ACTA UNIVERSITATIS CAROLINAE

AUC **GEOGRAPHICA**

58
2/2023

CHARLES UNIVERSITY • KAROLINUM PRESS

AUC Geographica is licensed under a Creative Commons Attribution License
(http://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution,
and reproduction in any medium, provided the original author and source are credited.

© Charles University, 2023
ISSN 0300-5402 (Print)
ISSN 2336-1980 (Online)

Effects of cancer mortality on life expectancy in European high-income countries between 1950 and 2019

Vitalie Stirba*

Charles University, Faculty of Science, Department of Demography and Geodemography, Czechia
* Corresponding author: vitalie.stirba@natur.cuni.cz

ABSTRACT
This article aims to analyze the effects of cancer mortality on life expectancy at birth in 15 European high-income countries between 1950 and 2019. To establish the 1950–2019 time series of deaths from cancer, mortality data were harmonized from the available datasets of the World Health Organization Mortality database, coded according to the International Classification of Diseases of the 7th, 8th, 9th, and 10th editions. The estimation of the cancer mortality effect on the life expectancy at birth was performed using the algorithm of stepwise replacement for the life expectancy decomposition. The increase in cancer mortality contributed to a decline in overall life expectancy growth until the mid-1990s, coinciding with the aging cohorts of heavy smokers and a long-term reduction in mortality from other non-communicable diseases. Subsequently, since the 1990s, the reduction in cancer mortality has contributed to a significant increase in life expectancy at birth, especially in males. Reduction in cancer mortality was the outcome of various factors, such as alcohol and tobacco control policies, advances in cancer prevention and its treatment, general increase in population well-being, and reduction in risk-factors.

KEYWORDS
cancer mortality; life expectancy; trends; Europe

Received: 18 April 2023
Accepted: 21 June 2023
Published online: 20 July 2023

Stirba, V. (2023): Effects of cancer mortality on life expectancy in European high-income countries between 1950 and 2019. AUC Geographica 58(2), 149–156
https://doi.org/10.14712/23361980.2023.11
© 2023 The Author. This is an open-access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0).

1. Introduction

During the 20th century, life expectancy (LE) sharply increased in most countries, while in those developed, it even doubled. This is due to changes in the distribution of cause-specific mortality during this period when a decline in infectious diseases was observed, continued by a cardiovascular revolution (Vallin and Meslé 2004) and a cancer mortality diminution. However, back then, despite the notable increase in LE, cancer mortality sharply increased, especially in males, and decreased in several countries only in the 1990s (DeVita and Rosenberg 2012; Bertuccio et al. 2019). Recently, due to noteworthy changes in mortality structure in developed countries, a shift occurred from cardiovascular diseases to cancer as the leading cause of death (Stringhini and Guessous 2018; Townsend et al. 2016).

Similar mortality trends in most frequent cancer sites were observed in many developed European countries (Hashim et al. 2016), where breast (Carioli et al. 2017; Malvezzi et al. 2019), lung (Islami et al. 2015; Remon et al. 2020), prostate (Baade et al. 2009; Wong et al., 2016), stomach (Balakrishnan et al. 2017), colon and rectal cancer (Araghi et al. 2018; Ait Ouakrim et al. 2015) declined in the last decades. This became possible in the context of reduction in exposure to cancer risk factors, diagnostics and medical care improvements, and enhancement of cancer treatment effectiveness (DeVita and Rosenberg 2012; Proctor 2013). However, despite the visible decline in cancer mortality during the last decades, the impact of the 2020–2021 COVID-19 pandemic may slow down future trends (Englum et al. 2021).

This study aims to describe the long-term trends in LE changes due to cancer mortality in selected European high-income countries and to measure the impact of cancer deaths on these trends. For this, the algorithm of stepwise replacement of the LE decomposition was used. Although cancer mortality trends have already been described previously in the literature, this study quantifies the contribution of cancer mortality trends to life expectancy using data from a large set of European countries.

2. Methods

This study is based on long-term datasets on cancer mortality in selected developed European countries that contain information on the distribution of deaths by sex, age, and cause of death between the 1950s and 2019. The following list of countries with periods of data availability were selected for the research: Belgium (1954–2018), Czechia (1950–2019), Denmark (1952–2018), Finland (1952–2018), France (1950–2017), Hungary (1955–2019), Ireland (1950–2018), Italy (1951–2017), Netherlands (1950–2019), Norway (1951–2016), Poland (1959–2019), Portugal (1955–2018), Spain (1951–2019), Sweden (1951–2018), and Switzerland (1951–2018).

Cancer deaths by sex and age were retrieved from WHO (World Health Organization) Mortality Database (WHO Mortality Database) and included cancer sites that, according to the International Classification of Diseases (ICD), corresponded to the codes 140–205 (ICD7), 140–209 (ICD8), 140–208 (ICD9), and C00–C97 (ICD10). The Human Mortality Database was used as a source for the population exposure data (HMD). To organize a comparable long series of data sets on the cause of death age distribution, ICD7, ICD8, ICD9, and ICD10 classifiers have

Tab. 1 Years of transition between the ICD7, ICD8, ICD9, and ICD10 classifications in analyzed countries.

Country	ICD7	ICD8	ICD9	ICD10
Belgium	until 1967	1968–1978	1979–1997	since 1998
Czechia	until 1967	1968–1978	1979–1993	since 1994
Denmark	until 1968	1969–1993	–	since 1994
Finland	until 1968	1969–1986	1987–1995	since 1996
France	until 1967	1968–1978	1979–1999	since 2000
Hungary	until 1968	1969–1978	1979–1995	since 1996
Ireland	until 1967	1968–1978	1979–2006	since 2007
Italy	until 1967	1968–1978	1979–2002	since 2003
Netherlands	until 1968	1969–1978	1979–1995	since 1996
Norway	until 1968	1969–1985	1986–1995	since 1996
Poland	until 1968	1969–1979	1980–1996	since 1997
Portugal	until 1970	1971–1979	1980–2001	since 2002
Spain	until 1967	1968–1979	1980–1998	since 1999
Sweden	until 1968	1969–1986	1987–1996	since 1997
Switzerland	until 1968	1969–1994	–	since 1995

Source: WHO Mortality Database

been harmonized. Tab. 1 shows the years of transition between classifications in the countries studied.

The age-standardized mortality rate was calculated using the direct standardization method, where the New European Standard Population (Eurostat 2013) with a 5-year age group and the last open-ended age interval 85+ was applied as a population standard.

LE calculation was based on an abridged life table construction with the last open-ended age interval of 85+. Therefore, LE decomposition was performed using the algorithm of stepwise replacement (Andreev et al. 2002), where the age components of LE were compared between time periods and measured the contribution of cancer mortality to the LE change. These calculations were made in two steps:

1. LE decomposition was performed based on a 5-year age group interval by using the following formula:

$$\delta_x^{2-1} = l_x^2(e_x^2 - e_x^1) - l_{x+1}^2(e_{x+1}^2 - e_{x+1}^1) \qquad (1)$$

Where δ_x is the difference in LE between two populations within age interval x; l_x is the number of life table survivors to age x; and e_x is the LE at the beginning of age interval x.

2. For measuring the contribution of cancer mortality to the LE change, was applied formula expressed as:

$$\Delta e_{i,x} = (m_{i,x}^2 - m_{i,x}^1) \div (m_x^2 - m_x^1) \times \delta_x^{2-1} \qquad (2)$$

Where $\Delta e_{i,x}$ is the contribution of cause-specific mortality to the life expectancy change within age interval x; and m_x and $m_{i,x}$ are age-specific and age-cause-specific mortality rates in the age interval x.

For a graphic presentation of the contribution of cancer mortality to the life expectancy change, the results have been assembled into more extended age groups: 0–34, 35–54, 55–74, and 75+, respectively.

3. Results

3.1 Cancer mortality trends

Since the 1950s, cancer mortality has continuously increased for both sexes in most European countries, reaching its peak in the late 1980s (Fig. 1). Such an increase in cancer mortality was evident in most countries despite the initial gap between them. Thus, in males in the 1950s, the age-standardized cancer mortality rate ranged between 200 deaths per 100 thousand population (in Ireland, Italy, Poland, and Portugal) and above 400 deaths per 100 thousand population in Finland and Switzerland. By the end of the 1980s (in Hungary, Poland, and Portugal by the end of the 1990s), cancer mortality in males grew by 25–200%. After reaching its peak, cancer mortality in males started to decline at a pronounced pace.

In females, between 1950–2019, cancer mortality levels out mainly reaching a plateau, with the highest values of mortality rate noted in the mid of the 1990s. A different situation was observed for females in Belgium, Denmark, the Netherlands and Switzerland, where the highest cancer mortality rates were recorded during the 1950s.

3.2 Contribution of cancer mortality to the life expectancy change

Between 1950 and 1990, in Norway, Czechia, Portugal, Spain, Poland, and Hungary, cancer mortality contributed with −0.2 years, −0.6 years, −0.6 years, −1.1 years, −1.2 years, and −1.4 to the overall LE growth. In the rest of the countries, except Finland, a peak in male cancer mortality was observed in the 1980s – deaths that from 1950 diminished the growth in LE in Switzerland (−0.2 years), Sweden (−0.3 years), Denmark (−0.5 years), Ireland (−0.5 years), Belgium (−0.6 years), France (−0.8 years), the Netherlands (−0.8 years), and Italy (−1.0 years). In Finland, due to cancer mortality, a slight decline in overall male LE change was found only between 1950–1960.

From 1980 to 1990, a decrease in males cancer mortality has led to a considerable gain in LE with a cumulative contribution by 2019 of 0.2 years in Portugal, 1.1 years in Ireland, Norway and Poland, 1.2 years in Hungary and Sweden, 1.4 years in Spain, 1.5 years in Denmark, 1.9 years in Finland, France and Italy, 2.0 years in the Netherlands and Czechia, and 2.2 years in Belgium.

Between 1950 and 1960, in most countries, LE slightly declined among females due to cancer mortality, except for Ireland, Italy, and Spain, where a decrease in LE was more pronounced by 0.1, 0.2, and 0.4 years, respectively. Among females in Norway, the Netherlands, Finland, Belgium, and Switzerland, cancer mortality has continuously decreased since the 1950s. This mortality trend increased LE to 2019 by 1.0, 1.2, 1.4, 1.5, and 1.8 years. In the period 1960–2019, female LE rose in Portugal (0.5 years), Spain (0.8 years), Italy (0.9 years), France (1.0 years), Ireland (1.0 years), Sweden (1.2 years), and Denmark (1.4 years) as a result of cancer mortality decline. In Poland, Hungary, and Czechia, female cancer mortality started to decline after the 1990s, increasing LE by 0.4, 0.6, and 1.2 years.

3.3 Age components of cancer mortality in the life expectancy change

In most countries, cancer mortality among males and females younger than 35 years contributes only negligibly to increase in LE in the period analyzed (Fig. 2). A significant decrease in LE due to cancer mortality was observed in males in the period 1950–1990,

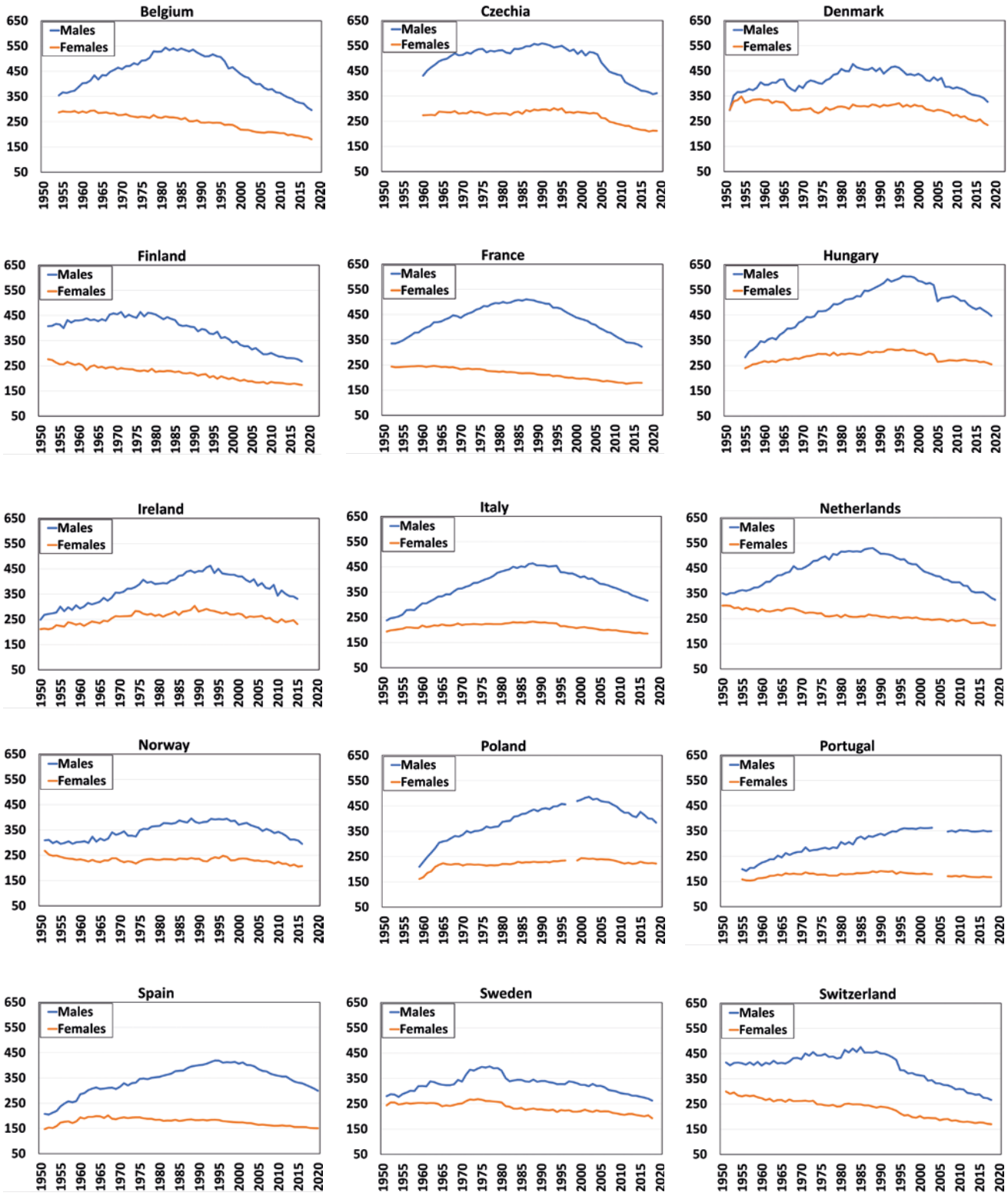


Fig. 1 Age-standardized cancer mortality rate (per 100,000 population) in 15 European high-income countries for males (A) and females (B), 1950–2019.
Source: WHO Mortality Database and Human Mortality Database

mainly in the age groups above 35, while in females, this decline was specific only to a few countries (Czechia, Hungary, Ireland, Italy, Poland, Portugal, and Spain). A considerable increase in females' LE for the 1950–1990 period was found in Belgium, the Netherlands, Switzerland and Nordic countries due to cancer mortality decline in all age groups.

After 1990 LE increase due to cancer mortality reduction was evident for males and females in all countries analyzed. This was mainly assured by the

Tab. 2 Contribution (in years) to the change in life expectancy by cancer mortality for males and females, 1950s–2019*.

	Years						
Country	1950s–1960	1960–1970	1970–1980	1980–1990	1990–2000	2000–2010	2010–2019*
Males							
Belgium	−0.28	−0.14	−0.19	0.25	0.56	0.64	0.77
Czechia	−0.15	−0.24	−0.12	−0.12	0.45	0.85	0.70
Denmark	−0.19	−0.03	−0.26	0.01	0.27	0.62	0.63
Finland	−0.07	0.09	0.26	0.44	0.44	0.31	0.37
France	−0.28	−0.21	−0.33	0.07	0.53	0.79	0.57
Hungary	−0.20	−0.27	−0.45	−0.52	−0.05	0.49	0.68
Ireland	−0.23	−0.20	−0.06	0.02	0.22	0.78	0.39
Italy	−0.36	−0.27	−0.34	0.08	0.64	0.64	0.49
Netherlands	−0.28	−0.31	−0.16	0.23	0.42	0.52	0.80
Norway	0.04	−0.16	0.05	−0.10	0.13	0.45	0.51
Poland	–	−0.56	−0.33	−0.29	0.12	0.49	0.53
Portugal	−0.18	−0.17	−0.14	−0.10	−0.01	0.09	0.10
Spain	−0.42	−0.16	−0.22	−0.30	0.10	0.59	0.66
Sweden	−0.17	0.04	−0.13	0.23	0.23	0.41	0.35
Switzerland	0.18	−0.05	0.04	0.12	0.65	0.63	0.53
Females							
Belgium	0.03	0.15	0.14	0.22	0.36	0.22	0.41
Czechia	−0.06	0.04	−0.01	−0.05	0.24	0.58	0.36
Denmark	−0.01	0.14	−0.06	0.04	0.13	0.55	0.62
Finland	0.17	0.27	0.17	0.07	0.24	0.24	0.24
France	−0.04	0.18	0.12	0.18	0.19	0.25	0.07
Hungary	−0.09	0.01	−0.11	−0.07	0.01	0.33	0.26
Ireland	−0.10	−0.16	0.03	0.04	0.15	0.62	0.26
Italy	−0.15	0.03	0.06	0.04	0.32	0.25	0.19
Netherlands	0.16	0.06	0.34	−0.04	0.09	0.16	0.38
Norway	0.20	0.06	0.11	−0.07	0.09	0.26	0.31
Poland	–	−0.31	−0.06	−0.02	0.00	0.27	0.17
Portugal	−0.07	−0.11	0.04	0.02	0.20	0.22	0.06
Spain	−0.36	0.02	0.12	0.07	0.19	0.22	0.19
Sweden	−0.08	0.17	0.04	0.29	0.14	0.29	0.26
Switzerland	0.24	0.10	0.26	0.11	0.44	0.37	0.24

Source: WHO Mortality Database and Human Mortality Database
* Last available year

age groups between 35 and 74 years, while in countries with the highest LE, a decline in cancer mortality was observed in the ages above 75.

4. Discussion

The results highlighted a considerable impact of cancer mortality on LE change during the last 70 years. The increase in cancer mortality observed until the end of the 1980s corresponds with the period of the aging cohorts of heavy smokers (Janssen and Van Poppel 2015) and occurred after a continuous decline in mortality from cardiovascular diseases (Vallin and Meslé 2004). Even though mortality from certain types of cancer (such as stomach, uterus, etc.) decreased since the 1950s, the cancers directly connected to smoking registered a visible increase until the end of the 1980s (Meslé 2002). Nevertheless, the age- and sex-differentiation in cancer mortality contribution to the LE change relied on divergence in the risk factors and predominant types of cancer in males and females (Radkiewicz et al. 2017; McCartney et al. 2011).

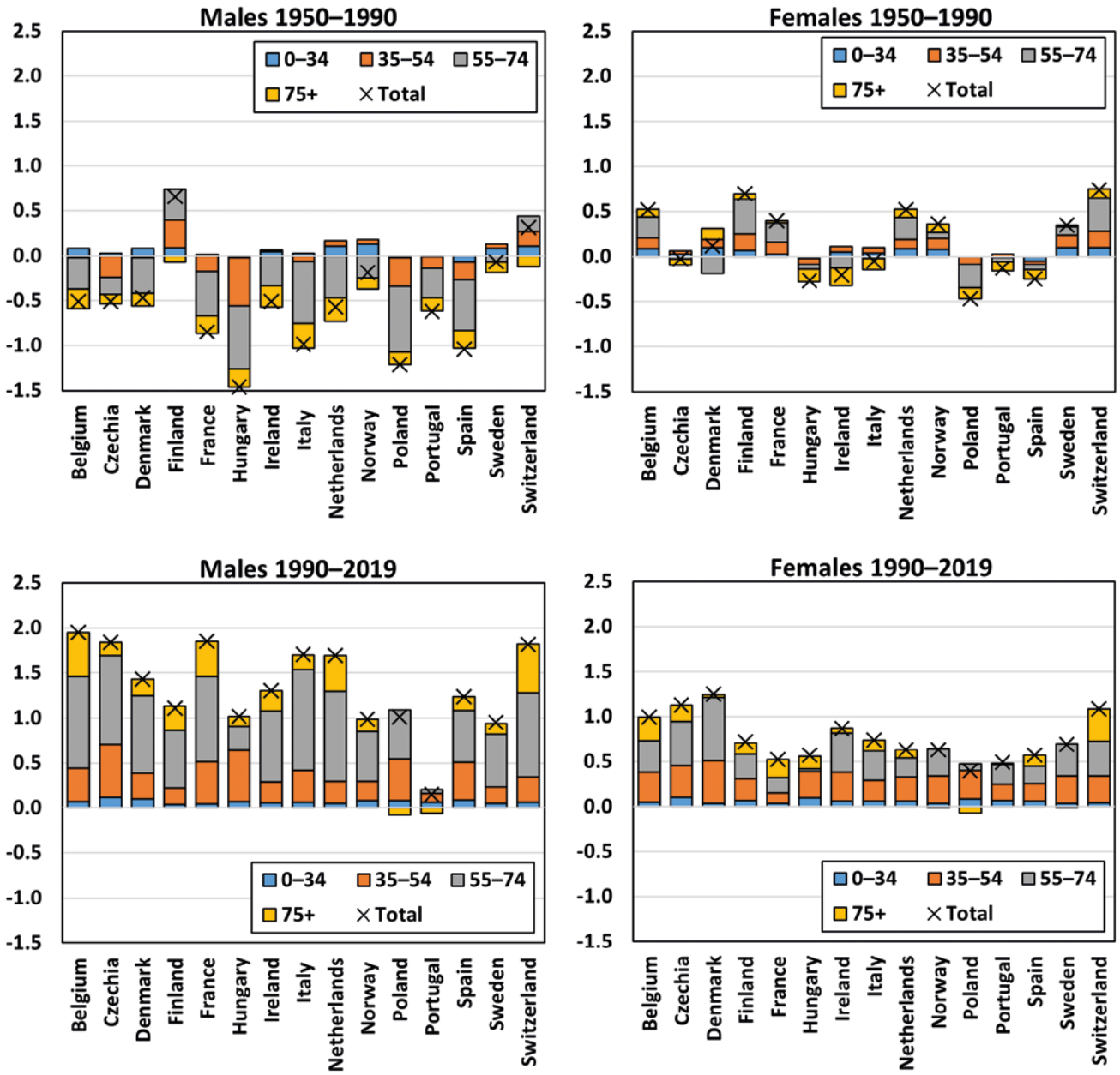


Fig. 2 Contribution (in years) of cancer mortality to the increase in life expectancy by age groups and sex, 1950–1990 and 1990–2019*.
Source: WHO Mortality Database and Human Mortality Database
* Last available year

In the 1990s, cancer mortality began to fall, which resulted from the mutual influence of cancer-preventive policies and significant improvements in cancer treatment. An outstanding contribution to reducing cancer mortality has been made by discovering new cancer treatment methods, including surgery, chemotherapy and radiation therapy (DeVita and Rosenberg 2012), introducing alcohol and tobacco control policies (Gredner et al. 2021), vaccination against hepatitis B (Lee et al. 1998) and efficient treatment of hepatitis C. An important contribution to cancer prevention and risk-factors reduction is made by a general increase in population education and welfare (Arık et al. 2021). However, such results in cancer mortality reduction were possible under an efficient bureaucratic system and high population responsiveness towards the health policies implemented in European countries (Mackenbach and McKee 2015).

The role of birth cohorts in determining cancer mortality trends has been distinctly described in earlier studies (La Vecchia et al. 1998). Consequently, the changes in behavioural factors that contribute to cancer mortality increase (e.g. high prevalence in alcohol and tobacco consumption), or the influence of cancer preventive policies and improvement in its treatment, will contribute to the changes in cancer mortality trends via cohort and period effects.

A substantial increase in LE due to a decrease in cancer mortality in the Central European countries

of the former socialist bloc might be explained by the so-called low base effect when in previous years, in Czechia, Hungary, and Poland, cancer mortality was significantly higher compared to Western European countries. Catching up economies of these countries allowed for adjusting cancer prevention and treatment standards, which contributed to a convergence in cancer mortality in the 15 European countries investigated (Bremberg 2017).

In recent years, during the COVID-19 pandemic, a decrease in the number of planned cancer treatments and screening procedures was observed (Englum et al. 2021; Mayo et al. 2021). This disruption might have a long-term impact on cancer mortality. However, continuous improvements in cancer treatment and early detection significantly contribute to cancer mortality reduction (DeVita and Rosenberg 2012).

5. Limitations

The study is based on WHO long-term data on age, sex, and cause-of-death mortality distribution, which primarily was provided by the national authorities, that not necessarily have an identical approach in codifying deaths according to the ICD classification (e.g. in diagnosing the underlying diseases that caused a fatal outcome). Additionally, the countries studied had different transition periods between the ICD classifications, which may affect the comparability results between the countries and periods analysed.

The method of decomposition used to measure the impact of cancer mortality on the LE change, in fact, measures the contribution of cause-specific mortality in the overall LE indicator. Thus, a decline in violent, infectious or non-communication disease mortality could contribute to an increase in exposure of cancer fatality, and, on the contrary, a decline in cancer mortality would increase in later ages mortality.

6. Conclusions

Cancer mortality continuously increased in 15 European countries until the late 1980s and mid-1990s, which coincided with a long period of growth in the population's lifespan, the aging of cohorts of heavy smokers, and an increase in behavioral and risk-factors incidence. After reaching its peak in the 1990s, cancer mortality has significantly declined, contributing to a considerable increase in LE, especially in males.

Recent decline in cancer mortality is one of the leading resources for increasing LE in most high-income countries. As cancer is nowadays one of the leading causes of death, understanding of past and future trends in cancer mortality and its contribution to the LE may be of importance for policymakers in order to set up preventive strategies.

References

Ait Ouakrim, D., Pizot, C., Boniol, M., Malvezzi, M., Boniol, M., Negri, E., Bota, M., Jenkins, M. A., Bleiberg, H., Autier, P. (2015): Trends in colorectal cancer mortality in Europe: retrospective analysis of the WHO mortality database. BMJ 351: h4970, https://doi.org/10.1136/bmj.h4970.

Andreev, E. M., Shkolnikov, V., Begun, A. (2002): Algorithm for decomposition of differences between aggregate demographic measures and its application to life expectancies, healthy life expectancies, parity-progression ratios and total fertility rates. Demographic Research 7(14), 499–522, https://doi.org/10.4054/DemRes.2002.7.14.

Araghi, M., Soerjomataram, I., Jenkins, M., Brierley, J., Morris, E., Bray, F., Arnold, M. (2018): Global trends in colorectal cancer mortality: projections to the year 2035. International Journal of Cancer 144(12), 2992–3000, https://doi.org/10.1002/ijc.32055.

Arık, A., Dodd, E., Cairns, A., Streftaris, G. (2021): Socioeconomic disparities in cancer incidence and mortality in England and the impact of age-at-diagnosis on cancer mortality. PLoS ONE 16(7): e0253854, https://doi.org/10.1371/journal.pone.0253854.

Baade, P. D., Youlden, D. R., Krnjacki, L. J. (2009): International epidemiology of prostate cancer: Geographical distribution and secular trends. Molecular Nutrition & Food Research 53(2), 171–184, https://doi.org/10.1002/mnfr.200700511.

Balakrishnan, M., George, R., Sharma, A., Graham, D. Y. (2017): Changing Trends in Stomach Cancer Throughout the World. Current Gastroenterology Reports 19, 36, https://doi.org/10.1007/s11894-017-0575-8.

Bertuccio, P., Alicandro, G., Malvezzi, M., Carioli, G., Boffetta, P., Levi, F., La Vecchia, C., Negri, E. (2019): Cancer mortality in Europe in 2015 and an overview of trends since 1990. Annals of Oncology 30(8), 1356–1369, https://doi.org/10.1093/annonc/mdz179.

Bremberg, S. G. (2017): Mortality rates in OECD countries converged during the period 1990–2010. Scandinavian Journal of Public Health 45(4), 436–443, https://doi.org/10.1177/1403494816685529.

Carioli, G., Malvezzi, M., Rodriguez, T., Bertuccio, P., Negri, E., La Vecchia, C. (2017): Trends and predictions to 2020 in breast cancer mortality in Europe. The Breast 36, 89–95, https://doi.org/10.1016/j.breast.2017.06.003.

DeVita, V. T., Rosenberg, S. A. (2012): Two Hundred Years of Cancer Research. New England Journal of Medicine 366(23), 2207–2214, https://doi.org/10.1056/NEJMra1204479.

Englum, B. R., Prasad, N. K., Lake, R. E., Mayorga-Carlin, M., Turner, D. J., Siddiqui, T., Sorkin, J. D., Lal, B. K. (2021): Impact of the COVID- 19 pandemic on diagnosis of new cancers: A national multicenter study of the Veterans Affairs Healthcare System. Cancer 128(5), 1048–1056, https://doi.org/10.1002/cncr.34011.

Eurostat (2013): Revision of the European Standard Population. Available online https://ec.europa.eu/eurostat/documents/3859598/5926869/KS-RA-13-028-EN.PDF/e713fa79-1add-44e8-b23d-5e8fa09b3f8f.

Gredner, T., Mons, U., Niedermaier, T., Brenner, H., Soerjomataram, I. (2021): Impact of tobacco control policies implementation on future lung cancer incidence in Europe: An international, population-based modeling

study. The Lancet Regional Health – Europe 4: 100074, https://doi.org/10.1016/j.lanepe.2021.100074.
Hashim, D., Boffetta, P., La Vecchia, C., Rota, M., Bertuccio, P., Malvezzi, M., Negri, E. (2016): The global decrease in cancer mortality: trends and disparities. Annals of Oncology 27(5), 926–933, https://doi.org/10.1093/annonc/mdw027.
HMD (n.d.): Human Mortality Database. Max Planck Institute for Demographic Research (Germany), University of California, Berkeley (USA), and French Institute for Demographic Studies (France). Available online https://mortality.org/.
Islami, F., Torre, L., Jemal, A. (2015): Global trends of lung cancer mortality and smoking prevalence. Translational Lung Cancer Research 4(4), 327–338, https://doi.org/10.3978/j.issn.2218-6751.2015.08.04.
Janssen, F., Van Poppel, F. (2015): The Adoption of Smoking and Its Effect on the Mortality Gender Gap in Netherlands: A Historical Perspective. BioMed Research International 2015:370274, https://doi.org/10.1155/2015/370274.
La Vecchia, C., Negri, E., Levi, F., Decarli, A., Boyle, P. (1998): Cancer mortality in Europe: effects of age, cohort of birth and period of death. European Journal of Cancer 34(1), 118–141, https://doi.org/10.1016/S0959-8049(97)00335-3.
Lee, M.-S., Kim, D.-H., Kim, H., Lee, H.-K., Kim, Ch.-Y., Park, T.-S., Yoo, K.-Y., Park, B.-Y., Ahn, Y.-O. (1998): Hepatitis B vaccination and reduced risk of primary liver cancer among male adults: a cohort study in Korea. International Journal of Epidemiology 27(2), 316–319, https://doi.org/10.1093/ije/27.2.316.
Mackenbach, J. P., McKee, M. (2015): Government, politics and health policy: A quantitative analysis of 30 European countries. Health Policy 119(10), 1298–1308, https://doi.org/10.1016/j.healthpol.2015.08.017.
Malvezzi, M., Carioli, G., Bertuccio, P., Boffetta, P., Levi, F., La Vecchia, C., Negri, E. (2019): European cancer mortality predictions for the year 2019 with focus on breast cancer. Annals of Oncology 30(5), 781–787, https://doi.org/10.1093/annonc/mdz051.
Mayo, M., Potugari, B., Bzeih, R., Scheidel, C., Carrera, C., Shellenberger, R. (2021): Cancer Screening During the COVID-19 Pandemic: A Systematic Review and Meta-analysis. Mayo Clinic Proceedings: Innovations, Quality & Outcomes 5(6), 1109–1117, https://doi.org/10.1016/j.mayocpiqo.2021.10.003.
McCartney, G., Mahmood, L., Leyland, A., Batty, G., Hunt, K. (2011): Contribution of smoking-related and alcohol-related deaths to the gender gap in mortality: evidence from 30 European countries. Tobacco Control 20(2), 166–168, https://doi.org/10.1136/tc.2010.037929.
Meslé, F., Vallin, J. (2002): Mortality in Europe: The Divergence Between East and West. Population 57(1), 157–197, https://doi.org/10.3917/popu.201.0171.
Proctor, R. (2013): The history of the discovery of the cigarette–lung cancer link: evidentiary traditions, corporate denial, global toll. Tobacco Control 21(2), 87–91, https://doi.org/10.1136/tobaccocontrol-2011-050338.
Radkiewicz, C., Johansson, A. L., Dickman, P. W., Lambe, M., Edgren, G. (2017): Sex differences in cancer risk and survival: A Swedish cohort study. European Journal of Cancer 84, 130–140, https://doi.org/10.1016/j.ejca.2017.07.013.
Remon, J., Reguart, N., García-Campelo, R., Conde, E., Lucena, C.-M., Persiva, O., Navarro-Martin, A., Rami-Porta, R. (2020): Lung Cancer in Spain. Journal of Thoracic Oncology 16(2), 197–204, https://doi.org/10.1016/j.jtho.2020.09.026.
Stringhini, S., Guessous, I. (2018): The Shift from Heart Disease to Cancer as the Leading Cause of Death in High-Income Countries: A Social Epidemiology Perspective. Annals of Internal Medicine 169(12), 877–878, https://doi.org/10.7326/M18-2826.
Townsend, N., Wilson, L., Bhatnagar, P., Wickramasinghe, K., Rayner, M., Nichols, M. (2016): Cardiovascular disease in Europe: epidemiological update 2016. European Heart Journal 37(42), 3232–3245, https://doi.org/10.1093/eurheartj/ehw334.
Vallin, J., Meslé, F. (2004): Convergences and divergences in mortality: a new approach of health transition. Demographic Research, Special Collection 2, 11–44, https://doi.org/10.4054/DemRes.2004.S2.2.
WHO Mortality Database (n.d.). Available online https://www.who.int/data/data-collection-tools/who-mortality-database.
Wong, M. C. S., Goggins, W. B., Wang, H. H. X., Fung, F. D. H., Leung, C., Wong, S. Y. S., Wong, S., Y., S., Ng, C. F., Sung, J. J. Y. (2016): Global Incidence and Mortality for Prostate Cancer: Analysis of Temporal Patterns and Trends in 36 Countries. European Urology 70(5), 862–874, https://doi.org/10.1016/j.eururo.2016.05.043.

Air pollution and topography in Tehran

Faezeh Afarideh[1,*], Mohammad Hossein Ramasht[1], Graham Mortyn[2]

[1] Faculty of Geography and Planning, University of Isfahan, Iran
[2] Department of Geography, Universitat Autònoma de Barcelona, Spain
[*] Corresponding author: faezehafarideh@yahoo.com

ABSTRACT

Tehran is one of the most polluted cities in the world with 48 days of air pollution exceeding the admissible threshold (AQI > 150) for 3 months of the 15 years studied. This period coincides with the time when Tehran's inversion reaches its maximum stability. The purpose of this study was to determine the height of air pollution in Tehran in the days when pollution exceeds the permissible limit. Continuing to study the pressure and temperature conditions of these days, we then considered the geographical and topographic conditions, and finally identified the best of these cells for potential theoretical air turbulence. The results of this study, based on the Harmonic Analysis method and based on Tehran temperature and pressure data over a 15-year period (2003–2017), show that the highest elevation of Tehran inversion does not exceed 1800 m on polluted days. Only within 6 days of those beyond the admissible threshold, temperature and pressure cells with the highest Newtonian mass are formed. The center of these cells formed with a compressive difference of 32 mg in November, 7 mg in January, 11 mg in December, and temperature difference of 1.1° in November, 4.4° in January, and 1.9° in December. Generally, we considered the formed cells by the temperature and pressure difference and the gradient between them, as well as the difference in height between the cells and their location. This information, combined with the local winds causing the differences in temperature and pressure, allows us to elucidate conditions for creating air turbulence in Tehran and mitigating the amount and degree of air pollution.

KEYWORDS

inversion; air pollution; thermal cells; pressure cells; air quality

Received: 9 November 2021
Accepted: 18 September 2023
Published online: 20 October 2023

Afarideh, F., Ramasht, M. H., Mortyn, G. (2023): Air pollution and topography in Tehran. AUC Geographica 58(2), 157–171
https://doi.org/10.14712/23361980.2023.12
© 2023 The Authors. This is an open-access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0).

1. Introduction

Air quality significantly impacts human life and the condition of ecosystems, both directly and through processes related to climate change.

Air pollution is associated with respiratory and eye diseases such as asthma, lung cancer and conjunctivitis, especially in the young and elderly (UNEP/WHO 1992; Patel 1994). Lead as a pollutant is particularly serious for children, since relatively low concentrations of lead in the blood may have a damaging and permanent effect on their mental development (Needleman et al. 1991). In the environment, air pollution is a major contributor to effects such as acid rain, which has been responsible for much damage to soil, fish resources and vegetation, often very far away from the source of the pollutant (Shahid and Hussain 2013).

Air pollution is identified by the World Health Organization (WHO) as responsible for several million deaths per year. In most industrialized countries, particulate matter emissions are regulated by severe environmental laws because smog crises can involve large territories and affect their residents. Several studies carried out in the U.S.A. and Europe indicate high risks for lung cancer due to small particulate matter (e.g. Vineis et al. 2006).

Tehran is one of the largest and the most crowded cities that suffer from air pollution. In some days of the year, the amount of contaminating and pollution elements increases so much that breathing is very difficult for inhabitants. Increasing atmospheric pollutants has made Tehran one of the most polluted cities in the world. Air pollution concentrations vary substantially within a year. The AQI varies over the course of a year in Tehran. During autumn and winter, Tehran becomes more polluted. Atmospheric temperature inversion worsens air pollution during that period. Spring and summer are usually less polluted. August, for example, was the cleanest month of the year in 2015. There are occasional dust storms at the end of spring and beginning of summer that increase PM concentrations and result in unhealthy air. The unhealthy days due to dust storms are relatively rare compared to those caused by combustion sources during the cold seasons (Sarraf and Heger 2018). To design an effective approach to air pollution management, it is important to diagnose the problem, determine its sources, and identify affordable and sustainable solutions (Gwilliam et al. 2004).

There are more than 17 million vehicular trips per day in Tehran (Shahbazi et al. 2017), and many of the vehicles have outdated technology. Thus, the air in Tehran is amongst the most polluted in the world. Topography and climate add to the pollution problem. Tehran is at a high altitude and is surrounded by the Alborz Mountain Range, which traps polluted air. Temperature inversion, a phenomenon particularly occurring during the winter months, prevents the pollutants from being diluted. Several recent trends indicate that reducing air pollution will not be straightforward: rapid population growth (partially due to migration from other cities and villages), industrial development, urbanization, and increasing fuel consumption are pressure points for clean air in Tehran.

The two factors of climate and topography are affecting air pollution in Tehran; these are 2 major issues emphasized in this research. In fact, using these factors, we are looking for a way to eliminate or at least decrease the pollution of Tehran's air.

All studies on atmospheric pollution in Tehran have focused on urban management, such as days free from school or work; avoidance of moving polluted cars; relocation of industrial factories, and so on.

This research focuses on vertical and horizontal exchanges via atmospheric mixing, by defining the good conditions for instability during the inversion periods in Tehran. If we have suitable mixing conditions (identified with cells of pressure and/or temperature), we should try to define the best status for instability. We must better know the differences of temperature and pressure that give rise to air turbulence.

2. Background

Related to air pollution many studies have been conducted and some of these have focused on physical factors such as climatology and topography.

Mage et al. (1996) made maps based on information from WHO and UNEP's creation of an air pollution monitoring network as part of the Global Environment Monitoring System. They described an increased population in the next century, mainly in the developing countries with a lack of capital resources for air pollution control, which means that there is a great potential that conditions will worsen in many more cities reaching megacity status[1]. This paper maps the potential for air pollution that cities will experience in the future unless control strategies are developed and implemented during the next several decades.

Thadathil and Gosh (1992) studied surface layer temperature inversion in the Arabian Sea during winter. They found that the inversion in this area is a stable seasonal feature and the occurrence is limited to the coastal regions. Finally, the possible forcing mechanism for such an advection (the causative factor for the inversion is identified to be the winter-time surface-advection of cold less saline Bay of Bengal water over the warm saline Arabian Sea water along the west coast of India) was examined using a

1 According to some recourses megacity is the city with more than 10 million people in habitants.

hydrographic section and wind observations along the west coast of India.

Kejna (2008) described the spatial differentiation of topoclimatic conditions in the vicinity of the Arctowski Station (King George Island, Antarctica). The paper analyzes meteorological elements such as air temperature, air humidity (eight sites) and wind direction and velocity (three sites). Significant topoclimatic diversities resulting from denivelation, exposure, ground properties and local air circulation were recorded in the study area.

Nowadays due to the increasing population, energy consumption, pollution caused by transport systems, and the increasing number of vehicles, industrial, and mining activities, in most big cities the important of environmental issues, has doubled, particularly with air pollution (Sarraf and Heger 2018).

Sánchez-Coyllo et al. (2002) studied the observed behavior of pollution concentrations due to prevailing meteorological conditions for the period from June 13 to September 2, 1994, for the Metropolitan Area of Sao Paulo (MASP). They observed that for both the synoptic conditions, namely the South Atlantic Subtropical High and Polar High, high concentration pollution values prevailed due to weak ventilation, low relative humidity and absence of precipitation.

Buchholz et al. (2010) studied impacts related to daily air quality index (DAQx), calculated for 15 air quality monitoring stations (traffic, background, and industry) in Belgium, France, Germany and Luxembourg, all compared to meso-scale atmospheric patterns between 2001 and 2007. The investigation of weather regimes indicates that zonal and mixed cyclonic circulation regimes are associated with better air quality than meridional and anticyclone weather regimes. In general, weather regimes with high daily precipitation lead to better air quality than dryer air masses because of lower contribution of PM10 to the air quality index.

Bityukova et al. (2013) in a research report in Russia, found that 60 million people live in the cities with high levels of air pollution. Comparative statistical analysis of pollutant emission and emission processes in 1099 cities in the country revealed the role of climate and other environmental factors, including fuel mix, and the impact of agglomeration effecton the distribution of pollutants in the cities' atmosphere.

Bidokhty and Hashem (1998) explained about the atmospheric boundary layer (ABL) and its effect on air pollution. They created a turbulent and integral model for prediction of air pollution in city.

Like all major cities, Tehran suffers from air pollution; however here it is particularly intensive. The amount of air pollutants in the warning state[2] has increased considerably. In order to find the solution for decreased air pollution many researchers have addressed the issue from a range of perspectives

Cheragi (1999) studied air quality in Tehran and Isfahan. He concluded that density of pollutants, in 329 days in Tehran and 34 days in Isfahan exceeded the limit. Shariepour et al. (2006) studied the density of pollution in Tehran. They concluded that Golhak and Tajrish stations are the most polluted stations and that reduced relative humidity was causing increased air pollution during the cold months of the year.

Zahedi et al. (2000) studied about factors that influence at air pollution in Tabriz. They concluded that the most important factor is industries located in the west and south west of this city. Other factors are topography, slope and wind.

