernicus High Resolutions layer (Büttner, 2012) or National Ecosystem Base map (Tanács et al., 2019) may provide better results.

The classification algorithms were applied to data from the Sentinel 2 satellite, but can be applied to any medium resolution multispectral satellite data set.

Table 2 Confusion matrix with the random forest predictions in the columns and the true values (test set) in the rows

RF	Deep water	Shallow water	Urban	Bare soil	Agricul-tural land	Grassland	Forest	Cloud	Total	Users acc
Deep water	1067	1	0	0	0	0	0	0	1068	0.9991
Shallow water	0	291	1	0	0	0	0	0	292	0.9966
Urban	8	0	3045	0	15	0	44	0	3112	0.9785
Bare soil	2	0	0	2971	0	0	0	0	2973	0.9993
Agricultural land	32	0	160	0	1689	0	0	0	1881	0.8979
Grassland	37	0	157	0	29	624	0	0	847	0.7367
Forest	44	0	251	0	0	0	1373	0	1668	0.8231
Cloud	0	0	3	0	0	0	0	1582	1585	0.9981
Total	1190	292	3617	2971	1733	624	1417	1582	13426	
Prod acc.	0.8966	0.9966	0.8419	1.0000	0.9746	1.0000	0.9689	1.0000		
									OA	0.9416
									Kappa	0.9299

Table 3 Confusion matrix with the Support Vector Machine predictions in the columns and the true values (test set) in the rows

SVM	Deep water	Shallow water	Urban	Bare soil	Agricul-tural land	Grassland	Forest	Cloud	Total	Users acc
Deep water	1066	1	0	0	0	0	1	0	1068	0.9981
Shallow water	0	289	1	0	0	1	1	0	292	0.9897
Urban	0	0	3000	0	29	0	83	0	3112	0.9640
Bare soil	0	0	0	2973	0	0	0	0	2973	1.0000
Agricultural land	0	0	130	0	1751	0	0	0	1881	0.9309
Grassland	0	0	71	0	103	673	0	0	847	0.7946
Forest	0	0	21	0	0	1	1646	0	1668	0.9868
Cloud	0	0	16	0	0	0	0	1569	1585	0.9899
Total	1066	290	3239	2973	1883	675	1731	1569	13426	
Prod acc.	1.0000	0.9966	0.9262	1.0000	0.9299	0.9970	0.9509	1.0000		
									OA	0.9658
									Kappa	0.9591

Table 4 Confusion matrix with the artificial neural network predictions in the columns and the true values (test set) in the rows

ANN	Deep water	Shallow water	Urban	Bare soil	Agricul-tural land	Grassland	Forest	Cloud	Total	Users acc
Deep water	1065	3	0	0	0	0	0	0	1068	0.9972
Shallow water	0	291	1	0	0	0	0	0	292	0.9966
Urban	0	0	2953	3	69	0	87	0	3112	0.9489
Bare soil	0	0	0	2973	0	0	0	0	2973	1.0000
Agricultural land	0	0	110	0	1771	0	0	0	1881	0.9415
Grassland	0	6	45	0	121	675	0	0	847	0.7969
Forest	2	37	28	0	0	0	1601	0	1668	0.9598
Cloud	0	0	3	0	0	0	0	1582	1585	0.9981
Total	1067	337	3140	2976	1961	675	1688	1582	13426	
Prod acc.	0.9981	0.8635	0.9404	0.9990	0.9031	1.0000	0.9485	1.0000		
									OA	0.9616
									Kappa	0.9542

CONCLUSION

Automatic supervised classification of multispectral satellite imagery is required to extract land use / land cover data for a wide range of applications. Machine learning algorithms are the most promising techniques to reach this goal. The field is developing rapidly, and new algorithms and implementations are becoming available continuously. The application of machine learning algorithms in LULC classification can result in high quality results, as the classification results of this research shows. Each presented methodology has an overall accuracy and a Cohen's Kappa of above 0.90. The deep water class could be detected almost perfectly, while there was some misclassification of the shallow water class. Clouds are detected very well, but their shadows cause the largest misclassifications. With the application of open source machine learning and scientific data processing libraries, it becomes straightforward to efficiently experiment with different algorithms and parameters to determine the optimal classification routine for a certain application.

With improved classification of inland excess water inundations based on satellite imagery covering large areas, this research supports the operational defense against the floods, and helps to understand their development. The

inland excess water maps can be used as input for scientific study of the phenomenon and to support sustainable water management.

Acknowledgements

This research was supported by the WATERatRISK project (HUSRB/1602/11/0057) and the Time series analysis of land-cover dynamics using medium and high-resolution satellite images project (OTKA, NKFIH 124648K).

References

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G.S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jozefowicz, R., Jia, Y., Kaiser, L., Kudlur, M., Levenberg, J., Mané, D., Schuster, M., Monga, R., Moore, S., Murray, D., Olah, C., Shlens, J., Steiner, B., Sutskever, I., Talwar, K., Tucker, P., Vanhoucke, V., Vasudevan,V., Viégas F., Vinyals, O., Warden, P., Wattenberg, M., Wicke, M., Yu, Y., Zheng, X. 2015. TensorFlow: Large-scale machine learning on heterogeneous systems. Software available from tensorflow.org.

Baamonde, S., Cabana, M., Sillero N., Penedo, M.G., Naveira, H., Novo, J. 2019. Fully automatic multi-temporal land cover classification using Sentinel-2 image data. *Procedia Computer Science 159*, 650–657. DOI: 10.1016/j.procs.2019.09.220

Balázs, B., Bíró, T., Dyke, G., Singh, S.K., Szabó, Sz. 2018. Extracting water-related features using reflectance data and principal component analysis of Landsat images. *Hydrological Sciences Journal* 63(2), 269–284. DOI: 10.1080/02626667.2018.1425802

Breiman, L. 2001. Random Forests. *Machine Learning* 45(5–32). DOI:10.1023/A:1010933404324

Büttner, G., 2012. Guidelines for verification and enhancement of high resolution layers produced under GMES initial operations (GIO) Land monitoring 2011–2013. EEA Report

Chatziantoniou, A., Petropoulos, G.P., Psomiadis E. 2017. Co-Orbital Sentinel 1 and 2 for LULC Mapping with Emphasis on Wetlands in a Mediterranean Setting Based on Machine Learning. *Remote Sensing* 9, 1259. DOI:10.3390/rs9121259

Chollet, F. 2015. Keras, https://keras.io [04-20-2020]

CLC, 2018. Corine Land Cover (CLC) 2018, Version 20. European Environment Agency. https://land.copernicus.eu/pan-european/corine-land-cover/clc2018 [04-20-2020]

Congalton, R.G., Green, K. 2008. Assessing the accuracy of remotely sensed data: principles and practices. CRC, Boca Raton London New York, 183 p

Csendes, B. Mucsi, L. 2016. Inland excess water mapping using hyperspectral imagery. *Geographica Pannonica* 20 (4), 191–196. DOI: 10.18421/GP20.04-01

Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P. Martimort, P., Meygret, A., Spoto, F., Sy, O., Marchese, F., Bargellini, P. 2012. Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services. *Remote Sensing of Environment* 120, 25–36, DOI: 10.1016/j.rse.2011.11.026

Feyisa, G.L., Meilby, H., Fensholt, R., Proud, S.R. 2014. Automated Water Extraction Index: A new technique for surface water mapping using Landsat imagery. *Remote Sensing of Environment* 140, 23–35

Foga, S., Scaramuzza, P.L., Guo, S., Zhu, Z., Dilley, R.D., Beckmann, T., Schmidt, G.L., Dwyer, D.J., Hughes, M.J., Laue, B. 2017. Cloud detection algorithm comparison and validation for operational Landsat data products. *Remote Sensing of Environment* 194, 379–390, DOI: 10.1016/j.rse.2017.03.026

Gudmann, A., Mucsi, L., Henits, L. 2019. A CORINE felszínborítási térkép automatikus előállításának lehetősége döntésifa-osztályozó segítségével. (Automatic land cover mapping using decision tree classifier). *Geodézia és Kartográfia* 71(2), 9–13. (in Hungarian)

