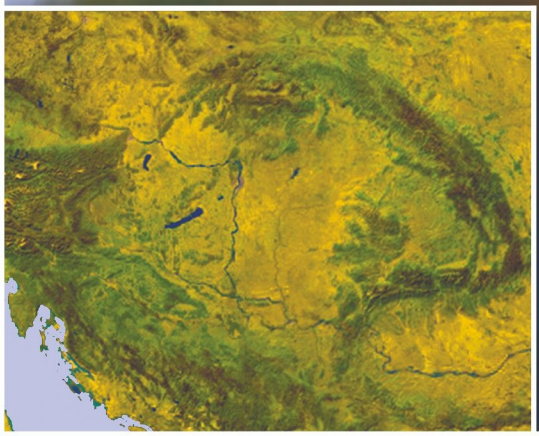
Hungarian Geographical Bulletin



Földrajzi Értesítő

Special issue:
Perspectives of quantitative analysis in modern physical and landscape geography

Edited by
Szilárd Szabó
and László Bertalan

Volume 67 Number 4 2018

MTA
Research Centre
for Astronomy and
Earth Sciences
GEOGRAPHICAL
INSTITUTE

HUNGARIAN GEOGRAPHICAL BULLETIN Volume 67 Number 4 2018

HUNGARIAN GEOGRAPHICAL BULLETIN 67 2018 (4)

CONTENT

Perspectives of quantitative analysis in modern physical and landscape geography

Book review section

DOI: 10.15201/hungeobull.67.4.1 Hungarian Geographical Bulletin **67** 2018 (4) 309–318.

Recent methods, sources and approaches in the study of temporal landscape change at different scales – a review

Anders LUNDBERG[1]

Abstract

Landscape change can be studied at different scales, from local to regional, e.g. from a biodiversity level to the level of land-use systems. Historical sources such as land taxation papers, cadastral maps, agricultural and population censuses are not very much used in physical geography studies but this article explores the potential of historical sources and gives examples of how they can be used. Studies of landscape change requires an approach to the temporal dimension and examples are given on how this can be solved. At the biodiversity level, much attention is presently paid to red-listed species. The article gives an account of criteria used to evaluate red-list status of species and critically reviews the way governmental nature management bodies presently deal with this. Several long-term studies have been carried out recently and huge variations in population size of species from year to year have been detected. As an alternative to red-list status based on population size based on static data from one or a few years, the concept of the *natural variation interval* of a species is introduced. The article demonstrates how this phenomenon can be identified based on a temporal approach.

Keywords: temporal landscape change, time-series, red-list criteria, natural variation interval, historical forest cover measurements

Introduction

Landscape change is more than ever a characteristic feature going on across the world (Heatherington, C. *et al.* 2017; Belén, M. *et al.* 2018). Changes in settlement patterns, land-use and climate change are well known reasons for changes in vegetation and biodiversity (Deák, B. *et al.* 2016, 2018; Kizos, T. *et al.* 2018; Török, P. *et al.* 2018). Loss of biodiversity has become a major concern all over the world and criteria to evaluate biodiversity loss developed by the International Union for the Conservation of Nature (IUCN) have been adopted by numerous countries. In Norway, 4,438 species are red-listed according to the recent up-date of 2015 (Henriksen, S. and Hilmo, O. 2015). Habitats and ecosystems also undergo change, and several nature types have become threatened. At present, 75 nature types are considered threatened in Norway, and 39 nature types are considered near threatened (Lindgaard, A. and Henriksen, S. 2018). In total, nature is under press, both species and ecosystems, even in a country like Norway, often valued for it's beautiful natural landscapes. The aim of the study is to elaborate how temporal landscape change can be studied at different scales: at landscape, population and species levels, emphasizing the temporal dimension.

Approaches to the study of landscape transformation through time

Changing landscapes is one major characteristic of recent environmental change across Europe. Urban sprawl is a dominant process, as well as changes in the settlement patterns due to urbanisation (Couch, C. *et al.* 2007). Agriculture is changing from small-scale to large-

[1] Department of Geography, University of Bergen, Fosswinckelsgt. 6. 5007 Bergen, Norway.
E-mail: anders.lundberg@uib.no

scale, from diversity of crops and livestock to specialisation and production with lots of input of external resources. Natural landscapes are transformed to recreational landscapes, some with new and modern technical installations, such as cable cars to mountain tops. Seascapes are also changing from natural fiords and coasts to fish farming production systems. To study these and similar questions belong to the core of academic geography, and such studies have links to both physical and human geography (Jones, M. 1988; Lundberg, A. 2005a, b; Skjeggedal, T. 2005).

Methods used to study landscape change include aerial photo interpretation, remote sensing, GIS, vegetation mapping and others. Comparison of aerial images of different age has proved highly useful to detect temporal changes in land-use, settlement patterns, roads and other types of infrastructure, forest increase/decrease, development of river meandering etc. (Plieninger, T. 2006; Hill, J. *et al.* 2008; Frondoni, R. *et al.* 2011; Novák, T.J. *et al.* 2014; Szabó, S. *et al.* 2015). Another supplementary source that can be used to analyse temporal landscape change is the use of historical sources, such as cadastral maps, historical land taxation papers, population and agricultural censuses. These historical sources are well known among historians but less so among geographers and ecologists. Some ecologists implement historical data in their studies of habitats and species diversity though (Helm, A. *et al.* 2006; Pitkänen, P.T. *et al.* 2016). They include information about land-use at certain times, extent of forests, meadows, pastures, cultivated fields, livestock and other types of spatial data that can be transformed and used in the analysis of changes in landscape (Török, P. *et al.* 2010). In other words, these historical data can be used in the reconstruction of past landscapes (Yang, Y. *et al.* 2017). Historical data can be used to study when physical landscape attributes appeared, how they developed, changed and disappeared and sometimes also reappeared. The static map can then be transformed to a complex and ever-changing mosaic of spatial phenomena that rise, meet, connect, separate and disappear (Hägerstrand, T. 1995; Lundberg, A. 2008).

Lundberg, A. (2005b) used historical sources to study forest development in Western Norway during the last 400 years. Forests invaded Western Norway in the late Holocene period (c. 12,000 B. P.) and different trees succeeded as climate became milder after the last Ice Age. The first tree to colonise Northern Europe after the last Ice Age was the birch, and thermophilous decedious trees were late invaders, such as elm, oak, lyme, and ash. The cultivation of land started in the Younger Stone Age period and production of cereals became common in the Bronze Age. In Medieval times the settlement pattern was consolidated but forests were still extensive and covered most parts of the land below the alpine tree limit. Forests were harvested and timber was used for building houses, boats and a number of things but open land was mostly minor patches in a dominant matrix of forest.

This situation changed rapidly due to the introduction of water-driven saws in the 14th century. Previously, the axe was the dominant tool in forestry but this was time-consuming. Water-driven saws were more efficient and effectively split logs into building planks. During the 14th and 15th century forestry became a major industry in Western Norway and timber and planks were even exported to Holland and Scotland. This trade is called the Scot forest trade.

At the same time cities developed along the coast, such as Bergen and Stavanger. Most houses were built of timber, a resource available in regional forests. However, sometimes Bergen and other cities were hit by devastative fires and major parts of the cities burnt and had to be rebuilt. Timber was still available regionally but Bergen and other cities burnt again and again (Bækken, I. *et al.* 2002). As a result of Bergen fires and the Scot forest trade Western Norway was deforested during the 15th century. Population increase and extensive livestock grazing along the coast, in fiord valleys and mountain pastures prevented the reestablisment of forest.

Land-use was intensified to a maximum and this lasted until the end of the 17th century when the emigration to America had it´s peak.

Only then forests started to re-develop, and this is why many forests in Western Norway can be dated back to the 1880s or so. Forests older than this can be found but they are scattered.

Lundberg, A. (2005a) used historical sources found in statal archives to reconstruct forest development in a part of Western Norway. Sources used were land taxation papers from 1665, 1723, 1867 and 1890, cadastral maps from 1825 and 1881, as well as areal photos (from 1956 and later), interviews and finally field registrations. The result is presented in *Figure 1.*

The result varies quite a bit from the general conclusions drawn by botanists prior to this research. Due to the luxourious vegetation the forest was interpreted as a primeval forest but thanks to the analysis of historical sources it turned out that the forest was a young, first generation forest (Lundberg, A. 2005a,b, 2010) . As can be imagined, this will very much impact the management of such a forest. This is because an old primeval forest would be in a mature state of development, while in a young, first generation forest more dynamics would be expected.

Core concepts in academic geography is the combination of time and space/place (Holt-Jensen, A. 2018). Historical geography has very much been involved in the study of landscape change and how different landscape attributes arise and develop. Any landscape is many-facetted and always include human and natural phenomena and combinations of those. Specialisation with focus on certain phenomena has been a de-

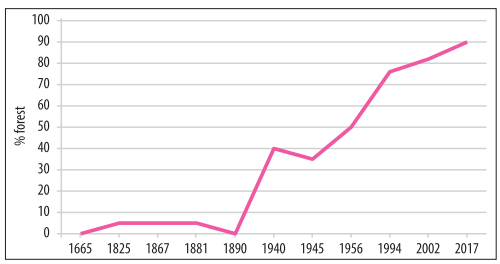


Fig. 1. Forested areas in percent of land area in a part of Western Norway during 1665 to present. *Source*: Based on Lundberg, A. (2005a), but revised and updated.

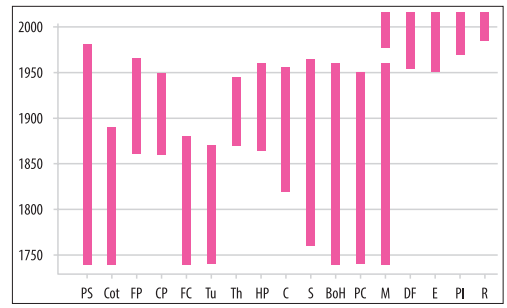


Fig. 2. Time-scale illustrating major factors, activities and land-use that characterised a landscape during certain periods of time. The approach is diverse more than limited (revised from Lundberg, A. 2002). – PS = permanent settlement; Cot = cottagers; FP = farm production; CP = cereal production; FC = fields cultivated with spades; Tu = tust; Th = threshing; HP = horse and plough; C = cattle; S = sheep; BoH = burning of heathlands; PC = peat cutting; M = mowing; DF = deciduous forest; E = electricity; Pl = plantations; R = recreation

veloping trend in recent landscape research. An attemt to reach beyond specialisation and achieve a wide scope on landscape and landscape change is presented in *Figure 2*.

Data supporting *Figure 2* has been collected using population and agricultural censuses, interviews of local informants, and aerial photo interpretation. A number of factors, activities and land-use that have influenced the landscape is presented along the x-axis, such as permanent settlement, cattle keeping, burning of heathlands, peat cutting, expansion of forests etc. The year different facors were initiated, continued and eventually came to an end is shown along the y-axis. This way of illustrating landscape development adopts a wide scope dealing with major elements that characterise a landscape during a given period of time.

Methods to monitor biodiversity dynamics

Criteria used in many national red-lists include population reduction, geographical range, small population and continuing decline in population, and very small or geographically very restricted population (Henriksen, S. and Hilmo, O. 2015).

One essential criteria for evaluation of the red-list categories is the trend during the last ten years. For many species we have information about distributions but one challenge is lack of data on temporal change in numbers and density. Consequently, red-listing has to be based on best knowledge and judgements, more than systematic empirical data on temporal population trends. Some taxonomic groups are known to have extensive changes in abundance from year to year, e.g. orchids. This might also be true for other taxonomic groups but empirical evidence on this is limited. This is a major challenge for red-listing of species. If information is based on the situation in an unfavourable year, a species might be given status as threatened; if information is based on the situation in a good year, the status might be considered near threatened. The actual status for that species may not have changed, but the evaluation of the situation might be very different depending on the situation in one or a few incidental year(s).

An example is the status given to the orchid *Coeloglossum viride* ssp. *islandica* in the Norwegian red-list of 2006. It was concluded that the taxon most likely was extinct in Norway. As demonstrated by the long-term study by Lundberg, A. (2015) this was luckily not the case. This was not because seeds had been dormant for a period. It was simply because the taxon was less known among botanists and also because the awerness of that subspecies among botanists was low.

All nature conservation areas have some purpose and they are often explicitly mentioned in management plans of protected areas, such as nature reserves, landscape protection areas and national parks. In modern nature conservation, the formulation of population measures for taxonomic groups present in the protected area has also become usual. This has been done for bird species and other groups but again one challenge is that lack of data on changes in the temporal abundance of the relevant groups.

The distinction between extinction and dormancy

To investigate the range of annual variations among orchid populations a study of *Dactylorhiza purpurella* was started in Norway in 2012 (Lundberg, A. and Frøland, T. 2016). All known Norwegian populations were visited and a monitoring program was started. In total, 48 populations are known and 30 of these are intact, 15 have been lost and three populations have unknown status. The species is considered to be critically threatened (CR) in Norway. Some of the populations are small and some are numerous with several hundreds to a few thousand plants in a year. In Norway, *D. purpurella* is found in sand dune meadows and dune slacks, in salt marshes, wetlands and other wet or moist habitats close to the sea. A few populations are also found in abandonded industrial sites, probably due to the open site with mineralic, calcareous soils. *D. purpurella* is a North Sea species known from the UK, Denmark, Norway and the Faroe Islands *(Figure 3).* The Norwegian distribution is on the western coast, in the south-west and the north-west part of Western Norway. Soils are often calcareous and the sites are usually rich in species.

During the monitoring period 2012 till present all populations showed considerable variation in numbers from year to year. Huge annual variations in abundance is the normal pattern, stability in the number of plants in any population has never been found (Lundberg, A. and Frøland, T. 2018). The distinct variation can not be explained by technical encroachments, because no physical change to the soil or habitat has been found. There is no one-way trend in decrease or increase so the fluctuations can not be due to climate change. The explanation we found is that lots of precipitation during late autum and the first part of spring is favorable for germination, particularily because the plants heavily depend on mycorrhiza infection in this part of the year. Intact leaves not damaged by drought the season before might also be necessary for production and storage of

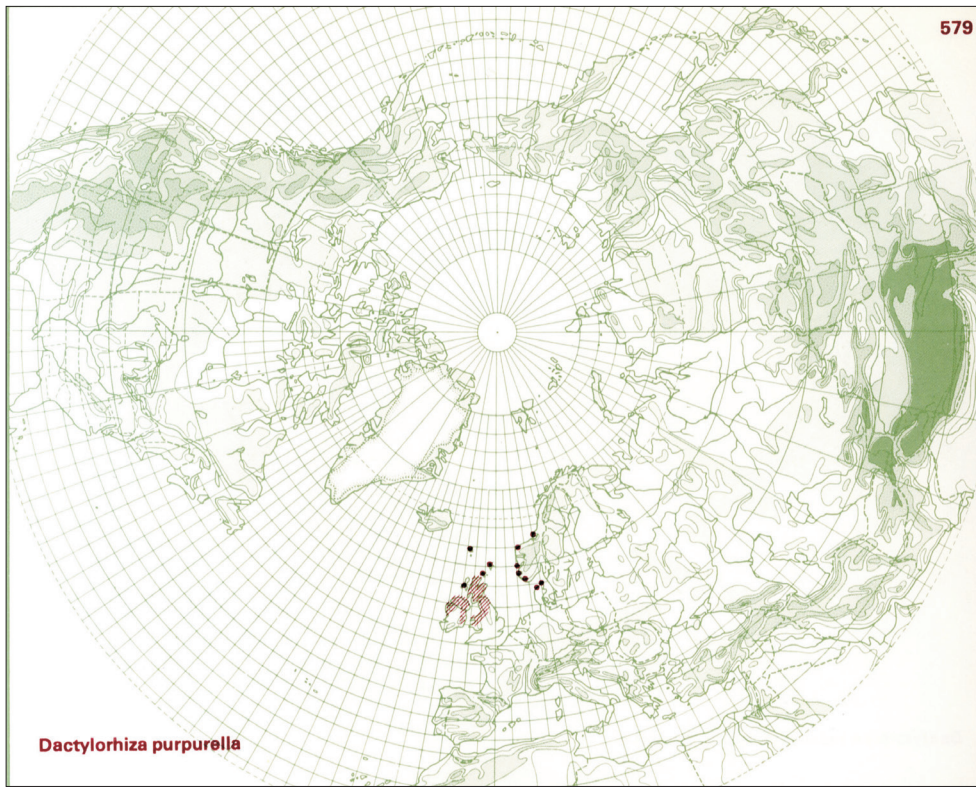


Fig. 3. The known distribution of *Dactylorhiza purpurella* (Hultén, E. and Fries, M. 1986).

Photo 1. Dactylorhiza purpurella (Photo by Lundberg, A.)

photosyntetic products to allow supply of nutrients in the following spring. If this is not the case, plants may enter hibernation. As other orchids, *D. purpurella* is a perennial species, and hibernation is not to be confused with decline or extinction *(Photo 1).*

An example of the huge range in annual fluctuations in one *D. purpurella* population is shown in *Figure 4.* Plants in this population grows in dune meadows and dune slacks. All plants were counted in six years, at the same time of the year each time. During the first four years number of plants varied from 88, 104, 63, and 98. In 2017 the number was 2,846. If red-list status had been given in

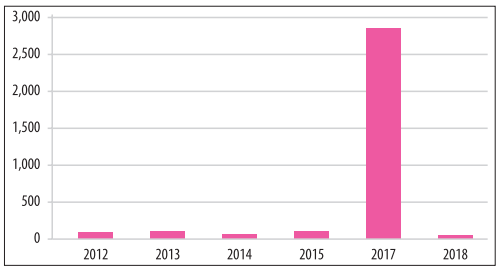


Fig. 4. Fluctuations in the number of plants of *D. purpurella* at Ognasanden, southwestern Norway. The first four years represent a normal situation, 2017 was a particularly favourable year for that species at that site.

2014, the judgement would have been heavily impacted by the low number of plants (63). The high number of plants registered in 2017 should not be considered as an increase but rather a peak within the natural fluctuation of that population.

Another example of the fluctuations in the number of *D. purpurella* plants that can be found in one population over time is from the site Kalveneset in Western Norway. This population has been monitored during four years, and the number of plants in those years were 546, 178, 1,274, and 682 *(Figure 5).* The site is untouched by technical influence and the variation from year to year has to be considered as natural fluctuation. The lesson learnt is that a time-series including several years is necessary to be able to identify the size of a population. The same will apply for measurements of the total number of national populations.

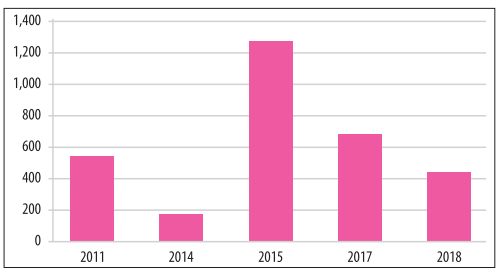


Fig. 5. Fluctuations in the number of plants of *D. purpurella* at Kalveneset, a population with no technical encroachments

The challenge of huge variations among annual plants

As mentioned, it has been known that the population of orchids may vary a lot from year to year, as also demonstrated in the examples given here. However, is this a phenomenon first and foremost characteristic of orchids or will this also be the case for other plant groups? The long-term study of *Aphanes australis* in Norway can shed light on this question. The study was initiated in 2009 as part of a monitoring program for several red-listed species in Norway (Lundberg, A. 2016). *Aphanes australis* is a small annual plant found on thin soil in open meadows. It is essential a European species. Its main distribution is found in continental Europe with the northernmost distribution in the Shetland Islands and in southwestern *Norway (Figure 6).*

All known Norwegian populations have been monitored since 2009. As mentioned, Norwegian populations of *Aphanes australis* are found on shallow ground, often on the fringe of pastures *(Photo 2).* Pastures are often heavily manured and this is a threat to *A. australis*. Several populations have become extinct because of this. On the other hand, some of the populations are not at all manured or just manured in limited quantities. Shallow soils are exposed to drought and this might be a problem for many other plant species. This is why *A. australis* can be found as a dominant species in appropriate environments. Populations might have a range of ½ x ½ to 1 x 10 m. In a good year the site may be close to totally covered by *A. australis*. However, in other years the species might not be present with adult plants at all.

This might easily be confused with extinction but long-term monitoring proved that a population may enter dormancy in certain years. In Norway, the species behaves as a summer annual plant but in certain years it may also behave as a winter annual.

Fig. 6. The known distribution of *Aphanes australis*, primarily a European species (Hultén, E. and Fries, M. 1986). Norwegian populations were not known in 1986 and are not included in the map.

Temperature is the major factor having an impact on germination, but also darkness, light and water. The rate of germination at different temperatures, humidities and light conditions proved to be governed by a continuum between dormancy and germination (Baskin, C.C. and Baskin, J.M. 2014; Lundberg, A. 2016). The total number of plants in Norway varies between <250 in a bad year and <20,000 in favourable years (Lundberg, A. 2016; *Figure 7*). It is easy to imagine that this can make a huge difference for the judgement of red-list status for such a species. The size of a population and the total number of populations in a country in one year cannot be used to decide red-list status. Long-term monitoring is essential to be able to identify the *normal and natural fluctuation interval* of a species. The problem is that such time-series are hard to find for most vascular plants and also other taxonomic groups. To obtain long-term time-series for most species is beyond any possibility but to develop some time-series of this kind should be a major target for governmental environmental bodies in all countries.

The recent Norwegian red-list for species takes into account precise terms and definitions as suggested by the IUCN, but when they are operationalised they cannot be supported by precise data, simply because such data hardly exist. Terms applied are "continuing decline in the extent of occurrence", "continuing decline in habitat quality", "continuing decline in the number of localities or subpopulations", "continuing decline in the number of mature individuals", and "continuing

Photo 2. Aphanes australis is an annual plant, in Norway usually behaving as a summer annual. Seeds often enter dormancy if spring temperatures and soil humidity are too low. This has a huge effect on the germination of seeds and is part of a normal cycle between dormancy and germination. The huge variation in adult plants present each year would have a major impact on the evaluation of red-list status but temporal data on the abundance and frequency of this and most other species hardly exist. (Photo by Lundberg, A.)

decline in population". The time period to be considered is the last ten years. Empirical evidence based on annual data developed during the last ten years does not exist for most Norwegian species. The knowledge gained from the limited number of national monitoring programs for a few species has developed some vital information but more work has to be done before the criteria for national red-listing set by the government can be met.

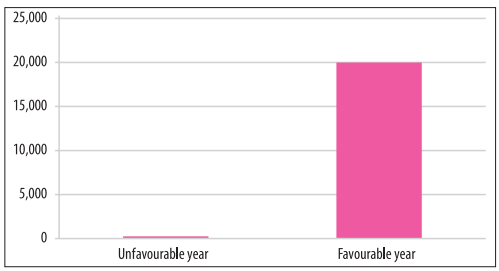


Fig. 7. Variations in spring temperature and soil moisture from year to year causes major changes in total national population size of *Aphanes australis* in Norway. The huge range between unfavourable and favourable years is a reflection of the *natural variation interval* of the total population size of that species.

Conclusions

The article discusses methods and approaches to study landscape change at different scales. Historical sources found in state archives were used to analyse the extent of deforestation and later reforestation through secondary succession. Quantitative measures found in the archives were used to identify the part of the land covered by forest through the last 400 years. Landscapes include a number of different types of phenomena, not just forests, and a way to illustrate major land-

scape attributes present in different periods of time is suggested.

At the biodiversity level, much attention is currently paid to red-listed species. Exact and precise concepts to help identify red-list status for species have been developed by national nature conservation bodies. Population trends of threatened species during a ten-year period should be applied when red-list status is identified. Although the concepts that should be at play in this evaluation are precise and clear, data to be used in the evaluation process hardly exist for most species.

This study suggests how this paradox could be dealt with and solved. Instead of a static model of population size based on information from one or a few years an alternative dynamic model is suggested. A temporal approach can avoid misinterpretations about species' extinction and can instead reveal periods of dormancy among populations. As an alternative to a static perspective on population size the identification of the natural variation interval of a species should be addressed. A temporal approach is vital in such an assessment.

REFERENCES

Baskin, C.C. and Baskin, J.M. 2014. *Seeds. Ecology, biogeography, and evolution of dormancy and germination*. San Diego, Academic Press.

Belén, M., Ortega, E., Martino, P. and Otero, I. 2018. Inferring landscape change from differences in landscape character between the current and a reference situation. *Ecological Indicators* 90. 584–593.

Bækken, I., Vihovde, A.B. and Nordby, H. 2002. *Bergen brenner. Om store branner i Bergen*. *Bergen* (Bergen on fire. About large fires in Bergen). Bergen, Det Hanseatiske Museum og Bryggens Museum.

Couch, C., Leontidu, L. and Petschel-Held, G. 2007. *Urban sprawl in Europe: landscapes, land-use change & policy*. Oxford, Blackwell.

Deák, B., Tóthmérész, B., Valkó, O., Sudnik-Wójcikowska, B., Bragina, T.M., Moysiyenko, I., Apostolova, I., Bykov, N., Dembicz, I. and Török, P. 2016. Cultural monuments and nature conservation: The role of kurgans in maintaining steppe vegetation. *Biodiversity & Conservation* 25. 2473–2490.

Deák, B., Valkó, O., Török, P., Kelemen, A., Bede, Á., Csathó, A.I. and Tóthmérész, B. 2018. Landscape and habitat and filters jointly drive richness and abundance of grassland specialist plants in terrestrial habitat islands. *Landscape Ecology* 33. 1117–1132.

Frondoni, R., Mollo, B. and Capotorti, G. 2011. A landscape analysis of land cover change in the Municipality of Rome (Italy): Spatio-temporal characteristics and ecological implications of land cover transitions from 1954 to 2001. *Landscape and Urban Planning* 100. (1): 117–128.

Heatherington, C., Jorgensen, A. and Walker, S. 2017. Understanding landscape change in a former brownfield site. *Landscape Research* 10. 1–16.

Helm, A., Hanski, I. and Partel, M. 2006. Slow response of plant species richness to habitat loss and fragmentation. *Ecology Letters* 9. 72–77.

Henriksen, S. and Hilmo, O. eds. 2015. *Norwegian Red List for Species 2015*. Trondheim, Artsdatabanken.

Hill, J., Stellmes, A., Udelhoven, Th., Röder, A. and Sommer, S. 2008. Mediterranean desertification and land degradation mapping related land use change syndromes based on satellite observations. *Global Planet Change* 64. 146–157.

Holt-Jensen, A. 2018. *Geography. History and concepts*. 5th edition, London, Sage.

Hägerstrand, T. 1995. Nature and society: The challenge of contemporary knowledge. In *Expanding environmental perspectives. Lessons of the past, prospects for the future*. Eds.: Lundgren, L.J., Nilsson, L.J. and Schlyter, P., Lund, Lund University Press, 165–172.

Hultén, E. and Fries, M. 1986. *Atlas of North European vascular plants north of the Tropic of Cancer I–III*. Königstein, Koeltz Scientific Books.

Jones, M. 1988. Land-tenure and landscape change in fishing communities on the outer coast of Central Norway, c. 1880 to the present. Methodological approaches and modes of explanation. *Geografiska Annaler* 70B. 197–204.

Kizos, T., Verburg, P.H., Bürgi, M., Gounaridis, D., Plieninger, T., Bieling, C. and Balatsos, T. 2018. From concepts to practice: combining different appraoches to understand drivers of landscape change. *Ecology and Society* 23. (25): 1–11.

Lindgaard, A. and Henriksen, S. eds. 2018. *Norsk rødliste for naturtyper 2018* (Norwegian red-list for nature types 2018). Trondheim, Artsdatabanken.

Lundberg, A. 2002. *Lindøy. Historia om folket som budde der og landskapet dei forma* (The island of Lindøy. The history of the people who lived there and the landscape they shaped). Nord-Rogaland og Sunnhordland Friluftsråd.

Lundberg, A. 2005a. Landscape transformation at Lindøy (Karmøy) 1955–2003 – a response to changes in land ownership and customary law. In *Landscape, Law and Justice*. Eds.: Jones, M. and Peil, N., Oslo, Novus forlag / The Institute for Comparative Research in Human Culture, 278–287.

Lundberg, A. 2005b. *Landskap, vegetasjon og menneske gjennom 400 år* (Landscape, vegetation and man through 400 years). Fagbokforlaget, Bergen.

Lundberg, A. 2008. Changes in the land and the regional identity of western Norway. The case of Sandhåland, Karmøy. In *Nordic landscapes. Landscape, Region and Belonging on the Northern Edge of Europe*. Eds.: Jones, M. and Olwig, K., Minneapolis, University of Minnesota Press, 344–371.

Lundberg, A. 2010. Conflicts between perception and reality in the management of alien species in forest ecosystems: a Norwegian case study. *Landscape Research* 35. (3): 319–338.

Lundberg, A. 2015. *Faggrunnlag for dvergmarikåpe, saronnellik, ekornsvingel, islandsgrønkurle, jærflangre, jærtistel og skredmjelt i Noreg. Årsrapport for 2015* (Scientific report on *Aphanes australis*, *Dianthus armeria*, *Vulpia bromoides*, *Coeloglossum viride ssp. islandica*, *Epipactis helleborine var. neerlandica*, *Serratula tinctoria* and *Oxytropis campestris ssp. scotica* in Norway. Annual report for 2015). Manuscript.

Lundberg, A. 2016. Dvergmarikåpe *Aphanes australis* i Noreg – utbreiing, økologi og tilstand (*Aphanes australis* in Norway – distribution, ecology and state). *Blyttia* 74. 241–251.

Lundberg, A. 2017. Jærflangre *Epipactis helleborine* (L.) Crantz var. *neerlandica* Verm. i Noreg – utbreiing, økologi og tilstand (*Epipactis helleborine* [L.] Crantz var. *neerlandica* Verm. in Norway – distribution, ecology and state). *Blyttia* 75. 221–232.

Lundberg, A. and Frøland, T. 2018. *Faggrunnlag for purpurmarihand Dactylorhiza purpurella. Årsrapport 2018.* (National monitoring of *Dactylorhiza purpurella* in Norway. Annual report for 2018). Miljødirektoratet, Trondheim .

Novák, T.J., Incze, J., Spohn, M., Glina, B. and Giani, L. 2014. Soil and vegetation transformation in abandoned vineyards of the Tokaj Nagy-Hill, Hungary. *Catena* 123. 88–98.

Pitkänen, P.T., Kumpulainen, J., Lehtinen, J., Sihvonen, M. and Käyhköa, N. 2016. Landscape history improves detection of marginal habitats on semi-natural grasslands. *Science of the Total Environment* 539. 359–369.

Plieninger, T. 2006. Habitat loss, fragmentation, and alteration – Quantifying the impact of land-use changes on a Spanish dehesa landscape by use of aerial photography and GIS. *Landscape Ecology* 21. 91–105.

Skjeggedal, T. 2005. The planning and building act outside urban areas – unbound ambitions, lilited possibilities. In *Landscape, Law and Justice*. Eds.: Jones, M. and Peil, N., Oslo, Novus forlag / The Institute for Comparative Research in Human Culture, 152–163.

Szabó, S., Bertalan, L., Kerekes, Á. and Novák, T.J. 2015. Possibilities of land use change analysis in a mountainous rural area: a methodological approach. *International Journal of Geographical Information Science* 30. 708–726.

Török, P., Deák, B., Vida, E., Valkó. O., Lengyel, S. and Tóthmerész, B. 2010. Restoring grassland biodiversity: Sowing low-diversity seed mixturescan lead to rapid favourable changes. *Biological Conservation* 143. 806–812.

Török, P., Helm, A., Kiehl, K., Buisson, E. and Valkó, O. 2018. Beyond the species pool: modification of species dispersal, establishment, and assembly by habitat restoration. *Restoration Ecology* 26. 65–72.

Yang, Y., Zhang, S., Liu, Y., Xing, X. and Sherbinin, A. 2017. Analysing historical land use changes using a Historical Land Use Reconstruction Model: a case study in Zhenlai Country, northeastern China. *Scientific Reports* 7. 41275. Doi: 10.1038/srep41275

DOI: 10.15201/hungeobull.67.4.2 Hungarian Geographical Bulletin **67** 2018 (4) 319–332.

Assessing environmental changes in abandoned German vineyards. Understanding key issues for restoration management plans

Jesús RODRIGO-COMINO[1], Martin NEUMANN[2], Alexander REMKE[3] and Johannes B. RIES[4]

Abstract

Land degradation in vineyards is a big concern which should be considered by farmers, enterprises and policymakers. Due to intense tillage, the use of herbicides and heavy machinery, vine plantations are registering a decrease in soil fertility and, subsequently, in productivity. Recently, farmers have decided to abandon the vineyards, but any restoration planning is being carried out to recover biodiversity or to reduce soil and water losses. Nowadays, there is no information about environmental changes after the abandonment in terms of possible soil property changes and erosion in Central European vineyards such as in Germany. Therefore, the main aims of this preliminary study were to compare: i) soil properties and soil profiles of one cultivated vineyard and an abandoned one; and, ii) to assess the activation of soil erosion processes using a small portable rainfall simulator. Our results showed that the vineyard registered several differences in soil properties among slope positions and soil profile characteristics due to tillage and trampling effects, showing clear marks of compaction and soil detachment in the lower parts. Also, in this cultivated field, higher means and maxima of soil losses (g m^{-2}) and sediment concentration (g l^{-1}) values than in the abandoned plot were quantified, being the main driving factors the vegetation cover and the inclination. On the other hand, in the abandoned vine plantation, a rapid homogenization of soil profiles and soil properties were found along the hillslope, where a deeper organic horizon was consistently developed above a compacted and rocky horizon, which was generated during the cultivation phase. Due to the high compaction due to the machinery cultivation and the difficulties for the roots to make deep into the soil, the infiltration defaulted and the amount of runoff and runoff coefficient were higher in the abandoned plots than in the cultivated ones.

Keywords: Soil erosion, German vineyards, abandonment, rainfall simulation, soil profile

Introduction

Soil erosion is a big concern for humankind because soils provide indispensable sources and goods for living creatures and human health (Smith, P. *et al.* 2015). However, negative human impacts on soils such as the intensification of the agricultural practices are generating a drastic decrease in soil fertility and quality (Szalai, Z. *et al.* 2016). Therefore, to solve these kinds of problems and achieve the best solutions, the scientific community and the policymakers should collaborate with the stakeholders, actively. In this way, physical geographers have to play an important role in developing research methods and tools which are able to design sustainable land plans and feasible measures.

Techniques such as modeling (Samani, A.N. *et al.* 2016; Balázs, B. *et al*. 2018), erosion plots (Kinnell, P.I.A. 2016) or isotopic markers (Bihari, Á. and Dezső, Z. 2008; Novara,

[1] Instituto de Geomorfologia y Suelos, Department of Geography, University of Malaga, 29071, Malaga, Spain. Physical Geography, Trier University, 54286 Trier, Germany. E-mail: rodrigo-comino@uma.es
[2] Czech Technical University in Prague, Faculty of Civil Engineering, Department of Landscape Water Conservation, Prague 6, Czech Republic. E-mail: martin.neumann@fsv.cvut.cz
[3] Physical Geography, Trier University, 54286 Trier, Germany. Diensleistungszentrum Ländlicher Raum Mosel, 54470 Bernkastel-Kues, Germany. E-mail: geo_trier@web.de
[4] Physical Geography, Trier University, 54286 Trier, Germany. E-mail: riesj@uni-trier.de

A. *et al.* 2016; Jakab, G. *et al*. 2018) are the most common methods applied to quantify soil erosion. However, to make reproducible and comparable results of water and soil losses, rainfall simulations can be also considered a valuable tool (Iserloh, T. *et al.* 2012; Szabó, J. *et al.* 2015).

In vineyards, the use of rainfall simulations to study initial soil erosion processes has increased because they are one of the most degraded landscapes. Rodrigo-Comino, J. *et al.* (2016a, b) qualitatively assessed different viticultural areas across Europe where distinct rainfall simulations showed high soil and water losses in Campo Real (Madrid, Spain), Champagne (France), the Pènedes (Lleida, Spain) or Ruwer-Mosel valley (Trier, Germany). Also, the use of small portable rainfall simulators has been applied to investigate different specific environmental characteristics in vineyards or erosion control measures (Blavet, D. *et al.* 2009; Morvan, X. *et al.* 2014). However, there is another process that is also affecting the rest of the European vineyards and which has not sufficiently been investigated: land abandonment (Lasanta, T. *et al.* 2015).

Vineyards´ soils are suffering from a high degradation as a consequence of intense tillage, the use of herbicides and heavy machinery, registering a decrease in soil fertility and, subsequently, also in productivity (Camps, J.O. and Ramos, M.C. 2012; García-Díaz, A. *et al.* 2017). Therefore, when the most fertile horizon is eliminated, vine growers decide to abandon the whole plantation. Also, as a consequence of the climate change, low lands are being abandoned, and hillslopes on the higher heights are being planted (Arnaez, J. *et al.* 2006; Galilea Salvador, I. *et al.* 2015). Recently, using rainfall simulations experiments (Martínez-Hernández, C. *et al.* 2017), it was observed that areas where there was no vegetation recovery at all, such as in almond trees, soil loss and runoff were higher than in the cultivated areas.

In Germany, the viticultural sector is reporting high benefits for wine producers and new planting is taking place (O.I.V. 2017). However, when a plantation is not productive or the next generation of farmers do not show any interest in vineyards, the abandonment process begins and a restoration plan should be conducted. Vegetation and biodiversity recoveries show positive benefits for both environment and humankind (Bienes, R. *et al.* 2016), but no incentives to carry out some kind of measures default this action.

In this way, there is no information about which environmental problems after the abandonment (e.g. soil erosion) in Central European vineyards such as in Germany could occur. We only found some precedents, for example in a study carried out in Eastern countries such as Hungary in the traditional Tokaj viticulture area, where the vegetation transformation and toposequences of the carbon storage after the abandonment and its influence on soil changes were studied (Novák, T.J. *et al.* 2014).

Therefore, the main goal of this preliminary research is to compare soils properties and initial soil erosion processes in a cultivated vineyard with an abandoned one. We pretend to show the main differences and transformations after the abandonment process in the same vineyard. To achieve this goal, soil profiles, soil analysis, and rainfall simulations were used.

Materials and methods

Study area

The localization of the two studied paired-plots can be observed in *Figure 1*. The selected vineyard and abandoned one are situated in the little village of Waldrach in the Ruwer-Moselle valley (Rhineland-Palatinate), Germany. The average elevation ranges from 200 m to 400 m a.s.l. and all are located on Devonian grauwackes, slates, and quartzites, which are in contact with Pleistocene fine materials transported by the Ruwer river, an affluent of Moselle river (Richter, G. 1979). The vine plantation is composed of 40-years old plants and was cultivated in the summit of a

Fig. 1. Study area and rainfall simulation localization of the experiments

hillslope. On the other hand, the abandoned study plot (cultivated during 1970 with similar tillage practices to the recently cultivated one) was abandoned during 1990. In general, the hillslopes are exposed to SW direction and mean inclinations reach maximum values of 30º, although the studied abandoned plot shows more gentle angles (15–25°). Annual total average rainfall is about 765 mm and mean annual temperatures approximately 9 °C (Rodrigo-Comino, J. *et al.* 2015).

The grape variety is Riesling and the main soil management practices are as follows: i) tilling with machinery before and after the vintage to 20 cm depth (beginning of spring, and autumn); ii) the use of vine training systems with a plantation framework of 90 x 140 cm; iii) a high amount of slate mulch to conserve soil temperature regime; and, iv) keeping soils bare as much as possible by applying pesticides and herbicides.

In both areas, on the embankments and inter-rows, rills, landslides, and ephemeral gullies as a consequence of soil erosion can be observed. The abandoned plot is cleaned from spontaneous vegetation once in the year to prevent the recolonization close to the roads and drainages as a part of maintenance practice.

Soil profile description and soil analysis

Soil samples were collected from three different slope positions (shoulder, backslope, and footslope), at two different depths (0–5 cm and 5–15 cm) in the rows and inter-row areas. All the samples were analysed with three replicates, being a total 36 samples and amounting to about 3–4 kg per soil sample. First, at all, soil samples were sieved (<2 mm) and basic soil properties were analysed in the laborato-

ry: Texture, total organic carbon (TOC), Calcium carbonates (Ca), electrical conductivity (EC), pH-value and soil water content (SWC).

Texture (sand, silt, and clays) was analysed by a Coulter LS230 device, by combining different diffraction patterns of a light beam. Total organic carbon was measured by loss of ignition (LOI) and its weight difference under 430 °C (24 h) in a muffle furnace (Davies, B.E. 1974; Rosell, R.A. *et al.* 2001). Electrical conductivity (EC) was measured by a digital conductivity-meter and carbonates with a Bernard calcimeter. pH-value in distilled water (1:5 proportion) using a digital pH-meter was obtained. Soil water capacity content at field capacity and wilting point were calculated with a pressure plate extractor.

Finally, three soil profiles at different slope positions (coinciding with the soil sample places) were described to classify soil types, using the methodology designed by FAO-WRB (IUSS Working Group WRB 2006, 2014).

Rainfall simulations

Nine rainfall simulations were carried out in the cultivated vineyard and fourteen in the abandoned one to compare soil loss, runoff, runoff coefficient, sediment concentration and infiltration. In *Figure 1*, the localisation of the experiments was mapped. We used a small nozzle-type rainfall simulator modified by Ries, J.B. *et al.* (2009). This device is characterized by i) a square metal frame (0.45 m x 0.45 m) with a Lechler 460 608 nozzle; ii) four telescopic aluminium legs in order to situate the nozzle two meters above the plot; iii) the aluminium linkage is covered by a rubber tarpaulin to eliminate wind interferences; iv) a circular test plot of 0.28 m^2 with a V-shaped outlet, which is put at the deepest point at surface level to collect the water and soil losses; v) a flow control and a 12 V low-pressure bilge pump that controls and make reproducible the simulated rainfall. The rainfall simulator was calibrated by Iserloh, T. *et al.* (2012) for a rainfall intensity of 40 mm h^{-1}.

Each experiment had a total duration of 30 minutes and was conducted in a randomized block at different slope positions. In five minutes´ intervals (six intervals in total), water and sediments were collected in plastic bottles, which were also changed at the beginning of a new interval. Prior starting the experiments, vegetation and rock fragment covers were perceptively quantified by taking the opinion of three experts, soil roughness was assessed with the chain method (Saleh, A. 1993), slope inclination was measured with a digital clinometer and antecedent soil moisture was calculated by taking a soil sample close to the ring plot and drying at air conditions in the laboratory. The collected water with sediments in each bottle was filtrated with circular fine-meshed filter papers (Munktell©, Prod.-Nr. 3.104.185, less than 2 μm mesh-width) and, then, filters were dried to constant weight at 105 °C. After that, they were weighted for determining soil loss (g) and runoff (l) for each measured interval. Final results were presented in g m^{-2} and l m^{-2} in order to be comparable with other study areas. Also, sediment concentration (g l^{-1}) was obtained by dividing the amount of soil loss and runoff. Runoff and infiltration coefficients were also calculated using the total area of the plot and rainfall intensity in each interval.

Statistical analysis

Descriptive statistics in boxplot graphics and tables to identify averages, maximum, minimum, median and outlier values were depicted and summarized, respectively. To compare soil properties obtained from both paired-plots (cultivated and abandoned), a nonparametric test at $p > 0.05$ was performed after testing the data normality (Shaphiro-Wilk test) and equal variance (F-test). They did not show a normal distribution. We used a Tukey test, where significant differences at $p < 0.001$ level were considered.