Lashkari and Hedayat (2007) studied synoptic patterns of inversions in Tehran and explained four synoptic patterns as causing intense inversions in Tehran. In two of these patterns, creating a ridge in 850 and 700 hectopascals and dynamical stability derived from it, are intense inversion conditions. In another pattern, transfer of cold weather by anticyclonic circulation at ground level creates a trough over Tehran of 850 and 700 hectopascals, thus causing intense inversions.

Esa et al. (2012) studied the air pollution distribution in Tehran. They claimed that concentrations of various air pollutants and inappropriate distribution of resources in the city causes Tehran to have different areas with different situations of air quality, and concluded that the western and southern parts of the city are the most polluted.

Saneie (2016) studied air pollution in Tehran from the perspective of urban planners. He concluded that density of pollution in Tehran follows the topography and that most air pollution is in low areas in Tehran. He considered it impossible to change climate factors that are influencing air pollution. He suggested that a changing pattern of city management and urban planning can better prepare Tehran for changing environmental conditions.

3. Main objective

The main motive of this research is to focus on changing atmospheric patterns for purposes of better management of air pollution risks in the Tehran basin. If this plan (to find a Model for Instability in Thermal and Pressure Cells and thus curtailing air pollution) succeeds, a better understanding of temperature and pressure conditions (quantities from synoptic stations and radiosonde launches, overall gradients both horizontally and vertically, etc.) will be achieved. This new knowledge should better prepare us for preventing air pollution risks in the Tehran region.

The main objective of this research is to find a model for instability in Thermal and Pressure Cells and thus curtailing Tehran air pollution. The more concrete and specific objectives are as follows

- Investigate air conditions in the days when the AQI (Air Quality Index) are greater than 150 in Tehran.

2 When the Air Quality Index represents more than 150.

- Determine temperature and pressure cells in different levels of the atmosphere. From this data we can construct maps that most clearly define instability conditions.
- Select a level for which thermal and pressure cells have the most differences, thus defining the most clearly unstable conditions suitable for mixing (and pollution).

4. Research method

To achieve these objectives, different stages will be pursued. Firstly, Tehran's air pollution data was collected from air quality sensors and, according to them, the days when the AQI is greater than 150, have been pre-selected as dangerous days of pollution.

Then, temperature and pressure data were collected for polluted days from synoptic stations and after that, the pressure and temperature maps were drawn at different levels of the atmosphere.

Further, based on these maps, it should be selected the levels that have the most number cells of pressure and temperature with the most gradient.

Then, maps of thermal and pressure cells that have the most differences will be selected for more concrete analysis and consideration.

In the final stage, it will be suggested the amount of differences in temperature and pressure that cells should have to create instability. Upon completion of this, it is anticipated that we will better understand the inversion processes that are major influences on air pollution in Tehran. This will hopefully progress toward knowledge and understanding of how to mitigate air pollution risks.

Tab. 1 Date of dangerous days of pollution.

Year	January	November	December
2017	1, 9		
2016		30	24, 23, 1
2015			30, 29, 14
2014	12		28
2013	8	27, 28	23, 24, 25, 26, 29
2012			2, 3, 4, 5, 20, 31
2011	20		
2010		22, 23, 27, 29	31, 30, 28, 27, 26, 19, 12, 11, 8, 4, 3, 2
2009			
2008			
2007			
2006			
2005	1		6, 7, 4
2004	5		
2003	13		

We have used the synoptic stations and the air pollution testing stations as well as Google earth, Arc GIS, Surfer and Voxler software.

5. Expected analyses

To achieve the research aims, at first, the days when the AQI (Air Quality Index) were greater than 150, during 15 years from 2003 until 2017 (just the three most polluted months of November, December and January) were defined. The reason why we select these 3 months is because they show the most polluted months in all the studied years.

In these days the values of temperature and pressure were defined from the synoptic stations and radiosonde measurements. Levels of inversion height were calculated in these 3 months according to these factors.

In the next step, we draw maps using Surfer software, according to inversion levels, the connections between height and temperature, and height and pressure, in each level.

In the following, from the maps of the above step, some maps will be selected that illustrate the largest gradients among thermal and pressure cells. Now the main aim is to find amounts of temperature and pressure gradients that cells should have between them to create instability.

6. Results and discussion

For the first step we find the days when the AQI (Air Quality Index) were greater than 150 as dangerous days of pollution from 2003 until 2017.

Based on the data in Tab. 1, during the statistical study period, a total of 48 critical days were identified, among which December was the most polluted month of the year with 33 days, followed by January with 8 and November with 7 days.

Temperature and pressure values were then extracted from the synoptic stations of Tehran province for the specified days. Fig. 1 shows the distribution of these stations taking into account their altitude.

According to Fig. 1, the Abali station has the highest altitude with 2462.2 m and the Varamin station has the lowest with 937 m.

Thereafter, upper atmosphere data were extracted for the critical pollution days. The data were collected at Mehrabad Airport, Tehran, and obtained from the University of Wyoming. We then analyzed these data and calculated and determined the ultimate inversion limit for November, December, and January.

In order to calculate the average inversion level using Radio sonde data, temperature and altitude were first evaluated for the critical days. Normally, with increasing altitude, temperature and pressure decrease. However, during inversion, the general

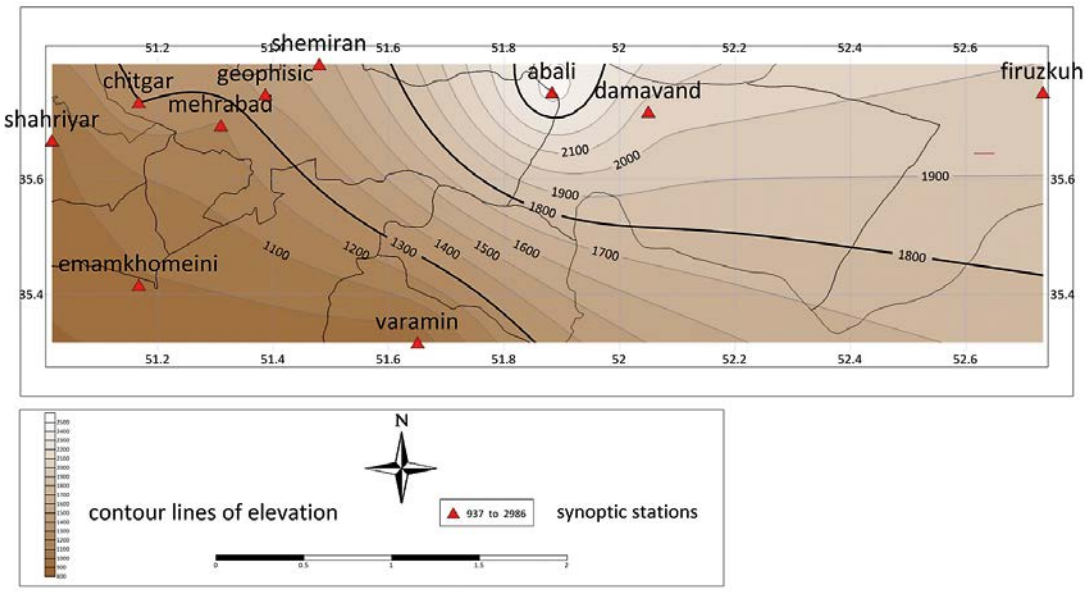


Fig. 1 Map of the synoptic stations in the study area.

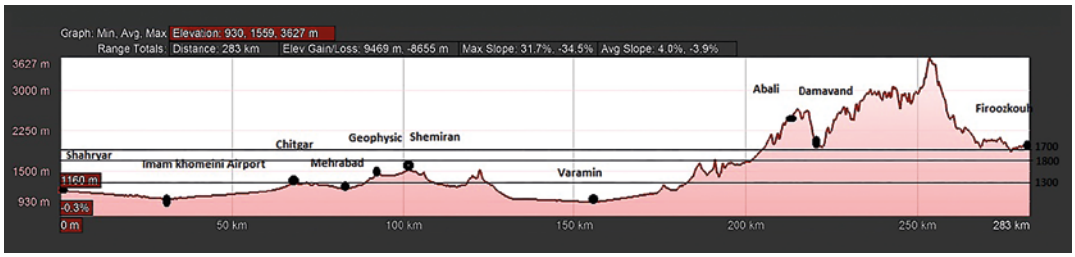


Fig. 2 Inversion height diagram for 3 months in the study area.

process changes and with increasing altitude, temperature also increases. This trend was studied on all critical days of air pollution in the three months in question. In this way, the trend of temperature changes was studied with increasing altitude, and the inversion limit was determined for each station up to the height at which temperature still increased. Accordingly, the inversion level was determined on the critical days for each month and then their averages were calculated for the desired month. The average inversion levels calculated for the three months in question are shown in Fig. 2.

In this diagram (Fig. 2), the position of each station relative to the average final height of inversion is shown for the three months on critical days. Accordingly, the inversion height for critical days is 1300 m and includes Varamin, Mehrabad, Chitgar, Imam Khomeini and Shahriar stations. Other stations fall outside the range of inversion (Damavand, Firouzkouh, Abali, Shemiran, Geophysics).

In December, the average final height of inversion is 1700 m. Therefore, only the Firoozkooh, Damavand, and Abali stations are outside of inversion. And in January, the average final height of inversion rises slightly to 1800 m, so, again, all lower-altitude stations fall into the inversion zone except for Firoozkooh, Damavand, and Abali stations. In fact, the two stations of Damavand and Abali, with the altitude of above 2000 m, will never experience inversion and air pollution because this phenomenon occurs at a maximum height of 2000 m.

The results of this part of study can be summarized as follows.

- The height of inversion phenomenon in Tehran is not the same in the target months (January, November, December).
- The highest inversion height in the target months is 1800 m and the lowest is 1300 m.
- The highest inversion is related to January (1800 m) and the lowest is related to November (1300 m).
- Exceedance of the AQI index or the pollution crisis threshold does not cover all areas of Tehran in the target months, that is, while some districts of Tehran experience higher pollutions than the thresholds, others do not.
- During December, the expanse of pollution in Tehran is wider than other target months.

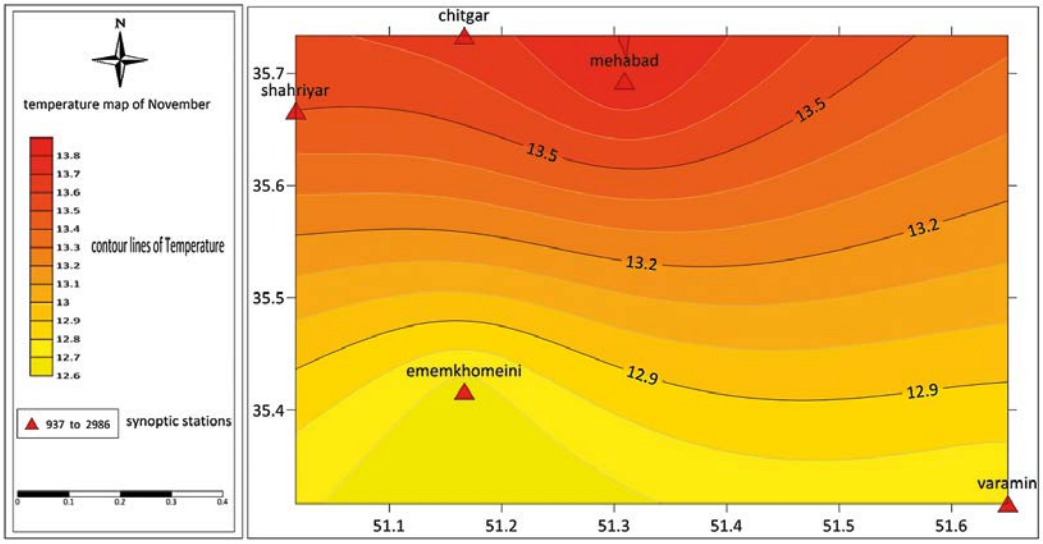


Fig. 3 Map of two temperature cells with the highest Newtonian mass in 28/11/2013 in November.

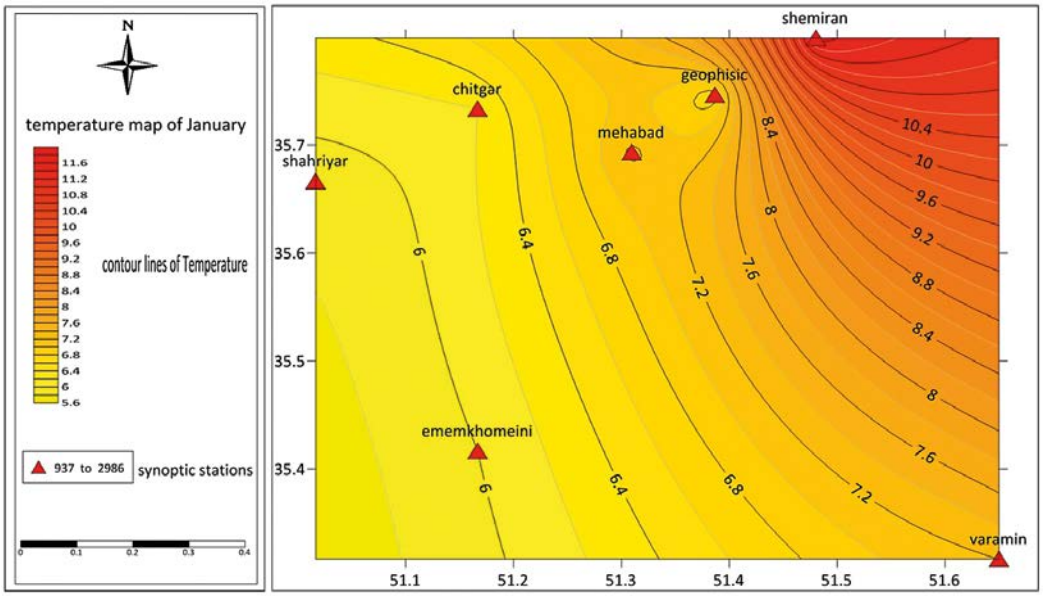


Fig. 4 Map of two temperature cells with the highest Newtonian mass in 8/1/2013 in January.

6.1 Determination of the conditions and location of temperature and pressure cells in critical days of pollution

In this study, after analyzing temperature and pressure data in critical days of pollution during the 15-year statistical period, the inversion was analyzed at different temperature and pressure levels and the mean inversion level was calculated for November, December and January.

Next, based on the determined inversion levels, zoning maps of pressure and temperature on critical days of pollution were drawn in the target months. In this regard, 48 maps of pressure cells and 48 maps of temperature cells were drawn for the critical days of pollution.

From among them, maps containing temperature and pressure cells were selected, and then a matrix was prepared for all cells in the selected maps and their Newtonian mass was calculated.

This matrix represents the cells that have the gradient, because the two factors of cell difference and distances play a major role in their triggering. In order to select the most favorable case from the obtained

Tab. 2 November temperatures Cells Matrix.

28/11/2013		
Temperature cell	Temperature	Newtonian's Mass
Mehrabad	13.75	0.1536
Imam Khomeini	12.65	

Tab. 3 January temperatures Cells Matrix.

8/1/2013			
Temperature cell	Temperature	Newtonian's Mass	
Geophysics	7.2	Shemiran-Geophysics	0.8011
Shemiran	11.6	Geophysics-Mehrabad	0.6124
Mehrabad	7.2	Mehrabad-Shemiran	0.2237

Tab. 4 December temperatures Cells Matrix.

12/12/2010			
Temperature cell	Temperature	Newtonian's Mass	
Varamin	8.8	Shemiran-Geophysics	0.7597
Imam Khomeini	8.3	Shemiran-Cheitger	0.0965
Chitgar	9.8	Shemiran-Imam Khomeini	0.0255
Geophysics	9.9	Shemiran-Varamin	0.0256
Shemiran	8.0	Geophysics-Chittger	0.2548
		Geophysics-Imam Khomeini	0.0472
		Geophysics-Varamin	0.0349
		Chitgar-Imam Khomeini	0.0598
		Chitgar-Varamin	0.0228
		Imam Khomeini-Varamin	0.0382

numbers, the Newtonian mass of the cells was calculated $f = (Mi \times Mj)/D2 \times 1000$.

Finally, for each month, two temperature and pressure cells with the highest Newtonian mass were selected. Following the temperature and pressure cells with the highest Newtonian mass in all three months, their matrix is provided.

Shown in Fig. 3 is the temperature map of November in extreme inversion up to the height of 1300 m, which indicates two cells at Imam Khomeini Airport Station and Mehrabad Station. These two cells have the highest Newtonian mass than other temperature cells in the other polluted days. The distance between the two existing cells is approximately 33.65 km. Imam Khomeini station is 990.2 m high and Mehrabad station is 1190.2 m high. The temperatures of cell centers are 13.75 °C for Mehrabad and 12.65 °C for Imam Khomeini Airport. Tab. 2 shows the matrix of the cells of the selected map (Fig. 3) and the Newtonian mass of its cells.

Shown in Fig. 4 is the temperature map of January in extreme inversion up to 1800 m. This map shows three temperature cells, with the largest Newtonian mass being related to Geophysics and Shemiran cells. The two cells are approximately 10.21 km apart. Shemiran station is at an altitude of 1548.2 m and Geophysics station at an altitude of 1423.8 m. Meanwhile, the temperature of cell centers is 7.2 ° for Geophysics station and 11.6 ° for Shemiran station.

Tab. 3 calculates the temperature of each cell, as well as the matrix between the cells and their Newtonian mass.

In Fig. 5, the temperature map of December is shown in the inversion up to 1700 m. This map shows the existence of 10 temperature cells, among which, Geophysics and Shemiran have the highest Newtonian mass. The two cells, with the altitude of about 1423.8 and 1548.2 m, respectively, are located at an approximate distance of 10.21 km from each other. Also the temperatures of their centers are 9.9 °C for Geophysics and 8 °C for Shemiran.

In Tab. 4 the Newtonian mass matrix of all temperature cells formed on 8/12/2010 is calculated.

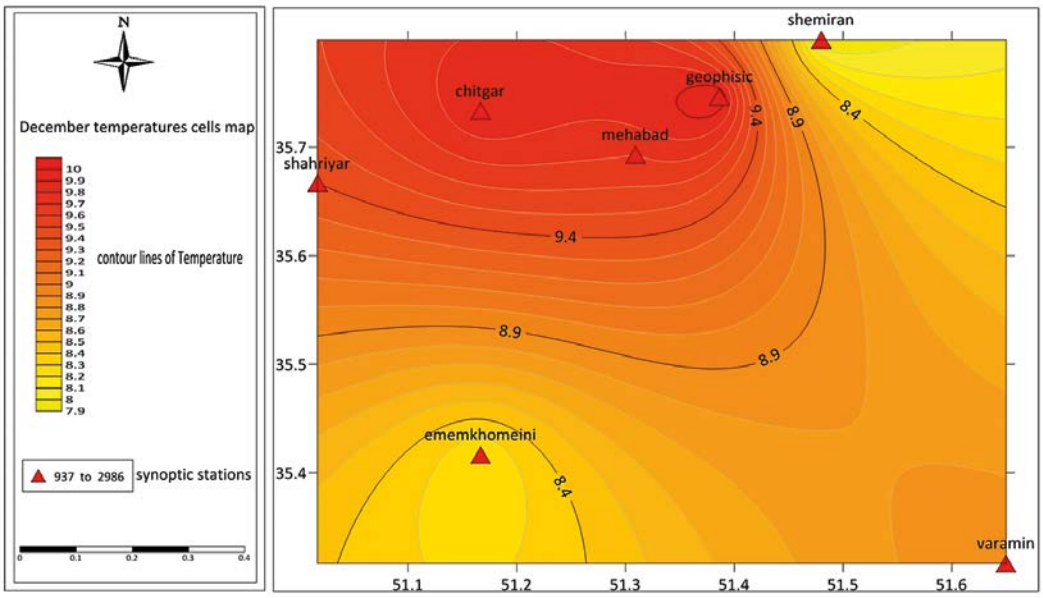


Fig. 5 Map of two temperature cells with the highest Newtonian mass in 12/12/2010 in December.

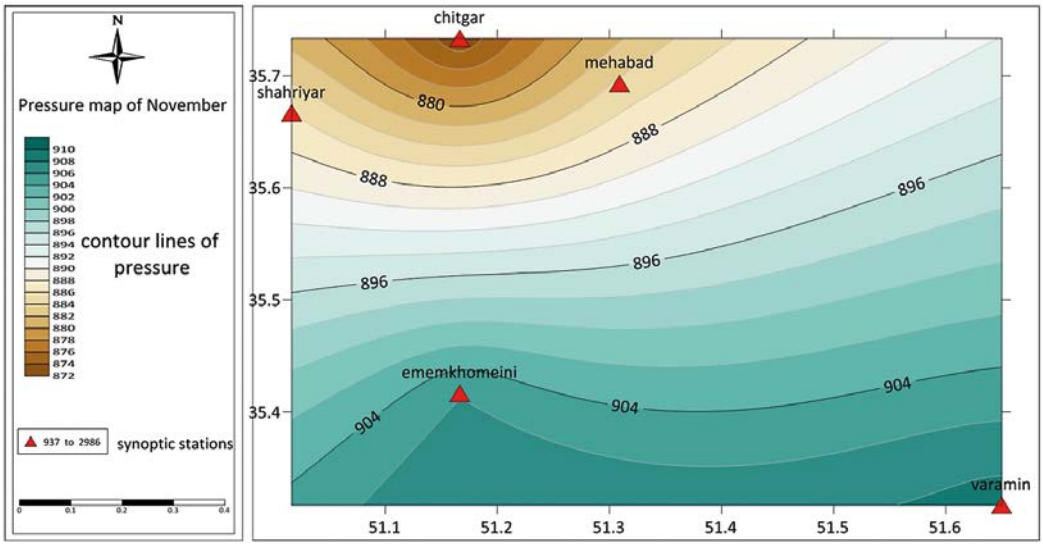


Fig. 6 Map of two Pressure cells with the highest Newtonian mass in 27/11/2013 in November.

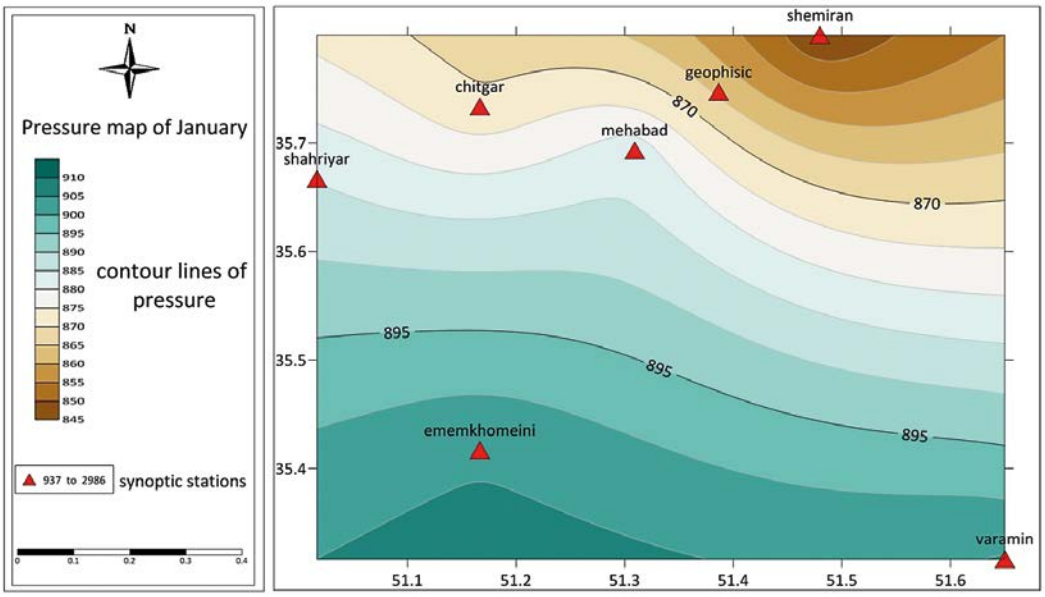


Fig. 7 Map of two Pressure cells with the highest Newtonian mass in 1/1/2017 in January.

Accordingly, the highest Newtonian masses are related to the Geophysics and Shemiran cells. In addition, the temperature values of each cell can also be seen in the table.

Fig. 6 shows a map of the pressure difference in the inversion at an altitude 1300 m in November. This map indicates the existence of two pressure cells, which have the highest Newtonian mass compared to other pressure cells on other polluted days. The two cells are approximately 36.88 km apart. Chitgar station is at the altitude of 1305.2 m and Imam Khomeini Airport station at 990.2 m. Pressure at the centers of Chitgar and Imam Khomeini cells is 874 and 906 hPa, respectively. Tab. 5 shows the temperature of the two cells of Chitgar and Imam Khomeini Airport, as well as their calculated Newtonian mass. Tab. 5 shows the temperature of the two cells of Chitgar and Imam Khomeini Airport, as well as their calculated Newtonian mass.

Fig. 7 shows a map of the pressure difference in the inversion at an altitude of 1800 in January. This map indicates four pressure cells, with Mehrabad and

Tab. 5 November pressure Cells Matrix.

27/11/2013		
Pressure cell	Pressure	Newtonian's Mass
Chitgar	874	616.8802
Imam Khomeini	906	

Tab. 6 January pressure Cells Matrix.

1/1/2017			
Pressure cell	**Pressure**	**Newtonian's Mass**	
Shemiran	850	Shemiran-Mehrabad	200.9521
Mehrabad	880	Shemiran-Imam Khomein	295.2137
Imam Khomein	903	Mehrabad-Imam Khomeini	701.7788
Cheitgar	873	Mehrabad-Cheitgar	3919.5918
		Shemiran-Cheitger	914.2148
		Imam Khomeini-Cheitger	579.5890

Tab. 7 December pressure Cells Matrix.

12/25/2013			
Pressure cell	**Temperature**	**Newtonian's Mass**	
Shemiran	895	Shemiran-Geophysics	6915.4999
Geophysics	904	Shemiran-Mehrabad	952.9616
Mehrabad	915	Geophysics-Mehrabad	8708.6483

Chittgar cells having the highest Newtonian mass. The two cells are about 14 km apart. Chitgar cell is at the altitude of 1305.2 m and Mehrabad cell at 1190.2 m. Pressure at the centers of Chitgar and Mehrabad cells was 873 and 880 hPa, respectively. In Tab. 6 the pressure of each cell and the Newtonian mass matrix of all cells are calculated.

In Fig. 8, three pressure cells formed in December can be seen at an altitude of 1700 m. According to calculations, the highest Newtonian masses belong to Geophysics and Mehrabad cells. The two cells, with the altitude of about 1423.8 and 1190.2 m, respectively, are located at an approximate distance of 9.2 km from each other. The pressure at the centers of the geophysical and Mehrabad cells is 904 and 915 hPa, respectively. In Table 7 the pressure of each cell and the Newtonian mass matrix of all cells are calculated.

6.2 Identification of general conditions of the study area on critical days of pollution during 15 years

In order to better analyze the conditions of the critical days of pollution, we calculated the mean of temperature, pressure and wind speed and then maps of these factors were plotted. The matrix of cells formed in each of these maps was then calculated to allow the comparison of the general conditions of the days of contamination with those of the cells with the highest Newtonian mass and temperature.

In the zoning map of the average temperature of the polluted days (Fig. 9), we observe the formation of four temperature cells, among which the Shemiran and Mehrabad cells have the highest Newtonian mass. The distance between the two cells is approximately 19.32 km. The central temperature of Shemiran cell is 16.05 °C and the central point temperature of Mehrabad cell is 18.4 °C. Tab. 8 shows the temperature of the cells in the map as well as their Newtonian mass values.

The following results are obtained by examining the maps of the average temperature of the critical days of pollution over 15 years, as well as a selection of daily temperature cell maps for the three target months.

a. The map of means shows the formation of four temperature cells including Shemiran, Mehrabad, Varamin and Imam Khomeini cells, with the largest Newtonian mass belonging to Shemiran and Mehrabad station.
b. We obtained these results from the daily temperature cell maps for the three months of November, January, and December.

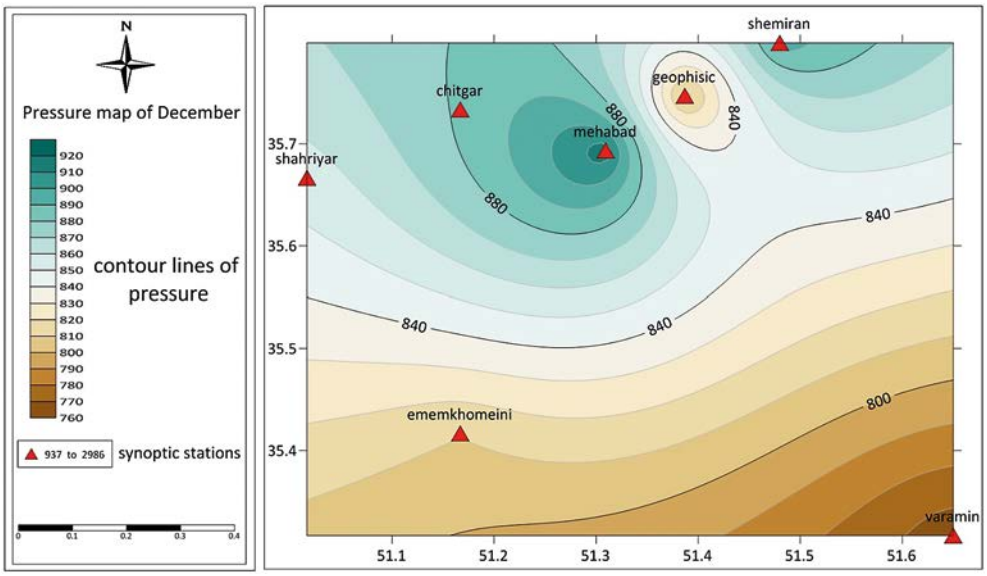


Fig. 8 Map of two Pressure cells with the highest Newtonian mass in 12/25/2013 in December.

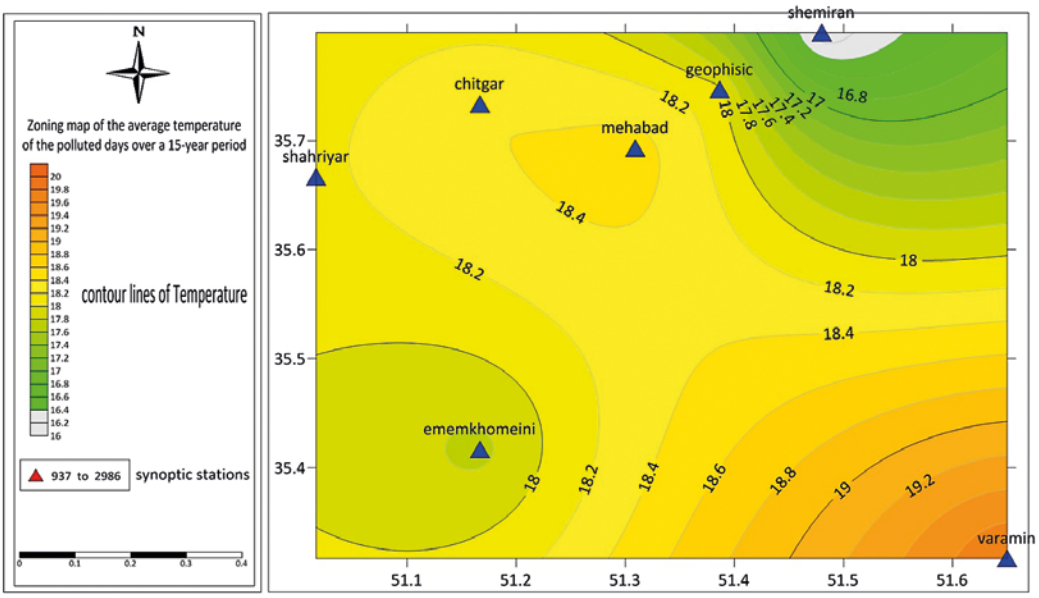


Fig. 9 Zoning map of the average temperature of the polluted days over a 15-year period.

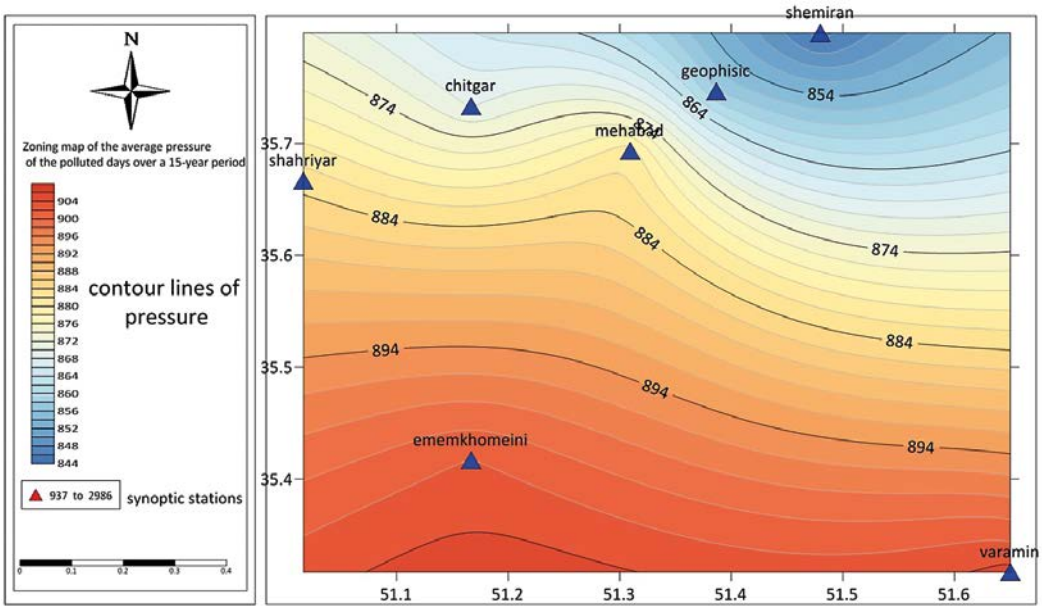


Fig. 10 Zoning map of the average pressure of the polluted days over a 15-year period.

Tab. 8 The average temperature Cells Matrix of the polluted days.

The average temperature over a 15-year period			
Temperature cell	Temperature	Newtonian's Mass	
Shemiran	16.05	Shemiran-Varamin	0.00500
Varamin	19.8	Shemiran-Imam Khomeini	0.00060
Imam Khomeini	17.7	Shemiran-Mehrabad	0.00620
Mehrabad	18.4	Varamin-Imam Khomeini	0.00109
		Varamin-Mehrabad	0.00590
		Imam Khomeini-Mehrabad	0.00061

Tab. 9 The average pressure Cells Matrix of the polluted days.

The average pressure over a 15-year period			
Pressure cell	Temperature	Newtonian's Mass	
Shemiran	845.16	Shemiran-Cheitger	0.030
Chitger	869.84	Shemiran-Imam Khomeini	0.021
Imam Khomeini	902.09	Shemiran-Mehrabad	0.096
Mehrabad	881.10	Chitgar-Imam Khomeini	0.023
		Chitgar-Mehrabad	0.057
		Imam Khomeini-Mehrabad	0.018

Mehrabad and Imam Khomeini cells have the highest Newtonian mass in November. In January, there are the Geophysics, Shemiran and Mehrabad temperature cells, among which, Geophysics and Shemiran have the highest Newtonian mass. December has five temperature cells, including Varamin, Imam Khomeini, Chitgar, Geophysics and Shemiran, among which, Geophysics and Shemiran have the highest Newtonian mass.

Accordingly, it can be concluded that in both types of maps (mean and daily), the location of the temperature cells is almost identical. Mehrabad, Shemiran and Imam Khomeini stations had the highest number of temperature cells in both types of maps.

There are four pressure cells in the mean pressure map of the 15-year period (Fig. 10). Shemiran and Mehrabad cells have the highest Newtonian mass, the values of which are shown in Table 9. The distance between the two cells is 19.32 km. The pressure at the centers of Shemiran and Mehrabad cells is 845.16 and 881.10 hPa.

The following results are obtained by examining the map of the mean pressure of the critical days of pollution over 15 years, and also investigation of the selected daily pressure cell maps for the three target months.

a. The map of means shows the formation of four pressure cells including Shemiran, Mehrabad, Chitgar and Imam Khomeini cells, with the largest Newtonian mass belonging to Shemiran and Mehrabad station.
b. We obtained these results from the daily temperature cell maps for the three months of November, January, and December.

November has Chitgar and Imam Khomeini cells with the highest Newtonian mass, January has four cells of Imam Khomeini, Shemiran, Chitgar and Mehrabad, among which, Chitgar and Mehrabad cells have the highest Newtonian mass, and December has three pressure cells including Mehrabad, Geophysics and Shemiran, with the highest Newtonian mass in Geophysics and Mehrabad cells.

Accordingly, it can be concluded that in both types of maps (mean and daily) the location of the pressure cells is almost identical. Mehrabad, Chitgar and Imam Khomeini stations had the highest number of pressure cells in both types of maps.

The results of comparing the two maps:

The location of the temperature cells is the same in both types of maps (the mean temperature maps of the polluted days in the 15-year period and the maps of heat cells of the polluted days in the three months in question), however, there is no such similarity between the maps in terms of the location of cells with the highest Newtonian mass.

This is also true for pressure cells (similarity of the location of cells and non-similarity of the location of cells with the highest Newtonian mass).

6.3 Winds of Tehran

The direction of winds in Tehran province is closely related to the geographical location and its terrains. General air currents in the province are subject to western winds. The general direction of these currents is almost parallel to the Alborz Range, and therefore the impact of these mountains appears as the lower average wind speed in the Southern valleys and foothills. The advance of the southern slopes of the eastern Karaj Heights diverts the surface winds to some parts of the southern plains, such as Shahriar, causing a relative increase in wind speed in these areas. Also, part of the southern slopes of Alborz Range in the north of Tehran has projections.

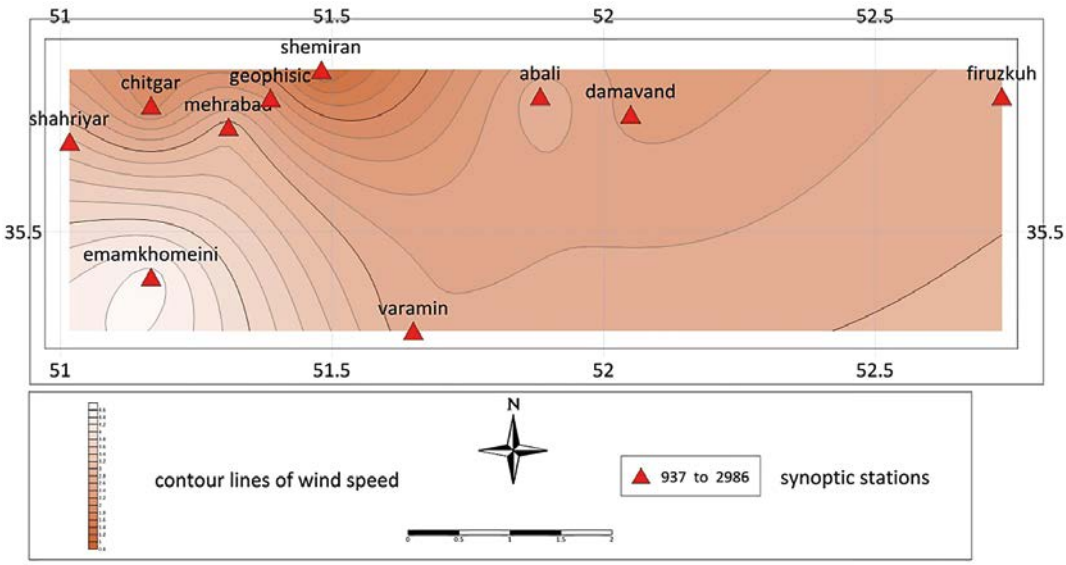


Fig. 11 Zoning map of the average wind speed of the polluted days over a 15-year period.

6.4 Wind direction

It can be said that 70% of Tehran's winds are weak (between 1 and 6 knots) and have no clear direction. Winds in Tehran are mainly from the west. As most of the winds in Tehran are weak , they do not have much effect on reducing air pollution.

In most months of the year, the wind direction is from west, with only two months from the southeast and two months from the north. Thus, for most months the western winds are highly volatile. The general direction of these currents is almost parallel to the Alborz Range; so, the effect of these mountains is mostly the reduction of the average wind speed in the Southern foothills valleys. Overall, Tehran's air is more stable than the surrounding areas and has little wind.

In the following, the zoning map of the average wind speed of the polluted days over the 15-year period is drawn to compare the wind speed at the stations in the study area.

In Fig. 11, the map of the average wind speed on polluted days during the 15-year period, we observe six wind cells that we can call wind holes.

Imam Khomeini cell has the highest average wind speed, followed by Abali, Mehrabad, Chitgar, Damavand and Shamiran stations, with the latter having the lowest average wind speed.

6.5 Topography

With an area of 800 km^2, the city of Tehran is located on the southern slope of Alborz Range. The altitude of the city reaches 1200 m in the south at Mehrabad airport and 2000 m in the north. Although the general slope of the city is southward, there are also lots of terrains. Alborz Range limits the city on the north and BibiShahrbanoo Mountains on the east. But the south and west areas are not so high.

The arrangement of the mountain dams has also led to local winds in Tehran. At night, for example, the mountain breeze brings pollutants downtown and raises the intensity of pollution. During the day, the breeze of the plains and the southern winds carry the pollutants northward and the northern areas become polluted.

6.6 Effect of topographic terrain features on temperature and pressure cells

In order to investigate the effect of topographic terrains on temperature and pressure cells, and to further understand the location of these cells, the temperature and pressure cells were overlain on the topography of the area.

For this purpose, a 3D map of the area's heights was plotted, and the synoptic stations and pressure and temperature cells overlapped for analysis and investigation.

Based on the three-dimensional map, in Fig. 12 two temperature cells of Mehrabad and Imam Khomeini can be observed. Mehrabad Airport cell is located in the projected area of southern Alborz Mountains and the Imam Khomeini Airport cell is located near the low heights of southern Shahriar.

Mehr Abad cell with a height of 1190.2 m and pressure 884.1375 and Imam Khomeini airport cell with a height of 990.2 m and a pressure of 905.5875 have a height difference of 200 m and a pressure difference of 21.45 hPa. These two temperature cells with the highest Newtonian mass have been identified in critical days of contamination for 15 years.

In Fig. 13 the overlapping the temperature cells of January with the topography of the study area, there can be seen three cells of Shemiran, Geophysics and Mehrabad. The two formed cells are located in the recesses of the southern slope of the Alborz Range, and it may be noted that the confinement of cell

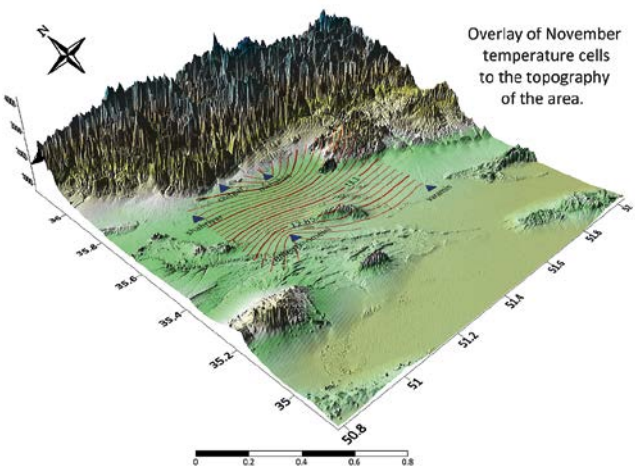


Fig. 12 Overlay of November temperature cells to the topography.