Huang, C., Davis, L.S., Townshend, J.R.G. 2002. An assessment of support vector machines for land cover classification. *International Journal of Remote Sensing* 23(4), 725–749, DOI: 10.1080/01431160110040323

Jin, Y., Liu, X., Chen, Y., Liang, X. 2018. Land-cover mapping using Random Forest classification and incorporating NDVI time-series and texture: a case study of central Shandong. *International Journal of Remote Sensing* 39 (23), 8703–8723, DOI: 10.1080/01431161.2018.1490976

Lacaux, J.P., Tourre, Y.M., Vignolles, C., Ndione, J.A., Lafaye, M. 2007. Classification of ponds from high-spatial resolution remote sensing: Application to Rift Valley Fever epidemics in Senegal. *Remote Sensing of Environment* 106, 66–74. DOI: 10.1016/j.rse.2006.07.012

Mezősi G. 2017. Physical Geography of Hungary. Heidelberg, London, New York, Springer, 334 p

Ming, D., Zhou, T., Wang, M., Tan, T. 2016. Land cover classification using random forest with genetic algorithm-based parameter optimization. *J. Appl. Remote Sens.* 10 (3), 035021. DOI: 10.1117/1.jrs.10.035021

Mucsi, L., Henits, L. 2010. Creating excess water inundation maps by sub-pixel classification of medium resolution satellite images. *Journal of Environmental Geography* 3 (1–4), 31–40.

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Köpf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., Bai, J., Chintala, S. 2019. PyTorch: An imperative style high-performance deep learning library. *Proc. Adv. Neural Inf. Process. Syst.* 32, 8024–8035.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O.,Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E. 2011. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* 12, 2825–2830

Rai, A.K., Mandal, N., Singh, A., Singh, K.K. 2020. Landsat 8 OLI Satellite Image Classification using Convolutional Neural Network. *Procedia Computer Science* 167, 987–993. DOI:10.1016/j.procs.2020.03.398

Rakonczai, J., Mucsi, L., Szatmári, J., Kovács, F., Csató, Sz. 2001. A belvizes területek elhatárolásának módszertani lehetőségei (Methods for delineation of inland excess water areas). A földrajz eredményei az új évezred küszöbén. Az I. Magyar Földrajzi Konferencia CD 14 p. (in Hungarian)

Shahtahmassebi, A., Yang, N., Wang, K., Moore, N., Shen, Z. 2013. Review of shadow detection and de-shadowing methods in remote sensing. *Chinese Geographical Science* 23, 403–420. DOI: 10.1007/s11769-013-0613-x

Shi D., Yang, X. 2015. Support Vector Machines for Land Cover Mapping from Remote Sensor Imagery. In: Li, J., Yang X. (eds.) Monitoring and Modeling of Global Changes: A Geomatics Perspective. Springer Remote Sensing/Photogrammetry, Dordrecht, DOI: 10.1007/978-94-017-9813-6_13

Szántó, G., Mucsi, L., van Leeuwen, B. 2008. Application of self-organizing neural networks for the delineation of excess water areas. *Journal of Env. Geogr.* 1 (3-4), 15–20.

Szantoi Z, Escobedo FJ, Abd-Elrahman A, Pearlstine L, Dewitt B, Smith S. 2015, Classifying spatially heterogeneous wetland communities using machine learning algorithms and spectral and textural features. *Environ Monit Assess* 187 (5), 262. DOI: 10.1007/s10661-015-4426-5

Szatmári, J., van Leeuwen, B. 2013. Inland Excess Water – Belvíz – Suvišne Unutrašnje Vode, Szeged, Újvidék, Szegedi Tudományegyetem, Újvidéki Egyetem, 154 p

Tanács E., Belényesi M., Lehoczki R., Pataki R., Petrik O., Standovár T., Pásztor L., Laborczi A., Szatmári G., Molnár Zs., Bede-Fazekas Á., Kisné Fodor L., Varga I., Zsembery Z., Maucha G. 2019. Országos, nagyfelbontású ökoszisztéma- alaptérkép: módszertan, validáció és felhasználási lehetőségek. (National high resolution ecosystem base map: Methodology, validation and possibilities for applications). *Természetvédelmi közlemények* 25, 34–58. DOI: 10.17779/tvk-jnatconserv.2019.25.34. (in Hungarian)

Thanh-Noi, P., Kappas, M. 2018. Comparison of Random Forest, k-Nearest Neighbor, and Support Vector Machine Classifiers for Land Cover Classification Using Sentinel-2 Imagery. *Sensors* 18 (2), 18. DOI: 10.3390/s18010018

van Leeuwen, B., Mezősi, G., Tobak, Z., Szatmári, J., Barta, K. 2012. Identification of inland excess water floodings using an artificial neural network. *Carpathian Journal of Earth and Environmental Sciences* 7 (4), 173–180.

van Leeuwen, B., Tobak, Z. 2014. Operational Identification of Inland Excess Water Floods Using Satellite Imagery, In: Vogler, R., Car, A., Strobl, J.Griesebner, G. (Eds.), GI_Forum 2014. Geospatial Innovation for Society. Herbert Wichmann Verlag, VDE Verlag GMBH, Berlin/Offenbach, 12–15. DOI: 10.1553/giscience2014s12

van Leeuwen, B., Tobak, Z., Kovács, F. 2020. Sentinel 1 and 2 based near real time inland excess water mapping for optimized water management. *Sustainability* 12 (7), 2854. DOI: 10.3390/su12072854

Zhu, X.X., Tuia, D., Mou, L., Xia, G-S.,Zhang, L., Xu, F., Fraundorfer, F. 2017. Deep Learning in Remote Sensing: A Comprehensive Review and List of Resources. *IEEE Geoscience and Remote Sensing Magazine* 5(4) 8–36. DOI: 10.1109/mgrs.2017.2762307

Journal of Environmental Geography 13 (1–2), 53–61.

DOI: 10.2478/jengeo-2020-0006
ISSN 2060-467X

IDENTIFICATION OF RIPARIAN VEGETATION TYPES WITH MACHINE LEARNING BASED ON LIDAR POINT-CLOUD MADE ALONG THE LOWER TISZA'S FLOODPLAIN

István Fehérváry[1]*, Tímea Kiss[2]

[1]Lower Tisza Hydrological Directorate, Stefánia 4, 6720 Szeged, Hungary
[2]University of Szeged, Department of Physical Geography and Geoinformatics, Egyetem str. 2-6, 6722 Szeged, Hungary
*Corresponding author, e-mail: FehervaryI@ativizig.hu

Research article, received 7 April 2020, accepted 5 May 2020

Abstract

The very dense floodplain vegetation on the artificially confined floodplains results in decreased flood conveyance, thus increase in flood levels and flood hazard. Therefore, proper floodplain management is needed, which must be supported by vegetation assessment studies. The aims of the paper are to introduce the method and the results of riparian vegetation classification of a floodplain area along the Lower Tisza (Hungary) based on automatized acquisition of airborne LiDAR survey. In the study area 15x15 m large training plots (voxels) were selected, and the statistical parameters of their LiDAR point clouds were determined. Applying an automatized parameter selection and 10-fold cross-validation he most suitable decision tree was selected, and following a series of classification steps the training plots were classified. Based on the decision tree all the pixels of the entire study area were analysed and their vegetation types were determined. The classification was validated by field survey. On the studied floodplain area the accuracy of the classification was 83%.

Keywords: airborne LiDAR, scikit-learn, Gini impurity, decision tree, riparian forest,

INTRODUCTION

During the last one and half century several environmental effects (e.g. climate change, land cover alterations) affected the floodplains and river channels altering their characteristics. However, these semi-natural effects were exceeded by the consequences of various river engineering works: the channel and floodplain regulations works altered the hydrological processes, and as one of the consequences the riparian vegetation changed too.

The riparian vegetation highly influences the channel-floodplain connections. For example, the vegetation along a river stabilises the banks (Abernethy and Rutherfurd, 1998), or decreases the overbank flow velocities (Kiss et al., 2019a). Along the river-banks the density of vegetation primarily influences the development of natural levees (Nagy et al., 2018), whilst in the distal floodplain areas the vegetation influences the flood flow directions and velocities (Rátky and Farkas, 2003; Zellei and Sziebert, 2003; Brooks, 2005; Corenblit et al., 2007; Geerling et al., 2008), thus the vertical aggradation pattern (Steiger et al., 2001; Kiss and Sándor, 2009).