Finally, to confirm which driving factor enhances soil erosion and makes a compari-

son between which environmental plot characteristic and erosion result shows possible interrelationships, a Spearman's rank correlation coefficient was conducted. SigmaPlot 12.0 (Systat Software Inc.) was the software used to carry out the statistical analysis.

Results and discussion

Soil analysis

Soil analysis results showed a higher proportion of coarse gravel (>2 mm) in the cultivated plots than in the abandoned one is registered. It is important to remark that in both plots more than 34 per cent of gravels were found. The rock fragment cover is a widely studied factor in vineyards (Rodrigo-Comino, J. *et al.* 2017) and in other crops or environments (Nyssen, J. *et al.* 2001; Jomaa, S. *et al.* 2012), because it shows a strong correlation with runoff and soil losses when it is embedded into the soil (Poesen, J.W. *et al.* 1998). If not, rock fragment cover uses to enhance infiltration (Zavala, L.M. *et al.* 2010) and biodiversity activity (Certini, G. *et al.* 2004). Moreover, both soils show a silty texture, but after the abandonment, a higher content of clays and fine silts can be registered. This process was also registered in other abandoned areas with schists as parent material (Martínez-Hernández, C. *et al.* 2017), although it appears more frequent in calcareous rocks, where a selective removal of fine particles occurs (Romero Díaz, A. *et al.* 2011), also affecting other soil properties such as water retention capacity and pH (Lesschen, J.P. *et al.* 2008; Bienes, R. *et al.* 2016). In our study area, this process could also be recognized. After the abandonment, increases in water retention capacity at the wilting point and at field capacity are observed. Moreover, pH also decreases, showing a more acidic trend, which is also an ecological indicator of soil quality registered after each land use change (Khaledian, Y. *et al.* 2017; Pahlavan-Rad, M.R. and Akbarimoghaddam, A. 2018). Statistical analysis proved that soil texture, organic carbon, carbonate, soil water retention capacity and pH show a significant difference among cultivated and abandoned plots, confirming the changes in soil properties after the abandonment *(Table 1).*

Soil profile descriptions and qualitatively assessment

In *Figure 2* and *3*, soil profiles described at different slope positions also show differences among each other and after the abandonment process. In Suppl. Material, the description of all soil profiles is included to observe more in detail these differences.

As we can observe in the cultivated plot *(Figure 2)*, soil profiles in the shoulder and backslope positions are characterized by a thin organo-mineral soil horizon (nearly 2 cm deep) with high alteration induced by tillage and compaction. This horizon can be signed as Ap. Underneath, a tilled soil horizon, which could be considered as Ap_2, has deeper mineral soil horizons (B/C). The horizon boundaries in Ap and B/C are abrupt (2–5 cm) and nearly plane in the compacted layers, and irregular in the rocky layer. Soil structure grade ranges from moderate to weak, with prismatic and crumb forms. However, in the footslope, one unique horizon can be described which is characterized by clear marks of compaction. Moreover, the surface horizon is removed as a consequence of the depletion, which was also confirmed by Rodrigo-Comino, J. *et al.* (2016b) using the stock unearthing method and topsoil level change maps. It is widely known in studies about connectivity processes that soil depth variations among slope positions can be linked to the fact of the mass movement processes, registering in the upper part erosion, in the middle erosion-deposition (transition) and, finally, in the lower part sedimentation (López-Vicente, M. *et. al.* 2015; Novara, A. *et al.* 2016; Rodrigo-Comino, J. *et al.* 2016a, b; Ben-Salem, N. *et al.* 2018). However, in tilled vineyards, the redistribution of materials (Follain, S. *et al.* 2012; Quiquerez, A. *et al.* 2014), tractor passes (Biddoccu, M. *et al.* 2017) and extreme

Table 1. Soil properties and differences among cultivated and abandoned plots

Plots		Soil particles		Sand	Silt	Clay	LOI	CaC	FC	WP	pH, H_2O	EC, dS m^{-1}
		>2 mm	<2 mm	in %								
Cultivated	x̄	37.7	62.3	44.9	45.5	9.6	4.7	0.7	24.4	10.3	7.1	0.2
	±	6.6	6.6	4.4	3.3	1.4	1.0	0.3	2.3	0.9	0.2	0.1
	Max	46.2	73.8	52.0	50.0	12.6	6.6	1.3	29.3	12.5	7.4	0.5
	Min	26.2	53.8	38.0	39.8	7.5	3.6	0.3	21.3	8.8	6.7	0.1
Abandoned	x̄	34.6	65.4	16.9	70.5	12.6	6.7	2.6	31.2	12.6	6.6	0.2
	±	17.9	17.9	7.1	4.3	3.3	1.6	0.7	4.0	2.4	0.7	0.1
	Max	65.7	93.2	33.5	80.3	20.4	9.5	5.0	37.2	16.7	7.9	0.4
	Min	6.8	34.3	2.1	59.5	7.0	3.9	2.2	25.4	8.8	6.0	0.1
Differences	p<	0.488		0.001						0.002	0.011	0.282

Notes: LOI = Los organic matter by ignition; CaC = Carbonate content; FC = Soil water content at the field point; WP = Soil water content at the wilting point; EC = Electrical conductivity.

rainfall events (Martínez-Casasnovas, J.A. *et al.* 2003; De Santisteban, L.M. *et al.* 2006) make difficult to predict topsoil level changes along the hillslope.

In the abandoned vine plantation *(Figure 3)*, soil profiles are characterized by 0 to 4 cm deep-rooted organic soil horizons (litter horizon O), where there are no rests of any Ap horizon and soil aggregates are similarly absent. The boundaries with underlying soil horizons are also abrupt. Underneath, we find a B/C horizon characterized by tilled and compacted mineral features. High rock fragment contents are noted in this layer with a weak soil structure characterized by prismatic and crumb forms of 20–50 mm size. Several authors confirmed that the recuperation of abandoned soils in semiarid and arid areas need long periods (Romero Díaz, A. *et al.* 2011; Kou, M. *et al.* 2016); however, in two decades we observe that Central European vineyards are able to generate a consistent A horizon relatively fast; although the compaction marks stay there yet.

Finally, these soils can be classified as *Leptic-Humic Regosols* according to the FAO/WRB soil classification (IUSS Working Group WRB, 2014).

Initial soil erosion processes

Rainfall simulation results can be observed in *Table 2* and *3*, where soil erosion results and environmental plot characteristics inside the ring plot are summarized, confirming high differences among plots. In *Figure 4*, mean, median, maximum, minimum values and outliers of measured soil erosion in both plots are depicted in box plots.

The most important differences inside the ring plots are found for vegetation cover and roughness, being higher in the abandoned plot than in the cultivated vineyard (97% vs 45%; 1.3 mm mm^{-1} vs 1.05 mm mm^{-1}). On the other hand, rock fragment covers and slope grades are higher in the cultivated vineyard than in the abandoned plot, reaching average values of 58, 28, 17 and 5 per cent, respectively.

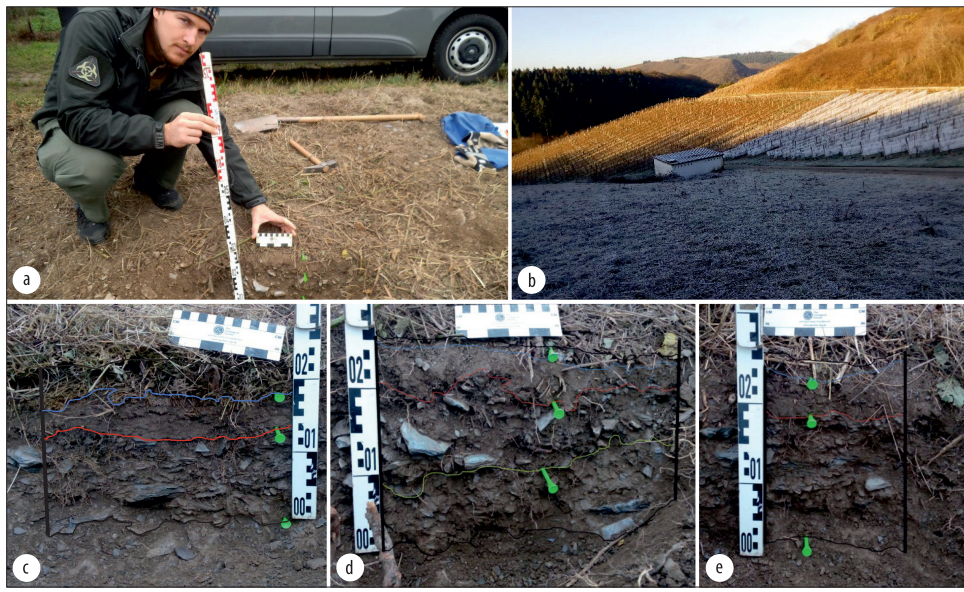


Fig. 2. Soil profiles in the cultivated plot. – a = soil profile elaboration; b = a general perspective of the plot; c = shoulder; d = backslope; e = footslope

Fig. 3. Soil profiles in the abandoned plot. – a–e = For explanation see *Fig. 2.*

Table 2. Rainfall simulation results and environmental plot characteristics in the cultivated vineyard

ID	Runoff			Soil loss		SC, g l^{-1}	Slope	VC	RFc	ASM	Roughness, mm mm^{-1}
	l	l m^2	%	g	g m^2					%	
1	0.2	0.7	3.9	0.6	2.2	3.3	26.0	0.0	100.0	12.4	1.1
2	0.0	0.1	0.5	0.1	0.4	5.0	36.5	35.0	90.0	14.3	1.0
3	1.0	3.6	15.2	7.2	25.8	7.8	23.0	10.0	60.0	16.5	1.1
4	0.3	1.2	6.7	1.5	5.3	7.0	27.0	70.0	75.0	17.0	1.1
5	0.0	0.0	0.0	0.0	0.0	0.0	27.1	40.0	60.0	18.0	1.0
6	0.0	0.0	0.0	0.0	0.0	0.0	37.0	7.5	85.0	25.7	1.0
7	0.0	0.0	0.0	0.0	0.0	0.0	26.0	85.0	15.0	27.8	1.0
8	0.0	0.0	0.0	0.0	0.0	0.0	24.0	75.0	15.0	29.1	1.0
9	0.0	0.0	0.0	0.0	0.0	0.0	21.0	80.0	20.0	28.2	1.3

Notes: SC = Sediment concentration; VC = Vegetation cover; RFc = Rock fragment covers; ASM = Antecedent soil moisture

Table 3. Rainfall simulation results and environmental plot characteristics in the abandoned vineyard

ID	Runoff			Soil loss		SC, g l^{-1}	Slope	VC	RFc	ASM	Roughness, mm mm^{-1}
	l	l m^2	%	g	g m^2		%				
1	0.0	0.0	0.0	0.0	0.0	0.0	10.0	100.0	5.0	23.0	1.1
2	0.4	1.5	7.7	0.0	0.1	0.0	15.0	100.0	5.0	19.0	1.1
3	0.0	0.0	0.0	0.0	0.0	0.0	15.0	100.0	0.0	29.7	1.2
4	0.6	2.2	10.9	0.1	0.4	0.1	18.0	100.0	5.0	33.1	1.1
5	0.0	0.0	0.0	0.0	0.0	0.0	18.0	100.0	5.0	22.8	1.5
6	0.1	0.5	2.6	0.0	0.0	0.1	16.0	90.0	30.0	26.1	1.5
7	1.7	6.0	28.1	3.6	13.0	3.3	20.0	100.0	0.0	30.6	1.6
8	0.0	0.0	0.0	0.0	0.0	0.0	25.0	100.0	0.0	29.3	1.3
9	0.0	0.0	0.0	0.0	0.0	0.0	13.0	100.0	10.0	31.6	1.6
10	0.0	0.0	0.0	0.0	0.0	0.0	18.0	90.0	15.0	13.4	1.8
11	0.0	0.0	0.0	0.0	0.0	0.0	15.0	100.0	0.0	16.1	1.2
12	0.0	0.0	0.0	0.0	0.0	0.0	20.0	100.0	0.0	20.5	1.4
13	0.0	0.0	0.0	0.0	0.0	0.0	15.0	80.0	5.0	29.6	1.4
14	0.0	0.0	0.0	0.0	0.0	0.0	15.0	100.0	0.0	16.1	1.2

Notes: SC = Sediment concentration; VC = Vegetation cover; RFc = Rock fragment covers; ASM = Antecedent soil moisture

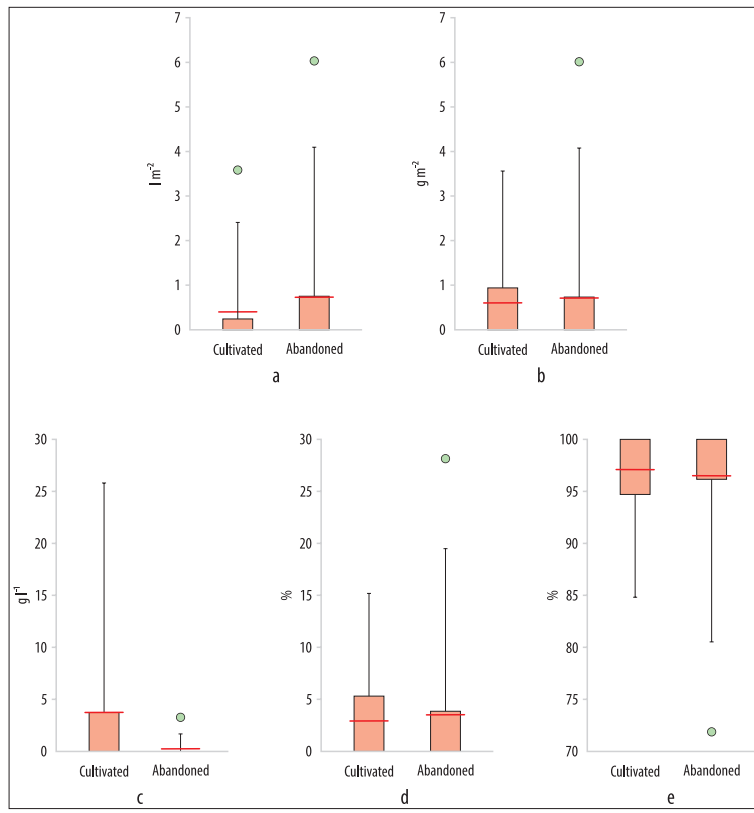


Fig. 4. Box plot of the soil erosion results in the cultivated and abandoned plots. – a = runoff; b = soil loss; c = sediment concentration; d = runoff coefficient; e = infiltration coefficient. Red dotted line represents the mean values.

In the vineyard, 56 per cent of the rainfall simulations do not obtain any runoff. On the other hand, in the abandoned field, a 71 per cent of the total experiments do not show water losses. In the cultivated plot, mean total runoff is 0.6 l m^2, reaching maximum values of 3.6 l m^2, meaning an average runoff coefficient of 3 per cent and maximums of 15 per cent. In the abandoned vine plantation, mean total runoff is 0.7 l m^2 and maximum values reach 6 l m^2. In terms of runoff coefficient, the abandoned plots show higher percentages, reaching mean values of 4 per cent and maximum amounts of 28 per cent.

These results confirm that after the abandonment and without a planned hillslope restoration, runoff can increase. These results are also coincident with other studied areas such as in the Mediterranean landscapes (Lasanta, T. *et al.* 2015). However, Mediterranean authors claim that this fact is due to bare soils, a decrease in porosity and an increase in soil Calcium carbonate crusts as a consequence of the high temperatures, low organic content and calcareous parent material (Romero Díaz, A. *et al.* 2007; Seeger, M. and Ries, J.B. 2008). In the studied vineyards, the main reason is that the vegetation cover, although very dense during winter and spring (>100%), is eliminated by the owners to maintain clean roads and drainages. As a consequence, the vegetation does not have enough time to act as a useful protection during the rainiest season of the year. Moreover, as we observed in the soil profiles´ description section, the sub-surface layers are

strongly compacted in specific slope positions and the roots cannot develop a stable net. As a result, the infiltration capacity and water retention capacity could be reduced (Botta, G.F. *et al.* 2012).

In the cultivated plot, mean total soil loss show values of about 4 g m^2 and a maximum of 25.8 g m^2. Our findings register a mean sediment concentration of 3 g l^{-1} and a maximum of 7.8 g l^{-1}. In the abandoned vineyard, 1 g m^2 and 13 g m^2 are the mean and maximum soil loss values, respectively. Sediment concentration results in the abandoned vineyards are also lower than in the cultivated plots, reaching mean values of 0.3 g l^{-1} and maximums of 3.3 g l^{-1}. These results confirm that the vineyards are more devoted to registering initial soil erosion processes than the abandoned plots, as other authors in several countries also registered (Marques, M.J. *et al.* 2008; Chevigny, E. *et al.* 2014; Biddoccu, M. *et al.* 2017; Ben-Salem, N. *et al.* 2018). However, it is important to remark that future research should be focused on studies over a long-term period to observe if they overpass tolerable soil erosion rates or not (Verheijen, F.G.A. *et al.* 2009).

Finally, in *Table 4* and *5*, Spearman rank´s coefficients are applied to observe which environmental factor acts as driving factor of soil erosion. In the cultivated plot, we observed that there is a strong correlation between the runoff generation and soil loss, and sediment concentration. These results also coincide with other crops, where bare soils and steep slopes generate a parallel increase in water and soil losses such as in olives or citrus orchards (Taguas, E.V. *et al.* 2015; Jianjun, W. *et al.* 2017). As recently mentioned, it exists a high correlation between bare soils and an increase of vegetation cover, which not only protect against soil and water losses, but also enhance nutrients retention (Olmstead, M.A. *et al.* 2001; Fourie, J.C. *et al.* 2016) and biodiversity development (Barrio, I.C. *et al.* 2012; Lopes, C. *et al.* 2015) as well. On the other hand, we observe that in the abandoned plot, only a high correlation is found with sediment concentration and antecedent soil moisture. This result also confirms that: i) when soils are saturated, soil and water losses are also activated, responding to a Hortonian dynamic (Imeson, A.C. and Lavee, H. 1998; Ziegler, A.D. *et al.* 2007); and, ii) vegetation cover reduces soil erosion activation, but a non-planned hillslope restoration modifies the hydrological dynamic, making it difficult to predict the spatial intra-plot variability.

Conclusions

This research pretends to demonstrate the significant changes in soil properties and initial soil erosion processes generated after vineyard´s abandonment. In *Figure 5,* we represented our findings which demonstrated that: i) in vineyards, there are several differences in soil properties and soil profiles among slope positions due to tillage and trampling effects, showing clear marks, features of compaction and soil depletion in the footslopes; ii) also, in the cultivated field, we registered higher mean and maximum values of soil loss and sediment concentration data

Table 4. Spearman rank´s coefficient in the cultivated vineyard

Indicators	Runoff	Soil loss	SC	RC	Slope	VC	RFc	ASM	Roughness
Runoff	–	0.994*	0.982*	1.000*	0.536*	-0.765*	0.755*	-0.152	0.126
Soil loss	–	–	0.994*	0.994*	0.531*	-0.755*	0.761*	-0.126	0.105
SC	–	–	–	0.982*	0.549*	-0.739*	0.755*	-0.121	0.110
RC	–	–	–	–	0.536*	-0.765*	0.755*	-0.152	0.126

Notes: SC = Sediment concentration; RC = Runoff coefficient; VC = Vegetation cover; RFc = Rock fragment covers; ASM = Antecedent soil moisture. *p<0.05.

Table 5. Spearman rank´s coefficient in the abandoned vineyard

Indicators	Runoff	Soil loss	SC	RC	Slope	VC	RFc	ASM	Roughness
Runoff	–	1.000*	0.854*	1.000*	0.270	0.048	0.134	0.369	-0.102
Soil loss	–	–	0.854*	1.000*	0.270	0.048	0.134	0.369	-0.102
SC	–	–	–	0.854*	0.381	-0.064	0.114	0.517*	0.077
RC	–	–	–	–	0.270	0.048	0.134	0.369	-0.102

Notes: SC = Sediment concentration; RC = Runoff coefficient; VC = Vegetation cover; RFc = Rock fragment covers; ASM = Antecedent soil moisture. *p<0.05.

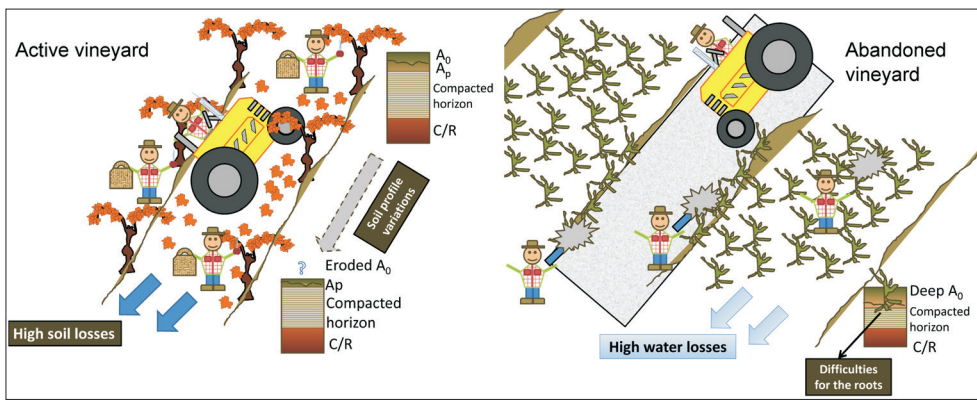


Fig. 5. Conclusions obtained from the cultivated and abandoned vineyards

than in the abandoned plot, being the vegetation cover and the steeper slopes the main driving factors; iii) on the contrary, in the abandoned plots a rapid homogenization of soil horizons and soil properties were found along the hillslope, where a deeper organic horizon was consistently developed on an underneath compacted and rocky horizon, which was developed during the cultivation phase; and, iv) due to high compaction and obstructed development of the roots, runoff and runoff coefficient were higher than in the cultivated plots.

Therefore, we claim that for the Central European vineyards under continental climate conditions, fortunately, at short-term periods, high facilities of a rapid recolonization and soil recuperation can be registered. However, any restoration plan that promotes a deep ploughing to remove the compacted sub-surface horizons and the prevention of annual pruning of vegetation after spring is not well suited for hillslope restoration and lessening soil and water losses.

REFERENCES

Arnáez, J., Ortigosa-Izquierdo, L.M., Ruiz-Flaño, P. and Lasanta, T.L. 2006. Distribución espacial del viñedo en la Comunidad Autónoma de La Rioja: influencia de la topografía y de las formas del relieve (Spatial distribution of the vineyards in the Autonomous Community La Rioja: the influence of the topography and relief). *Polígonos: Revista de geografía* 16. 11–34.

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

Barrio, I.C., Villafuerte, R. and Tortosa, F.S. 2012. Can cover crops reduce rabbit-induced damages in vineyards in southern Spain? *Wildlife Biology* 18. 88–96.

Ben-Salem, N., Álvarez, S., López-Vicente, M. 2018. Soil and water conservation in rainfed vineyards with common sainfoin and spontaneous vegetation under different ground conditions. *Water* 10(8) 1058. Doi:10.3390/w10081058

Biddoccu, M., Ferraris, S., Pitacco, A. and Cavallo, E. 2017. Temporal variability of soil management effects on soil hydrological properties, runoff and erosion at the field scale in a hillslope vineyard, North-West Italy. *Soil and Tillage Research* 165. 46–58.

Bienes, R., Marques, M.J., Sastre, B., García-Díaz, A. and Ruiz-Colmenero, M. 2016. Eleven years after shrub revegetation in semiarid eroded soils. Influence in soil properties. *Geoderma* 273. 106–114.

Bihari, Á. and Dezső, Z. 2008. Examination of the effect of particle size on the radionuclide content of soils. *Journal of Environmental Radioactivity* 99. (7): 1083–1089.

Blavet, D., De Noni, G., Le Bissonnais, Y., Leonard, M., Maillo, L., Laurent, J.Y., Asseline, J., Leprun, J.C., Arshad, M.A. and Roose, E. 2009. Effect of land use and management on the early stages of soil water erosion in French Mediterranean vineyards. *Soil and Tillage Research* 106. 124–136.

Botta, G.F., Tolon-Becerra, A., Tourn, M., Lastra-Bravo, X. and Rivero, D. 2012. Agricultural traffic: Motion resistance and soil compaction in relation to tractor design and different soil conditions. *Soil and Tillage Research* 120. 92–98.

Camps, J.O. and Ramos, M.C. 2012. Grape harvest and yield responses to inter-annual changes in temperature and precipitation in an area of north-east Spain with a Mediterranean climate. *International Journal of Biometeorology* 56. 853–864.

Certini, G., Campbell, C.D. and Edwards, A.C. 2004. Rock fragments in soil support a different microbial community from the fine earth. *Soil Biology and Biochemistry* 36. 1119–1128.

Chevigny, E., Quiquerez, A., Petit, C. and Curmi, P. 2014. Lithology, landscape structure and management practice changes: Key factors patterning vineyard soil erosion at metre-scale spatial resolution. *Catena* 121. 354–364.

Davies, B.E. 1974. Loss-on-Ignition as an estimate of soil organic Matter. *Soil Science Society of America Journal* 38. 150–151.

De Santisteban, L.M., Casalí, J. and López, J.J. 2006. Assessing soil erosion rates in cultivated areas of Navarre (Spain). *Earth Surface Processes and Landforms* 31. 487–506.

Follain, S., Ciampalini, R., Crabit, A., Coulouma, G. and Garnier, F. 2012. Effects of redistribution processes on rock fragment variability within a vineyard topsoil in Mediterranean France. *Geomorphology* 175–176. 45–53.

Fourie, J.C., Louw, P.J.E. and Agenbag, G.A. 2016. Cover Crop Management in a Sauvignon Blanc/ Ramsey Vineyard in the Semi-Arid Olifants River Valley, South Africa. 2. Effect of different cover crops and cover crop management practices on grapevine performance. *South African Journal of Enology and Viticulture* 28. 81–91.

Galilea Salvador, I., Lasanta Martínez, T., Arnáez-Vadillo, J. and Ortigosa, L.M. 2015. Evolución y desfragmentación del paisaje del viñedo en la Rioja alta (España) en el periodo 1956–2000 (*Evolution and defragmentation of the vineyard l*andscapes in la Rioja alta [Spain] in the period 1956–2000). Boletín de la Asociación de Geógrafos Españoles 69. 315–331.

García-Díaz, A., Marqués, M.J., Sastre, B. and Bienes, R. 2017. Labile and stable soil organic carbon and physical improvements using ground-covers in vineyards from central Spain. *The Science of the Total Environment* 621. 387–397.

Imeson, A.C. and Lavee, H. 1998. Soil erosion and climate change: the transect approach and the influence of scale. *Geomorphology* 23. 219–227.

Iserloh, T., Fister, W., Seeger, M., Willger, H. Ries, J.B. 2012. A small portable rainfall simulator for reproducible experiments on soil erosion. *Soil and Tillage Research* 124. 131–137. https://doi.org/10.1016/j.still.2012.05.016

IUSS Working Group WRB 2006. *Guidelines for constructing small-scale map legends using the WRB*. 2nd edition. Rome, FAO.

IUSS Working Group WRB 2014. *World Reference Base for Soil Resources 2014*. Rome, FAO.

Jakab, G., Hegyi, I., Fullen, M., Szabó, J., Zacháry, D. and Szalai, Z. 2018. A 300-year record of sedimentation in a small tilled catena in Hungary based on δ13C, δ15N, and C/N distribution. *Journal of Soils Sediments* 18. 1767–1779.

Jianjun, W., Quansheng, L. and Lijiao, Y. 2017. Effect of intercropping on soil erosion in young citrus plantation – a simulation study. *Chinese Journal of Applied Ecology* 8. 143–146.

Jomaa, S., Barry, D.A., Brovelli, A., Heng, B.C.P., Sander, G.C., Parlange, J.-Y. and Rose, C.W. 2012. Rain splash soil erosion estimation in the presence of rock fragments. *Catena* 92. 38–48.

Khaledian, Y., Kiani, F., Ebrahimi, S., Brevik, E.C. and Aitkenhead-Peterson, J. 2017. Assessment and monitoring of soil degradation during land use change using multivariate analysis. *Land Degradation and Development* 28. 128–141.

Kinnell, P.I.A. 2016. A review of the design and operation of runoff and soil loss plots. *Catena* 145. 257–265.

Kou, M., Jiao, J., Yin, Q., Wang, N., Wang, Z., Li, Y., Yu, W., Wei, Y., Yan, F. and Cao, B. 2016. Successional trajectory over 10 years of vegetation restoration of abandoned slope croplands in the Hill-Gully region of the Loess Plateau. *Land Degradation and Development* 27. 919–932.

Lasanta, T., Nadal-Romero, E. and Arnáez, J. 2015. Managing abandoned farmland to control the

impact of re-vegetation on the environment. The state of the art in Europe. *Environmental Science and Policy* 52. 99–109.

Lesschen, J.P., Cammeraat, L.H. and Nieman, T. 2008. Erosion and terrace failure due to agricultural land abandonment in a semi-arid environment. *Earth Surface Processes and Landforms* 33. 1574–1584.

Lopes, C., Monteiro, A., Rückert, F.E., Gruber, B., Steinberg, B. and Schultz, H.R. 2015. Transpiration of grapevines and co-habitating cover crop and weed species in a vineyard. A "snapshot" at diurnal trends. *VITIS – Journal of Grapevine Research* 43. 111.

López-Vicente, M., Quijano, L., Palazón, L., Gaspar, L. and Navas, A. 2015. Assessment of soil redistribution at catchment scale by coupling a soil erosion model and a sediment connectivity index (Central Spanish Pre-Pyrenees). *Cuadernos de Investigación Geográfica* 41. 127–147.

Marques, M.J., Bienes, R., Pérez-Rodríguez, R. and Jiménez, L. 2008. Soil degradation in central Spain due to sheet water erosion by low-intensity rainfall events. *Earth Surface Processes and Landforms* 33. 414–423.

Martínez-Casasnovas, J.A., Antón-Fernández, C. and Ramos, M.C. 2003. Sediment production in large gullies of the Mediterranean area (NE Spain) from high-resolution digital elevation models and geographical information systems analysis. *Earth Surface Processes and Landforms* 28. 443–456.

Martínez-Hernández, C., Rodrigo-Comino, J. and Romero-Díaz, A. 2017. Impact of lithology and soil properties on abandoned dryland terraces during the early stages of soil erosion by water in south-east Spain. *Hydrological Processes* 31. 3095–3109.

Morvan, X., Naisse, C., Malam Issa, O., Desprats, J.F., Combaud, A. and Cerdan, O. 2014. Effect of ground-cover type on surface runoff and subsequent soil erosion in Champagne vineyards in France. *Soil Use and Management* 30. 372–381.

Novák, T.J., Incze, J., Spohn, M., Glina, B. and Giani, L. 2014. Soil and vegetation transformation in abandoned vineyards of the Tokaj Nagy-Hill, Hungary. *Catena* 123. 88–98.

Novara, A., Keesstra, S., Cerdà, A., Pereira, P. and Gristina, L. 2016. Understanding the role of soil erosion on CO_2-c loss using 13c isotopic signatures in abandoned Mediterranean agricultural land. *Science of The Total Environment* 550. 330–336.

Nyssen, J., Haile, M., Poesen, J., Deckers, J. and Moeyersons, J. 2001. Removal of rock fragments and its effect on soil loss and crop yield, Tigray, Ethiopia. *Soil Use and Management* 17. 179–187.

O.I.V. 2017. *2017 World vitiviniculture situation*. Paris, International Organization of Vine and Wine.

Olmstead, M.A., Wample, R.L., Greene, S.L. and Tarara, J.M. 2001. Evaluation of Potential Cover Crops for Inland Pacific Northwest Vineyards. *American Journal of Enology and Viticulture* 52. 292–303.

Pahlavan-Rad, M.R. and Akbarimoghaddam, A. 2018. Spatial variability of soil texture fractions and pH in a flood plain (case study from eastern Iran). *Catena* 160. 275–281.

Poesen, J.W., van Wesemael, B., Bunte, K. and Benet, A.S. 1998. Variation of rock fragment cover and size along semiarid hillslopes: a case-study from southeast Spain. *Geomorphology* 23. 323–335.

Quiquerez, A., Chevigny, E., Allemand, P., Curmi, P., Petit, C. and Grandjean, P. 2014. Assessing the impact of soil surface characteristics on vineyard erosion from very high spatial resolution aerial images (Côte de Beaune, Burgundy, France). *Catena* 116. 163–172.

Richter, G. 1979. *Bodenerosion in Rebanlagen des Moselgebietes. Ergebnisse quantitativer Untersuchungen 1974–1977.* Trier, Forschungsstelle Bodenerosion der Universität Trier.

Ries, J.B., Seeger, M., Iserloh, T., Wistorf, S. and Fister, W. 2009. Calibration of simulated rainfall characteristics for the study of soil erosion on agricultural land. *Soil and Tillage Research* 106. 109–116.

Rodrigo Comino, J., Bogunovic, I., Mohajerani, H., Pereira, P., Cerdà, A., Ruiz Sinoga, J.D. and Ries, J.B. 2017. The impact of vineyard abandonment on soil properties and hydrological processes. *Vadose Zone Journal* 16. 12.

Rodrigo Comino, J., Iserloh, T., Morvan, X., Malam Issa, O., Naisse, C., Keesstra, S.D., Cerdà, A., Prosdocimi, M., Arnáez, J., Lasanta, T., Concepción Ramos, M., José Marqués, M., Ruiz Colmenero, M., Bienes, M., Ruiz Sinoga, J.D., Seeger, M. and Ries, B.J. 2016a. Soil erosion processes in European vineyards: A qualitative comparison of rainfall simulation measurements in Germany, Spain and France. *Hydrology* 3. (1): 6. https://doi.org/10.3390/hydrology3010006

Rodrigo Comino. J., Quiquerez, A., Follain, S., Raclot, D., Le Bissonnais, Y., Casalí, J., Giménez, R., Cerdà, A., Keesstra, S.D., Brevik, E.C., Pereira, P., Senciales, J.M., Seeger, M.,Ruiz Sinoga, J.D. and Ries, J.B. 2016b. Soil erosion in sloping vineyards assessed by using botanical indicators and sediment collectors in the Ruwer-Mosel valley. *Agriculture, Ecosystems & Environment* 233. 158–170.

Rodrigo-Comino, J., Brings, C., Lassu, T., Iserloh, T., Senciales, J., Martínez Murillo, J., Ruiz Sinoga, J., Seeger, M. and Ries, J. 2015. Rainfall and human activity impacts on soil losses and rill erosion in vineyards (Ruwer Valley, Germany). *Solid Earth* 6. 823–837.

Romero Díaz, A., Lasanta, T., Regües, D., Lana-Renault, N. and Cerdà, A. 2011. Hydrological response and sediment production under different land cover in abandoned farmland fields in a Mediterranean mountain environment. *Boletín de la Asociación de Geógrafos Españoles* 55. 303–323.

Romero Díaz, A., Marín Sanleandro, P., Sánchez Soriano, A., Belmonte Serrato, F. and Faulkner,

H. 2007. The causes of piping in a set of abandoned agricultural terraces in southeast Spain. *Catena* 69. 282–293.

Rosell, R.A., Gasparoni, J.C. and Galantini, J.A. 2001. Soil organic matter evaluation. In *Assessment methods for soil carbon.* Eds.: Lal, R., Kimble, J., Follet, R. and Stewart, B., US, Lewis Publishers, 311–322.

Saleh, A. 1993. Soil roughness measurement: Chain method. *Journal of Soil and Water Conservation* 48. 527–529.

Samani, A.N., Wasson, R.J., Rahdari, M.R. and Moeini, A. 2016. Quantifying eroding head cut detachment through flume experiments and hydraulic thresholds analysis. *Environmental Earth Sciences* 75. 14–24.

Seeger, M. and Ries, J.B. 2008. Soil degradation and soil surface process intensities on abandoned fields in Mediterranean mountain environments. *Land Degradation and Development* 19. 488–501.

Smith, P., Cotrufo, M.F., Rumpel, C., Paustian, K., Kuikman, P.J., Elliott, J.A., McDowell, R., Griffiths, R.I., Asakawa, S., Bustamante, M., House, J.I., Sobocká, J., Harper, R., Pan, G., West, P.C., Gerber, J.S., Clark, J.M., Adhya, T., Scholes, R.J. and Scholes, M.C. 2015. Biogeochemical cycles and biodiversity as key drivers of ecosystem services provided by soils. *SOIL* 1. 665–685.

Szabó, J., Jakab, G. and Szabó, B. 2015. Spatial and temporal heterogeneity of runoff and soil loss dynamics under simulated rainfall. *Hungarian Geographical Bulletin* 64. (1): 25–34.

Szalai, Z., Szabó, J., Kovács, J., Mészáros, E., Albert, G., Centeri, Cs., Szabó, B., Madarász, B., Zacháry, D. and Jakab, G. 2016. Redistribution of soil organic carbon triggered by erosion at field scale under subhumid climate, Hungary. *Pedosphere* 26. 652–665.

Taguas, E.V., Guzmán, E., Guzmán, G., Vanwalleghem, T. and Gómez, J.A. 2015. Characteristics and importance of rill and gully erosion: a case study in a small catchment of a marginal olive grove. *Cuadernos de Investigación Geográfica* 41. 107–126.

Verheijen, F.G.A., Jones, R.J.A., Rickson, R.J. and Smith, C.J. 2009. Tolerable versus actual soil erosion rates in Europe. *Earth-Science Reviews* 94. 23–38.

Zavala, L.M., Jordán, A., Bellinfante, N. and Gil, J. 2010. Relationships between rock fragment cover and soil hydrological response in a Mediterranean environment. *Soil Science and Plant Nutrition* 56. 95–104.

Ziegler, A.D., Giambelluca, T.W., Plondke, D., Leisz, S., Tran, L.T., Fox, J., Nullet, M.A., Vogler, J.B., Minh Troung, D. and Vien, T.D. 2007. Hydrological consequences of landscape fragmentation in mountainous northern Vietnam: Buffering of Hortonian overland flow. *Journal of Hydrology* 337. 52–67.

DOI: 10.15201/hungeobull.67.4.3 Hungarian Geographical Bulletin **67** 2018 (4) 333–341.

Quantification of nitrate fluxes to groundwater and rivers from different land use types

Dmytro DIADIN[1], Yuliya VYSTAVNA[1,2] and Yuri VERGELES[1]

Abstract

Nitrate enters aquatic systems from anthropogenic and natural sources affecting drinking water supply and surface water eutrophication. Conventional hydro-chemical measurements have been used together with the geographic information system (GIS) and stable isotopes techniques to track nitrate origin, sources distribution and quantify their fluxes from various land use types to ground and surface waters in East Ukraine. Average fluxes of nitrate in groundwater are estimated at 356 kg $year^{-1}$ km^{-2} from settlements (mostly rural), 214 kg $year^{-1}$ km^{-2} – from agricultural lands and 73 kg $year^{-1}$ km^{-2} – from forested areas. According to the mass balance estimation, nitrogen input (150 kg $year^{-1}$ km^{-2}) occurs mainly in the upper part of the Seversky Donets River basin and is attributed to the discharge of untreated municipal wastewater to rivers as well as groundwater contamination by leaking septic tanks and pit latrines from residential areas.

Keywords: land use; nitrate contamination; GIS; Ukraine; stable isotopes of nitrate

Introduction

Nitrate contamination of groundwater is a rising problem worldwide (Rivett, M.O. *et al.* 2008; WHO 2011) that affects drinking water supply and leads to eutrophication of surface water that, in turn, poses environmental and human health risks (Slomp, C.P. and Van Cappellen, P. 2004; WHO 2011). Nitrates can enter both ground and surface water from both anthropogenic (sewage leakages, storages of fertilizers, manures, landfills, runoff from crop fields) and natural (forest soil, wetlands) sources (Katz, B.G. *et al.* 2004; Lockhart, K.M. *et al.* 2013; Vystavna, Y. *et al.* 2017a). Besides, nutrient input from agricultural soils to runoff and further to water bodies may be facilitated by soil erosion processes (Göndöcs, J. *et al.* 2015). Depending on land use patterns, hydrogeological and denitrification conditions, nitrate ion in groundwater may be highly mobile and persistent (Ouyang, Y. 2012; Yakovlev, V. *et al.* 2015; Vystavna, Y. *et al.* 2017a, b). Conventional hydro-chemical measurements are not sufficient to track nitrate origin, sources distribution and to quantify fluxes from various land use types. However, additional tools such as the geographic information systems (GIS) allowing to define the land use types and drainage area for contaminants pathways (Vystavna, Y. *et al.* 2017a, b), and stable isotopes of nitrate allowing to identify dominant sources, can be useful to quantify nitrate fluxes in relation to the land use (Matiatos, I. 2016). Nitrate balance in soil-water systems together with other nutrient related indicators is also used for modelling ecosystem services (Makovníková, J. *et al.* 2017). While some studies (Ouyang, Y. 2012; Lockhart, K.M. *et al.* 2013; Matiatos, I. 2016) have provided good insights into nutrient loads from groundwater to surface waters, nitrate fluxes from different land use types are still poorly understood and rarely quantified.

[1] Department of Urban Environmental Engineering and Management, O.M. Beketov National University of Urban Economy in Kharkiv, 6 1002, Marshala Bazhanova Street 17, Kharkiv, Ukraine. E-mail: dmdyadin@gmail.com
[2] Biology Centre of the Czech Academy of Sciences, Institute of Hydrobiology, Na Sádkách 7, 370 05 České Budějovice, Czech Republic. E-mail: yuliya.vystavna@hbu.cas.cz

The combination of conventional hydrochemical measurements (water temperature, discharge, major ions) with stable isotopes of nitrate and the application of GIS tools has been used to quantify nitrate fluxes from the Kharkiv region in East Ukraine – the largest one, in a term of size and population. Our previous investigations (Vystavna, Y. *et al.* 2015; Yakovlev, V. *et al.* 2015; Vystavna, Y. *et al.* 2017a) have shown that about 20 per cent of groundwater samples taken in the Kharkiv region featured nitrate content at more than 50 mg L^{-1} being above a water quality limit established by the World Health Organization (WHO). The studied region is a relatively water scarce one and has very limited local runoff, with substantial part of water supply provided by water translocation from other Ukraine's river basins via artificial canals (Vystavna, Y. and Diadin, D. 2015). Additionally, the region is considered to be influenced by climate change that results in the deterioration of water quality (Pidlisnyuk, V. *et al.* 2016). Thus, knowledge on water contamination sources is urgently needed to secure drinking water supply, prevent the spread of water-related diseases and protect natural aquatic systems (Canter, L.W. 1997; WHO 2008). The objective of the study was to quantify nitrate fluxes from settlements, forested and agricultural land use to ground and surface waters used for drinking water supply applying geochemical, stable isotopes of nitrates and GIS-based techniques.

Methods and materials

Study area

The area of the studied Kharkiv region of Ukraine (the population is ca. 2.7 million) is 31,415 km^2. The dominant land use type is agricultural lands that cover 77 per cent of the territory. The rest of the area is covered by forests, dense urban, sub-urban and rural settlements *(Figure 1).* Grey forest and thick grassland soils, called chernozems, dominate in the study area (Vystavna, Y. *et al.* 2012) with average water infiltration rate about 2 m $year^{-1}$ (Maslov, B.S. 2009).