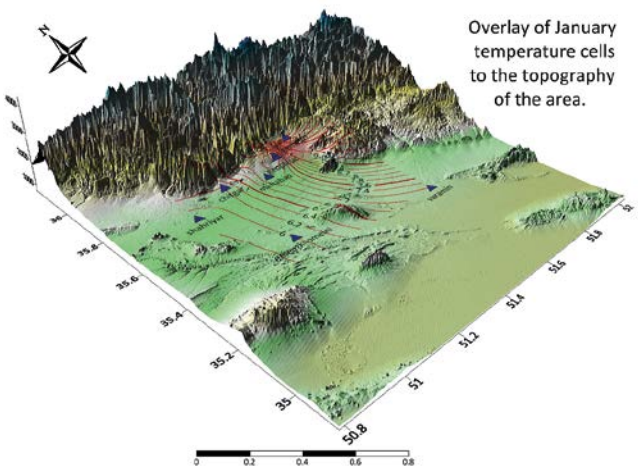


Fig. 13 Overlay of January temperature cells to the topography.

formation zones may influence the formation of these temperature cells.

Meanwhile, Geophysics and Shemiran cells have a higher Newtonian temperature than other cells. Geophysics cell has an altitude of 1423.8 m and a pressure of 854.1. The two cells, with a height difference of 124.4 m and a pressure difference of 9.8, are located on a sloping terrain at the lower end of the Alborz Range.

Based on Fig. 14, five temperature cells of Varamin, Imam Khomeini, Chitgar, Geophysics and Shemiran, are located on Tehran topography. Among these cells, Geophysics and Shemiran cells have the highest Newtonian mass.

These two cells are located in the recesses of the southern slope of the Alborz Range, and the confinement of areas where the cells are located may influence the formation of these temperature cells.

The Geophysics cell with a height of 1423.8 m and pressure of 862.9 and Shemiran cell with a height of 1548.2 m and pressure of 850.2 have a height difference of 124.4 m and a pressure difference of 12.7 hPa.

The Fig. 15 shows two pressure cells at Chitgar and Imam Khomeini Airport Stations. The Chitgar cell lies on the southern slope of the Alborz Range, where the mountain have eaves Imam Khomeini Airport cell is located near the low heights of southern Shahriar. The formation of these pressure cells at the above-mentioned areas may be affected by the air currents of the area. Due to the, eaves of the southern slopes of the eastern mountain these currents divert the surface winds to the southern plains and increase the relative wind speed at these points.

Chitgar station with a height of 1305.2 meters and temperature of 11.55 °C and Imam Khomeini Airport station with a height of 990.2 meters and a temperature of 12.15 °C have a height difference of 315 m and a temperature difference of 0.6 °C.

In Fig. 16, the map of overlapping of pressure cells of January on the heights of the study area, we observe

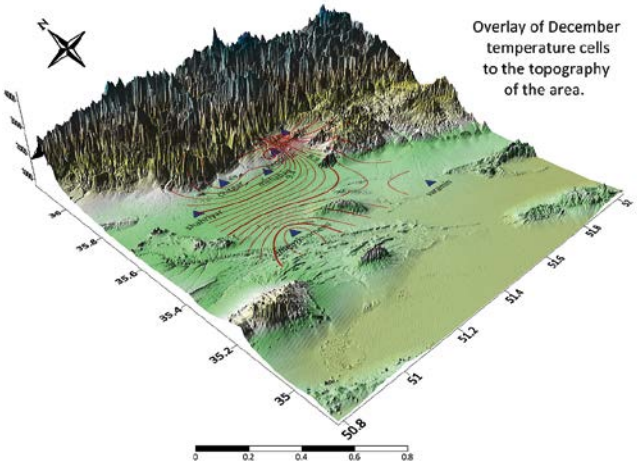


Fig. 14 Overlay of December temperature cells to the topography.

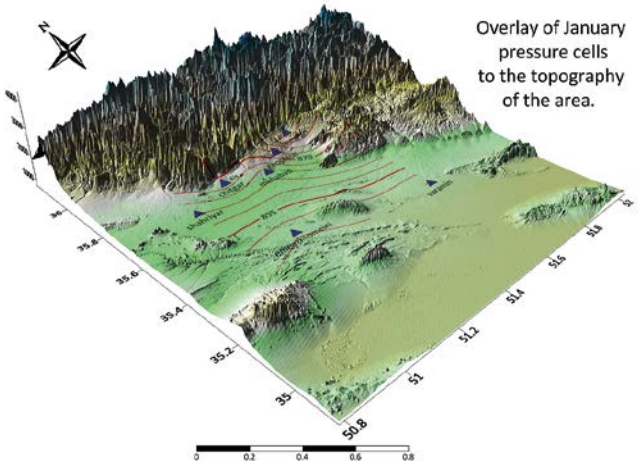


Fig. 16 Overlay of January pressure cells to the topography.

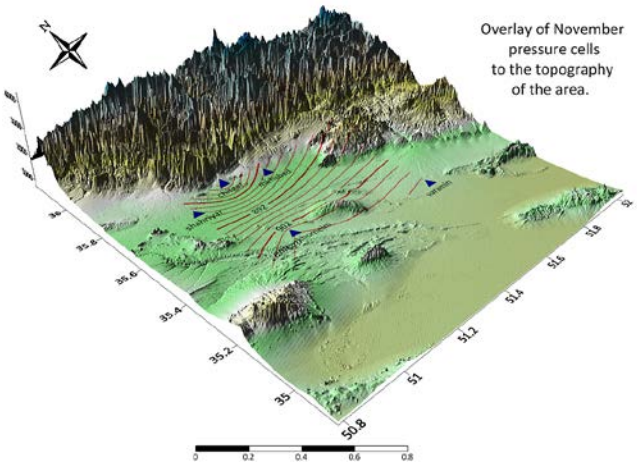


Fig. 15 Overlay of November pressure cells to the topography

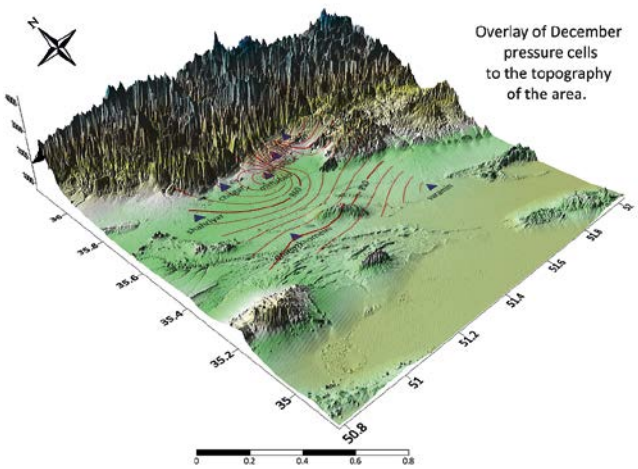


Fig. 17 Overlay of December pressure cells to the topography.

four cells of Shemiran, Mehrabad, Imam Khomeini and Chitgar, which are located at the desired altitudes.

In terms of pressure, the two cells of Mehrabad and Chitgar have the highest Newtonian mass. The two cells are located in the projections of the southern Alborz Range. Chitgar cell with height of 1305.2 m and temperature of 11.05 °C have a height difference of 115 m and a temperature difference of 0.05 °C.

Based on Fig. 17, the three pressure cells of Mehrabad, Geophysics and Shemiran are located on Tehran terrain, among which Mehrabad and Geophysics have the highest Newtonian mass in terms of pressure.

These two cells are located on the northern heights of Tehran. In fact, this part of the southern slope of Alborz Range has a recession that can affect the wind directions and also the formed cells. The Geophysics cell with a height of 1423.8 m and temperature of 3.8 °C and Mehrabad cell with a height of 11190.2 m and a temperature of 3.7 °C have a height difference of 233.6 m and a temperature difference of 0.1 °C.

7. Conclusions

Tehran's air pollution is considered as one of most important environmental problems and so far numerous studies have been done to mitigate this problem. Among the factors that cause air pollution are human factors as well as geographical ones and conditions such as Tehran's topography and the occurrence of inversions, especially during the cold months of the year.

In this study, with emphasis on Tehran geographical factors, we attempted to determine the critical days of air pollution using the data of air pollution measuring stations and synoptic stations in Tehran province from 2003 to 2017. Then after investigating and determining the level of inversion, we studied probability of the existence of closed temperature and pressure cells in these days. Then, by calculating and measuring the cellular properties as well as the existing conditions, the theoretical possibility of their turbulence was studied in order to cause turbulence in Tehran air to reduce air pollution. The results of the determination of the level of inversion showed that AQI > 150 in the target months did not cover all areas of Tehran. This means that while some areas of Tehran are more polluted than others, exceptions to this rule are excluded.

Generally considering the formed cells by the temperature and pressure difference and the gradient between them as well as the difference in height between the cells and their location and pointing out that the local winds cause the difference of temperature and pressure, it is more possible to elucidate conditions for creating air turbulence in Tehran within the study area to control the contamination amount. Results of this study provide new knowledge of the natural conditions in the study area, and there is no uniform pattern for all areas in the subject area. Since this study was conducted only for a limited period of 15 years (from 2003 to 2017) in Tehran province, and also all analyses were performed on the basis of statistics measured in synoptic stations in this area, the results should be interpreted with caution. However, it can be emphasized that all reviews and results are based on this range and data and cannot be over-generalized.

Acknowledgements

This article has done by funding of Iran National Science Foundation

References

Abbey, D. E., Nishino, N., McDonnell, W. F., Burchette, R. J., Knutsen, S. F., Beeson, W. L., Yang, J. X. (1999): Long-term inhalable particles and other air pollutants related tomortality in nonsmokers. American Journal of Respiratory and Critical Care Medicine 159(2), 373–382, https://doi.org/10.1164/ajrccm.159.2.9806020.

Beelen, R., Hoek, G., van der Brandt, P., Goldbohm, R. A., Fisher, P., Shouten, L. J., Jerret, M., Hughes, E., Armstrong, B., Brunekreef, B. (2008): Long-term effects of traffic-related air pollution on mortality in a Dutch cohort (NLCS-Air Study). Environmental Health Perspectives 116(2), 196–202, https://doi.org/10.1289/ehp.10767.

Bidokhty, A. A., Tajedin, B. H. (1998): Urban mix layer and air pollution. Geography Research Quartery 20(3), 41–62.

Bityukova, V. R., Kasimov, N. S. (2012): Atmospheric pollution of Russia's cities: Assessment of emissions and immissions based on statistical data. Geofizika 29(1), 53–67, https://hrcak.srce.hr/84995.

Buchholz, S., Junk, J., Krein, A., Heinemann, G., Hoffmann, L. (2010): Air pollution characteristics associated with meso scale atmospheric patterns in northwest continental Europe. Atmospheric Environment 44(39), 5183–5190, https://doi.org/10.1016/j.atmosenv.2010.08.053.

Cheragi, M. (1999: Study and comparison air quality between Tehran and Isfahan in 1998 and presentation solution for improvement. MA thesis Environment. Faculty of Natural Resources, University of Tehran.

Crosignani, P., Borgini, A., Cadum, E., Mirabelli, D., Porro, E. (2002): New directions: air pollution – how many victims? Atmospheric Environment 36(29), 4705–4706, https://doi.org/10.1016/S1352-2310(02)00381-3.

Dockery, D. W., Pope III, C. A., Xu, X., Spengler, J. D., Ware, J. H., Fay, M. E., Ferris, B. G., Speizer, F. E. (1993): An association between Air Pollution and Mortality in Six U.S. Cities. New England Journal of Medicine 329, 1753–1759, https://doi.org/10.1056/NEJM199312093292401.

Esa, L., Ali, A., Shah, M., Behzad, B., Siravan, A. M. (2012): Special vulnerability and air pollution disaster in Tehran. Third National Urban Development Conference, Islamic Azad University. Sanandaj.

Gwilliam, K., Kojima, M., Johnson, T. (2004): Reducing Air Pollution from Urban Transport. The International Bank for Reconstruction and Development/THE WORLD BANK, https://www.esmap.org/sites/default/files/esmap-files/urban%20pollution%20entire%20report.pdf.

Kejna, M. (2008): Topoclimatic conditions in the vicinity of the Arctowski Station (King George Island, Antarctica) during the summer season of 2006/2007. Polish Polar Research 29(2), 95–116, https://journals.pan.pl/Content/110568/PDF/PPR29-095.pdf.

Laden, F., Schwartz, J., Speizer, F. E., Dockery, D. W. (2006): Reduction in Fine Particulate Air Pollution and Mortality. Extended Follow-up of the Harvard Six Cities Study. American Journal of Respiratory and Critical Care Medicine 173(6), 667–672, https://doi.org/10.1164/rccm.200503-443OC.

Lashkari, H., Hedayat, P. (2007): Analysis model of synoptic extreme inversion in Tehran. Geography Research Quarterly 56(4), 65-82.

Mage, D., Ozolins, G., Peterson, P., Webster, A., Orthofer, R., Vandeweerd, V., Gwynne, M. (1996): Urban air pollution in megacities of the world. Atmospheric Environment 30(5), 681–686, https://doi.org/10.1016/1352-2310(95)00219-7.

Needleman, H. L., Schell, A., Beelinger, D., Leviton, A., Allred, E. N. (1990): The Long-Term Effects of Exposure to Low Doses of Lead in Childhood – An 11-Year Follow-up Report. New England Journal of Medicine 322, 83–88, https://doi.org/10.1056/NEJM199001113220203.

Patel, T. (1994): Killer smog stalks the Boulevards. New Scientist 1947.

Prüss-Ustün, A., Wolf, J., Corvalan, C., Bos, R., Neira, M. (2016): Preventing disease through healthy environments: a global assessment of the burden of disease from environmental risks. World Health Organization, https://www.who.int/publications/i/item/9789241565196.

Sánchez-Coyllo, O. R., de FátimaAndrade, M. (2002): The influence of meteorological conditions on the behavior of Siapaolo Brazil. Environmental Pollution 116(2), 257–263, https://doi.org/10.1016/s0269-7491(01)00129-4.

Saneie, R. (2016): Disaster management and planning of Tehran air pollution. PhD thesis. University of Isfahan.

Sarraf, M., Heger, M. (2018): Air Pollution in Tehran: Health Costs, Sources, and Policies. Environment and Natural Resources Global Practice Discussion Paper; No. 6. World Bank, Washington, DC, http://hdl.handle.net/10986/29909.

Shabazi, H., Ghaniazad, R., Hossieni, V., Hamedi, M. (2017): Investigation the influence of traffic emission reduction plans on Tehran air quality using WRF/CAMx modeling tools. Transportation Research Part D: Transport and Environment 57, 484–495, https://doi.org/10.1016/j.trd.2017.08.001.

Shahid, M. A. K., Hussain, K. (2013): Atmospheric air pollution and mapping of Faisalabad city using syntax map method. International Journal of Chemistry and Material Science 2(1), 001-009, http://www.academeresearchjournals.org/download.php?id=438603387359772169.pdf&type=application/pdf&op=1.

Shariepour, Z., Bidokhti, A. A. (2006): Investigating air pollution in Tehran and its relationship with meteorological characteristics. The first conference on air pollution and its effects on health, Clean Environment Studies Institute, Tehran.

Thadathil, P., Gosh, A. K. (1992): Surface layer temperature inversion in the Arabian Sea during winter. Journal of Oceanography 48, 293–304, https://doi.org/10.1007/BF02233989.

UNEP/WHO (1992): Urban air pollution in megacities of the world. Blackwell Reference, Oxford, https://iris.who.int/handle/10665/39902.

Vineis, P., Hoek, G., Krzyzanowski, M., Vigna-Taglianti, F., Veglia, F., Airoldi, L., Autrup, H., Dunning, A., Garte, S., Hainaut, P., Malaveille, C., Matullo, G., Overvad, K., Raaschou-Nielsen, O., Clavel-Chapelon, F., Linseisen, J., Boeing, H., Trichopoulou, A., Palli, D., Peluso, M., Krogh, V., Tumino, R., Panico, S., Bueno-De-Mesquita, H. B., Peeters, P. H., Lund, E. E., Gonzalez, C. A., Martinez, C., Dorronsoro, M., Barricarte, A., Cirera, L., Quiros, J. R., Berglund, G., Forsberg, B., Day, N. E., Key, T. J., Saracci, R., Kaaks, R., Riboli, E. (2006): Air pollution and risk of lung cancer in a prospective study in Europe. International Journal of Cancer 119(1), 169–174, https://doi.org/10.1002/ijc.21801.

Zahedi, M., Zolfagari, H., Torabi, S. (2000): Study effect of climatology factor on formation nucleus air pollution in northern of Tabriz. Second Zonal Meeting on Climate Change. Meteorological Organization, Tehran.

Assessment of forest cover and forest loss using satellite images in Thua Thien Hue province, Vietnam

Bui B. Thien[1,*], Vu T. Phuong[2,3], Akinola A. Komolafe[4]

[1] Institute of Earth Sciences, Southern Federal University, Russia
[2] Faculty of Social Sciences, Hong Duc University, Vietnam
[3] Innovation Startup Support Center, Hong Duc University, Vietnam
[4] Department of Remote Sensing and GIS, Federal University of Technology Akure, Nigeria
* Corresponding author: buibaothienha@gmail.com

ABSTRACT

Deforestation in the tropics continues inexorably with severe implications for biodiversity conservation, climate regulation and ecosystem services. This study investigated variation in forest cover in Thua Thien Hue province, Vietnam, using the Landsat Thematic Mapper and Operational Land Imager satellite images over the period 1989–2021. Imageries were classified using the maximum likelihood classification technique for the years 1989, 2006, and 2021 and were evaluated for accuracy using the kappa coefficient for each year. Furthermore, forest cover losses and gains were evaluated using the Normalized Difference Vegetation Index and Soil Adjusted Vegetation Index, which were compared with the output of the supervised classification. Results showed that the forest cover of Thua Thien Hue province has drastically declined over the years. The forest cover, which was estimated at 68.88% (3461.46 km^2) of the total land area in 1989, increased to 69.04% (3469.51 km^2) in 2006 and subsequently decreased to 57.55% (2891.81 km^2) in 2021. Severely reduced forest cover is often associated with the expansion of agriculture on the forest edge; other contributing factors include logging, illegal production land, and forest fires. Overall, our results show the necessity of forest management, rational land-use planning policy, and increased community awareness of conservation and sustainable development of forest resources in the study area in the future.

KEYWORDS

forest cover; vegetation index; remote sensing; GIS; Thua Thien Hue province

Received: 13 March 2023
Accepted: 27 September 2023
Published online: 1 November 2023

Thien, B. B., Phuong, V. T., Komolafe, A. A. (2023): Assessment of forest cover and forest loss using satellite images in Thua Thien Hue province, Vietnam. AUC Geographica 58(2), 172–186
https://doi.org/10.14712/23361980.2023.13
© 2023 The Author. This is an open-access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0).

1. Introduction

Particularly significant natural resources on the planet are forests. In order to effectively shield the riverbank and inland areas from natural disasters, it does create a sort of natural barrier (hurricanes, cyclones, and tsunamis) (Alongi 2008). It makes a substantial contribution to broad ecological protections such as soil preservation, improved biodiversity, and climate change mitigation. By providing supplies of medicine and materials for manufacturing and building, it effectively aids in the expansion of the wider national economy. The world's forest cover has seen substantial and unprecedented changes in recent decades as a result of a number of human factors, including increasing urbanisation, industrialization, agricultural expansion, logging, and mining (Atmiş et al. 2007; Agarwal et al. 2010; Hor et al. 2014; Ahammad et al. 2019). Logging operations, it was previously believed, had become the primary factor in the loss of forest cover in emerging countries with tropical climates. In these regions, logging, unemployment, migration rates, population pressure, infrastructure development, and agriculture are some of the main forces that alter forest cover (Ranjan 2019).

Vietnam, one of the nations with the almost largest sections of unquestionably main forests and annually changing plantations, is well known for its highly diverse and unique tropical forest ecology (Mermoz et al. 2021). Since 1990, Vietnam's forests have undergone the transition from net deforestation to reforestation (Meyfroidt and Lambin 2008). Although forest cover in Vietnam has increased, deforestation and forest degradation particularly continue (JICA, VNFOREST 2012). In an effort to prevent deforestation and forest degradation, Vietnam participates in the Reduction of Emissions from Deforestation and Forest Degradation (REDD+) (Pham et al. 2012). The total area of natural forests and planted forests with reserves in the entire Thua Thien Hue province is 2884.02 km^2, and the forest coverage rate is 57.38%, according to the results of the announcement of the forest status in Thua Thien Hue province in 2020 (People's Committee of Thua Thien Hue Province 2021). The management, preservation, prevention, and suppression of forest fires in Thua Thien Hue province are complicated by a sizable forest area that is dispersed throughout 9 districts and towns and stretches from the Truong Son mountain range to the East Sea. The complicated problem of illegal deforestation for agricultural land, timber, and other products, as well as the ongoing problem of forest fires brought on by the dry season's heat, the burning of votive paper, the burning of vegetation for ritual purposes, and other factors, have all contributed to the province's loss of forests.

The application of science and technology in detecting forest loss is essential to contribute to saving time and effort for the patrol and supervision of forest resources as well as improving the efficiency of management, conservation, forest protection, forest fire prevention and fighting. In Thua Thien Hue province, there is very little research related to the use of remote sensing and geographic information systems (GIS) to analyse the change in forest cover over the years (Dien 2004; Yen et al. 2005). However, these studies are conducted only in a small area and have not fully applied the vegetation index to analyse the change in forest cover more deeply.

Globally, there has been significant development in the use of remote sensing and GIS to track changes in forest areas and other types of land cover. Evidence is derived from remote sensor data using a variety of image analysis and change detection techniques (Lu et al. 2011; Chowdhury et al. 2020; Thien et al. 2022). Various sensors are available as a data source to study land use/land cover (LULC) variability. Among them, Landsat images provide an acceptable mean accuracy for analysis in the LULC study, which is generally accepted in the scientific community (Bakr et al. 2010). Therefore, remote sensing and GIS can help stakeholders identify the main areas where changes are occurring and obtain understanding of how various human behaviours and climate changes impact growth patterns. Consider your current actions and policies, foresee the future, and make the required preparations in light of the possibility that natural weather patterns and seasonal landscapes will trend through time (Mubako et al. 2018). Analysis of the Normalized Difference Vegetation Index (NDVI) and the Soil Adjusted Vegetation Index (SAVI) is used to map forest basically cover and for the most part is also useful for monitoring plant health (Huete 1988; Matsushita et al. 2007; Spadoni et al. 2020).

The study specifically focuses on Thua Thien Hue province in Vietnam. While there may have been previous studies on forest cover change at regional or national scales, this study narrows down the scope to a specific province (Jadin et al. 2013; Tran et al. 2022). By doing so, it provides detailed insights into the dynamics of forest cover change in this particular area, which may differ from broader-scale patterns observed in previous studies. In this study, our particularly goal was to generally integrate open-access remote sensing data and GIS to examine the spatial trends of forest and other land-cover changes in Thua Thien Hue province, Vietnam, over 32 years (1989–2021). By utilizing advanced remote sensing techniques, the study can provide new insights and potentially more accurate assessments of forest cover change in the study area. During the period from 1989 to 2021, many important changes in forest policy and management occurred in Vietnam and Thua Thien Hue province (Pham et al. 2021). In addition, this period has also witnessed significant environmental changes and human impacts on the forest cover in the study area. Factors such as agricultural expansion, urban development, logging, and climate

change may have influenced the appearance of forest cover. Research during this period will help to better understand the forest management measures already in place, such as forest protection policies, sustainable forest development and land use planning. This study provides basic information to support land-use planning to basically avoid degradation and deforestation in the study area in a big way. The main objectives of this study mostly are defined as follows: (1) detecting forest cover in a typical Vietnamese area of Thua Thien Hue province from 1989 to 2021, (2) investigating detailed spatial and temporal variations in forest cover and other major cover types, and (3) linking vegetation indices with changes in forest cover.

2. Materials and Methods

2.1 Study area

Thua Thien Hue province is located in the coastal strip of central Vietnam, geographical position of the study area lies in the coordinates of latitude 15°59′30″N–16°44′30″N and longitude 107°00′56″E–108°12′57″E, covering an area of 5025.30 km^2 (Fig. 1) (Thua Thien Hue General Statistical Office 2019). The study area is 675 km south of Hanoi capital, 94 km north of Da Nang city with the natural boundary of Bach Ma mountain range. The province of Thua Thien Hue features a variety of geography, including plains, mountains, hills, and coastal regions. One-fourth of the province's land area, or 250 to 1800 m, is made up of mountainous terrain, which is distributed in the west of the region. Half of the area is covered by hills with a height of 10 to 250 m. With a total size of around 1400 km^2, the plain was an abrasive, agglomerated, sandy, and lagoon plain (People's Committee of Thua Thien Hue Province 2005). The study area is located in tropical monsoon climate, so the weather occurs in a four-season cycle. The average temperature of the whole year is 25 °C, the number of sunny hours a year is 2000 hours. In terms of economic scale, the gross domestic product per person in 2018 was $1793. Agriculture, forestry, and fishery made up 11.60% of the economy, followed by the service sector at 31.20%, the service sector at 50.20%, and product taxes minus product subsidies at 7.00% (Thua Thien Hue General Statistical Office 2019).

2.2 Data used and sources

Satellite images were used to map LULC in Thua Thien Hue province from 1989 to 2021 and assess changes in forest cover. Landsat 5 TM images were used for the years 1989 and 2006 and Landsat 8 OLI/TIRS images were used for 2021 in this study. The Landsat image dataset was downloaded from the USGS EarthExplorer (https://earthexplorer.usgs.gov) and USGS Glovis websites (https://glovis.usgs.gov). A detailed data summary is given in Tab. 1.

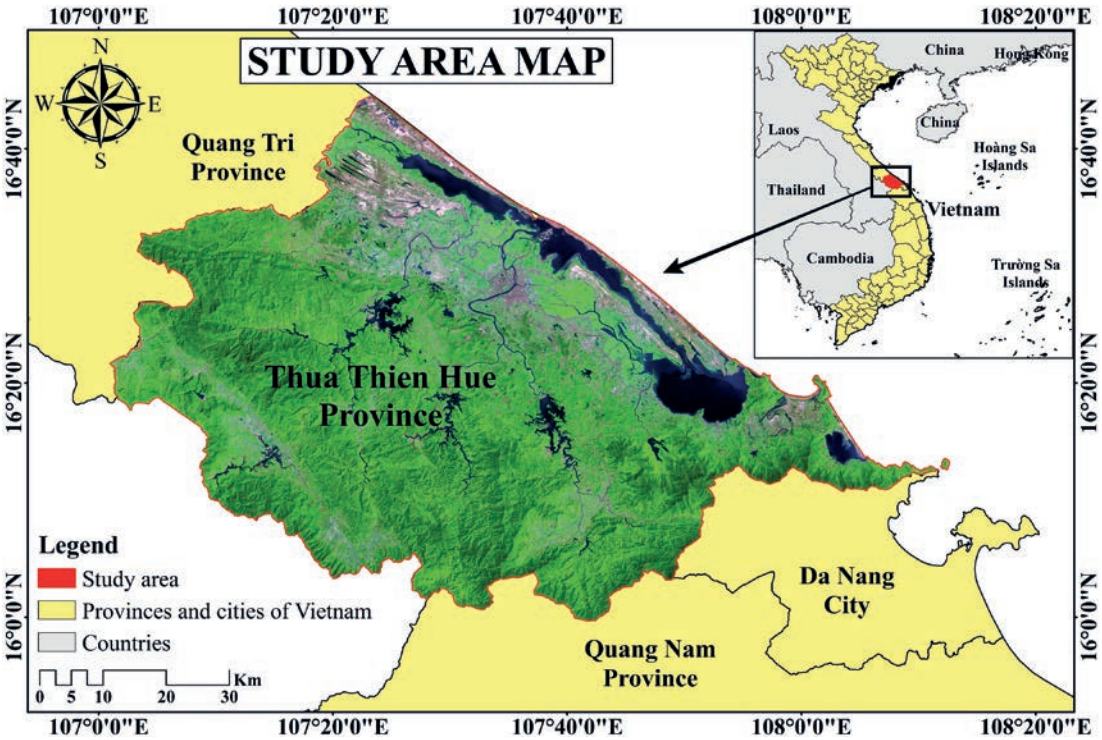


Fig. 1 Map showing the study area in Thua Thien Hue province.

Tab. 1 Detailed data summary of satellite imagery used in the study.

Landsat Scene ID	Acquisition data	Satellite	Path/row	Resolution (m)	Source
LT51240491989009BKT00	09/01/1989	Landsat 5 TM	124/049	30	USGS Glovis
LT51250481989048BKT01	17/02/1989	Landsat 5 TM	125/048	30	USGS Glovis
LT51250491989048BKT01	17/02/1989	Landsat 5 TM	125/049	30	USGS Glovis
LT51250482006095BKT00	05/04/2006	Landsat 5 TM	125/048	30	USGS Glovis
LT51250492006095BKT00	05/04/2006	Landsat 5 TM	125/049	30	USGS Glovis
LC81250492021056LGN00	25/02/2021	Landsat 8 OLI	125/049	30	USGS EarthExplorer

2.3 Image pre-processing and classification

Based on Google Earth Pro satellite image data, modified Anderson LULC scheme level I (Anderson et al. 1976), Vietnam's regulation on LU, the existing condition of the study area, specific pixel values of different landscape elements and reference to related documents, five LULC classes were mapped using Landsat satellite image classification as agricultural land, barren land, forest, settlement, and water (Tab. 2).

In order to meet the main objective of this study which was to investigate the variation of the forest cover, detailed LULC classes of Thua Thien Hue province, Vietnam were considered from 1989 to 2021; this enabled comparison between the change in forest cover and other classes. Fig. 2 illustrates the primary steps taken during the LULC changes analysis and detection of forest cover change.

Image processing tasks were performed using ArcGIS 10.8 software. Atmospheric correction (dark-object subtraction (DOS) correction) was applied to remove the dust and haze effect from each image. The band set was defined (combination RGB colour (Red, Green, Blue) corresponds to bands 5, 4, 3 for Landsat 5 TM and bands 6, 5, 4 for Landsat 8 OLI), and the training input file was created. The choice of these specific bands for the RGB colour representation is based on their spectral characteristics and their ability to capture different aspects of the landscape, including vegetation health and density. The combination of these bands allows for visual interpretation and analysis of forest cover and forest loss within the study area (Dimyati et al. 2018; Osio et al. 2018). Training samples were selected for each predefined LULC class (Tab. 2) by delineating polygons around representative locations. We used the pixels enclosed by these polygons to generate the spectral signatures for the respective land coverings recorded by satellite images over time. We employed the spectral signature in the classification process using the maximum likelihood method when it was deemed to be satisfactory. To classify the research area's land cover and assess land use, supervised classification was used (Manandhar et al. 2009, Forget et al. 2018). This is a type of image classification approach that is mainly controlled by the analyst, who selects pixels that represent the desired classes (Butt et al. 2015; Shivakumar and Rajashekararadhya 2018). To improve classification accuracy and reduction of misclassifications, post-classification refinement was therefore used for simplicity and effectiveness of the method (Harris and Ventura 1995). The selected LULC categorization features unambiguous class borders based on changes in natural and anthropogenic elements within the examined area, consistency in the description of each category, and consistency across categories. This classification strategy is also scale-independent, making it practicable to use it at any level of spatial detail or size. The LULC differences between the categorised photos were found utilising a post-classification comparison method that used change detection comparison (pixel by pixel) to measure the change in forest cover.

Tab. 2 Identified classes by supervised classification.

Class	Description
Agriculture	Cultivated outfields, homestead garden fields, aquaculture, salt field and small scattered plots of grazing land
Barren land	Fallow land, sands, earth dumps
Forest	Forestry, natural forests, individual trees
Settlement	Residential buildings, industrial use, roads, villages, and other impervious surfaces
Water	Rivers, canals, lakes, artificial ponds

2.4 Analysis of vegetation indices to detect changes in forest cover

Several vegetation indices have been developed to assess forest vegetation and are used in forest cover detection and analysis (Huete 1988; Matsushita et al. 2007; Huete 2012; Liang et al. 2018; Spadoni et al. 2020; Pesaresi et al. 2020). In this study, the Normalized Difference Vegetation Index (NDVI) and the Soil Adjusted Vegetation Index (SAVI) were used to detect forest cover in Thua Thien Hue province. NDVI is widely used to analyse vegetation health and has been shown to be a relevant indicator of plant greenness. SAVI is often used in arid and semi-arid regions to help reduce the influence of soil reflectance. In this case, we extracted the NDVI and SAVI values from

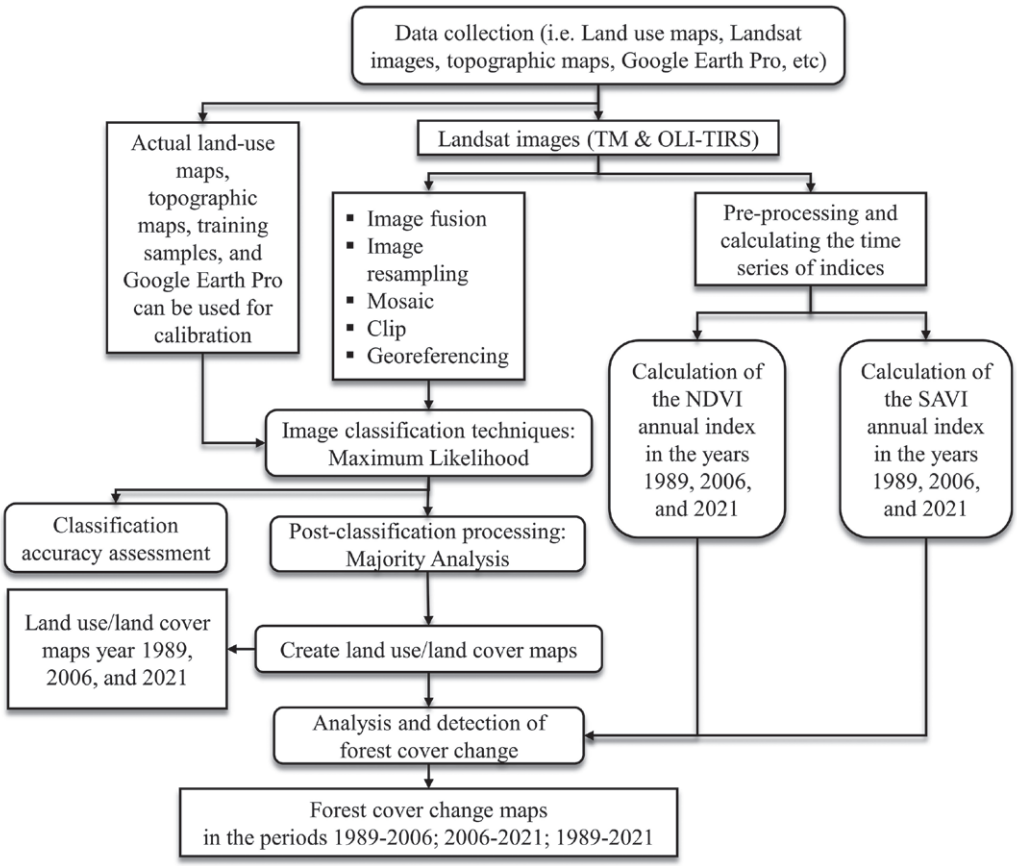


Fig. 2 Overall process of land use/land cover change technique.

all the forest polygons that we detected through the supervised classification of the Landsat 8 OLI images, then used the maximum and minimum values of NDVI and SAVI to reclassify the classification table to map the area of forest land cover for all other periods. NDVI and SAVI were calculated using equations (1) and (2), respectively, given below:

$$\mathrm{NDVI} = \frac{NIR - RED}{NIR + RED} \tag{1}$$

where NIR is the reflectance radiated in the near-infrared wave band, and RED is the reflectance radiated in the visible red wave band of the satellite radiometer.

$$\mathrm{SAVI} = \left(\frac{NIR - RED}{NIR + RED + 1}\right) \times (1 + \mathrm{L}) \tag{2}$$

where L is 0.5, the default value.

2.5 Classification accuracy assessment

We assessed the accuracy of the thematic maps produced to determine the quality of the information derived from the data and to reflect the actual discrepancy between our classification and the map or reference data (Owojori et al. 2005, March; Disperati et al. 2015; Tsutsumida et al. 2015). The accuracy of the classified images for the years 1989, 2006, and 2021 was assessed using error matrices, overall accuracy, producer accuracy, user accuracy, and kappa coefficients. With the help of a stratified random technique, we chose 150 random points from each image that was classed, and we digitally compared them with the corresponding pixels of the original images in Google Earth Pro as reference data. These spots, which represent all of the LULC categories in the research area, were found and located using Google Earth Pro maps, ground truth information, and topographic maps. By creating an error categorization matrix for each LULC map, the overall accuracy was evaluated. The equations for the kappa coefficient, overall accuracy, user accuracy, and producer accuracy shown in equations (3), (4), (5), and (6), respectively are among the best quantitative measurements for classifying satellite images (Chowdhury et al. 2020; Hasan et al. 2020; Thakur et al. 2021):

$$\mathrm{Kappa} = \frac{\Sigma_{i=1}^{k} n_{ii} - \Sigma_{i=1}^{k} n_{ii}\,(G_i C_i)}{\mathrm{n}^2 - \Sigma_{i=1}^{k} n_{ii}\,(G_i C_i)} \tag{3}$$

where, i is the class number, n is the total number of classified pixels that are being compared to actual data, n_{ii} is the number of pixels belonging to the actual data class i, that were classified with a class i, C_i is the

total number of classified pixels belonging to class i and G_i is the total number of actual data pixels belonging to class i. Number of correctly classified pixels in each category / Total number of reference pixels in each category (row total)

$$\text{Overall Accuracy} = \frac{\text{Number of correctly classified pixels (diagonal)}}{\text{Total number of reference pixels}} \times 100 \quad (4)$$

$$\text{User's Accuracy} = \frac{\text{Number of correctly classified pixels in each category}}{\text{Total number of reference pixels in each category (row total)}} \times 100 \quad (5)$$

$$\text{Producer's Accuracy} = \frac{\text{Number of correctly classified pixels in each category}}{\text{Total number of reference pixels in each category (column total)}} \times 100 \quad (6)$$

3. Results

3.1 Land use/land cover classification and accuracy

The LULC classification map of Thua Thien Hue province for the years 1989, 2006, and 2021 are presented in Fig. 3. In 1989, the majority of the study area was covered by forest, accounting for over 68% of the total area. Forest covered an area of 3461.46 km^2, followed by agriculture at 14.34% (720.75 km^2), barren land at 11.17% (561.22 km^2), water at 5.03% (252.77 km^2), and finally, settlement occupying only 0.58% (29.10 km^2) of the total study area (Tab. 3). By 2006, the forest area decreased to 3469.51 km^2, representing 69.04% of the total study area, and further declined to 2891.81 km^2 (57.55%) by 2021. The agricultural class increased to 980.33 km^2 (19.51%) in 2006 and continued to expand to 1296.62 km^2 (25.80%) in 2021 (Tab. 3). The area occupied by settlement and water also steadily increased to 182.46 km^2 (3.63%) and 260.09 km^2 (5.18%), respectively, in 2006. By 2021, these two classes had further expanded to 465.15 km^2 (9.26%) and 295.74 km^2 (5.89%), respectively. However, the barren land class continuously decreased to 132.91 km^2 (2.64%) in 2006 and 75.98 km^2 (1.51%) in 2021 (Tab. 3).

The accuracy rating reflects the actual difference between our classifier and the reference data. The overall accuracy scores for the three years 1989, 2006, and 2021 were 92.00%, 91.33%, and 92.72%, respectively, with kappa coefficients of 0.888, 0.885, and 0.903 (Tab. 4). Kappa values greater than 0.8

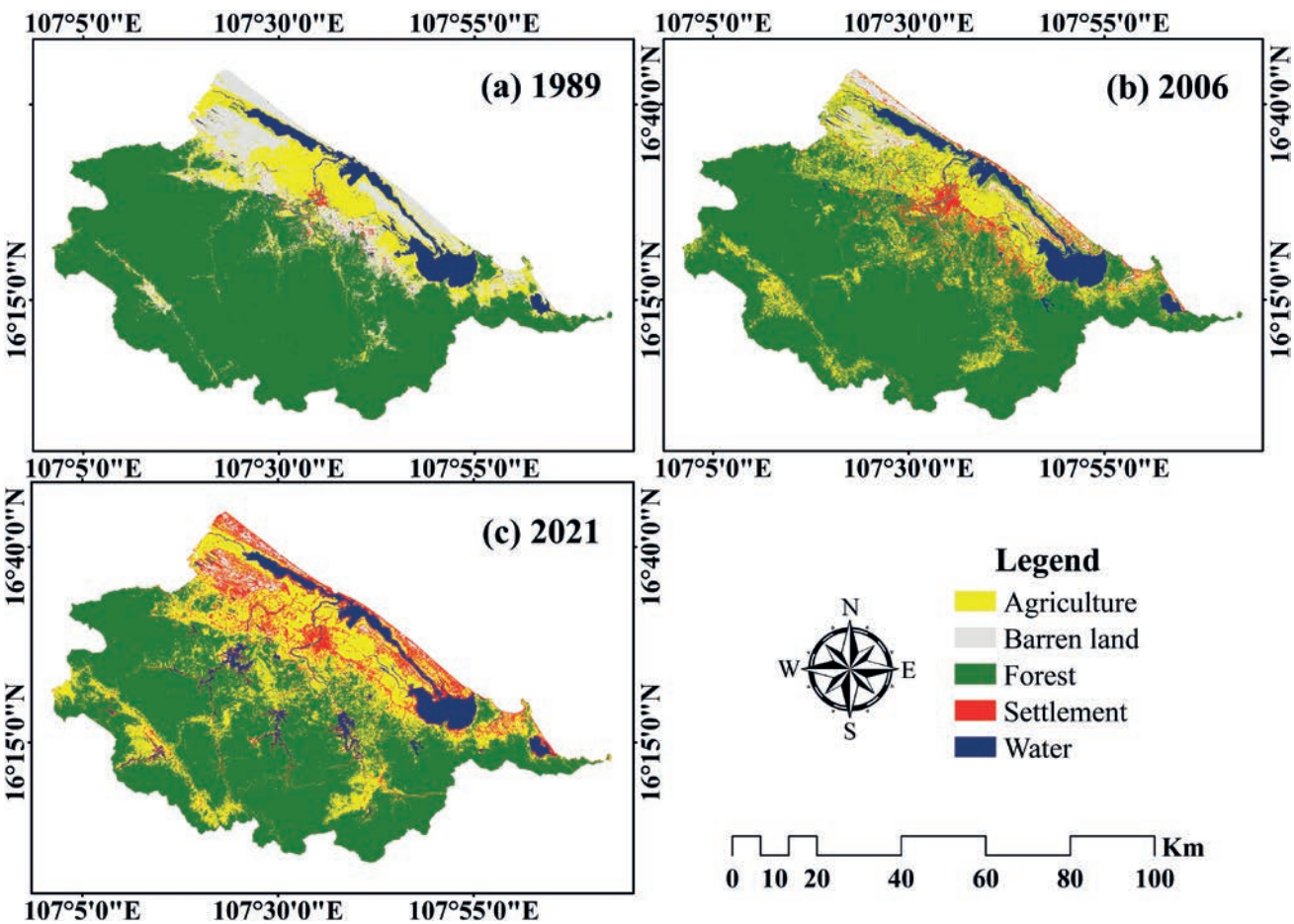


Fig. 3 Land use/land cover of Thua Thien Hue province in (a) 1989, (b) 2006, and (c) 2021.