These processes are related to the roughness controlling function of the vegetation, which influences the flood stages too (Jalonen et al., 2015; Kiss et al., 2019a). The highest peak flood level on record was set in 1970 along the Tisza River, Hungary, however along the Lower Tisza this record was overprinted in 2000 and also in 2006 by higher stages of 80 cm (Kiss et al., 2019b), while the discharge of these record floods were lower than in 1970 (Kovács and Váriné Szöllősi, 2003). These hydrological changes draw attention to flood conductivity decline of the floodplain: nowadays only 13% of the flood discharge is drained on the floodplain while it was 23% in 1970 (Kovács and Váriné Szöllősi, 2003). It could be partly related to the very dense floodplain vegetation, therefore proper floodplain management is needed, which must be supported by vegetation assessment studies. Their first step should be the identification of riparian vegetation types; therefore, our aim is to apply the latest methods for this identification using automatized acquisition of LiDAR survey data.

Lately several researches relied on the statistical analysis of point-clouds of LiDAR surveys to identify vegetation characteristics and types. Most of these researches were made in the field of forestry or ecology. For example, Hudak et al. (2008) identified various tree species in forest patches based on various statistical parameters of the LiDAR points representing the canopy. Heurich and Thoma (2008) did similar research, but they measured some dendrological parameters too (e.g. tree height, canopy width) based on LiDAR data. Naesset et al. (2004) combined airborne and terrestrial LiDAR data to calculate the main parameters (i.e. number of stems, volume of harvestable wood) of forest units. Jung et al. (2011) calculated not only the parameters of trees (i.e. tree height, lower canopy height, canopy volume, stem diameter), but also the temporal changes between two survey campaigns including both LiDAR technologies. Though these researches slightly differ in the identification of vegetation types and in defining their parameters, it could be concluded, that some high-

resolution parameters could be applied just on small areas or on individual level, while the measurements on larger areas have limited resolution and less accuracy.

During the last years the LiDAR based vegetation analysis has been applied also in hydrological studies. For example, Vetter et al. (2011) determined the vegetation roughness based on the spatial connection of voxels (3D pixels) and the rate of reflections, using an airborne LiDAR survey with very high point density (>25 point/m^2). The resulted vegetation density values were applied in a 2D hydraulic model, and the derived hydrological data were compared to modelled data based on classical land-use category maps. Vetter et al. (2011) concluded, that the LiDAR based modelling gave much more reliable results, as the hydrological data were closer to the actually measured ones. Similarly, Manners et al. (2011) determined the role of Tamarix bushes in vegetation roughness using terrestrial LiDAR survey.

During the late 20th century forest became the dominant land cover type along the Tisza River and invasive species became widespread in the undergrowth, thus the vegetation roughness of the floodplain drastically increased, as it was indicated by point-based classical vegetation surveys (Kiss et al., 2019a). As the high vegetation roughness fundamentally decreases the overbank flow velocities and increases flood peak levels, high-resolution and up-to-date data would be needed for precise flood modelling. However, no such a dataset exists.

Therefore, our aims are to introduce the method and the results of riparian vegetation classification, based on automatized acquisition of airborne LiDAR survey. Within this article our goals are to describe the detailed methodology of the automatized classification, to classify the vegetation on a study area located along the Lower Tisza River, and to evaluate the feasibility of the method.

STUDY AREA

The research was conducted on a 3 km^2-large floodplain area of the Lower Tisza (197-194 fkm) between the settlements of Algyő and Szeged (Fig. 1). The floodplain is artificially confined by artificial embankments to 800 m; and the river channel is 130 m wide in average. At Algyő gauging station the greatest flood-stage was at 84.65 m asl, while the lowest stage was measured at 71.55 m asl, thus the absolute change in water level is ca. 13 m. The flooding of the floodplain starts when the water is at 80 m asl, and up to 5-6 m deep water column develops over the floodplain at the time of record floods. The slope of the water is very low (2.9 cm/km), therefore the overbank flow velocity is also low (max. 0.1-0.2 m/s).
In the mid-19th century the river regulation works (e.g. artificial cut-offs, building artificial levees) resulted in considerable land-use changes of the floodplain (Kiss et al. 2019a). At the beginning of the 20th century the former wetlands were replaced by meadows, pastures and plough fields, while the proportion of forests remained low. However, as the result of intensive afforestation in the 1970-80s the proportion of forests increased above 70% in the 1980s, and nowadays it is above 80%. These land-cover changes resulted in fourfold increase in vegetation roughness (from 0.02 to 0.08), which is even higher (0.13) if the dense stands of invasive species in forest and on fallow lands are considered. Among the invasive species the *Amorpha fruticosa* is the most abundant (11%), which creates impenetrable shrubbery. According to our latest

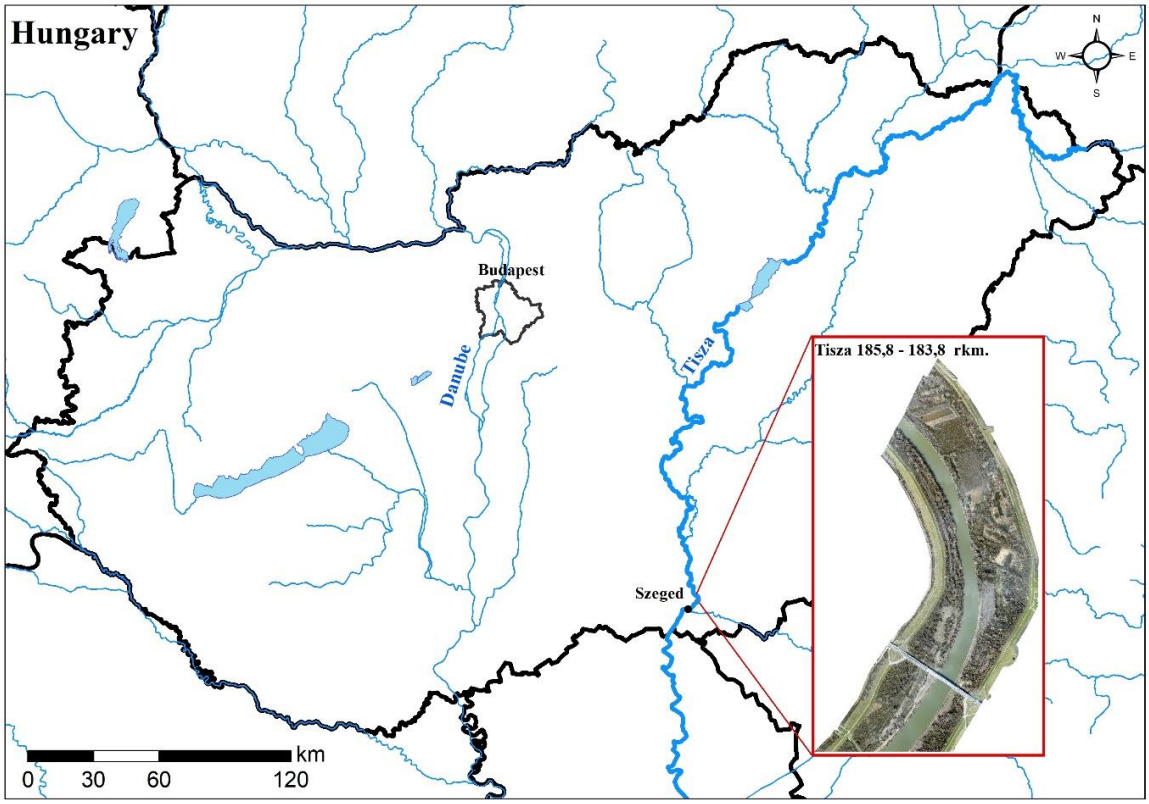

Fig. 1 The study area is located along the Lower Tisza, north of the city of Szeged

results its thickets decrease the flood flow velocity to one third, creating 20-30 cm increase in flood levels at our study area (Kiss et al., 2019a).