Agricultural lands include arable areas (62% of agricultural lands area), grasslands and pastures (27%) and others (orchards, farms, etc.) (ECO, 2015). Inorganic fertilizers, mainly nitrogen, are applied at the rate of ca. 50 kg ha^{-1}, where most of them (88%) are three-component (N, P and K) and ammonium nitrate fertilizers (Vystavna, Y. *et al.* 2017a). Manure is applied as an organic fertilizer at small farms and individual rural and suburban households. About 90 per cent of wastewater in rural area is not treated and disposed to poor-isolated septic tanks and pit latrines from which it seeps into shallow groundwater (Vystavna, Y. *et al.* 2017a, b).

Approximately 2/3 of the study area belongs to the Seversky Donets catchment, the Sea of Azov's basin, and the rest 1/3 belongs to the Dnipro River basin. The Seversky Donets River is a transboundary Russian Federation/Ukraine waterway with an average annual discharge of 12 m^3 s^{-1} (upstream of the Kharkiv region). The Seversky Donets River is used for drinking and industrial water supply, but also for receiving both treated and untreated wastewaters. Untreated wastewater discharge is about 4–5 mln^3 $year^{-1}$ (ECO 2015). The river is of alluvial character and is mainly recharged by precipitation (65%) and groundwater (up to 35%). Within the study area, the surficial geology up to 120 m of depth is composed mostly of permeable and loose sedimentary materials – sands, loams and clayey loams of Quaternary, Neogene and Paleogene ages. The presence of hydraulic connectivity between surficial Quaternary aquifers including the Upper Cretaceous formations results in a wide range of groundwater geochemistry (Vystavna, Y. *et al.* 2015). The shallow aquifer lies at 5–30 m below the surface in sands or sandy loams and outflows as numerous springs to main and tributary river valleys. High share of groundwater in the river's recharge indicates that surface water chemistry is influenced by chemical fluxes from the shallow aquifer (Yakovlev, V. *et al.* 2015).

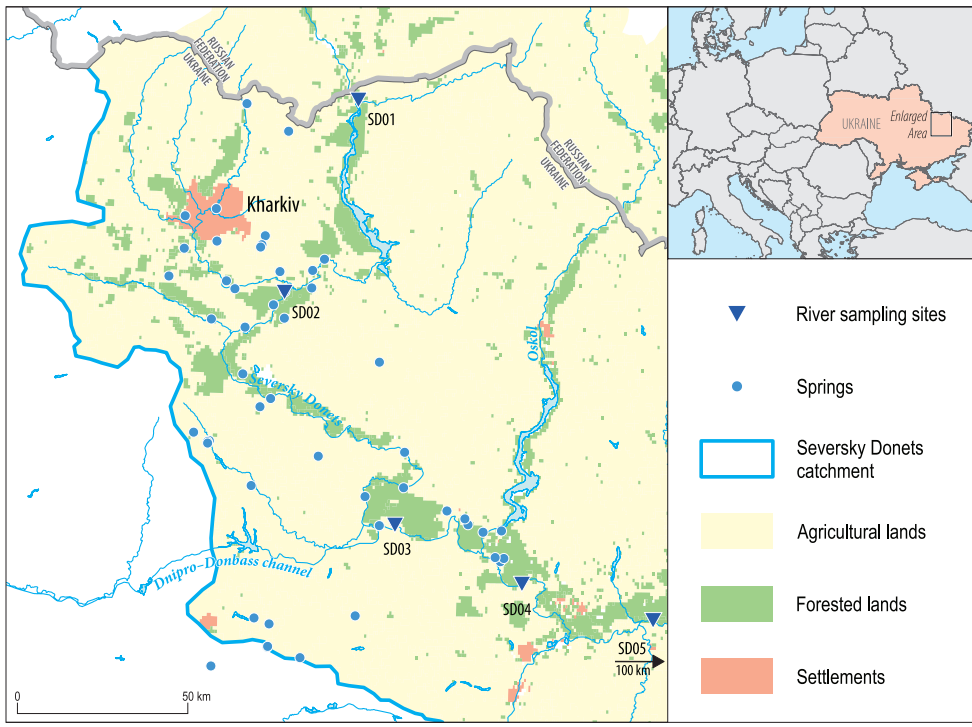


Fig. 1. The study sites location and major land uses in the Kharkiv region, East Ukraine

Sampling and analysis

In total, 50 groundwater-driven springs and 4 surface water sites were sampled in the studied area. Groundwater discharge rates were measured with a stopwatch and a calibrated container during the sampling. Water temperature, pH, electrical conductivity (EC) and redox potential (ORP) were measured in the field applying HI-98130 Multi-parameter tester and ORP (Redox) Tester 98121 (Hanna Instruments ®). Major ions (calcium, hydrocarbonate, magnesium, chloride, sulphate, sodium, potassium and nitrate) were analysed using the potentiometric method (described in Yakovlev, V. *et al.* 2015, and Vystavna, Y. *et al.* 2017a). In order to trace dominant nitrate sources in the study area, 8 springs selected from different land use types were additionally sampled on stable isotopes of nitrate. For this purpose, water samples were filtered (0.22 μm) in the field and transferred into 50 mL plastic bottles. Nitrogen ($\delta^{15}N–NO_3$) and oxygen ($\delta^{18}O–NO_3$) isotopes of nitrate were analysed using the denitrifying method (described in Vystavna, Y. *et al.* 2017a). N and O isotope ratios were reported as δ values as a part per thousand (‰) deviations relative to the standards: atmospheric N_2 (AIR) for nitrogen and Vienna Standard Mean Ocean Water (V-SMOW) for oxygen:

$$\delta\ (‰) = [(R_{sample} – R_{standard})/R_{standard}] \times 1000, \quad (1)$$

where R is the ratio of the heavy to light isotopes as $^{15}N/^{14}N$ or $^{18}O/^{16}O$.

Data treatment and nitrate fluxes estimation

The GIS software (ArcGIS 10.2.2) was used to delineate catchment area of each sampling site basing on digital elevation model (Rawat,

K.S and Singh, S.K. 2018). Land cover types were defined from the Global Land Cover database of Europe (GLC 2000). The groundwater sampling sites were grouped according to the sub-catchments of the Seversky Donets River: SD01–SD02 represents the upper part of the basin, from the entrance to Ukraine and before the conjunction with the Udy River (flowing through the Kharkiv metropolitan area); SD02–SD03 represents the area downstream the city of Kharkiv, with treated wastewater discharges and upstream inflow of the artificial canal that transfers water masses from the Dnipro River basin; SD03–SD04 is a part of basin that lies between the inflow of the Dnipro-Donbas artificial canal and the conjunction with the Oskol River at the border of the Kharkiv and Donetsk regions; SD04–SD05 is a lower part of the studied basin, downstream the conjunction with the Oskol River *(see Figure 1).*

The studied groundwater springs were classified into three groups according to dominant land use types in their catchment areas: 16 groundwater-driven springs with drainage areas within the agricultural lands (AG), 18 springs – within the forest areas (FR) and 16 springs – within the settlements (SE). Software package Golden Software Grapher 12.0 was applied for data plotting.

Nitrate fluxes from a ground to surface waters were estimated using the following equation (Yakovlev, V. *et al.* 2015):

$$F_{NO3} = Q \times C \times K/S, \quad (2)$$

where, F_{NO3} is an annual flux of nitrate, tkm^{-2} y^{-1}, Q is a measured groundwater discharge, Ls^{-1}, C is the analysed nitrate concentration, mgL^{-1}, K is seconds per year (31,536,000 s), and S is an estimated drainage area, km^2, following the equation:

$$S = Q/M, \quad (3)$$

where M is the groundwater flow module, projected according to the average river flow rate of 90 per cent of water availability (M = 1.25 Ls^{-1}) (Yakovlev, V. *et al.* 2015).

Specific fluxes of nitrate were estimated for the springs grouped according to the dominant land use in their drainage areas. The values of specific fluxes were calculated as average values for the group of springs in each land use type. Fluxes for the nitrate nitrogen were calculated according to the molecular ratio (14/62) and normalized by the sub-catchment area.

Results and discussions

General chemistry of groundwater

Among 50 studied springs, more than 60 per cent had groundwater discharge less than 0.5 Ls^{-1} *(Figure 2*, a*)*. Groundwater temperature, EC and ORP were highly variable *(Figure 2*, b*)*

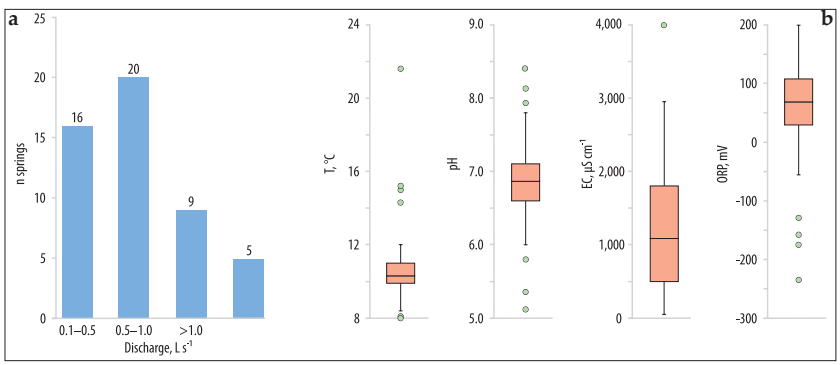


Fig. 2. Distribution chart of measured groundwater discharge rates, Ls^{-1} (a), and box plots of measured hydro-chemical parameters of groundwater (b)

indicating different hydrogeological and land use conditions (Vystavna, Y. *et al.* 2015).

Major ion composition of groundwater was significantly varying, too, with the prevalence of Ca–HCO_3 (46%) and Na–Ca Cl–SO_4 (42%) groundwater types *(Figure 3*, a*)*.

The studied sub-catchments clear differed by their hydrochemistry *(Figure 3)* with prevalence of Ca–HCO_3-type groundwater in the upper SD01–SD02 and Ca–Cl–SO_4-type groundwater in the lower SD04–SD05 sub-basins. However, in the SD02–SD03 and SD03–SD04 sub-catchments, both sulphate- and bicarbonate-dominated groundwater types were found *(Figure 3*, b*)*.

Among cations, calcium dominated in the majority of samples (86%), while sodium dominated only in 14 per cent of samples *(Figure 3*, a*)*. Among anions, the gradual increase of Cl^- and SO_4^{2-} was detected from upstream to downstream sub-catchments *(Figure 3*, b*)*. This can be attributed to the lithology of vadose zone and aquifers, but also to the continuous discharge of wastewaters. The Ca–HCO_3 groundwater type is consistent with the lithology nature of the river beds and explains water-rock interaction with the Cretaceous carbonate rocks (Vystavna, Y. *et al.* 2015). At the lower part of the water basin, Cl^- and SO_4^{2-} composition of groundwater can be associated with natural (interaction with halite and gypsum inclusions in clays) and anthropogenic factors (wastewater discharges). Previous estimations of the groundwater saturation index within the study basin indicated its under-saturation with halite confirming the minor role of this geological structure in the groundwater chemistry (Vystavna, Y. *et al.* 2015). Besides, upward recharge from deeper salty Triassic aquifers would be an additional source of chlorides in this area. In southern part of the region, sulphate minerals (such as gypsum and anhydrite) were found in the uppermost layers of Quaternary deposits and could be suggested an additional source for this anion in groundwater (Vystavna, Y. *et al.* 2015). However, the water-rock interaction and upward recharge are needed to be additionally studied. Similarly, to springs, Cl^- increases downstream the Seversky Donets River together with rising of total dissolve solids (TDS) values *(Figure 3*, a*)*.

Land use and dominant sources of nitrate

GIS-analysis of Global Land-Cover datasets indicated three dominating land use types in the studied basin: agricultural lands (including arable lands, pastures and grasslands), forested areas and settlements (both rural and urban) (*Figure 1*). Measuring contents of stable isotopes of nitrate in groundwater from settlements, agricultural and forested areas revealed that manure and sewage were dominant sources of nitrate in the Seversky Donets basin *(Figure 4)*, in the agreement with our previous study (Vystavna, Y. *et al.* 2017a).

Domination of manure and sewage among other groundwater pollution sources is a result of improper wastewater treatment in rural settlements and excessive application of organic fertilizers at domestic kitchen-gardens and farmlands. Only few rural settlements are supplied with centralized sewerage, and construction of permeable pit latrines instead is still a common practice in rural areas (Skryzhevska, Y. and Karácsonyi, D. 2012).

Nitrate fluxes in groundwater and rivers

Highest-rate and most variable nitrogen fluxes were calculated for settlement areas while the lowest and the most stable values were obtained for forested lands *(Figure 5).*

Nitrate fluxes in settlements were derived from such point sources as pit latrines, private gardens and livestock sheds. Sewage and polluted runoff infiltrates unevenly, but poses high nitrate pollution in shallow ground waters (Bermudez-Couso, A. *et al.* 2013). Besides, the surface permeability is highly variable within the study area (Yakovlev, V. *et al.* 2015; Vystavna, Y. *et al.* 2015) leading to sewage percolation from pit latrines directly to the aquifer passing

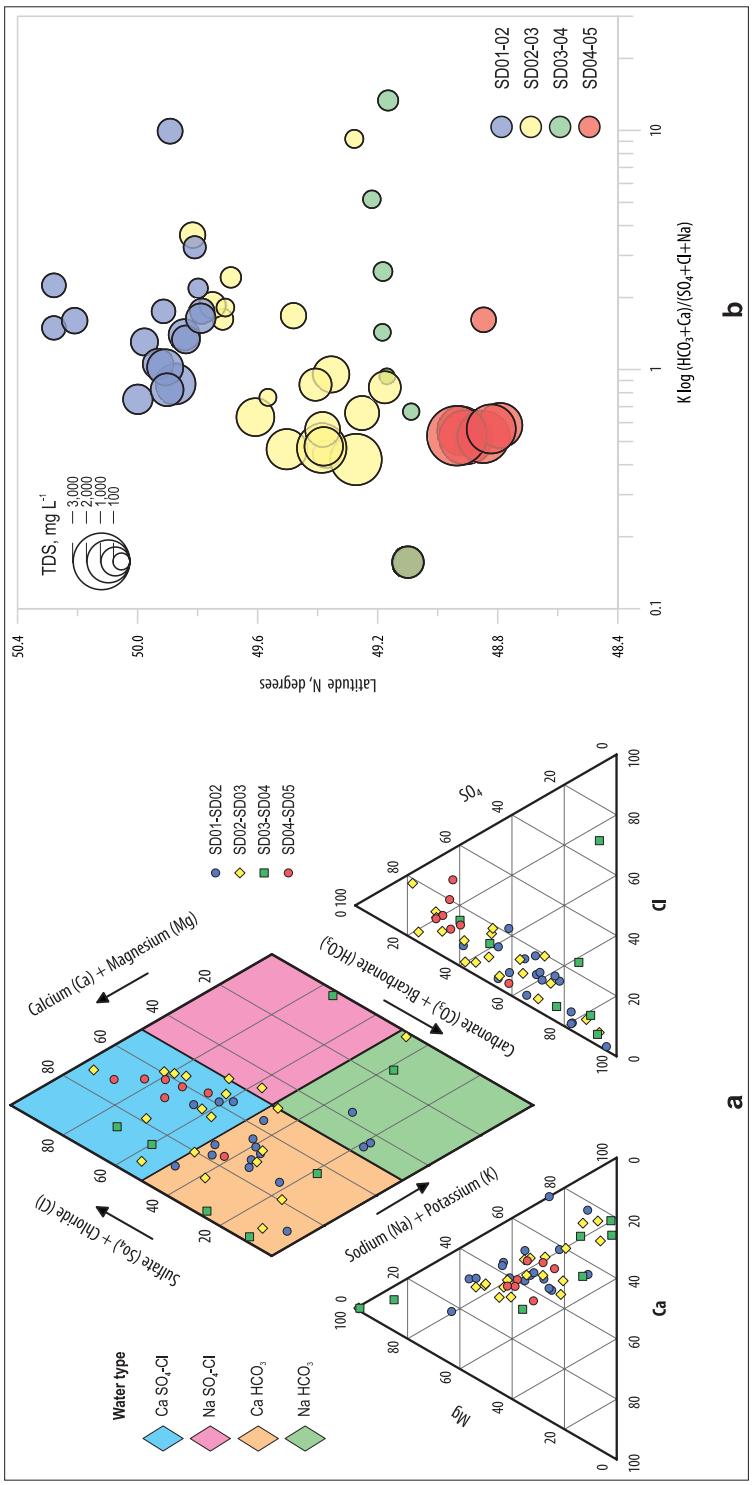


Fig. 3. The Piper diagram for major ions contents in groundwater samples (a), and latitudinal distribution of major ions in groundwater of the studied area (b)

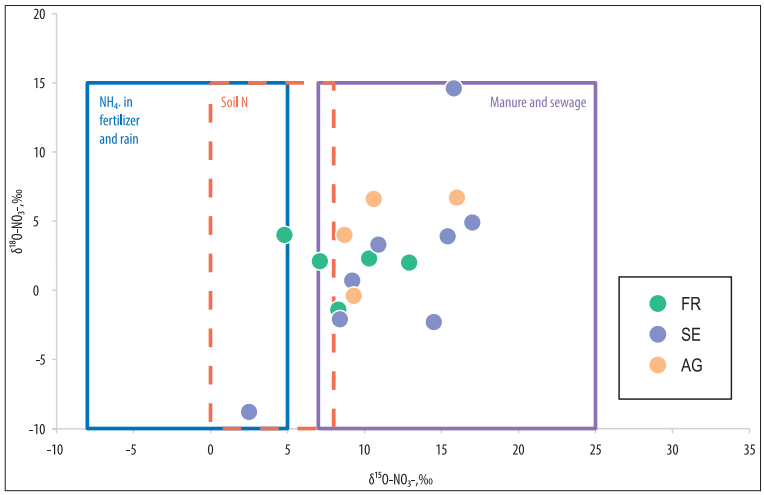


Fig. 4. Isotopic signature of nitrate in selected groundwater of agricultural (AG), forest (FR) and settlement (SE) land uses in the Kharkiv region

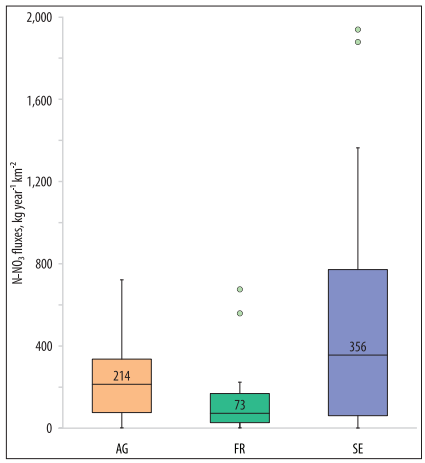


Fig. 5. Annual nitrogen (N-NO_3) fluxes from agricultural (AG), forest (FR) and settlement (SE) land uses (median values are shown on boxplots)

by soil layer and vadose zone. In some cases, nitrogen is uptaken by plants, and bacterial denitrification took place; the both processes influence nitrate variations in groundwater (Vystavna, Y. *et al.* 2017b). The same mechanisms of intensive uptake of nitrogen by crops as well as denitrification processes explain the less amount of nitrate in groundwater in agricultural areas. Among three types of land use, forested area featured the lowest values of nitrate fluxes and less variation of nitrate concentration *(see Figure 5)* that can be explained by combination of denitrification, deposition and plant uptake processes in less impacted by anthropogenic activity ecosystem (Teller, A. *et al.* 2012).

The composition of surface water in the Seversky Donets River reflects the overall distribution of nitrate in the basin brought by groundwater, runoff and wastewater discharge *(Table 1).*

Calculated mass balance of substances (chloride, nitrate-N, potassium and TDS) indicated that nitrogen input occurs mainly in the upper part of the basin reaching at 150 kg $year^{-1}$ km^{-2} while depletion of nitrate concentrations and decrease of nitrate fluxes occur downstream. The negative values of nitrogen

Table 1. Chloride, nitrate and potassium concentrations and discharge rates in the Seversky Donets River

Site	n	Average concentration, mg/L^{-1}				Water discharge, m^3 s^{-1}	Substances discharge, t $year^{-1}$			
		Cl^-	NO_3^-	K^+	TDS		Cl^-	N-NO_3^-	K^+	TDS
SD01	11	58±24	17.4±26.8	10.6±9.5	719±145	12	22,063	1,511	4,011	272,055
SD02	6	58±5	16.5±17.7	7.6±7.9	765±281	24	43,671	2,866	5,752	578,623
SD03	6	75±13	7.0±4.6	7.0±5.9	922±85	42	99,338	2,128	9,272	1,221,333
SD04	8	70±15	2.8±2.3	6.9±4.6	795±86	80	175,592	1,621	17,408	2,006,446
SD05	2	187±36	5.5±1.5	n/m	1,227±83	62	367,592	2,476	n/m	2,406,416

in the middle part of the basin indicate the potential retention of nutrients (sedimentation, denitrification, plants uptake) in the ecosystem *(Table 2)*.

Table 2. Estimated fluxes to river water from sub-catchments

Sub-catchment	S, km^2	Estimated fluxes, t year^{-1} km^{-2}			
		Cl^-	$N\text{-}NO_3^-$	K^+	TDS
SD01–SD02	9,013	2.40	0.150	0.193	34.0
SD02–SD03	7,943	7.01	– 0.093	0.443	80.9
SD03–SD04	10,663	7.15	– 0.048	0.763	73.6
SD04–SD05	15,721	12.20	0.054	n/m	25.4

Beside regional lithology differences, such a distribution may be caused by significant anthropogenic inputs of nitrate and other compounds to the river. These are discharges of treated municipal wastewater from the Kharkiv metropolitan area (SD01–SD02) and inflow of salty chloride-sulphate water from the Dnipro River via the artificial canal (SD02–SD03 and SD03–SD04) *(Figure 1)*. The increase of Cl^- from upstream to downstream can be as well explained by inflow of chloride-contained wastewaters and contaminated groundwater into the surface waters. Along with that, the average concentrations of nitrate and potassium are less at downstream sites (SD03–SD05) than upstream the basin (SD01–SD02) *(see Table 1)*.

Conclusions

Nitrate fluxes well reflected general land use structure in the Seversky Donets basin with the highest values of nitrate concentration in groundwater and surface water in settlements and agricultural lands. The application of stable isotopes of nitrate revealed the input from sewage and manure application as the most important sources of nitrate in the watershed that is used for drinking water supply. Average estimated fluxes of nitrate in groundwater are 356 kg year^{-1} km^{-2} from settlements (mostly rural) and 214 kg year^{-1} km^{-2} from agricultural lands. Nitrate fluxes from forested lands are estimated at 73 kg year^{-1} km^{-2} and derived mostly from soil nitrification and – to a less extend – from scattered households located there. Domination of manure and sewage among other groundwater pollution sources as shown by applying the stable isotope techniques is viewed as a result of improper wastewater treatment in rural settlements and excessive application of organic fertilizers at private gardens and farmlands.

***Acknowledgements*:** The research has been carried out in the framework of the Research Project CRP F33021 "Evaluation of human impacts on water balance and nutrients dynamics in the transboundary Russia/Ukraine river basin" funded partly by the International Atomic Energy Agency (IAEA) and the O.M. Beketov National University of Urban Economy in Kharkiv (2014–2017), with partial financial support from the Shell Exploration and Production Ukraine Investments (IV) B.V. through the Donation Agreement UI55229 for studying and improvement of rural people access to clean groundwater sources in 8 districts of the Kharkiv region (2016).

REFERENCES

Bermudez-Couso, A., Fernandez-Calvino, D., Alvarez-Enjo, M.A., Simal-Gandara, J., Novoa-Munoz, J.C. and Arias-Estevez, M. 2013. Pollution of surface waters by metalaxyl and nitrate from non-point sources. *Science of the Total Environment* 461–462. 282–289.

Canter, L.W. 1997. *Nitrates in Groundwater*. New York, CRC Press Inc.

ECO 2015. *Ecological and Environmental Passport of the Kharkiv region, 2015*. Kiev, Ministry of the Environmental Protection of Ukraine. (in Ukrainian).

GLC 2000. *Global land cover (GLC) 2000 – Europe*. European Environment Agency. URL: https://www.eea.europa.eu/data-and-maps/data/global-land-cover-2000-europe

Göndöcs, J., Breuer, H., Horváth, Á., Ács, F. and Rajkai, K. 2015. Numerical study of the eff ect of soil texture and land use distribution on the convective precipitation. *Hungarian Geographical Bulletin* 64. (1): 3–15.

Katz, B.G., Chelette, A.R. and Pratt, T.R. 2004. Use of chemical and isotopic tracers to assess nitrate contamination and groundwater age, Woodville karst Plain, USA. *Journal of Hydrology* 289. 36–61.

Lockhart, K.M., King, A.M. and Harter, T. 2013. Identifying sources of groundwater nitrate contam-

ination in a large alluvial groundwater basin with highly diversified intensive agricultural production. *Journal of Contaminant Hydrology* 151. 140–154.

Makovníková, J., Kanianska, R. and Kizeková, M. 2017. The ecosystem services supplied by soil in relation to land use. *Hungarian Geographical Bulletin* 66. (1): 37–42.

Maslov, B.S. 2009. *Agricultural land improvement. Amelioration and reclamation.* Vol. II. EOLSS Publications.

Matiatos, I. 2016. Nitrate source identification in groundwater of multiple land-use areas by combining isotopes and multivariate statistical analysis: A case study of Asopos basin (Central Greece). *Science of The Total Environment* 541. 802–814.

Ouyang, Y. 2012. Estimation of shallow groundwater discharge and nutrient load into a river. *Ecological Engineering* 38. (1): 101–104.

Pidlisnyuk, V., John Harrington, J.R., Melnyk, Y., Vystavna, Y. 2016. Fluctuations of annual precipitation and water resources quality in Ukraine. *Chemistry and Chemical Technology* 10. (4): 621–629.

Rawat, K.S. and Singh, S.K. 2018. Water Quality Indices and GIS-based evaluation of a decadal groundwater quality. *Geology, Ecology, and Landscapes* 2. (4): 240–255. https://doi.org/10.1080/24749508.2018.1452462

Rivett, M.O., Buss, S.R., Morgan, P., Smith, J.W.N. and Bemment, C.D. 2008. Nitrate attenuation in groundwater: a review of biogeochemical controlling processes. *Water Research* 42. (16): 4215–4232.

Skryzhevska, Y. and Karácsonyi, D. 2012. Rural population in Ukraine: assessing reality, looking for revitalization. *Hungarian Geographical Bulletin* 61. (1): 49–78.

Slomp, C.P. and Van Cappellen, P. 2004. Nutrient inputs to the coastal ocean through submarine groundwater discharge: controls and potential impact. *Journal of Hydrology* 295. (1–4): 64–86.

Teller, A., Mathy, P. and Jeffers, J.N.R. 2012. *Responses of forest ecosystems to environmental changes.* Springer Science and Business Media.

Vystavna, Y. and Diadin, D. 2015. Water scarcity and contamination in Eastern Ukraine. *IAHS-AISH Proceedings and Reports* 366. 149–150.

Vystavna, Y., Diadin, D., Grynenko, V., Yakovlev, V., Vergeles, Y., Huneau, F., Rossi, P., Hejzlar, J. and Knöller, K. 2017a. Determination of dominant sources of nitrate contamination in transboundary (Russia/Ukraine) catchment with heterogeneous land use. *Environmental Monitoring and Assessment* 189(10):509. DOI: 10.1007/s10661-017-6227-5

Vystavna, Y., Diadin, D., Yakovlev, V., Hejzlar, J., Vadillo, I., Huneau, F. and Lehmann, M.F. 2017b. Nitrate contamination in a shallow urban aquifer in East Ukraine: Evidence from hydrochemical, stable nitrate isotope, and land use analysis. *Environmental Earth Sciences* 76(13):463. DOI: 10.1007/s12665-017-6796-1

Vystavna, Y., Huneau, F., Schafer, J., Motelica-Heino, M., Blanc, G., Larrose, A., Vergeles, Y., Diadin, D. and Le Coustumer, P. 2012. Distribution of trace elements in waters and sediments of the Seversky Donets transboundary watershed (Kharkiv region, Eastern Ukraine). *Applied Geochemistry* 27. (10): 2077–2087.

Vystavna, Y., Yakovlev, V., Diadin, D., Vergeles, Y. and Stolberg F. 2015. Hydrochemical characteristics and water quality assessment of surface and ground waters in the transboundary (Russia/Ukraine) Seversky Donets basin. *Environmental Earth Sciences* 74. (1): 585–596.

WHO 2008. *Guidelines for drinking water quality*. 3rd edition. World Health Organization. Geneva, WHO Press.

WHO 2011. *Nitrate and nitrite in drinking water. Background document for development of WHO guidelines for drinking water quality*. World Health Organization. Geneva, WHO Press.

Yakovlev, V., Vystavna, Y., Diadin, D. and Vergeles, Y. 2015. Nitrates in springs and rivers of East Ukraine: Distribution, contamination and fluxes. *Applied Geochemistry* 53. 71–78.

DOI: 10.15201/hungeobull.67.4.4 Hungarian Geographical Bulletin **67** 2018 (4) 343–359.

Analysis of shallow turbulent flows using the Hilbert-Huang transform: a tool for exploring the characteristics of turbulence and coherent flow structures

Kory M. KONSOER[1] and Bruce RHOADS[2]

Abstract

The Hilbert-Huang transform (HHT) is a method of spectral analysis that is suitable for application to non-stationary and non-linear signals that holds enormous potential for the analysis of turbulent flows in fluvial, aaeolian, and coastal systems. HHT begins with decomposition of the signal into Intrinsic Mode Functions (IMFs) using the Empirical Mode Decomposition method. A Hilbert transform is then applied to each IMF, enabling the calculation of the local spectral characteristics of the signal. Four applications of the HHT are used to demonstrate the utility of this method for spectral analysis of turbulent flows. The method is applied to: (1) velocity measurements of unidirectional flow with high suspended sediment concentration (laboratory), (2) velocity measurements from a combined uni-i-direction and wave flow over a mobile, evolving bed (laboratory), and (3) temperature measurements from the mixing interface of a large river confluence (field). Comparisons among HHT, Fourier, and wavelet analysis are provided, and we identify a number of major benefits of HHT based on these four applications. The results presented show that the spectral method of HHT provides a very useful tool for analysis of turbulence in natural flows and can greatly enhance signal analysis in addition to traditional methods such as Fourier and wavelet analysis.

Keywords: turbulence, spectral analysis, Hilbert-Huang Transform, signal processing, river flow

Introduction

The characteristics of fluid turbulence in rivers, estuaries, coastal environments and the atmosphere are a primary focus of the study of Earth surface processes. Characterization of coherent turbulent structures is of considerable interest because these structures govern the exchange of energy and momentum over a wide range of spatial and temporal scales, thereby influencing sediment transport rates, patterns of erosion and deposition, rates of fluid and constituent mixing, and interactions between flow structure and vegetation (e.g. Niño, Y. and Garcia, M.H. 1996; Buffin-Bélanger, T. *et al*. 2000; Zedler, E. and Street, R. 2001; Cellino, M. and Lemmin, U. 2004; Wu, F. and Yang, K. 2004; Singh, A. *et al*. 2012; Nepf, H. *et al.* 2013; Venditti, J.G. *et al*. 2013; Lewis, Q.W. and Rhoads, B.L. 2015). Thus, coherent structures play an important role in the morphologic evolution of fluvial, aeolian, and coastal systems.

Analysis of velocity time series has become an important tool in attempts to identify coherent structures in turbulent flows and to characterize the properties of these structures. Fourier analysis, which yields power spectra illustrating the distribution of turbulent energy associated with different frequencies, is a common method of analysing velocity signals for natural turbulent flows (Nikora, V.I.

[1] Department of Geography and Anthropology, Louisiana State University; Coastal Studies Institute, Louisiana State University, 227 Howe-Russell, Baton Rouge, LA 70803 USA. E-mail: konsoer1@illinois.edu

[2] Departments of Geography and Geographic Information, Science, Geology, and Ven Te Chow Hydrosystems Laboratory, University of Illinois at Urbana-Champaign, 1301 W Green Street, Natural History Building, Second Floor; MC-150, Urbana, IL 61801 USA. E-mail: brhoads@illinois.edu

and Smart, G.M. 1997; Uijttewaal, W.S.L. and Tukker, J. 1998; Sukhodolov, A. and Rhoads, B. 2001; Rhoads, B. and Sukhodolov, A. 2004; Singh, A. *et al.* 2010). More recently, wavelet analysis, which yields amplitude-frequency-time spectra, has been used to study the time-dependent behaviour of turbulent flows (Chen, J. and Hu, F. 2003; Hardy, R.J. *et al*. 2009; Wu, X. *et al*. 2013). Applications of these two methods have greatly enhanced our understanding of turbulence by identifying important frequency scales of velocity fluctuations in different types of turbulent flows and by providing empirical evidence in support of theoretical analyses, such as energy cascade theory (Kolmogorov, A.N. 1941; Sukhodolov, A. 1998; Sukhodolov, A. and Rhoads, B. 2001; Rhoads, B. and Sukhodolov, A. 2004).

Despite these successes, Fourier and wavelet analyses are subject to limitations. To be suitable for Fourier analysis signals must be both stationary and linear, while for wavelet analysis signals must be linear (Huang, N.E. *et al.* 1998). Moreover, both methods decompose the velocity signal using pre-defined waveforms. In the case of Fourier analysis, signals are decomposed into sine waves. If the fluctuations in a velocity signal are not well described by sine waves, this approach can spread energy over spurious harmonics that do not necessarily reflect real, physical motions of the fluid. Fourier analysis, since it treats sine-wave oscillations over the entire velocity signal, also cannot capture well (if at all) localized fluctuations in a signal produced by turbulent events that occur only for a short duration of the total signal length. Wavelet analysis, by yielding amplitude-frequency-time spectra, can capture events localized in time; however, the waveform used to analyse the signal is pre-defined by the user. Common forms include the Morlet, Gaussian, or Haar wavelets. Because results depend on the type of waveform selected, analysis of a signal using this method is not unique, complicating comparisons among studies.

To address the limitations of Fourier and wavelet analysis, Huang, N.E. *et al.* (1998) proposed a new method for spectral analysis of nonlinear and non-stationary data. The cornerstone of this method is Empirical Mode Decomposition (EMD), which decomposes a signal into a finite number of modes, termed Intrinsic Mode Functions (IMFs), using the local timescales present in the signal itself. An important attribute of EMD is that the resulting IMFs have physically meaningful Hilbert transforms. The set of Hilbert transforms of the IMFs is used to generate the amplitude-frequency-time distribution of a signal. This distribution may then be integrated over time to yield the power spectrum of frequency. It is therefore possible to investigate the frequency spectrum for the entire signal as well as the time-evolution of the frequency spectrum.

When Huang, N.E. *et al.* (1998) first proposed the Hilbert-Huang transform (HHT) method, they used a time series of wind speed to demonstrate its capabilities. Since then, the Hilbert-Huang method has been used to investigate periodicities in many different types of data, including time series of rainfall, lake temperature, and wind speed (e.g. Rao, A.R. and Hsu, E.C. 2010), stream gage data (e.g. Huang, Y. *et al*. 2009; Rao, A.R. and Hsu, E.C. 2010), ground motion during earthquakes (Loh, C. *et al*. 2000, 2001), and oxygen isotope data from ice cores (Thomas, E.R. *et al.* 2009). Application of the HHT method to the study of ocean waves has been especially useful because such waves are non-linear, which limits the applicability of Fourier and wavelet analysis (Huang, N.E. and Wu, Z. 2008). Additionally, efforts have been made to establish confidence limits for the EMD (Huang, N.E. *et al.* 2003) and to determine the nature of noise in the EMD (Flandrin, P. *et al.* 2004, 2005; Flandrin, P. and Gonçalves, P. 2004; Wu, Z. and Huang, N.E. 2004, 2005). A full summary of previous applications and advances of the HHT method is beyond the scope of this paper; a more complete description can be found in Huang, N.E. and Wu, Z. (2008).

Although the HHT method has been successfully applied in a wide variety of fields, applications of the HHT method to turbulent shear flows, such as those that can occur in riv-

ers, streams, estuaries and coastal zones, are rare and limited in scope (e.g. Schmitt, F.G. *et al.* 2009; Kanani, A. *et al.* 2010). Given its robust capacity to deal with nonlinear, nonstationary signals, the HHT method has enormous potential to improve our understanding of turbulent shear flows. The purpose of this paper is to illustrate the value of the HHT method for analysing turbulence in water flows through applications including: (1) velocity measurements of unidirectional flow over a flat sand bed below the threshold for sediment motion (laboratory), (2) velocity measurements of unidirectional flow with high suspended sediment concentration (laboratory), (3) velocity measurements from a combined flow (e.g. a flow with both a unidirectional and oscillatory component of velocity) over a mobile, evolving bed (laboratory), and (4) temperature measurements from the shear layer of a large river confluence (field). These examples were specifically chosen to highlight the advantages of the HHT method. The applications show that the method can be used to identify intermittency in an otherwise stationary velocity or temperature signal, analyse time-dependent flows, and capture fluctuations in a signal that are not sinusoidal. EMD also provides a means of noise-removal and signal de-trending. The HHT method is *not* a replacement for Fourier or wavelet analysis, but it can provide useful additional information on the character of turbulent flows when used in conjunction with those methods.

Methods

The Hilbert-Huang transform (HHT) determines the *local* frequency spectrum (i.e. the frequency spectrum at each individual time step) of a signal by applying a Hilbert transform to a complete set of locally orthogonal modes, or intrinsic mode functions (IMFs), of the signal. The HHT method consists of two main steps. First, the original signal is decomposed into a set of IMFs using a method termed Empirical Mode Decomposition (EMD). Next, a Hilbert transform is applied to each IMF and the frequency spectrum of the signal is calculated.

Empirical Mode Decomposition

The method of EMD is used to transform the signal into a set of IMFs (Huang, N.E. *et al.* 1998), which satisfy the following two conditions: (1) the number of extrema (maxima or minima) in the IMF must be equal to the number of zero crossings (e.g. where the signal crosses the x-axis) or differ by one, and (2) the average of the envelope curve of the IMF maxima and the envelope curve of the IMF minima must equal zero at each time step. The HHT method calculates these modes from the signal itself, with no waveform defined a priori. Each IMF represents characteristic oscillations over a narrow range of frequencies, which occur locally, intermittently, or persistently throughout the sample length of the signal. The decomposition procedure involves several steps (for full details of the method, see the original work by Huang, N.E. *et al.* 1998):

1. Create upper and lower envelope curves $e_{max}(t)$ and $e_{min}(t)$ for the signal by fitting cubic splines to the local maxima and minima in the time series.
2. Average the values of the upper and lower envelope curves at each time to get $m_1(t)$ (Eqn. 1).

$$m_1(t) = \frac{e_{max}(t)+e_{min}(t)}{2} \quad (1)$$

3. Subtract $m_1(t)$ from the signal $x(t)$ to get $h_1(t)$ (Eqn. 2).

$$h_1(t) = x(t) - m_1(t) \quad (2)$$

Check to see if $h_1(t)$ satisfies the two conditions to be an IMF. If $h_1(t)$ satisfies the requirements, it is $IMF_1(t)$. If not, treat $h_1(t)$ as the new signal and repeat steps 1–3 for "j" iterations until if $h_{1j}(t)$ is an IMF. This process is referred to as "sifting." Although satisfying the first condition is absolutely necessary for an IMF, if iterations are carried to an extreme to satisfy

the second condition, physically meaningful amplitude fluctuations can be obliterated, resulting in a pure frequency modulated signal of constant amplitude (Huang, N.E. *et al.* 1998). It is therefore necessary to introduce a stopping criterion for the sifting process. This criterion involves imposing a constraint on the magnitude of the standard deviation (Eqn. 3) between two consecutive "siftings"; in the analyses presented in this paper a value of 0.2 is used (Huang, N.E. *et al.* 1998):

$$SD = \sum_{t=0}^{T} \left[\frac{\left| \left(h_{1(k-1)}(t) - h_{1k}(t) \right) \right|^2}{h_{1(k-1)}^2(t)} \right] < 0.2 \qquad (3)$$

4. Calculate the residual $r_1(t)$ between the first IMF and the signal (Eqn. 4).

$$r_1(t) = x(t) - IMF_1(t) \qquad (4)$$

The residual $r_1(t)$ is now the new signal.
5. Repeat steps 1–4 using $r_1(t)$ to determine $m_2(t)$, $h_{2j}(t)$, and the second intrinsic mode function $IMF_2(t)$.
6. Calculate the new residual $r_2(t)$ (Eqn. 5).

$$r_2(t) = r_1(t) - IMF_2(t) \qquad (5)$$

7. Repeat steps 1–7 to obtain IMF_3, IMF_4, … IMF_n until $r_n(t)$ becomes a monotonic function or has only one extreme.

The EMD process produces n orthogonal IMFs and a residual $r(t) = r_n(t)$. The sum of the n IMFs and the residual $r(t)$ equals the original signal $x(t)$ (Eqn. 6). Performing this summation is a check to determine that the decomposition is complete.

$$x(t) = \sum_{j=1}^{n} IMF_j + r_n \qquad (6)$$

During each iteration of the decomposition, progressively lower frequencies of characteristic oscillations remaining in the signal are "sifted" out into a new IMF. The resulting IMFs encompass a narrow range of the frequencies present in the original signal, and as mode number increases, the mean frequency associated with that mode decreases. As with Fourier analysis, the sample length and the Nyquist frequency (half the sampling frequency) set the limit on the lowest and highest frequencies, respectively, that can be identified in the signal. The residual, which must be monotonic or contain only one extreme, represents the trend of the signal over the sample length.

Hilbert transform

The Hilbert transform is used to extract the local frequency and amplitude from each IMF for spectral analysis. The IMFs, as extracted from EMD, satisfy the two conditions required to define meaningful frequencies using the Hilbert transform. The Hilbert transform is the convolution of a function $f(t)$ with $1/\pi t$ (Eqn. 7).

$$\tilde{f}(t) = \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{f(\tau)}{t-\tau} d\tau \qquad (7)$$

A real function $f(t)$ and its Hilbert transform $f(t)$ form a strong complex analytic signal, Z(t) in which the real (r) part of Z(t) is $f(t)$ and the imaginary (i) part is given by $\tilde{f}(t)$ (Eqn. 8).

$$Z(t) = f(t) + i\tilde{f}(t) \qquad (8)$$

The Hilbert transform is applied to each of the n IMFs, forming n strong complex analytic signals (Eqn. 9a), which can be also be written in a polar form (Eqn. 9b).

$$Z_j(t) = IMF_j(t) + i\widetilde{IMF_j}(t) = Z_{rj}(t) + iZ_{ij}(t) \qquad (9a)$$

$$Z_j(t) = A_j(t)e^{i\theta(t)} \qquad (9b)$$

The local amplitude $A_j(t)$ is given by Eqn. 10 and the local angular frequency ω_n and frequency $f_j(t)$ for each IMF are computed as shown in Eqn. 11a and Eqn. 11b.

$$A_j(t) = \sqrt{Z_{rj}^2 + Z_{ij}^2} \qquad (10)$$

$$\omega_j(t) = \frac{d\theta_j(t)}{dt} \qquad (11a)$$

$$f_j(t) = \frac{\omega_j(t)}{2\pi} \qquad (11b)$$

Thus, there are n values of frequency and amplitude at every time step, which can be expressed as $A_j(f,t)$. The amplitude of an IMF at any given time step, for example $A_3(t_m)$, reflects the degree to which the corresponding local IMF frequency, $f_3(t_j)$, is present in the original signal *at that time step*, t_m. The combined set of all n $A_j(f,t)$ gives the Hilbert spectrum $H(f,t)$ (Eqn. 12).

$$H(f,t) = \{A_1(f,t), A_2(f,t) \dots A_n(f,t)\} \quad (12)$$

If two or more IMFs happen to have the same frequency at the same time step, the Hilbert spectrum value for that frequency is the sum of those IMF amplitudes at that time. The Hilbert spectrum can be plotted as discrete points in time-frequency space using a colour scale to represent H. Previous studies have applied a weighted Gaussian filter to the discrete points of the Hilbert spectrum to produce a visually continuous plot (Huang, N.E. *et al*. 1998).