Tab. 3 Results of land use/land cover classification in Thua Thien Hue province from 1989 to 2021.

Class	Land cover in 1989		Land cover in 2006		Land cover in 2021	
	Area (km²)	%	Area (km²)	%	Area (km²)	%
Agriculture	720.75	14.34	980.33	19.51	1296.62	25.80
Barren land	561.22	11.17	132.91	2.64	75.98	1.51
Forest	3461.46	68.88	3469.51	69.04	2891.81	57.55
Settlement	29.10	0.58	182.46	3.63	465.15	9.26
Water	252.77	5.03	260.09	5.18	295.74	5.89
Total	**5025.30**	**100.00**	**5025.30**	**100.00**	**5025.30**	**100.00**

Tab. 4 Accuracy assessments for 1989, 2006 and 2021.

Land cover class	1989		2006		2021	
	Producers accuracy (%)	Users accuracy (%)	Producers accuracy (%)	Users accuracy (%)	Producers accuracy (%)	Users accuracy (%)
Agriculture	85.71	90.91	87.50	90.32	88.37	92.68
Barren land	88.24	88.24	82.61	90.48	87.50	87.50
Forest	96.97	95.52	94.83	94.83	94.23	96.08
Settlement	88.89	80.00	94.44	89.47	100.00	90.32
Water	91.30	91.30	94.74	85.71	90.00	90.00
Overall accuracy	92.00		91.33		92.72	
Kappa coefficient	0.888		0.885		0.903	

indicate strong agreement between our classification and the reference data (Lea et al. 2010). The user accuracy results show that, for the year 1989, the maximum accuracy was achieved for the forest class (95.52%) and the minimum for the settlement class (80.00%). For 2006, user accuracy ranged from a minimum of 85.71% (water class) to a relatively precise classification of 94.83% (forest class). For 2021, user accuracy was 96.08% for the forest classes and was lowest for the barren land class (87.50%). The results of the producer accuracy assessment show that classification was relatively accurate for the forest class in 1989 and 2006 (96.97% and 94.83%, respectively) (Tab. 4). In 2021, classes achieved 100% accuracy (settlement land class). The lowest accuracy ratings in 1989, 2006, and 2021 were for agriculture land class (85.71%), barren land class (82.61%), and barren land class (87.50%), respectively (Tab. 4).

3.2 Land use/land cover change from 1989 to 2021

The area statistics and model on land-cover changes in Thua Thien Hue province in each period (1989–2006, 2006–2021, and 1989–2021) are shown in Tab. 5 and Fig. 4. During the period 1989–2006, the area of barren land decreased the most with 8.52% (428.31 km²), while agricultural land increased the most with 5.17% (259.58 km²). In addition, the forest, settlement and water classes also increased by 0.16% (8.05 km²), 3.05% (153.36 km²) and 0.15% (7.32 km²). By the period 2006–2021, the area of agricultural land and settlement classes has increased continuously by 6.29% (316.29 km²) and 5.63% (282.69 km²), respectively. The water class area during this period also increased by 0.71% (35.65 km²). As for the area of barren land class has continued to decrease by 1.13% (56.93 km²). The area of forest

Tab. 5 Change statistics of land use/land cover in Thua Thien Hue province from 1989 to 2021.

Class	Period					
	1989–2006		2006–2021		1989–2021	
	Area (km²)	%	Area (km²)	%	Area (km²)	%
Agriculture	259.58	5.17	316.29	6.29	575.87	11.46
Barren land	−428.31	8.52	−56.93	1.13	−485.24	9.66
Forest	8.05	0.16	−577.70	11.50	−569.65	11.34
Settlement	153.36	3.05	282.69	5.63	436.05	8.68
Water	7.32	0.15	35.65	0.71	42.97	0.86

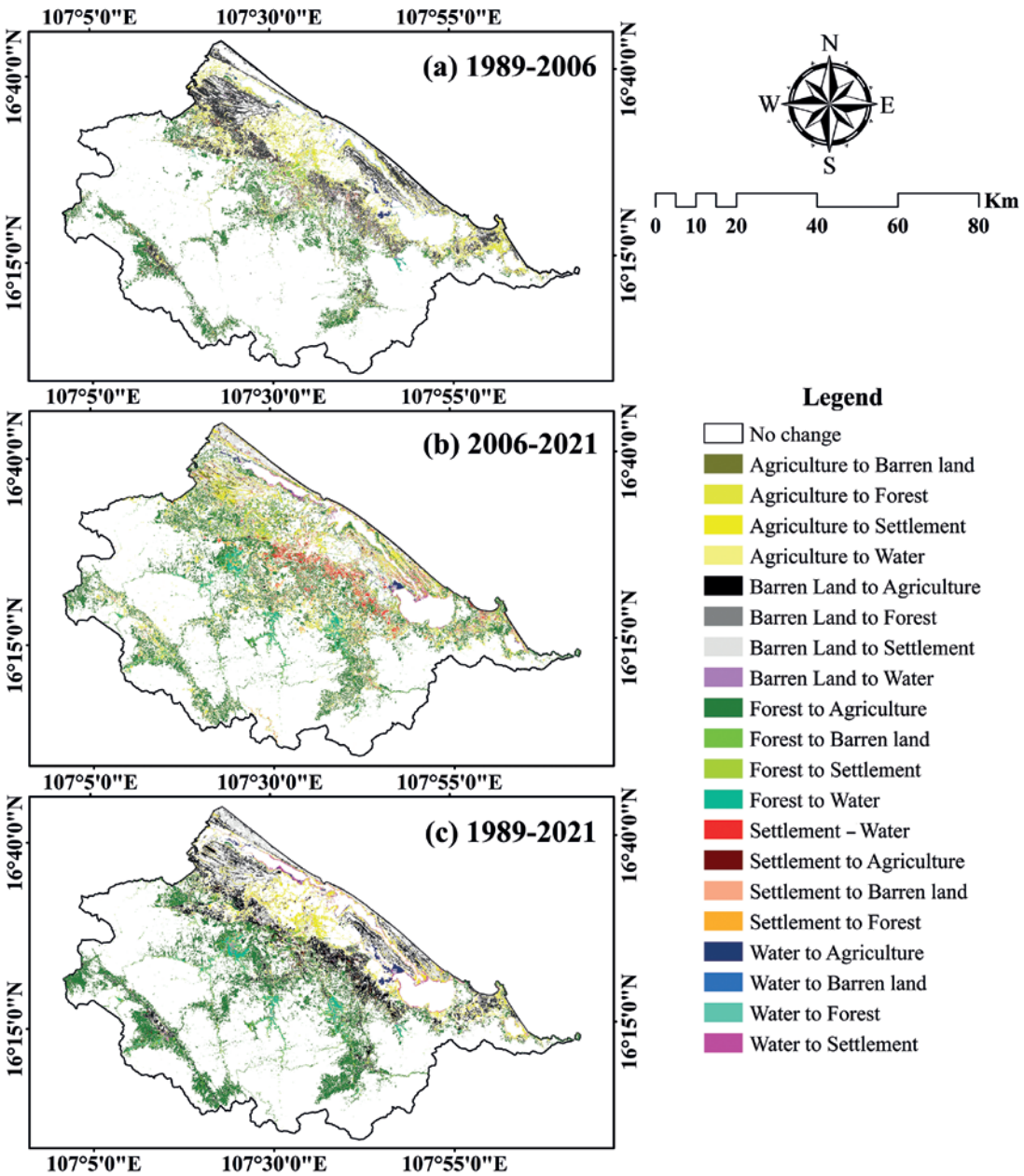


Fig. 4 Land use/land cover changes maps in Thua Thien Hue province (a) 1989–2006, (b) 2006–2021, and (c) 1989–2021.

class in the period 2006–2021 has tended to decrease sharply compared to the period 1989–2006, with the area reduced by 11.50% (577.70 km^2). Overall, over the past 32 years, agricultural land, settlement and water classes have increased continuously by 11.46% (575.87 km^2), 8.68% (436.05 km^2) and 0.86% (42.97 km^2), respectively. For the barren land class in this period (1989–2021) has decreased continuously by 9.66% (485.24 km^2). Particularly for the forest class, in the first period (1989–2006) there was increased but still very low compared to the area lost in the second period (2006–2021), so in general, the forest area in over 32 years has decreased by 11.34% (569.65 km^2).

3.3 Relationship between vegetation indices and decadal forest cover changes

Maps showing the NDVI and SAVI in Thua Thien Hue province from 1989 to 2021 are shown in Fig. 5 and Fig. 6, respectively. During this process, we looked at all NDVI and SAVI pixel values from our graded image in 2021 and observed that NDVI values greater than 0.26 and SAVI values greater than 0.41 all belong to the forest polygons. Considering these NDVI and SAVI values, we reclassified all other classified images for 1989, and 2006 as forest and non-forest. The area covered by forest from 1989 to 2021 according to the vegetation indices is presented in Tab. 6.

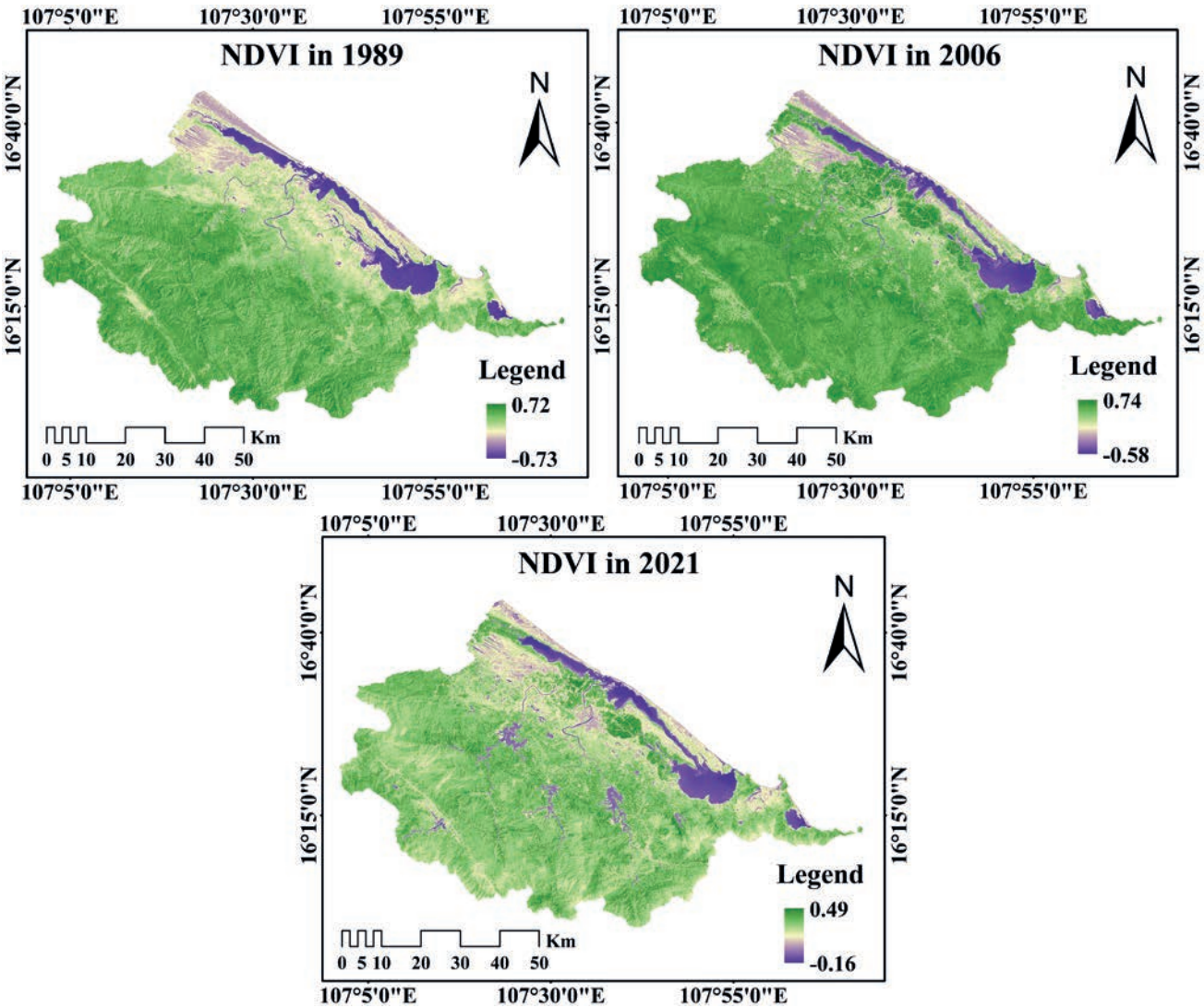


Fig. 5 Spatial distribution of NDVI for 1989, 2006 and 2021.

Forest cover from 1989 to 2021 based on these vegetation indices is presented in Tab. 6. The trend of year-over-year changes during the study period is also shown in Fig. 7. The results of the analysis of forest class reclassification based on the NDVI index through each of the years 1989, 2006 and 2021 are 3527.49 km^2, 3580.58 km^2 and 2905.27 km^2 respectively, corresponding to 70.19%, 71.25% and 57.81% (Tab. 6). In addition, based on the SAVI index, the forest classification results are quite similar to the NDVI index with 3501.11 km^2 (69.67%) in 1989, 3295.26 km^2 (65.57%) in 2006 and 2619.65 km^2 (52.13%) in 2021 (Tab. 6).

3.4 Forest cover changes from 1989 to 2021

Forest cover patterns can be summarized in terms of net change and gross gains and losses, with a

Tab. 6 Based forest cover area analyzed by vegetation indices (NDVI and SAVI) from 1989 to 2021.

Category		Distribution in 1989		Distribution in 2006		Distribution in 2021	
		Area (km^2)	(%)	Area (km^2)	(%)	Area (km^2)	(%)
NDVI	Forest	3527.49	70.19	3580.58	71.25	2905.27	57.81
	Other	1497.81	29.81	1444.72	28.75	2120.03	42.19
	Total	5025.30	100.00	5025.30	100.00	5025.30	100.00
SAVI	Forest	3501.11	69.67	3295.26	65.57	2619.65	52.13
	Other	1524.19	30.33	1730.04	34.43	2405.65	47.87
	Total	5025.30	100.00	5025.30	100.00	5025.30	100.00

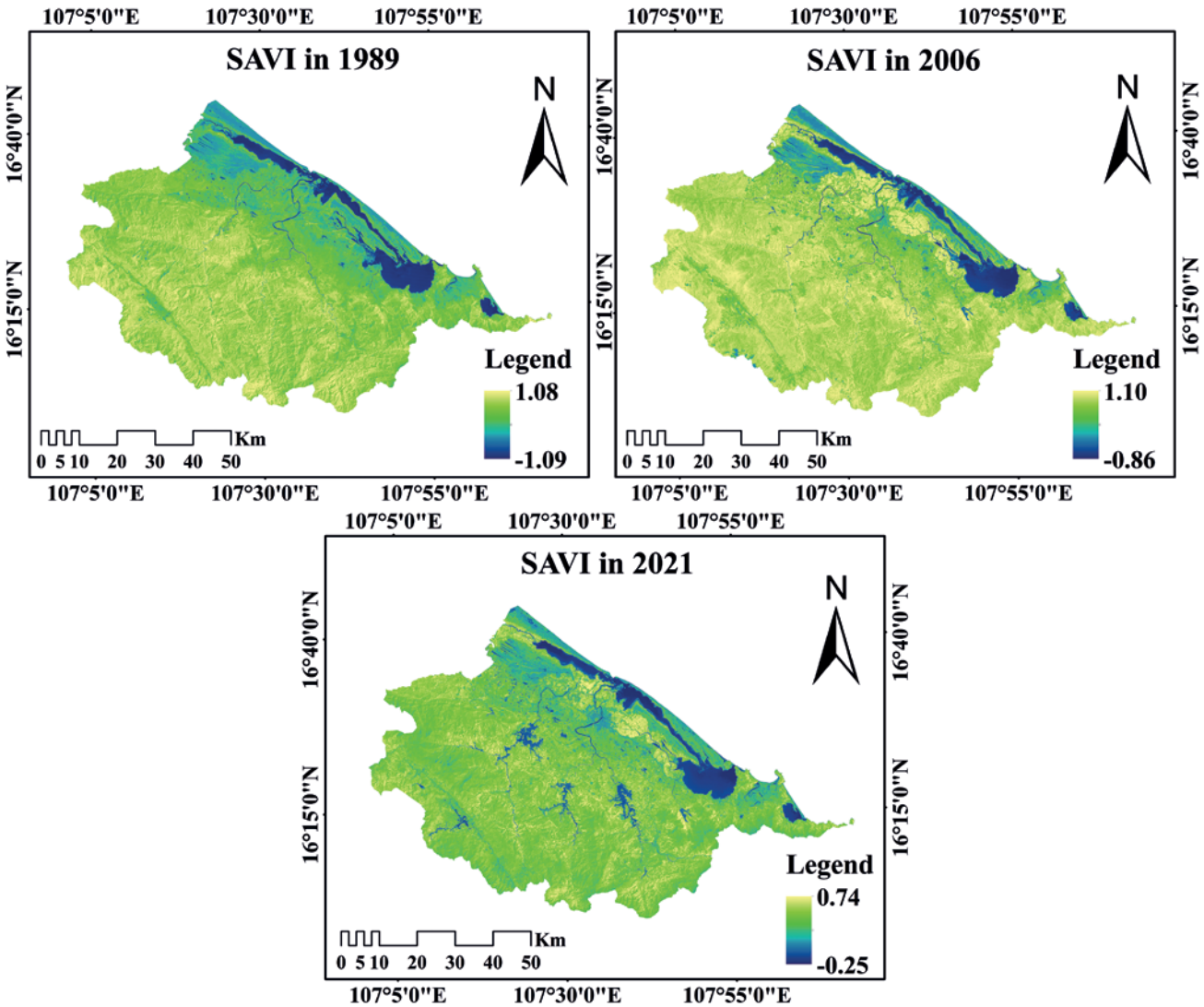


Fig. 6 Spatial distribution of SAVI for 1989, 2006 and 2021.

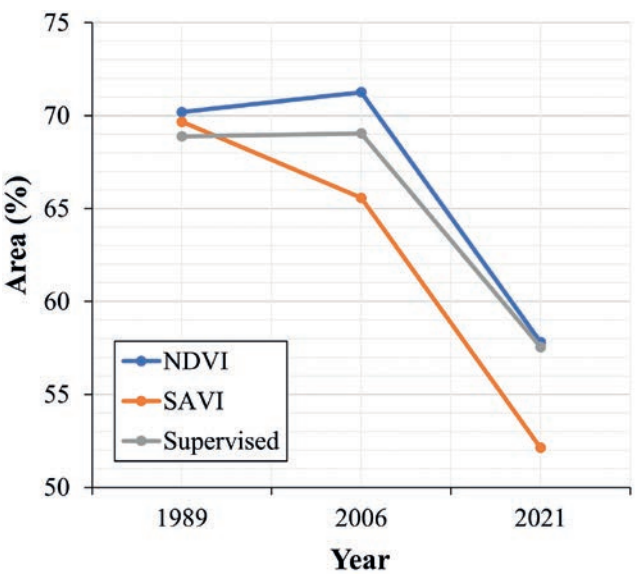


Fig. 7 Comparison of forest cover from 1989 to 2021 through NDVI, SAVI and supervised classification.

special focus on persistence and swaps (Tab. 7). A survey of forest-cover changes showed that, of a total 3461.46 km^2 of forest cover in 1989, 91.66% existed until the end of the period 1989–2006. In the next period (2006–2021), of the total 3469.51 km^2 of forest that existed in 2006, 80.03% existed in 2021 (Tab. 7). The remaining forest was converted to non-forest use.

The model showing the gain, losses, and persistence of forest cover in the study area in each period is shown in Fig. 8. The first period (1989–2006) observed a more significant increase in total forest area by 301.91 km^2, while the second period (2006–2021) saw an increase of only 120.63 km^2. The area of forest loss in the first phase was 293.86 km^2, which was lower than the increase in forest area, while in the second phase, the area of forest loss was greatly increased compared to the increased forest area of 698.33 km^2. More forest was lost during the period 1989–2006 than during the period 2006–2021 saw the least deforestation (Tab. 7). That results in the annual loss and gain in each

Tab. 7 Summary of forest cover change in Thua Thien Hue province between 1989 and 2021.

Forest Cover	Period					
	1989–2006		2006–2021		1989–2021	
	Area (km²)	%	Area (km²)	%	Area (km²)	%
Initial Year	3461.46		3469.51		3461.46	
Final Year	3469.51		2891.81		2891.81	
Persistence	3172.76	91.66	2776.70	80.03	2824.61	81.60
Loss	293.86	8.49	698.33	20.13	642.00	18.55
Gain	301.91	8.72	120.63	3.48	72.35	2.09
Annual Loss	17.29		46.56		20.06	
Annual Gain	17.75		8.04		2.26	
Annual Change	35.04		54.60		22.32	
Swap	587.72	16.98	1396.66	40.26	1284.00	37.09
Net Change	8.05	0.23	−577.70	16.65	−569.65	16.46

period also changing. During the period 1989–2006, the annual loss and gain were roughly equivalent at 17.29 km² and 17.75 km², respectively. From 2006 to 2021, the annual loss and gain area had a clear difference of 46.56 km² and 8.04 km², respectively. Forest cover increased by 8.04 km² in the first period and decreased in the second period by 577.70 km². In terms of annual variation, the annual change in forest cover in the first period was 35.04 km²; and in the second period, it increased to 54.60 km² (Tab. 7).

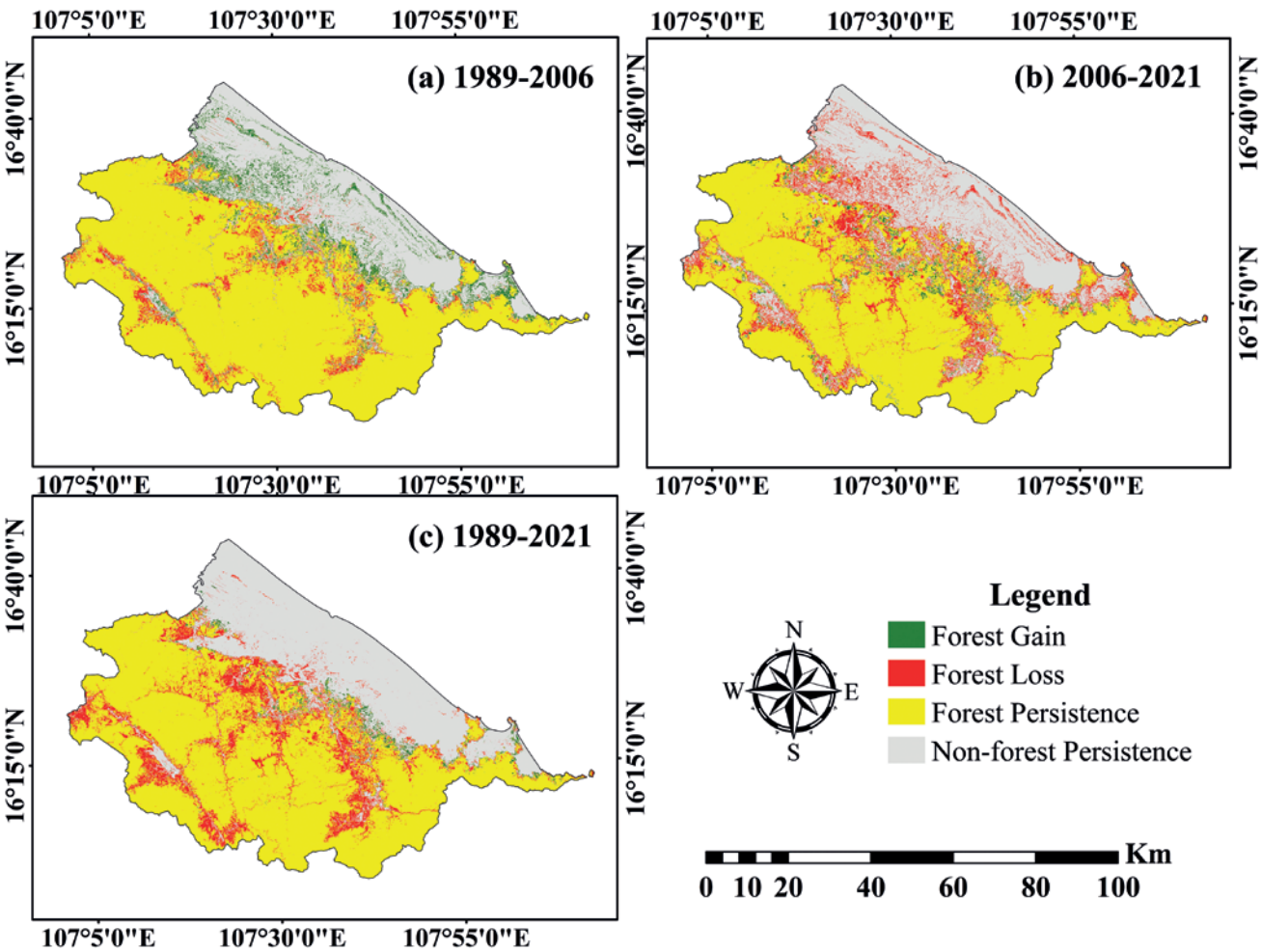


Fig. 8 Forest cover changes in Thua Thien Hue province during: (a) 1989–2006; (b) 2006–2021; and (c) 1989–2021

4. Discussion

Multi-spectral satellite images (Landsat TM and OLI) for the years 1989, 2006, and 2021 were used to classify LULC and evaluate changes in forest cover in Thua Thien Hue. These images were divided into five LULC classes: agricultural land, barren land, forest, settlement, and water (Tab. 2). The classification of these images was used the maximum likelihood algorithm for the three years is shown in Fig. 3 (Faruque et al. 2022; Thien and Phuong 2023). The results of the LULC distribution over the three focal years indicate that forests are the predominant class in the study area (Tab. 3). Forests are more common in the southwest region, while agricultural land and settlements are located near freshwater lakes, coastal areas, and plains. Generally, over the past 32 years, the forest cover has decreased, while the areas of other LULC classes have increased. Furthermore, it is crucial to assess the accuracy and validate the reliability of the resulting land cover map after classifying the LULC (Hasan et al. 2020; Thakur et al. 2021; Thien et al. 2023). This involves evaluating the errors associated with each land cover category and the overall accuracy of the classified image. The kappa coefficient values serve as a metric for measuring the consistency or precision between the reference data and the classified LULC classes. These coefficients range from −1.00 to +1.00, where values between 0.80 and 1.00 indicate a high level of agreement between the classification results and the actual data, approaching near-perfect consistency (Thakur et al. 2021). In this study, the kappa coefficient values for all land cover exceed 0.80 (Tab. 4), indicating an excellent agreement between the classified results and the reference data. This suggests that the classification process has provided highly accurate and reliable land cover information, increasing confidence in the obtained results.

The analysis of Thua Thien Hue province's multi-year LULC maps reveals significant changes that have occurred over a span of 32 years, from 1989 to 2021. These changes in LULC have both positive and negative impacts and are influenced by various natural and human factors (Phuong and Thien 2023). In particular, developing countries have experienced LULC changes that have resulted in the depletion of important natural resources such as vegetation, land, and water. Therefore, studying LULC alterations requires a comprehensive understanding and continuous monitoring of all contributing factors. In this study, our objective was to understand the changes in the forest class and identify the factors driving these changes over the 32 year period (1989–2021). We achieved this by comparing the statistical data of the forest class group in Tab. 3 and Tab. 5 and analyzing the spatial changes in the distribution of the LULC classes, as shown in Fig. 4. Through this examination, we aimed to provide a comprehensive overview of the changes in the forest class and their underlying causes during the study period.

Forests play a vital role in human life and the environment, providing not only valuable resources like timber and firewood but also playing a significant role in climate regulation and natural disaster protection (Thien and Phuong 2023). However, the findings presented in Tab. 5 indicate a severe decline in forest area during the study period from 1989 to 2021, while agricultural land and settlements have continuously expanded. Fig. 4 visually illustrates that a significant portion of the lost forest area has been converted into agricultural land and settlements, with a smaller portion transformed into water bodies. This suggests that multiple factors contribute to the decrease in forest cover in the study area, largely driven by human activities such as unsustainable logging practices, deforestation, wildfires, and land use conversion. The conversion of forest land into highly productive agricultural areas highlights the local community's dependence on agriculture as their primary livelihood (Muhati et al. 2018; Pendrill et al. 2022). Furthermore, over the past 32 years, the study area has experienced population growth, resulting in urbanization and the development of new residential areas accompanied by the necessary public infrastructure. This has led to an increased demand for food and subsequently expanded agricultural land. Water bodies have emerged in areas previously covered by forests in 1989 and 2006, serving the irrigation needs of farmers in agricultural areas. Unsustainable logging practices, forest fires for land clearance, and a lack of awareness among the local population regarding forest conservation have contributed to the significant decline in forest area. Additionally, indirect factors such as high agricultural commodity prices and ineffective forest management practices have also contributed to deforestation and forest degradation (Dan et al. 2018; Islam 2021).

The NDVI and SAVI are widely used vegetation indices for vegetation detection and mapping (Liang et al. 2018; Truong et al. 2018; Spadoni et al. 2020). In the study area, the NDVI values ranged from −0.73 to 0.74 (Fig. 5), and the SAVI values ranged from −1.09 to 1.10 (Fig. 6) across the assessment years. Higher values indicate forested areas, lower positive values represent sparse or no vegetation cover, and negative values indicate water. The increasing values of NDVI (0.74 > 0.72) and SAVI (1.10 > 1.08) in 2006 compared to 1989 indicate an increased in forest cover. However, the values of these two indices have decreased a lot in 2021 compared to 2006 with NDVI (0.49 < 0.74) and SAVI (0.74 < 1.10) which partly confirms the previous evidence about the deforestation and forest conversion resulted in a reduced forest area (Tab. 5). The estimation of forest area using the NDVI showed a positive trend in the first period (1989–2006) and a negative trend in the later period (2006–2021), similar to but slightly higher than the estimates obtained using the supervised classifier. On the other hand, the

SAVI provided similar estimates for 1989 but yielded much lower estimates for 2006 and 2021 (Fig. 7). This discrepancy may be due to sparse vegetation in 2006 and 2021, rather than dense forest, as the SAVI does not account for soil reflectivity. Based on these findings, the NDVI is a reliable indicator for detecting and monitoring forest cover in Thua Thien Hue province, especially in the context of rapid forest cover assessment (Chakraborty et al. 2018; Pesaresi et al. 2020; Huang et al. 2021; Faruque et al. 2022). In addition, Tab. 7 and Fig. 8 provide more detailed information on forest area loss and gain from 1989 to 2021. The total gross gain in the forest class is accompanied by a total loss in the other land use, particularly in the agricultural land class. Conversely, gross gains primarily came from barren and agricultural land in both periods, with the highest gains observed in the first period.

5. Conclusion

This study examines changes in forest cover and other LULC classes in Thua Thien Hue province, Vietnam, over a period of 32 years (1989–2021). By combining multi-temporal remote sensing data and GIS techniques, we quantified and analyzed the spatial and temporal patterns of LULC changes, with a particular focus on forest cover. The LULC classification results achieved an overall accuracy of over 90% during the research phase. The findings reveal significant changes in LULC classes over the study period. Forest cover and barren land decreased by 11.34% (569.65 km^2) and 9.66% (485.24 km^2) respectively, while agricultural land, settlements, and water areas increased by 11.46% (575.87 km^2), 8.68% (436.05 km^2), and 0.86% (42.97 km^2), respectively. These results indicate a strong correlation between the increase and decrease in these LULC classes. Notably, the study area experienced substantial deforestation, with 577.70 km^2 of forest being lost, particularly between 2006 and 2021. Although there was a small increase in forest cover from 1989 to 2006 (8.05 km^2), it was outweighed by the overall loss of forest cover during the 32 year period. The main drivers of changes in forest cover are agricultural expansion, urbanization and logging for housing and firewood. The study highlights the suitability of using the NDVI for mapping forest cover during rapid monitoring.

Based on these findings, we propose several recommendations to enhance forest protection and management. Firstly, policymakers should reconsider the conversion of forest land into agricultural areas, which has significantly contributed to the reduction in forest cover over the decades. Additionally, strong laws and regulations should be implemented to address both legal and illegal logging activities. Secondly, the relevant agencies should increase their personnel to effectively monitor and protect forest cover from illegal logging. By implementing these recommendations, we can better preserve the forest areas in the study region. Ultimately, we hope that this study can provide valuable insights to forest managers and agencies, assisting them in creating informative and educational materials, formulating policies, and making informed decisions regarding the conservation and monitoring of forest ecosystems. These efforts are crucial for preventing forest loss and degradation, particularly in relation to illegal logging activities.

Acknowledgements

The authors would like to thank the editor and anonymous reviewers for their thoughtful comments and efforts toward improving our manuscript.

References

Agarwal, M., Fatima, T., Freedman, H. I. (2010): Depletion of forestry resource biomass due to industrialization pressure: A ratio-dependent mathematical model. Journal of Biological Dynamics 4(4), 381–396, https://doi.org/10.1080/17513750903326639.

Ahammad, R., Stacey, N., Eddy, I. M., Tomscha, S. A., Sunderland, T. C. (2019): Recent trends of forest cover change and ecosystem services in eastern upland region of Bangladesh. Science of the Total Environment 647, 379–389, https://doi.org/10.1016/j.scitotenv.2018.07.406.

Alongi, D. M. (2008): Mangrove forests: resilience, protection from tsunamis, and responses to global climate change. Estuarine, Coastal and Shelf Science 76(1), 1–13, https://doi.org/10.1016/j.ecss.2007.08.024.

Anderson, J. R., Hardy, E. E., Roach, J. T., Witmer, R. E. (1976): A land use and land cover classification system for use with remote sensor data. Geological survey professional paper, U.S. government printing office. Washington DC 964, 1–28.

Atmiş, E., Özden, S., Lise, W. (2007): Urbanization pressures on the natural forests in Turkey: An overview. Urban Forestry & Urban Greening 6(2), 83–92, https://doi.org/10.1016/j.ufug.2007.01.002.

Bakr, N., Weindorf, D. C., Bahnassy, M. H., Marei, S. M., El-Badawi, M. M. (2010): Monitoring land cover changes in a newly reclaimed area of Egypt using multi-temporal Landsat data. Applied Geography 30(4), 592–605, https://doi.org/10.1016/j.apgeog.2009.10.008.

Butt, A., Shabbir, R., Ahmad, S. S., Aziz, N. (2015): Land use change mapping and analysis using Remote Sensing and GIS: A case study of Simly watershed, Islamabad, Pakistan. The Egyptian Journal of Remote Sensing and Space Science 18(2), 251–259, https://doi.org/10.1016/j.ejrs.2015.07.003.

Chakraborty, A., Seshasai, M. V. R., Reddy, C. S., Dadhwal, V. K. (2018): Persistent negative changes in seasonal greenness over different forest types of India using MODIS time series NDVI data (2001–2014). Ecological Indicators 85, 887–903, https://doi.org/10.1016/j.ecolind.2017.11.032.

Chowdhury, M., Hasan, M. E., Abdullah-Al-Mamun, M. M. (2020): Land use/land cover change assessment of Halda watershed using remote sensing and GIS. The Egyptian Journal of Remote Sensing and Space Science 23(1), 63–75, https://doi.org/10.1016/j.ejrs.2018.11.003.

Dan, K. O., David, P. K., Pierre, N. L. J., Chérif, A. Y. (2018): Analysis of the Causes of Deforestation and Degradation of the Forest of Katako Village. Environment and Forestry 123(2018), 51945–51948.

Dien, V. T. (2004): Susceptibility to forest degradation a case study of the application of remote sensing and GIS in Bach Ma National Park, ThuaThien Hue Province-Vietnam. International Institution for Geo-Information and Earth Observation, Enschede.

Dimyati, R. D., Danoedoro, P., Dimyati, M. (2018): Digital interpretability of annual tile-based mosaic of landsat-8 OLI for time-series land cover analysis in the Central Part of Sumatra. Indonesian Journal of Geography 50(2), 168–183, https://doi.org/10.22146/ijg.27954.

Disperati, L., Virdis, S. G. P. (2015): Assessment of land-use and land-cover changes from 1965 to 2014 in Tam Giang-Cau Hai Lagoon, central Vietnam. Applied Geography 58, 48–64, https://doi.org/10.1016/j.apgeog.2014.12.012.

Faruque, M. J., Vekerdy, Z., Hasan, M. Y., Islam, K. Z., Young, B., Ahmed, M. T., Monir, M. U., Shovon, S. M., Kakon, J. F., Kundu, P. (2022): Monitoring of land use and land cover changes by using remote sensing and GIS techniques at human-induced mangrove forests areas in Bangladesh. Remote Sensing Applications: Society and Environment 25, 100699, https://doi.org/10.1016/j.rsase.2022.100699.

Forget, Y., Linard, C., Gilbert, M. (2018): Supervised classification of built-up areas in sub-Saharan African cities using Landsat imagery and OpenStreetMap. Remote Sensing 10(7), 1145, https://doi.org/10.3390/rs10071145.

Harris, P. M., Ventura, S. J. (1995): The integration of geographic data with remotely sensed imagery to improve classification in an urban area. Photogrammetric Engineering and Remote Sensing 61(8), 993–998.

Hasan, M. E., Nath, B., Sarker, A. R., Wang, Z., Zhang, L., Yang, X., Nobi, M. N., Røskaft, E., Chivers, D. J., Suza, M. (2020): Applying multi-temporal Landsat satellite data and markov-cellular automata to predict forest cover change and forest degradation of Sundarban reserve forest, Bangladesh. Forests 11(9), 1016, https://doi.org/10.3390/f11091016.

Hor, S., Saizen, I., Tsutsumida, N., Watanabe, T., Kobayashi, S. (2014): The impact of agricultural expansion on forest cover in Ratanakiri Province, Cambodia. Journal of Agricultural Science 6(9), 46, https://doi.org/10.5539/jas.v6n9n46.

Huang, S., Tang, L., Hupy, J. P., Wang, Y., Shao, G. (2021): A commentary review on the use of normalized difference vegetation index (NDVI) in the era of popular remote sensing. Journal of Forestry Research 32(1), 1–6, https://doi.org/10.1007/s11676-020-01155-1.

Huete, A. R. (1988): A soil-adjusted vegetation index (SAVI). Remote Sensing of Environment 25(3), 295–309, https://doi.org/10.1016/0034-4257(88)90106-X.

Huete, A. R. (2012): Vegetation indices, remote sensing and forest monitoring. Geography Compass 6(9), 513–532, https://doi.org/10.1111/j.1749-8198.2012.00507.x.

Islam, M. S. (2021): Assessing the dynamics of land cover and shoreline changes of Nijhum Dwip (Island) of Bangladesh using remote sensing and GIS techniques. Regional Studies in Marine Science 41, 101578, https://doi.org/10.1016/j.rsma.2020.101578.

Jadin, I., Vanacker, V., Hoang, H. T. T. (2013): Drivers of forest cover dynamics in smallholder farming systems: the case of northwestern Vietnam. Ambio 42, 344–356, https://doi.org/10.1007/s13280-012-0348-4.

JICA, VNFOREST (2012): The study on Potential Forests and Land Related to ‘‘Climate Change and Forests” in The Socialist Republic of Viet Nam. Final report.

Lea, C., Curtis, A. C. (2010): Thematic accuracy assessment procedures: National Park Service vegetation inventory, version 2.0. Natural resource report NPS/2010/NRR–2010/204. National Park Service, Fort Collins, Colorado.

Liang, L., Chen, F., Shi, L., Niu, S. (2018): NDVI-derived forest area change and its driving factors in China. PloS One 13(10), e0205885, https://doi.org/10.1371/journal.pone.0205885.

Lu, D., Moran, E., Hetrick, S., Li, G. (2011): Land-use and land-cover change detection. Advances in environmental remote sensing sensors, algorithms, and applications. CRC Press Taylor & Francis Group, New York, pp 273–290, https://doi.org/10.1201/b10599-18.

Manandhar, R., Odeh, I. O., Ancev, T. (2009): Improving the accuracy of land use and land cover classification of Landsat data using post-classification enhancement. Remote Sensing 1(3), 330–344, https://doi.org/10.3390/rs1030330.

Matsushita, B., Yang, W., Chen, J., Onda, Y., Qiu, G. (2007): Sensitivity of the enhanced vegetation index (EVI) and normalized difference vegetation index (NDVI) to topographic effects: a case study in high-density cypress forest. Sensors 7(11), 2636–2651, https://doi.org/10.3390/s7112636.

Mermoz, S., Bouvet, A., Koleck, T., Ballère, M., Toan, T. L. (2021): Continuous Detection of Forest Loss in Vietnam, Laos, and Cambodia Using Sentinel-1 Data. Remote Sensing 13(23), 4877, https://doi.org/10.3390/rs13234877.

Meyfroidt, P., Lambin, E. F. (2008): The causes of the reforestation in Vietnam. Land Use Policy 25(2), 182–197, https://doi.org/10.1016/j.landusepol.2007.06.001.

Mubako, S., Belhaj, O., Heyman, J., Hargrove, W., Reyes, C. (2018): Monitoring of land use/land-cover changes in the arid transboundary middle Rio grande basin using remote sensing. Remote Sensing 10(12), 2005, https://doi.org/10.3390/rs10122005.

Muhati, G. L., Olago, D., Olaka, L. (2018): Land use and land cover changes in a sub-humid Montane forest in an arid setting: A case study of the Marsabit forest reserve in northern Kenya. Global Ecology and Conservation 16, e00512, https://doi.org/10.1016/j.gecco.2018.e00512.

Osio, A., Lefèvre, S., Ogao, P., Ayugi, S. (2018): OBIA-based Monitoring of Riparian Vegetation Applied to the Identification of Degraded Acacia Xanthophloea along Lake Nakuru, Kenya. In GEOBIA 2018-From pixels to ecosystems and global sustainability hal-01960341, version 1, 1–22.

Owojori, A., Xie, H. (2005, March): Landsat image-based LULC changes of San Antonio, Texas using advanced atmospheric correction and object-oriented image analysis approaches. In 5th international symposium on remote sensing of urban areas, Tempe, AZ.

Pendrill, F., Gardner, T. A., Meyfroidt, P., Persson, U. M., Adams, J., Azevedo, T., Lima, M. G. B., Baumann, M., Curtis, P. G., Sy, V. D., Garrett, R., Godar, J., Goldman, E. D., Hansen, M. C., Heilmayr, R., Herold, M., Kuemmerle, T., Lathuillière, M. J., Ribeiro, V., Tyukavina, A., Weisse, M. J., West, C. (2022): Disentangling the numbers behind agriculture-driven tropical deforestation. Science 377(6611), eabm9267, https://doi.org/10.1126/science.abm9267.