DATA SOURCE

The analysed point cloud is based on a LiDAR (full-waveform) survey was made at the early spring of 2015 during low stage of the Tisza River. There were no leaves on the trees, thus the canopy structure of the individual trees is nicely visible on the survey. Simultaneously to the LiDAR survey an ortho-photo was also made with 10 cm resolution. The study area is represented by 22.5 million reflected points on the airborne LiDAR survey, which are stored in eight .las files. The lowest points represent the bank of the river at the actual water-surface (at 75.1 m asl), while the highest indicate the topmost points of the 30-35 m high trees. The water surfaces (e.g. open water, wetlands) do not reflect the LiDAR beams, therefore, these areas were excluded from the analysis. After this exclusion the point-density of the study area is 9 point/m^2, thus it could be considered as suitable for the analysis, as according to Laes et al. (2008) the analysis of forests requires at least 4 point/m^2 point-density. Only 3% of the study area does not fulfil this requirement, thus the quality of the survey was good for further analysis. The digital elevation model (DEM) made of the LiDAR data has 0.5 m resolution.

METHODS

For the analysis the Fusion 3.8 and ArcMap10.6.1 software were used, while the algorithm of the vegetation classification (decision tree) was written in Python using the scikit-learn (0.22.1) library (see Pedregosa et al., 2011). The decision tree were generated with *DecesionTreeClassifier* class-based on the Gini impurity (Grabmeier and Lambe 2007) which is in the *sklearn.tree* module. To find the most ideal and relevant parameters for the decision tree algorithm, the *GridsearchCV* class was applied using K-fold cross validation method. The detailed description of the decision tree can be found in the Results 5.1. chapter. Based on the results of the decision tree the vegetation types of all pixels in the study area were determined automatically, and finally the results were checked and validated based on a field survey.

Data preparation

As the DEM was in a *.flt* format, it had to be converted to *.dtm*, thus it could be used in the Fusion software for further calculations. As a first step the DEM was exported in ArcMap software to .ascii format using the *RastertoAscii* tool, then in Fusion it was converted to *.dtm* applying the *ASCII2DTM* tool.

In the next step the quality assessment of the LiDAR point-clouds stored in *.las* files had to be done, analysing the extreme values and the number of reflections. The quality check was made in Fusion software applying the *Catalog* tool. As a result, the software made a quality report of each .las file, including the number of points, their minimum and maximum heights, and the point-density (point/voxel). Using the *FilterData* tool the extreme values were deleted from the file.

Determine the spatial resolution

To select the most suitable spatial resolution is crucial point of the research. If it is too high, the point cloud will be over dissected, and the typical parameters of a given vegetation type could not be distinguished. However, if the spatial resolution is too low, the spatial differences could disappear and there will be a greater chance to have mixed classes. Laes et al. (2008) suggested, that the spatial resolution should be fitted to the mean canopy width. Therefore, in the study area 15x15 m spatial resolution was selected, thus the point-cloud was split into voxels with 15x15 m cell-size, and the height of the voxels was determined by the highest point of the vegetation.

Calculation of statistical parameters (metrics)

In the following step the statistical parameters of the LiDAR point-clouds representing the vegetation were calculated applying 15 m resolution for the voxels. The calculation was made by Fusion software using its *GridMetrics* tool. The input data included the filtered and quality-checked point-cloud and the DEM in *.dtm* format. In the programme a *heightbreak* could be set, thus the program could recalculate some statistical parameters of the voxels split into two parts by a given elevation (for example the proportion of reflected points above a given height). In our case we had selected 6.0 m as a height brake, and the voxel parameters for the forests were calculated above this value. This value was selected because (i) the bushes (especially invasives) are never taller than 6 m, and we wanted to classify the vegetation regardless of the rate of invasives in the underwood; and (ii) this is the limit of the overbank flood height. The calculated statistical parameters of each voxel were stored in .csv files. The programme provided 74 parameters for each voxel. Not all these parameters are introduced in this paper (following McGaughey 2018), only those, which were used to classify the vegetation types of the study area applying the decision tree.

The *canopy relief ratio* (CRR) was calculated as the ratio of the difference between the mean and minimum heights and the maximum and minimum height of the points of a given voxel [(mean-min)/(max-min)]. Thus, the greater the difference is between the mean and maximum values, the CRR is lower. This parameter refers to the spreading of the canopy: in case of large and wide canopy the mean and maximum values have relatively small differences. At the study area the old white poplars have huge canopy with 0.2-0.3 CRR values, whereas the young planted black poplars have slender canopy with 0.03-0.04 CRR values. The open surfaces (e.g. short grasslands) have the highest CRR values (0.4-0.5), as their mean and the maximum height values are almost similar.

The standard deviation of the height values of a voxel (*Elev_std*) refers to the diversity of the points, thus to the vertical dissection and density of the canopy: the flat and at a given elevation dense canopy is reflected by almost homogenous point-cloud at a given elevation of

the voxel, thus it is characterised by low standard deviation. For example, in the study area the open surfaces, where almost every point represents the land or the grass, the standard deviation is 0.03 m, while in case of the riparian willow stands the standard deviation is higher by two orders (3-4 m). The greatest standard deviation (8-10 m) was measured at the *Populus alba* stands which has variegated canopy.

The 99th percentile value for a voxel (Elev_P99) refers to that elevation, where the height values reach the 99% from the ground, thus it refers almost to the maximum height of the voxel. For its calculation the height values of the point-cloud of a voxel are ordered, and that value is selected which represents the 99% of the dataset. For example, on the study area this value is much lower for the willow stands (~15 m) than for the white poplar trees (~25 m).

The 95th percentile value for a voxel (Elev_P95) refers to that elevation, where the height values reach the 95% from the ground. It is calculated similarly as the previous parameter (Elev_P99). This parameter is useful during the calculations, because it is a good indication of the maximum height of the vegetation, however it does not contain points with survey errors.

Skewness of the heights for the points in the voxel (Elev_skewness) provides some indication of how asymmetric the distribution of the values is. In case of symmetric (standard normal) distribution the skewness is zero. The skewness is influenced by abundant and extreme values: if they are at low values, then the skewness will be negative, while in case of higher values it will be positive. In the study area high (4-6) and positive skewness characterises the young poplar plantations, and the lonely and slim trees: in their case the canopy is not perfectly closed, therefore high proportions of the reflected points are from the ground or from the top of the canopy. In case of grasslands the skewness is low (0.1-0.4), as no extreme values are present, the reflected points originate from almost a flat surface.

Selection of training-plots and definitions of vegetation types

Before the training plots were selected, the main vegetation types had to be determined at the study area. The following categories were identified based on our preliminary field survey and the Forestry WebMap of Hungary (https://erdoterkep.nebih.gov.hu) (1): open surface, *Amorpha* thicket, young poplar plantation, poplar plantation, riparian willow forest, and riparian poplar forest with *Populus alba*.

Based on our field-survey and the available ortho-photo we had selected 15x15 cell-sized pixels with homogenous vegetation as training plots. They were selected for each vegetation types, at least 40-50 cells per type. During the selection of training plots we aimed to select homogenous pixels, thus the pixel should not be affected by side effects of other vegetation types. The selection of the training plots was supported by the ortho-photo providing idea on the character of the cell, while the point-cloud of the voxel gave idea about the height conditions of the vegetation and the shape of the canopy (Fig. 2).