The plot of the Hilbert spectrum is similar to a wavelet plot in that it displays the time evolution of the frequency spectrum. The main difference between $H(f,t)$ and a wavelet plot is in the method of calculating the amplitude at each frequency-time pair. Wavelet analysis performs a correlation of the signal with a wavelet of varying size at each time-step. The wavelet scale (a) is converted into a "pseudo-frequency" (F_a) using the centre frequency of the wavelet (F_c) and the sampling period of the signal (Δ) (Eqn. 13).

$$F_a = \frac{F_c}{a\Delta} \quad (13)$$

The form of the wavelet is defined a priori and can influence the results of the analysis substantially. Thus, comparisons between wavelet analyses are limited to those which use the same wavelet. In contrast, the Hilbert-Huang transform does not require any a priori choice of waveform, as the Hilbert spectrum is calculated directly from empirically derived IMFs. The only tuning parameter that affects HHT is the standard deviation tolerance in the shifting process, and, if kept at or below 0.2 (Huang, N.E. *et al*. 1998), primarily influences effects the calculation time with negligible changes in the IMFs and residual.

Integrating the Hilbert spectrum over time gives the marginal Hilbert spectrum h(f) (Eqn. 14).

$$h(f) = \int_0^T H(f,t)dt \quad (14)$$

This spectrum is similar to the Fourier spectrum in that it represents the contribution of a given frequency to the original signal. The marginal Hilbert spectrum differs substantially from the Fourier spectrum in that the HHT method can capture periodicities that occur only locally or intermittently in a signal – a capability that Fourier analysis lacks. Thus, peaks in the marginal Hilbert spectrum may represent frequencies that are present locally, intermittently or throughout the whole sample length. In contrast, the Fourier spectrum only captures periodicities that are present persistently over the entire sample length. Peaks in the marginal Hilbert spectrum $h(f)$ can be examined in detail via the Hilbert spectrum $H(f,t)$ to determine if these peaks represent persistent periodicities in the signal or intermittent variations in signal amplitude at certain frequencies. The frequency spectrum of the signal can also be investigated in more detail by integrating each $A_j(f,t)$ separately, yielding the marginal spectrum of each IMF. Such an approach is useful for identifying the relative contributions of characteristic oscillatory modes in a signal to the overall spectral characteristics of the signal and for comparing characteristic modes among different signals (Huang, N.E. *et al*. 1998; Konsoer, K.M. and Rhoads, B.L. 2014). The Hilbert spectrum may also be integrated over frequency to give the "Instantaneous Energy" (IE) at each time step (Eqn. 15). Large values of IE reflect local high amplitude oscillations, a metric of signal energy, of unspecified frequency:

$$IE(t) = \int_0^{f_{max}} H(f,t)df \quad (15)$$

In summary of this method section, HHT provides an alternative approach to analys-

ing temporal or spatial signals in spectral space compared to Fourier and wWavelet analyses, and offers provides an additional methods of decomposing a the signal and/or integrating it over time or frequency space. Importantly, HHT does not require the selection of a waveform for the spectral analysis a priori. In the next sections, we illustrate using present through various hydrodynamic examples the utility of the HHT method.

Results and discussion

Application to clay-laden flows

To illustrate the capabilities of the method to characterize complex turbulent flows with high degrees of intermittency, HHT analysis is applied to velocity data for clay-laden water flows, which exhibit flow characteristics that vary both with the concentration of suspended sediment and the mean velocity. Previous work has identified five types of clay flows based on the effect of increasing clay concentration on the characteristics of fluid turbulence: turbulent flow (TF), turbulence-enhanced transitional flow (TETF), lower and upper transitional plug flow (LTPF/UTPF) and quasi-laminar plug flow (PF) (Baas, J.H. *et al.* 2009). HTT analysis was performed on streamwise velocities measured in four types of clay-laden flows (Runs 10–1, 10–19, 10–22 and 10–26 from Baas, J.H. *et al*. 2009) *(Table 1).* These flows have similar depth-averaged velocities and depths, but different clay concentrations. Velocity signals for all four flows were measured 5 mm above the bed with an ultrasonic Doppler velocity profiler (UDVP) operating at a sampling frequency of 125 Hz (for details see Baas, J.H. *et al.* 2009). The UDVP measures the component of velocity parallel to the instrument beam at 128 points along the beam length.

Velocity time series for each of the four types of flows have distinctive characteristics (Baas, J.H. *et al.* 2009) (*Figure 1*). Signals for the TF and the TETF flows are similar in form, but amplitudes of turbulent fluctuations for the TETF are greater in magnitude than those for the TF. The LTPF exhibits distinct low-frequency oscillations compared to TF and TETF, whereas UTPF is characterized by intermittent "saw-tooth" shaped oscillations (Baas, J.H. *et al.* 2009). The Fourier and marginal HHT spectra for each of the four flows are generally similar in form, but the Fourier spectra exhibit more variability than the HHT spectra *(Figure 2).* The Fourier and Hilbert spectra of the LTPF and UTPF have slightly steeper slopes than spectra for the TF and TETF, but otherwise all four spectra have slopes close to -5/3 – the value expected for the energy cascade associated with the inertial subrange of turbulent flows (Kolmogorov, A.N. 1941; Sukhodolov, A.N. and Uijttewaal, W.S.J. 2010).

By contrast, the full Hilbert spectra of the four flows, plotted as filled contours, differ considerably from one another *(Figure 3).* The TF Hilbert spectrum and the TETF Hilbert spectrum are similar in that the Hilbert amplitude fluctuates intermittently, with a few localized high-amplitude sections. The TETF Hilbert spectrum exhibits consistently higher Hilbert amplitudes than the TF Hilbert spectrum over the whole range of frequencies and over the entire duration of the signal.

As expected based on the apparent periodicity of the velocity signal, the Hilbert

Table 1. Details for the four clay-laden flows

Run*	Flow type	Clay concentration, vol. %	Depth-averaged velocity, cm/s	Mean streamwise velocity**, cm/s	Flow depth, cm
10–1	TF	0.03	140.3	97.85	14.3
10–19	TETF	12.48	143.3	96.12	14.0
10–22	LTPF	14.45	140.4	75.15	14.1
10–26	UTPF	16.44	134.9	38.83	14.8

*From Baas, J.H. *et al.* 2009. **At 5 mm above the bed.

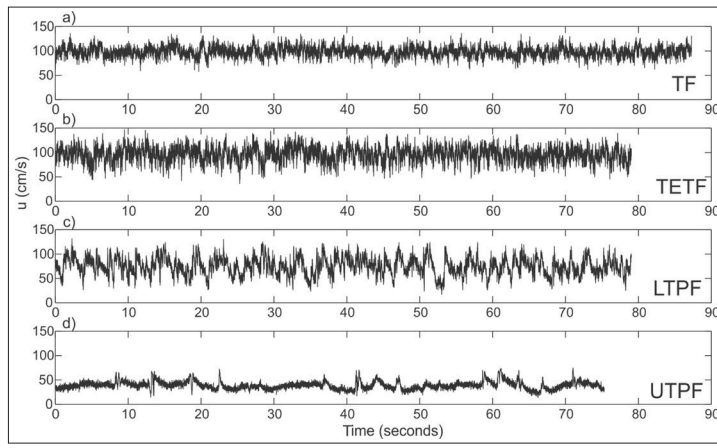


Fig. 1. Time series of streamwise velocity for different concentrations of suspended sediment. Classifications as identified by Baas, J.H. *et al*. (2009). – TF = turbulent flow; TETF = turbulence enhanced transitional flow; LTPF = lower transition plug flow; UTPF = upper transition plug flow

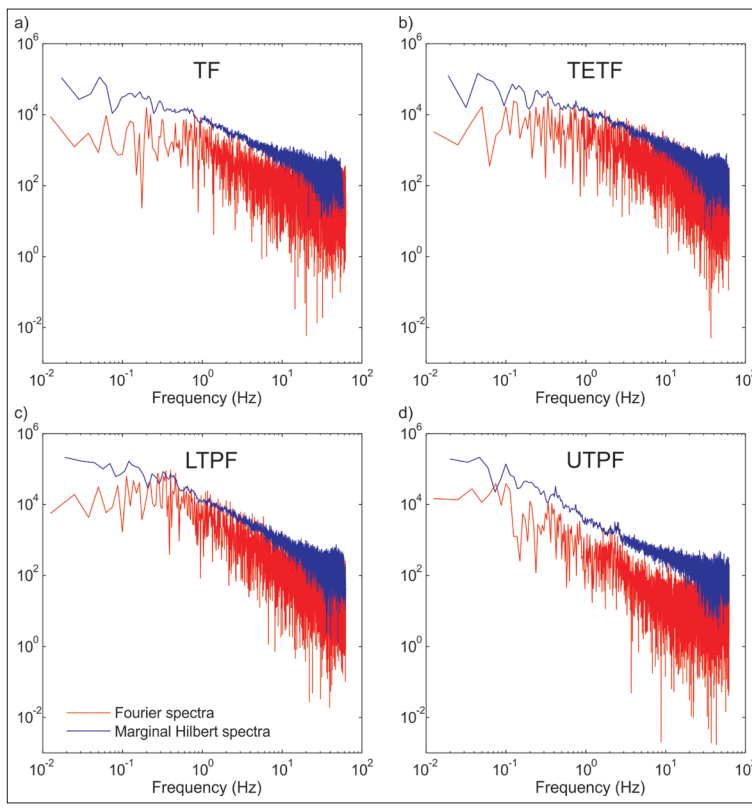


Fig. 2. Comparison of the Fourier spectra to the marginal Hilbert spectra for the four time series of clay-laden flows. – TF, TETF, LTPF and UTPF: For explanation see *Fig. 1.*

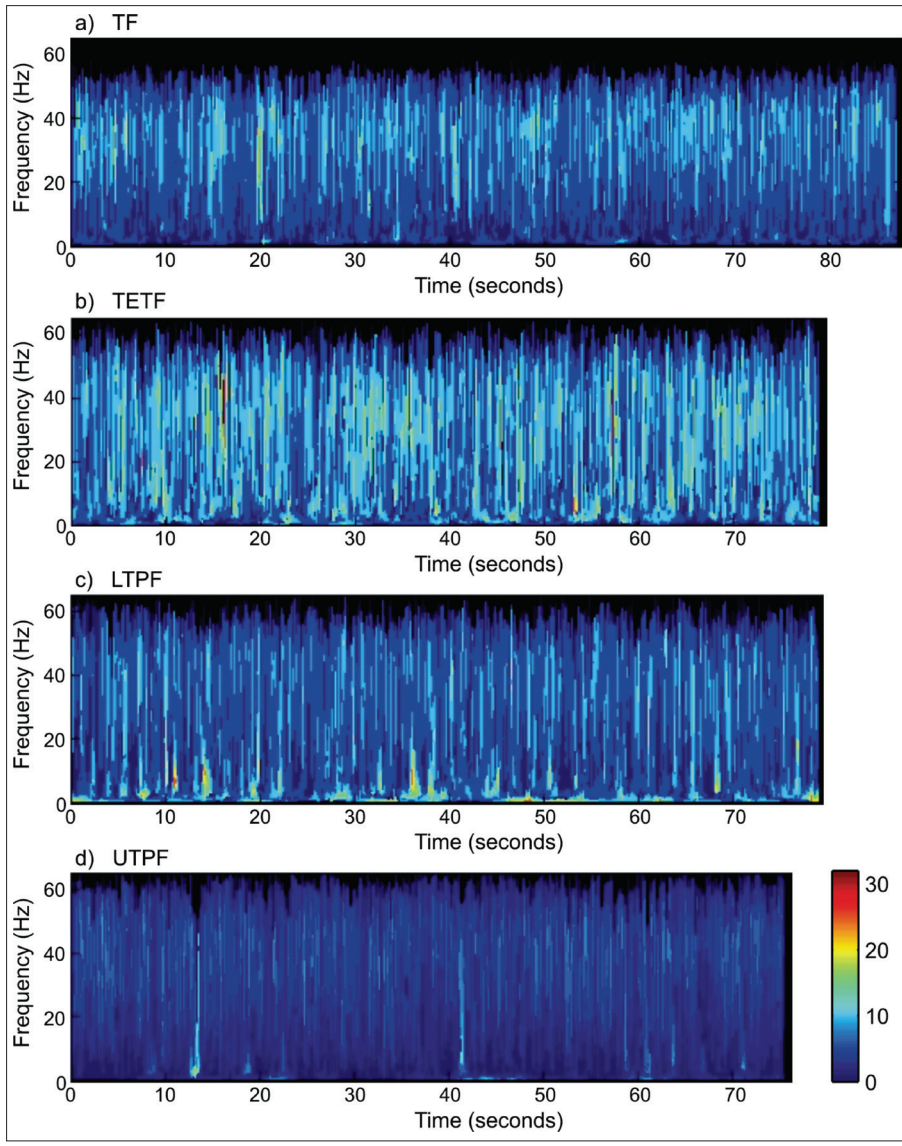


Fig. 3. Full Hilbert spectra for the four time series of clay-laden flows. – TF, TETF, LTPF and UTPF: For explanation see *Fig. 1.*

spectrum of the LTPF is characterized by intermittent high amplitudes at frequencies less than ~10 Hz. Values of amplitude for these low-frequency velocity variations are the greatest of any of the four flows. Strong low-frequency turbulence for LTPF has been attributed to Kelvin-Helmholz instabilities along an internal basal shear layer that develops between highly turbulent near-bed flow and less turbulent flow in the upper portion of the fluid column (Baas, J.H. *et al*. 2009). The Hilbert spectrum of the UTPF is charac-

terized by small amplitudes at all frequencies, indicating nearly pervasive damping of turbulence by high sediment concentrations. Weak turbulence near the bed, where the vertical gradient in streamwise velocity is largest, produces a number of highly isolated high amplitude fluctuations of the signal at low and intermediate frequencies (e.g. at ~13 seconds and ~42 seconds).

The Instantaneous Energy (IE) plots of the four signals reflect the differences in the Hilbert spectra *(Figure 4).* Because these plots show the sum of the local Hilbert amplitudes over all frequencies, high values of IE at a particular time may signify a strong local fluctuation of a single frequency or synchronous local fluctuations over multiple frequencies. Averaging the Instantaneous Energy over time gives a measure of the overall oscillatory energy in the signal. The mean Instantaneous Energy of the TETF (mean IE = 493) is higher than that of the TF (mean IE = 221), which is consistent with the higher turbulence intensity of the TETF compared to the TF (Baas, J.H. *et al*. 2009). The IE of both the TETF and TF fluctuate over time, reflecting that the strength of the turbulence fluctuates over time. The LTPF Instantaneous Energy consists of systematic low-frequency fluctuations about a relatively high mean value of 690. These low frequency fluctuations in IE probably reflect Kelvin-Helmholz instabilities within the basal shear layer of the LTPF, which produce high magnitude, low frequency fluctuations in streamwise velocity. The mean Instantaneous Energy of UTPF is the lowest of the four signals (mean IE = 143); a result that indicates suppression of turbulence related to gelling of the flow (Baas, J.H. *et al.* 2009). The IE of the UTPF is also characterized by local spikes, which occur at the same times as the "saw-tooth" velocity fluctuations in the raw velocity signal.

The distinct forms of the Hilbert spectra and Instantaneous Energy plots illustrate nicely the different behaviours of clay-enriched flows described by Baas, J.H. *et al.* (2009). Moreover, the full Hilbert spectrum provides a means of distinguishing among different types of clay-laden flows even

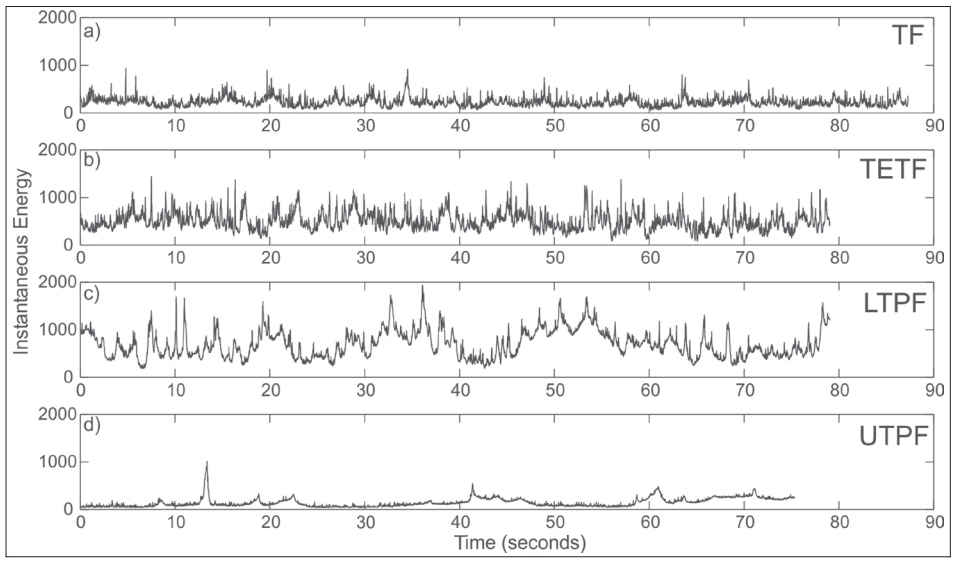


Fig. 4. Instantaneous energy plots for the time series of clay-laden flows. – TF, TETF, LTPF and UTPF: For explanation see *Fig. 1.*

though the marginal Hilbert spectra and Fourier spectra of these flows do not differ substantially. Results of the analysis can also be used to characterize differences in the intermittency of these flows, both for individual frequencies and over the entire range of frequencies. The applications of the HHT method to these clay-laden flows shows that it is a robust tool for examining the characteristics of complex turbulent flows, especially when velocity fluctuations of these flows are highly intermittent.

Application to combined flow over a mobile bed

The second application examines laboratory measurements of non-stationary time series of streamwise and vertical velocity components from a combined flow over an evolving mobile bed. The experiments were conducted in the Large Oscillatory Water-Sediment Tunnel facility at the Ven Te Chow Hydrosystems Laboratory at the University of Illinois, Urbana-Champaign (Perillo, M.M. *et al.* 2014). The combined flow is composed of an oscillatory component with a period of 5 seconds and a maximum oscillatory velocity of 30 cm/s, and a unidirectional component with mean velocity of 10 cm/s. The experiment began with a flat, mobile sand bed (D_{50} = 250 μm) and continued for 640 seconds through the development of bedforms. Velocities were measured with an ADV at a distance of 1 cm above the initial flat bed at a sampling frequency of 25.6 Hz. The time series were de-spiked prior to analysis with the HHT method (Goring, D.G. and Nikora, V.I. 2002).

Time series of streamwise (u) and vertical (w) velocities clearly show the non-stationarity of the signal, particularly the variance of the vertical velocities *(Figure 5).* The marginal Hilbert spectra of the streamwise and vertical velocities both have peaks centred at a frequency of 0.2 Hz, but the peak is much more pronounced for the streamwise velocities *(Figure 6).* In this case, the peak at 0.2 Hz is much sharper for the Fourier spectrum than for the marginal Hilbert spectrum because the persistent 0.2 Hz oscillatory component of the flow over the entire measurement is particularly well suited for Fourier analysis. Overall the marginal Hilbert spectra and the Fourier spectra of the two velocity components are similar in shape (see *Figure 6*). However, the marginal Hilbert spectrum of streamwise velocity exhibits a broad secondary peak centred on a frequency of 1 Hz, not visible in the Fourier spectrum, suggesting

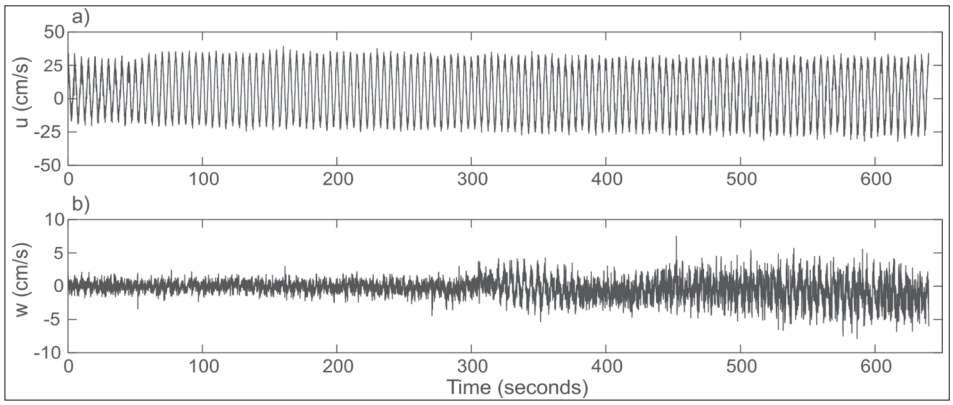


Fig. 5. Time series of (a) streamwise and (b) vertical velocity components of a combined oscillatory and unidirectional flow. (Data courtesy of Perillo, M.M.)

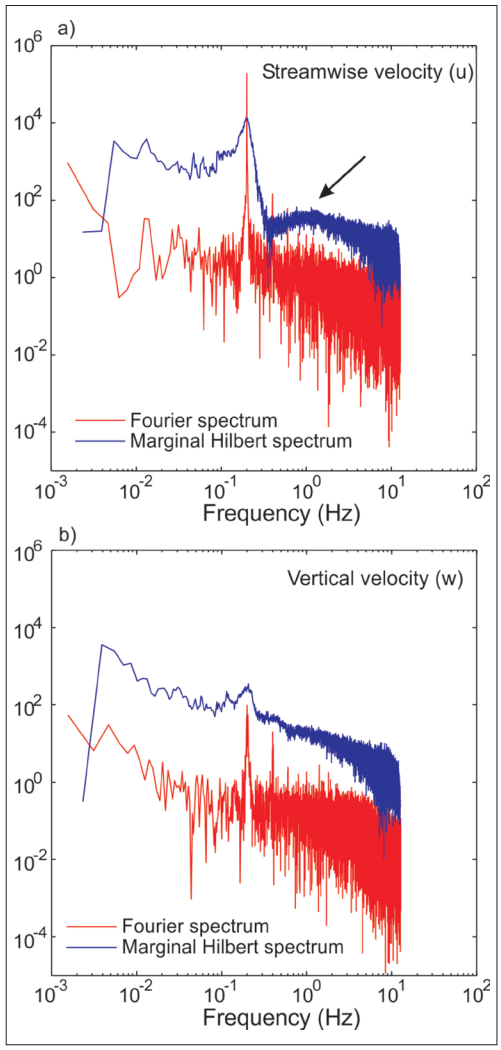


Fig. 6. Comparison of Fourier spectra and marginal Hilbert spectra for (a) streamwise velocity and (b) vertical velocity components for the combined flow described.

that high-frequency turbulent fluctuations are an important part of the signal and that the Fourier spectrum does not capture these high frequencies as well as the marginal Hilbert spectrum (see *Figure 6*).

A major advantage of the HHT method is its capacity to yield marginal Hilbert spectra for each IMF *(Figure 7)*. The IMFs represent major modes of variability over a narrow range of frequencies within the original time series. For the time series of streamwise velocity, the individual marginal Hilbert spectra of the IMFs show that IMFs 4–6 capture the primary oscillatory component of the combined flow, IMFs 7–11 represent large scale fluctuations not associated with the oscillatory motion, and IMFs 2 and 3, as well as the high-frequency tails of IMFs 4 and 5, represent turbulence associated with high frequency fluctuations (see *Figure 7*). The first IMF can be associated with measurement noise (Kanani, A. *et al*. 2010).

The Hilbert spectrum for the streamwise velocity of the combined flow (plotted as discrete points) shows a high-amplitude band centred on 0.2 Hz *(Figure 8,* a*)*. By isolating IMFs 2 and 3, which do not include the 0.2 Hz oscillation, it is possible to examine the time evolution of high frequency turbulent fluctuations independent of the low-frequency oscillatory component of the combined flow *(Figure 8,* b*)*. To aid interpretation of the time evolution of the signal, the data in the Hilbert spectrum were contoured. After an initial period (0–50 seconds) of increased Hilbert amplitudes associated with turning on the pumps for the experiment, amplitudes diminish as flow over the initially flatbed stabilizes and attains its target oscillatory streamwise velocities. At approximately 290 seconds, the Hilbert amplitudes begin to steadily increase until the end of the time series. This increase in Hilbert amplitude represents an increase in high-frequency turbulent energy associated with the initiation and growth of bedforms. The influence of the initiation of bedforms at roughly 290 seconds is also apparent in the time series of vertical velocity (see *Figure 5,* b) and its Hilbert spectrum *(Figure 8,* c*)*. The full Hilbert spectrum for streamwise velocity is in good agreement with the wavelet spectrum for lower frequencies *(Figure 8,* d*)*.

A comparison of the Hilbert spectrum and wavelet spectrum computed using the Mexican hat waveform illustrates some of the major differences between these two meth-

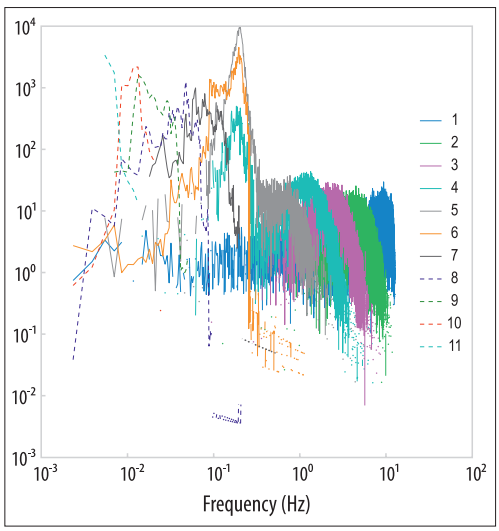


Fig. 7. Marginal Hilbert spectra for each IMF from the streamwise velocity signal of combined flow. Numbers 1–11: For explanation see the text.

ods (see *Figure 8*). The wavelet spectrum clearly shows an increase in the strength of the 0.2 Hz oscillation over the first 50 seconds, as expected from turning on the pumps for the experiment. However, in the wavelet spectrum, the 0.2 Hz oscillatory component dominates the entire signal and it is difficult to isolate the temporal evolution of the high frequencies. Thus, the ability to isolate individual IMFs is a unique advantage of the HHT method. Moreover, while the Hilbert spectrum provides discrete values of the frequency associated with these high amplitude events (see *Figure 8*, a), the wavelet spectrum spreads high amplitudes over a range of frequencies (see *Figure 8*, d).

Detrending and noise-removal with HHT

Because IMFs represent different oscillatory modes of characteristic frequencies contained within the original signal, the highest and lowest frequency components can be used to de-trend or remove noise from a signal. The HHT method is a robust tool for de-trending because it does not require an initial assumption about the type of trend in the signal and because it can easily remove nonlinear trends. The trend in a signal is simply the residual component from EMD. Subtracting this residual from the signal yields a stationary signal that can be analysed with either the Hilbert transform or the Fourier transform.

A time series of surface water temperatures collected from a mixing interface of a large river confluence of the Wabash and White rivers in Illinois *(Figure 9,* a–c*)* is used to demonstrate the de-trending capabilities of HHT (Konsoer, K.M. and Rhoads, B.L. 2014). The data were collected using a boat-mounted RDI acoustic Doppler current profiler (ADCP) at a sampling frequency of 1 Hz over a sampling interval of 117 minutes. The residual trend in the data, as extracted by EMD, shows temperatures increasing nonlinearly throughout the sampling interval *(Figure 9,* b*)*. This increase in temperature is associated with diel fluctuations of heating. Subtracting the residual from the original time series produces a detrended signal that is stationary over the full length of the time series *(Figure 9,* c*)*. In this manner, the diel trend in heating is isolated from the fluctuations in water surface temperature that result from the dynamics of the mixing interface at the confluence without assuming a certain type of trend a priori.

Assuming the surface water temperature acts as a passive marker of the fluid from the two tributaries, it is possible to extract information about the mixing interface from the frequency spectrum of the detrended temperature data. The Fourier spectrum of the original signal contains the energy associated with the low frequency diel fluctuation in temperature *(Figure 10,* a*)*, and is therefore greater than the Fourier spectrum of the detrended signal at frequencies less than ~10^{-2}. The Fourier spectrum of the detrended signal is nearly identical to the marginal Hilbert spectrum of the original signal, in which the residual trend is automatically excluded from calculations. The power-law form of the frequency spectrum of the detrended signal re-

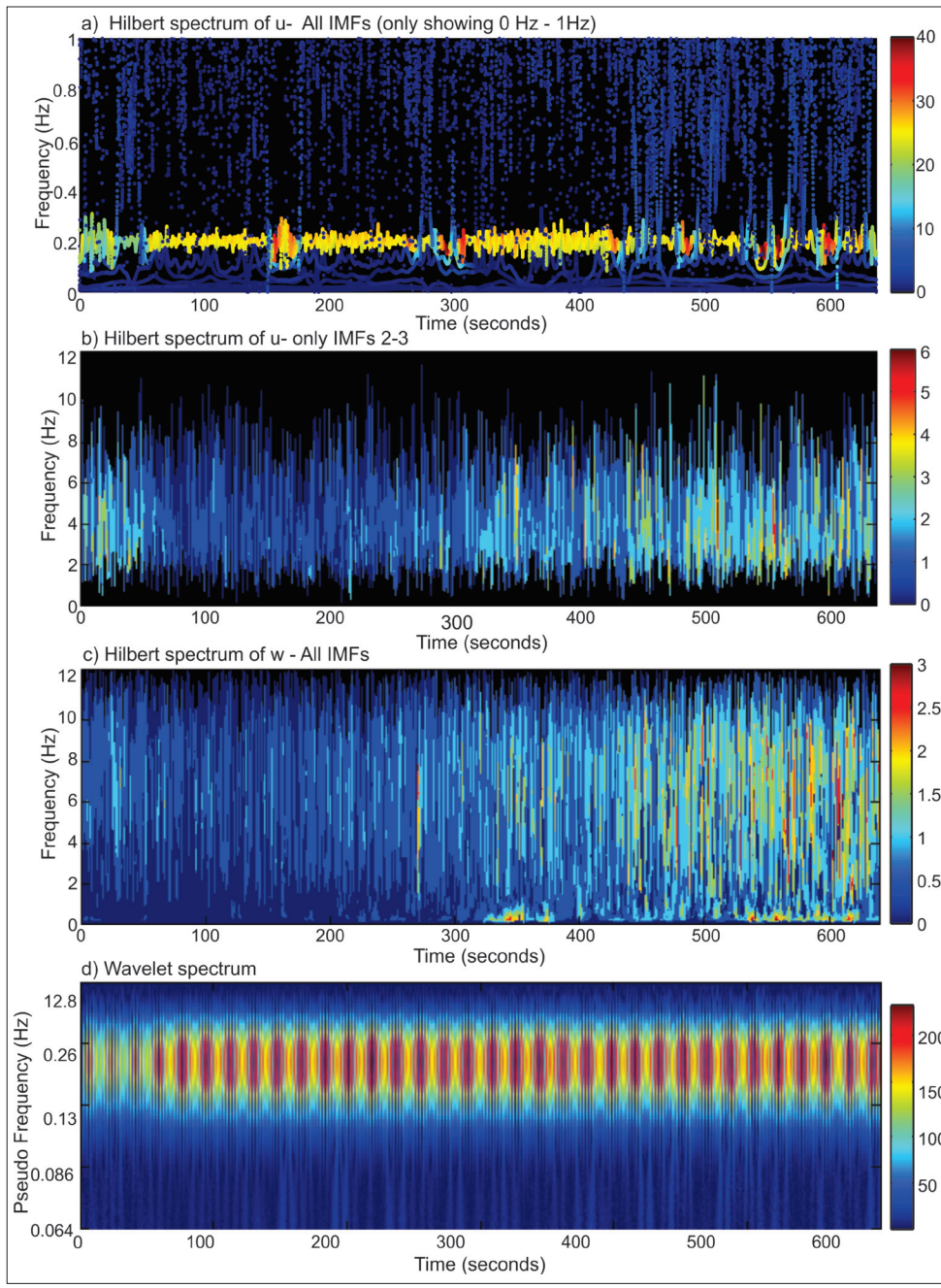


Fig. 8. The full Hilbert spectrum for the streamwise component of the combined flow signal – shown in *Figure 5*, upper part – plotted as discrete points (a); the Hilbert spectrum for only IMFs 2–3, plotted as filled contours (b); Full Hilbert spectrum for the vertical velocity component of combined flow, plotted as filled contours (c); and wavelet spectrum calculated using the Mexican Hat waveform (d) with scale converted to pseudo-frequency using Eqn. 13.

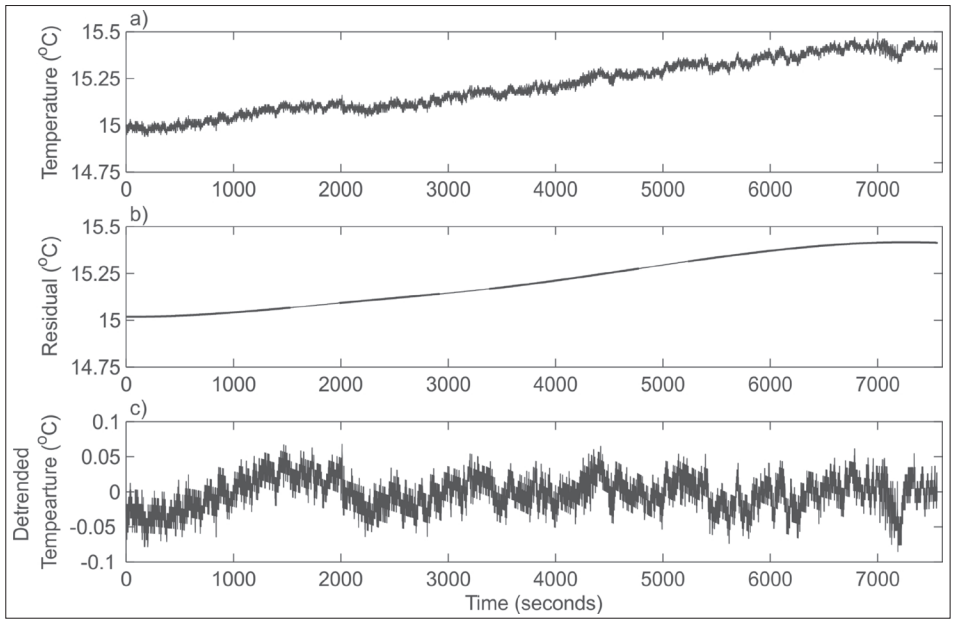


Fig. 9. Surface temperature at a fixed location along a mixing interface at a large river confluence, measured over the course of ~2 hours (a); Residual of the temperature signal as extracted from the EMD process (b); Detrended surface temperature, calculated by subtracting the residual from the original signal (c).

flects the variety of scales at which fluid is exchanged across the mixing interface at the confluence. A more detailed interpretation of the flow structures at the Wabash-White confluence is presented in Konsoer, K.M. and Rhoads, B.L. (2014). In addition to detrending hydraulic signals, the HHT method can also be used to de-trend and analyse morphologic data, as was shown for outer bank roughness characteristics along the meandering Wabash River (Konsoer, K.M. *et al.* 2017).

The first IMF, which represents the highest frequency fluctuations, contains most of the measurement noise in the signal (Huang, E.N. *et al.* 2003; Kanani, A. *et al*. 2010). To remove this noise, the first IMF can be subtracted from the original signal prior to Fourier analysis *(Figure 10,* b*)* or the marginal Hilbert spectrum of the first IMF can be excluded from the total marginal Hilbert spectrum (see *Figure 10,* b). In either case, this adjustment to the data removes the flat, high frequency tail of the spectrum that is dominated by signal noise.

Conclusions

The examples presented herein demonstrate the unique capabilities of the Hilbert-Huang transform method to shed light on the characteristics of turbulent shear flows. In particular, Hilbert-Huang transform analysis is well-suited for identification of intermittent or localized turbulent events and for analysis of turbulence characteristics over specific frequency ranges. The major benefits of the Hilbert-Huang transform include:

– The Hilbert spectrum is discrete, permitting precise identification of the frequency of a distinct high-amplitude turbulent event occurring at a specific time.
– The Hilbert spectrum allows for the analysis of non-stationary time-evolving flows, enabling exploration of how the spectral characteristics of flow properties vary with changing boundary conditions.
– The Hilbert spectrum is useful in identifying intermittency and determining the character-

istics of intermittent turbulent events in signals that otherwise appear to be stationary.
– The Hilbert spectrum provides an additional means of distinguishing types of flows that may show similar marginal Hilbert spectra or Fourier spectra (e.g. the four types of clay-laden flows).
– The Hilbert-Huang transform allows for isolation of individual IMFs, enabling detailed analysis of key frequency ranges.
– The EMD process provides a means of detrending and noise-removal that requires no a-priori assumptions about the signal.

Because of these attributes, Hilbert-Huang transform analysis is useful as an *additional* tool for spectral analysis of data from these types of flows, but is not a replacement for Fourier or wavelet analysis. Further application of the Hilbert-Huang transform method to turbulent shear flows may provide new insights into how the frequency spectrum of these flows is influenced by strong intermittency, non-stationarity, and non-linearity.

***Acknowledgements*:** The authors would like to acknowledge and thank Jessica LeRoy for her major contributions to the data analysis and writing of this manuscript. This manuscript was improved through thoughtful reviews by the editor and anonymous reviewer. The authors also gratefully thank Mauricio M. Perillo as well as Jaco Bass, Jim Best, Jeff Peakall and Mi Wang for use of their datasets. Matlab routines for the empirical mode decomposition were provided by Patrick Flandrin (http://perso.ens-lyon.fr/patrick.flandrin/emd.html). Work conducted for this project was supported by NSF Grant BCS-0453316.

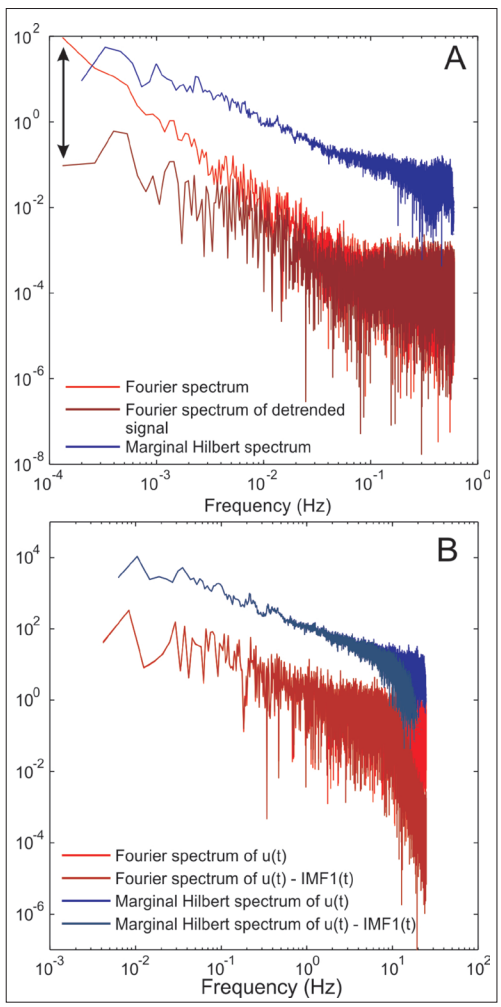


Fig. 10. Fourier spectrum of the original temperature signal (red), Fourier spectrum of the detrended temperature signal (dark red) and the marginal Hilbert spectrum of the original temperature signal (blue) – double-arrow line indicates low frequency energy associated with residual trend in Fourier spectrum (a); Examples of noise removal using EMD for the streamwise velocity of the unidirectional flow. The light red line and light blue line show the Fourier spectrum and the marginal Hilbert spectrum, respectively, prior to noise removal and the darker lines show the spectra following noise-removal (b).

REFERENCES

Baas, J.H., Best, J.L., Peakall, J. and Wang, M. 2009. A phase diagram for turbulent, transitional, and laminar clay suspension flows. *Journal of Sedimentary Research* 79. 162–183. Doi:10.2110/jsr.2009.025.

Buffin-Bélanger, T., Roy, A.G. and Kirkbride, A.D. 2000. On large-scale flow structures in a gravel-bed river. *Geomorphology* 32. (3–4): 417–435. Doi:10.1016/S0169-555X(99)00106-3.

Cellino, M. and Lemmin, U. 2004. Influence of Coherent Flow Structures on the Dynamics of Suspended Sediment Transport in Open-Channel Flow. *Journal of Hydraulic Engineering* 130. (11): 1077–1088.

Chen, J. and Hu, F. 2003. Coherent structures detected in atmospheric boundary-layer turbulence using wavelet transforms at Huaihe River basin, China. *Boundary Layer Meteorology* 107. 429–444.

Flandrin, P., Rilling, G. and Gonçalves, P. 2004. Empirical mode decomposition as a filter bank. *IEEE Signal Processing Letters* 11. 112–114. Doi:10.1109/LSP.2003.821662.

Flandrin, P. and Gonçalves, P. 2004. Empirical mode decompositions as data-driven wavelet-like expansions. *International Journal of Wavelets, Multiresolution and Information Processing* 2. 477–496. Doi:10.1142/S0219691304000561.

Flandrin, P., Gonçalves, P. and Rilling, G. 2005. EMD equivalent filter banks, from interpretation to applications. In *Hilbert-Huang Transform and Its Applications.* Eds.: Huang, N.E. and Shen, S.S.P., Singapore, World Scientific Publishing Co., 57–74.

Goring, D.G. and Nikora, V.I. 2002. Despiking acoustic Doppler velocimeter data. *Journal of Hydraulic Engineering* 128. 117–126. Doi:10.1061/(ASCE)0733-9429(2002)128:1(117).

Hardy, R.J., Best, J.L., Lane, S.N. and Carbonneau, P.E. 2009. Coherent flow structures in a depth-limited flow over a gravel surface: The role of near-bed turbulence and influence of Reynolds number. *Journal of Geophysical Research* 114. 1–18. Doi:10.1029/2007JF000970.

Huang, N.E., Shen, Z., Long, S.R., Wu, M.C., Shih, H.H., Zheng, Q., Yen, N., Tung, C.C. and Liu, H.H. 1998. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society A*, 454. 903–995. Doi:10.1098/rspa.1998.0193.

Huang, N.E., Wu, M.C., Long, S.R., Shen, S.S.P., Qu, W., Gloersen, P. and Fan, K.L. 2003. A confidence limit for the empirical mode decomposition and Hilbert spectral analysis. *Proceedings of the Royal Society A*, 459. 2317–2345. Doi:10.1098/rspa.2003.1123.

Huang, N.E. and Wu, Z. 2008. A review on Hilbert-Huang Transform: Method and its applications to geophysical studies. *Reviews of Geophysics* 46. 1–23. Doi:10.1029/2007RG000228.