People's Committee of Thua Thien Hue Province (2005): Thua Thien hue chorography – part nature. Social Sciences Publishing House.

People's Committee of Thua Thien Hue Province (2021): Decision to announce the forest status of Thua Thien Hue province in 2020. Available at: https://stnmt.thuathienhue.gov.vn/UploadFiles/TinTuc/2021/3/3/00.00.h57439qdubnd2021pl2_signed_1.pdf.

Pesaresi, S., Mancini, A., Quattrini, G., Casavecchia, S. (2020): Mapping mediterranean forest plant associations and habitats with functional principal component analysis using Landsat 8 NDVI time series. Remote Sensing 12(7), 1132, https://doi.org/10.3390/rs12071132.

Pham, T. T., Moeliono, M., Nguyen, T. H., Nguyen, H. T., Vu, T. H. (2012): The context of REDD+ in Vietnam: drivers, agents and institutions. CIFOR Occasional, pp 75, https://doi.org/10.17528/cifor/003737.

Pham, T. T., Ngo, H. C., Dao, T. L. C., Hoang, T. L., Moeliono, M. (2021): Participation and influence of REDD+ actors in Vietnam, 2011–2019. Global Environmental Change 68, 102249, https://doi.org/10.1016/j.gloenvcha.2021.102249.

Phuong, V. T., Thien, B. B. (2023): Using Landsat Satellite Images to Detect Forest Cover Changes in the Northeast Region of Vietnam. Bulletin of the Transilvania University of Brasov. Series II: Forestry – Wood Industry – Agricultural Food Engineering 16(1), 19–36, https://doi.org/10.31926/but.fwiafe.2023.16.65.1.2.

Ranjan, R. (2019): Assessing the impact of mining on deforestation in India. Resources Policy 60, 23–35, https://doi.org/10.1016/j.resourpol.2018.11.022.

Shivakumar, B. R., Rajashekararadhya, S. V. (2018): Investigation on land cover mapping capability of maximum likelihood classifier: a case study on North Canara, India. Procedia Computer Science 143, 579–586, https://doi.org/10.1016/j.procs.2018.10.434.

Spadoni, G. L., Cavalli, A., Congedo, L., Munafò, M. (2020): Analysis of Normalized Difference Vegetation Index (NDVI) multi-temporal series for the production of forest cartography. Remote Sensing Applications: Society and Environment 20, 100419, https://doi.org/10.1016/j.rsase.2020.100419.

Thakur, S., Maity, D., Mondal, I., Basumatary, G., Ghosh, P. B., Das, P., De, T. K. (2021): Assessment of changes in land use, land cover, and land surface temperature in the mangrove forest of Sundarbans, northeast coast of India. Environment, Development and Sustainability 23(2), 1917–1943, https://doi.org/10.1007/s10668-020-00656-7.

Thien, B. B., Phuong, V. T. (2023): Using Landsat satellite imagery for assessment and monitoring of long-term forest cover changes in Dak Nong province, Vietnam. Geographica Pannonica 27(1), 69–82, https://doi.org/10.5937/gp27-41813.

Thien, B. B., Sosamphanh, B., Yachongtou, B., Phuong, V. T. (2022): Land use/land cover changes in the period of 2015–2020 in AngYai Village, Sikhottabong District, Vientiane Capital, Lao PDR. Geology, Geophysics and Environment 48(3), 279–286, https://doi.org/10.7494/geol.2022.48.3.279.

Thien, B. B., Yachongtou, B., Phuong, V. T. (2023): Long-term monitoring of forest cover change resulting in forest loss in the capital of Luang Prabang province, Lao PDR. Environmental Monitoring and Assessment 195(8), 1–17, https://doi.org/10.1007/s10661-023-11548-4.

Thua Thien Hue General Statistical Office (2019): Statistical yearbook 2018. Thuan Hoa Publishing House.

Tran, D. X., Tran, T. V., Pearson, D., Myint, S. W., Lowry, J., Nguyen, T. T. (2022): Spatiotemporal analysis of forest cover change and associated environmental challenges: a case study in the Central Highlands of Vietnam. Geocarto International 37(25), 9277–9297, https://doi.org/10.1080/10106049.2021.2017013.

Truong, N. C. Q., Nguyen, H. Q., Kondoh, A. (2018): Land use and land cover changes and their effect on the flow regime in the upstream Dong Nai River Basin, Vietnam. Water 10(9), 1206, https://doi.org/10.3390/w10091206.

Tsutsumida, N., Comber, A. J. (2015): Measures of spatio-temporal accuracy for time series land cover data. International Journal of Applied Earth Observation and Geoinformation 41, 46-55, https://doi.org/10.1016/j.jag.2015.04.018.

Yen, P., Ziegler, S., Huettmann, F., Onyeahialam, A. I. (2005): Change detection of forest and habitat resources from 1973 to 2001 in Bach Ma National Park, Vietnam, using remote sensing imagery. International Forestry Review 7(1), 1–8, https://doi.org/10.1505/ifor.7.1.1.64163.

Tuning spatial parameters of Geographical Random Forest: the case of agricultural drought

Daniel Bicák*

Department of Applied Geoinformatics and Cartography, Faculty of Science, Charles University, Czechia
* Corresponding author: daniel.bicak@natur.cuni.cz

ABSTRACT
Machine learning algorithms are widely used methods in geographical research. However, these algorithms are not properly exploiting the underlying spatial relationships present in the geographical data. One of the approaches, which addresses this problem, is based on an ensemble of local models, which are constructed from samples in close proximity to the location of prediction. This concept was applied to the Random Forest (RF) algorithm, creating a Geographical Random Forest (GRF). This study aims to further develop GRF by tuning the spatial parameters for each location in case of agricultural drought. In addition to tuning, the explanatory property of RF within the framework GRF is explored. Four machine learning models were constructed; regular RF, regular RF with spatial covariates, GRF, and GRF with the tuning of spatial parameters. Models were evaluated using Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE). Although the decrease in RMSE in this very case is relatively small, the method may provide higher improvement with different datasets.

KEYWORDS
machine learning; Random Forest; Geographical Random Forest

Received: 9 December 2023
Accepted: 29 September 2023
Published online: 1 November 2023

Bicák, D. (2023): Tuning spatial parameters of Geographical Random Forest: the case of agricultural drought.
AUC Geographica 58(2), 187–199
https://doi.org/10.14712/23361980.2023.14
© 2023 The Author. This is an open-access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0).

1. Introduction

Machine learning algorithms are increasingly being used in various academic and commercial research fields. Geographical topics are no exception. The difference between machine learning in geography and other research fields is the input data. Geographical data are located in space, therefore are often denoted as spatial data. Location in space is described either in absolute terms (geographical coordinates) or in relative, for example, adjacency to neighborhood entities. This property of data can be exploited to achieve a higher degree of accuracy. Such an approach can be described as spatially sensitive. This study aims to apply a spatially sensitive machine learning model to the complex phenomena of agricultural drought.

The traditional approach, which does not take into account spatial patterns inside data, suffers from the inability to properly model spatial relationships. This inability is called "spatial non-stationarity" (Fotheringham, Charlton and Brunsdon 1996, 605). Fotheringham, Brunsdon and Charlton (2003, 9–10) list three reasons which cause spatial non-stationarity. Firstly, there is sampling variation, which relates to statistical artifacts. Secondly, some relationships are intrinsically different across space, especially for social processes. And lastly, there is a possibility that one or more important variables are missing from the model. The second point can be exemplified by Simpson's Paradox (Simpson 1951) which refers to incorrect estimation of function when data are analyzed separately and then aggregated.

One option to capture spatial non-stationarity is to include spatial covariates, the most popular and easiest to use being geographical coordinates. A comprehensive study and evaluation of such an approach were conducted (Hengl et al. 2018) with Random Forest (RF) algorithm. The second option is to create an ensemble of local models, which encompass only a portion of all samples depending on their location. For each location where prediction takes place, a local model is created including n closest samples. In addition, one global model is created and final predictions are weighted averages of global and local models. This method has been applied to Linear Regression creating Geographical Weighted Regression (Brunsdon, Fotheringham and Charlton 1996) and to the RF algorithm creating Geographical Random Forest (GRF) (Georganos et al. 2019).

1.1 Random Forest

The Random Forest was developed by Breiman (2001) and belongs to a family of decision trees. Decision tree-based models make predictions by dividing prediction space into several subregions and have a tree-like hierarchical structure. The building of decision trees follows two steps; firstly, at each split, divide the feature space (range of values for each feature) into several distinct regions. Secondly, for each observation that falls into the same region, a prediction is made – the mean of values of the predicted variable. The problem is to find value by dividing predictor space most efficiently. The threshold value is calculated so that the overall sum of square errors is minimized (Kuhn and Johnson 2013, 175). However, it is not computationally feasible to find optimal partition for features. The algorithm begins with one region and then successively divides the feature space at each split. At each split, the best partition is made at that particular split. This is also known as a top-down greedy approach. Unfortunately, the variance of trees is very high. The application of bootstrap aggregation (bagging) decreases the variance by averaging many similar trees from bootstrapped datasets. The new datasets are sampled with replacements from the original dataset. On average, one-third of all samples are not used during the tree-building process and are called out-of-bag (OOB). RF algorithms further develop this concept by incorporating an ensemble of decision trees. In addition, the algorithm considers only part of the features at every splitting, which decreases correlation among trees and therefore decreases variance.

The Random Forest algorithm achieves one of the highest forecasting accuracies compared to other algorithms for the broad field of tasks (Berk 2020, 288). One of the advantages of RF is its great performance for high dimensional data when the amount of predictors is higher than the amount of observation. Another reason to choose RF is its great computational performance, which is native to all tree-based algorithms. "Compared to bagging, RF is more computationally efficient on a tree-by-tree basis since the tree-building process only needs to evaluate a fraction of the original predictors at each split" (Kuhn and Johnson 2013, 200). RF can be running simultaneously on more cores and results can be aggregated afterward (Liaw and Wiener 2002, 22).

Hyperparameters of machine learning algorithms control the training process. Values for each hyperparameter need to be set before the start of the training phase. Hyperparameter optimization is necessary to construct a stable and accurate machine learning model. RF has several important hyperparameters. A number of randomly drawn features during the splitting phase often denoted as mtry, influences the stability and prediction accuracy. Lower values tend to boost the stability of the model, on the other hand, the accuracy is slightly lower (Probst, Wright and Boulesteix 2018, 3). Lower values also decrease the computational complexity, as the algorithm does not need to calculate as many thresholds. The next parameter, the number of trees in the forest should be set to at least 100. According to Probst and Boulesteix (2017), the accuracy increases with diminishing returns when inputting higher values. However, the computational complexity increases as well.

The Hyperparameter Minimal Number of Samples describes how many samples are used for training each tree and its effect is similar to hyperparameter number of randomly drawn features. Additional parameters are Node size (minimum number of samples in a terminal node) and Splitting rule (function to assess the quality of the split). Generally, RF performs well with tuning only mtry hyperparameter (Fernández-Delgado et al. 2014, 3175). Hyperparameters can be tuned with traditional methods, for example, Grid Search or Random search. Existing OOB samples can be utilized to evaluate the model, which saves time.

In addition to hyperparameters native to RF, GRF brings hyperparameters bandwidth and local weight. Bandwidth describes the size of the kernel for each local model. In other words, a number of closest samples of which local models are trained. There are two types of kernels – adaptive and fixed. The former encompasses n closest samples in the vicinity where prediction takes place. The latter is the circle, in which the radius is the bandwidth (Fotheringham, Brunsdon and Charlton 2003, 44). The final prediction for location is made from a weighted average of the local model and global model, where the weight for the local model is a tunable parameter. The combination of two models results in higher accuracy – the local model ensures low bias and the global one has low variance (Georganos et al. 2019, 7). The drawback is higher computational complexity, GRF needs to compute a new model for each predicted location.

1.2 Agricultural drought

Environmental hazards are natural phenomena that negatively affect human society regarding economic and social losses. Drought hazard belongs to the most damaging and widespread causes of huge economic and human losses. The severity of drought depends on the environment's (or society's) ability to cope with hazards. For example, in developed countries, drought's direct impact is almost invisible, and indirect impact projects to higher consumption of water to irrigate agricultural plants. In developing countries, drought might cause crop failure and subsequent instability. However, climate change will worsen many aspects of drought, including its recurrence, severity, and timespan (Mukherjee et al. 2018).

The gravity of drought hazard is reflected in the abundance of research studies focusing on identifying vulnerable locations or factors. Various methods have been applied; including the subjective weighting of drought drivers (Wilhelmi and Wilhite 2002) or the analytical hierarchy process (Hoque et al. 2020). More recently, machine learning algorithms are utilized to model drought. For example, Rahmati et al. (2019) employed RF, Support vector machines, and others to create a vulnerability map of Queensland, Australia. A similar study utilized an Artificial Neural Network (Rahmati et al. 2020).

The machine learning model requires independent variables, which influence the drought intensity, and a dependent variable, which functions as a drought indicator. According to Mishra and Singh (2010, 207), the drought indicator is a prime variable for assessing the effect of drought and defining different drought parameters, which include intensity, duration, severity, and spatial extent. The selection of an appropriate indicator is essential as it will be the dependent variable, which will be predicted by the model. Soil moisture, especially within the root system of the plants, is an accurate indicator of agricultural drought. Soil moisture-based indicators are used in similar studies concerning agricultural drought e.g. Rahmati et al. (2019) and Rahmati et al. (2020).

The severity of agricultural drought is influenced by various factors. The most profound is meteorological. The connection between agricultural drought and meteorological patterns is clear. Precipitation is the only source of moisture for the environment with the exception of irrigation, which is available for a fraction of cultivated areas. Temperature influences the rate of transpiration, higher temperatures increase the transpiration rate. A region with higher temperatures is, therefore, more prone to drought. However, precipitation deficit impacts are greater than high temperatures in general (Yang et al. 2020, 9).

Topographic characteristics refer to the quantitative descriptions of the physical features of land. Vegetation in mountainous regions subscribes to different patterns of climatic conditions and develops specific adaptations. The slope of an area affects the run-off, recharge, and movement of surface water. Flat terrain areas have relatively high infiltration rates, on the other hand, areas with steeper slopes have low infiltration rates and higher run-off (Shekhar and Pandey 2015, 409). Another topographic factor is aspect, which refers to the orientation of the slope. The aspect of a slope can influence local climate because of the length of the exposure to sun rays. West and south-facing slopes are warmer than east and north-facing slopes, therefore having lower soil moisture and higher evaporation rate (Magesh and Chandrasekar 2010, 375). The topographical Wetness Index (TWI) (Beven and Kirkby 1979) describes the proclivity of a place to accumulate water based on topographic information. TWI is a widely used indicator to obtain information on the spatial distribution of wetness conditions, since only a terrain model is required for calculation. Soil properties are important factors influencing the environment's ability to cope with drought. Soil acts as a substrate for plants' roots, providing them with water and nutrients. Soil characteristics influence these functions to various degrees. Soil texture refers to the size of solid particles, that soil is composed of. The size of particles determines the amount of water that can be stored for plants. Organic matter is one of the most important soil characteristics. According to Bot and Benites

(2005, 35–36), organic content increases water infiltration and water holding capacities, increasing the diversity and activity of soil organisms and providing nutrient availability. Land cover is intertwined with water demand and the coping abilities of the environment to drought hazards. Land use describes how society uses land, land cover refers to the physical features of the land. In case of vulnerability to drought, the scientific community classifies several types – mainly agricultural fields, grasslands, forests, barren lands, urban areas, and water bodies (Jain, Pandey and Jain 2014; Thomas et al. 2016; Hoque et al. 2020).

This study continues to develop GRF by tuning spatial parameters for each location. The hypothesis is that tuning the spatial hyperparameters for each location will improve the accuracy of the GRF model. This hypothesis is based on the assumption that those spatial hyperparameters are spatially correlated. The study aims to confirm the hypothesis by completing three subtasks; firstly, creating an accurate statistical model based on the RF algorithm of drought hazard which consists of many local models and one global model and subsequently evaluating the accuracy metrics for both models. Secondly, performing a tuning of parameters for each local GRF model, answering the question of whether it is possible to improve the accuracy of the model further. Lastly, providing insight into the vulnerability modeling from the feature's importance of local models.

2. Methodology and Data

This section describes the study area, datasets used to build the machine learning model, and method, which facilitates the local tuning of spatial hyperparameters.

2.1 Study Area

The problem is studied within the agricultural landscape in the Czech Republic, Central Europe. The extent reaches approximately 50 km beyond the border north to Poland, west to Germany, south to Austria, and east to Slovakia. The study area was limited to an agricultural landscape with these conditions:

- Forest should not cover more than 20% of the pixel area.
- Built-up areas should not cover more than 20% of the pixel area.

The study are is shown in Fig. 1. The total area of locations, that met the conditions is 53,860 km^2.

2.2 Data

The drought model requires many features (independent variables) to work properly. This is reflected in the variety of data sources, from which data were collected. The independent variable – the drought predictor was chosen to be the Soil Water Index (SWI). The selection was influenced by the availability of data in terms of spatial and temporal resolution. SWI is available within Copernicus Global Land Service (Bauer-Marschallinger et al. 2018). SCATSAR-SWI (Scatterometer Synthetic Aperture Radar Soil Water Index) is computed from the data fusion of products Sentinel-1 SSM (Surface Soil Moisture) and ASCAT SSM/SWI, which assess soil moisture. The dataset includes layers with various temporal parameters T, which correspond to different soil depths. The layer with a T value of 20 was chosen as it correlates best with the subsoil conditions (10–20 cm below the surface) (Paulik et al. 2014, 5) and has uniform quality scores across the study area.

Meteorological features were acquired from the E-OBS dataset maintained by the European Climate

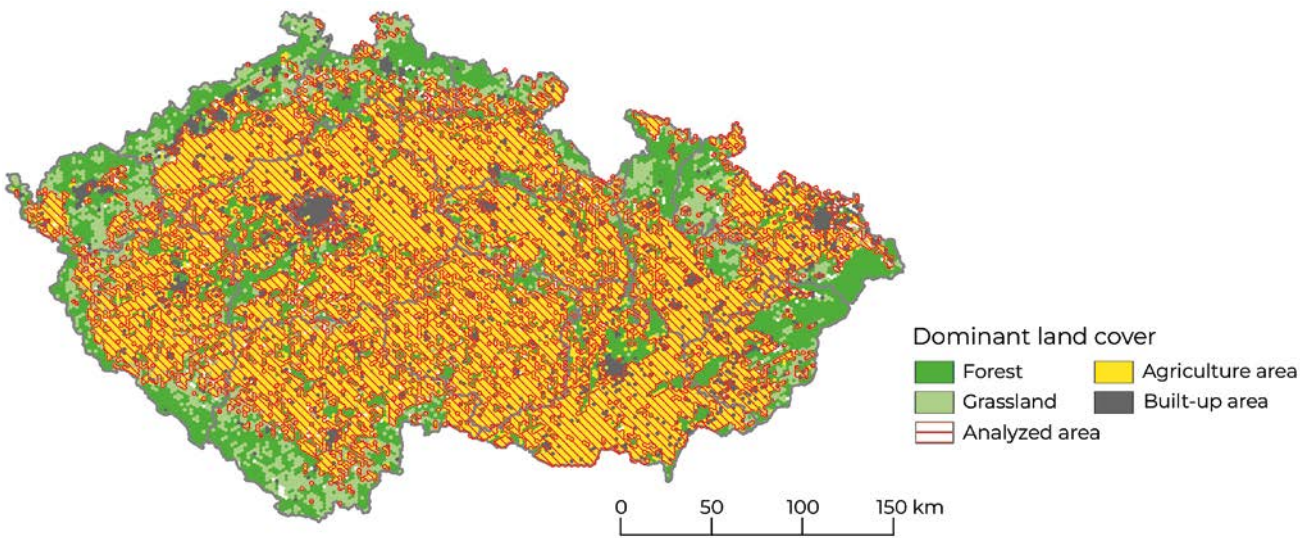


Fig. 1 Study area.

Tab. 1 Aggregated thematic classes and their corresponding Land cover categories.

Aggregated class	Former classes
Built-up areas	Continuous urban fabric, Discontinuous urban fabric, Industrial or commercial units, Road and rail networks and associated land, Port areas, Airports, Mineral extraction sites, Dump sites, Construction sites
Agricultural areas	Non-irrigated arable land, Vineyards, Fruit trees and berry plantations, Annual crops associated with permanent crops, Complex cultivation patterns
Grasslands	Pastures, Natural grasslands, Moors and heathland
Forests	Broad-leaved forest, Coniferous forest, Mixed forest, Transitional woodland-shrub

Assessment & Dataset project. E-OBS is interpolated from point data gathered from national meteorological stations across Europe. According to the project website (Cornes et al. 2018), Czechia has an above-average density of stations (770 km^2 for precipitation and 913 km^2 for temperature per station). All topographic-related features were extracted from European Digital Elevation Model (EU-DEM).

All soil properties except Soil organic matter were acquired from the "Topsoil Physical Properties for the Europe" dataset, which is based on Land Use and Cover Area frame Statistical Survey (LUCAS) dataset. LUCAS is the largest harmonized soil dataset in Europe overseen by the Statistical Office of the European Union, which consisted of in situ measurements from more than 22,000 locations (Orgiazzi et al. 2018). Another dataset derived from LUCAS is Soil Organic Matter (SOM) fractions (Cotrufo et al. 2019), which utilized more than 9400 points, to interpolate point data to a grid with a 1 km spatial resolution using the RF algorithm. Organic matter is divided by size into particulate and mineral-associated organic matter (less than 53 µm). Datasets are delivered in GeoTiff format and ETRS89-LAEA coordinate system. Both datasets, TPPE and SOM are distributed by the European Soil Data Centre (Panagos et al. 2012).

Land cover information was obtained from Corine Land Cover (European Environment Agency (EEA), 2019). Land cover categories were aggregated into four thematic classes – built-up areas, agricultural areas, grasslands, and forests. Land cover categories and their corresponding thematic classes are listed in the table below (Tab. 1).

Metadata of datasets used in the study are listed in following table (Tab. 2).

2.3 Methods

A new variant of GRF was developed – Locally Tuned GRF (LT GRF). Values for hyperparameters bandwidth and local weight are universal for every sample across space. LT GRF aims to find optimal values for each location. The optimal values are found for each location during the training process. Values are interpolated for the whole study area by linear interpolation. In case there are several different values in one place, the mean value is used. We assume, that there exists a spatial autocorrelation in the model parameter's weights and bandwidth. Six models in total were developed; RF, GRF, and LT GRF with coordinates and RF, GRF, and LT GRF without coordinates. LT GRF should be the most accurate model created. Algorithm LT GRF can be described by the following pseudocode.

Algorithm 1 Geographical Random Forest with local tuning.

1: **for** training observation **do**
2: **for** bandwidth, local weight in the kernel, weights **do**
3: Perform Random Forest.
4: Perform Random Forest with bandwidth number of samples.
5: **end for**
6: Select optimal values of bandwidth and local weight.
7: **end for**

Tab. 2 Metadata of datasets.

Product name	Original temporal resolution	Temporal resolution used in the study	Original spatial resolution	Format	Reference
SCATSAR-SWI	1 day	14 days	1 km	NetCDF	Copernicus Global Land service (2023)
E-OBS	1 day	14 days	0.1°	NetCDF	Cornes et al. (2018)
EU-DEM	/	/	25 m	GeoTiff	European Environment Agency (2016)
Topsoil Physical Properties for the Europe	/	/	500 m	GeoTiff	Panagos et al. (2022)
Soil Organic Matter	/	/	1 km	GeoTiff	Lugato et al. (2021)
Corine Land Cover	/	/	100 m	GeoTiff	Copernicus Land Monitoring Service (2019)

8: Interpolate bandwidth and local weight values for the location of testing observations.
9: **for** testing observation **do**
10: Perform Random Forest.
11: Perform Geographical Random Forest.
12: Compute the weighted average of the output of Random Forest and Geographical Random Forest.
13: **end for**

The process of evaluating models consists of several steps; pre-processing of the data, model building, and performance evaluation by accuracy metrics. Firstly, the time periods were chosen for SWI. Five time periods – the first two weeks of August from 2015 to 2019. Each period has a different distribution of values, therefore combined dataset contains observations with low and high values for the same place.

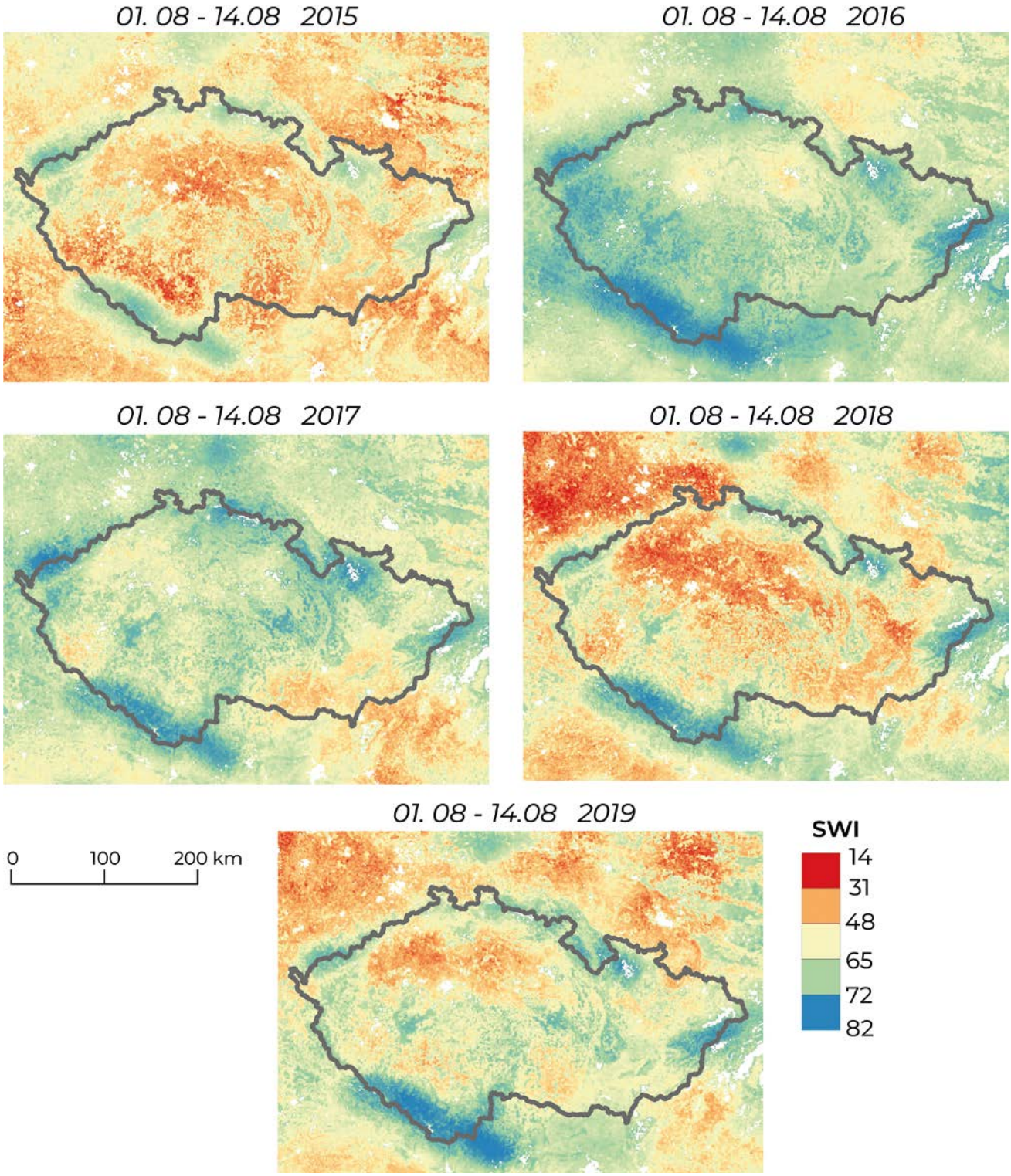


Fig. 2 SWI values for each selected period.

Various values of SWI for the same place can facilitate a robust and accurate model. SWI values are shown in Fig. 2. Secondly, drought predictors were chosen; geographical coordinates (X and Y), elevation, slope, aspect, TWI from category terrain characteristics, temperature and precipitation from meteorological characteristics, soil texture, organic matter content, soil bulk ratio, and AWC from soil properties. Organic matter content was created as a sum of both layers of the SOM dataset. Landcover features are represented by their proportion in each location. Four land cover classes were chosen – built-up areas, agricultural areas, grasslands, and forests. In addition to the listed features, the distance to large water bodies (rivers and reservoirs) was added. SWI and meteorological features are available for each day, therefore need to be aggregated. SWI and temperature are averaged, and precipitation is summed. Three periods of temperature and precipitation are selected. Two weeks period, which is identical to the SWI period, a one-month period (two weeks before the start of the SWI period), and a three-month period.

All datasets were resampled to sample size with a resolution of 3 km^2 using linear interpolation. Values of all features (independent variables) were scaled from 0 to 1. The dataset was split into training and testing sets with a ratio of 0.33 (two-thirds were used for tuning and one-third for testing) using random sampling. Three hyperparameters of RF were tuned using Grid Search with OOB samples – a number of randomly drawn features, a number of trees, and a minimal number of samples. Subsequently, GRF hyperparameters bandwidth and local weight were tuned using a grid search cross-validation method. Several bandwidth values were tested; 50, 100, 150, 200, 250, 500, 750, 1000, 1500, 2000 and 5000. The distances are not equal for all locations, because of the use of an adaptive kernel. For a bandwidth of size 50, the average distance is 5205 m. For maximum bandwidth of 5000, the average distance is approximately 45 km. The parameter of local weight values from 0 to 1 with increment 0.1 were tested.

The performance of statistical models is evaluated by metrics – Root Mean Square Error (RMSE) and Mean Absolute Error (MAE). These metrics were chosen because of their wide use in the scientific community. The RMSE is calculated by the formula:

$$RMSE = \sqrt{\sum_{i=1}^{n} \frac{(\hat{y}_i - y_i)^2}{n}}$$

and MAE is expressed by formula:

$$MAE = \frac{1}{n} \sum_{i=1}^{n-1} |y_i - \hat{y}_i|$$

In addition to the RMSE and MAE, relative accuracy in % is used. Relative accuracy is calculated as a ratio of error (RMSE or MAE) to the range of values of the dependent variable without outliers. Outliers are understood to be values less than one percentile and higher than 99 percentil of SWI values.

The import, processing and building of the models took place using the python programming language. Libraries used include scikit-learn, NumPy, Pandas and Xarray. Map outputs were created using QGIS software.

3. Results

This section describes the results of tuning the machine learning models, performance assessment, and the feature importances of GRF LT model shown graphically in the maps.

The hyperparameters were tuned using the Grid Search method with OOB samples. The optimal value for the number of randomly drawn features was found to be 16. The minimum number of samples was set to 5. The number of trees is 200, and the RMSE decreases with diminishing returns with increasing the value of the hyperparameter. The RMSE of values of spatial hyperparameters is depicted in Figure 3. The optimal value for local weight was found to be 0.7 (0.7 local model and 0.3 global model). The decrease in error between regular RF (local weight is 0) is approximately 0.297 RMSE. In comparison, the difference between the default value of hyperparameter Randomly Drawn Features (typically one-third of all available features) is 0.2 RMSE. The optimal bandwidth is 100 observations. The decrease in RMSE in comparison to the global model is very small 0.09, smaller than the decrease of hyperparameter local weight.

The optimal values of the bandwidth and local weight are found for each location. The count of each value of the parameter is displayed in a histogram below (Fig. 4). The most numerous value for bandwidth is 50 constituting 20% of all values. The second place belongs to value 100 with a 12.7% share. Other values constitute a portion smaller than 10%. GRF assigns one universal value to all locations, however, as can be seen, it is not optimal for the vast majority of locations. In the case of local weight, the situation is more uneven. The most numerous local weight value is 1 (only the local model is employed) which constitutes 56.4% of all values. The second most numerous is value 0 (only the global model is employed) with an 8% contribution to all values. Other values are represented less, the count decreases with lower local weight. However, the best value achieved by GRF tuning is 0.8. This value is not optimal for more than 92% of all locations.

Each tuned model was trained and tested. RF without coordinates achieved an RMSE of 4.48 (MAE of

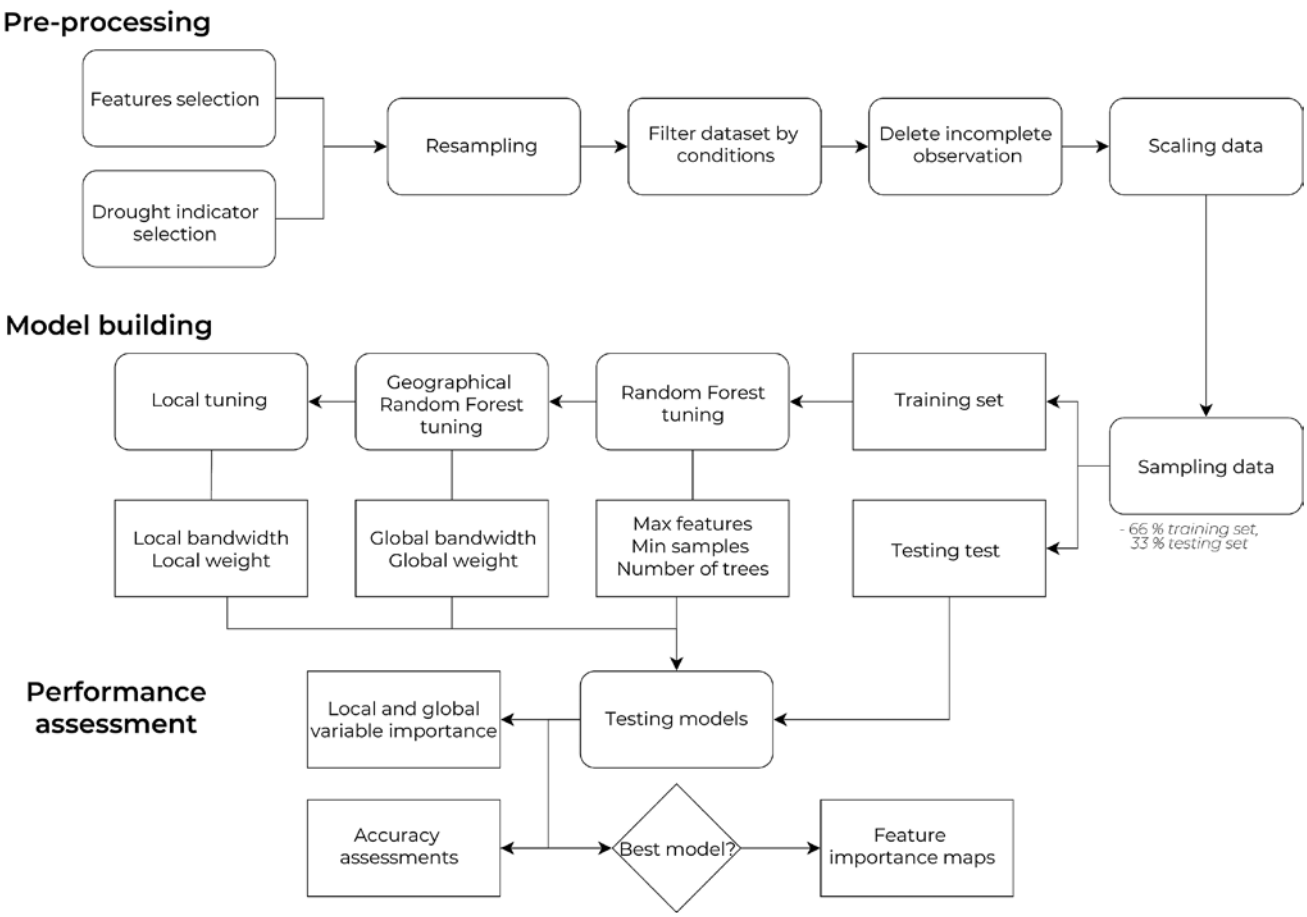


Fig. 3 Flow chart of the research.

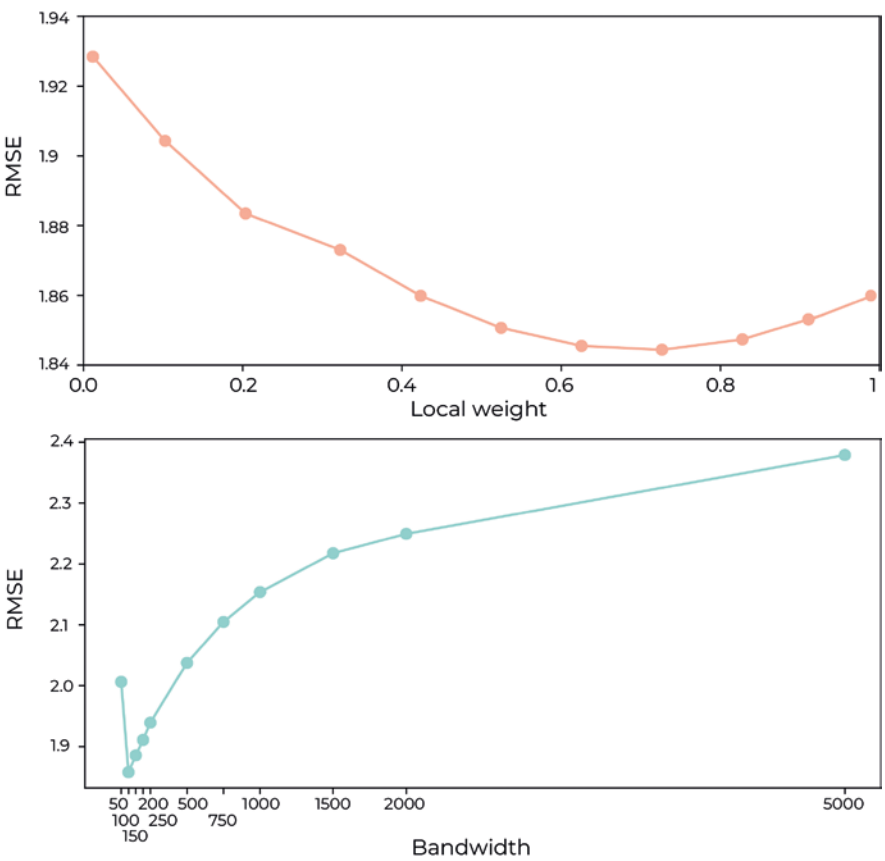


Fig. 4 RMSE of spatial parameters, extracted from tuning of GRF with spatial coordinates.

Tab. 3 Accuracy metrics for each tested model expressed in relative and absolute values. Models with spatial coordinates are denoted with 'XY'.

	RMSE		MAE	
	abs	rel [%]	abs	rel [%]
RF model	4.48	90.86	3.40	93.02
RF XY model	2.70	94.46	2.05	95.77
GRF model	2.60	94.66	1.96	95.97
GRF XY model	2.42	95.03	1.83	96.24
LT GRF model	2.58	94.70	2.02	95.85
LT GRF XY model	2.41	95.05	1.80	96.30

3.4), RF with spatial coordinates achieved an RMSE of 2.7 (MAE of 2.05), GRF without coordinates of 2.6 (MAE of 1.96), GRF with spatial coordinates of 2.42 (MAE of 1.83), LT GRF without coordinates of 2.58 (MAE of 2.02) and LT GRF with spatial coordinates of 2.41 (MAE of 2.41). Values are listed in the table below (Tab. 3).

Feature importance of the RF and GRF models are displayed in bar plots below (Fig. 5 and Fig. 6). The most important features of the RF (without spatial coordinates) are precipitation features. Together they account together for almost approximately 60% of the importance. Summed precipitation over 1-month accounts for 38% and is the most important feature. The fourth place belongs to elevation with 6% of importance. The elevation is followed by temperature features, each accounting for 5% of importance. Other features have less than 3% of importance.

Feature importances of the GRF LT model (without spatial coordinates) were aggregated by the mean value (Fig. 7). Values are similar to the RF model. The

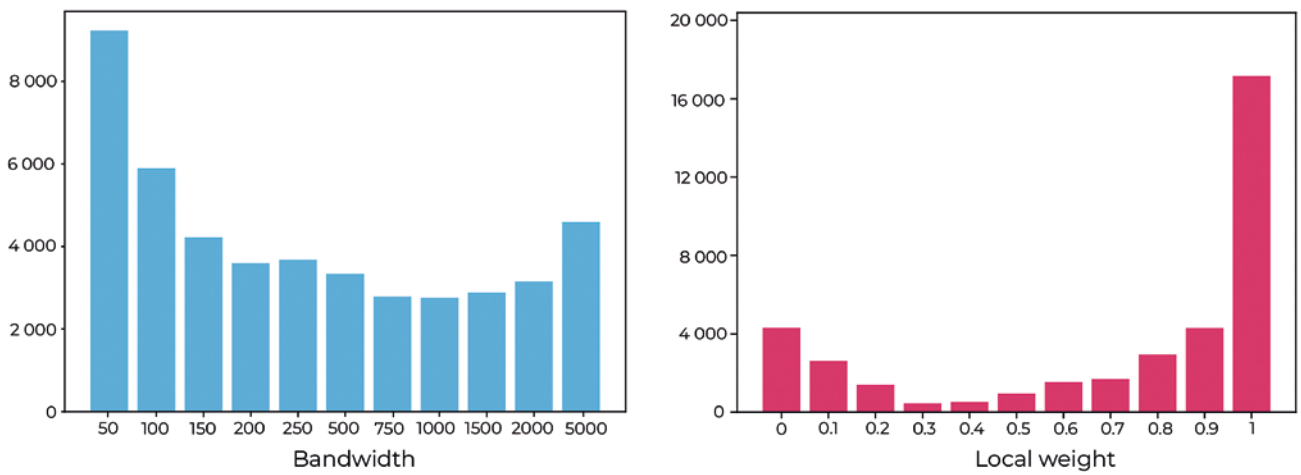


Fig. 5 Histograms for bandwidth and local weight extracted from tuning of GRF with spatial coordinates.

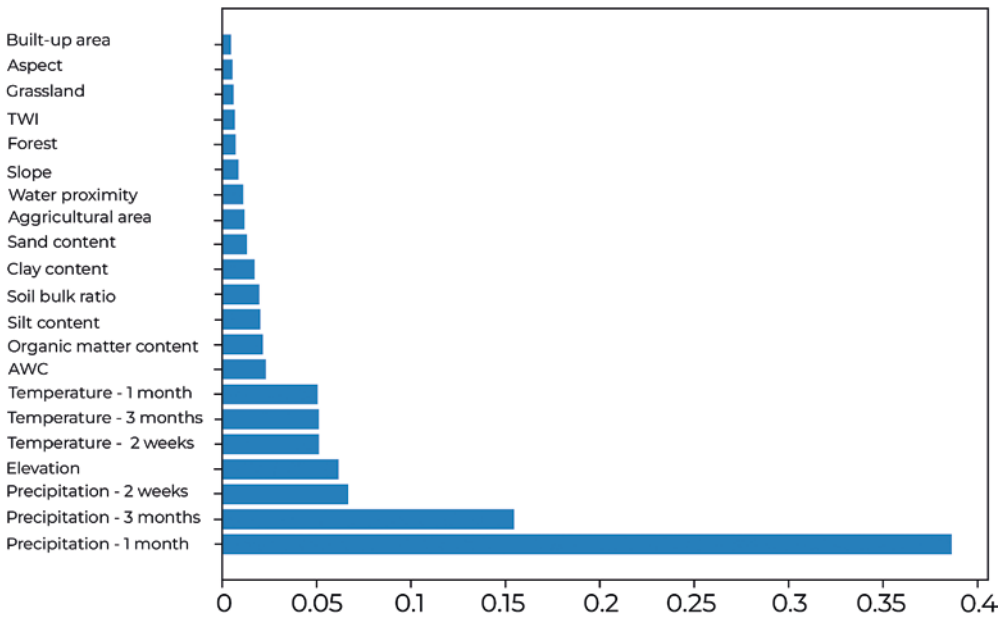


Fig. 6 Feature importances for RF model (without spatial coordinates).