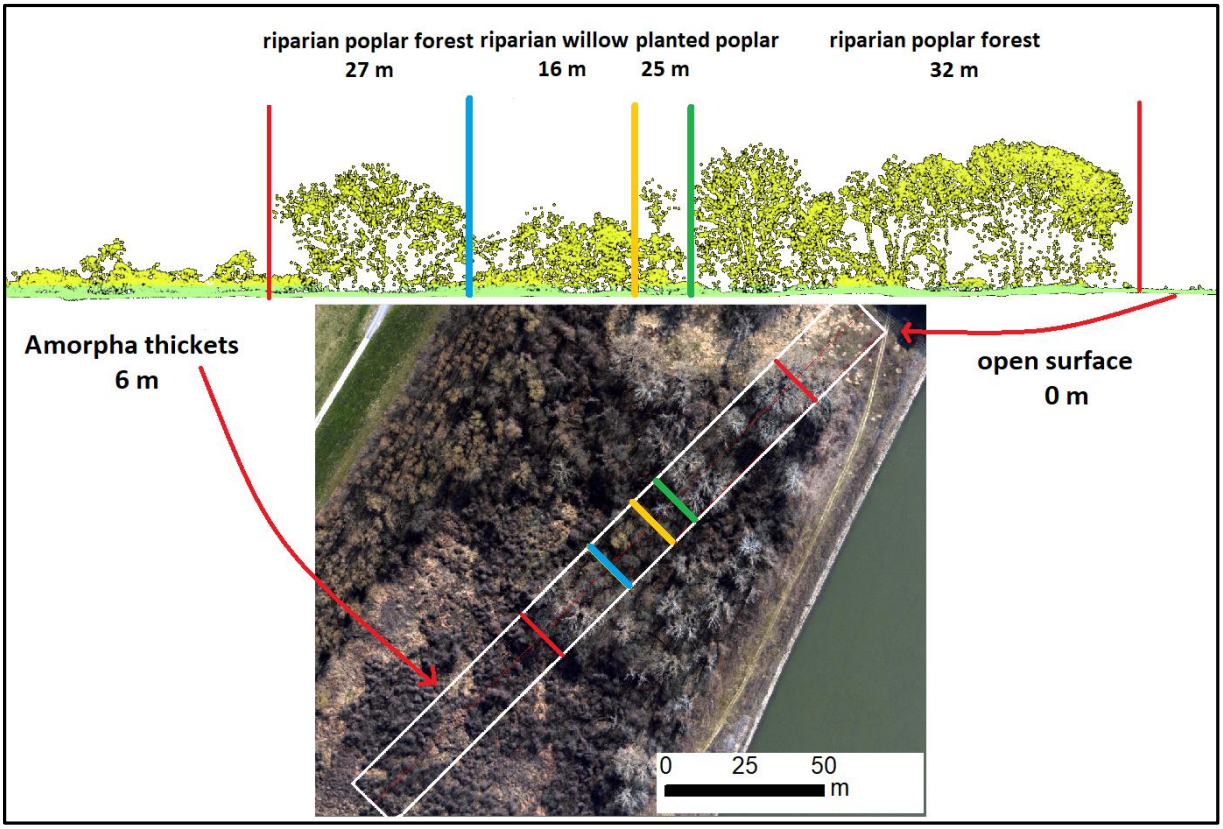

Fig. 2 Various vegetation types represented by the vertical view of the LiDAR point-cloud, and its appearance on the ortho-photo

Grasslands and open surfaces are mostly located on the artificial levees. The LiDAR survey was made in early spring when the grass was very short and dead, therefore the reflected points actually represent the land surface, therefore here the height of the vegetation is almost 0 m. The dense *Amorpha thickets* have height points at 1-6 m, the top of the canopy is almost flat, and great proportion of the LiDAR point-cloud is at the upper part of the canopy. The *riparian willow forests* are high (16-20 m), and on the ortho-photo the almost bunding willow-branches have brownish-orange colour. The *poplar plantations* are planted in rows, thus the skewness of their point-cloud is high, as lots of points are reflected from the ground and also from the top of the canopy. On the LiDAR point-clouds their canopy structure is quite specific, as their branches are thin, most of the points are around the stem, and on the ortho-photo the top of these plantations is quite homogenous. Within the poplar plantations we distinguished the group of *young poplar plantations*. Their trees are shorter (>5.6 m), and more points are reflected from the ground due to their undeveloped canopy. This forest type also includes lonely trees with sparse bushes. The *riparian poplar forest* patches are characterised by tall and easily distinguishable *Populus alba* trees. The white poplar has greyish-white branches on the ortho-photo, and it has very special canopy structure both on the ortho-photo and the LiDAR point-cloud (Fig. 2).

Creating the decision tree

The statistical parameters of the training plots were saved in a *.csv* file. In this file the names of vegetation types and all of the statistical parameters were given. The next step was the parametrisation of the decision tree algorithm. The decision tree classifies the elements – in our case the attributes of the voxels – based on the series of classification steps, aiming to have the most homogenous classes. The selection of attributes and the thresholds between the classes are based on a calculation algorithm. In our case the threshold values were based on the Gini impurity, because it could be run faster than the entropy based calculations, besides, there is no qualitative difference between the accuracy of these classifications. The Gini impurity refers to the probability of the classification of an element to a wrong class (Grabmeier and Lambe, 2007). If its value is zero, it means that the given selection criteria perfectly divided a class from the main population, while 1.0 refers to a totally diverse class.

The setting of the parameters of the decision tree was made automatically applying the *GridsearchCV* module, considering the (i) maximum depth of the decision tree referring to the number of decision levels; (ii) the minimum element number of the leaves; and (iii) the minimum element number for split. To precisely determine the above mentioned settings of the decision tree, each setting was set to an interval (e.g. decision tree depth: 1-10; minimum element number of leaves: 2-10; minimum split: 2-20), and finally best setting combination was selected, which resulted in the most precise decision tree.

To check the accuracy of the decision tree algorithm, a cross-validation method was applied, which is very common at automatized learning technologies. We had selected the method of *10-fold cross-validation*. As a first step the dataset (training plot voxels) were divided to 10 groups, and one of them was selected for validation by the algorithm. The cross-validation lasts for 10 iterations, until every group will be used exactly once as a training set (Bengio and Grandvalet 2004). To estimate the accuracy, the average of 10 results were needed. The advantage of this kind of cross-validation is that each point (voxel) of the dataset will be used for automatic learning and for validation too, however its disadvantage is that it is a quite long process, as in our case the automatic learning was repeated 10 times.

The accuracy of the final classification based on the decision tree is expressed in percentage, referring to the rate of well-classified elements, though this value does not refer to the efficiency of the classification. During the application of the decision tree two kind of methodological mistakes could be made. In the first case the classification process and the applied thresholds do not determine the classes homogenously, thus the decision tree will have low accuracy and it is underfit, however if there are lots of data and several parameters, this case is quite rare. Much more often the decision tree will be overfit, thus the selection criteria will be too specific and valid just for some elements of the dataset. In this case the accuracy is very high (>95%), however the decision tree could not be applied on other datasets (Schaffer 1993).

The aim of the automatized parameter selection and of the *10-fold cross-validation* was to find the most suitable decision tree, which eliminates the errors of the overfitting. During the cross-validation runs it became obvious, that the decision tree is greatly influenced by its depth. Our results suggest, that if the depth of the decision tree is greater than 4, the accuracy won't be considerable better, however the risk of overfitting increases (Fig. 3). Therefore, the depth of decision tree was determined to be 4. The minimum element number and minimum split number was determined to be 2. In this way the accuracy of the created decision tree based on the *10-fold cross-validation* was 92%.

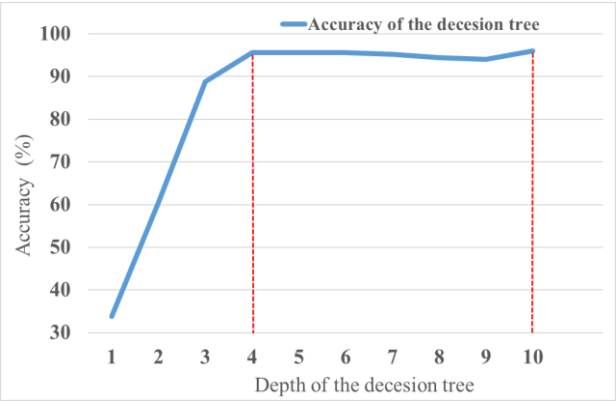

Fig. 3 Relationship between the depth of the decision tree and the accuracy of the classification

Fitting the decision tree to the entire study area and its validation

The decision tree was created based on training plots. In the next step, based on the decision tree all the 11656 pixels of the study area were analysed and their vegetation types were determined. The result and precision of the automatized classification were determined by field-survey. For the field validation aerial photos made by DJI Phantom III Pro drone were taken on 72 cells. The survey was made at 30-60 m height from orthogonal position (applying 90º camera axis). On these low-aerial photos the vegetation types of the cells were identified by expert judgement. The geo-coordinates of the photos were extracted; thus the identified vegetation types could be compared to the automatized classification of the same cell. The results of the comparison were evaluated applying a confusion matrix. The titles of columns and rows in the matrix refer to the vegetation categories. In the main diagonal line, the numbers refer to the proportion of precisely classified cells, while the other cells refer to the proportion of false vegetation classes.