Huang, Y., Schmitt, F.G., Lu, Z. and Liu, Y. 2009. Analysis of daily river flow fluctuations using empirical mode decomposition and arbitrary order Hilbert spectral analysis. *Journal of Hydrology* 373. 103–111. Doi:10.1016/j.jhydrol.2009.04.015.

Kanani, A., Ahmari, H. and Ferreira da Silva, A.M. 2010. Investigation of horizontal coherent structures in a shallow open-channel flow using velocity signal decomposition. In *Proceedings from the International Conference on Fluvial Hydraulics River Flow.* September 8–10, 2010. Braunschweig, Germany, 1059–1066.

Kolmogorov, A.N. 1941. Dissipation of energy in locally isotropic turbulence. *Doklady Akademii Nauk SSSR 32.* 19–21. (in Russian)

Konsoer, K.M. and Rhoads, B.L. 2014. Spatial-temporal structure of mixing interface turbulence at two large river confluences. *Environmental Fluid Mechanics* 14. 1043–1070. Doi:10.1007/s10652-013-9304-5.

Konsoer, K.M., Rhoads, B., Best, J., Langendoen, E., Ursic, M., Abad, J. and Garcia, M. 2017. Length scales and statistical characteristics of outer bank roughness for large elongate meander bends: The influence of bank material properties, floodplain vegetation and flow inundation. *Earth Surface Processes and Landforms* 42. 2024–2037.

Lewis, Q.W. and Rhoads, B.L. 2015. Rates and patterns of thermal mixing at a small stream confluence under variable incoming flow conditions. *Hydrological Processes* 29. (20): 4442–4456.

Loh, C., Lin, C. and Huang, C. 2000. Time domain identification of frames under earthquake loadings. *Journal of Engineering Mechanics* 126. 693–703. Doi:10.1061/(ASCE)0733-9399(2000)126:7(693).

Loh, C., Wu, T. and Huang, N.E. 2001. Application of the empirical mode decomposition-Hilbert spectrum method to identify near-fault ground-motion characteristics and structural responses. *Bulletin of the Seismological Society of America* 91. 1339–1357. Doi:10.1785/0120000715.

Nepf, H., Rominger, J. and Zong, L. 2013. Coherent Flow Structures in Vegetated Channels. In *Coherent Flow Structures at Earth's Surface.* Eds.: Venditti, J.G., Best, J.L., Church, M. and Hardy, R.J., Chichester, UK, John Wiley & Sons Ltd. Doi: 10.1002/9781118527221.ch9.

Nikora, V.I. and Smart, G.M. 1997. Turbulence characteristics of New Zealand gravel-bed rivers. *Journal of Hydraulic Engineering* 123. 764–773.

Niño, Y. and Garcia, M.H. 1996. Experiments on particle – turbulence interactions in the near-wall region of an open channel flow: implications for sediment transport. *Journal of Fluid Mechanics* 326. 285–319. Doi:10.1017/S0022112096008324.

Perillo, M.M., Best, J.L. and Garcia, M.H. 2014. A unified model for bedform development and equilibrium under unidirectional, oscillatory and combined-flows. *Sedimentology* 61. (7): 2063–2085.

Rao, A.R. and Hsu, E.C. 2010. *Hilbert-Huang Transform Analysis of Hydrological and Environmental Time Series.* Dordrecht, Springer.

Rhoads, B. and Sukhodolov, A. 2004. Spatial and temporal structure of shear layer turbulence at a stream confluence. *Water Resources Research* 40. 1–13. Doi:10.1029/2004WR002811.

Schmitt, F.G., Huang, Y., Lu, Z., Liu, Y. and Fernandez, N. 2009. Analysis of velocity fluctuations and their intermittency properties in the surf zone using empirical mode decomposition. *Journal of Marine Systems* 77. 473–481. Doi:10.1016/j.jmarsys.2008.11.012.

Singh, A., Porte-Agel, F. and Foufoula-Georgiou, E. 2010. On the influence of gravel bed dynamics on velocity power spectra. *Water Resources Research* 46. 1–10, Doi:10.1029/2009WR008190.

Singh, A., Foufoula-Georgiou, E., Porte-Agel, F. and Wilcock, P.R. 2012. Coupled dynamics of the co-evolution of bed topography, flow turbulence and sediment transport in an experimental flume. *Journal of Geophysical Research* 117. F04016. Doi:10.1029/2011JF002323.

Sukhodolov, A. 1998. Turbulence structure in a river reach with sand bed. *Water Resources Research* 34. 1317–1334. Doi:10.1029/98WR00269.

Sukhodolov, A. and Rhoads, B. 2001. Field investigation of three-dimensional flow structure at stream confluences. *Water Resources Research* 37. 2411–2424.

Sukhodolov, A.N. and Uijttewaal, W.S.J. 2010. Assessment of a river reach for environmental fluid dynamics studies. *Journal of Hydraulic Engineering* 136. 880–888. Doi:10.1061/(ASCE)HY.1943-7900.0000267.

Thomas, E.R., Dennis, P.R., Bracegirdle, T.J. and Franzke, C. 2009. Ice core evidence for significant 100-year regional warming on the Antarctic Peninsula *Geophysical Research Letters* 36. 1–5. Doi:10.1029/2009GL040104.

Uijttewaal, W.S.L. and Tukker, J. 1998. Development of quasi two-dimensional structures in a shallow free-surface mixing layer. *Experiments in Fluids* 24. 192–200.

Venditti, J.G., Hardy, R.J., Church, M. and Best, J.L. 2013. What is a Coherent Flow Structure in Geophysical Flow? In *Coherent Flow Structures at Earth's Surface*. Eds.: Venditti, J.G., Best, J.L., Church, M. and Hardy, R.J., Chichester, UK, John Wiley & Sons Ltd. Doi:10.1002/9781118527221.ch1

Wu, F. and Yang, K. 2004. Entrainment Probabilities of Mixed-Size Sediment Incorporating Near-Bed Coherent Flow Structures. *Journal of Hydraulic Engineering* 130. (12): 1187–1197.

Wu, X., Yu, B. and Wang, Y. 2013. Wavelet analysis on turbulent structure in drag-reducing channel flow based on direct numerical simulation. *Advances in Mechanical Engineering* 5. (1): 1–11. Doi:10.1155/2013/514325.

Wu, Z. and Huang, N.E. 2004. A study of the characteristics of white noise using the empirical mode decomposition method. *Proceedings of the Royal Society A*, 460. 1597–1611. Doi:10.1098/rspa.2003.1221.

Wu, Z. and Huang, N.E. 2005. Statistical significance test of intrinsic mode functions. In *Hilbert-Huang Transform and Its Applications.* Eds.: Huang, N.E. and Shen, S.S.P., Singapore, World Scientific Publishing Co., 107–127.

Zedler, E. and Street, R. 2001. Large-Eddy Simulation of Sediment Transport: Currents over Ripples. *Journal of Hydraulic Engineering* 127. (6): 444–452.

DOI: 10.15201/hungeobull.67.4.5 Hungarian Geographical Bulletin **67** 2018 (4) 361–373.

Evaluation of debris flow susceptibility in El Salvador (CA): a comparison between Multivariate Adaptive Regression Splines (MARS) and Binary Logistic Regression (BLR)

Edoardo ROTIGLIANO, Chiara MARTINELLO
Valerio AGNESI and Christian CONOSCENTI[1]

Abstract

In the studies of landslide susceptibility assessment, which have been developed in recent years, statistical methods have increasingly been applied. Among all, the BLR (Binary Logistic Regression) certainly finds a more extensive application while MARS (Multivariate Adaptive Regression Splines), despite the good performance and the innovation of the strategies of analysis, only recently began to be employed as a statistical tool for predicting landslide occurrence. The purpose of this research was to evaluate the predictive performance and identify possible drawbacks of the two statistical techniques mentioned above, focusing in particular on the prediction of debris flows. To this aim, an inventory of debris flows triggered by the passage of the hurricane IDA and the low-pressure system associated with it 96E, on 7th and 8th November 2009, in an area of about 26 km^2 close to the Caldera Ilopango, El Salvador (CA), was employed. Two validation strategies have been applied to both statistical techniques, thus obtaining four models – BLR (I), MARS (I), BLR (II) and MARS (II) – to be compared in pairs. Model performance was assessed in terms of AUC (area under the receiver operating characteristic (ROC) curve), Sensitivity, Specificity, Positive Prediction Value and Negative Prediction Value. Moreover, to evaluate the robustness of the modelling procedure, 50 replicates were created for each model and standard deviation was calculated for each of them. The results show that both techniques allow for obtaining good or excellent performances so that it is not possible to define one of the two techniques as absolutely better. However, the validation procedure reveals slightly better performance of the MARS models, with greater sensitivity and greater discrimination among True Negatives (TNs).

Keywords: landslide susceptibility, debris flows, Multivariate Adaptive Regression Splines (MARS), Binary Logistic Regression (BLR), hurricane Ida, El Salvador

Introduction

Landslide susceptibility is defined as the probability of the occurrence of landslides in a given area according to local geo-environmental variables (Brabb, E.E. 1984; Carrara, A. *et al.* 1995; Guzzetti, F. *et al*. 1999).

For defining landslide susceptibility, both direct and indirect methods can be used. Direct methods are based on analyses performed by expert geomorphologists who divide the territory into areas with different susceptibility, defining the latter in qualitative terms (Verstappen, H.T. 1983; Van Westen, C.J. *et al.* 2003). By using indirect, deterministic or statistical methods, it is instead possible to obtain a quantitative classification of landslide susceptibility (Chung, C.J.F. *et al.* 1995; Ohlmacher, G.C. and Davis, J.C. 2003; Conoscenti, C. *et al.* 2008).

Landslide susceptibility assessment, developed in recent years, has seen an increasingly frequent use of the statistical approach. This method is based on the assumption

[1] Dipartimento di Scienze della Terra e del Mare, Univestità degli Studi di Palermo, Via Archirafi 22. 90123 Palermo, Italy. Corresponding author's e-mail: christian.conoscenti@unipa.it

that new landslides occur in areas with geo-environmental conditions that have caused landslides in the past (Carrara, A. *et al.* 1995; Guzzetti, F. *et al*. 1999; Van Westen, C.J. *et al.* 2003, 2008). In light of this, geo-environmental variables can be considered as predictors or independent variables while the past/present distribution of landslides as the dependent variable. The covariates are selected to reflect the variability of geo-environmental factors that are considered related to the activation of landslides and, moreover, the choice is also a function of the quality and resolution of available data. The landslide archive can be generated through field mapping or detection from high resolution remotely-sensed images. In this study, we employed the landslide archive produced by Rotigliano, E. *et al.* (2018), which was obtained through remote mapping of the Google Earth™ image dated 11/21/2009 (DigitalGlobe Catalog ID: 101001000AA5D801).

The statistical method is therefore aimed at determining relationships existing between the covariates and the dependent variable. For the determination of relationships, statistical analysis of bivariate (logistic regression, binary logistic regression [BLR]) or multivariate type (e.g., Multivariate Adaptive Regression Splines [MARS], cluster analysis, discriminant analysis) can be used. In the literature, a number of examples of landslide susceptibility studies performed using bivariate (e.g., Guzzetti, F. *et al.* 1999; Rotigliano, E. *et al.* 2012; Costanzo, D. *et al.* 2014) or multivariate analysis (e.g., Vorpahl, P. *et al.* 2012; Felicísimo, Á.M. *et al.* 2013; Conoscenti, C. *et al.* 2015, 2016) can be found. Comparative studies of the two methods are however rare (e.g. Conoscenti, C. *et al.* 2015) and, as far as we know, no comparison has ever been made in the case of evaluation of susceptibility from debris flow.

The main objective of the study is to show the difference in terms of predictive performance of the two methods, i.e. BLR and MARS, by using two different validation schemes (Chung, C.J.F. and Fabbri, A.G. 2003). Firstly, we created 50 datasets consisting of balanced samples of event and non-event pixels. In the first validation scheme, each dataset was exploited for both calibration and validation of the models. In the second scheme, each archive was split in two: the first, containing 75 per cent of the positive and negative cases, was used for the calibration of the models; the second, containing the remaining 25 per cent of the balanced archive, was used for validation. The performance of the models was assessed through the analysis of AUC (area under the receiver operating characteristic (ROC) curve) values and the confusion matrices.

The two validation schemes have been developed for both BLR and MARS models. The creation of 50 models for each procedure also allowed us to evaluate the robustness of the analysis (Costanzo, D. *et al.* 2014).

Materials and methods

Study area

The study area is a small drainage basin of about 26 km^2, located along the slopes of the Caldera Ilopango, El Salvador, CA *(Figure 1).* The slopes of the area are covered by levels of tefra and ignimbrite, derived from the most recent Quaternary eruptions of the Caldera *(Figure 2).* In the area are found deep V-shaped valleys *(Figure 3)*, whose formation is linked to intense weathering and mass movements that affected the volcanic bedrock. The latter, characterized by poor mechanical properties, can be in fact easily eroded by water especially during extreme meteorological events (cyclones and hurricanes), which occur very frequently in the region.

The climate regime of El Salvador is tropical-humid, with average annual rainfall above 1,500 mm and average annual temperature between 20 °C and 30 °C. Under these conditions, physical degradation is favoured, with further deterioration of the geotechnical properties of the pyroclastics. The meteorological phenomena are then responsible for the saturation of the degraded material and thus for a signifi-

Fig. 1. Location of the study area

cant decrease of cohesion. Due to these environmental conditions, the study area is particularly prone to landsliding, especially to debris flow type landslides. Unregulated deforestation and intensive cultivation of the area, even in very steep slopes, further promotes soil erosion as well as landslide processes.

Landslide inventory

Between 7th and 8th November 2009, El Salvador was affected by the simultaneous passage of Hurricane Ida and the 96/E low-pressure system. The greatest damage occurred in an area of about 400 km^2 around the Caldera Ilopango, where more than 300 mm of rainfall in 12 hours were recorded at the Ilopango rain gauge station. Such an intense rainfall event triggered more than two thousand debris flows and caused flooding of the valleys, causing approximately 200 fatalities and immense economic loss with destruction of houses, roads and crops (MARN 2010).

The landslide archive used in this study is a database of landslide phenomena that occurred in the catchment due to the concomitant passage of Ida and 96/E. The archive has already been used in Rotigliano, E. *et al.* (2018). The recognition of the landslides and their mapping has been carried out remotely, using a high-resolution satellite image available on the Google Earth software, which is dated 11/21/2009 (DigitalGlobe Catalog ID: 101001000AA5D801). This image, acquired only 2 weeks after the passage of Ida-96/E, allowed the identification and mapping of 2231 debris flows triggered by the aforementioned rainfall event.

Each failure has been mapped by using a landslide identification point (LIP), located at the point of origin of the movement. In the case of evaluation of susceptibility to debris flow, according to Rotigliano, E. *et al.* (2011), LIPs allow us to obtain the most reliable landslide prediction as their environmental characteristics are those that best represent pre-failure conditions and thus can be considered the best diagnostic areas for calibrating (and validating) landslide predictive models (Rotigliano, E. *et al.* 2011, Lombardo, L. *et al.* 2014; Cama, M. *et al.* 2015). For this reason, it was decided to use the archive without making any changes with respect to the initial characteristics.

Statistical modelling

In the last 20 years, many studies have dealt with landslide susceptibility modelling and a number of them have used a statistical modelling approach. Binary logistic regression (BLR) is among the most frequently used statistical techniques in geomorphology, and in particular in the field of landslide susceptibility assessment (e.g. Bai, S.B. *et al.* 2010; Atkinson, P.M. and Massari, R. 2011; Costanzo, D. *et al.* 2014). On the other hand, Multivariate Adaptive Regression Splines (MARS) (Friedman, J.H. 1991), has been employed only a few times in geomorphology (e.g., Gómez-Gutiérrez, Á. *et al.* 2009, 2015; Conoscenti, C. *et al.* 2016, 2018; Garosi, Y. *et al.* 2018).

Both BLR and MARS make it possible to identify relationships between a set of independent variables (predictors), both continuous or categorical, and a dependent dichotomic variable, which is usually coded as 0 (non-event) or 1 (event).

The aim of BLR is to describe the linear relationship between the logit (or log odds) of the dependent variable and the set of n independent variables (Hosmer, D.W. and

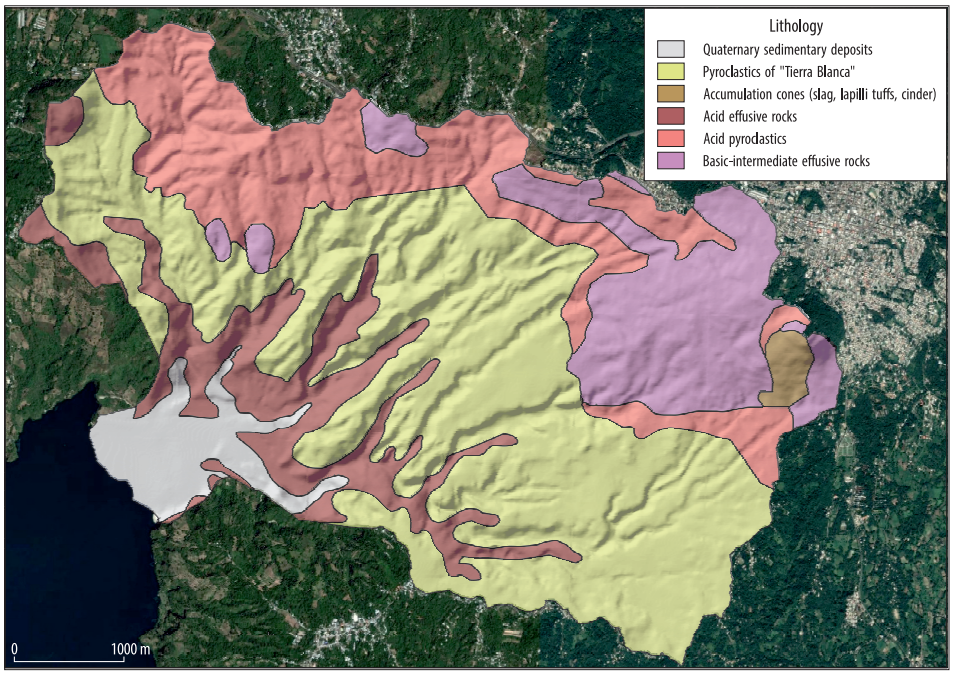


Fig. 2. Outcropping lithology in the study area (from Weber, H.S. *et al.* 1978)

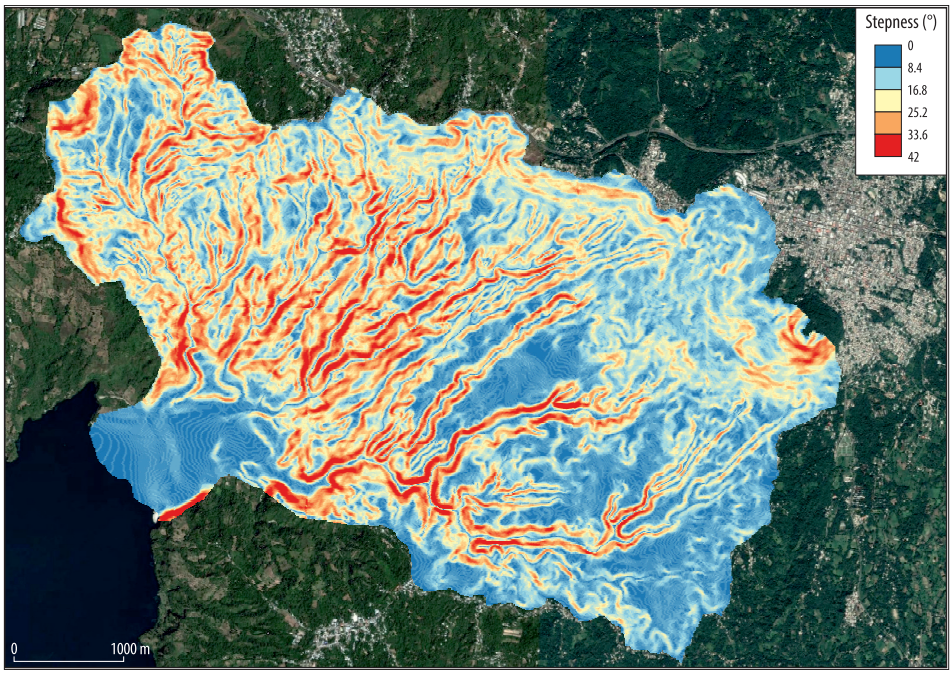


Fig. 3. Slope gradient map of the study area

Lemeshow, S. 2000). This is described by the following equation:

$$y = \ln\left(\frac{\pi(x)}{1-\pi(x)}\right) = \alpha + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n$$

where π (x) is the conditional mean of the response given specific values of x, α is the constant or intercept, x_i $(i = 1, 2, \ldots n)$ is the n^{th} independent variable and β_i $(i = 1, 2, \ldots n)$ is the n^{th} coefficient of the independent variable.

To optimize the values of y having certain independent variables, or rather to identify the value of α and β_i, the maximum likelihood technique is used, actually the log-likelihood (LL) function (Menard, S. 1995).

MARS is a non-parametric regression technique capable of identifying non-linear adaptation relationships between independent variables and dependent variables. MARS divides the range of the predictor values into regions and generates a linear regression equation for each region. The "nodes" are the extreme values of each region while each distinct interval is called "basis function" (BF). The latter can take the form:

$$\max(0, x - c) \text{ o}$$
$$\max(0, c - x),$$

where x is an independent variable and c is a constant corresponding to a knot. The general expression of MARS can be written as follows:

$$y = \alpha + \sum_{n=1}^{N} \beta_n h_n (x_n),$$

where y is the dependent variable, α is the constant, N is the number of terms, β_n is the coefficient of the n^{th} term and h_n (x) is a single basic fiction or a product of two or more BFs.

MARS builds the model in two phases. In the first step (forward pass) a complex model is produced in which basis functions generated for each variable are added. In the second phase (backward pass) a leaner model is established through generalized cross validation (GCV) (Craven, P. and Wahba, G. 1979). Basically, MARS removes the least influential pair in the creation of the best model and the GCV allows the identification, among the models generated, of the one offering the best compromise between adaptation (low RSS) and complexity/completeness of the model (Briand, L.C. *et al.* 2004; Gómez-Gutiérrez, Á. *et al.* 2009).

Both regression techniques and related analysis have been implemented using the software R (R CoreTeam, 2017). For the BLR analysis the "stats" package has been used, for MARS analysis the "earth" package (Milborrow, S. *et al.* 2011; Milborrow, S. 2015).

Predictors

The predictor variables were chosen according to their expected influence on slope instability and to their control on slope failure mechanisms (e.g., Conoscenti, C. *et al.* 2015; Cama, M. *et al.* 2016; Prokos, H. *et al.* 2016).

A set of ten geo-environmental variables was employed to predict debris flow susceptibility in the Caldera Ilopango. This set includes lithology (LIT) and land use (USE), in addition to the following eight terrain attributes: landform classification (LCL), elevation (ELE), slope steepness (STP), slope aspect (ASP), plan curvature (PLN), profile curvature (PRF), topographic wetness index (TWI) and terrain ruggedness index (TRI).

The outcropping lithology was obtained through the acquisition and processing of a 1:100,000 geological map (Weber, H.S. *et al.* 1978), which was derived from field survey 1: 25,000 scale. An up-to-date land use map was created on the basis of field surveys and analysis of ASTER and Google Earth images. The terrain attributes were extracted from a digital elevation model (DEM) with a ground resolution of 10-m, by using the software SAGA-GIS. All terrain attributes were extracted as continuous variables with the exception of LCL, which was classified into 10 classes, namely: streams, mid-slope drainages, upland drainages, valleys, plains, open slopes, local ridges, mid-slope ridges, high ridges. The 10-m cells of the DEM were employed as mapping units of the debris flow susceptibility in the studied area.

In order to detect collinearity among the chosen covariates, we calculated the Variance Inflation Factor (*VIF*) by using the "usdm" package (Naimi, B. 2015) implemented in the software R. A *VIF* value equal or higher than 10 indicates collinearity among the selected covariates (Heckmann, T. *et al.* 2014; Jebur, M.N. *et al.* 2014; Bui, D.T. *et al.* 2016). As *VIF* values calculated for our variables were below this threshold, all were included in both BLR and MARS models.

Model building and validation strategy

Landslide susceptibility assessment requires a validation procedure in order to evaluate the accuracy of the predictive models. This is generally performed in two steps: i) calibration of the models and ii) validation of the models (Chung, C.J.F. and Fabbri, A.G. 2003).

In this study, we evaluated adaptation, accuracy and robustness of the models generated with BLR and with MARS. To this aim, two validation strategies were developed, applying a random partition to the same landslide archive.

First, the study area was divided into 249,994 10-m grid cells corresponding to the pixels of the employed DEM. This data set includes 2,231 "event" or "positive" cells (i.e. cells hosting at least one LIP) and 247,763 "non-event" or "negative" cells (i.e. cells not intersecting any LIP). Through random selection, 50 balanced data sets were created, each of them containing all event cells and an equal number of randomly selected negative cells (Conoscenti, C. *et al.* 2016), thus including in total 4,462 cells.

The first validation strategy involved the calibration and validation of one model for each of the 50 data sets. Therefore, each data set was exploited as both learning and validation data set. In the second validation scheme, each of the 50 data sets was randomly divided into two balanced subsets: a training set, including 75 per cent of the cases, and a test set, including the remaining 25 per cent of the cases.

For both the validation schemes, it was possible to obtain a pair of models, one generated with the BLR and one with MARS, for each balanced data set *(Figure 4* and *5)*. This allowed us to analyse the difference in terms of performance and robustness between the two employed statistical techniques. As training and test datasets were the same, these differences were assumed as due only to the different characteristics of the two statistical techniques. Statistical analyses were carried out to evaluate and quantify the goodness of fit, the prediction skill and the robustness of the models.

By comparing the prediction image of each model with the spatial occurrence of the event cells, the confusion matrix and thus the number of true positive, true negative, false positive and false negative cases (TP, TN, FP and FN, respectively) for each model, applying a Youden index optimized cut-off (Youden, W.J. 1950).

To evaluate the goodness of fit and prediction skill of the susceptibility models the AUC (area under the receiver operating characteristic [ROC] curve) (Goodenough, D.J. *et al.* 1974; Hanley, J.A. and McNeil, B.J. 1982; Lasko, T.A. *et al.* 2005) was used. A ROC curve plots the true positive rate (sensitivity) against the false negative rate (1 – specificity), at any given cut-off value. For the AUC values, Hosmer, D.W. and Lemeshow, S. (2000) identify the threshold values of 0.7, 0.8 and 0.9 corresponding to acceptable, excellent and outstanding predictions respectively.

Finally, to evaluate the robustness of the models, the validation procedures have been applied to all the model runs (50 for BLR and 50 for MARS, for each validation strategy) in order to analyse the accuracy and reliability of the models through the study of the average and standard deviation of the AUC values. These validation tools have already been successfully used in previous studies with the aim of comparing different methods and models (e.g., Von Ruette, J. *et al.* 2011; Conoscenti, C. *et al.* 2015, 2016; Cama, M. *et al.* 2017).

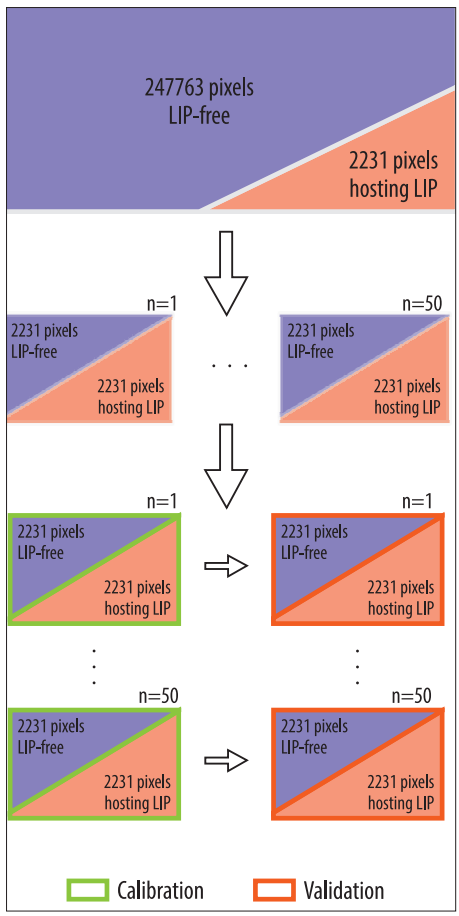


Fig. 4. Graphical summary scheme of the first adopted validation strategy

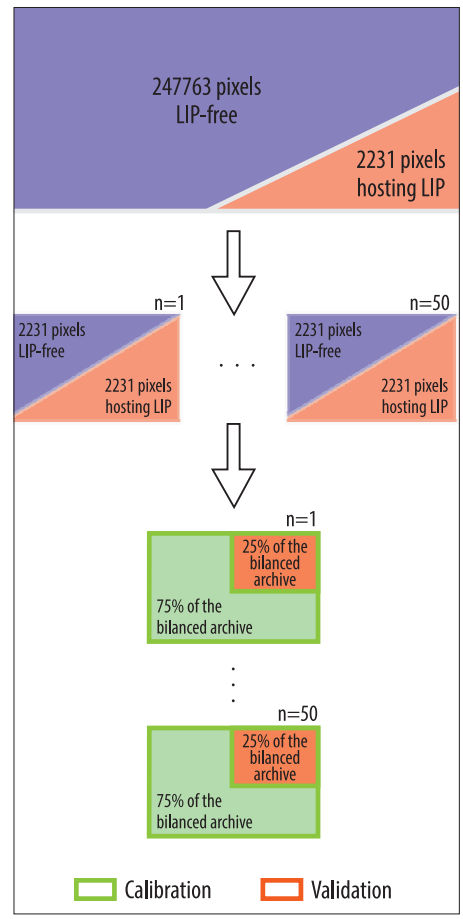


Fig. 5. Graphical summary scheme of the second adopted validation strategy

Results

For the description of the developed models and the relative results, a subscript (I) is adopted for those generated through the first validation strategy, while subscript (II) is used for those created with the second validation strategy.

The mean AUC values of the BLR (I), MARS (I), BLR (II) and MARS (II) models are 0.796, 0.821, 0.789 and 0.811, respectively. According to the classification proposed by Hosmer, D.W. and Lemeshow, S. (2000), these values indicate excellent (> 0.8) and acceptable (> 0.7) performance of the models. As shown by the AUC standard deviation values *(Table 1)*, the performance of both modelling techniques is quite stable. The boxplots of *Figure 6* show a low degree of dispersion in the AUC values, which, as expected, appears slightly higher for the second validation strategy. *Figure 7* shows the ROC curves obtained from the replicates of each model (grey) while the average ROC curves are plotted in red.

Table 2 shows the cumulative confusion matrices extracted by applying the models to the 50 validation data sets of both validation strategies. *Table 3* shows the average values

Table 1. Characteristics of the AUC values for the four susceptibility models

Models	Accuracy	AUC-mean	AUC-min	AUC-max	AUC SD*
BLR (I)	0.720	0.796	0.783	0.806	0.005
MARS (I)	0.744	0.822	0.805	0.833	0.006
BLR (II)	0.716	0.789	0.754	0.815	0.012
MARS (II)	0.736	0.811	0.779	0.836	0.012

*SD = Standard deviation.

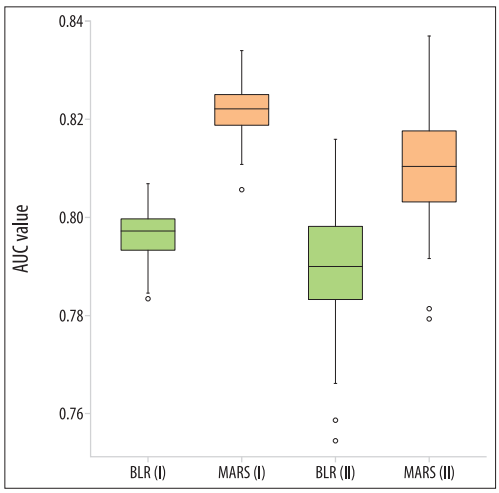


Fig. 6. AUC boxplots for the four models

of sensitivity, specificity, positive prediction value (PPV) and negative prediction value (NPV) and the relative Youden index cut-off.

The accuracy of the models can be considered good, with values between 0.71 and 0.74. Sensitivity values between 0.76 and 0.82, attest to a good predictive power of positive cases while slightly lower is the ability to discriminate the true negatives (specificity in the range 0.66–0.67). On the other hand, it is noteworthy that the NPV values, which are between 0.74 and 0.78, reveal acceptable predictions of the TNs whereas the PPV values, which are approximately 0.7, attest to a slightly worse ability to predict the TPs.

Discussion

For the discussion of the results of model validations we have to take into account that in the first validation strategy the training and the test data sets coincide whereas in the second strategy, learning and validation sets do not share any pixels and they are randomly extracted from the training/test data sets employed in the first procedure.

Models validation using the first validation strategy

MARS (I) demonstrate slightly better performance than BLR (I). It should be noted that the difference in terms of AUC is very small, being between 0.02 and 0.05. The accuracy of MARS (I) is only 0.02 higher than the accuracy of BLR (I), whereas the difference of average AUCs is only 0.03. Regarding the ability to predict event cells, a greater difference is recorded: the average sensitivity of MARS (I) is indeed 0.82 whereas that of BLR (I) is 0.77. However, PPV values reveal the same ability for both BLR (I) and MARS (I). On the other hand, although the specificity values suggest similar abilities of BLR (I) and MARS (I) to predict the non-event cells, NPV values demonstrate better performance of MARS (I) (0.78) compared to that of BLR (I) (0.74).

Models validation using the second validation strategy

Also the second validation strategy reveals a slightly better performance of MARS compared to that of BLR, although the observed differences are once again weak. The difference of both accuracy and AUC values are indeed approximately 0.02. Again, the difference in terms of sensitivity between MARS (II) (0.81) and BLR

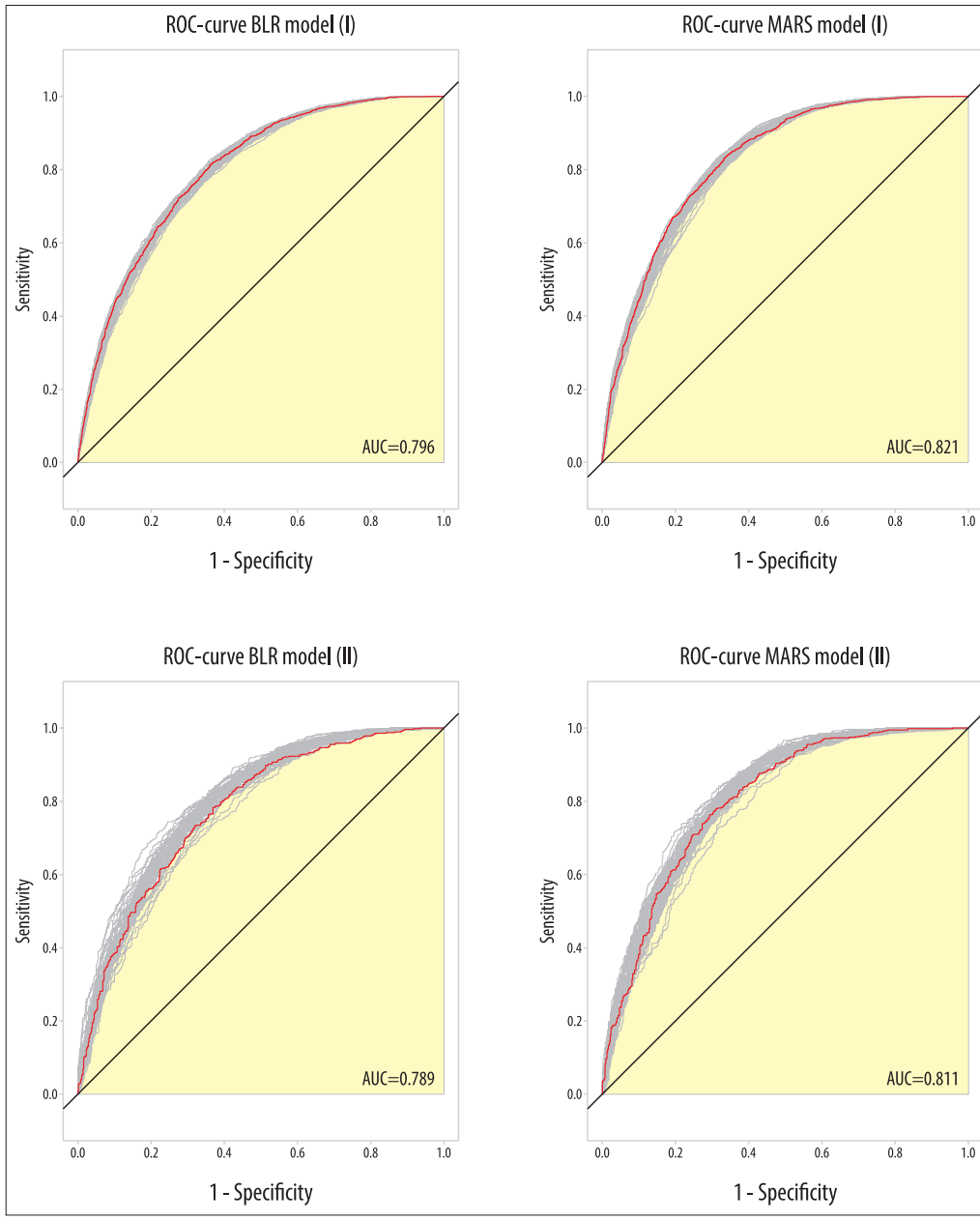


Fig. 7. ROC-plots for the four models

Table 2. Confusion matrices of the four susceptibility models

BLR (I)	Reference			MARS (I)	Reference		
		0	1			0	1
Prediction	0	74,786	25,594	Prediction	0	74,558	20,004
	1	36,764	85,956		1	36,992	91,546
BLR (II)	Reference			MARS (II)	Reference		
		0	1			0	1
Prediction	0	18,569	6,490	Prediction	0	18,453	5,234
	1	9,331	21,410		1	9,447	22,666

Table 3. Summary of the validation metrics for the four susceptibility models

Models	Youden index cut-off	Sensitivity	Specificity	Positive prediction value	Negative prediction value
BLR (I)	0.48	0.77	0.67	0.70	0.74
MARS (I)	0.46	0.82	0.66	0.71	0.78
BLR (II)	0.48	0.76	0.66	0.69	0.74
MARS (II)	0.46	0.81	0.66	0.70	0.77

(II) (0.76) does not result in a greater discriminatory power of TP (the difference of PPV is 0.01). Finally, the two techniques show the same specificity (0.66), but the discriminatory ability of TN is higher for MARS (II), with NPV values equal to 0.77 versus 0.74 of BLR (II).

Concluding remarks

The use of statistical methods in landslide susceptibility assessment raises the problem of the type of analysis to perform and which one is the best modelling approach and technique. BLR has been proven a useful technique for achieving reliable assessment of landslide susceptibility. In recent years, however, several other statistical techniques have also demonstrated equally good, and sometimes even better, performance. MARS, which is a relatively new technique, has been employed in few cases for assessing landslide susceptibility but it has already been demonstrated to provide very good accuracy in predicting the occurrence of slope failures. However, as far as we know, MARS has never been employed to predict debris flows.

The aim of this study was to highlight the differences in terms of predictive performance between BLR and MARS and, thus, identify the best method for the assessment of debris flow susceptibility in the area of Ilopango Caldera.

The obtained results show that both methods achieve good to excellent predictive performances. Although MARS demonstrated slightly better performance, the difference is too small to be able to define this technique as clearly better than BLR.

Rotigliano, E. *et al.* (2018) hypothesize that in the 2009 dataset there is a problem related to a secondary triggering of a number of phenomena due to incision or lateral erosion produced by debris flows activated directly by the storm event. In fact, even in this study, the models obtained are affected by this problem, as shown by the low specificity values. In light of this, however, the performance in terms of NPV is higher than expected. MARS, in fact, with the same dataset, is able to discriminate TN with better ability than BLR. This is probably due to the ability of MARS of identifying different relationships between the dependent and the independent variables, for different regions of the predictors' ranges. This allows MARS to overcome, even if only slightly, the problem of secondary triggering of landslides,

certainly with a better distinction of cases with respect to BLR. Furthermore, both validation strategies, albeit with subtle results, show a greater ability of MARS to identify positive cases compared to BLR.

In light of this, although the differences are not marked and certainly the results do not allow the definition of a modelling technique as absolutely better than the other, it is possible to identify more merits in the MARS technique than in the BLR.

***Acknowledgments*:** The present research was supported by a project funded by the Ministry of the Foreign Affairs of the Italian Government and carried out by the University of Palermo (resp. Prof. G. Giunta). All the authors equally contributed to the research.

REFERENCES

Atkinson, P.M. and Massari, R. 2011. Autologistic modelling of susceptibility to land sliding in the Central Apennines, Italy. *Geomorphology* 130. 55–64.

Bai, S.B., Wang, J., Lü, G.N., Zhou, P.G., Hou, S.S. and Xu, S.N. 2010. GIS-based logistic regression for landslide susceptibility mapping of the Zhongxian segment in the Three Gorges area, China. *Geomorphology* 115. 23–31.

Brabb, E.E. 1984. Innovative approaches to landslide hazard and risk mapping. In *Proceedings 4th International Symposium on Landslides* 1. Toronto, Canadian Geotechnical Society, 307–324.

Briand, L.C., Freimut, B. and Vollei, F. 2004. Using multiple adaptive regression splines to support decision making in code inspections. *Journal of Systems and Software* 73. (2): 205–217.

Bui, D.T., Tuan, T.A., Klempe, H., Pradhan, B. and Revhaug, I. 2016. Spatial prediction models for shallow landslide hazards: a comparative assessment of the efficacy of support vector machines, artificial neural networks, kernel logistic regression, and logistic model tree. *Landslides* 13. 361–378. https://doi.org/10.1007/s10346-015-0557-6

Cama, M., Lombardo, L., Conoscenti, C., Agnesi, V. and Rotigliano, E. 2015. Predicting storm-triggered debris flow events: Application to the 2009 Ionian Peloritan disaster (Sicily, Italy). *Natural Hazards and Earth System Sciences* 15. (8): 1785–1806. https://doi.org/10.5194/nhess-15-1785-2015

Cama, M., Conoscenti, C., Lombardo, L. and Rotigliano, E. 2016. Exploring relationships between grid cell size and accuracy for debris-flow susceptibility models: a test in the Giampilieri catchment (Sicily, Italy). *Environmental Earth Sciences* 75. (238): 1–21. https://doi.org/10.1007/s12665-015-5047-6

Cama, M., Lombardo, L., Conoscenti, C. and Rotigliano, E. 2017. Improving transferability strategies for debris flow susceptibility assessment. Application to the Saponara and Itala catchments (Messina, Italy). *Geomorphology* 288. 52–65. https://doi.org/10.1016/j.geomorph.2017.03.025

Carrara, A., Cardinali, M., Guzzetti, F. and Reichenbach, P. 1995. GIS technology in mapping landslide hazard. In *Geographical Information Systems in Assessing Natural Hazards*. Eds.: Carrara, A. and Guzzetti, F., Dordrecht, Kluwer, 135–175.