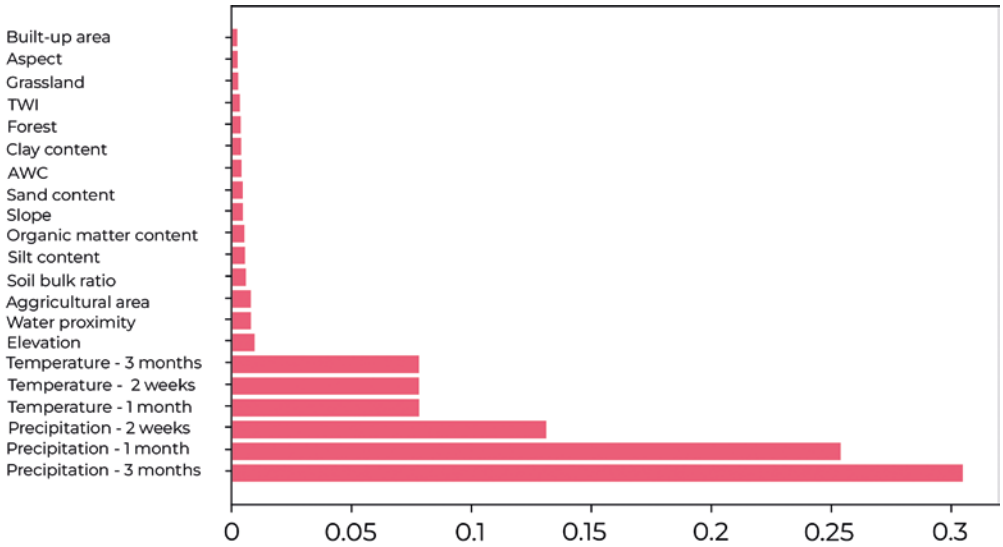


Fig. 7 Feature importances for GRF LT without spatial coordinates.

most important features are precipitation features (1 month with 38%, 3 months with 15%, and 2 weeks with 6%) followed by temperature features (each with approximately 7%). Elevation and water proximity account for approximately 1%. The rest of the features account for only 6% of feature importance.

The top eight features were visualized in maps (Fig. 8). Precipitation features show different spatial distributions. The longest precipitation period (3-months) shows strong importance in the northwestern part of the area, 1-month precipitation in western and southwestern parts of the area, and the shortest period in the eastern part. Temperature features show a similar pattern for each period. Strong importance can be assessed in southern part of the area.

3. Discussion

The tuning of spatial parameters showed that there is no optimal uniform value for all training locations. However, GRF or LT GRF did not achieve a significantly higher degree of accuracy. The minuscule difference in error between models can be explained in several ways. Firstly, regular RF achieves very good results. Accuracy of more than 90.86% is very high and the possibility for improvements is limited. It is most likely that model accuracy cannot be significantly improved any further for the given modeled problem and available input data sets. Secondly, spatial non-linearity is explained well by spatial coordinates, which are input features in the RF model. In other words, the global model (RF model with spatial coordinates) has not left any space for local models (GRF and GRF LT models) to improve more significantly.

The reduction in RMSE error between the GRF and GRF LT models (models with spatial coordinates) is surprisingly low. During the tuning phase, GRF LT achieved an RMSE of 1.8, which is significantly less than the resulting error in the testing phase. The limited improvement of the GRF corresponds with the visual examination of bandwidth. There is no or very little spatial correlation between bandwidth and local weight and a decrease in error. Values are localized randomly as a residue of random error.

GRF creates local models on a subset of original datasets. This process can be reinterpreted as a huge number of created decision trees with a very small number of observations. A similar situation can be recreated with regular RF with parameter maximum samples set to a value of best bandwidth (100). However, experiments show that such a model is very inaccurate (RMSE of 9.5) and this hypothesis can be rejected.

The performance of GRF and LT GRF was compared to the performance in other studies. The GRF with spatial covariates in the study by Georganos et al. (2021) achieved an RMSE of 0.606, the global model achieved an RMSE of 0.65. The error decreased by 6.76%. Master thesis by Hokstad and Tiganj (2020) compared RF with spatial covariates to GRF. RF achieved an RMSE of 17,944 and GRF of 16,705, a 6.9% decrease in error. In both studies, a decrease in error between RF with spatial coordinates and without them is more significant.

Improvement of GRF or LT GRF over regular RF is small and computational runtime is much higher. A desktop PC is not sufficient for larger datasets (more than 100,000 samples) and a more expensive solution needs to be employed. Therefore, it may seem that GRF might not be advantageous over classical RF based on this case study given the computational requirements and not significant improvements in model performance.

The feature importance results provide unique insight into the drivers of agricultural drought.

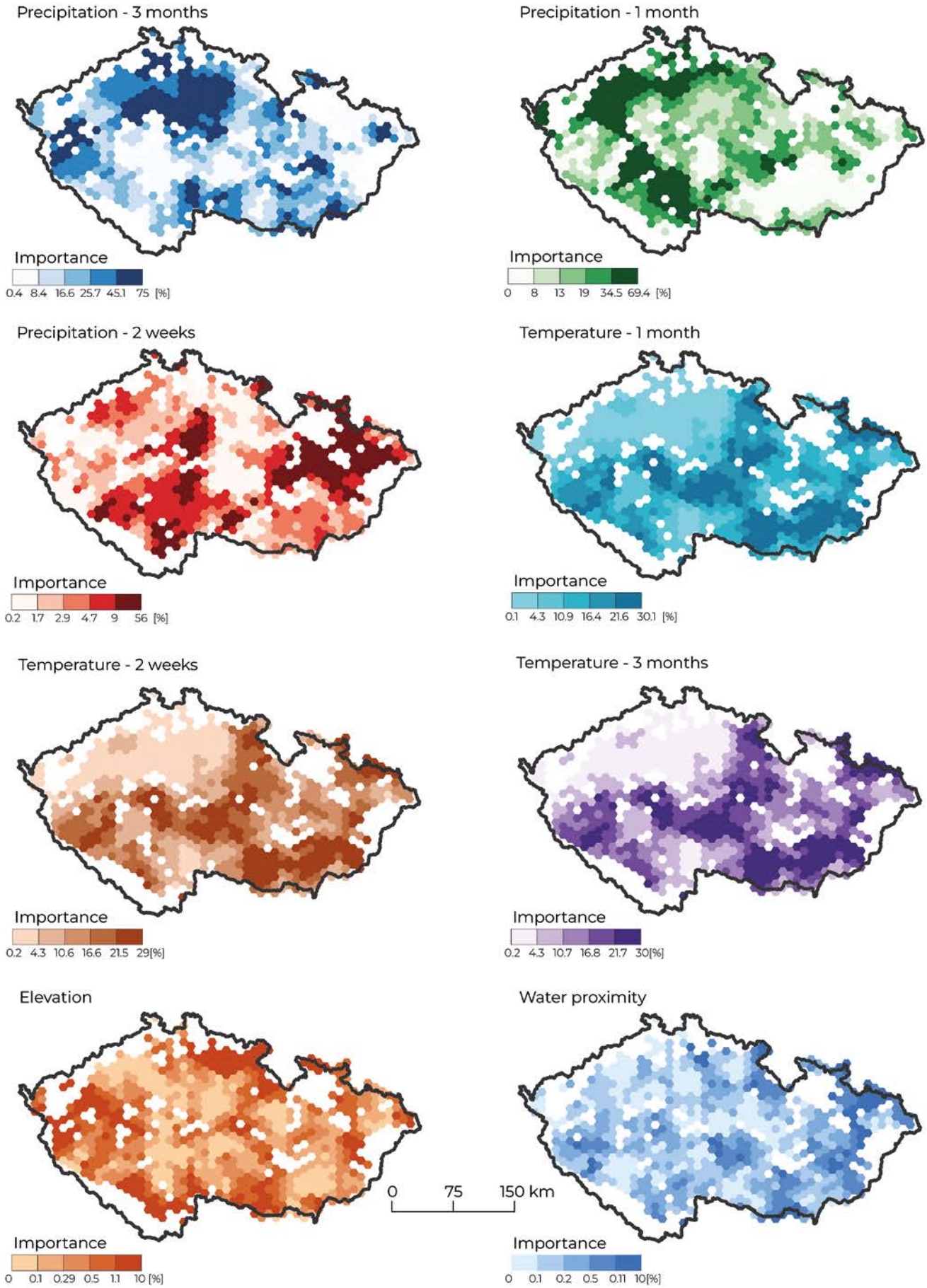


Fig. 8 Top most important features of GRF LT (without spatial coordinates) visualized on the map.

However, to draw conclusions from maps, deeper knowledge of local conditions is needed.

The concept of local sub-models and their parameters can be studied further. Despite the increase in accuracy being small, local tuning might deliver more beneficial results in different use cases. Therefore, it would be useful to find and evaluate spatial patterns in various datasets, which would benefit most from this method. The use of more sophisticated spatial interpolation methods such as kriging when obtaining unknown values of local parameters can increase the accuracy of models. Such an approach would be particularly advantageous for a sparse dataset. The GRF and GRF LT models used use a binary kernel – records up to a certain distance are included in the local model. Another approach is to use a function that would assign weights to records based on their distance.

The explanatory function of the model (features importance) has the potential to provide additional insight into geographical phenomena. Results from GRF or LT GRF can be compared with more established methods such as Geographic Weighted Regression. This concept can be also extended to other Machine Learning algorithms. As mentioned by Georganos et al. (2019), Support Vector Machines are suitable methods, because of their lower computational complexity.

4. Conclusion

The study developed six machine learning algorithms; RF with and without spatial coordinates, GRF with and without spatial coordinates and LT GRF with and without spatial coordinates. LT GRF in contrast to GRF tunes the local parameters – bandwidth and local weight for each location. The models were applied and evaluated in the case of agricultural drought. A total of 21 features were used to predict drought using a soil moisture-based index as the dependent variable. In addition, the study provides insight into the feature importance property of GRF. The increase in accuracy is relatively small in this very case, however, different datasets may provide more desirable results.

References

Bauer-Marschallinger, B., Paulik, C., Hochstöger, S., Mistelbauer, T., Modanesi, S., Ciabatta, L., Massari, C., Brocca, L., Wagner, W. (2018): Soil moisture from fusion of scatterometer and SAR: Closing the scale gap with temporal filtering. Remote Sensing 10(7), 1030, https://doi.org/10.3390/rs10071030.

Berk, R. A. (2008): Statistical Learning as a Regression Problem. In: Statistical Learning from a Regression Perspective. Springer Series in Statistics. Springer, New York, NY, https://doi.org/10.1007/978-0-387-77501-2_1.

Beven, K. J., Kirkby, M. J. (1979): A physically based, variable contributing area model of basin hydrology/Un modèle à base physique de zone d'appel variable de l'hydrologie du bassin versant. Hydrological Sciences Journal 24(1), 43–69, https://doi.org/10.1080/02626667909491834.

Bot, A., Benites, J. (2005): The importance of soil organic matter: Key to drought-resistant soil and sustained food production. In: Food & Agriculture Org. Breiman, L. (2001). Random forests. Machine Learning 45, 5–32, https://doi.org/10.1023/A:1010933404324.

Brunsdon, C., Fotheringham, A. S., Charlton, M. (1998): Geographically weighted regression. Journal of the Royal Statistical Society: Series D (The Statistician) 47(3), 431–443, https://doi.org/10.1111/1467-9884.00145.

Copernicus Global Land service (2023): Soil Water Index. https://land.copernicus.eu/global/products/swi.

Cornes, R. C., van der Schrier, G., van den Besselaar, E. J., Jones, P. D. (2018): An ensemble version of the E-OBS temperature and precipitation data sets. Journal of Geophysical Research: Atmospheres 123(17), 9391–9409, https://doi.org/10.1029/2017JD028200.

Cotrufo, M. F., Ranalli, M. G., Haddix, M. L., Six, J., Lugato, E. (2019): Soil carbon storage informed by particulate and mineral-associated organic matter. Nature Geoscience 12, 989–994, https://doi.org/10.1038/s41561-019-0484-6.

Hengl, T., Nussbaum, M., Wright, M. N., Heuvelink, G. B. M., Gräler, B. (2018): Random forest as a generic framework for predictive modeling of spatial and spatio-temporal variables. PeerJ 6: e5518, https://doi.org/10.7717/peerj.5518.

Hokstad, V., Tiganj, D. (2020): Spatial modelling of unconventional wells in the Niobrara Shale play: a descriptive, and a predictive approach. Master's thesis. Norwegian School of Economics.

Hoque, M. A., Pradhan, B., Ahmed, N., Sohel, Md. S. I. (2021): Agricultural drought risk assessment of Northern New South Wales, Australia using geospatial techniques. In: Science of the Total Environment 756: 143600, https://doi.org/10.1016/j.scitotenv.2020.143600.

European Union, Copernicus Land Monitoring Service (2019): Corine Land Cover. European Environment Agency (EEA), https://land.copernicus.eu/en.

European Union, Copernicus Land Monitoring Service (2016): EU – Digital Elevation Model 1.1. European Environment Agency (EEA), https://land.copernicus.eu/en.

Fernández-Delgado, M., Cernadas, E., Barro, S., Amorim, D. (2014): Do we need hundreds of classifiers to solve real world classification problems? Journal of Machine Learning Research 15, 3133–3181, https://jmlr.org/papers/volume15/delgado14a/delgado14a.pdf.

Fotheringham, A. S., Charlton, M., Brunsdon, C. (1996): The geography of parameter space: an investigation of spatial non-stationarity. International Journal of Geographical Information Systems 10(5), 605–627, https://doi.org/10.1080/02693799608902100.

Fotheringham, A. S., Brunsdon, C., Charlton, M. (2003): Geographically weighted regression: the analysis of spatially varying relationships. John Wiley & Sons. Chichester.

Georganos, S., Grippa, T., Gadiaga, A. N., Linard, C., Lennert, M., Vanhuysse, S., Mboga, N., Wolff, E., Kalogirou, S. (2019): Geographical random forests: a spatial

extension of the random forest algorithm to address spatial heterogeneity in remote sensing and population modelling. Geocarto International 36(2), 121–136, https://doi.org/10.1080/10106049.2019.1595177.

Jain, K. V., Pandey, R. P., Jain, M. K. (2015): Spatio-temporal assessment of vulnerability to drought. Natural Hazards 76, 443–469, https://doi.org/10.1007/s11069-014-1502-z.

Kuhn, M., Johnson, K. (2013): Applied predictive modeling. Springer. New York, https://doi.org/10.1007/978-1-4614-6849-3.

Liaw, A., Wiener, M. (2002): Classification and regression by randomForest. Race News 2, 18–22, https://cogns.northwestern.edu/cbmg/LiawAndWiener2002.pdf.

Lugato, E., Lavallee, J. M., Haddix, M. L., Panagos, P., Cotrufo, M. F. (2021): Different climate sensitivity of particulate and mineral-associated soil organic matter. Nature Geoscience 14, 295–300, https://doi.org/10.1038/s41561-021-00744-x.

Magesh, N. S., Chandrasekar, N., Soundranayagam, J. P. (2011): Morphometric evaluation of Papanasam and Manimuthar watersheds, parts of Western Ghats, Tirunelveli district, Tamil Nadu, India: a GIS approach. Environmental Earth Sciences 64, 373–381, https://doi.org/10.1007/s12665-010-0860-4.

Sourav, M., Mishra, A., Trenberth, K. E. (2018): Climate Change and Drought: A Perspective on Drought Indices. Current Climate Change Reports 4, 145–163, https://doi.org/10.1007/s40641-018-0098-x.

Paulik, C., Dorigo, W., Wagner, W., Kidd, R. (2014): Validation of the ASCAT Soil Water Index using in situ data from the International Soil Moisture Network. International journal of Applied Earth Observation and Geoinformation 30, 1–8, https://doi.org/10.1016/j.jag.2014.01.007.

Orgiazzi, A., Ballabio, C., Panagos, P., Jones, J., Fernández-Ugalde, O. (2018): LUCAS Soil, the largest expandable soil dataset for Europe: a review. European Journal of Soil Science 69(1), 140–153, https://doi.org/10.1111/ejss.12499.

Panagos, P., Liedekerke, M. V., Jones, A., Montanarella, L. (2012): European Soil Data Centre: Response to European policy support and public data requirements. Land Use Policy 29(2), 329–338, https://doi.org/10.1016/j.landusepol.2011.07.003.

Panagos, P., Liedekerke, M. V., Borrelli, P., Köninger, J., Ballabio, C., Orgiazzi, A., Lugato, E. (2022): European Soil Data Centre 2.0: Soil data and knowledge in support of the EU policies. European Journal of Soil Science 73(6), e13315, https://doi.org/10.1111/ejss.13315.

Probst, P., Boulesteix, A. L. (2017): To tune or not to tune the number of trees in random forest. Journal of Machine Learning Research 18, 1–18, https://doi.org/10.48550/arXiv.1705.05654.

Probst, P., Wright, M. N., Boulesteix, A. L. (2018): Hyperparameters and tuning strategies for random forest. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery 9(3), e1301, https://doi.org/10.1002/widm.1301.

Mishra, A. K, Singh, V. P. (2010): A review of drought concepts. Journal of Hydrology 391(1–2), 202–216, https://doi.org/10.1016/j.jhydrol.2010.07.012.

Rahmati, O., Falah, F., Dayal, K. S., Deo, R. C., Mohammadi, F., Biggs, T., Moghaddam, D. D., Naghibi S. A., Bui, D. T. (2019): Machine learning approaches for spatial modeling of agricultural droughts in the south-east region of Queensland Australia. Science of the Total Environment 699, 134230, https://doi.org/10.1016/j.scitotenv.2019.134230.

Rahmati, O., Panahi, M., Kalantari, Z., Soltani, E., Falah, F., Dayal, K. S., Mohammadi, F., Deo, R. C., Tiefenbacher, J., Bui, D. T. (2020): Capability and robustness of novel hybridized models used for drought hazard modeling in southeast Queensland, Australia. Science of the Total Environment 718, 134656, https://doi.org/10.1016/j.scitotenv.2019.134656.

Shashank, S., Pandey, A. C. (2015): Delineation of groundwater potential zone in hard rock terrain of India using remote sensing, geographical information system (GIS) and analytic hierarchy process (AHP) techniques. Geocarto International 30(4), 402–421, https://doi.org/10.1080/10106049.2014.894584.

Simpson, E. H. (1951): The interpretation of interaction in contingency tables. Journal of the Royal Statistical Society: Series B (Methodological) 13(2), 238–241, https://doi.org/10.1111/j.2517-6161.1951.tb00088.x.

Thomas, T., Jaiswal, R. K., Galkate, R., Nayak, P. C., Ghosh, N. C. (2016): Drought indicators-based integrated assessment of drought vulnerability: a case study of Bundelkhand droughts in central India. Natural Hazards 81, 1627–1652, https://doi.org/10.1007/s11069-016-2149-8.

Wilhelmi, O. V., Wilhite, D. A. (2002): Assessing vulnerability to agricultural drought: a Nebraska case study. Natural Hazards 25, 37–58, https://doi.org/10.1023/A:1013388814894.

Yang, M., Mou, Y., Meng, Y., Liu, S., Peng, C., Zhou, X. (2020): Modeling the effects of precipitation and temperature patterns on agricultural drought in China from 1949 to 2015. Science of the Total Environment 711, 135139, https://doi.org/10.1016/j.scitotenv.2019.135139.

Related variety and state-sponsored R&D collaboration: a geographical and industrial analysis in Czechia

Petr Horák[1,2,*]

[1] Department of Social Geography and Regional Development, Faculty of Science, Charles University, Czechia
[2] Technology Agency of the Czech Republic, Czechia
[*] Corresponding author: horakpe3@natur.cuni.cz

ABSTRACT

This paper aims to explore the influence of related variety on direct state-supported R&D cooperation across various geographical levels to understand regional performance differentiation and economic base restructuring in Czechia by employing Frenken et al.'s (2007) methodological approach to calculate a related and unrelated variety for all NACE and NACE C-Manufacturing. Findings indicate that the city of Prague has the highest unrelated and related variety, followed by the cities of Brno, Ostrava, and Pilsen. Calculation just for C-Manufacturing changes the ordering significantly. Furthermore, intra-regional and extra-regional pairwise R&D cooperation in joint projects is calculated. The cluster analysis of Czech microregional data (SO ORP) reveals patterns such as emerging collaborators and collaboration powerhouses. Linear regression analyses established a strong positive association between R&D collaboration intensity and related variety, while a negative link was observed with unrelated variety. Similar relationships were observed in the manufacturing sector (NACE-C).

KEYWORDS

related variety; unrelated variety; cluster analysis; Czech microregional data (SO ORP); state-supported R&D collaboration

Received: 21 July 2023
Accepted: 30 October 2023
Published online: 21 November 2023

Horák, P. (2023): Related variety and state-sponsored R&D collaboration: a geographical and industrial analysis in Czechia. AUC Geographica 58(2), 200–213
https://doi.org/10.14712/23361980.2023.15
© 2023 The Author. This is an open-access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0).

1. Introduction

The differentiation of regional performance and the different restructuring of the economic base of the region began to be explained in the last 15 years through the concept of related variety developed by the Dutch school of evolutionary economic geography (Frenken and Boschma 2007; Frenken et al. 2007). The concept follows and to a certain extent overcomes the traditional dual concept of the process of spillover of knowledge between companies and institutions, where on the one hand the advantages resulting from the concentration of a certain industry in space (so-called Marshall-Arrow-Romer externality) and on the other the second benefits resulting from the creation of knowledge spillovers within a diversified economic structure (the so-called J. Jacobs externality). The work of Nooteboom (2000) highlighting the role of cognitive proximity in various spheres of communication and interaction and indirectly in the production process became an equally important source of inspiration.

The related variety allows us to analytically capture the potential for cooperation and knowledge transfer in various geographical units. Moreover, the contribution of related variety to the overall growth and economic development of the region has been documented in existing studies (Frenken et al. 2007). This approach works with the diversity of industries within the region, which are cognitively connected and can maximize the potential of opportunities, growth of existing industries, and the potential of local resources for new industries.

In Czechia, research analysis mapping related / unrelated varieties is very limited. An exception is the research by Květoň and Šafr (2019), who measured unrelated variety and regional embeddedness of interregional and intersectoral relations in Czechia. Blažek et al. (2016) tried to clarify different methods of calculating related variety using the example of Czech R&D projects. Furthermore, geographical and cognitive proximity was clarified in the example of R&D collaborative projects (Květoň et al. 2022), however, these authors did not directly apply the concept of related variety. Therefore, this paper is the first attempt to calculate the unrelated and related variety for the whole of Czechia on regional, district, and municipalities with extended competence levels. The article also provides a partial reflection on the critique of the methodological approach to related variety as presented by Bathelt and Storper (2023).

The overarching goal of the article is to investigate the factors influencing the intensity of R&D collaborations in state-supported projects, focusing specifically on the role of related and unrelated variety across sectors and regions. By conducting multiple linear regression analyses, the article tests a series of hypotheses to elucidate how the diversity and specificity of industries within a region impact the propensity for research and development collaborations.

The article is designed as follows. First, the concept of related and unrelated variety is described, and current knowledge from the Czech environment is also emphasized. Subsequently, the research question is presented. Next, the methodical approach to the measurement of related variety and also to R&D cooperation is described in detail. In the following section, the hypotheses are tested and empirical results are interpreted.

2. Conceptual departures for related variety assessment

Related variety refers to the co-location of different sectors sharing commonalities and complementary competencies, which is conducive to knowledge spillovers underpinning regional growth and innovation (Corradini and Vanino 2022).

Research on related and unrelated variety has been ongoing for several years, with many studies exploring the effects of these concepts on regional growth, innovation, and entrepreneurship. Regarding the related variety and innovation process, previous research suggests that related variety can enhance regional innovation. When industries in a region are cognitively similar and have inter-industry knowledge spillovers, it becomes easier for innovation to occur (Martynovich and Taalbi 2022; Ejdemo and Örtqvist 2022). Moreover, Innocenti et al. (2021) emphasized that local related variety enhances the overall innovation rate and can contribute to recombination or incremental innovation. In terms of technological breakthroughs, related variety would raise the likelihood of innovations in general, but unrelated variety would raise the likelihood of breakthrough innovations, which in themselves are rare (Castaldi et al. 2015).

Furthermore, researchers have explored the role of related variety in regional diversification and path development (Yeung 2020), the relevance of relatedness research in economic diversification and regional competitiveness (Ferraz et al. 2021), the integration of related variety and strategic coupling in understanding regional industrial diversification and economic resilience (Yeon et al. 2022) or the relationship between relatedness, growth, and industry clustering (Bond-Smith and McCann 2019).

These studies have contributed to a deeper understanding of the role of related variety in economic geography research. They have highlighted the importance of relatedness and diversity in economic activities for regional competitiveness, growth, employment, and resilience.

Researchers have also emphasized the need to consider the social, cultural, and institutional dimensions of economic activities and the importance of context sensitivity in economic-geographic theorizing. The current knowledge in the related variety research

provides insights into the complexities and dynamics of economic systems within different spatial contexts.

Researchers use various methods to measure related and unrelated variety in their studies. One methodology used to compute related and unrelated variety is based on entropy measures (e.g. Frenken et al. 2010) and this method is also applied in this paper. However, it is necessary to point out that the whole calculation of kin diversity is to some extent an "ex-ante approach to the evaluation of cooperation and knowledge transfer" (see Blažek et al. 2017), and a higher kin diversity does not guarantee a more effective transfer of knowledge and information but expresses a certain assumption for such cooperation. This methodological approach presupposes that relatively diverse firms are cognitively close enough to understand each other and cooperate, but at the same time far enough away not to compete with each other. Companies and their representatives can therefore "understand" each other and have something to offer, but at the same time, they do not threaten each other on the market.

3. Research question

The central inquiry of this study aims to unravel the complex interplay between regional R&D collaborations within state-supported projects and the potential influence of the related variety within the SO ORP. Utilizing cluster analysis, the research seeks to delineate distinct patterns or clusters of regions based on their R&D collaborative dynamics. The central research question driving this investigation is: "How does R&D collaboration, as manifested by collaborating firms and research organizations in state-supported projects, relate to the related variety in Czech microregions (SO ORP)?"

Following this primary research question and based on the current state of related variety knowledge (e.g. Bathelt and Storper 2023), several hypotheses have been developed to provide a structured approach to addressing the research question:

Hypothesis 1: Based on Květoň and Horák (2018), who clarified the differentiation of R&D capacities at the regional NUTS 3 level in Czechia when subjected to k-means clustering based on relative joint R&D projects, related variety and unrelated variety, the SO ORP (Czech microregions) will yield more than two distinct clusters.

Hypothesis 2: Based on the current state of knowledge about related variety in different countries (Wise and Anderson 2017, Boschma and Iammarino 2009) we expect that the intensity of R&D collaboration in state-supported projects will be positively associated with the related variety.

Hypothesis 3: The intensity of R&D collaboration of state-supported projects in the manufacturing sector (NACE-C) is positively associated with the related variety specific to manufacturing.

4. Methodology

This study draws upon the methodological approach of "Related variety" presented by Frenken et al. (2007). This approach allows us to analytically capture the potential for cooperation and knowledge transfer in the geographical unit. Moreover, the contribution of related variety to the overall growth and economic development of the region has been documented in existing studies (Frenken et al. 2007). This approach works with the diversity of industries within the geographical unit, which is cognitively connected and can maximize the potential of opportunities, growth of existing industries, and the potential of local resources for new industries.

4.1 Data

Underlying data for the assessment of related and unrelated variety are drawn from the Register of Economic Subjects (RES). This source provides information on all economically active entities in Czechia. From this data, it is possible to filter out the legal persons engaged in business, i.e. firms. In this paper, all legal forms of business are selected. From the MagnusWeb database, information is secured on the number of employees of individual firms. Data for cooperation in R&D are drawn from the IS VAVAI database. The related and unrelated variety is calculated for:

1) the full breadth of NACE 2-digit industries
2) for NACE C-Manufacturing.

The results and underlying calculation are published on GitHub to enable further research: https://github.com/ph1559/related-variety/.

4.2 Data limitation

The sources used, despite being the best publicly available, have serious limitations of which the analysis in this paper is aware. First, the number of employees in MagnusWeb may not be available for all firms listed in RES. For this reason, the available sample of data is listed below. Tab. 1 presents all firms with more than one employee by RES compared to the number of available employees from MagnusWeb. The version of RES and number of employees is relative to the year 2021. In the absence of information on the number of employees for 2021, the nearest available value is used. Firms with available data on the number of employees were further used to calculate the unrelated/related variety.

Second, it should also be noted that some larger concerns do not split the number of employees by production facilities. The employers that stand out the most are Škoda Auto, Siemens, Bosch, Honeywell.

Škoda Auto a.s. is classified in the location Mladá Boleslav (CZ020), with the listed number of employees as 35,063. The listed production plants of Škoda Auto a.s. are located in Mladá Boleslav, Kvasiny

Tab. 1 Availability of firm employee data from MagnusWeb.

Region	NUTS 3 region name	Firms with information about the number of employees (MagnusWeb)	Total firms with one or more employees (RES)	Proportion of available employee data
CZ010	Prague	31,863	69,448	45.9%
CZ020	Central Bohemian Region	10,227	17,445	58.6%
CZ031	South Bohemian Region	5,968	9,262	64.4%
CZ032	Pilsen Region	5,579	8,242	67.7%
CZ041	Karlovy Vary Region	2,308	3,645	63.3%
CZ042	Ústí nad Labem Region	6,087	8,797	69.2%
CZ051	Liberec Region	3,645	5,893	61.9%
CZ052	Hradec Králové Region	5,567	7,566	73.6%
CZ053	Pardubice Region	4,834	7,132	67.8%
CZ063	Vysočina Region	4,058	5,914	68.6%
CZ064	South Moravian Region	14,919	25,424	58.7%
CZ071	Olomouc Region	5,543	8,742	63.4%
CZ072	Zlín Region	6,534	9,021	72.4%
CZ080	Moravian-Silesian Region	12,134	17,328	70.0%
Total	All regions	119,266	203,859	58.5%

Source: Own calculations based on data drawn from RES and MAGNUS.

(CZ041) and Vrchlabí (CZ041). The division has approximately 20,000 employees in Mladá Boleslav, 9,000 in Kvasiny and 6,000 in Vrchlabí. Siemens, s.r.o. has 9,691 employees according to MagnusWeb and is divided into seven legal entities. Five of these entities are listed in Prague (CZ010). However, it also has production plants in Brno (CZ064), Drásov (CZ064), Frenštát pod Radhoštěm (CZ072), Trutnov (CZ041), Letohrad (CZ053) and Mohelnice (CZ071). Brno and Drásov has its legal entity and therefore the employee number corresponds to the correct NUTS 3 region.

Next concern is Robert Bosch. There are Robert Bosch subsidiaries in eight cities in Czechia. They have a total of four production plants, one service centre and one logistics warehouse. In total, they are divided into five legal entities. These partly reflect the territorial division of the Group. The Honeywell Group is divided into 4 legal entities in the RES. The biggest shortcoming in this case is that MagnusWeb does not provide the number of employees for its largest entity Honeywell Aerospace s.r.o. However, its spin-off plant in Olomouc (CZ071) is a separate entity. The activities of this concern are still concentrated in Prague (CZ010) and Brno (CZ064).

The data of other larger companies and foreign concerns might be subject to similar problems with the difference between the location of the legal entity and the location of the production facilities. The data are also sensitive to the reporting of agricultural production, which will play a lesser role in the following calculations. These limitations need to be reflected in the interpretation of the results obtained.

Third, information about the CZ-NACE sector is important for the following unrelated and related variety calculations. The main CZ-NACE code is used for the calculation of five and two places. This indicator is assigned by the Czech Statistical Office (CZSO) and is based on the largest reported sales volume of own sales of goods and services, change in inventories of own operations and capitalization. These three items are grouped in the CZSO accounts under one heading of output. The CZ-NACE classification is therefore not an answer to the general classification of an enterprise, but rather a description of its economic activity. For example, Honeywell Aerospace Olomouc s.r.o. has listed the main CZ-NACE 30900 equivalent to C30.9 – Manufacture of transport equipment n.e.c. but is considered for CZ-NACE 30300 with the equivalent of C30.3 – Manufacture of air and spacecraft and related machinery. As the CZ-NACE performance reporting methodology is uniform, it can be expected that similar nuances will be evenly spread across the national sample and thus partially cancel each other out. However, it is imperative that this shortcoming is taken into account when interpreting the results.

Fourth, the data used from IS VaVaI contain only R&D collaborations under direct public support (in the form of collaborative projects). Unfortunately, data for private R&D collaboration are not available and therefore the dataset used is not exhaustive. These limitations will be taken into account when interpreting the results.

4.3 Unrelated and Related Variety Calculation

In the first step, an unrelated diversity index was calculated using the provided formulas by Frenken et al. (2007). The following calculations, P_S represents the

share of employees in industry S (section) compared to the total number of employees Z in the territorial unit i in period t. For P_S NACE industries are used in two places.

$$P_s = (\frac{Z_S}{Z})$$

S takes the values of all five-digit NACE sector codes.

$$UVar_i = \sum_{s=1}^{S} P_s (\frac{1}{P_s})$$

UVar is the resulting value of unrelated variety for the geographical unit. The temporal aspect (the year to which the unrelated diversity value relates) is not considered in this view. The latest available employment data is used (the most common year is 2021).

In the next step, the related variety index was calculated according to the formulas below:

$$RVar_i = \sum_{s=1}^{S} P_s H_s$$

and

$$H_s = \sum_{d \in S_S} \frac{p_i}{P_s} (\frac{1}{\frac{p_i}{P_S}})$$

Where p_i represents the proportion of NACE employees per five locations relative to the total number of employees in a geographical unit. *RVar* equations are used for obtaining related variety indexes for regions, districts and SO ORP. P_S and H_S are calculated separately for each geographic level.

The subsequent results of related and unrelated diversity are comparable only at the same level of the territorial unit due to the nature of the calculation. Subsequently, the related and unrelated variety is calculated for three samples of enterprises according to the main NACE sector indicated: firstly, the calculation was carried out for the whole range of NACE sectors, from Section A – Agriculture, forestry, and fishing to Section S – Other activities. Furthermore, the calculation was performed on the NACE range falling within NACE section C – Manufacturing.

4.4 R&D cooperation

To measure the amount of R&D cooperation in a geographic context four indicators are calculated:

1) Internal R&D cooperation within the region.
2) External R&D cooperation outside of the region.
3) Internal R&D cooperation within the region taking into consideration only firms.
4) External R&D cooperation outside of the region taking into consideration only firms.

The focus on firms is done by subsetting the dataset only for firms. Cooperation between organizations taken into account is within the years 2006–2021. Internal R&D cooperation is calculated as the number of two or more firms in the same region in a collaborative research project. External R&D cooperation is calculated as the number of projects with cooperation outside of the region. The point for the project is granted to all participating regions.

4.5 Cluster analysis

Before initiating the clustering process, the data was subject to a preliminary examination to ensure it was suitable for cluster analysis. Any missing values were addressed, and potential outliers were either rectified or justified. The variables were also normalized to ensure equal weightage during clustering. Normalization was achieved using the min-max scaling method, which transforms the data into a range between 0 and 1, ensuring that each variable contributes equally to the clustering process (Virmani, Taneja, Malhotra, 2015). The formula for min-max scaling is given by:

$$normalized(x) = \frac{x - \min(x)}{\max(x) - \min(x)}$$

In the article's cluster analysis methodology, three fundamental metrics are used for clustering: Joint state-supported projects, related variety, and unrelated variety. It was observed that there's a pronounced correlation between related variety, unrelated variety, and the number of project collaborations to the number of employees. Therefore, to avoid potential biases, these variables are adjusted by dividing them by the latter metric. Without such a modification, the clustering could inadvertently emphasize primarily the population size of SO ORPs, rather than the intended nuances of the regions.

To address the related and unrelated diversity bias mentioned by Bathelt and Storper (2023), the derived metrics are divided by a number of employees. This reduces the importance of large cities and towns in favor of microregions with higher related and unrelated diversity per employee.

Three primary metrics were chosen for clustering on the level of SO ORP:

$$\frac{joint\ projects_r}{number\ of\ employees_r} \quad (1)$$

$$\frac{related\ variety_r}{number\ of\ employees_r} \quad (2)$$

$$\frac{unrelated\ variety_r}{number\ of\ employees_r} \quad (3)$$

The underscored r signifies the Czech microregion (SO ORP). The appropriate number of clusters determined using the Elbow Method was 3. This involved running the k-means clustering on the dataset for a

Tab. 2 Regional unrelated and related variety for Czechia.

Region code	NUTS 3 region name	Unrelated variety	Related variety	Unrelated variety only NACE-C	Related variety only NACE-C
CZ010	Prague	2.028	0.535	0.674	0.245
CZ020	Central Bohemian Region	0.819	0.191	0.778	0.252
CZ031	South Bohemian Region	0.451	0.090	0.471	0.142
CZ032	Pilsen Region	0.461	0.090	0.508	0.145
CZ041	Karlovy Vary Region	0.183	0.027	0.181	0.040
CZ042	Ústí nad Labem Region	0.490	0.109	0.493	0.165
CZ051	Liberec Region	0.317	0.058	0.401	0.094
CZ052	Hradec Králové Region	0.404	0.081	0.454	0.137
CZ053	Pardubice Region	0.400	0.076	0.442	0.125
CZ063	Vysočina Region	0.354	0.075	0.453	0.137
CZ064	South Moravian Region	0.869	0.221	0.707	0.274
CZ071	Olomouc Region	0.400	0.082	0.443	0.141
CZ072	Zlín Region	0.492	0.121	0.539	0.201
CZ080	Moravian-Silesian Region	0.812	0.181	0.771	0.268

Source: Own calculations based on data drawn from RES and MAGNUS.

range of values of k (e.g., k from 1 to 10), and then for each value of k computing the sum of squared distances from each point to its assigned centre. The 'elbow' of the curve represents an optimal value for k (a balance between precision and computational cost) (MacQueen 1967). In the article, the final number of clusters is selected to be 5 to better represent the granularity of Czech microregions (SO ORP).

With the selected value of k, k-means clustering was applied to the dataset using the chosen metrics. The k-means algorithm seeks to minimize the squared sum of Euclidean distances from the mean of each cluster (Ismkhan 2017). The iterative algorithm divides the microregions into k clusters based on the similarity in their R&D collaborative dynamics. The k-means clustering was executed using the following command in R:

```
kmeans(data_normalized, centers = k_optimal,
nstart = 25)
kmeans_result=kmeans(data_normalized, centers =
k_optimal, nstart = 25)
```

The nstart parameter ensures that the algorithm is initialized multiple times to avoid local optima (Hartigan and Wong, 1979).

5. Initial results

The results of calculations in the previous part are structured as follows: First unrelated and related variety for NUTS 3 regions, districts, and microregions, and second internal and external cooperation for the same geographical units are presented.

5.1 Unrelated and Related Variety in NUTS 3 regions, districts, and microregions

The main NUTS 3 region that dominates in unrelated and related variety is Prague, the capital of Czechia. Because Prague is a capital city many firms have domicile there even though most of their employees and production is located elsewhere. The results demonstrate its greatest general diversity in the nomenclature of economic activities in both related and unrelated fields. The second place is usually occupied by the second largest city, Brno, and its surrounding NUTS 3 region, the South Moravian Region. Significant differences in ranking become apparent when evaluating related varieties and unrelated varieties for the NACE-C manufacturing only.

5.2 Regional unrelated and related variety

Tab. 2 shows the related and unrelated variety dominance of the three NUTS 3 regions where the three largest cities in Czechia are located: Prague, South Moravian Region, and Moravian-Silesian Region. Following this table further data are visualized as cartograms. The change in order can be seen when only manufacturing (NACE-C) is taken into consideration.

For related variety only in manufacturing (NACE-C), Prague lags behind the South Moravian Region, Moravian-Silesian Region, and even the Central Bohemian Region, but it is not surprising. These regions show higher nomenclature specialization. It shows that even the economically most important NUTS 3 region in Czechia (Prague) may be most diverse in terms of broad nomenclature (NACE) but not in terms of industry (Fig. 1).

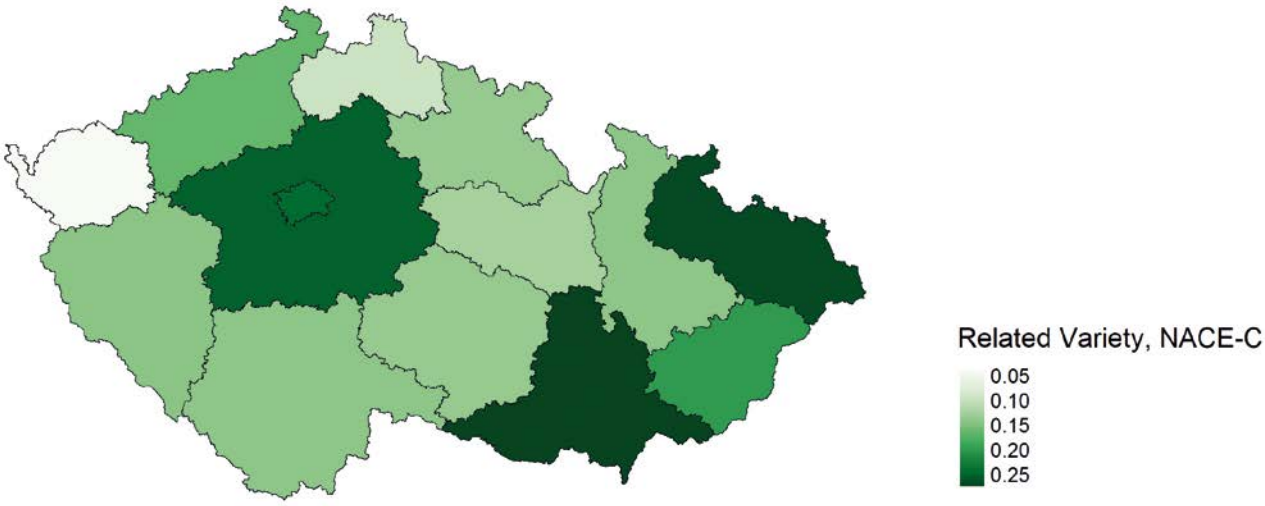


Fig. 1 Related variety only for manufacturing (NACE-C) in Czechia.

5.3 Unrelated and related variety in districts of Czechia

On the level of districts (okres) in Czechia, Prague shows the highest variety in all measured aspects. In the Tab. 3, the first 10 districts by related variety of manufacturing (NACE-C) are shown. The highest-scoring districts are cities and towns.