RESULTS

The decision tree

The decision tree automatically selected the parameters of the voxels (see 4.5. chapter) for the identification of vegetation types (see 4.3. chapter). First, the *young poplar plantations* were selected based on *CRR≤0.039* criteria. This parameter well sunders the young and short poplar trees with undeveloped canopy from the higher trees with more complex canopy and from the grassland/open surfaces. The next step followed the false (no-)branch of the decision tree (*CRR≥0.039*). Here, based on the standard deviation of the voxel's point-cloud the *Amorpha* thickets and the grassland/open surfaces were identified. The two vegetation classes could be divided based on their height (*Elev_P99*). To identify the *grassland/open surfaces* the voxels had to be fulfil the following criteria: *CRR≥0.039* and *Elev_std≤1.783,* and *Elev_P99≤2.119*. The *Amorpha thickets* were identified by *CRR≥0.039* and *Elev_std≤1.783,* and *Elev_P99>2.119*. The Gini impurity (0.0) of these three categories reflects that they were identified with the greatest accuracy (Fig. 4.).

On the false (no-)branch of the standard deviation (*Elev_std≥1.783*) of the decision tree the older poplar plantations, the riparian poplar forests and riparian willow forests remained. On the floodplain the riparian poplar forests are characterised by tall *Populus alba* trees, thus they could be selected based on their height conditions (*Elev_P95>17.987*). However, this selection criterion is not totally clear, as some voxels with tall planted poplars also fall into this class. The natural and planted poplar forest could be divided based on the CRR parameter: the poplar plantations have less complex canopy than of the natural poplars, therefore the plantations have smaller CRR values, thus their selection criterion is *CRR≤0.103*. The *riparian poplar forests* were selected following *Elev_std>1.783* and *Elev_P95>17.987* and *CRR>0.103*. Some of the poplar plantation's cells (with older and higher trees) also fall into this class, only their CRR values differed (*Elev_std≥1.783* and *Elev_P95>17.987* and *0.039<CRR≤0.103*). Based on the tests the tall (≤18 m) and *old poplar plantations* were clearly separated (Gini impurity = 0.0) from the riparian poplar forests with *Populus alba* (Fig. 4.).

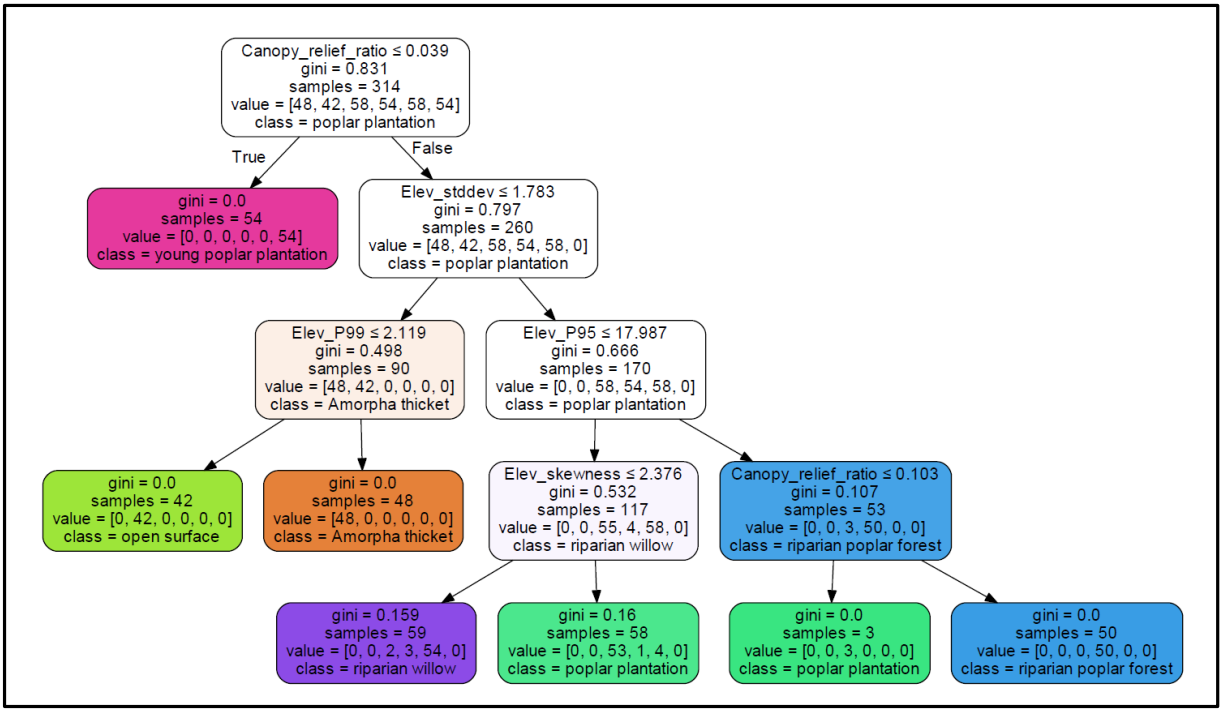

Fig. 4 Decision tree built up based on training plots of the study area. This decision tree was applied to classify the vegetation types on the entire study area

On the true branch of the *Elev≤17.987* criteria the riparian willow forests and the less-old and shorter poplar plantations remained. The planted poplars are slim with column-like canopy, and the reflected points from the ground between the tree-rows result in asymmetric distribution of the height points. As the skewness reflects well the asymmetric distribution of the values, the criteria of Elev_skewness> 2.376 were applied to identify the two vegetation classes. The identification of the ***poplar plantations*** followed the criteria of *CRR>0.039* and *Elev_std>1.783* and *Elev_P95≤17.987* and *Elev_skewness>2.376*. The parameters of the ***riparian willow forests*** are almost the same as of poplar plantations, only their skewness is different: *CRR>0.039* and *Elev_std>1.783* and *Elev_P95≤17.987* and *Elev_skewness≤2.376*. However, at the study site these vegetation types could be mixed even within a 15x15 m cell, which influences the accuracy of the classification, however, the identification of the classes was still effective (Gini impurity < 0.16).

Vegetation types of the study area based on the automatized classification

The decision tree created on training plots were applied on the entire, 3 km^2-large study area. The vegetation types of 11656 voxels were identified, and the land-cover map of the area was created (Fig. 5.).

In the study area the most abundant land-cover category (Fig. 6.) is riparian willow forest (30%, 80 ha). Willow patches appear mostly in deeper lying areas, like in front of the artificial levees and on the edges of clay-pits. The grasslands/open surfaces (24%, 63 ha) mainly cover the artificial levees, but in this category also some plough-fields are and clay-pits, where the dead herbaceous vegetation covers the surface like a mat. Planted poplar forests (15%, 40 ha) appear in great units. From the point-of view of flood conductivity the proportion of *Amorpha* thickets (10%, 25 ha) is crucial. They usually appear along the edges of other vegetation types and on the fallow lands. In the study area the smallest area is occupied by young poplar plantations (9%, 23 ha), but this category contains those patches as well, where young trees are mixed with bushes, but they do not create dense stands.

Validation of the results

The accuracy of the automatized classification was determined by comparing its results to field-surveys on 72 randomly selected cells. The results are summarised in a confusion matrix (Table 1.).