Chung, C.J.F., Fabbri, A.G. and Van Westen, C.J. 1995. Multivariate regression analysis for landslide hazard zonation. In *Geographical Information Systems in Assessing Natural Hazards*. Eds.: Carrara, A. and Guzzetti, F., Dordrecht, Kluwer, 107–133. https://doi.org/10.1007/978-94-015-8404-3_7

Chung, C.J.F. and Fabbri, A.G. 2003. Validation of spatial prediction models for landslide hazard mapping. *Natural Hazards* 30. (3): 451–472. https://doi.org/10.1023/B:NHAZ.0000007172.62651.2b

Conoscenti, C., Di Maggio, C. and Rotigliano, E. 2008. GIS analysis to assess landslide susceptibility in a fluvial basin of NW Sicily (Italy). *Geomorphology* 94. 325–339. Doi:10.1016/j.geomorph.2006.10.039

Conoscenti, C., Ciaccio, M., Caraballo-Arias, N.A., Gómez-Gutiérrez, Á., Rotigliano, E. and Agnesi, V. 2015. Assessment of susceptibility to earth-flow landslide using logistic regression and multivariate adaptive regression splines: a case of the Belice River basin (western Sicily, Italy). *Geomorphology* 242. 49–64. https://doi.org/10.1016/j.geomorph.2014.09.020

Conoscenti, C., Rotigliano, E., Cama, M., Caraballo-Arias, N.A., Lombardo, L. and Agnesi, V. 2016. Exploring the effect of absence selection on landslide susceptibility models: A case study in Sicily, Italy. *Geomorphology* 261. 222–235. https://doi.org/10.1016/j.geomorph.2016.03.006

Conoscenti, C., Agnesi, V., Cama, M., Caraballo-Arias, N.A. and Rotigliano, E. 2018. Assessment of Gully Erosion Susceptibility Using Multivariate Adaptive Regression Splines and Accounting for Terrain Connectivity. *Land Degradation and Development* 29. 724–736. https://doi.org/10.1002/ldr.2772

Costanzo, D., Chacón, J., Conoscenti, C., Irigaray, C. and Rotigliano, E. 2014. Forward logistic regression for earth-flow landslide susceptibility assessment in the Platani river basin (southern Sicily, Italy). *Landslides* 11. (4): 639–653. https://doi.org/10.1007/s10346-013-0415-3

Craven, P. and Wahba, G. 1979. Smoothing noisy data with spline functions. *Numerische Mathematik* 31. (4): 377–403. https://doi.org/10.1007/BF01404567

Felicísimo, Á.M., Cuartero, A., Remondo, J. and Quirós, E. 2013. Mapping landslide susceptibility

with logistic regression, multiple adaptive regression splines, classification and regression trees, and maximum entropy methods: a comparative study. *Landslides* 10. 175–189.

Friedman, J.H. 1991. Multivariate adaptive regression splines. The *Annals of Statistics* 19. (1): 1–141. http://www.jstor.org/stable/2241837

Garosi, Y., Sheklabadi, M., Porghasemi, H.R., Besalatpour, A.A., Conoscenti, C. and Van Oost, K. 2018. Comparison of differences in resolution and sources of controlling factors for gully erosion susceptibility mapping. *Geoderma* 330. 65–78. https://doi.org/10.1016/j.geoderma.2018.05.027

Gómez-Gutiérrez, Á., Schnabel, S. and Lavado Contador, F. 2009. Using and comparing two nonparametric methods (CART and MARS) to model the potential distribution of gullies. *Ecological Modelling* 220. (24): 3630–3637. https://doi.org/10.1016/j.ecolmodel.2009.06.020

Gómez-Gutiérrez, Á., Conoscenti, C., Angileri, S.E., Rotigliano, E. and Schnabel, S. 2015. Using topographical attributes to evaluate gully erosion proneness (susceptibility) in two Mediterranean basins: advantages and limitations. *Natural Hazards* 79. (1): 291–314. https://doi.org/10.1007/s11069-015-1703-0

Goodenough, D.J., Rossmann, K. and Lusted, L.B. 1974. Radiographic applications of receiver operating characteristic (ROC) curves. *Radiology* 110. 89–95.

Guzzetti, F., Carrara, A., Cardinali, M. and Reichenbach, P. 1999. Landslide hazard evaluation: a review of current techniques and their application in a multi-scale study, Central Italy. *Geomorphology* 31. 181–216. http://dx.doi.org/10.1016/S0169-555X(99) 00078-1

Heckmann, T., Gegg, K., Gegg, A. and Becht, M. 2014. Sample size matters: investigating the effect of sample size on a logistic regression susceptibility model for debris flows. *Natural Hazards and Earth System Sciences* 14. 259–278. http://dx.doi.org/10.5194/nhess-14-259-2014

Hanley, J.A. and McNeil, B.J. 1982. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 143. 29–36.

Hosmer, D.W. and Lemeshow, S. 2000. *Applied logistic regression*. Wiley Series in Probability and Statistics, Wiley. http://dx.doi.org/10.1198/tech.2002.s650

Jebur, M.N., Pradhan, B. and Tehrany, M.S. 2014. Optimization of landslide conditioning factors using very high-resolution airborne laser scanning (LiDAR) data at catchment scale. *Remote Sensing of Environment* 152. 150–165. http://dx.doi.org/10.1016/j.rse.2014.05.013

Lasko, T.A., Bhagwat, J.G., Zou, K.H. and Ohno-Machado, L. 2005. The use of receiver operating characteristic curves in biomedical informatics. *Journal of Biomedical Informatics* 38. 404–415. http://dx.doi.org/10.1016/j.jbi.2005.02.008

Lombardo, L., Cama, M., Maerker, M. and Rotigliano, E. 2014. A test of transferability for landslides susceptibility models under extreme climatic events: Application to the Messina 2009 disaster. *Natural Hazards* 74. (3): 1951–1989. https://doi.org/10.1007/s11069-014-1285-2

MARN 2010. *Síntesis de los informes de evaluación técnica de las lluvias del 7 y 8 de noviembre 2009 en El Salvador: Análisis del impacto físico natural y vulnerabilidad socio ambiental* (Technical report of the tropical storm event of 7–8 November 2009 in El Salvador: Physical impact and social vulnerability analysis). San Salvador, Ministerio de Medio Ambiente y Recursos Naturales de Salvador. http://www.oas.org/summit/sisca/Download.aspx?type=C&lang=es&id=728

Menard, S. 1995. *Applied Logistic Regression Analysis*. London, Sage Publications.

Milborrow, S. 2015. *Notes on the earth package*. An online www document. http://www.milbo.org/doc/earth-notes.pdf

Milborrow, S., Hastie, T. and Tibshirani, R. 2011. *Earth: Multivariate Adaptive Regression Spline Models.* R Software Package.

Naimi, B. 2015. *Uncertainty Analysis for Species Distribution Models*. R Software Package. Doi: 10.1111/j.1365-2699.2011.02523.x

Ohlmacher, G.C. and Davis, J.C. 2003. Using multiple logistic regression and GIS technology to predict landslide hazard in northeast Kansas, USA. *Engineering Geology* 69. 331–343.

Prokos, H., Baba, H., Lóczy, D. and El Kharim, Y. 2016. Geomorphological hazards in a Mediterranean mountain environment – Example of Tétouan, Morocco. *Hungarian Geographical Bulletin* 65. (3): 283–295. https://doi.org/10.15201/hungeobull.65.3.6

R Core Team 2017. *R: A language and environment for statistical computing*. Vienna, R Foundation for Statistical Computing. https://www.R-project.org/

Rotigliano, E., Agnesi, V., Cappadonia, C. and Conoscenti, C. 2011. The role of the diagnostic areas in the assessment of landslide susceptibility models: a test in the Sicilian chain. *Natural Hazards* 58. (3): 981–999. https://doi.org/10.1007/s11069-010-9708-1

Rotigliano, E., Cappadonia, C., Conoscenti, C., Costanzo, D. and Agnesi, V. 2012. Slope units-based flow susceptibility model: Using validation tests to select controlling factors. *Natural Hazards* 61. (1): 143–153. https://doi.org/10.1007/s11069-011-9846-0

Rotigliano, E., Martinello, C., Hernandéz, M.A., Agnesi, V. and Conoscenti, C. 2018. Predicting the landslides triggered by the 2009 96E/Ida tropical storms in the Ilopango caldera area (El Salvador, C.A.): optimizing MARS-based model building and validation strategies. *Environmental Earth Sciences* (in press)

Van Westen, C.J., Rengers, N. and Soeters, R. 2003. Use of geomorphological information in indirect landslide susceptibility assessment. *Natural Hazards* 30. (1): 399–419.

Van Westen, C.J., Castellanos, E. and Kuriakose, S.L. 2008. Spatial data for landslide susceptibility, hazard, and vulnerability assessment: an overview. *Engineering Geology* 102. 112–131.

Verstappen, H.T. 1983. Geomorphology of the Agri valley, Southern Italy. *ITC Journal* 4. 291–301.

Von Ruette, J., Papritz, A., Lehmann, P., Rickli, C. and Or, D. 2011. Spatial statistical modeling of shallow landslides-validating predictions for different landslide inventories and rainfall events. *Geomorphology* 133. 11–22. http://dx.doi.org/10.1016/j.geomorph.2011.06.010

Vorpahl, P., Elsenbeer, H., Märker, M. and Schröder, B. 2012. How can statistical models help to determine driving factors of landslides? *Ecological Modelling* 239. 27–39.

Weber, H.S., Wiesemann, G., Lorenz, W. and Schmidt-Thome, M. 1978. *Mapa geologico de la Republica de El Salvador/America Central, 1:100,000* (Geological map of the El Salvador Republic/Central America, 1:100,000). Hannover, Bundesanstalt für Geowissenschaften und Rohstoffe.

Youden, W.J. 1950. Index for rating diagnostic tests. *Cancer* 3. (1): 32–35. Doi: 10.1002/1097-0142(1950)3:1<32::AID-CNCR2820030106>3.0.CO;2-3

DOI: 10.15201/hungeobull.67.4.6 Hungarian Geographical Bulletin **67** 2018 (4) 375–392.

Identification of roofing materials with Discriminant Function Analysis and Random Forest classifiers on pan-sharpened WorldView-2 imagery – a comparison

Dávid ABRIHA[1], Zoltán KOVÁCS[1], Sarawut NINSAWAT[2], László BERTALAN[1], Boglárka BALÁZS[1] and Szilárd SZABÓ[1]

Abstract

Identification of roofing material is an important issue in the urban environment due to hazardous and risky materials. We conducted an analysis with Discriminant Function Analysis (DFA) and Random Forest (RF) on WorldView-2 imagery. We applied a three- and a six-class approach (red tile, brown tile and asbestos; then dividing the data into shadowed and sunny roof parts). Furthermore, we applied pan-sharpening to the image. Our aim was to reveal the efficiency of the classifiers with a different number of classes and the efficiency of pan-sharpening. We found that all classifiers were efficient in roofing material identification with the classes involved, and the overall accuracy was above 85 per cent. The best results were gained by RF, both with three and with six classes; however, quadratic DFA was also successful in the classification of three classes. Usually, linear DFA performed the worst, but only relatively so, given that the result was 85 per cent. Asbestos was identified successfully with all classifiers. The results can be used by local authorities for roof mapping to build registers of buildings at risk.

Keywords: remote sensing, pan-sharpening, asbestos, machine learning

Introduction

We are witnessing the increasing relevance of remote sensing in all areas of life. The first applications aimed at the analysis of land use and land cover (LULC), and then, parallel with the wider palette of satellite and aerial images, the detection of changes became the focus of research (Szabó, S. *et al.* 2016; Gulácsi, A. and Kovács, F. 2018). The geometric resolution of images has improved from 80–100 m to about 1 m over the last 30–40 years; furthermore, there are images (e.g. Landsat and Sentinel) which can be obtained for free. Another tendency is the increase in thematic resolution due to the larger number of spectral bands. Consequently, while first studies attempted to classify aggregated land cover classes (e.g. forests, grasslands or artificial surfaces), nowadays it is possible to produce species-level habitat maps (Burai, P. *et al.* 2015, 2016; Deák, M. *et al.* 2017).

Beside LULC mapping new research targets have emerged with the improved possibilities of remotely sensed data. One of these new areas is the mapping of roofing materials (Nagyváradi, L. *et al.* 2011; Mucsi, L. *et al.* 2017). The topic has its relevance from various perspectives: materials can be flammable (wooden, hay) or can be risk factors of 'carcinogenicity' (asbestos). In this study we focus on traditional roofing materials and asbestos (Cilia, C. *et al.* 2015; Wilk, E. *et al.* 2015;

[1] Department of Physical Geography and Geoinformatics, University of Debrecen, Debrecen, Hungary. Correspondent author's e-mail: szabo.szilard@science.unideb.hu

[2] Department of Remote Sensing and Geographic Information Systems (RS-GIS), Asian Institute of Technology, Bangkok, Thailand.

Gibril, M.B.A. *et al.* 2016; Krówczyńska, M. *et al.* 2016). Asbestos mapping has a wide literature, and several authors have attempted to identify this dangerous type of roofing. If the material is damaged or weathered, asbestos causes diseases such as asbestosis, mesothelioma or lung-cancer, due to its microfiber content (Mándi, A. *et al.* 2000; Petja, P.M. *et al.* 2010). Most municipalities have to face this environmental issue, but, even at settlement-level, comprehensive inventories do not exist; therefore, a supervised classification of remotely sensed images with an appropriate accuracy assessment provides an accessible alternative (Comber, A. *et al.* 2012).

Roof mapping uses both aerial photographs and satellite images. While aerial surveys are usually conducted at an altitude of 1–6 km with 10–50 cm geometric resolution, satellite images are taken at 700–800 km with a coarser (2–30 m) resolution. In addition, quantum (photon) energy is in inverse ratio to the wavelength, which means that sensors have to scan larger areas for larger wavelengths to collect reliable data from the surface. Accordingly, panchromatic (a single spectral band with 400–500 nm bandwidth) bands always have finer resolution than narrow multispectral bands (in the case of Landsat satellites the geometric resolution of the panchromatic band is 15 m, while the multispectral bands are 30 m). Geometric resolution is a limiting factor when using satellite images: if a pixel is larger than a potential house, its pixel values mix with the pixel values of the surrounding environment and the image cannot be used for this purpose. A potential solution can be to apply the pan-sharpening method, when we improve the geometric resolution of the multispectral bands with the finer resolution of the panchromatic band. The procedure distorts the spectral profiles of the objects, but improves the spatial characteristics (Yuhendra, Alimuddin, I. Y. *et al.* 2012).

Previous studies have applied aerial hyperspectral images of different sensors (APEX, AISA, MIVIS; Taherzadeh, E. and Shafri, H.Z.M. 2013; Książek, J. 2014; Szabó, S. *et al.* 2014) and satellite images (WorldView; Taherzadeh, E. and Shafri, H.Z.M. 2013; Taherzadeh, E. *et al.* 2014; Samsudin, S.H. *et al.* 2016) and have had different degrees of success in the identification of the roof types. We aimed to reveal whether pan-sharpening can improve the classification results when using a WorldView-2 satellite image. Fine resolution raises the question of the inhomogeneous reflectance of small surfaces: roof segments with different irradiation status; i.e. those in the sun and those in the shade. We distinguished between these roof segments and studied the efficiency of this kind of reference data collection.

Materials and methods

We performed our investigations in Debrecen, which is the second largest city in Hungary (*Figure 1*). Its population is 203,000 and has various built-in areas, ranging from blocks of flats to detached houses with gardens. We selected an area where the roofing materials were diverse and contained asbestos.

Datasets and reference data

We applied a WorldView-2 (WV2) satellite image. WV2 operates at a 770 km altitude and has eight multispectral bands (coastal blue [400–450 nm], blue [450–510 nm], green [510–580 nm], yellow [585–625 nm], red [630–690 nm], red-edge [705–745 nm] and two near infrared [770–895 and 860–1,040 nm]) with a 2 m geometric resolution and a panchromatic band (450–800 nm) with a 0.5 m resolution. The image was captured at 24.07.2016 without cloud cover.

Three types of roofing types were collected: red tile, brown tile and asbestos. Red and brown tiles were both concrete tiles, the only difference was their color, and both types are popular in Hungary. Asbestos roofing was also popular due to its low price; furthermore, it has an indisputable advantage: resistance to heat and fire (Kang, D. *et al.* 2013). However, this material is a serious threat to health and both production and use has

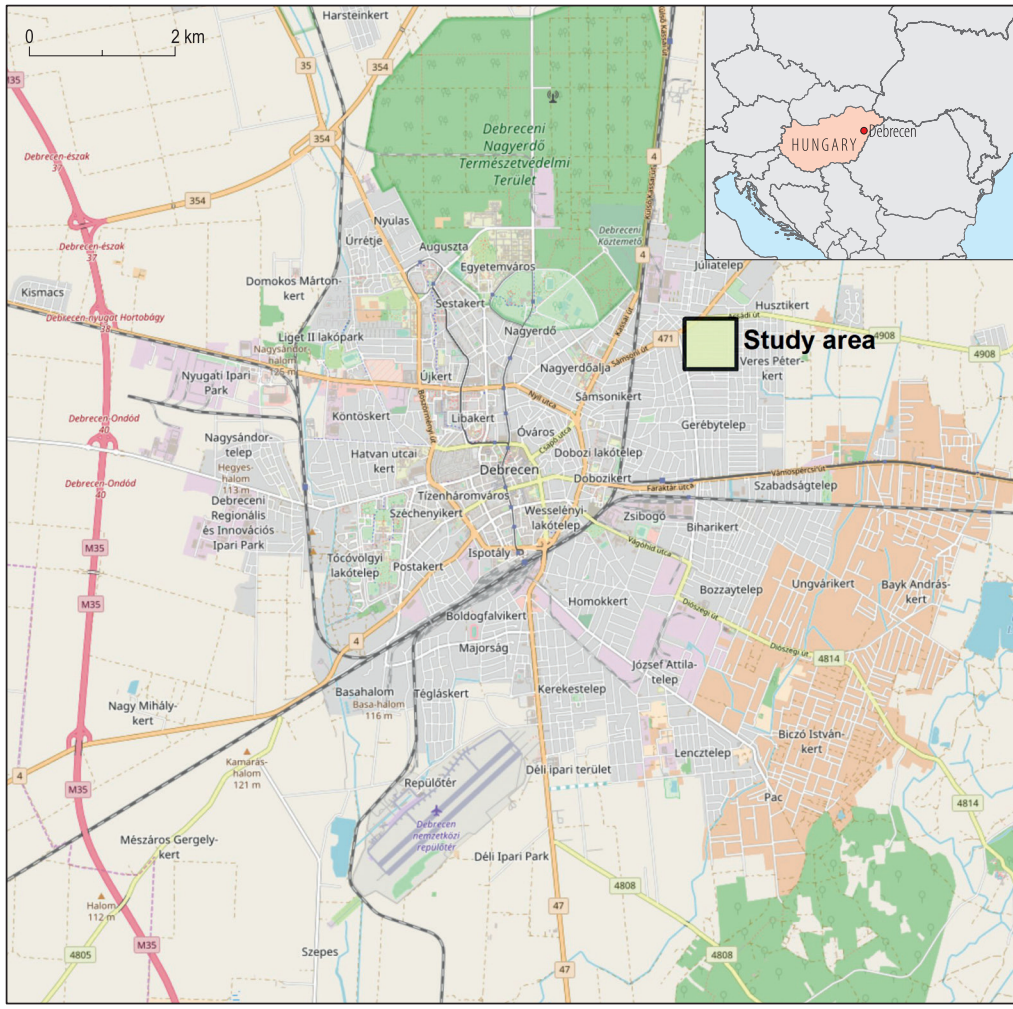


Fig. 1. Location of the study area

been banned in Hungary since 2005 (decree 41/2000 EüM-KöM).

We collected the roofing types of 350 houses as ground truth (reference) data in a field survey using a Stonex S9 RTK GPS. Later, in the GIS Laboratory, we assigned the roofs to the types observed in the field in ENVI 5.3 software (Harris Geospatial Solutions, 2017). We discriminated the different segments of the roofs considering the irradiation based on visual interpretation of the satellite image's panchromatic band. During the classifications, we intended to test whether the application of three categories or six categories (with the sunny and shadowed segments) were more efficient.

We conducted pan-sharpening with the panchromatic band fusing with the lower resolution multispectral bands: Gram-Schmidt method (Maurer, T. 2013) in ENVI 5.3 software (Harris Geospatial Solutions, 2017) was applied.

Image classification

We classified the images using two approaches. Firstly, we applied the traditional multivariate statistical classification method, the Linear Discriminant Function Analysis (LDFA) and Quadratic Discriminant Function Analysis (QDFA). This approach is an ordination (dimension reduction) technique which substitute the original variables (i.e. bands) with discriminant function (DF) scores. The resulting DF scores are derived along new axes to maximize the discrimination among the *a priori* groups (i.e. we have prior knowledge on the groups as reference data unlike in case of other ordination techniques such as Principal Component Analysis; Podani, J. 2000). DFs are calculated in the m-dimensional space defined by the input variables (m-1 dimension, where m is the number of *a priori* categories; i.e. roof types) based on "decision boundaries", depending on the input reference data (i.e. *a priori* groups). Decision boundaries or surfaces can be defined with linear or quadratic functions (Tharwat, A. 2016).

As we used more than two categories in the classification, the applied method is called multiple or "Canonical" Discriminant Function Analysis (but we did not apply it in the nomenclature). DFA supposes multivariate normality, homogeneity of covariance matrices (note: QDFA allows that issue) and, similarly to regression, is sensitive to multicollinearity (nevertheless, some researchers have found DFA to be robust when assumptions were violated; Stevens, J. 1996) and outlier data. It is not a common image classification technique and we intended to examine how efficient the usage of these approaches might be.

In contrast to this, the classifier applied, the Random Forest (RF) is a robust machine learning technique, and has no prerequisites regarding distribution or the variables involved (Ho, T.K. 1995; Pal, M. 2005; Pásztor, L. *et al.* 2015). RF is calculated from a large number of decision trees: in our study 500 decision trees were generated. Data was taken from the training dataset with a random selection for each decision tree; the number of variables involved was the square root of the number of the possible maximum: in our case we had 8 bands; thus, the algorithm also used 2 variables in each tree applying random selection (Breiman, L. 2001; Louppe, G. *et al.* 2013). A critical remark can be, that according to the random sampling, each run of the algorithm provides (slightly) different outcomes; i.e. the reproducibility can be an issue. Although it is true for lots of software implementations, it can be eliminated if the parameters of random sampling are also fixed such in case of R software.

We applied a building-mask layer which was produced with the help of the NDVI (Normalized Difference Vegetation Index, Rouse, J.W. *et al.* 1974) values (<0.1) and a normalized digital surface model (>3 m), which was derived from the Digital Terrain Model and the Digital Surface Model using a LiDAR survey conducted in 2013. Thus, all misclassifications concerning the non-building areas were omitted.

In order to hold the conditions constant for all classifications, we did not apply variable selection, all bands were involved for all models.

Summary of data procession is presented in *Figure 2.* Image classification was performed in R 3.4 (R Core Team, 2018) with the caret (Kuhn, M. *et al.* 2018; model building and evaluation), MASS (Venables, W.N. and Ripley, B.D. 2002; LDFA and QDFA classification), rpart (Therneau, T. and Atkinson, B. 2018; RF classification) and tidyverse (Wickham, H. 2017; data preparation) packages; and, for the visualization we used the raster package (Hijmans, R.J. 2017).

Accuracy assessment

Accuracy assessment was carried out with the reference data: we separated the reference dataset into training and testing subsets in an 80–20 per cent ratio with random selection. We used the confusion matrix for the evaluation of the classification results *(Table 1).* We reported overall accuracy (OA; Eq. 1), precision (Eq. 2), sensitivity (True Positive Rate; Eq. 3) and specificity (True Negative Rate; Eq. 4; Powers, D.M.W. 2007).

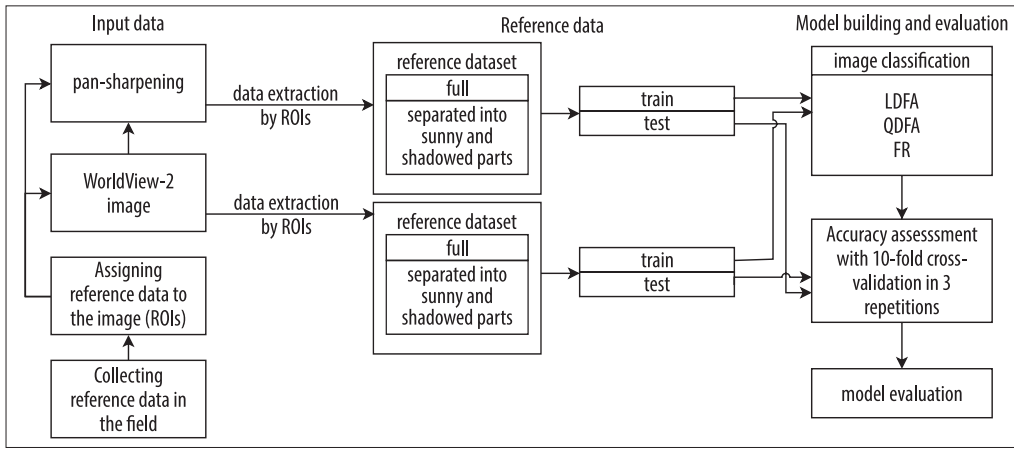


Fig. 2. Workflow of the data preparation and image classification

These equations concern the binomial approach, but we had three and six classes to predict. Therefore, we applied the "one vs. all" approach: we calculated the indices for one class to all other classes (Kassambra, A. 2018). These indices differ from the common and widely known approach of Congalton, R.G. (1991), but we intended to provide a deeper analysis of the thematic accuracy.

We applied resampling with a 10-fold cross-validation to estimate the statistical parameters of OA and used the 95 per cent confidence interval to describe the uncertainty of the outcomes. We split the training dataset into 10 subgroups and used 9 at a time to train and test on 1; the procedure was then repeated until all subgroups were used as a test dataset. Finally, the whole procedure was repeated three times. Thus, statistical parameters were derived from a resampling dataset of 30 elements for each classification algorithm (Kassambra, A. 2018). Accuracy assessment was performed in R 3.4 with the caret package (Kuhn, M. *et al.* 2018) following the methodology of Brownlee, J. (2016).

Table 1. A confusion matrix with explanations

		Observed	
	Total population	Negative	Positive
Predicted	Negative	True negative/correct omission (TN)	False negative/false omission (FN)
	Positive	False positive/false discovery (FP)	True positive/correct discovery (TP)
$Accuracy\ (OA) = \frac{TP + TN}{TP + TN + FP + FN}$			Eq. 1.
$Precision = \frac{TP}{TP + FP}$			Eq. 2.
$Sensitivity = \frac{TP}{TP + FN}$			Eq. 3.
$Specificity = \frac{TN}{TN + FP}$			Eq. 4.

Results

Pixel values and satellite bands

Coastal blue and blue bands (B1, B2) were not appropriate to identify the roofing types, as their range of the pixel values were similar. The green band (B3) was the first band which made a limited differentiation possible, but asbestos and red tiles were still similar. However, from the electromagnetic range of the yellow band (B4; from 585 nm), all the three types of roofing materials had a distinct range of pixel intensity values *(Figure 3)*. The discrimination of the sunny and shadowed roof planes caused more overlap between the pixel intensity ranges *(Figure 4)*. B1 and B2 bands were still very similar, and the different irradiation segments did not help to distinguish them. A common feature of the similarity was that the shadowed segments of asbestos were very similar to the sunny segments of brown tiles in almost every band, except for the B8 band, where the shadowed segments of red tiles were similar to this class.

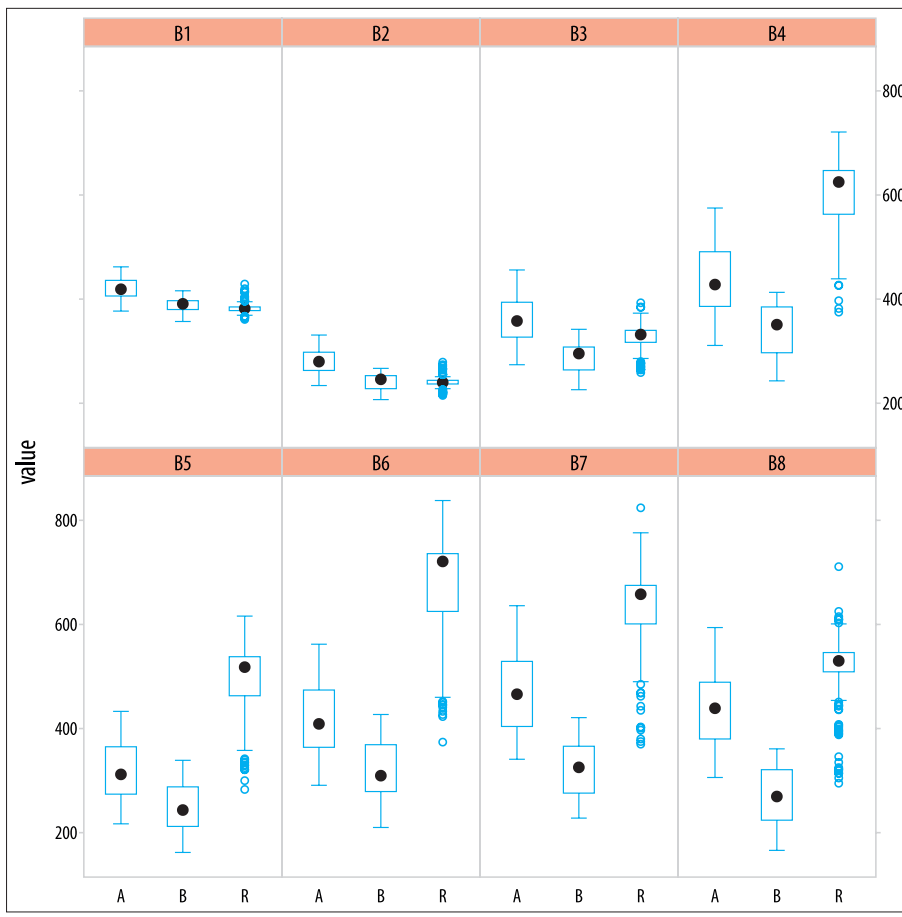


Fig. 3. Pixel value distribution of the roofing materials by bands (B1–B8) of WV-2 using three classes. – A = asbestos; B = brown tile; R = red tile

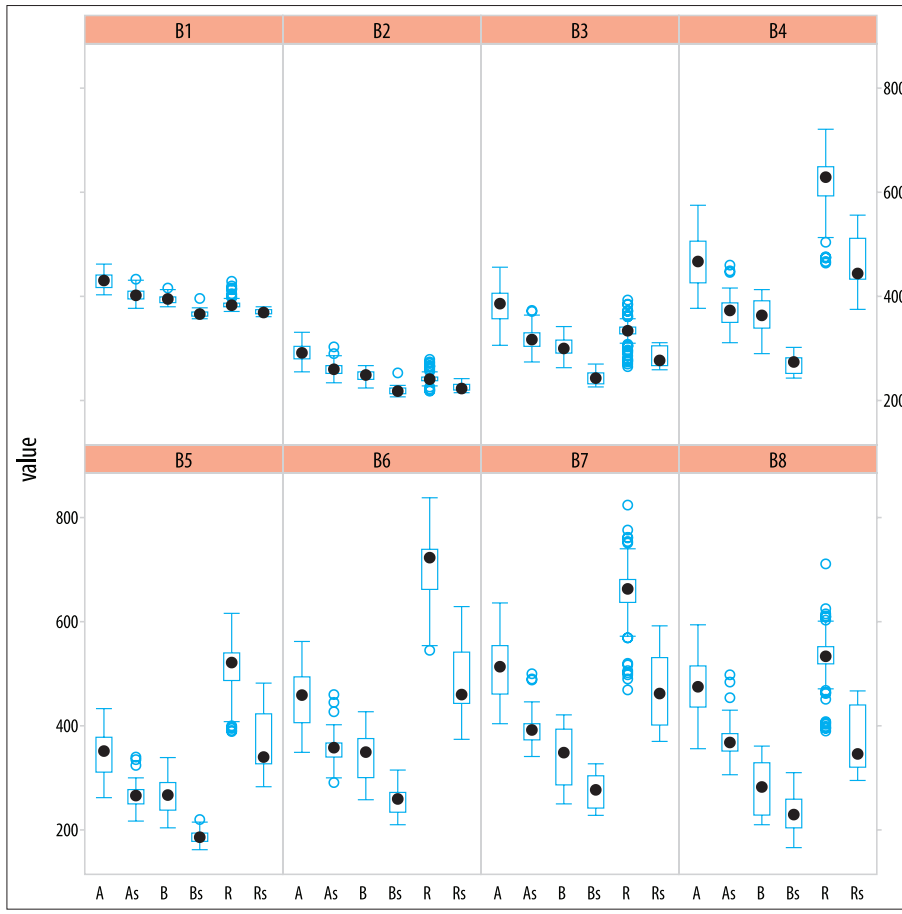


Fig. 4. Pixel value distribution of the roofing materials by bands (B1–B8) of WV-2 using six classes. – A = asbestos; B = brown tile; R = red tile; As, Bs, Rs = A, B and R with shadowed subclasses, respectively

Evaluation of the classification performance

Classifications provided the maps of the roofing materials *(Figures 5* and *6)* with varying accuracy and reliability. Although RF classifier seemed more reliable visually, we evaluated the results with the indices of accuracy assessment, too.

Generally, LDFA performed the worst, while QDFA and RF provided similar better results, with only a few percentage points of difference *(Tables 2* and *3).* However, classifications were successful in every case; considering the OA, the worst result was 0.848 and the best 0.996. All indices of accuracy assessment showed good results; however, there were some lower values, too.

Using the original bands resulted in 2–3 per cent worse OAs compared to the pan-sharpened input data. Besides, indices of class level accuracy also indicated good classifications. In the case of LDFA the improvement with the pan-sharpened images were not obvious because even if there were better results, some others became worse (e.g. precision changed from 0.970 to 0.955 in the case of asbestos).

Table 2. Accuracy assessment indices of the classification conducted with all original satellite bands with three classes

Method	Indicator	Red tile	Brown tile	Asbestos
LDFA	OA	0.949	–	–
	Precision	1.000	0.823	0.970
	Sensitivity	0.933	1.000	0.943
	Specificity	1.000	0.953	0.977
QDFA	OA	0.974	–	–
	Precision	1.000	0.928	0.972
	Sensitivity	0.966	0.984	1.000
	Specificity	1.000	0.928	0.977
RF	OA	0.962	–	–
	Precision	1.000	0.886	0.971
	Sensitivity	1.000	0.929	0.943
	Specificity	1.000	0.969	0.977

Table 3. Accuracy assessment indices of the classification conducted with all pan-sharpened satellite bands with three classes

Method	Indicator	Red tile	Brown tile	Asbestos
LDFA	OA	0.961	–	–
	Precision	1.000	0.893	0.955
	Sensitivity	0.963	0.948	0.963
	Specificity	1.000	0.975	0.967
QDFA	OA	0.995	–	–
	Precision	0.998	0.977	1.000
	Sensitivity	0.989	0.995	1.000
	Specificity	0.998	0.994	1.000
RF	OA	0.996	–	–
	Precision	0.998	0.991	0.996
	Sensitivity	1.000	0.986	0.997
	Specificity	0.999	0.998	0.996

Table 4. Accuracy assessment indices of the classification conducted with all original satellite bands with six classes (sunny and shadowed sides of roof planes)

Method	Indicator	Red tile	Red tile in shadow	Brown tile	Brown tile in shadow	Asbestos	Asbestos in shadow
LDFA	OA	0.848	–	–	–	–	–
	Precision	1.000	0.285	0.769	1.000	0.954	0.785
	Sensitivity	0.814	0.666	0.909	1.000	0.913	0.916
	Specificity	1.000	0.934	0.955	0.333	0.982	0.955
QDFA	OA	0.924	–	–	–	–	–
	Precision	1.000	0.600	0.833	0.666	1.000	0.9167
	Sensitivity	0.925	1.000	0.909	0.986	0.956	0.9167
	Specificity	1.000	0.973	0.970	0.666	1.000	0.9851
RF	OA	0.886	–	–	–	–	–
	Precision	1.000	0.600	0.769	1.000	0.916	0.777
	Sensitivity	0.963	1.000	0.909	0.666	0.956	0.583
	Specificity	1.000	0.973	0.955	1.000	0.964	0.970

Table 5. Accuracy assessment indices of the classification conducted with all pan-sharpened satellite bands with six classes (sunny and shadowed sides of roof planes)

Method	Indicator	Red tile	Red tile in shadow	Brown tile	Brown tile in shadow	Asbestos	Asbestos in shadow
LDFA	OA	0.877	–	–	–	–	–
	Precision	0.959	0.833	0.897	0.788	0.935	0.653
	Sensitivity	0.947	0.735	0.852	0.820	0.850	0.859
	Specificity	0.979	0.990	0.981	0.990	0.976	0.933
QDFA	OA	0.941	–	–	–	–	–
	Precision	0.979	0.800	0.975	0.980	0.974	0.811
	Sensitivity	0.962	0.882	0.987	0.980	0.908	0.926
	Specificity	0.989	0.986	0.996	0.999	0.990	0.968
RF	OA	0.988	–	–	–	–	–
	Precision	1.000	0.985	0.987	1.000	0.977	0.979
	Sensitivity	1.000	1.000	0.987	1.000	0.994	0.933
	Specificity	1.000	0.999	0.998	1.000	0.990	0.997

However, in the case of QDFA and RF the decrease in the indices was only between 1–2 per cent, although the increase was more than 10 per cent (e.g. in the case of brown tile, precision increased from 0.886 to 0.991).

When we discriminated the roof segments based on the irradiation, the thematic accuracy reflected the observations reported in the previous section: the higher number of classes caused lower OA values *(Tables 4* and *5).* LDFA had the weakest performance with its 0.848 OA value, while QDFA provided a very efficient solution (OA: 0.924) with the original bands. However, the accuracy improved by 10.2 per cent for RF when we applied the pan-sharpened bands.

We experienced the worst performance with LDFA with the original bands: precision was only 0.285 for the shadowed red tiles. Pan-sharpening improved it to a relevant degree; the new outcome with pan-sharpened bands became 0.833 and an increase was observed in each class (see *Tables 4* and *5*). RF and QDFA had similar results with three categories, and QDFA performed better with the original bands, the application of the pan-sharpened images resulting in a 4.7 per cent better performance. This result was somewhat below the best classification with three classes.

Evaluation of pan-sharpening and classification performance

According to the evaluation of the indices of accuracy assessment by the performance of classifiers in identifying the different roofing materials, we can observe that brown tiles usually fell outside 95 per cent accuracy (4 occurrences), while red tiles and asbestos had only one and two occurrences, respectively *(Figure 7).* Along the sensitivity and precision indices LDFA's performance was the worst, with five occurrences outside the 95 per cent limit, but the relatively good RF and QDFA also had two and one occurrences, respectively. Furthermore, pan-sharpened images were the most accurate considering thematic accuracy, only one occurrence was outside the 95 per cent quadrant. Sensitivity measures were usually higher than precision, ranging from 0.93, while precision had the lowest value at 0.82.

From another point of view, plotting the accuracy assessment indices along the original and pan-sharpened bands, we observed that pan-sharpened images were clustered in the upper 95 per cent quadrant *(Figure 8)* with only two exceptions. The range and the ratio of indices outside the 95 per cent limit were higher in the case of original bands, too.

Discussion

In total we built 12 types of model and their performance varied by their efficiency in discriminating the 3 or 6 classes using the original or pan-sharpened satellite bands *(Figure 9).* The first eight classifications had an OA higher than 95 per cent.

Considering the classified outputs (see *Figures 5* and *6*), we can spot error generated from misclassifications visually, i.e. salt and pepper appearance of different roofing classes within a dominant patch of roofs. This phenomenon is acceptable in pixel-based techniques and should be ignored in the interpretation.

Multiple Discriminant Function Analysis is a common classifier in remote sensing, but its usage is overshadowed by robust machine learning techniques; thus, nowadays this technique is not a common one in this filed. Several authors applied it but usually used its extensions or modifications (Chhikara, R.S. and Odell, P.L. 1973; Switzer, P. 1980; Du, Q. and Nekovei, R. 2005; Du, Q. and Younan, N. 2008; Wina, Herwindiati, D.E. and Isa, S.M. 2014). Authors sometimes apply ordination techniques (e.g. Principal Component Analysis) as a data preparation method prior to DFA to eliminate the issues raised by multicollinearity or, when using hyperspectral images, to reduce the number variables (Bandos, T.V. *et al.* 2009); however, we did not mix the two types of dimension reduction, following Martínez, A.M. and Kak, A.C. (2001).

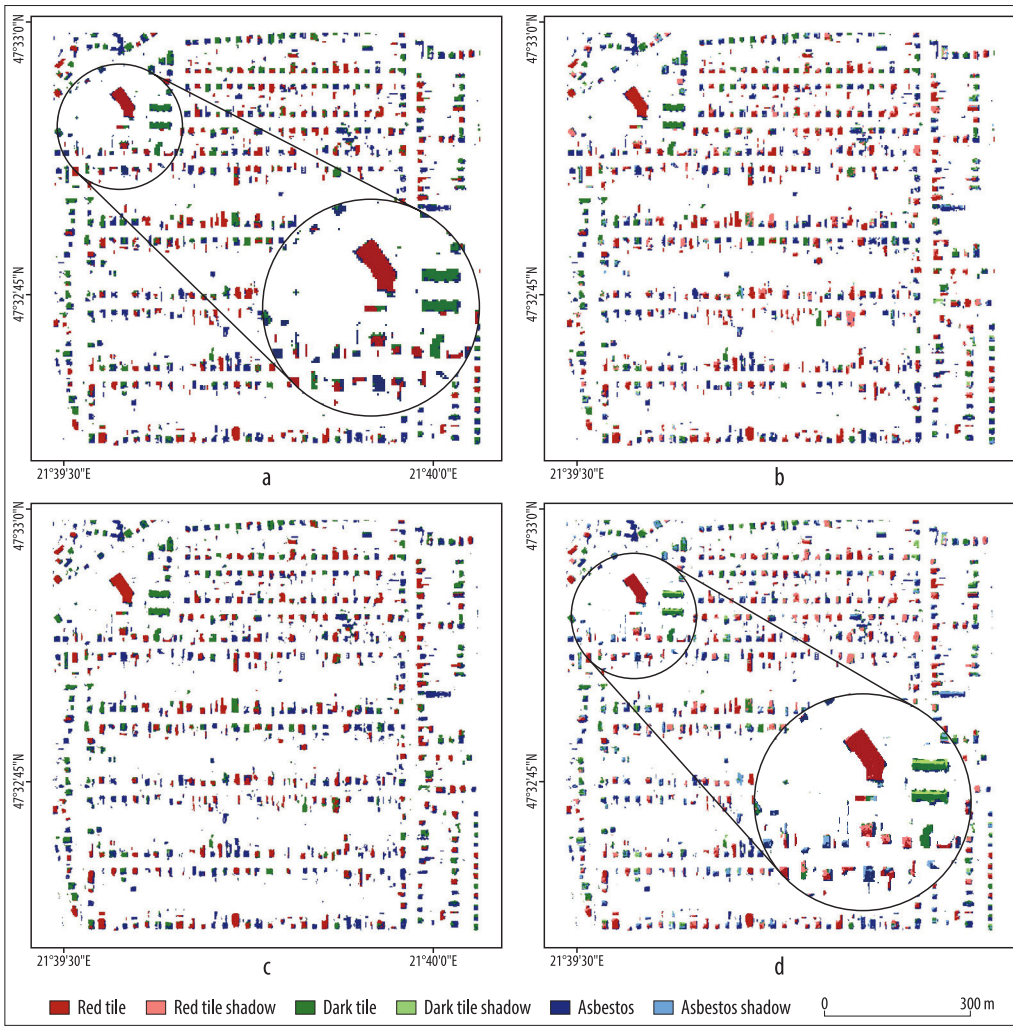


Fig. 5. Classified roofs of the study area with Linear Discriminant Function Analysis. – a = original bands with 3 classes; b = with 6 classes; c = pan-sharpened bands with 3 classes; d = with 6 classes

The linear type of Discriminant Analysis we applied is one of the most basic types and its performance was below the quadratic type and RF classifiers (see *Figure 9*). However, we must point out that even the worst result (84.8%) with the six categories reached the desired 80 per cent OA (ESRI, 1994). Nevertheless, although LDFA performed well, QDFA provided better thematic accuracy, both in the case of original or pan-sharpened bands, a result reported by other authors, too (Tharwat, A. 2016; Siqueira, L.F.S. *et al.* 2017). However, we can find exceptions, when the two types of DFA perform almost identically (Hallouche, F. *et al.* 1993; Manickavasagan, A. *et al.* 2008; Vadivambal, R. *et al.* 2010). Our results support the findings of studies revealing the better performance of QDFA: the second-best result was gained by QDFA with 3 classes.