Importantly the industrial districts also show higher unrelated variety when only NACE-C is considered. The outlier that wasn't particularly visible in other measurements is the district of Mladá Boleslav which hosts large car manufacturing capacities. With a general unrelated variety of 0.171, the unrelated variety only for NACE-C is 0.245, contrary to this the related variety is 0.012 and if only manufacturing is taken into consideration it is 0.019 (Fig. 2).

5.4 Unrelated and related variety on the Czech microregional level

The lowest presented geographic level of Czechia in this paper is a municipality with extended powers (SO ORP). There are 206 such units in Czechia. The unrelated and related variety follows the expected trend. The capital Prague scores highest and Brno. Ostrava and Pilsen follow it. In unrelated variety of manufacturing Mladá Boleslav holds 2nd place. That doesn't correspond to its size in the population (19th). Fig. 3 also shows that in terms of industry, the spots of nearly zero related variety are not located along borders. This unexpected phenomenon may lead one to think about the location of the inner peripheries. Furthermore, strong SO ORPs in terms of related variety in manufacturing are often bordered by SO ORPs that

Tab. 3 Related and unrelated variety of Czech districts, first 10 ordered by Related variety NACE-C.

District code	District name	Unrelated variety	Related variety	Unrelated variety NACE-C	Related variety NACE-C
CZ0100	Praha	2.028	0.535	0.674	0.245
CZ0642	Brno-město	0.518	0.099	0.292	0.082
CZ0724	Zlín	0.222	0.044	0.229	0.069
CZ0806	Ostrava-město	0.363	0.055	0.238	0.049
CZ0534	Ústí nad Orlicí	0.132	0.020	0.197	0.045
CZ0323	Plzeň-město	0.242	0.036	0.236	0.044
CZ0804	Nový Jičín	0.136	0.021	0.215	0.043
CZ0723	Vsetín	0.142	0.021	0.185	0.041
CZ0643	Brno-venkov	0.152	0.025	0.174	0.038
CZ0722	Uherské Hradiště	0.137	0.022	0.162	0.037
CZ0207	Mladá Boleslav	0.171	0.012	0.245	0.019

Source: Own calculations based on data drawn from RES and MAGNUS.

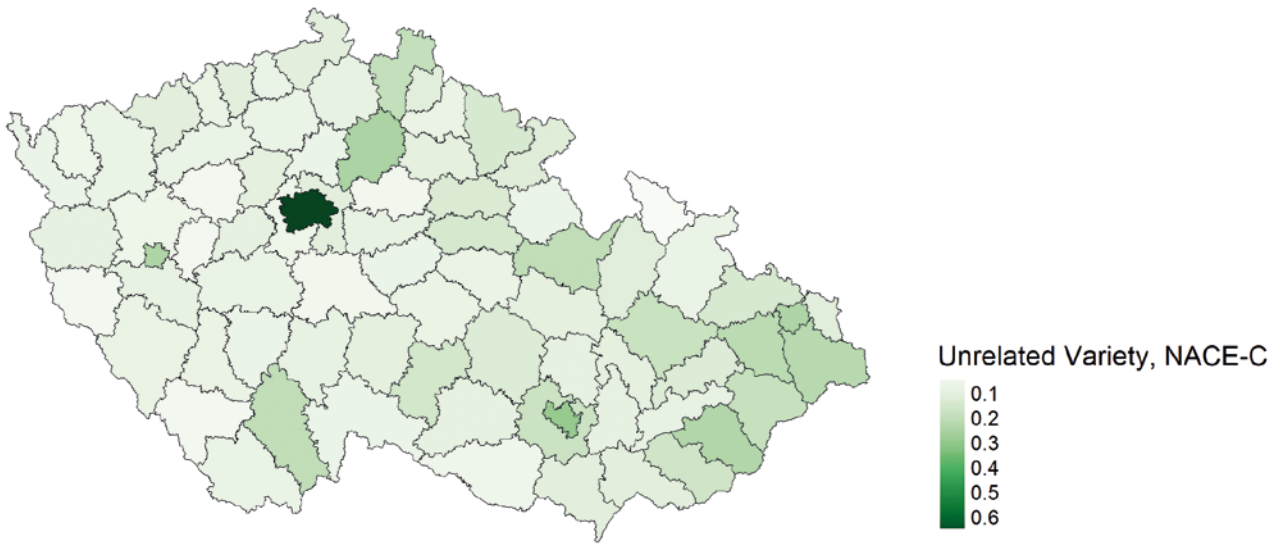


Fig. 2 Unrelated variety only for manufacturing (NACE-C).

have no related variety in manufacturing. This offers room for further research and discussion of inner peripheries.

5.5 R&D Cooperation in NUTS 3 regions, districts and municipalities

To measure R&D and collaboration in the NUTS 3 regions the CEP database of collaborative projects is used. The CEP data contains partially state-supported R&D projects. Some of them are joint projects of R&D collaboration. Project data for supported projects with collaboration that started in the years 2006–2022 are used. These are 12,577 unique projects with collaborations with a total of 3,852 unique organizations, including 3,006 unique firms. Together these projects represent support from the state budget of 155.65 billion CZK, which is 56.27% of all projects for the same period in CEP. It is 28.11% of the public budget dedicated to R&D in 2006–2022 and about 11.3% of total R&D expenditure in Czechia (GERD). Thus, the scope of the analysis of R&D cooperation is limited to this slice of approximately 11.3% of the total R&D expenses. In the article, R&D cooperation is examined using the CEP dataset and a pairwise collaboration methodology. This approach identifies every possible two-region combination in which organizations are jointly engaged in a project. Every organization, including each faculty as a distinct entity, is taken into account. Collaborations are classified into those occurring within the same region (intra-regional) and those bridging different regions (extra-regional). This

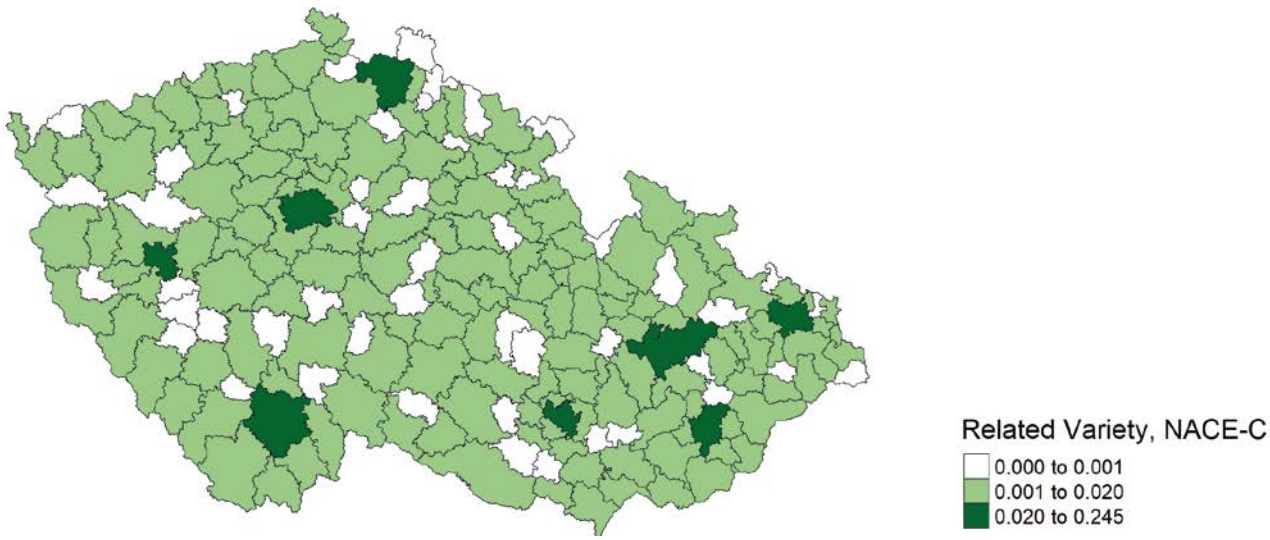


Fig. 3 Related variety only for manufacturing (NACE-C) in Czech municipalities with extended powers.

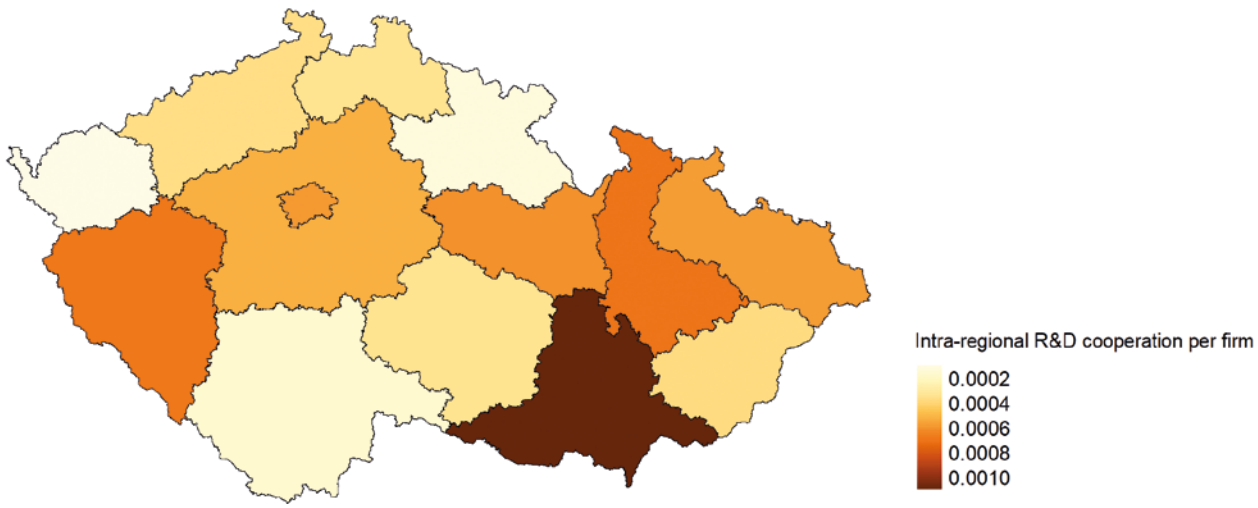


Fig. 4 R&D cooperation of all organizations.

method offers a thorough insight into the landscape of regional R&D collaborations. The values are calculated at three geographic levels: a) NUTS 3 region b) district c) Czech microregions (SO ORP). 75,774 participations with collaboration is observed. Of these, 28,474 (37.58%) are intra-regional and 47,300 (62.42%) are extra-regional. The collaborations that did not leave the Prague borders (Prague-Prague) are 23.46% of the direct collaborations. In the case of only firm-firm cooperations, we observe 15,200 direct links between firms within projects. Of these, 2,964 (19.50%) are intra-regional and 12,236 (80.50%) are extra-regional. The collaborations that did not leave the Prague (Prague-Prague) border account for 10.11% of the direct collaborations. Next, we conduct pairwise connections for organizations that are not in the set of firms and are in RVVI's list of research organizations. For these research organizations only, we observe 11,688 direct connections of which 5,138 (43.95%) are intra-regional and 6,550 (56.04%) extra-regional. Direct Prague-Prague connections accounted for 35.06% of the connections (Fig. 4).

6. Hypothesis testing

6.1 Hypothesis 1

To test the first hypothesis: "Based on Květoň and Horák (2018), who clarified the differentiation of R&D capacities at the regional NUTS 3 level in Czechia when subjected to k-means clustering based on relative joint R&D projects, related variety and unrelated variety, the SO ORP (Czech microregions) will yield more than two distinct clusters." cluster analysis was employed on the provided data. Cluster analysis groups data points into clusters so that data points in the same cluster are more similar to each other than to those in other clusters.

In the article, k-means clustering was applied to segment SO ORP regions using three primary metrics: shared state-funded projects, unrelated variety and related variety.

Given the significant variance between clusters in terms of the three metrics, it can be concluded that distinctive patterns indeed emerge among SO ORP regions when characterized by their R&D collaborative dynamics in terms of shared state-funded projects, related, and unrelated variety. The appropriate number of clusters determined using the k-means (Elbow Method) was 3. Therefore, based on this cluster analysis, hypothesis H1 fails to be rejected. The results indicate that distinctive patterns (clusters) emerge, aligning with the hypothesis's premise. The results indicate that distinctive patterns (clusters) emerge, aligning with the hypothesis's premise, with the chosen five clusters providing more granularity and detail in understanding the distinctive patterns in the R&D collaborative dynamics within the Czech microregions:

Cluster 1: Microregions classified under "Emerging Collaborators'' present a promising picture. They show a higher related variety over unrelated variety when metrics are divided by the number of employees. The average related variety divided by number of employees can compete with larger and sophisticated microregions such as Prague. These microregions are also involved in a commendable number of state-funded projects. They cater to a moderate population and have a substantial employment base, positioning them as areas that are budding and showing promise in their collaborative endeavors.

Cluster 2: "Collaboration Powerhouses". The microregions classified under this cluster are characterized by a harmonised interplay of related and unrelated variety metrics, especially when contextualised against the number of employees. What

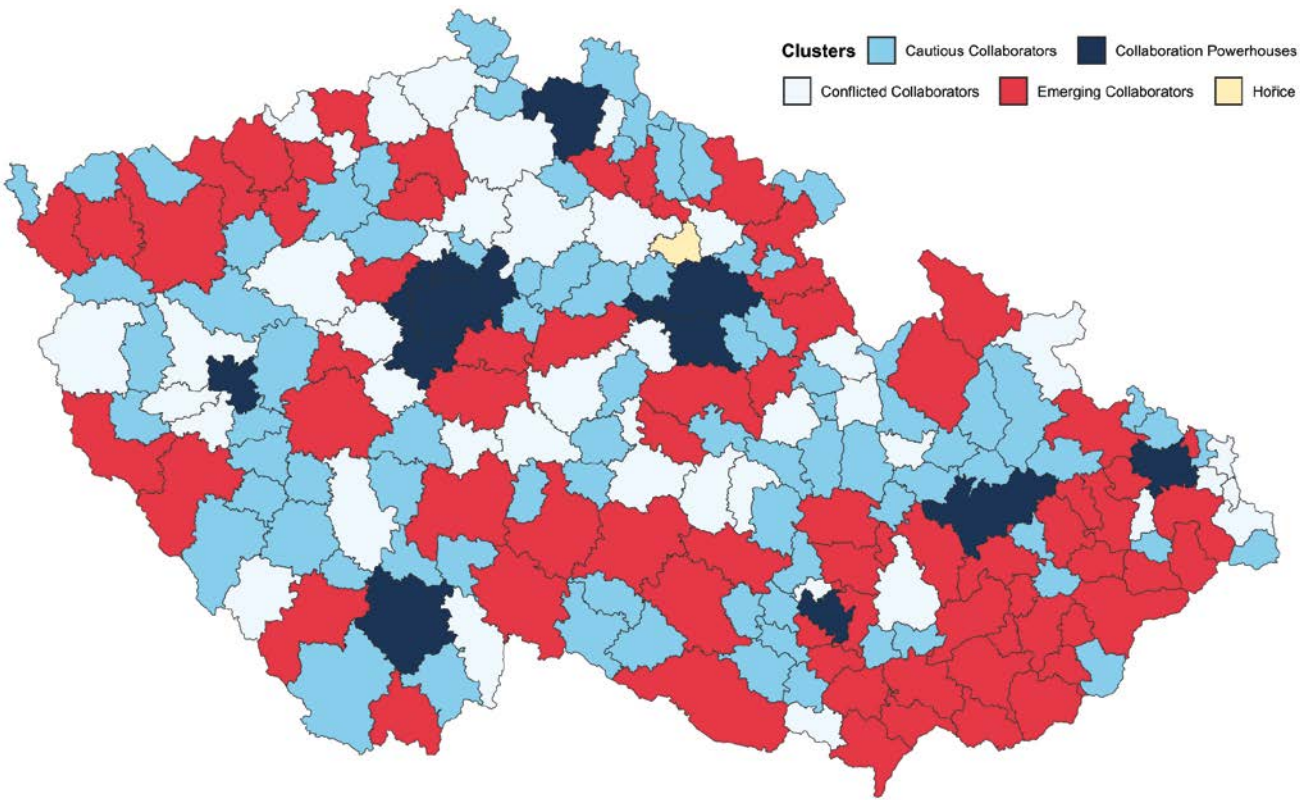


Fig. 5 Clustering of Czech microregions (SO ORP) based on related variety and direct state-supported R&D cooperation in Czechia.

sets them apart is the impressive level of collaboration per employee. With a pronounced related variety, these microregions have participated in a substantial number of state-funded projects. They hold a significant population and are supported by a large workforce. These microregions can aptly be described as the nexus of collaborative activities, making them true collaboration epicenters.

Cluster 3: "Hořice". This microregion did fall into one unique cluster as there is a research organization "Výzkumný a šlechtitelský ústav ovocnářský Holovousy s.r.o." that was supported by 128 projects and within this project created 218 extra-regional and 26 intra-regional connections.

Cluster 4: "Cautious Collaborators". Microregions in this cluster are low in related variety and collaborations, particularly when metrics are proportioned against the number of employees. Their absolute related variety and involvement in state-funded projects are also very low. The SO ORPs are inhabited by a smaller population and supported by a moderate workforce. They have established a modest footprint in the collaboration arena and have room to explore further synergies.

Cluster 5: "Conflicted Collaborators". Microregions in this cluster present an intriguing dichotomy. Despite their relatively low related variety when adjusted for the number of employees, they exhibit a heightened engagement in state-funded R&D projects. This suggests a distinct focus on select, specialized areas of expertise or perhaps a concentration of knowledge within certain domains. The participation rate in collaborations remains consistently high, indicating an active pursuit of partnerships and shared initiatives. The SO ORPs in question have a moderate population. These microregions are navigating a path that, while conflicted between specialized knowledge and broad collaboration, holds the potential for unique growth trajectories (Fig. 5).

6.2 Hypothesis 2

"Based on the current state of knowledge about related variety in different countries (Wise and Anderson 2017, Boschma and Iammarino 2009) we expect that the intensity of R&D collaboration in state-supported projects will be positively associated with the Related Variety."

In the article, a linear regression analysis was conducted to examine the relationship between the intensity of R&D collaborations and the related variety among SO ORP. The linear model results suggest a strong positive association between related variety and the intensity of R&D collaborations. In contrast, unrelated variety exhibited a significant negative relationship with R&D collaboration intensity. Additionally, certain clusters, population density, and the total number of employees in a microregion also influenced R&D collaborations, though not all were

Tab. 4 Linear regression of R&D collaboration and related variety.

Variable	Estimate	Standard Error	p-value	Significance
(Intercept)	198.900	51.78160	0.00017	***
Related variety	98338.300	11719.52476	0	***
Unrelated variety	−17679.100	1619.86162	0	***
Employees	0.029	0.00784	0.00031	***
cluster 1	−139.200	52.73475	0.00897	**
cluster 2	503.700	127.00239	0.00010	***
cluster 3	90.100	233.80022	0.70026	
cluster 4	−104.100	49.48022	0.03673	*
popdens	0.130	0.13156	0.31688	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 229.3 on 197 degrees of freedom
Multiple R-squared: 0.9955, Adjusted R-squared: 0.9953
F-statistic: 5407 on 8 and 197 DF, p-value: < 2.2e−16
Source: Author.

Tab. 5 Linear regression of R&D collaboration and related variety in manufacturing.

Variable	Estimate	Standard Error	p-value	Significance
(Intercept)	16.5000	4.750	0.00064	***
Related variety NACE-C	5171.2000	248.200	< 2e−16	***
Unrelated variety NACE-C	−1097.5000	185.500	1.43e−08	***
Employees in NACE-C	0.0055	0.001	0.00001	***
cluster1	−13.7000	3.900	0.00062	***
cluster2	−22.2000	8.700	0.01144	*
cluster3	−11.9000	19.200	0.53702	
cluster4	−6.6000	4.100	0.10548	
popdens	−0.0180	0.010	0.07771	.

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 18.79 on 197 degrees of freedom
Multiple R-squared: 0.9414, Adjusted R-squared: 0.939
F-statistic: 395.5 on 8 and 197 DF, p-value: < 2.2e−16
Source: Author.

statistically significant. The model captures approximately 99.55% of the variation in total cooperations. Given these findings, the article concludes that the data supports the hypothesis, emphasizing the role of related industries in fostering R&D collaborative dynamics across microregions. Based on these results, the article fails to reject hypothesis 2, affirming that microregions with higher related variety tend to have intensified R&D collaborations (Tab.4).

Multiple Linear Regression Equation 1:
R&D pairwise cooperation
$= \beta_0 + \beta_1 RV + \beta_2 UV + \beta_3 EMPL + \beta_4 C1 + \beta_5 C2 + \beta_6 C3 + \beta_7 C4 + \beta_8 C5 + \beta_9 Popdens + \in$

6.3 Hypothesis 3

"The intensity of R&D collaboration of state-supported projects in the manufacturing sector (NACE-C) is positively associated with the related variety specific to manufacturing."

In the article, a linear model is employed to test the second hypothesis that investigates the relationship between the intensity of R&D collaboration in manufacturing (NACE-C) and the related variety specific to manufacturing. This model takes into account various control variables, integrating factors such as unrelated variety, clusters, population density, and the total number of employees in the manufacturing sector. Through linear regression, the model provides a robust statistical framework to determine how the diversity of manufacturing activities, both related and unrelated, along with microregional characteristics, influences collaborative R&D efforts in the sector (Tab. 5).

Multiple Linear Regression Equation 2:
R&D pairwise cooperation NACE-C
$= \beta_0 + \beta_1 RV_{NACE-C} + \beta_2 UV_{NACE-C} + \beta_3 EMPL_{NACE-C} + \beta_4 C1 + \beta_5 C2 + \beta_6 C3 + \beta_7 C4 + \beta_8 C5 + \beta_9 Popdens + \in$

This assertion is underlined by the positive and highly significant coefficient for the variable related variety for NACE-C (manufacturing). As the related variety specific to manufacturing increases, the intensity of R&D collaboration in the sector also witnesses a marked increase.

Notably, while the related variety presents a positive relationship with collaboration intensity, the unrelated variety for NACE-C displays a negative and significant relationship. This suggests that a higher unrelated variety in the manufacturing sector could act as a detriment to the intensity of R&D collaborations. The observed negative association might indicate that when activities are too diversified or unrelated in a microregion, it becomes challenging to find common ground or mutual benefits, thereby reducing collaborative endeavors even in partially state-funded R&D projects.

Considering control variables, it's evident that certain clusters, notably cluster1 (Emerging Collaborator) and cluster2 (Collaboration Powerhouses), exhibit a negative statistical relationship with R&D collaborations in manufacturing. This suggests that SO ORPs belonging to these clusters might have some inherent characteristics or challenges impeding collaboration. Conversely, the total number of employees in the NACE-C shows a positive and significant relationship with collaboration intensity, pointing to the fact that microregions with a larger workforce in manufacturing have heightened collaborative activities.

Lastly, the population density presents a marginally negative influence on collaborations. This might imply that in densely populated areas, the nature of industrial activities could be more fragmented or diverse, possibly diluting the intensity of focused R&D collaborations in manufacturing.

7. Discussion of empirical findings

In this section, the article delves into the influence of related variety on regional development, particularly within the context of collaboration in state-supported R&D projects. This analysis is contrasted with established research, highlighting both similarities and differences in approach and findings. Notably, Frenken et al. (2007), Boschma and Iammarino (2009), and Boschma, Minondo, and Navarro (2013) all underscored that regions with a pronounced related variety tend to witness enhanced employment growth. This observation, while aligning with the broader theme of this article, diverges in its primary approach and objectives. While these studies primarily focused on employment growth as a direct outcome of related variety, this article pivots towards understanding the dynamics of collaboration within the context of related variety. In this topic, Ebersberger, Herstad, and Koller (2014) further explored the connection between regional knowledge bases, collaborations, regional technological specialization and related variety. The specialization reduced domestic collaborations, while related technological variety bolstered international innovation ties.

The sector-specific effects of related variety, as highlighted by Bishop and Gripaios (2010) and Hartog, Boschma, and Sotarauta (2012), offer another dimension of comparison. These studies suggest that the influence of related variety can vary significantly across different sectors. In the context of this article, the use of NACE-C (manufacturing) classifications provides a broader lens to dissect the specific and industrial part of the economy. Bishop and Gripaios (2010) emphasized the potential oversimplification of broadly categorizing sectors into manufacturing and services. They argue that these sectors, in their essence, are heterogeneous, leading to varied mechanisms and extents of spillovers between them. Driven by this perspective, they employed a more granular approach, examining employment growth in individual 2-digit sectors. Given this critique by Bishop and Gripaios, it seems prudent for future research to delve deeper into the manufacturing NACE-C classification, breaking it down further into specific 2-digit sectors for a more nuanced understanding.

The results of the hypothesis testing, when viewed through the lens of economic geography and regional development literature, offer a nuanced understanding of R&D collaborative dynamics in SO ORP microregions. The summarising work by Content and Frenken (2016) underscores the importance of related variety in economic development such as employment growth. Their comprehensive literature review suggests that regions with a diverse yet related set of industries tend to exhibit higher levels of innovation (observed through labour productivity) and economic growth. The critique by Bathelt and Storper (2023) on measuring related and unrelated variety as entropy which leads to a strong statistical link between related variety, unrelated variety, and the population of a (micro)region, is addressed by controlling for these factors and using the number of employees as a denominator.

On the other note, the observed negative relationship between unrelated variety and R&D collaboration intensity in our results resonates with the foundational principles of the related variety literature. As highlighted by Marek and Blažek (2016) and Květoň, Novotný, Blažek, Marek (2022), a region with technologically related industries often benefits from enhanced knowledge spillovers, learning, and growth. However, when activities become overly diversified or unrelated, it can pose challenges in finding synergies, potentially reducing collaborative endeavors. This perspective aligns with the argument that spatial externalities are most potent among firms with related but distinct knowledge. Yet, it's crucial to consider the insights from Grillitsch et al. (2018), who emphasize the potential of unrelated diversification in fostering radical innovations, especially in regions with strong human capital.

In light of the findings from the article's hypothesis testing and the insights from Květoň et al. (2022), it becomes evident that the dynamics of R&D collaboration in SO ORP microregions are multifaceted. The observed negative relationship between unrelated variety and R&D collaboration intensity underscores the challenges of excessive diversification in hindering synergies. Květoň et al. (2022) further illuminate this by revealing that while R&D collaborations often span large cognitive distances, they are not arbitrary. Firms, in their pursuit of innovation, tend to collaborate with partners that, although unrelated, share closer cognitive proximity than other potential collaborators. This intricate balance between diversification and the quest for synergies is further complicated by the geographical dynamics, as seen in the predominant inter-regional linkages in Czech regional innovation systems.

8. Conclusions

In this paper, the concepts of unrelated variety and related variety were introduced and empirically analyzed, first separately and then for manufacturing (NACE-C). Furthermore, concepts of intra-regional and extra-regional R&D cooperation within state-supported joint projects were introduced and assessed.

They are divided into R&D cooperation of firms only, research organizations only and then all pairwise connections for all types of organizations. The mentioned indicators were calculated at three geographical levels, NUTS 3 regional, district, and microregional.

The indicators yield the following results: Prague and the South Moravian Region ranked first in unrelated and related variety. In the case of the related variety of the NUTS 3 region in industry (NACE-C) Prague drops out of the first place. Analysis of unrelated variety at the district level for industry (NACE-C) shows that districts with a large industrial presence tend to have higher unrelated variety rather than related variety. Microregional level analysis shows that municipalities with the highest related variety for industry (NACE-C) are often adjacent to municipalities with almost zero related variety. Such results can be further explored in detail and built on the research of the inner peripheries of Czechia. In the case of the number of collaborations counted by joint R&D projects, it is evident that although Prague has the highest number of absolute collaborations, in a relative view (number of collaborations divided by the number of companies in the region) Prague falls into the background and the highest values are reported by the Pilsen Region, South Moravian Region and Pardubice Region.

A final comparison of these indicators with each other at the regional level shows a high degree of correlation between R&D cooperation and the general unrelated and related variety, but lower only in the case of the unrelated and related variety for industry (NACE-C). The research question of the article is elaborated in three hypotheses. The first hypothesis posited that SO ORP regions, when characterized by their R&D collaborative dynamics in terms of shared state-funded projects and related and unrelated variety, would show distinctive patterns. To investigate this, k-means clustering was applied to group the SO ORP regions based on shared state-funded projects and related variety. After identifying five distinct clusters, the analysis revealed considerable differences between these clusters based on the given metrics. Consequently, the hypothesis was not rejected, suggesting that distinctive patterns were indeed evident among the regions.

Among the clusters identified, a standout group was labeled "Collaboration Powerhouses". These SO ORP microregions were distinctive due to their dominant related variety, especially when adjusted for the number of employees. Their collaboration intensity per employee was also noteworthy. They possess a significant population and are backed by a substantial workforce, these microregions stood out as central hubs of collaborative activities, solidifying their reputation as true epicenters of collaboration. Next "Emerging Collaborators" microregions showed promise with a commendable number of state-funded projects and a higher related variety, hinting at their potential growth. In contrast, "Hořice" was a singular microregion due to a unique research organization significantly supported by 128 R&D state-supported projects. The "Cautious Collaborators" with both related variety and collaborations being modest, suggesting they might be lacking their footing in the R&D landscape. Lastly, "Conflicted Collaborators" displayed an interesting dichotomy, showing potential for unique growth trajectories while navigating a balance between specialized knowledge and broad collaboration.

The second hypothesis suggested a positive association between the intensity of R&D collaboration in state-supported projects and the related variety. A linear regression analysis was conducted to explore this relationship. The outcome showcased a robust positive correlation between related variety and the intensity of R&D collaborations. Conversely, unrelated variety had a significant negative relationship with R&D collaboration intensity. Some other variables, like certain clusters, population density, and the total number of employees in a microregion, also influenced the collaborations, though not all significantly. Overall, the data provided strong support for the hypothesis.

For the third hypothesis, it was proposed that the intensity of R&D collaboration in state-supported projects in the manufacturing sector (NACE-C) would be positively correlated with the related variety specific to manufacturing. A linear regression model was employed, factoring in several control variables. The results underscored a significant positive relationship between the related variety in manufacturing and the intensity of R&D collaborations. However, a notable discovery was that a higher unrelated variety in manufacturing is negatively associated with the intensity of R&D collaborations.

Above all, the study identifies clear clusters within the SO ORP microregions based on their collaborative tendencies and sector closeness. Furthermore, a distinct positive relationship emerges between the intensity of related variety, also in the manufacturing sector, and the extent of cooperation within state-supported joint projects. Interestingly the unrelated variety relates to the extent of cooperation negatively.

References

Bathelt, H., Storper, M. (2023): Related Variety and Regional Development: A Critique, Economic Geography, https://doi.org/10.1080/00130095.2023.2235050.

Bishop, P., Gripaios, P. (2010): Spatial Externalities, Relatedness and Sector Employment Growth in Great Britain. Regional Studies 44(4), 443–454, https://doi.org/10.1080/00343400802508810.

Blažek, J., Marek, D., Květoň, V. (2016): The Variety of Related Variety Studies: Opening the Black Box of Technological Relatedness via Analysis of Inter-firm R&D

Cooperative Projects. Papers in Evolutionary Economic Geography (online), http://econ.geo.uu.nl/peeg/peeg1611.pdf.
Bond-Smith, S., McCann, P. (2019): A Multi-sector Model of Relatedness, Growth and Industry Clustering. Journal of Economic Geography 5(20), 1145–1163, https://doi.org/10.1093/jeg/lbz031.
Boschma, R., Iammarino, S. (2009): Related Variety, Trade Linkages, and Regional Growth in Italy. Economic Geography 85(3), 289–311, https://doi.org/10.1111/j.1944-8287.2009.01034.x.
Boschma, R., Minondo, A., Navarro, M. (2013): The Emergence of New Industries at the Regional Level in Spain: A Proximity Approach Based on Product Relatedness. Economic Geography 89(1), 29–51, https://doi.org/10.1111/j.1944-8287.2012. 01170.x.
Castaldi, C., Frenken, K., Los, B. (2015): Related Variety, Unrelated Variety and Technological Breakthroughs: An analysis of US State-Level Patenting. Regional Studies 49(5), 767–781, https://doi.org/10.1080/00343404.2014.940305.
Corradini, C., Vanino, C. (2022): Path dependency, regional variety and the dynamics of new firm creation in rooted and pioneering industries, Journal of Economic Geography 22(3), 631–651, https://doi.org/10.1093/jeg/lbab021.
Ebersberger, B., Herstad, S. J., Koller, C. (2014): Does the composition of regional knowledge bases influence extra-regional collaboration for innovation? Applied Economics Letters 21(3), 201–204, https://doi.org/10.1080/13504851.2013.848019.
Frenken, K., Boschma, R. (2007): A theoretical framework for Evolutionary Economic Geography: Industrial dynamics and urban growth as a branching process. Journal of Economic Geography 7(5), 635–649, https://doi.org/10.1093/jeg/lbm018.
Frenken, K., van Oort, F., Verburg, T. (2007): Related variety, unrelated variety and regional economic growth. Regional Studies 41(5), 685–697, https://doi.org/10.1080/00343400601120296.
Grillitsch, M., Asheim, B., Trippl, M. (2018): Unrelated knowledge combinations: the unexplored potential for regional industrial path development. Cambridge Journal of Regions, Economy and Society 11(2), 257–274, https://doi.org/10.1093/cjres/rsy012.
Hartigan, J. A., Wong, M. A. (1979): Algorithm AS 136: A K-Means Clustering Algorithm. Journal of the Royal Statistical Society. Series C (Applied Statistics) 28(1), 100–108, https://doi.org/10.2307/2346830.
Hartog, M., Boschma, R., Sotarauta, M. (2012): The Impact of Related Variety on Regional Employment Growth in Finland 1993–2006: High-Tech versus Medium/Low-Tech. Industry and Innovation 19 (6), 459–476, https://doi.org/10.1080/13662716.2012.718874.
Ismkhan, H. (2017): An initialization method for the k-means using the concept of useful nearest centers. Journal of Advanced Engineering Research and Science, https://doi.org/10.48550/arXiv.1705.03613.
Květoň, V., Novnotný, J., Blažek, J., Marek, D. (2022): The role of geographic and cognitive proximity in knowledge networks: The case of joint R&D projects. Papers in Regional Science 101(2), 351–372, https://doi.org/10.1111/pirs.12656.
Květoň, V., Šafr, K. (2019): Regional embeddedness, relatedness and inter-regional linkages among less developed regions in Central Europe. European Planning Studies 27(5), 862–884, https://doi.org/10.1080/09654313.2019.1576591.
MacQueen, J. (1967): Some methods for classification and analysis of multivariate observations. In Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, eds L. M. Le Cam, J. Neyman 1, 281–297. Berkeley, CA: University of California Press.
Nooteboom, B. (2000): Learning by Interaction: Absorptive Capacity, Cognitive Distance and Governance. Journal of Management and Governance 4, 69–92, https://doi.org/10.1023/A:1009941416749.
Virmani, D., Taneja, S., Malhotra, G. (2015): Normalization based K means Clustering Algorithm. Journal of Advanced Engineering Research and Science, https://doi.org/10.48550/arXiv.1503.00900.
Yeon, J., Jun, B. (2022). The spillover effect of neighboring port on regional industrial diversification and regional economic resilience. General Economics, https://doi.org/10.48550/arxiv.2204.00189.
Yeung, H. (2020): Regional worlds: from related variety in regional diversification to strategic coupling in global production networks. Regional Studies 55(6), 989–1010, https://doi.org/10.1080/00343404.2020.1857719.

The typology of countryside architectonical forms in South-Moravia, a region of Czechia

Andrea Lešková, Antonín Vaishar*

Mendel University in Brno, Faculty of AgriSciences, Institute of Applied and Landscape Ecology, Czechia
* Corresponding author: A.Vaishar@seznam.cz

ABSTRACT
This article focuses on the typology of countryside architectonical forms in the region of South-Moravia in southeastern Czechia and on the expression of village identity through architecture in case study villages. Original folk architecture has been altered by new types of constructions built in rural areas since the 1950s, followed by a more recent wave of new architectural forms that have developed since the 1990s. The number of architectural types in case study villages was predominantly calculated using the panoramic sceneries on mapy.cz. The coefficients of countryside identity were allocated to architectural types based on basic folk house features. The value of countryside identity is higher in smaller villages except for suburbanized settlements of the regional capital of Brno.

KEYWORDS
folk architecture; identity of the countryside; aesthetic value; recreational potential; attractiveness; Moravia

Received: 31 May 2022
Accepted: 2 November 2023
Published online: 14 December 2023

Lešková, A., Vaishar, A. (2023): The typology of countryside architectonical forms in South-Moravia, a region of Czechia. AUC Geographica 58(2), 214–224
https://doi.org/10.14712/23361980.2023.16
© 2023 The Authors. This is an open-access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0).

1. Introduction

Rural sustainability is generally understood as the maintenance of environmental, economic and social conditions for the preservation of rural settlements, often seen from the agricultural viewpoint (Bosshard 2000). However, the sustainability of the countryside could be seen as a distinctive space in the processes of globalization (Gallent, Shucksmith, and Tewdwr-Jones 2003). It is a matter of how much the countryside remains rural, keeping its essential attributes. Maintaining at least some aspects of the rural lifestyle is crucial. This includes work, leisure, community life and other aspects including the preservation of rural architecture and construction style.

Urban architecture is traditionally a branch on the boundary between technology and art – later completed with social, environmental, economic and other aspects. Historical urban architecture in European culture is usually structured in time periods reflecting individual cultural styles. On the other hand, rural constructions develop on a different basis. They have always been focused on practical use, developing step-by-step based on trial and error. The material of the buildings corresponded to the local raw materials, the way of their construction relates to the climatic conditions and the way of use. Folk constructions differ more regionally than by cultural styles. Consequently, rural buildings could be considered a part of regional identity in the process of globalization.

Where the elements of folk architecture have been preserved, they become part of the cultural heritage and, in the future, goals of cultural tourism. There are 61 village conservation areas and 211 village monument zones in Czechia. The South-Bohemian village of Holašovice has become a part of the UNESCO World Heritage.

However, preserving the original folk buildings is becoming an increasing problem, as it is increasingly in conflict with the requirements of modern (i.e. urban) way of life. The original character of villages is changing through suburbanization and other urban processes in rural areas and urban-rural divisions are being blurred (Stringer 2017). Individual houses and purpose-built buildings are adapted to modern requirements. Paradoxically, increasing wealth of the population leads to the gradual disappearance of rural architecture and thus the identity of both the region and perhaps the countryside itself.

The paper aims at an analysis of a representation of remaining folk architecture and typology of new rural architecture types of houses, and a discussion of relations between folk architecture, regional identity and cultural tourism. The study should bring a proposal of a methodology for assessing the degree of preservation of rural architecture, which may be applicable in other areas.

2. Theory: Folk architecture and the identity of the countryside

In connection with the promotion of human rights, individual identity is increasingly asserted. However, A. Paasi (2003) draws attention to the new content of regional identity. Marek (2020) specifies that crucial to the existence of region are the subjective images that can be identified with perceptual regions. Regional identity becomes an inseparable part of geographical research (Chromý 2003) and plays an important role in the socialization of regions, and thus also comes into the perspective of regional planning (Raagmaa 2002) and it could also be applied in the administrative division of the state (Melnychuk and Gnatiuk 2018) – if there is political will to do so.

However, this contribution is not about the identity of individual regions, but about the identity of the countryside. This is also related to one view of the immaterial definition of the countryside (Halfacree 2008) as a space where people feel they are in the countryside. It is obvious that individual human settlements remain in their places, but their character changes under the influence of urbanization. Urban conveniences are penetrating the rural area (and partly also vice versa) and with them, the way of life and the urban structure of the villages are changing. There is a risk that the countryside ceases to be the countryside and turns into a difficult-to-define urbanized rural space. Moreover, rural areas are sometimes related to national or regional identity (Woods 2001) whereas cities are more symbols of globalization. That is why we asked ourselves the question of the relationship between the identity of the countryside and the representation of individual types of buildings in the urban structure of villages.

The identity of the countryside consists of a number of different factors. It can be a rural landscape (Scazzosi 2018), rural habits, customs, and traditions (Kwiatkowski et al. 2018), the design of rural settlements (Soukup and Sýkora 2020). These aspects form a complex and can be considered part of cultural heritage which can be used as a tourist attraction (Silva and Leal 2015).

One of the most important tangible factors of the identity of the countryside is the architecture that creates the overall image of the village. In addition to the dominant elements – churches, monasteries, fortresses, chateaux, folk architecture forms villages. Folk architecture has evolved for hundreds of years with respect to local specifics. Silva (2010) speaks about the folk architecture heritagization.

Traditional folk architecture is usually presented as an attraction for tourists (Copeţchi-Kopecký 2018), often concentrated in open-air museums or ethnic theme parks. However, folk buildings and architecture can also be important for their inhabitants, as they are part of the identity of their village and region and create a sense of home. In EU, vernacular architecture

belongs to intangible traits of heritage (Aytuğ and Mikaeili 2016; Namičev, Vuksanović-Macura, and Petrevska 2018).

Folk architecture is significantly threatened by modernization trends. Today's rural residents require the same conveniences as urban residents – all technical networks, garages, satellite receivers, large windows, modern materials. These requirements are often in conflict with the traditional buildings. Therefore, sophisticated protection, high level of understanding between property owners, architects, cultural heritage institutions and local authorities is needed (Alcindor and Coq-Huelva 2019). However, this protection is limited by financial resources and bargaining power. The matter is a subject of intensive discussion (Ilies et al. 2018; Sala, Trombadore, and Fantacci 2019). Šťastná, Vaishar, and Pakozdiová (2015) even state that the use of folk architecture and traditional way of life for the development of tourism in the Romanian Banat has led to improvement the situation of local people who used the funds to enhance their homes. However, this means the disappearance of the main motivation for tourism.

Folk architecture corresponds to the centuries-old experience of the locals. It uses local materials and is adapted to climatic and meteorological conditions (e.g. Philokyprou et al. 2017 or Lopez-Besora, Coch, and Pardal 2019). Its sensitive revitalization may even bring economic benefits in certain regions (Stival et al. 2020). Olukoya and Atanda (2020) studied various aspects of the sustainability of traditional architecture in the example of Cyprus. They concluded that the investigated vernacular architecture ranked lowly in physical resilience, accessibility and satisfaction but demonstrated sufficient lessons in the context of health and safety; participation and control, social equity; social cohesion; and cultural value.

Several authors have dealt with the issue of traditional folk architecture in Czechia since the 1970s. Among the most beneficial ones are Mencl (1980), who outlined the main areas of folk architecture and the characteristics of stylistic types throughout Czechoslovakia, Frolec (1974), who focused on folk architecture in Moravia and Silesia, Škabrada (1999), who also deals with the development of a traditional folk house, its construction and urban structure of settlements. Langer (2005) also devoted himself to open-air museums throughout Europe and thus placed the Czech and Slovak folk architecture in the European context.

Fewer authors deal with the evaluation and permanent documentation of contemporary architectural works in the countryside. In the second half of the last century, new residential units began to be built in villages, which were more similar to urban than rural development. After 1990, suburbanization started to develop.