The accuracy of the automatized vegetation classification based on the decision tree algorithm was 83%. The open surfaces were classified with the lowest accuracy (75%). Some open surfaces were classified by

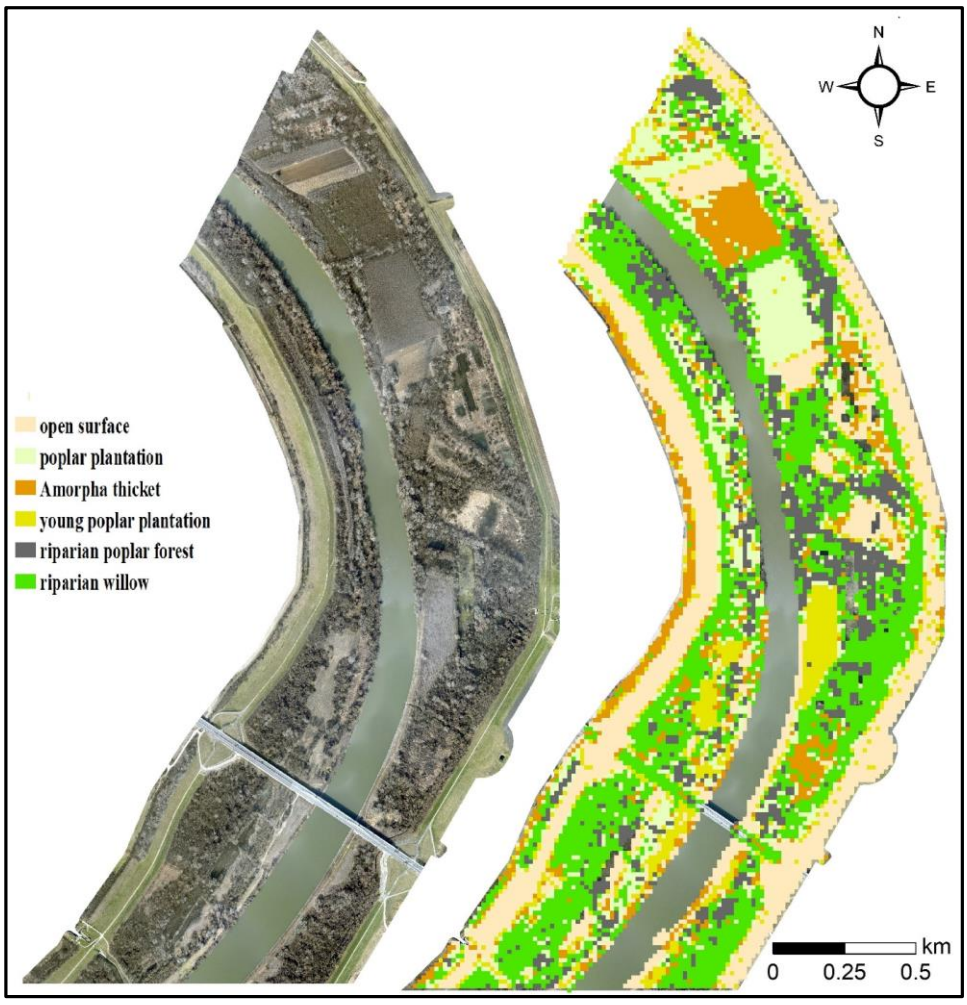

Fig. 5 Ortho-photo of the study area (A) and the automatized vegetation type map of the same area based on the classes of the decision tree (B)

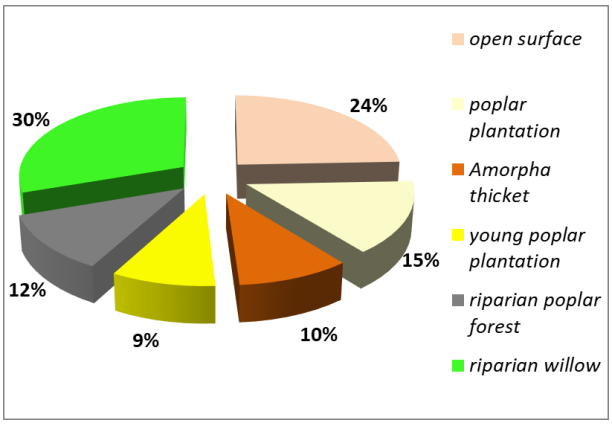

Fig. 6 Proportion of the different vegetation types in the study area

the algorithm as *Amorpha* ticket (8% of the cases) or as young poplar plantation (17%). Most of these mis-identifications occurred at the boundary between the grass-covered artificial levee and the arboreous vegetation, where many sprouts and low branches stretches over the grassland. The identification accuracy of the riparian willow forests is 85 %, in reality, the mis-identified patches belong to poplar plantations (5%), *Amorpha* thickets (5%) or riparian poplar forests (5%). This error has multiple sources: (i) between the LiDAR and the field surveys some forests were cleared and the clearances were colonised by *Amorpha*; (ii) these vegetation types could be mixed on a 15x15 m-sized cell; (iii) depending on the age of the forest patch the various vegetation types could have similar height and even similar canopy size. The *Amorpha* thickets were the most accurately (92%) classified by the algorithm. Only 8% of them were mis-classified, and got to the class of riparian willow forest. However, this error is not considerable, as it was detected on cells where the willow forest was highly invaded by *Amorpha*. The identification accuracy of riparian poplar forest was 83%, as some of their patches were identified as planted poplar (8%) and as young poplar plantation (8%). These mis-identifications were in

Table 1 Confusion matrix summarizing the validation results. Green colour indicates % of well -classified voxels (%), red colours refer to percent of unwell-classified voxels

		Based on decision tree					
		open surface	*riparian willow*	*Amorpha thicket*	*riparian poplar forest*	*young poplar plantation*	*poplar plantation*
Based on field work	*open surface*	0.75	0.00	0.08	0.00	0.17	0.00
	riparian willow	0.00	0.84	0.05	0.05	0.00	0.05
	Amorpha thicket	0.00	0.08	0.92	0.00	0.00	0.00
	riparian poplar forest	0.00	0.00	0.00	0.83	0.08	0.08
	young poplar plantation	0.00	0.00	0.17	0.00	0.83	0.00
	poplar plantation	0.00	0.00	0.00	0.18	0.00	0.82

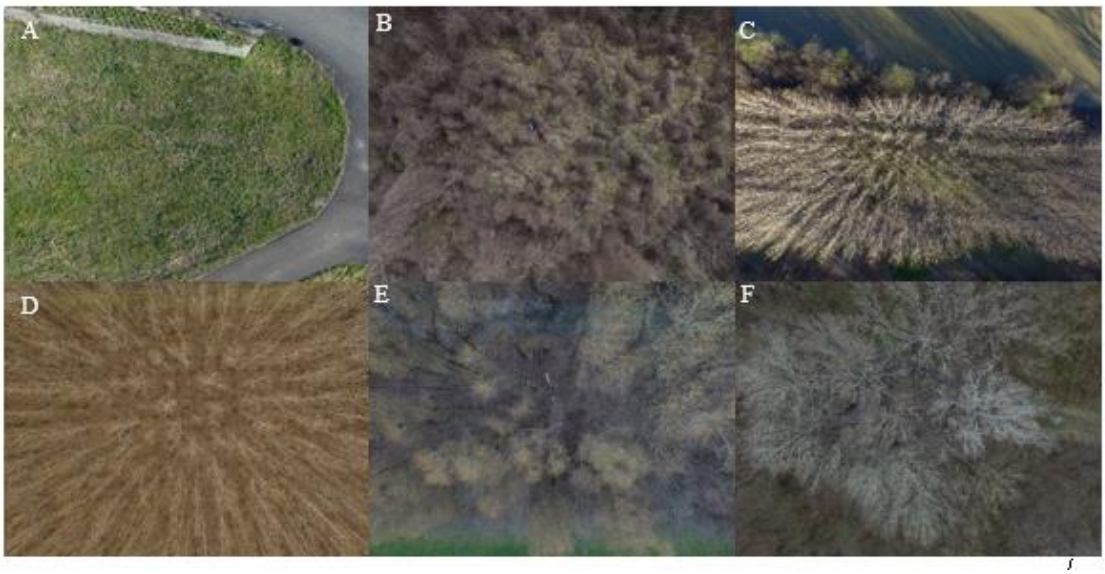

Fig. 7 Vegetation types based on drone photos. A: open surface, B: Amorpha thicket, C: young poplar plantation, D: poplar plantation, E: riparian willow forest, and F: riparian poplar forest with Populus alba

those cells, where the natural poplars were mixed to planted poplars. The young poplar plantations were identified by 83% accuracy too, as some of the cells were identified by the algorithm as *Amorpha* thickets. It could be explained by the fact, that very young plantations have similar height and density conditions as the thickets, besides, if the plantation is not managed properly, *Amorpha* could invade them very quickly. The accuracy of poplar plantations was 82%, as the greatest errors occurred when the algorithm classified them as riparian poplar forests. It could be explained by the similar height of old plantations and younger *Populus alba* trees.