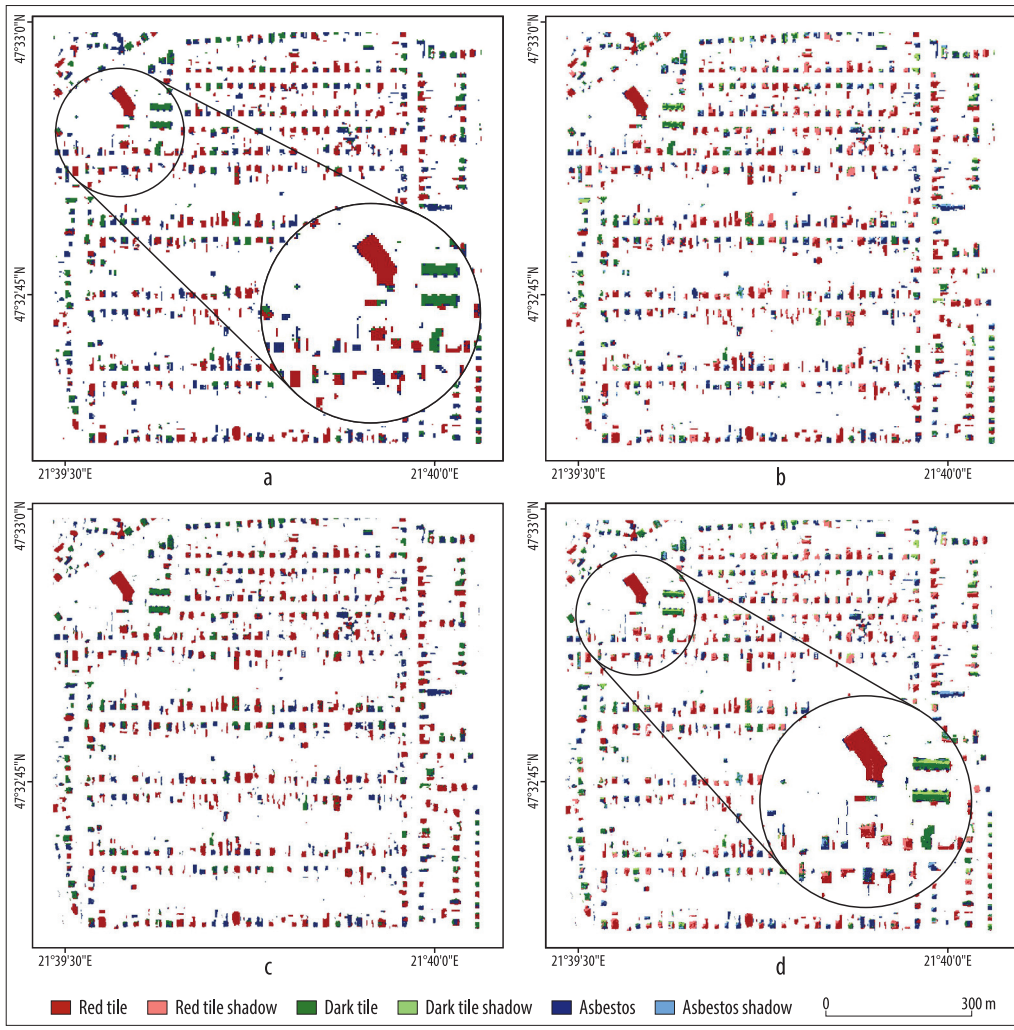


Fig. 6. Classified roofs of the study area with Random Forest – a–d = For explanation see *Fig 5.*

LDFA's best performance was only the 5th in the ranking. Both types of discriminant function classifiers were outperformed by the RF.

RF classifier resulted in good thematic accuracy from the applied algorithms, regardless of the number of classifiers. Sometimes it also provided weaker results with the original bands, but when the spatial resolution was increased, all indices of classification performance were above 95 per cent, and mostly above 98 per cent, indicating efficiency and robustness, as the distribution of the variables (i.e. bands) were not normal. Although we can find examples of a weaker performance of RF than other machine learning methods such as Support Vector Machine or Artificial Neural Networks (Statnikov, A. *et al.* 2008; Prančkevicius, T. and Marcinkevičius, V. 2017), we have to note that the difference between them

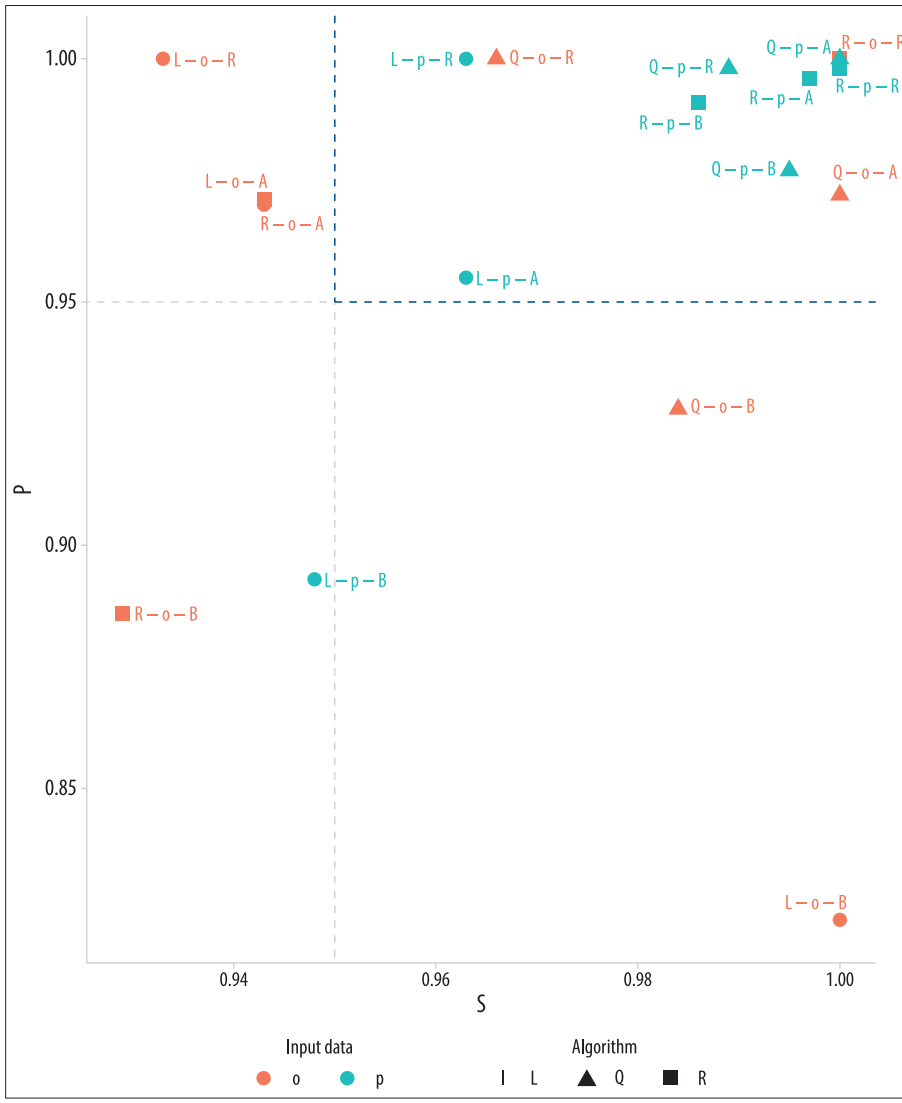


Fig. 7. Thematic accuracy based on Sensitivity (S) and Precision (P), highlighting the upper 95 per cent quadrant. – First letter of the labels: L = LDFA; Q = QDFA; R = RF. Second letter: o = original bands; p = pan-sharpened bands. Third letter: A = asbestos; B = brown tile; R = red tile

was below 3–5 per cent and it was considerably low only in some cases (Raczko, E. and Zagajewski, B. 2017).

These outcomes were the results of classifications conducted on only two classes (i.e. true/false), but when we include more classes RF can outperform the other classifiers (Fernández-Delgado, M. *et al.* 2014; Balázs, B. *et al.* 2018). In this case, RF was very efficient; nevertheless, the comparison revealed that QDFA can be very efficient, too. Considering the rank of the resampled OAs,

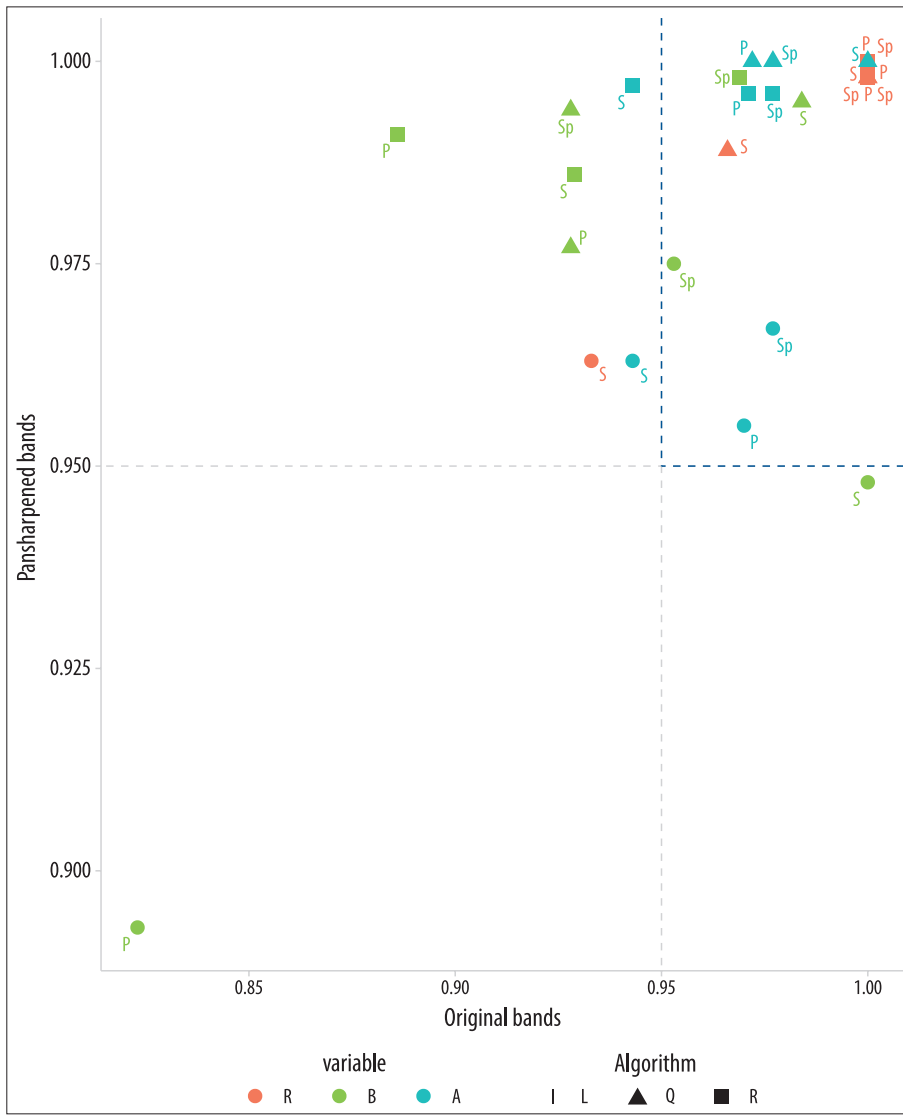


Fig. 8. Thematic accuracy based on the types of the bands involved, highlighting the upper 95 per cent quadrant – S, P, L, Q, R and coloured circles A, B and R = for explanation see *Fig 7.*

the most important result is that RF also performed well, both with three and six classes (see *Figure 9),* and except for the RF there was no other classifier in the first half of the ranking regarding the models of six classes which could achieve an OA as high as RF. Another important observation is the small range of 95 per cent confidence of RF with both three and six classes; thus, the reliability of the results was also excellent.

Most of the results can be explained by the advantages of pan-sharpening. Roofs have

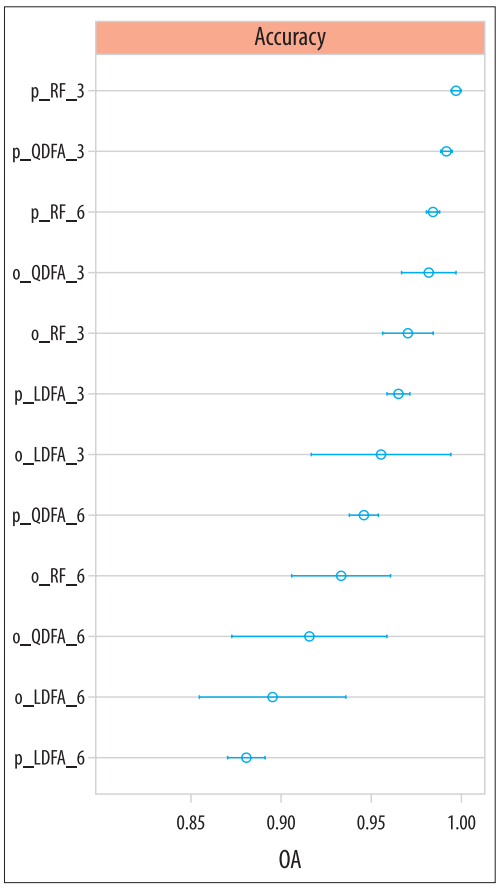


Fig. 9. Decreasing rank of Overall Accuracy (OA) of the applied classifiers. – o = original bands; p = pan-sharpened bands; LDFA = Linear Discriminant Function Analysis; QDFA = Quadratic Discriminant Function Analysis; RF = Random Forest; 3 or 6 = number of classes

large surfaces but also have several roof planes depending on the geometry types (e.g. flat, gable, pyramid hip, mansard etc.) and all roof planes can have different spectral profiles. In the studied area the roof area of most houses is about 100 m^2, and the most common geometry types are pyramid and gable roofs. Geometry can be more complex when cross gabled or cross hipped roofs or gable roofs are combined with valley roofs. Furthermore, these roof planes are fractioned with dormer windows, skylights, roof windows, chimneys or different vents. Therefore, the probability that reflectance of a 2×2 m pixel is biased by roof plane geometry is very high. Besides, this 4 m^2 pixel does not necessarily cover only the roof; it is also probable that the reflectance of the roofs and the surroundings of the roofs constitute the spectral profile. Accordingly, when we improve the geometric resolution to 0.5 m with the pan-sharpening procedure, the resultant pixel's area will be 0.25 m^2.

Although several authors have reported that pan-sharpening alters the spectral consistency (Alparone, L. *et al.* 2004; Ehlers, M. *et al.* 2010), Padwick, C. *et al.* (2010) developed a method to overcome this issue for WV-2 satellite images. We also found that there was no statistical difference between the original and pan-sharpened bands considering the reference database. Furthermore, we achieved the best results with the pan-sharpened input bands: out of the best six classifications five were with pan-sharpened bands, and only QDA had sufficient accuracy to reach 4th place.

Considering the class-level accuracy measures, we revealed that brown tiles were outside the 95 per cent accuracy quadrat. From the perspective of classification, this shows that the spectral features were similar to the asbestos roofing materials; nevertheless, asbestos had smaller issues with misclassification (see *Figures 7* and *8*). Previous studies produced different results regarding the identification of roofing materials. In the work of Szabó, S. *et al.* (2014) the OA was between 60–80 per cent, and the asbestos was identified with accuracies of 23–98 per cent (considering error omission and commission) with different approaches, which is worse than the accuracy achieved in this study, although they applied 10 classes of roofing types. The results of Abriha, D. (2017) were similar; he achieved 66–79 per cent accuracy in roof identification, and the asbestos was identified with an accuracy of 67–100 per cent. In this case, the discrimination of shadowed and sunny roof parts was

ambiguous: while the sunny asbestos roof segments were identified with 98–100 per cent accuracy, in the case of shadowed parts the omission error was high (72%). Barakat, D. *et al.* (2017) developed a rule-based method and achieved 93 per cent OA in asbestos identification with WV-2 data. Krówczyńska, M. *et al.* (2016) achieved 95 per cent OA with hyperspectral data, but the asbestos identification was biased by 38 per cent omission and 27 per cent commission errors.

Although the investigation was performed within a small study area, the method can be regarded as a general methodology: data collection, modelling and accuracy assessment can be generalized and applied in any types area regardless of the extent. Furthermore, traditional statistical analysis or machine learning can provide valuable data for all types of geographical analyses (e.g. Allen, C. *et al*. 2016; Szabó, Z. *et al*. 2017; Balázs, B. *et al*. 2018; Enyedi, P. *et al.* 2018). Our study focused on image classification, but the procedure also works with tabular data. In our case the large number of roofs identified, and the careful segmentation of the roof planes yielded the relatively high accuracy. However, it is not a general panacea as the outcome depends on the reference data. This can be promising for municipalities when they decide to refine roof registers based on remotely sensed data, as satellite images are cheaper than unique aerial hyperspectral surveys.

Conclusions

We conducted an analysis on a WorldView-2 satellite image with LDFA, QDFA and RF classifiers. We investigated the effect of the number of classes and the potential efficiency of pan-sharpening. Our results revealed that:

- discriminating the shadowed and sunny roof tiles did not improve the classification accuracy: results were up to 6–7 per cent worse when compared to the simple approach where the training dataset contained both the shadowed and sunny pixels;
- pan-sharpening was an effective technique to improve the classifications: it usually caused a 2–3 per cent better overall accuracy, but in the case of RF with the six classes the improvement was 10 per cent;
- regarding the classification algorithms, all of them performed well, but the best results were gained with Random Forest; besides, Random Forest was the most effective classifier with six classes;
- DFA-techniques performed better with fewer classes and QDFA outperformed LDFA;
- the resampling technique with the 10-fold cross-validation is an effective tool for the comparison of different classifiers.

Acknowledgement: The research was supported by the European Union and the State of Hungary, co-financed by the European Regional Development Fund (EFOP-3.6.1-16-2016-00022) and the TNN 123457 project. Abriha, D. was supported by the ÚNKP-18-2-I-DE-81 New National Excellence Program of the Ministry of Human Capacities.

REFERENCES

Abriha, D. 2017. *Roofing material determination with hyperspectral data.* Student V4 Geoscience Conference and Scientific Meeting GISÁČEK, Ostrava, Czech Republic, 5–10.

Allen, C., Tsou, M-H., Aslam, A., Nagel, A. and Gawron, J-M. 2016. Applying GIS and Machine Learning Methods to Twitter Data for Multiscale Surveillance of Influenza. *PLoS ONE* 11. (7): e0157734. https://doi.org/10.1371/journal.pone.0157734

Alparone, L., Baronti, S., Garzelli, A. and Nencini, F. 2004. A global quality measurement of pan-sharpened multispectral imagery. *IEEE Geoscience and Remote Sensing Letters* 1. (4): 313–317.

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

Bandos, T.V., Bruzzone, L. and Camps-Valls, G. 2009. Classification of hyperspectral images with regularized linear discriminant analysis. *IEEE Transactions on Geoscience and Remote Sensing* 47. (3): 862–873.

Barakat, M.A., Shafri, H.Z.M. and Hamedianfar, A. 2017. New semi-automated mapping of asbestos cement roofs using rule-based object-based image analysis and Taguchi optimization technique from WorldView-2 images. *International Journal of Remote Sensing* 38. (2): 467–491.

Breiman, L. 2001. Random Forests. *Machine Learning* 45. (1): 5–32.

Brownlee, J. 2016. *Machine Learning Mastery with R. Get Started, Build Accurate Models and Work Through Projects Step-by-Step.* First edition, Machine Learning Mastery. https://machinelearningmastery.com/machine-learning-with-r/

Burai, P., Lénárt, Cs., Valkó, O., Bekő, L., Szabó, Zs. and Deák, B. 2016. Fátlan vegetációtípusok azonosítása légi hiperspektrális távérzékelési módszerrel (Vegetation mapping in an alkali landscape – application of airborne hyperspectral data). *Tájökológiai Lapok* 14. (1): 1–12. (In Hungarian)

Burai, P., Tomor, T., Bekő, L. and Deák, B. 2015. Airborne hyperspectral remote sensing for identification grassland vegetation. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 40. (3): 427–431. Doi: 10.5194/isprsarchives-XL-3-W3-427-2015

Chhikara, R.S. and Odell, P.L. 1973. Discriminant analysis using certain normed exponential densities with emphasis on remote sensing application. *Pattern Recognition* 5. (3): 259–272.

Cilia, C., Panigada, C., Rossini, M., Candiani, G., Pepe, M. and Colombo, R. 2015. Mapping of Asbestos Cement Roofs and Their Weathering Status Using Hyperspectral Aerial Images. *ISPRS International Journal of Geo-Information* 4. (2): 928–941.

Comber, A.J., Fisher, P., Brunsdon, C. and Khmag, A. 2012. Spatial analysis of remote sensing image classification accuracy. *Remote Sensing of Environment* 127. 237–246. Doi:10.1016/j.rse.2012.09.005

Congalton, R.G. 1991. A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment* 37. 35–46.

Deák, M., Telbisz, T., Árvai, M., Mari, L., Horváth, F., Kohán, B., Szabó, O. and Kovács, J. 2017. Heterogeneous forest classification by creating mixed vegetation classes using EO-1 Hyperion. *International Journal of Remote Sensing* 38. (18): 5215–5231.

Du, Q. and Nekovei, R. 2005. Implementation of real-time constrained linear discriminant analysis to remote sensing image classification. *Pattern Recognition* 38. (4): 459–471.

Du, Q. and Younan, N.H. 2008. Dimensionality Reduction and Linear Discriminant Analysis for Hyperspectral Image Classification. In *Knowledge-Based Intelligent Information and Engineering Systems.* Eds.: Lovrek, I., Howlett, R.J. and Jain, L.C., KES 2008. Lecture Notes in Computer Science 5179. 392–399.

Ehlers, M., Klonus, S., Åstrand, P.J. and Rosso, P. 2010. Multi-sensor image fusion for pansharpening in remote sensing. *International Journal of Image and Data Fusion* 1. (1): 25–45.

Environmental Systems Research Institute, National Center for Geographic Information and Analysis, and The Nature Conservancy 1994. *Accuracy Assessment Procedures: NBS/NPS Vegetation Mapping Program.* Santa Barbara, CA, and Arlington, VA. Report prepared for the National Biological Survey and National Park Service, Redlands.

Enyedi, P., Pap, M., Kovács, Z., Takács-Szilágyi, L. and Szabó, S. 2018. Efficiency of local minima and GLM techniques in sinkhole extraction from a LiDAR-based terrain model. *International Journal of Digital Earth*, online version: https://doi.org/10.1080/17538947.2018.1501107

EüM-KöM 2000. EüM-KöM decree on the limitations of the activities with dangerous materials and products. *Magyar Közlöny* 41. 125. (In Hungarian)

Fernández-Delgado, M., Cernadas, E., Barro, S. and Amorim, D. 2014. Do we need hundreds of classifiers to solve real world classification problems? *Journal of Machine Learning Research* 15. 3133–3181.

Gibril, M.B.A., Shafri, H.Z.M. and Hamedianfar, A. 2016. New semi-automated mapping of asbestos cement roofs using rule-based object-based image analysis and Taguchi optimization technique from WorldView-2 images. *International Journal of Remote Sensing* 38. (2): 467–491.

Gulácsi, A. and Kovács, F. 2018. Drought monitoring of forest vegetation using MODIS-based normalized difference drought index in Hungary. *Hungarian Geographical Bulletin* 67. (1): 29–42.

Hallouche, F., Adams, A.E., Hinton, O.R., Surtees, D.P., Wadehra, V. and Sherbet, G.V. 1993. Discriminant analysis for classification of murine melanomas and human cervical epithelial cells. *Analytical and Quantative Cytollogy and Histology* 15. 50–60.

Hijmans, R.J. 2017. *raster: Geographic Data Analysis and Modeling. R package version 2.6-7.* https://CRAN.R-project.org/package=raster

Ho, T.K. 1995. *Random Decision Forests*. Proceedings of the 3rd International Conference on Document Analysis and Recognition, Montreal, QC, 14–16. August 1995. 278–282.

Kang, D., Myung, M-S., Kim, Y-K. and Kim, J-E. 2013. Systematic Review of the Effects of Asbestos Exposure on the Risk of Cancer between Children and Adults. *Annals of Occupational and Environmental Medicine* 25(1) 10. Doi: 10.1186/2052-4374-25-10.

Kassambara, A. 2018. *Machine Learning Essentials: Practical Guide in R.* 1st edition. CreateSpace Independent Publishing Platform.

Krówczyńska, M., Wilk, E., Pabjanek, P., Zagajewski, B. and Meuleman, K. 2016. Mapping asbestos-cement roofing with the use of APEX hyperspectral airborne imagery: Karpacz area, Poland – a case study. *Miscellanea Geographica* 20. (1): 41–46.

Książek, J. 2014. Methods for Detection of Asbestos-Cement Roofing Sheets. *Geomatics and Environmental Engineering* 8. (3): 59–76.

Kuhn, M., Wing, J., Weston, S., Williams, A., Keefer, C., Engelhardt, A., Cooper, T., Mayer, Z., Kenkel, B., R Core Team, Benesty, M., Lescarbeau, R., Ziem,

A., Scrucca, L., Tang, Y., Canadan, C. and Hunt, T. 2018. *caret: Classification and Regression Training.* R package version 6.0–79. https://CRAN.R-project.org/package=caret

Louppe, G., Wehenkel, L., Sutera, A. and Geurts, P. 2013. *Understanding variable importances in forests of randomized trees.* Proceedings of the 26th International Conference on Neural Information Processing Systems 1. 431–439.

Mándi, A., Posgay, M., Vadász, P., Major, K., Rödelsperger, K., Tossavainen, A., Ungváry, G., Woitowitz, H.J., Galambos, E., Németh, L., Soltész, I., Egerváry, M. and Böszörményi Nagy, G. 2000. *Role of occupational asbestos exposure in Hungarian lung cancer patients.* International Archives of Occupational and Environmental Health 73. 555–560.

Manickavasagan, A., Jayas, D.S. and White, N.D.G. 2008. Thermal imaging to detect infestation by *Cryptolestes ferrugineus* inside wheat kernels. *Journal of Stored Products Research* 44. (2): 186–192.

Martínez, A.M. and Kak, A.C. 2001. PCA versus LDA. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 23. (2): 228–233.

Maurer, T. 2013. *How to Pan-Sharpen Images Using the Gram-Schmidt Pan-Sharpen Method – a Recipe.* International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XL-1/W1. 239–244.

Mucsi, L., Liska, Cs.M., Henits, L., Tobak, Z., Csendes, B. and Nagy, L. 2017. The evaluation and application of an urban land cover map with image data fusion and laboratory measurements. *Hungarian Geographical Bulletin* 66. (2): 145–156.

Nagyváradi, L., Gyenizse, P. and Szebényi, A. 2011. Monitoring the changes of a suburban settlement by remote sensing. *Acta Geographica Debrecina Landscape and Environment* 5. 76–83.

Padwick, C., Deskevich, M., Pacifici, F. and Smallwood, S. 2010. *WorldView-2 Pan-Sharpening.* Proceedings of ASPRS 2010 Annual Conference, San Diego, CA, USA, 26–30.

Pal, M. 2005. Random forest classifier for remote sensing classification. *International Journal of Remote Sensing* 26. 217–222.

Pásztor, L., Laborczi, A., Takács, K., Szatmári, G., Dobos, E., Illés, G., Bakacsi, Zs. and Szabó, J. 2015. Compilation of novel and renewed, goal oriented digital soil maps using geostatistical and data mining tools. *Hungarian Geographical Bulletin* 64. (1): 49–64.

Petja, P.M., Twumasi, Y.A., Tengbeh, G.T. and Atanasova, M. 2010. Spatial epidemiology risk assessment for rehabilitated former asbestos mining areas in Limpopo Province, South Africa, using remote sensing and conventional analytical methods. *South African Journal of Epidemiology and Infection* 25. 32–39.

Podani, J. 2000. *Introduction to the Exploration of Multivariate Biological Data*. Leiden, Backhuys Publisher.

Powers, D.M.W. 2007. *Evaluation: From Precision, Recall and F-Factor to ROC, Informedness, Markedness & Correlation.* Technical Report SIE-07-001. Adelaide, South Australia. School of Informatics and Engineering, Flinders University Adelaide.

Pranckevičius, T. and Marcinkevičius, V. 2017. Comparison of Naive Bayes, Random Forest, Decision Tree, Support Vector Machines, and Logistic Regression Classifiers for Text Reviews Classification. *Baltic Journal of Modern Computing* 5. (2): 221–232.

Raczko, E. and Zagajewski, B. 2017. Comparison of support vector machine, random forest and neural network classifiers for tree species classification on airborne hyperspectral APEX images. *European Journal of Remote Sensing* 50. (1): 144–154.

R Core Team 2018. *R: A language and environment for statistical computing.* Vienna, R Foundation for Statistical Computing. https://www.R-project.org/

Rouse, J.W. Jr., Haas, R.H., Deering, D.W., Schell, J.A. and Harlan, J.C. 1974. *Monitoring the Vernal Advancement and Retrogradation (Green Wave Effect) of Natural Vegetation*. NASA/GSFC Type III Final Report, Greenbelt, MD.

Samsudin, S.H., Shafri, H.Z.M. and Hamedianfar, A. 2016. Development of spectral indices for roofing material condition status detection using field spectroscopy and WorldView-3 data. *Journal of Applied Remote Sensing* 10(2) 025021. Doi: 10.1117/1.JRS.10.025021

Siqueira, L.F.S., Araújo Júnior, R.F.A., de Araújo, A.A., Morais, C.L.M. and Lima, K.M.G. 2017. LDA vs. QDA for FT-MIR prostate cancer tissue classification. *Chemometrics and Intelligent Laboratory Systems* 162. 123–129.

Statnikov, A., Wang, L. and Aliferis, C. 2008. A comprehensive comparison of random forests and support vector machines for microarray-based cancer classification. *BMC Bioinformatics* 9. 319. Doi:10.1186/1471-2105-9-319

Stevens, J. 1996. *Applied multivariate statistics for the social sciences*. 3rd edition. Mahwah, NJ, USA, Lawrence Erlbaum.

Switzer, P. 1980. Extensions of linear discriminant analysis for statistical classification of remotely sensed satellite imagery. *Journal of the International Association for Mathematical Geology* 12. (4): 367–376.

Szabó, S., Gácsi, Z. and Balázs, B. 2016. Specific features of NDVI, NDWI and MNDWI as reflected in land cover categories. *Acta Geographica Debrecina Landscape and Environment* 10. (3–4): 194–202.

Szabó, Sz., Burai, P., Kovács, Z., Szabó, Gy., Kerényi, A., Fazekas, I., Paládi, M., Buday, T. and Szabó, G. 2014. Testing of algorithms for the identification of asbestos roofing based on hyperspectral data.

Environmental Engineering and Management Journal 143. (11): 2875–2880.

Szabó, Z., Tóth, C.A., Tomor, T. and Szabó, S. 2017. Airborne LiDAR point cloud in mapping of fluvial forms: a case study of a Hungarian floodplain. *GIScience and Remote Sensing* 54. 862–880.

Taherzadeh, E. and Shafri, H.Z.M. 2013. Development of a Generic Model for the Detection of Roof Materials Based on an Object-Based Approach Using WorldView-2 Satellite Imagery. *Advances in Remote Sensing* 2. (4): 312–321.

Taherzadeh, E., Shafri, H.Z.M. and Shahi, K. 2014. Roof Material Detection Based on Object-Based Approach Using WorldView-2 Satellite Imagery. *International Journal of Environmental and Ecological Engineering* 8. (10): 1826–1829.

Tharwat, A. 2016. Linear vs. quadratic discriminant analysis classifier: a tutorial. International *Journal of Applied Pattern Recognition* 3. (2): 145–180.

Therneau T. and Atkinson, B. 2018. *rpart: Recursive Partitioning and Regression Trees. R package version 4.1–13.* https://CRAN.R-project.org/package=rpart

Vadivambal, R., Vellaichamy, C., Jayas, D.S. and White, N.D.G. 2010. Detection of Sprout-Damaged Wheat Using Thermal Imaging. *Applied Engineering in Agriculture* 26. 999–1004.

Venables, W.N. and Ripley, B.D. 2002. *Modern Applied Statistics with S.* Fourth edition. New York, Springer.

Wickham, H. 2017. *Tidyverse: Easily Install and Load the 'Tidyverse'. R package version 1.2.1.* https://CRAN.R-project.org/package=tidyverse

Wilk, E., Krówczyńska, M. and Pabjanek, P. 2015. Determinants influencing the amount of asbestos-cement roofing in Poland. *Miscellanea Geographica* 19. (3): 82–86.

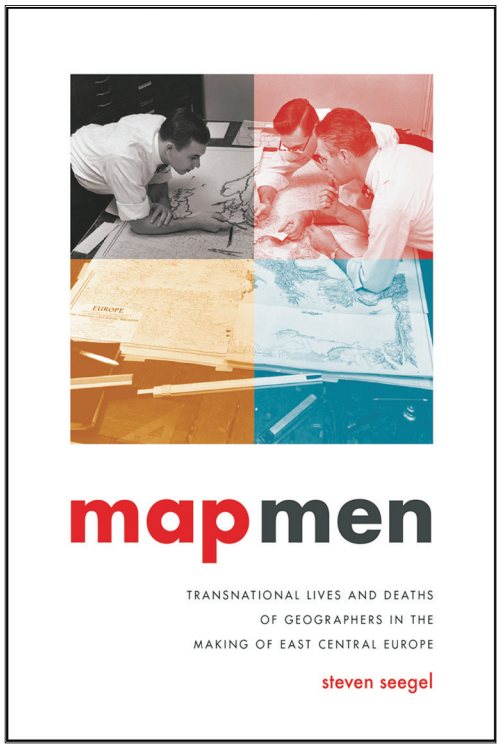
Wina, Herwindiati, D.E. and Isa, S.M. 2014. *Robust discriminant analysis for classification of remote sensing data.* International Conference on Advanced Computer Science and Information System, Jakarta, 18–19. Oct. 2014. Jakarta, IEEE Indonesia Section, 454–458.

Yuhendra, Alimuddin, I., Sumantyo, J.T.S. and Kuze, H. 2012. Assessment of pan-sharpening methods applied to image fusion of remotely sensed multi-band data. *International Journal of Applied Earth Observation and Geoinformation* 18. 165–174.

BOOK REVIEW SECTION

Seegel, S.: Map Men: Transnational Lives and Deaths of Geographers in the Making of East Central Europe. Chicago–London, The University of Chicago Press, 2018. 320 p.

One hundred years ago, fighting ceased on the major fronts of the Great War. Both the victorious Allied Powers as well as the defeated Central Powers started to prepare for peace. Though the guns had fallen silent, the peace preparations mobilised armies of experts, and in particular geographers, for a new struggle, one that required geographical expertise and scientific reasoning to justify new borders and to defend the integrity of state territory. It was in this geopolitically-charged post-war context that geographers who had been working in academia were called upon to serve their nation. The one-hundredth anniversary of the end of the Great War, therefore, is perfect timing for the publication of Steven Seegel's "Map Men: Transnational Lives and Deaths of Geographers in the Making of East Central Europe". Focusing on four East Central European geographers and one American geographer, Seegel's book tells the story of five scholars well known in the history of geography who either contributed to the making of the new borders or fought to defend the old ones in the immediate post-war period.



The oldest of Seegel's main characters is Albrecht Penck, born in Leipzig in 1858. He studied at Leipzig and from the mid-1880s served as a professor of physical geography at the University of Vienna for two decades. He was appointed chair of geography at the University of Berlin at the peak of his career, and became without doubt the most influential German geographer after the turn of the century. Polish geographer Eugeniusz Romer was born in 1871 in Lemberg/Lwów/Lviv, a city that then belonged to the Austro-Hungarian Empire. He studied in Kraków, and it was on leave to Berlin and Vienna that he was introduced to the ideas of Ferdinand von Richthofen and Penck respectively. Romer was appointed chair of geography at the university in his home town a few years before the outbreak of the Great War. Like Romer, the Ukrainian Stepan Rudnyts'kyi was also a son of Galicia and an Austrian citizen. Born in Peremyshl/Przemyśl in 1877, he first studied philosophy at the University of Lemberg/Lwów/Lviv, and later turned to geography. He attended Penck's lectures in Vienna, habilitated at Lemberg, and was allowed as a *Privatdozent* to teach in geography in Ukrainian. Next among Seegel's geographers is Isaiah Bowman, who was born in Ontario, Canada, and who was a descendant of German Swiss emigrants. His family moved from Canada to the USA when he was young. At Harvard, he studied under William Morris Davis, and after defending his thesis, he taught at Yale. The youngest among the five geographers that Seegel studies is Count Pál Teleki, who was born to a Transylvanian-Hungarian aristocratic family in Budapest in 1879. Teleki studied law at the University of Budapest, and worked at the Institute of Geography at the university under the supervision of Lajos Lóczy. His book on the history of the cartography of the Japanese Islands brought him considerable renown as a geographer, and he was elected general secretary of the Hungarian Geographical Society in 1910.

Seegel's is a story of the several links that connect these geographers. All of them loved to travel, to be outdoors, and to do fieldwork. They were connected through an international scholarly network (Bowman, Romer, and Teleki participated in the

Transcontinental Excursion of the AGS in 1912), and many of them were in correspondence for decades. Above all, they shared a passionate love of maps. According to Seegel's interpretation, theirs was not a harmless romance. Hidden behind a veil of objectivity and civility, the maps they created and revered were tools of nation building, imperialism, and propaganda. The maps themselves were not independent of their makers, Seegel argues, and so through the maps we can catch a glimpse of the personality of their authors. As Seegel writes, "the book's core argument is that interest in maps was often pathological, a sign of frustration and unfulfilled personal ambition" (p. 3). "Map Men" is not the kind of story that has a happy ending. It is, instead, a dark drama whose protagonists' lives end in tragedy.

The characters of this drama are most definitely not positive heroes in Seegel's estimation. According to Seegel, the "mobile yet place sensitive" map men were "illiberal, provincial, pre-1914 hyphenated Anglophile Germans … who envisaged geography as a new megascience" (p. 3). They were "Germans" in the sense that all of them spoke German fluently, and all of them were close to German science and German geography. "Their grasp of maps and geography," Seegel claims, "was largely antimodern, anti-urban, and, in some cases anti-Semitic" (p. 3). More than this, maps and geography were mobilised by these men "as defence of privilege and Europe's grand explorer tradition in East Central Europe" (p. 3). Both Penck and Teleki, moreover, were dedicated to irredentism after World War I. The maps that they produced and circulated were neither neutral nor purely scientific products. They were, in fact, "affective, not just rational tools" (p. 4).

Prior to the Great War, these map men had taken the opportunity to participate in a flourishing international network of scholarly life. With respect to "scientific pursuits," it was a time in which "the men saw little contradiction between nations and internationalism" (p. 40). The War, however, changed everything. The geographers became "stateside experts" (p. 227). Contradictory national interests turned them against one another. In his efforts to help create a new Poland, for example, Romer confronted both Penck and Rudnyts'kyi, with the Germans supporting the Ukrainian claims much more than the Polish ones. Teleki, advocating the Hungarian cause, trusted in the power of personal relationships in vain in 1918-20. Romer, by contrast, was successful in bending Bowman's ear. As collegial bonds started to weaken, friendships were replaced by cold emotions, even open hostility. Perhaps as consolation, service to the state opened up new vistas for these geographers after the war. Almost all of them participated in organising and leading geographical and cartographical projects that were closely connected to political goals, though none of them were as successful at fulfilling political ambitions as Teleki, who twice served as prime minister of interwar Hungary.

With the notable exception of the American Bowman, the lives of the East Central European geographers ended in tragedy. Rudnyts'kyi was the first to suffer a tragic fate. Immediately after the war he worked in Vienna and in Prague, but later accepted the invitation of the Ukrainian SSR to serve as a professor in Kharkov in the mid-1920s. When Stalinisation gained momentum, he was deemed to be a Ukrainian nationalist and "propagator of fascism in geography" (p. 143). (Rudnyts'kyi had never joined the Communist Party.) The Ukrainian Academy of Sciences expelled him, after which he was stripped of his professorship, imprisoned, and finally executed in 1937. Four years later Teleki would die at his own hands. As prime minister at the outbreak of World War II, Teleki presided over some tangible revisionist successes, but he was not able to restrain the growing Nazi German influence in Hungary, and he committed suicide when the German troops crossed the Hungarian border to attack Yugoslavia in April 1941. Both Penck and Romer, by contrast, lived out the remainder of their lives in a form of exile. Though not a member of the Nazi Party, Penck agreed more or less with its goals, but when the heavy air raids started, he fled from Berlin to the relative safety of his daughter's home in Prague, and died there in the spring of 1945. Romer, in turn, hid in a Catholic monastery as the war raged all around him. In the end he could never return to Lwów/Lviv, as the city was annexed to the Soviet Union after the war. He died in Kraków in 1954.

Seegel's work on the life and career of the five geographers is an outstanding scholarly achievement. The array of sources and literature used for the book is really impressive, even more so if we take into account that Seegel's extensive use of archival sources and printed material required a fluency in German, Polish, Ukrainian, and Hungarian that is truly rare. Only very few scholars are able to do such comparative work. More importantly, however, Seegel has produced a strong and convincing argument that clearly illustrates how politics and ideology were intertwined with scientific work. At the same time, "Map Men" also underlines how strongly the personality, faith, ambition, and fantasies of these five scholars influenced their scientific results. Scientific knowledge is inseparable from its maker, and maps are no exception. It has, in fact, become somewhat of a truism that a map is not an impartial mirror of the Earth's surface or of the spatial distribution of human phenomena, but rather is a socially-constructed scientific product that delivers its message visually. Seegel's work illustrates this reality in novel and convincing ways.

It is worth noting that the above-mentioned relationship between an author and his or her work is true in case of book reviews, too. I do not in any way want to pretend that my opinion of Seegel's recent work is definitive or final. I can review this book only from the perspective of an East Central European, or more precisely, a Hungarian geographer. In my opinion this work

is important for us, as it destroys myths, and forces us to face up to many inconvenient truths. Compared to the usual Hungarian interpretation, Seegel's study clearly illustrates the value of conclusions that are drawn by a researcher who interprets East Central European history from an outsider's perspective. Many good essays and books have been written on Teleki in Hungarian, for example (chief among them Balázs Ablonczy's book published both in Hungarian and English: Ablonczy, B. 2007). Seegel adds much to our knowledge. Particularly valuable is Seegel's analysis of Teleki's romantic phantasies of the American West. Largely based on Karl May's novels (which remained favourite reading for Teleki even as an adult), fantasies of the American West had a profound impact on Teleki's geographical and political thought (pp. 16, 36, 51, 114).

Despite the many strengths of Seegel's study, what is missing from a geographer's standpoint is geography itself. We are well informed about how the bodies and souls of these five geographers influenced their scientific work, but we gain little knowledge with respect to their scholarly thoughts. Readers, therefore, might understandably come to the conclusion that geography was a kind of pseudo-science in the early twentieth century, and that it was nothing more than a useful vehicle for the articulation of political dreams and personal ambitions. I am sure that the geography of the first half of the twentieth century was more than that, and it is perhaps misleading if we judge too harshly one-hundred-year-old scientific concepts using early-twenty-first century norms. Seegel also frequently refers to the map men's 'illiberalism,' which seems anachronistic. Speaking from the Hungarian perspective, nobody used this word in Hungary before 2010. It would therefore be prudent to deploy this notion cautiously in a historical context, unless of course the goal is to speak directly to present political issues. To be fair, this is indeed one of Seegel's goals. Writing of one of Teleki's books, for instance, Seegel argues that: "*On Europe and Hungary* went beyond conservativism, for it imagined an exceptionalist Hungary-led Europe united in an illiberal vision" (p. 149). This interpretation seems more relevant to an analysis of a contemporary Hungarian politician's declarations than it is to Teleki's thinking, and I would find it tenuous to parallel two historically separated politicians in this way. Seegel's narrative, moreover, is very dark, and succeeds in creating the atmosphere of a real *géographie noire* (p. 135). In many ways, he has rendered his drama excessively bleak. In my opinion, these geographers may not have been irreproachable heroes, but neither were they excessively evil malefactors. They were mortal human beings, who found themselves in a very desperate situation, and they did what they thought had to be done.

There are some details concerning Teleki and interpretations of Hungarian history that may have been approached differently by a Hungarian author. Most Hungarian geographers will admit that Teleki's famous "Carte Rouge" was a biased visual representation of statistical data, and that Teleki chose a cartographical method that supported the Hungarian point of view concerning the integrity of Greater Hungary at the Paris (Trianon) Peace Talks in 1920. But only a few would agree with Seegel's opinion that "Teleki asserted population density even ahead of nationality. He followed the linear logic of modernisation, that the density of assimilated Magyars increased as peasants moved to the city, became literate and settled naturally into St. Stephen enclaves" (p. 65). Similarly, a number of Hungarian scholars would question the following: "In the map's surreal and subliminal codes, its message was the defeat of the Little Entente, marginalisation of Romanians and Jews, and omission of rural lands and indigenous peoples, whisked out of history" (p. 69). Seegel's interpretation implies that the presence of the ethnic Hungarians, who lived in the territories ceded from Hungary between 1918 and 1920, was a result of Hungarian assimilation and colonisation, and, further, that Teleki exaggerated their importance by utilising inappropriate and ultimately falsifying cartographical tools. There is, of course, some truth in the claim that the increase of the urban population in the non-Hungarian regions or the settlements in Délvidék (today Vojvodina, Serbia) exemplified these processes. But the truth is that the vast majority of the 3.3 million ethnic Hungarians who lived in the territories that Hungary lost were themselves inhabitants of the "rural lands," and that the Hungarian people were arguably one of the "indigenous peoples" in these regions. Unfortunately, Seegel also provided inaccurate data concerning the territorial losses of Hungary. He writes that "the country lost roughly two-thirds of its population, one-third of its territory" (p. 86), when in fact Hungary lost two-thirds of its territory, roughly 60 per cent of its population, and one-third of its Hungarian speaking population.

Teleki has been a controversial figure in Hungary's political history. As Seegel indicates, his commemoration was a matter of debate even after the turn of millennium (pp. 220–221). Teleki's geographical thought was also complex. Seegel puts Teleki among the "Anglophile German map men" and emphasises the impact of Ratzel's theories on Teleki's geography (pp. 37, 50, 128). At the same time, he writes little about the fact that contemporary French geography also had a significant influence on him, and that it was at least as important (if not even more so) as the German influence on Teleki's thinking. Historians of Hungarian geography usually stress that it was Teleki who broke with the strong German orientation in Hungarian geography, turning instead to the French *géographie humaine* (Krasznai, Z. 2012, pp. 73–74). Teleki and his disciples were wary, for example, of Gyula Prinz's geographical work, and in par-

ticular of "his highly unusual Germanophilia within the Hungarian academia" (Fodor, F. 2006, p. 718). By the same token, nobody at the time regarded Teleki as a Germanophile. The same goes for Teleki's political credo. Despite his apparent anti-Semitism, he was dismissive of Nazi ideology, and as prime minister he endeavoured to shield Hungary from the influence of Nazi Germany. Teleki was aware of the German plans for the future of Eastern Europe. In contrast to Seegel, I do not believe that "Teleki was too blind to grasp that he was being mapped colonially from an imperial Berlin in 1938-39" (p. 182). Teleki's political and geographical legacy will no doubt remain controversial, and it is in light of this that Seegel's research is an important contribution to our knowledge of him. At the same time, however, Seegel's harsh judgment of Teleki and his geography is exaggerated. Can it truly and accurately be said of him that "[t]his hodgepodge geography of the Transylvanian count, not without hackneyed ideas or prejudice draped in science, was surely characteristic of an insecure man who dabbled in studies of the natural world" (p. 131)?

In the final analysis, Seegel's compelling new book should become compulsory reading for those who are interested in the history of twentieth-century geography, and above all, for those who are scholars of East Central European history and geography. Seegel's comparative study on the life and work of five geographers is an outstanding scholarly achievement. Collecting and reading the archival and printed sources was a demanding task in itself, and Seegel has constructed a strong and convincing narrative. "Map Men" is a dark and tragic story devoid of positive heroes and a soothing ending. Although I do not agree with every detail of the book, the book's core argument is sound. "Map Men" is a sad but true story, which in itself is a pity, as it points to uncomfortable truths about our own past.

Róbert Győri[1]

REFERENCES

Ablonczy, B. 2007. *Pál Teleki. The Life of a Controversial Hungarian Politician*. Wayne (NJ), Hungarian Studies Publications.

Fodor, F. 2006. *A magyar földrajztudomány története* (The history of Hungarian geography). Budapest, MTA Földrajztudományi Kutatóintézet.

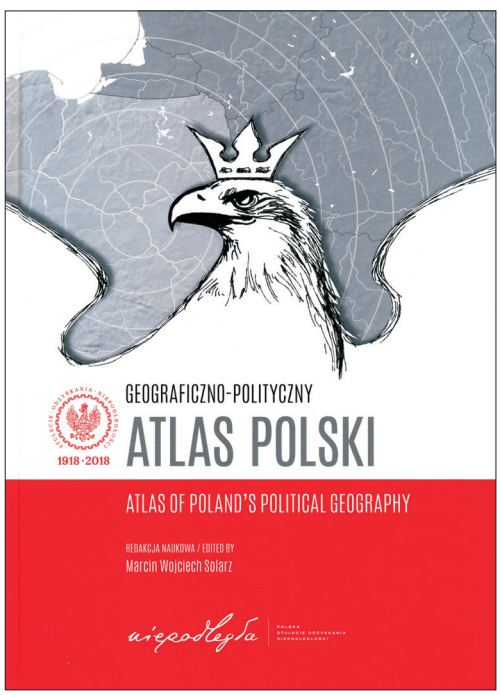
Krasznai, Z. 2012. *Földrajztudomány, oktatás és propaganda (A nemzeti terület reprezentációja a két világháború közti Magyarországon)* (Geography, education and propaganda [The representation of national territory in Hungary between the world wars]). Pécs, IDResearch Kft/Publikon Kiadó.

[1] Department of Social and Economic Geography, Institute of Geography and Earth Sciences, Eötvös Loránd University (ELTE), Budapest, Hungary. E-mail: gyorirobert@caesar.elte.hu. The research has been supported by the National Research, Development and Innovation Office–NKFIH, contract number K 125001.

Solarz, M.W. (ed.): Poland in the modern world: Atlas of Poland's Political Geography. Warsaw, University of Warsaw, Trzecia Strona, and Faculty of Geography and Regional Studies University of Warsaw, 2018. 248 p.

The map is a political tool (Batuman, B. 2010, 222–223) – if we take this now established position as a starting point then we can consequently see an atlas as a (geo)political toolkit. Accordingly, "maps are discursive tools socially produced to persuade others", and as a representational tool the map has been utilised towards maintaining political power and constructing identities (*ibid*, p. 222). Especially with the rise of nation-states, the map has emerged as a powerful sign of national unity and a cultural product materialising nationalist discourse (*ibid*). Thus, the presentation of the national territory in the form of maps within textbooks and atlases serves for the rationalisation and naturalisation of the relationship between the territory and the people, provoking a sense of "territorial bonding" (Herb, G.H. 2004). Put differently, maps and charts fix and legitimise; they "produced, and are produced by arguments of legitimacy" (Reynolds, P.R.A. 2008, p. 72).

That critical perspective barely infiltrated the Atlas of Poland's Political Geography. This otherwise in many respects fine work fulfils a rather different purpose: it is part of this year's commemoration efforts in Poland to celebrate the centenary of the state's regained independence (following its disappearance during 1795–1918). The importance of that event for Poles and others needs perhaps no detailed explanation here, but the background and explicit goal of the publication ought to be borne in mind while reading.



The title page is equipped with the official logo of the centenary, complemented with a line indicating that this work enjoys the "National Patronage of the President of the Republic of Poland Andrzej Duda to mark the Centenary of Regaining Independence" (p. 3). Apart of *Trzecia Strona* (a Warsaw-based publisher) the University of Warsaw and additionally the Faculty of Geography and Regional Studies of that university are indicated as the publishers, which provide the scientific quality – if less a critical approach – that such a volume requires.

The fact that the atlas begins with an introductory letter by Polish President Andrzej Duda (p. 6) further proves its significance as not just a pure academic or popular scientific undertaking: "this interesting scientific publication … responds to enormous needs…". Accordingly, the following quote reflects a view that probably many statesmen have of political geography: "I believe that this … piece of work will become a stimulus for the development of political geography in our country, as this is a science without which no rational plans and forecasts concerning the future of Poland and Europe could be outlined". According to the President, "until the mid-20th century this discipline had been developing freely… It undertook subjects that are of crucial importance for the Republic of Poland and its key interests". Finally, Mr. Duda expresses his ardent hope that the atlas "will become one of the most essential pieces of reading to those who think, write and actively strive for the security and successful development of contemporary Poland" (*ibid*).

The short prologue (p. 7) of editor Marcin Wojciech Solarz – Professor of the University of Warsaw – confirms that the publication was carried out "with a sense of obligation" to two anniversaries: "the first miracle of restored independence" and "the eve of the 30th anniversary of the second". The fact that the regained independence of the Polish state is referred to as a miracle (not just here but also elsewhere in the book) is highly illustrative of the strong symbolic value ascribed to these events. One may add that comparing 1918 and 1989 can appear a little far-fetched (something the author alludes to later, on page 12), since despite the undoubtedly strong Soviet dominance over Poland during the Cold War the latter was still formally an independent state, which put it in a different situation compared to the Baltic Soviet republics, for instance.

Finally, the prologue informs that the European Union (EU) within its 2017 borders serves as the basic reference point for Poland in the international environment, which deserves two short remarks. On the one hand, from the perspective of the EU it can be reassuring that it remains the key reference point for Poland – in line with Polish public opinion, but somewhat in contrast to recent EU-scepticism among parts of the country's political establishment. On a more practical note, the strong focus on the EU – rather than Europe as a whole, or possibly some other space – has resulted in many maps on which countries like Norway, Switzerland, Serbia, Ukraine, etc. are missing. Yet Poland has intensive links with these countries as well (e.g. migration), and Eurostat has many data on not just EU Member States but also the candidates, as well as countries of the European Economic Area. Moreover, the atlas has not just worked from EU- or European databanks but others as well, such as the Human Development Index.

The big chunk of the atlas can be said to be divided into five parts: an introductory text on the political geography of Poland and four chapters containing maps 'only'. I find the former easier – and hopefully more constructive – to comment. The chapter is titled "Poland – politics and space" (pp. 11–30) and is divided into five (unnumbered) sub-chapters: an (untitled) introduction; "The State"; "Location"; "Geopolitics"; and "Borders, territory, sovereignty". The chapter starts with two quotes, the first of which by Eugeniusz Romer, the founder of Polish political geography, whose 1916 "Atlas of Poland" was the foremost cartographical work on the Polish territories cited by the Western Allies at the Paris Peace Conference (Labbé, M. 2018, p. 94). The second quote comes from a 1982 speech by Ronald Reagan in the British Parliament and ends with the sentences: "Poland is not East or West. Poland is at the centre of European civilization. It has contributed mightily to that civilization" (p. 11).

The introduction starts with several lengthy though illustrative quotes by political emigrant and publicist Karol Zbyszewski, the first of which is saying that the "boundaries between fields are more permanent than the borders of Poland" – cited as expressing "a profound truth" (p. 11). At least

> [t]he Vistula has always been within its [Poland's] borders… The Vistula finds its faithful reflection in the character of the Polish people. Flamboyance without adequate means is seen in the nonsensical breadth of the Vistula which suffers a chronic lack of water; fickleness, inconsistency, hysteria – it switches channel constantly, lunging first here then there, one day ominous, and the next listless; these magnificent outbursts and ignominious downfalls – the Vistula rises, rushes with power, and a week later scarcely murmurs, languishes at every step on the shallows; this capriciousness, this continual, irritating pose of greatness, this lack of stability, this charm, this melancholy, this unpredictability … these are what characterize the Poles and the Vistula. (pp. 11–12)

Whereas such geographical narratives and national self-images are fascinating to read, the academically or critically inclined reader may miss a certain distance to them by the author. That also goes for some of his own statements intended to emphasise the (relative) greatness of Poland: "Poland has a significantly larger population than other countries in the region (with the exception of Russia, Germany and Ukraine)"; or, "[w]ith the exception of its periods of eclipse Poland has consistently been a force to be reckoned with…" (p. 13).

A little later it is stated that "[t]erritorial and national lack of cohesion were key problems for Poland between 1918 and 1939. These were resolved after 1945 as a result of the redrawing of the borders of the Polish state, the accompanying population resettlement, and the assimilation policy pursued by the communist authorities" (pp. 13–14). I find this formulation problematic for more than one reason. One, if the verb 'resolve' is correct to use here at all, I would have at least put it in inverted commas. More crucially, the "national lack of cohesion" was "resolved" not just after but also during World War II… And following the war, Poland was one of the few countries where Jews were still harassed and, in some cases, even killed (*cf*. Gross, J.T. 2006). Finally, one may mention that the assimilation policy was described in the quote as having been pursued by "communist authorities" rather than Polish ones (both of which are correct, but the choice of terms is telling). Regarding the territorial lack of cohesion Solarz notes that

> [w]ith the collapse of the Soviet Union and the increasing dependence of Belarus on Russia, the "Suwałki isthmus" which separates the Kaliningrad region from Belarus has *de facto* become a new „Polish Corridor" akin to the Pomeranian corridor in its past forms (1657–1772 and 1918–1939). The last of these was a source of conflicts, and, in 1939, it was one of the reasons for Germany's aggression against Poland and consequently the outbreak of the Second World War. (p. 14, original emphasis)

It is true that the "Polish corridor" was a casus belli for Germany in 1939, although – as the author alludes to – it is hardly realistic to have been a key reason for attacking Poland. What is interesting (from a social scientific perspective) is not primarily whether the quoted fears and historical parallel-drawings are realistic or exaggerated, but the fact that they apparently continue to shape Polish geopolitical thinking. Illustratively, to Solarz "[i]t seems that contemporary Poland faces new challenges because after the reunification of Germany (1990) and the coming to power of Vladimir Putin in Russia (1999), we are observing the renewed formation of the two poles of political power in the direct vicinity of Poland" (p. 20).

Elsewhere, it is stated that Poland "is now classified as a highly developed country" (p. 15). While there is

no reference indicated here, it is quite possible that Poland is nowadays ranked in this group of countries according to some – even established – indices such as the Human Development Index. In any case, a fading belief in the narrative of convergence between Europe's East and West, which characterised collective hopes up until about the 2008 economic crisis, means that few perceive Central and East European countries to be "highly developed". It is interesting that the editor himself recently published critically on the United Nations' designation of 'least developed countries' (Solarz, M.W. and Wojtaszczyk, M. 2017) and yet adopts the same vocabulary for his own country uncritically.

The author is taking a political stand (which is by no means *per se* illegitimate) in saying that "in the late 1980s and early 1990s there was presumably no alternative route to effective systemic transformation by which the high social and economic costs could have been avoided" (p. 15). There are of course alternative approaches to this question (*cf*. Buchowski, M. 2001), and the socio-economic costs could likely have at least been mitigated. Solarz too is trying to nuance the overall picture by providing some critical remarks that I find praiseworthy:

> … it cannot be doubted that Polish success came at a price. Among the ills that Poland has experienced are mass emigration, a demographic catastrophe, social injustice and deindustrialization… the prioritization of special interests over those of the community… We present Poland as a model for democratic transformations… but shouldn't low voter turnout and deep political division in society rather prompt a very critical assessment of the quality of Polish political life, of the Polish political elites and ourselves as citizens? These and other ruptures require speedy and careful remedy, to be carried out first and foremost by the elites of free Poland. (p. 15)

The sub-chapter "Location" makes clear that Poland's situatedness "on a flat and open plain without any natural barriers" has been seen as its most important geopolitical characteristic, translating into numerous threats and challenges but also some opportunities (p. 16): "[f]rom the moment of its birth, Poland has been a borderland country squeezed between the querns of great worlds which here intersect and collide" (p. 17). I am missing a reference to Oskar Halecki (1980) here, but we get acquainted with other historical Polish thinkers such as Piotr Grabowski and Wacław Nałkowski.

According to Solarz, "Poland's place on the map of Europe can be described using the comparison of an hourglass. Poland occupies the narrow tube connecting two large glass spheres which, on opposite sides, contain Western and Eastern Europe, Europe and Asia…" (p. 17). One may criticise this metaphor for assuming an image of Europe without Hungary, Southeast Europe, and Scandinavia. Numerous are the cities, regions, and countries in Central and Eastern Europe (and beyond) that claim to be *the* intersection of West and East…

I personally find the sub-chapter "Geopolitics" the most interesting, and perhaps also less emotionally loaded than the other sub-chapters. This section introduces the reader to three key Polish geopolitical meta-concepts historically developed. The first of these may by now seem familiar: the narrative of 'Poland as a transitional land' emphasises the "indeterminate character" of Polish territories – i.e. lacking physical geographic borders – that poses a permanent threat (p. 20).

The second concept envisions "Poland as a bridge country" between the Baltic and Black Seas, an area which Nałkowski saw as "a separate and distinct geographic whole ascribed to Poland" (p. 21). Solarz writes that this narrative "encourages the Polish state to develop activity in the bridge region and seek the role of leader of the smaller countries located between Germany and Russia" (*ibid*). I find it strange that neither Piłsudski's interwar concept of *Intermarium* nor the much more recent Three Seas Initiative is mentioned here, both of which clearly followed the logic of 'Poland as a bridge' – their meagre results notwithstanding. In any case, "Poland's actions in support of Ukraine in 2004–2005 and after 2013" (p. 22) are mentioned.

The third narrative characterises "Poland as a bulwark of Christendom, the West, Europe" – as "a shield which protects a certain community of states, variously defined in different periods, but which in general can be described as Western Europe" (*ibid*). Importantly, "[t]he Polish bulwark concept is currently being manifested in Poland's fulfilment of the obligations arising from its location on the eastern borders of the European Union and NATO" (p. 23).

The remainder of the book is as mentioned a collection of maps, divided into four chapters by the following titles: "International relations"; "The state"; "Society"; and "Development". Apart of the minor criticism I made above regarding the EU-centric maps, there is little constructive input I can provide (I am sure the atlas will also be reviewed by a cartographer). The maps appear carefully done, in high quality, providing excellent visuality. It is progressive that some maps were included on gender and socio-economic inequalities (pp. 140–144), even if all these maps compare disparities between EU-countries i.e. none within Poland. As I made clear in the beginning of the review, I tend to miss at least some critical remarks on the exercise of mapping in general, but that would have been in place more in the text part rather than on the maps.

Despite the critical remarks formulated here, I overall find the Atlas of Poland's Political Geography a great achievement. I believe it could function very well as an introduction to Polish geography, history, or to Poland more generally; in academic courses or for a broader audience; thanks to its bilingual character, in Polish- and English-language environments

alike. While this volume is understandably a Polish project, at least some regional parallels could have been drawn: a number of countries in Central and Eastern Europe are celebrating their centenaries this year. Moreover, geopolitical narratives such as the 'Christian bulwark' are by no means unique to Poland, but also exist in other countries of the region and beyond (Tazbir, J. 2005). It was less surprising to learn that a fear of Russia is vivid in Polish geopolitical thinking, but more so that German reunification in 1990 can still be referred to as a challenge, despite strongly improved Polish-German relations ever since (Balogh, P. 2014, 25–27). Indeed, one is unlikely to see the Bundeswehr marching across the Polish border for other reasons than shared NATO-exercises within the foreseeable future. But continued presence of various fears and an enhanced need for security, stability, and peace are some of the reasons why political geography should continue to be studied. Thus, a toast is in place to the next one hundred years of Polish geography: *Sto lat!*

Péter Balogh[1]

REFERENCES

Balogh, P. 2014. *Perpetual borders: German-Polish cross-border contacts in the Szczecin area*. PhD dissertation. Stockholm, Department of Human Geography, Stockholm University.

Batuman, B. 2010. The shape of the nation: visual production of nationalism through maps in Turkey. *Political Geography* 29. (4): 220–234.

Buchowski, M. 2001. *Rethinking transformation: an anthropological perspective on post-socialism*. Poznań, Wydawnictwo Fundacji Humaniora.

Gross, J.T. 2006. *Fear: anti-Semitism in Poland after Auschwitz*. New York and Princeton, Random House and Princeton University Press.

Halecki, O. 1980. *Borderlands of Western civilization: A history of East Central Europe.* 2nd edition. Safety Harbor, Simon Publications.

Herb, G.H. 2004. Double vision: territorial strategies in the construction of national identities in Germany, 1949–1979. *Annals of the Association of American Geographers* 94. (1): 140–164.

Labbé, M. 2018. Eugene Romer's 1916 Atlas of Poland: creating a new nation state. *Imago Mundi* 70. (1): 94–113.

Reynolds, P.R.A. 2008. *Transmission and recall: the use of short wall anchors in the wide world*. PhD dissertation. York, Department of Archaeology, University of York.

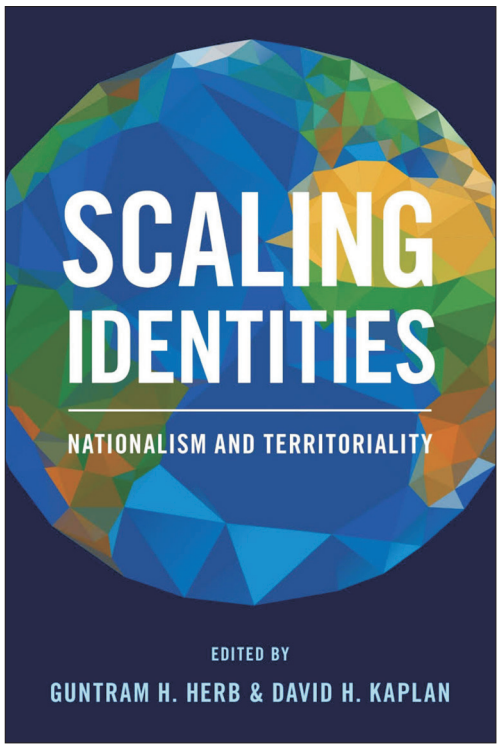
Solarz, M.W. and Wojtaszczyk, M. 2017. Are the LDCs really the world's least developed countries? *Third World Quarterly* 38. (4): 805–821.

Tazbir, J. 2005. The bulwark myth. *Acta Poloniae Historica* 91. 73–97.

[1] Central and North Hungarian Research Department, Institute for Regional Studies, CERS, Hungarian Academy of Sciences (MTA), Budapest, Hungary. E-mail: Balogh.peter@krtk.mta.hu. Research for this publication has been supported from the National Research, Development and Innovation Fund (NKFIA) grant nr. 124543 (program: PD_17), and by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences.

Herb, G.H. and Kaplan, D.H. (eds.): Scaling Identities: Nationalism and Territoriality. Lanham–Boulder–New York–London, Rowman & Littlefield, 2018. 294 p.

Both scale and identity tend to be increasingly popular, albeit contested concepts in the realm of human geography. Discussions around scale have come a long way from the 1980s. From early on, scale seemed to be an appropriate tool to interpret an increasingly complex world, and has become widely used in a myriad of investigations. However, increased scholarly attention cast doubt on the explanatory power of scalar divisions first, then called into question their ontological status. It was not until the first decade of the new millennium that, in a programmatic call, the elimination of scale was proposed (Marston, S.A. *et al.* 2005). While such concepts like actor-network theory (Latour, B. 2005) and flat ontology (Jones, J.P. III. *et al.* 2007) seem to fill the void left in the downsizing of scale, one may find that scales are not neglected at all, but a re-conceptualisation has been going on. Likewise, the notion of identity has been debated for quite a long time as scholars attached different associations to it. Prior to the post-structuralist turn (see among others Deleuze, G. and Guattari, F. 1988), identity was deemed socially and institutionally embedded, so the dominant view conceptualised it as a collective or institutional mechanism of control and domination (Malešević, S. 2003). However, as the postmodern condition favoured a relational approach (Murdoch, J. 2006), identities have come to be viewed as multiple, fluid and flexible. Despite the notion's conceptual ambiguity, as we are currently witnessing that identity politics, being articulated within the wider context of culture wars, appear to be gaining momentum again, identity inevitably remains a trivial matter in social sciences.



Such rapid changes in scholarly literature in relation to post-structural perspectives demand prudent attitude towards scale and identity, something the editors of „Scaling Identities", Guntram H. Herb and David H. Kaplan, are completely aware of. The editors are well-known experts in the field, their long-term commitment has already been proven since this volume's predecessor „Nested Identities" (Herb, G.H. and Kaplan, D.H. 1999). Right in the introduction, the editors insist on the ever-changing nature of scales, and they acknowledge that in the previous volume a conceptualisation in a dynamic manner, regrettably, was missing. This is indeed a sympathetic attitude which deserves respect. In the introduction the editors explicitly express that a move beyond Eurocentric national identity narratives is desired, and we could not be more delighted with this intention as nine out of twelve case studies take place outside of Europe.

The book, on its surface, explores topics around the very notions of scale and identity. Nevertheless, it becomes clear that it is national identity and its spatial articulation what is aimed to be discussed. This does not mean automatically, however, that identities coalesce around the national scales, because identities are in a constant flux. This volume contains five thematic sections, consisting of a total of 14 chapters.

The first two chapters, marked by the editors themselves, offer profound theoretical insights into national identities, emphasising their interrelatedness with place and scale. Herb is primarily concerned with issues revolving around identity and territory. As he traces their relationship throughout multiple examples, we may feel that a tangible and quantifiable subject of power can only have a troublesome interaction with its elusive counterpart. Speaking more clearly, an ethnic group's nation-building strategy attempts to exert control over an utmost piece of territory, but so do others, and if these territories overlap, conflicts arise. Herb concludes that these will continuously occur as far as 'us-them' distinction will remain a central characteristic of national identity. Thereafter,

Kaplan examines why a sense of belonging is so often related to nations. However, the national scale does not exclude either lower or higher scales, but a nested coexistence is given. Moreover, in the wake of globalisation, flows in the form of transnationalism, diasporas and hybridity transcend scales to an increasingly larger extent.

Part Two pays tribute to the multi-faceted social processes through which identities are being consolidated. This part is opened by Corey Johnson's intriguing study about German nationalism. For the collective memory, relation between Germany and nationalism is practically identical to the Third Reich's short-lived but destructive cultural policies. Johnson deliberately steps out of this historical period, drawing attention to German discourses of nation and identity prior to World War II. German nationalism, as Johnson points it out, has not come from nowhere. In the aftermath of the successful unification in 1871, a common political-territorial umbrella was given for German people, but „allegiances and identity politics in the new empire were highly parochial and cleavages along regional, class-based, urban-rural, as well as linguistic and religious lines were pervasive" (p. 55). So German-ness needed to be discovered through a rescaling of local affairs and through internal othering (e.g. orientalised Catholics). The latter has been repeatedly brought about after the reunification in 1990, treating the economically backward regions as a risk to well-being of the nation. The author uses the case of the Monument to the Battle of the Nations located in Leipzig to reflect on scalar relationships between the local and the abstract (supra) national. Being one of the largest battle monuments, it was originally intended to represent German pride, but, particularly during the Communist era, became the symbol of friendship between German and Russian people as they have been fighting together against the Napoleonic regime. Nowadays, festivities around it tend to celebrate European peace, marking a new shift in the scaling of German identity.

In the next chapter, David Keeling addresses the tensions between territorial and sociocultural aspects of identity, epitomised by the modern identity of Argentina (*Argentinidad*). Given the vast territories, loose connections between inhabitants – including both indigenous people and those coming from Europe – hindered the solidification of a modern state. However economic progress might have taken place from the 1880s, the challenge was to „define an acceptable state-based territorial, rather than local, identity" (p. 74) What brought on a constant headache for national elites has been reconciling a cosmopolitan, Europeanised cultural identity with an authentic one, imbued with ethnic and colonial heritage. The latter signifies Latin American commonalities too, binding together various scales from the local to supranational.

In chapter 5, Kefa M. Otiso illuminates why Tanzania is being perceived as having the most developed national identity in Africa (on the basis of an Afro-barometer survey). This fact could cause more surprise if we take into account that an ethnically diverse country is given, where nation-building projects took quite different trajectories between Zanzibar and the mainland. Otiso argues that under external threat identities operating on a micro-scale (e.g. families and households) are being forced to create meso-scale units (e.g. clans and ethnic communities). In that vein, taking advantage of consecutive colonial rules' subjection, the Tanzanian state has been proclaimed on favourable grounds. The next key element lies in Julius Nyerere's presidency, under whom almost every aspect of the society was nationalised and de-ethnicised. However, the so-called Utanzania's (Tanzian-ness) bright idea nowadays faces major challenges as the transition from socialism to capitalism – paralleled by the resignation of Nyerere – surfaced such issues as the growing tension between Muslims and Christians.

Part Three sets forth supranational identities. As first within this section, Alexander B. Murphy explores a meaningful European identity. As we may all assume, the very idea of a European identity has yet to be espoused by its population. That is to say, Europeans do not seem themselves as clearly Europeans, but they rather attach to the identities operating at national or regional scale. Nonetheless, it would be misleading to compare European identity with other identities, because „European identity coexists with other identities – state, regional, ethnic and local – in a way that is not strictly hierarchical", hence it is no more than a „cultural-territorial construct to which meaning is attached" (p. 112). Diffusion of such a meaningful European identity requires a smoothly functioning and cohesive European Union first. But growing EU-scepticism, Brexit, enduring impacts of the economic crisis, and falling turnouts at European elections, to name but a few, do not confirm that. On the other hand, in Murphy's opinion, the European Union was able to establish a discursive space, in which a European identity can unfold.

In the next chapter, Gary S. Elbow investigates whether a regional Caribbean identity does exist. He vividly deconstructs the homogenous image of the Caribbean by distinguishing several scales and subregions, where identities are being expressed. We may find how the colonial history and associated languages shape the elusive notion of Caribbean identity. Unsurprisingly, as Elbow argues, geographical location has little to do with this scale of identity. Then in the search of a general sense of Caribbean-ness, he turns to material and performative aspects of identity. For example, carnivals show that „the specific elements vary from place to place, but their general characteristics make up an Afro-European tradition expressed at the macroscale" (p. 129). Another case worth to underline are the writings of Gabriel García

Márquez that both retain his Colombian coastal origins and grasp a macro-level Caribbean identity.

What does the notion Arab Homeland cover, and what role does Palestine play in it? Karen Culcasi's chapter revolves around such crucial questions in a macro-region characterised by never-ending struggles over territory and identities. Neither the Arab Homeland nor Palestine are internationally recognised or clearly defined, so he resorts to official maps. The included maps reveal altogether that no singular delineation of the Arab Homeland exists, though Culcasi assesses that the prevalent use of Arab language and Arab League membership decide commonly whether a certain country belongs to it. This empirical study is completed by another compelling inquiry as Culcasi does not settle for official cartographic texts. Drawing on interviews with Palestinians living abroad, he asserts that for most of the respondents, Arab unity is just a desired myth. Furthermore, even if they are rhetorically linked to the Arab Homeland, they retain a strong sense of localism taking shape in the Palestine identity. This relationship becomes even more strained when we took into consideration the different mappings of Palestine, which invoke hybrid scalar conceptions as the ordering of territories into scales „can be simultaneously conformed to and defied" (p. 151). Drawing on his observations, Culcasi stands firm against hierarchical and universalist approaches, because the concerned region is made up by „multiscaled, transnational and hybrid territorial entities" (p. 152).

The fourth part, entitled "Connecting Identities", begins with Steven E. Silvern's lucidly written study about Native American identities. This chapter asserts that indigenous people, after quite a few centuries of suppression, have once reclaimed their agency in determining their future within a settler colonial state. This unprecedented cultural resurgence has been unfolding simultaneously at three scales. In this vein the local scale has been the subject to re-imagine place-based identity through reconnecting indigenous people with their physical environment. Regional scale sets the stage for another opportunity of expanding indigenous identity as the formerly imposed borders of the reservations are getting permeable through treaty rights. On a national level, Silvern presents in a subtle way how identities are being scaled up by the spread of information and communication technologies. Notwithstanding these positive multi-scalar dynamics, the author shares his concern about colonialist imaginaries' prevailment, with a wink to the prioritisation of neoliberal economic interests.

At this point of the book, we may have collected some sort of knowledge about the often-manifested tensions between minority regionalism and majority nationalism. In Chapter 10, Takashi Yamazaki adds more flavour to this issue by turning our attention to Okinawa Prefecture, where a wide array of influences has been shaping local identity. A cursory investigation would suggest that Okinawans (though characterised by multiple socio-spatial cleavages) were 'Japanised', but with the islands' several re-borderings and vast distances from Japanese centres, Okinawan identity has far more layers than a centralised, nationally imposed one. Therefore, Yamazaki distinguishes four different, albeit intertwined scales of identity, each of which having a collective category (islander, Okinawan, Ryukyuan, and Japanese, respectively). He also sheds light on the re-instutionalisation of national identity through the education system, not solely on ideological grounds but in the threatening shadow of Chinese geopolitical endeavours of socioeconomic interests as well.

In the next chapter, Susan M. Walcott scrutinises the almost inscrutable idea of Chinese national identity. Sketching an overview of China's history, she attempts to grasp what the multi-faceted Chinese characteristics would entail. Thereafter, Walcott calls into question this essentialist position by numerous arguments. First, periphery identities, especially in Tibet and Xinjiang, are hard to fit into historical meta-narrative, through which a construction of congruent Chinese identity is sought. Second, the globally scattered Chinese diaspora also poses a constant challenge to Chinese-ness via the emergence of hybrid identities. Walcott differentiates the ostensible mass of Chinese migrants, inter alia, by the historical period when migration took place, and the place of origin. Additionally, in the midst of interstate turmoil, both Hong Kong and Taiwan has seen a reification of distinct identities.

Kurdistan set the tone for fragmented identities in the fifth part of the volume. Carl T. Dahlman and Sanan Moradi trace the evolution of the deeply flawed Kurdish history, seeking to answer how the Kurds have become the largest (approximately 33 million) nation without a state. The authors illuminate that in the medieval ages Kurdistan has been stuck between the Turkish Ottoman Empire and the Persian Empire, a situation getting even more complicated with the appearance of British and Russian imperial interests. Thus, the region has come to be seen as a frontier, with no singular Kurdish geopolitical orientation. The expression 'Kurds have no friends but the mountains' indicate not just the harsh consequences of *Realpolitik* but intra-Kurdish rivalries as well. More contemporarily, we are still witnessing a fractured Kurdish nationalism: „while oil, gas, and Islamist militants might flow easily across the partition, Kurdish unity does not" (p. 236). In concluding, authors pour fantasies around Kurdish unity into a pessimistic mould because territorialised self-interests and contesting geopolitical agendas will probably keep Kurdish people divided.

In Chapter 13, George W. White explores the contested identities in Transylvania, a well-known topic among Hungarian geographers. In theorising the discussion, he moves away from the popular view of ethnic differences, proposing instead a more fluid conceptualisation that does not see the multiethnic

character doomed to be the subject of conflicts. In that vein he argues for a Transylvanian identity based on a blend of different cultures, hoping that „if shared history is a component of shared identity, then a shared Transylvanian identity exists at some level" (p. 246). This bright prospect has been over-shadowed by mutually exclusive nationalist discourses though, as Transylvania is considered a special constituent of both Romania's and Hungary's national consciousness.

Fragmentation remains the core issue in Pablo Bose's chapter, drawing the attention to Indian diasporas and their linkages to national identity. It is no wonder that one could find it difficult to illustrate a monolithic Indian national identity among diasporas, exemplified in this chapter by the great variations between those from Punjab and Kerala. Yet another scene where scale and place matters. Bose explores the so-called Khalistan movement (a Sikh-centric political project), the goal of which is to establish a sovereign country, with territories encompassing not only Punjab but the broader region, now being part of Pakistan. What makes this case particularly interesting is the process through which a regional issue is being pushed to a broader scale by the international Sikh community's militancy. Kerala represents an adequate counter case. Being a remittance-based economy (in fact, one out of every four Indians working in countries of the Persian Gulf originate from there), Bose addresses another spatial relation, where 'outsiders' are involved directly in internal affairs.

Juoni Häkli is in charge of the afterword, which is entitled "Transcending scale". Based on his concluding words, despite "none is willing to go as far as to simply reject the concept of scale tout court (…), scale, it seems, has been somewhat scaled down in terms of theoretical ambition and vigor" (pp. 274–275). As Häkli duly recognises, scale could retain its useful role for ordering analytical observation, but one need to be aware of future post-structural methodologies.

The value of this book is beyond doubt. In an inextricably complex world, where nationalisms tend to be a cornerstone in geopolitics, fluid notions of territory, nation and identity need to be addressed. Where the volume most successful is in weaving together wholly different place contexts around clear-cut lens of investigation. Accordingly, every chapter places scale and identity at its centre in a thought-provoking way. The structure of the book is well-balanced, and the easy-to-read case studies have the potential (and even the danger) to sustain the interest for quite a long time. This comprehensive volume will surely serve as a point of departure for anyone who is interested in geographies of identity.

Tamás Illés[1]

[1] Doctoral School of Earth Sciences, Eötvös Loránd University (ELTE), Budapest, Hungary. E-mail: tamas.illes92@gmail.com. The research has been supported by the National Research Development and Innovation Office (NKFI), contract no. K 124291.

REFERENCES

Deleuze, G. and Guattari, F. 1988. *A Thousand Plateaus: Capitalism and Schizophrenia*. London, Athlone Press.

Herb, G.H. and Kaplan, D.H. (eds.) 1999. *Nested Identities: Nationalism, Territory, and Scale*. Lanham, Rowman & Littlefield Publishers.

Jones, J.P. III, Woodward, K. and Marston, S.A. 2007. Situating flatness. *Transactions of the Institute of British Geographers* 32. (2): 264–276.

Latour, B. 2005. *Reassembling the Social: An Introduction to Actor-Network Theory*. Oxford, Oxford University Press.

Malešević, S. 2003. Researching social and ethnic identity: A sceptical view. *Journal of Language and Politics* 2. (2): 265–287.

Marston, S.A., Jones, J.P. III and Woodward, K. 2005: Human geography without scale. *Transactions of the Institute of British Geographers* 30. (4): 416–432.

Murdoch, J. 2006. *Post-Structuralist Geography*. London–Thousand Oaks–New Delhi, SAGE.

GUIDELINES FOR AUTHORS

Hungarian Geographical Bulletin (formerly Földrajzi Értesítő) is a double-blind peer-reviewed English-language quarterly journal publishing open access **original scientific works** in the field of physical and human geography, methodology and analyses in geography, GIS, environmental assessment, regional studies, geographical research in Hungary and Central Europe. In the regular and special issues also discussion papers, chronicles and book reviews can be published.

Manuscript requirements

We accept most word processing formats, but MSWord files are preferred. Submissions should be single spaced and use 12pt font, and any track changes must be removed. The paper completed with abstract, keywords, text, figures, tables and references should not exceed **6,000 words**.

The Cover Page of the article should only include the following information: title; author names; a footnote with the affiliations, postal and e-mail addresses of the authors in the correct order; a list of 4 to 8 keywords; any acknowledgements.

An abstract of up to **300 words** must be included in the submitted manuscript. It should state briefly and clearly the purpose and setting of the research, methodological backgrounds, the principal findings and major conclusions.

Figures and tables

Submit each illustration as a separate file. Figures and tables should be referred in the text. Numbering of figures and tables should be consecutively in accordance with their appearance in the text. Lettering and sizing of original artwork should be uniform. Convert the images to TIF or JPEG with an appropriate resolution: for colour or grayscale photographs or vector drawings (min. 300 dpi); bitmapped line drawings (min.1000 dpi); combinations bitmapped line/photographs (min. 500 dpi). Please do not supply files that are optimised for screen use (e.g., GIF, BMP, PICT, WPG). Size the illustrations close to the desired dimensions of the printed version. Be sparing in the use of tables and ensure that the data presented in tables do not duplicate results described elsewhere in the article.

REFERENCES

Please ensure that every reference cited in the text is also present in the reference list (and vice versa).

Reference style

Text: In the text refer to the author's name (small capitals with initials) and year of publication. References should be arranged first chronologically and then further sorted alphabetically if necessary. More than one reference from the same author(s) in the same year must be identified by the letters 'a', 'b', placed after the year of publication.
Examples: (Ridgewell, A.J. 2002; Maher, B.A. *et al.* 2010) or Ridgewell, A.J. (2002); Maher, B.A. *et al.* (2010).

Journal papers:
Aagaard, T., Orford, J. and Murray, A.S. 2007. Environmental controls on coastal dune formation; Skallingen Spit, Denmark. *Geomorphology* 83. (1): 29–47.

Books:
Pye, K. 1987. *Aeolian Dust and Dust Deposits.* London, Academic Press.

Book chapters:
Kovács, J. and Varga, Gy. 2013. Loess. In *Encyclopedia of Natural Hazards.* Ed.: Bobrowsky, P., Frankfurt, Springer, 637–638.

Book reviews

Book reviews should be between 2,000 and 3,000 words (including references).

Submission

Submission to this journal occurs online. Please submit your article via http://ojs3.mtak.hu/index.php/hungeobull

All correspondence, including notification of the Editor's decision and requests for revision, takes place via the electronic submission system.