CONCLUSION

The applied automatized machine learning-based classification is suitable to identify various vegetation types based on airborne LiDAR survey data. Not only land-cover types (e.g. forest), but various vegetation and forest types could be identified using the method, which has acceptable accuracy. For example, on the studied floodplain area the accuracy of the classification was 83% (based on 72 observation).

The data acquisition of LiDAR surveys combined to automatized machine learning enables us to precisely, effectively and quickly map the vegetation even on large, remote or impenetrable areas. In case of repeated surveys, the algorithm easily could be trained to the new dataset, thus the temporal changes in vegetation could be quickly and automatically detected, which is a great advantage for both researchers, stakeholders and decision makers.

The application of the resulted vegetation type map is quite wide. For example, it could be used by hydrologists, as up-to-date vegetation maps are needed during the planning and maintenance of flood-conductivity zones, or during the modelling of floods, when up-to-date data are needed on vegetation roughness to determine the Manning coefficient. In forestry these LiDAR-based vegetation maps could be used too, as the statistical parameters of forests could be calculated, and forest clearance plans could be supported.

Acknowledgement

The LiDAR data are owned and provided by the ATIVIZIG (Lower Tisza Hydrological Directorate). The research was supported by the Hungarian Research Fund (OTKA 119193).

References

Abernethy, B., Rutherfurd, I.D. 1998. Where along a river's length will vegetation most effectively stabilise stream banks? *Geomorphology* 23, 55–75. DOI: 10.1016/s0169-555x(97)00089-5

Bengio, Y., Grandvalet, Y. 2004. No Unbiased Estimator of the Variance of K-Fold Cross-Validation. *Journal of Machine Learning Research* 5, 1089–1105.

Brooks, G.R. 2005. Overbank deposition along the concave side of the Red River meanders, Manitoba, and its geomorphic significance. *Earth Surface Processes and Landforms* 30, 1617–1632. DOI: 10.1002/esp.1219

Corenblit, D., Tabacchi, E., Steiger, J., Gurnell, A.M. 2007. Reciprocal interactions and adjustments between fluvial landforms and vegetation dynamics in river corridors: A review of complementary approaches. *Earth-Science Reviews* 84, 56–86. DOI: 10.1016/j.earscirev.2007.05.004

Geerling, G.W., Kater, E., van den Brink, C., Baptist, M.J., Regas, A.M.J., Smits, A.J.M. 2008. Nature rehabilitation by floodplain excavation: The hydraulic effect of 16 years of sedimentation and vegetation succession along the Waal River, NL. *Geomorphology* 99, 317–328. DOI: 10.1016/j.geomorph.2007.11.011

Grabmeier, J. L., Lambe, L. A. 2007. Decision trees for binary classification variables grow equally with the Gini impurity measure and Pearson's chi-square test. *International Journal of Business Intelligence and Data Mining* 2(2), 213. DOI:10.1504/ijbidm.2007.013938

Heurich, M., Thoma, F. 2008. Estimation of forestry stand parameters using laser scanning data in temperate, structurally rich natural European beech (*Fagus sylvatica*) and Norway spruce (*Picea abies*) forests. *Forestry* 81, 645–661. DOI: 10.1093/forestry/cpn038

Hudak, A., Crookston, N., Evans, J., Hall, D., Falkowski, M. 2008. Nearest neighbour imputation of species-level, plot-scale forest structure attributes from lidar data. *Remote Sensing of Environment* 112, 2232–2245. DOI: 10.1016/j.rse.2007.10.009

Jalonen, J., Järvelä, J., Virtanen, J.P., Vaaja, M., Kurkela, M., Hyyppä, H. 2015. Determining Characteristic Vegetation Areas by Terrestrial Laser Scanning for Floodplain Flow Modelling. *Water* 7(2), 420–437. DOI: 10.3390/w7020420

Jung, S.E., Kwak, S.A., Park, T., Lee, W.K, Yoo, S. 2011. Estimating Crown Variables of Individual Trees Using Airborne and Terrestrial Laser Scanners. *Remote Sensing* 3, 2346–2363. DOI: 10.3390/rs3112346

Kiss, T., Fiala, K., Sipos, Gy., Szatmári, G. 2019b. Long-term hydrological changes after various river regulation measures: are we responsible for flow extremes? *Hydrology Research* 50(2), 417–430. DOI: 10.2166/nh.2019.095

Kiss, T., Nagy, J., Fehérváry, I., Vaszkó, Cs. 2019a. (Mis)management of floodplain vegetation: The effect of invasive species on vegetation roughness and flood levels. *Science of the Total Environment* 686, 931–945. DOI: 10.1016/j.scitotenv.2019.06.006

Kiss, T., Sándor, A. 2009. Land-use changes and their effect on floodplain aggradation along the Middle-Tisza River, Hungary. *AGD Landscape and Environment* 3(1), 1–10.

Kovács, S., Váriné Szöllősi, I. 2003. A Vásárhelyi Terv Továbbfejlesztését megalapozó hidrológiai és hullámtér hidraulikai vizsgálatok eredményei a Közép-Tiszán. MHT XXI. 2/12. 1–11.

Laes, D., Reutebuch, S., McGaughey, B., Maus, P., Mellin, T., Wilcox, C., Anhold, J., Finco, M., Brewer, K. 2008. Practical lidar acquisition considerations for forestry applications. RSAC-0111-BRIEF1. Salt Lake City, UT: U.S. Department of Agriculture, Forest Service, Remote Sensing Applications Center. 32 p.

Manners, R., Schmidt, J., Wheaton, M. J. 2013. Multiscalar model for the determination of spatially explicit riparian vegetation roughness. *Journal of Geophysical Research: Earth Surface* 118, 65–83. DOI: 10.1029/2011jf002188

McGaughey, R. 2018. Users Manual of Fusion/LDV: Software for LIDAR Data Analysis and Visualization. United States Department of Agriculture, Forest Service, Pacific Northwest Research Station.

Naesset, E., Gobakken, T., Holmgren, J., Hyyppa, J., Maltamo, M., Nilsson, M., Olsson, H., Persson, A., Doderman, U. 2004. Laser scanning of forest resources: the Nordic experience. Scandinavian. *Journal of Forest Research* 19, 482–499. DOI: 10.1080/02827580410019553

Nagy, J., Kiss, T., Fiala, K. 2018. Hullámtér-feltöltődés vizsgálata az Alsó-Tisza mentén. II. Folyóhátak (parti hátak) feltöltődését befolyásoló tényezők. *Hidrológiai Közlöny* 98(1), 33–40.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, É. 2011. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* 12(85), 2825–2830. DOI: 10.3389/fninf.2014.00014

Rátky, I., Farkas, P. 2003. A növényzet hatása a hullámtér vízszállító képességére. *Vízügyi Közl.* 85(2), 246–264.

Schaffer, C. 1993. Overfitting Avoidance as Bias. *Machine Learning* 10, 153–178. DOI: 10.1007/bf00993504

Steiger, J, Gurnell, A.M., Ergenzinger, P., Snelder, D.D. 2001. Sedimentation in the riparian zone of an incising river. *Earth Surf. Process. Landforms* 26, 91–108. DOI: 10.1002/1096-9837(200101)26:1<91::aid-esp164>3.0.co;2-u

Vetter, M., Höfle, B., Hollaus, M., Gschöpf, C., Mandlburger, G., Pfeifer, N. 2011. Vertical vegetation structure analysis and hydraulic roughness determination using dense ALS point cloud data–a voxel based approach. *Int. Arch. Photogr. Remote Sens. Spat. Inf. Sci.* 38(5), 200–206. DOI: 10.5194/isprsarchives-xxxviii-5-w12-265-2011

Zellei, L., Sziebert, J. 2003. Árvizi áramlásmérések tapasztalatai a Tiszán. In: Szlávik L. (szerk.): Elemző és módszertani tanulmányok az 1998-2001. évi ár- és belvizekről. Vízügyi Közlemények különszám 4, 133–144.