In 1990s, several architects and urban planners reflected on the questions of the village's appearance, identity of the countryside, an image of the village and its landscape and assessed great changes that villages have undergone since the middle of the century. In this period, Knopp (1994) and Škabrada (1999) formulated ten basic urban and architectural principles in the village:

1) Rectangular floor plan of houses.
2) Location of the house on the plot – either perpendicular to the street or longitudinally, but always following the street line and usually directly connecting to the neighbouring house.
3) Saddle roof with a slope of about 45°.
4) Three-part floor plan of buildings, manifesting itself from the outside.
5) Binding situation of the entrance to the hall, the middle part of the floor plan, not the entrance through the gable wall.
6) Accessibility of the ground floor without the use of stairs.
7) The base of the roof was identical to the level of ceiling of the adjacent lower floor.
8) Threshold along the side of the house with the entrance protected by the overhanging roof resulting in an asymmetrical shield.
9) Chimney at the wall opposite the entrance.
10) Addition of the second wing of the house – narrower parts are added at the rear longitudinal wall.

Blažek (2004) describes the current works of folk constructions, tastelessness and fake searching for the future shape of the countryside, which will not lose its identity. In Moravia, Kašparová and Rozehnalová (2008) present specialized methodologies for the planned development of villages, where they deal with the spatial aesthetic aspect of the rural environment and with the preservation of the traditional image of the village.

If we start from the conceptual definition of cultural tourism as the movement of people from their place of permanent residence in order to obtain new information and experiences that satisfy their cultural needs (Richard, 2003), we can also consider folk architecture as one of the attractions of this form of tourism. In this case, the indicator of the identity of the countryside can serve as one of the measures of the attractiveness of individual settlements. Although vernacular architecture is mainly offered in open-air museums set up for that purpose, living villages that have at least partially retained their original rural character can represent a valuable element of cultural tourism, because it offers a live experience.

3. Methodology

This study is focused on the western surroundings of Brno where there is the lowest number of village monument reserves, zones and historically valuable villages. For the case study, 8 large municipalities with

a population of about 1000 inhabitants and 17 small municipalities with a population of about 100 inhabitants were selected (Fig. 1). The division was based on the assumption that large municipalities have a different proportion of public buildings, technical buildings and probably apartment buildings in their structure. In the second half of the last century, small settlements developed only minimally, and therefore it can be assumed that they were less influenced by new and architectural elements. The selection of villages for analysis was made with regard to whether elements of vernacular architecture combined with newer types of houses occur in their urban structure.

The methodology used is related to the concept of image use in geography (Yarwood 2005). In each case study village, all family houses were assessed. This study uses web portals with aerial maps and panoramic images. The main source of data for research was portal mapy.cz (©Seznam.cz, a.s., ©GEODIS BRNO, s r.o.), which has mapped not only the main roads, but also all side streets of villages. The mapping took place in 2013–2018. Partial sources of information were Google maps and Street View. The advantage of this approach is that it is possible to carry out analyzes of any settlement in this way without the need for financially, time- and labor-intensive field research. At the same time, in Czechia, in addition to a very detailed street view type display, other materials are available, such as geographical maps, historical maps, aerial photographs and aerial views, oblique images. Disadvantages include the fact that only objects that can be directly seen from the road or path are directly observable. Sometimes objects are either intentionally shadowed by tall fences or were shadowed by another object (such as a passing truck) at the time of the image. Therefore, the analysis was supplemented by field research conducted by students as part of their seminar practices. The field survey allows to capture the overall atmosphere of the place, sounds, bustle, people's mood or intangible spirit of the place among others.

Based on findings from field research, which were confronted with the available literature, individual types of country houses were singled out. The typology was made on the basis of similar features of houses that occur repeatedly and in different villages. Types of architecture were recorded in the map base for GIS with buildings accessible from https://services.cuzk.cz. Based on the elementary features of folk architecture such as rectangular floor plan, entrance orientation, number of storeys and others, each type of rural architecture has a "coefficient of identity of the countryside" (Fig. 2).

The valuable folk architecture has the highest coefficient of identity of the countryside (1). The devalued folk architecture and the new architecture in the style of folk architecture are evaluated by lower coefficients (0.8). Two-storey houses are rated with a coefficient of 0.6. Catalogue family houses are rated with a medium coefficient (0.5). Bungalows, as they do not have the next floor and often are not so visible from public spaces are rated with a higher coefficient of 0.7. Cottages and summerhouses are rated also with a higher coefficient of 0.7 whereas they are also often invisible from public spaces and situated mainly on the outskirts of the villages, even though their architecture is sometimes wild.

On the other hand, the lowest coefficient is for apartment buildings (0.1), large and high-reaching buildings often with flat roofs which would be more

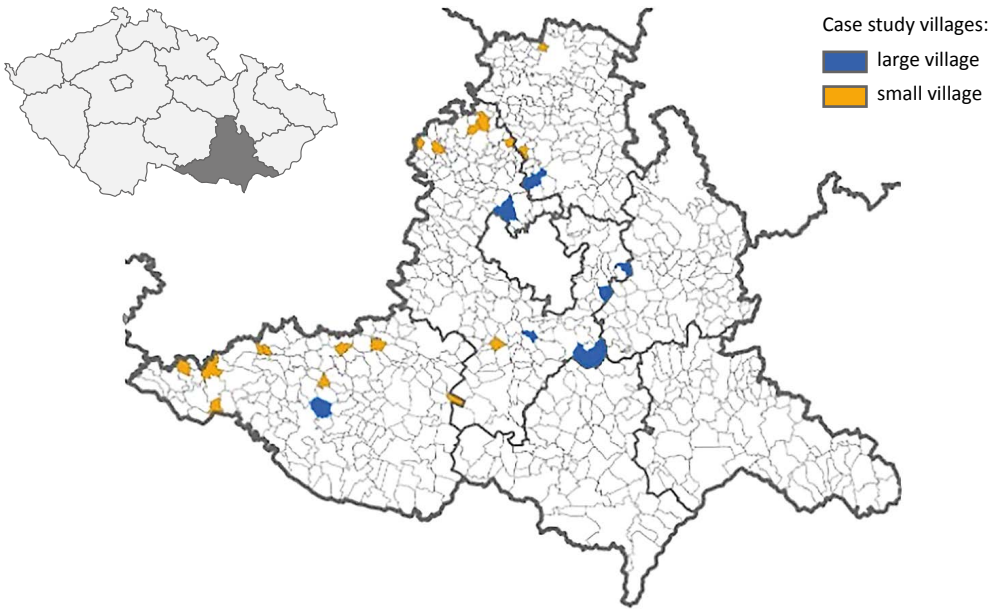


Fig. 1 Case study of small and large villages in three districts of the South Moravia region.
Source: © ArcČR ARCDATA PRAHA, ZÚ, ČSÚ, 2016, own elaboration.

Folk architecture devalued	Folk architecture valuable	Two-storey houses with saddle or hip roof	"Cube" houses	Apartment houses
– coefficient: 0.8 – until 1950s – rectangular floor plan – saddle roof – 1 (2) storey	**– coefficient: 1.0** – until 1950s – rectangular floor plan – saddle roof – 1 (2) storey	**– coefficient: 0.6** – 1960s, 1970s – square or rectangular floor plan – saddle or hip roof – 2 storeys	**– coefficient: 0.2** – 1960s, 1970s (present) – square floor plan – flat roof – 2 storeys	**– coefficient: 0.1** – 1970s, 1980s (present) – rectangular floor plan, large – saddle or a flat roof – 2–3 storeys
Catalogue houses	**Bungalows**	**McMansion**	**New houses in the manner of folk**	**Individual architecture**
– coefficient: 0.5 – 1990s–present – irregular floor plan – complex hip roof with dormers – 2 storeys	**– coefficient: 0.7** – 1990s–present – square floor plan – tent or hip roof – 1 storey	**– coefficient: 0.2** – 1990s–2000s – irregular floor plan – complex roof – 2–3 storeys	**– coefficient: 0.8** – 2000s–present – rectangular floor plan – saddle roof – 2 storeys	**– coefficient: 0.2** – 2000s–present – irregular floor plan – complex roof – 1–3 storeys

Fig. 2 Main types of rural architecture in the case study villages, their basic characters and coefficients of identity of the countryside. Source: own research.

suitable more for towns than rural areas. Similarly, as the apartment buildings, also the cube houses and houses of the "Šumperák" type are those which have no common feature with the traditional folk house, their coefficient of identity of the countryside is also low (0.2). Distinctive architecture as the artistic expression of individual architects with totally new shapes has also a very low coefficient of rural identity (0.2). McMansion is rated for its excessive pomp as 0.2 too.

Outside the residential buildings, there are some civic and technical buildings in the countryside. The highest coefficient is given to profane and sacral buildings (1) which are also a part of the cultural heritage and a source of cultural tourism in the place, too. A higher coefficient is also given to civic buildings (0.8) as they are often treated in the local folk style, but some stores and municipal offices come from the period of socialism. They are located in the centre of the village and have flat roofs which put overall emphasis on the functionality rather than on the visual appearance of these buildings. Agricultural and technical buildings are rated with a low coefficient of 0.1. They are often large untreated buildings on the edge of municipalities owned by companies. In larger villages, about 4% of buildings are garages which are also rated with a coefficient of 0.1.

The value of identity of the countryside index of the village is calculated as:

$$RI = \frac{A_1 \times c_1 + A_2 \times c_2 + \cdots + A_n \times c_n}{B}$$

where:

RI = value of identity of the countryside index of the village

A_1 = total buildings of architecture type 1 in the village

c_1 = coefficient given to architecture type 1

A_2 = total buildings of architecture type 2 in the village

c_2 = coefficient given to architecture type 2

…

…

A_n = total buildings of architecture type n in the village

C_n = coefficient given to architecture type n

B = total identified buildings in the village

Using the formula for calculating rural architecture identity of villages is a pilot attempt to express identity of the countryside in a quantitative way. There are some factors which are not included in the formula and which can also affect the overall impression of the village and its rural identity. Some of them are landscape features and landscaping of public spaces, front gardens and also the residents' own gardens. For example, large public spaces and the village greenery with tall trees can mask aesthetic shortcomings of buildings under the condition that domestic trees are used. Mareček (2005) suggests to cover unsuitable types of cube-shaped houses in rows by planting broad-crowned trees in front gardens or closing the visual axis of the street with conspicuous, handsome objects. Furthermore, the appearance of the village centre and its main busiest streets has the greatest impact on the overall impression of the village compared with areas on the outskirts. The distance of buildings from the village centre and main roads is not taken into account in the formula.

The discussion summarizes the possibility of using the formula in practice as well as the limitations of the formula and possible further research to specify a formula for calculating the identity of the countryside.

4. Analytical results

4.1 Valuable and devalued folk architecture

In this research, folk architecture is considered something that preserves the historical development of the village and has basic features such as a rectangular floor plan, single-storey structure, traditional orientation of the entrance to the house and others. In this research, however, we distinguish between the devalued and valuable folk architecture. Valuable folk architecture has also preserved other typical features such as the decoration of windows, doors, gates, the typical division of windows and other details that make this architecture aesthetically pleasing and therefore valuable. Devalued architecture often has modern windows with large undivided glass sheets, it does not have typical decorations around windows, doors, or it has a completely new decoration, a new facade, or possibly bright colours. It is often also a time-degraded architecture that needs to be repaired. Many of buildings in the Moravian villages (up to 35%) still retain their original location along the street line. However, the façade or exterior is changing, the floors are being adjusted, which makes this architecture less valuable in terms of preserving its typical character even in smaller decorative elements. Valuable folk architecture can be seen particularly in smaller villages (2%).

4.2 Two-storey houses, houses with wall gabled dormer and wall hipped dormer

Frequent type of houses represented mainly in larger villages (about 23%) are two-storey houses. Two-storey houses with wall gabled dormer were popular mainly in the 1960s and 1970s. The dormer is most often located in the middle of the front part of the house and accommodates one window. The house has two other windows on the ground floor, sometimes more. The entrance to the house is from the side and it is often advanced. The wall gabled dormer can also protrude from the facade and form a risalit, or the entire dormer is not placed symmetrically in the middle of the front of the house, but is located on the edge of the house.

Another type of two-storey house are house with the wall hipped dormer. This dormer is usually placed symmetrically in the middle of the front of the house and accommodates one full window. The other two windows are located on the ground floor. Variations of this type of house have the dormer located on the right or left. The house is also entered from the side. The hipped dormer allows to insertion a triangular element above the window, either in the form of another smaller window or in the form of a triangular decoration in the plaster of the house. Another variation is the advance of this dormer from the front facade in the form of a risalit. The hipped dormer does not have to be connected to the wall, it can also form a small protrusion from the roof of the house with a smaller window. The popularity of two-storey houses probably has not disappeared even today, as the elements of wall gabled dormer or wall hipped dormer placed symmetrically in the middle of the house front are still found in many new buildings.

4.3 Two-storey houses with the flat roof, Šumperák

Houses with a flat roof have been appearing since the 1970s. It was a very fundamental change influenced by functionalism. This type of house is very often a two-storey building with a square floor plan, so that it resembles a cube. It often has large windows, especially in the front side. These were the first types of houses with balconies in villages. A typical example is a cube with four large windows in the front, or a cube with balconies on both floors, while under the balcony there is a garage. Sometimes these houses were built in rows.

A specific type of house with a flat roof, repeating in Czechia and Slovakia, is a so-called "Šumperák" house. Its design was created in the 1960s by architect Josef Vaněk for the director of hospital in the town Šumperk. The design of this two floor house is famous for the balcony, which has two oblique walls on the sides and often with five circular windows. This house is so popular that certain elements of the bent side walls with circular windows are also found in completely different houses.

4.4 Apartment houses

The construction of apartment houses in the 1970s penetrated villages, too. Even in small villages, two-storey and three-storey apartment houses can be found. Older types often have a saddle or hip roof, newer types have a flat roof, which, however, further intensifies the conflict with the traditional visual of the village house. Apartment houses are not only remnants of the last century, but they are still popular today in the countryside, often with lower garages or balconies. The roof is no longer just flat or saddle, but also of other forms.

4.5 Catalogue family houses and bungalows

In 1990s, catalogue houses offer various versions of a comfortable family house. This type of houses, occurring mainly on the edges of suburban villages, sometimes forms entire new village districts, often not fitting into the original village part with the traditional buildings. They have various types of floor plan, various types of roofs (saddle, hip, half-hip), dormers, balconies and often various polygonal protrusions of the floor plan. They are often situated in the middle of the piece of land surrounded with an architecturally designed garden. They often have a high fence separating them from the street and neighbours.

Single-storey family houses – bungalows represent a frequent category of catalogue houses. The category of catalogue houses is the most represented category of houses under construction, so today it is still one of the most popular family houses.

4.6 Villas, McMansion

While the term villa refers to a large representative family house that often reflects the architecture of profane buildings of various architectural styles, the term McMansion is pejorative designation for a combination of different building styles. McMansion evokes the appearance of luxury, where the original noble materials are replaced by more affordable replacements (plastic windows, concrete interlocking paving). However, neither Villas nor McMansion are as widespread in the vicinity of Brno. In our case study villages, there are only buildings with the hints of baroque elements – a turret with arcades, arcades in the front of the house, entrances to the house highlighted by arches and columns.

4.7 Modern country houses

In recent decades, people's awareness of identity of the countryside and elements of architecture that make the countryside harmonious, pleasant and attractive has also been growing. More architects and owners are now interested in architecture that would be a representative for some kind of modern return to the traditional folk house (Nguyen et al. 2019). However, this architecture also accepts modern housing requirements such as larger windows, glass doors, built-in garages, floors, or dormers.

Another modern trend is the return to wood as the basic material of the house. New log cabins are being built such as residential houses, wooden houses, or even houses made of straw bales as a filling material for wooden constructions. Owners are attracted by the vision of living surrounded by natural materials, as well as the idea of sustainable architecture (Ragheb, el-Shimy, and Ragheb 2016). The idea of ecology and sustainable architecture is also the driving force behind energy-passive houses (Schnieders et al. 2020).

4.8 Cottages, second houses

Czechia is specific in its high number of cottages. Cottages and second houses began to develop at the beginning of the last century, when the first tramp settlements were established. Zapletalová (2007) presents a collection of 500 photographs of Czech cottages, in which their architecture is illustrated. The architecture of the first tramp huts was simple wooden (log cabins). In the period of socialism, cottages represented an escape from the collectivism, they compensated for closed borders and the impossibility of travelling. A large number of cottages can be found on the outskirts of smaller villages in our case study villages (7% of all buildings).

4.9 Representation of individual types of houses in the villages under research

In Moravian villages, houses of folk architecture predominate, devalued by later modifications or unsatisfactory maintenance. In small municipalities their share is 46%, in large municipalities 37%. The second most common type of residential buildings are two-storey houses of various variants, which occupy 10% in small municipalities and 23% in large municipalities. In small villages, the third most common house is a cottage (7%) and new houses in the ruralizing style. In large municipalities, the third most common type are catalogue family houses (10%) and cube houses with flat roofs.

Other types of buildings occur rather sporadically. Nevertheless, they influence the rural character of the villages either by their mass (apartment buildings) or ecstasy (villas, McMansions). Unfortunately, valuable houses of folk architecture occupy only about 2% in small municipalities and about half a percent in large municipalities.

The most common non-residential buildings are agricultural and technical buildings, which make up 18% of buildings in small villages and 9.5% of buildings in large villages. Among them we can find objects of large-scale agricultural production, created during the period of collectivization, small industrial areas, transport and military facilities functioning sometimes as entrepreneurial zones now. Some of them represent rural brownfields. In small villages, agricultural buildings are followed by civic facilities and sacral buildings, while in large villages, civic amenities are complemented by garage areas. Although non-residential buildings form a minority, we usually find dominants among them (Tab. 1).

The overall value of identity of the countryside index of the case study villages is about 0.6. Small villages have a higher value of the identity of the countryside index. However, the difference between small and large municipalities is not so big. In small villages, agricultural and technical buildings represent a high percentage (18%) compared to large villages (8%). This fact causes a decrease of the countryside index of small villages.

5. General results and discussion

Valuable folk architecture is a significant element of cultural tourism in Czechia. Many village monument reservations and village monument zones with official protection have information boards guiding tourists from the main roads to individual attractions. The most visited open-air museum of folk architecture in

Tab. 1 The number of buildings categorized into individual types of architecture in the selected case study villages. Percentages of the architecture types in small and large villages are calculated.

		Residential buildings												Civic and technical buildings							
Municipality	Population	Folk architecture devalued	Folk architecture valuable	Two-storey house (gable, hip roof)	Cube house (two storey, flat roof)	„Šumoerák"	Apartment house	Catalogue family house	Bungalow	Villa, McMansion	New architecture as folk	Individual architecture	Cottage	Civic amenities	Agricultural and technical buildings	Sacral and profane architecture	Garage	Total classified buildings	Nor identified	Total buildings	Total value for the village
Coefficient of rural identity		0.8	1	0.6	0.2	0.2	0.1	0.5	0.7	0.2	0.8	0.2	0.7	0.8	0.1	1	0.1				
Blanné	74	31		4	2				1			2		3	20			63		63	0.53
Bukovice	76	19	2	11	4			4				1	3	4	18	1	4	71		71	0.50
Němčičky	83	48	1	5	1						1	1		4	21	1		83		83	0.60
Čermákovice	90	59	4	2	1			6					6	3	18	1		100		100	0.65
Tišnovská Nová Ves	91	13	5	16			1	7	1		1	2	6	4	16	1	2	75		75	0.54
Prokopov	92	65	2	1	1			3						2	2	1		77		77	0.77
Vysočany	92	47	3	1		1	1		1	1		4	54	3	21	1	2	140		140	0.62
Rudlice	94	50	1	4			2		1	1				1	9	1		70	12	82	0.67
Lubě	95	33		9				1				9		1	7			60		60	0.59
Podmyče	99	13	3	8			1				15			2	30	1		73		73	0.49
Vratislávka	100	22	3	4	1		6	2	7		19			2	16	1		83		83	0.59
Přeskače	104	22	2	18								1		9	13	2		67	19	86	0.61
Chvalatice	107	90		20	8			2			1		20	4	10	1		156		156	0.68
Skrchov	111	36		16	1						4				13			70	12	82	0.62
Ochoz u T.	112	28		10					1		5		2	2	15	2		65		65	0.61
Synalov	114	53	1	11				2	1		10		13	4	9	1		105		105	0.70
Trnové Pole	115	32		4				5	1		2			1	20			65	4	69	0.55
Total number of types of architecture in very small villages		661	27	144	19	1	11	32	14	2	58	20	104	49	258	15	8	1423	47	1470	0.62
Percentage of types of architecture in very small villages		46.5	1.9	10.1	1.3	0.1	0.8	2.2	1.0	0.1	4.1	1.4	7.3	3.4	18.1	1.1	0.6				
Mor. Knínice	958	153	6	48	13	3		53	16	5	26	7	3	7	73	3	13	429	10	439	0.56
Holasice	1161	154		90	26		1	126	6			25	2	10	49	1	60	550		550	0.50
Moutnice	1170	203		160	13	1		16	3		2	2		10	43	1	5	459	2	461	0.63
Blažovice	1221	270	2	63	4	1	8	18	8	2			3	11	30	2		422	35	457	0.68
Těšany	1230	193	7	68	47		2	63	3			14		38	62	9	27	533		533	0.56
Únanov	1266	158		74	56	2	10	3	51		1	1		23	17	1	30	427	119	646	0.58
Lipůvka	1294	168		120	49	7	12	58	4	2		1	11	25	63	5		525		525	0.55
Vin. Šumice	1343	107	6	242	2		5	35	5	2	38	3		11	29			485	117	602	0.62
Total number of types of architecture in large villages		1406	21	865	210	14	38	372	96	11	67	53	19	135	366	22	135	3830	383	4213	0.58
Percentage of types of architecture in large villages		36.7	0.5	22.6	5.5	0.4	1.0	9.7	2.5	0.3	1.7	1.4	0.5	3.5	10.0	0.6	3.5				
Percentage of types of architecture in all villages		39.4	0.9	19.2	4.4	0.3	0.9	7.7	2.1	0.2	2.4	1.4	2.3	3.5	11.9	0.7	2.7				

Source: own research

Rožnov pod Radhoštěm is annually visited by 350,000 persons.

Every year, the Czech Republic evaluates the Village of the Year competition on the basis of criteria, including architectural and urban principles supporting traditional architecture and the traditional image of the village among others. Results of this research show that even small villages may hide a certain cultural potential. Due to their non-expansion in the last century, small villages have mostly retained their typical traditional expression. A lot of valuable folk architecture is located in these small villages even today. Creation of a brand of the most beautiful and valuable municipalities, information boards and map applications to support tourism in these municipalities could help their overall development.

Although municipalities with a high degree of preservation of traditional folk buildings are popular tourist attractions, having permanent housing in such municipalities is usually not popular. Their residents are bound by a number of restrictions. They cannot build new objects on their own land, repairs of old buildings are also tied to a number of official permits. Repair of old buildings is often costly, even more if they want to be done precisely, with original techniques and materials. Nevertheless, even today, the number of people who consciously and voluntarily choose a more modest way of life in a family house with the traditional elements of folk architecture is still rising.

The authors are aware of the subjective contribution to determining the identity of the countryside coefficients. The methodology is universal in terms of the use of data sources that are available and comparable at least throughout Europe. However, the typology and assessment of individual types of buildings would need to be modified for each region, to the extent that folk architecture buildings differ regionally according to natural conditions and historical development. Thus, specific typologies and coefficients may be valid for the territory of South Moravia (not necessarily identical to the administrative borders of the South Moravian Region), but local conditions should be taken into account when applying the methodology in different regions in Czechia or abroad.

6. Conclusions

This research focuses on the typology of architectural forms in the Moravian countryside and on the expression of identity of the countryside of case study villages. About 10 main architectural types of residential buildings were defined in the countryside, using predominately panoramic sceneries at mapy.cz. Coefficients of identity of the countryside were given to the architectural types based on agreement with the basic folk house features. A new methodology for calculating the value of identity of the countryside is presented in this research. The value of rural identity is somewhat higher in smaller villages which are more apart from suburbanization of Brno. These villages could have a hidden tourist potential.

Rural architecture as well as the overall image of the countryside has changed very much since the 1950s. Our results show that valuable folk architecture is almost lost and its impact on rural sustainability is minimal. It fits into the outgoing vision rural idyll (Shucksmith 2018). One possible solution is to protect the remaining heritage. The protection in the open-air museums and declared monument zones does not address the issue of rural sustainability. The protection of the folk architecture of permanently inhabited buildings encounters an understandable interest of their inhabitants in modernization and is not sustainable in the long run.

A more promising option would be not to allow the construction of buildings in the rural environment that are not justified in the countryside. This principle was broken in the collectivization period. Unfortunately, the possibility of stopping this process in the 1990s was not used. The countryside was flooded with buildings that have no place here either because of their inhumane scale or foreign origin.

However, there is some hope of maintaining the sustainability of the rural character of the architecture. In the second half of the last century, this was caused by cottagers (Fialová et al. 2010), whose activities prevented the destruction of a large part of country houses and who maintain their buildings precisely to escape the city and therefore do not have such demands for modernization as the locals. Some small and very small municipalities are now experiencing migratory increases in population. It can be presupposed that it is people who prefer the rural way of life, including the character of their homes, who move to them.

It is obvious that the countryside will differentiate not only according to the distance from regional centres, according to physical-geographical characteristics, human potential and historical development, but also according to the degree of urban transformation of the housing stock. This circumstance will be important for rural planning at the level of municipalities, their associations or local action groups, as well as for the potential development of cultural tourism.

Acknowledgements

The paper is a result of the project “Analysis of rural identity in regions of traditional folk architecture in Moravia” Nr. AF-IGA2020-IP057of the Internal Grant Agency of the Faculty of AgriSciences of Mendel University in Brno. The authors would like to thank students of the course Renewal of Countryside 2020 for their help with determining the rural architectural types. The authors declare that they have no conflict of interest.

References

Alcindor, M., Coq-Huelva, D. (2019): Refurbishment, vernacular architecture and invented tradition: the case of the Empordanet (Catalonia). International Journal of Heritage Studies 26(7), 684–699, https://doi.org/10.1080/13527258.2019.1678054.

Aytuğ, H. K., Mikaeili, M. (2016): Evaluation of Hopa's Rural Tourism Potential in the Context of European Union Tourism Policy. Procedia Environmental Sciences 37, 234–245, https://doi.org/10.1016/j.proenv.2017.03.039.

Blažek, B. (2004): Venkovy: Anamnéza, diagnóza, terapie. Brno: ERA.

Bosshard, A. (2000): A methodology and terminology of sustainability assessment and its perspectives for rural planning. Agriculture, Ecosystems & Environment 77(1–2), 29–41, https://doi.org/10.1016/S0167-8809(99)00090-0.

Chromý, P. (2003): Formování regionální identity: nezbytná součást geografických výzkumů. In Jančák, V., Chromý, P., Marada, M., eds., Geografie na cestách poznání. Praha.

Copețchi-Kopecký, B. (2018): The functions and the impact of the vernacular architecture from Vaideeni village on the development of agro-mountain tourism in the Căpățânii mountains. Knowledge Horizons – Economics 10(4), 8–11, http://www.orizonturi.ucdc.ro/arhiva/KHE%20nr.%204%20-%202018/1.%20THE%20VERNACULAR%20ARCHITECTURE%20FROM%20%20VAIDEENI.pdf.

Fialová, D., Chromý, P., Kučera, Z., Spilková, J., Štych, P., Vágner, J. (2010): The forming of regional identity and identity of regions in Czechia – introduction to the research on the impact of second housing and tourism. AUG Geographica 45(1–2), 49–60, https://doi.org/10.14712/23361980.2015.56

Frolec, V. (1974): Lidová architektura na Moravě a ve Slezsku. Brno: Blok. 399p.

Gallent, N., Shucksmith, M., Tewdwr-Jones, M., eds. (2003): Housing in the European countryside. London: Routledge, https://doi.org/10.4324/9780203451717.

Halfacree, K.H. (2008): Locality and social representation: space, discourse and alternative definitions of the rural. In Munton, R. (ed.). The Rural: Critical Essays in Human Geography. London: Routledge, 245–260, https://doi.org/10.4324/9781315237213.

Ilies, G., Ilies, M., Hotea, M., Simion, A. S., Bumbak, S. V. (2018): What future for the land of the wooden civilisation? Vernacular architecture from Maramures as subject of a long standing debate. GeoJournal of Tourism and Geosites 22(2), 585–596, https://doi.org/10.30892/gtg.22226-313.

Kašparová, L., Rozehnalová, E. (2008): Vesnice. Nové stavby pro venkov. Brno: Institute for Spatial Development.

Kwiatkowski, G., Oklevik, O., Hjalager, A. M., Maristuen, H. (2020): The assemblers of rural festivals: organizers, visitors and locals, European Planning Studies 28(2), 255–272, https://doi.org/10.1080/09654313.2019.1651829.

Langer, J. (2005): Atlas památek. Evropská muzea v přírodě. Praha: Baset.

Lopez-Besora J., Coch H., Pardal C. (2019): Contemporary Roof Design Concepts: Learning from Vernacular Architecture. In Sayigh, A. (eds.), Sustainable Vernacular Architecture. Innovative Renewable Energy 357–378, Cham: Springer, https://doi.org/10.1007/978-3-030-06185-2_16.

Marek, P. (2020): Region as a social construct and the critical discussion of the Paasi's conceptualization of regional identity. Geografie 125(1), 47–68, https://doi.org/10.37040/geografie2020125010047.

Melnychuk, A., Gnatiuk, O. (2018): Regional identity and the renewal of spatial administrative structures: The case of Podolia, Ukraine. Moravian Geographical Reports 6(1), 42–54, https://doi.org/10.2478/mgr-2018-0004.

Mencl, V. (1980): Lidová architektura v Československu. Praha: Academia.

Namičev, P., Vuksanović-Macura, Z., Petrevska, B. (2018): Vernacular architecture in Macedonia and Serbia: a comparative study. In Brankov, J., Drobnjaković, M. (eds.) The Balkan penninsula of Jovan Cvijić 243–252. Beograd: Geographical Institute Jovan Cvijić.

Nguyen, A. T., Ha Truong, N. S., Rockwood, D., Tran Le, A. D. (2019): Studies on sustainable features of vernacular architecture in different regions across the world: A comprehensive synthesis and evaluation, Frontiers of Architectural Research 8(4), 535–548, https://doi.org/10.1016/j.foar.2019.07.006.

Olukoya, O. A. P., Atanda, J. O. (2020): Assessing the Social Sustainability Indicators in Vernacular Architecture – Application of a Green Building Assessment Approach. Environments 7(9), 67, https://doi.org/10.3390/environments7090067.

Paasi, A. (2003): Region and place: regional identity in question. Progress in Human Geography 27(4), 475–485, https://doi.org/10.1191/0309132503ph439pr.

Philokyprou, M., Aimilios, M., Malaktou, E., Savvides, A. (2017): Environmentally responsive design in Eastern Mediterranean. The case of vernacular architecture in the coastal, lowland and mountainous regions of Cyprus. Building an Environment 111, 91–109, https://doi.org/10.1016/j.buildenv.2016.10.010.

Raagmaa, G. (2002): Regional Identity in Regional Development and Planning1, European Planning Studies 10(1), 55–76, https://doi.org/10.1080/09654310120099263.

Ragheb, A., El-Shimy, H., Ragheb, G. (2016): Green Architecture: A Concept of Sustainability, Procedia – Social and Behavioral Sciences 216, 778–787, https://doi.org/10.1016/j.sbspro.2015.12.075.

Richards, G., ed. (2003): Cultural tourism in Europe. Wallingford Oxfordshire, UK, CABI.

Sala, M., Trombadore, A., Fantacci, L. (2019): The Intangible Resources of Vernacular Architecture for the Development of a Green and Circular Economy. In Sayigh, A., ed., Sustainable Vernacular Architecture. Innovative Renewable Energy 229–256. Cham: Springer, https://doi.org/10.1007/978-3-030-06185-2_12.

Scazzosi, L. (2018): Rural Landscape as Heritage: Reasons for and Implications of Principles Concerning Rural Landscapes as Heritage ICOMOS-IFLA 2017. Built Heritage 2, 39–52, https://doi.org/10.1186/BF03545709.

Schnieders, J., Eian, T. D., Filippi, M. et al. (2020): Design and realisation of the Passive House concept in different climate zones. Energy Efficiency 13, 1561–1604, https://doi.org/10.1007/s12053-019-09819-6.

Shucksmith, M. (2018): Re-imagining the rural: From rural idyll to Good Countryside. Journal of Rural Studies 59, 163–172, https://doi.org/10.1016/j.jrurstud.2016.07.019.

Silva, L. M. S. (2010): Folk architecture heritagization in rural Portugal. In Roigé Ventura, X., Frigolé Reixach, J. (eds.), Constructiong Cultural and Natural Heritage 121–131. Girona: Documenta Universitaria.

Silva, L., Leal, J. (2015): Rural tourism and national identity building in contemporary Europe: Evidence from Portugal. Journal of Rural Studies 38, 109–119, https://doi.org/10.1016/j.jrurstud.2015.02.005.

Škabrada, J. (1999): Lidové stavby: Architektura českého venkova. Praha: ARGO.

Soukup, P., Sýkora, K. (2020): Development of rural architecture – design and creation of a web database. Civil Engineering Journal 29(4), https://doi.org/10.14311/CEJ.2020.04.0042.

Šťastná., M., Vaishar, A., Pákozdiová, M. (2015): Role of tourism in the development of peripheral countryside. Case studies of Eastern Moravia and Romanian Banat. Forum Geographic 14(1), 84–94, https://doi.org/10.5775/fg.2067-4635.2015.198.i.

Stival, C. A., Berto, R., Morano, P., Rosato, P. (2020): Reuse of Vernacular Architecture in Minor Alpine Settlements: A Multi-Attribute Model for Sustainability Appraisal. Sustainability 12: 6562, https://doi.org/10.3390/su12166562.

Stringer, B. (2017): Architecture and Culture: A Villages and Globalization Issue. Architecture and Culture 5(1), 1–4, https://doi.org/10.1080/20507828.2017.1299434.

Vondrušková, A., Vondruška, V. (2014): Vesnice. Průvodce českou historií. Praha: Vyšehrad.

Yarwood, R. (2005): Beyond the Rural Idyll: Images, countryside change and geography, Geography 90(1), 19–31, https://doi.org/10.1080/00167487.2005.12094114.

Zapletalová, V. (2007): Chatařství: architektura lidských snů a možností. Summerhouses: the architecture of human dreams and possibilities. Brno: ERA.

Consequences of capital drain among EU member states

Peter Deďo, David Hána*

Charles University, Faculty of Science, Department of Social Geography and Regional Development, Czechia
* Corresponding author: david.hana@natur.cuni.cz

ABSTRACT
The aim of this article is to analyze the capital drain among individual European Union (EU) member states and its cohesive and political consequences. Since the capital drain has not yet been calculated at the individual country level, the methodological part of this article delves into this calculation in more detail. Between 1999 and 2018, Ireland and Luxembourg had the highest capital drain due to their tax haven policies. Apart from these extremes, Czechia experienced the largest capital drain during this period. Inequalities among EU member states were gradually decreasing in terms of gross domestic product and gross national disposable income, suggesting that the EU's cohesion policy has partially been successful in reducing inequalities among EU countries. However, capital drain and its populist interpretations may become a significant political problem for the most negatively affected countries.

KEYWORDS
capital drain; spatial inequalities; gross disposable income; gross domestic product; European Union

Received: 30 June 2023
Accepted: 9 November 2023
Published online: 29 November 2023

Deďo, P., Hána, D. (2023): Consequences of capital drain among EU member states. AUC Geographica 58(2), 225–237
https://doi.org/10.14712/23361980.2023.17
© 2023 The Authors. This is an open-access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0).

1. Introduction

Economic inequalities are one of the most widely discussed topics among the world's leading economists and geographers (such as Sala-i-Martin 2002; Smith 2008; Harvey 2010; Piketty 2014; Piketty and Saez 2014). In the perspective that sees inequalities as a problem escalating over time (see Keynes 2018), a state or a multinational organization must intervene to reduce disparities through capital reallocations. The European Union (EU), or previous organizations respectively, has focused on this objective since its establishment when it adopted the Treaty of Rome in 1957 with one of the specific goals being to "reduce the economic and social differences between the EEC's [European Economic Community] various regions". An outcome of this stance was the conceptualization and implementation of the European cohesion policy in 1980s (Molle 2017). Nowadays, the EU spends about one third of its budget on this policy, implemented through EU funds every year. Its importance is evident from the turbulent discussions in countries of Central, Eastern, and Southern Europe about the proposed plan for cuts in the cohesion policy (e.g. AP News 2019).

However, there is a reverse capital flow exceeding EU funds (Kučera 2016; Keller 2017) which may have a significant effect on EU cohesion policy. This "capital drain" (Hakenes, Schnabel 2010) is gaining political priority mainly in the countries experiencing problems with massive capital outflow (e.g. Chmelař et al. 2016). In a simplified dichotomous view at the European level, these are peripheral states in Central, Eastern, and Southern Europe that are sources of cheaper labor and production in sectors with lower added value compared to the core states in Northwestern Europe where more advanced technology and more profitable economic activities are concentrated (Storper 2018; Pavlínek 2022a). In the literature, it is quite frequently debated at the level of intranational regional disparities (Hakenes and Schnabel 2010; Bečicová and Blažek 2015; Hána, Hellebrandová 2018). Although such discussion is practically non-existent among states. From the perspective of several theories (Myrdal 1957; Wallerstein 2011), capital should move from the periphery to the core (Wallerstein 2011).

By acceding to any economic union that aims to remove trade barriers, a new member state exposes its market to competition from the old members. In peripheral regions (for the difference compared to FDI in core regions, see Pavlínek 2022b), the inflow of foreign direct investments (FDI) often has a positive impact on the growing macroeconomic indicators (Hlaváček and Bal-Domanska 2016). After a certain time, however, it facilitates the draining of profits from the host economy (Kučera 2016; Grela et al. 2017). It is a process that follows the logic of capitalism because FDI is primarily a tool to achieve the profits of TNCs (Pavlínek 2022b). Therefore, it is not perceived negatively at first glance, but to a certain extent, it may have fundamental cohesive and political consequences. In this context, EU funds can be perceived as compensation for capital flows aimed at not increasing disparities (Keller 2017). It is important to find out how individual states stand in this process. The aim of this paper is, therefore, to analyze the balance of the capital drain among individual EU member states. The article then discusses its cohesive and political consequences.

2. Capital drain and its impact on EU member states cohesion

There are specific spatial capital flows known as "regional drainage", explaining the flows on the intranational regional level (Bečicová and Blažek 2015), or more generally as "capital drain" (Hakenes and Schnabel 2010). We can explain both essentially using world systems theory with hierarchical relationships between the dominating 'core' and the dependent 'periphery' (in a simplified and illustrative dichotomous view; in reality, there is a continuous spatial transition through the semi-periphery areas; Pavlínek 2022a). Due to the dichotomy of producers with lower-priced inputs and lower returns in peripheral regions and producers with higher-priced inputs and higher returns in core regions, the world system is characterized by mechanisms that cause value redistribution from the periphery to the core (Holubec 2009; Sorinel 2010; Wallerstein 2011). Small possibilities of safe and high-return investments in peripheries are not suitable for producers' savings, which could be invested in core regions where a lack of available finance is in contrast. Consequently, with no regulation between regions or states, capital freely flows and accumulates in the core (Myrdal 1957; Wallerstein 2011). This is one of the factors contributing to growing regional disparities and devaluation of capital in the peripheries, as documented several times at the regional level (Hakenes and Schnabel 2010; Bečicová and Blažek 2015).

Since core and peripheries may exist on many hierarchical levels, we can study these flows on an international level. In this view, the organization of the global economy based on global value chains (Gereffi 2005) and global production networks (Henderson et al. 2002) is essential. Peripheral regions are characterized (among other things) by a high degree of foreign ownership and control due to the strong position of core companies in comparison to peripheral ones, the lowering of trade barriers (including the establishment and expansion of the common market in the EU), deregulation of FDI, and various government policies (Pavlínek 2022a). Leading firms from the economic core then make higher profits in lower-cost peripheries (Pavlínek 2022a) and control

value creation and capital flows throughout the entire value chain. All decisions are tailored to their interests, respecting conditions in specific countries that may differ according to their involvement in the global economy (Gereffi 2005; Pavlínek and Ženka 2011). The created and enhanced value can be captured in the territory where it was created (Henderson et al. 2002), but it can also flow elsewhere, which can be influenced by both corporate policies (e.g., profit shifting to tax-advantaged countries, see Nerudová et al. 2023) and government policies (e.g., tax regulations). Corporations thus have a significant influence on the places where they operate, including through the capturing or the transfer of values, which can take place in both directions to and from the host economy (Henderson et al. 2002; Coe et al. 2004). This process then has a significant influence on strengthening the position of the core and perpetuating the peripheral status (Pavlínek 2022a).

Therefore, it makes sense to compare the long-term balance of each country's transfers in Europe to understand the position of countries from different types of regions in the core-periphery dichotomy. Moreover, this level of capital drain may have significant political effects, as this topic can be used and abused for political goals at the state level. It can be present as an uneven labor burden on behalf of wealthier countries, which creates not only economic but also fundamental political inequalities that can lead to political tension and conflicts (Piketty 2014). Therefore, there is a significant gap in our knowledge about this capital drain on the international level, which should be filled for a better understanding of European disparities and their political consequences.

In the literature, capital flows such as FDI are commonly studied, either from an international perspective (Borensztein et al. 1998), including their relationship with political regimes and democracy (Jensen 2003; Li and Resnick 2003), or with a more focused view on their ability to be used in a destination country (Alfaro et al. 2004) and their impacts on the destination country (Javorcik 2004; Shahbaz et al. 2018; Pavlínek 2022b). Similarly, within the EU, flows of EU funds are often studied, with a focus on their impact on mitigating intra-Union convergence both at the international (Puigcerver-Peñalver 2007) and state regional levels (Lolos 2009; Kyriacou and Roca-Sagalés 2012). The reason is that the objective of reducing regional disparities was already established in the Treaty of Rome, and later emphasized with each individual accession (Magrini 1999). In one view, the common market, as one of the building blocks of the EU, is not enough to alleviate economic and social inequalities, which is why EU regional policy was introduced (Fiala et al. 2018: 604). According to another view, it is precisely the common market that creates these inequalities, which need to be addressed by EU regional policy (Fiala et al. 2018: 606). The main EU tool of the regional and cohesion policy is EU funds (Puigcerver-Peñalver 2007) with a significant amount of financing (350 billion euros in 2014–2020, Fiala et al. 2018: 622), representing approximately one third of the EU budget (Goulet 2011; European Commission 2014).

Their efficiency has been intensively discussed, but no consensus has been reached concerning their impact on Europe-wide regional convergence (Ederveen et al. 2006). Becker et al. (2010) demonstrated that the funds have a relatively significant positive impact on economic growth, which is also mentioned by Cuaresma et al. (2008). Some authors observe a positive impact on the growth of regional incomes (Lolos 2009) or intra-state reduction of regional disparities (Kyriacou and Roca-Sagalés 2012), but they do not evaluate the contribution to Europe-wide convergence. In this perspective, Dall'Erba and Fang (2017) claim that the allocation of resources has become more efficient in recent years, and Grela et al. (2017) state that the countries of (semi-) peripheral Central and Eastern Europe are successfully converging to the GDP levels of the core older EU member states. On the other hand, Mihaljek (2018) applies different methods to evaluate convergence and expresses a more skeptical view, claiming that no significant convergence has been demonstrated. An interesting observation is advanced by Gros (2018) in his study: there is convergence between east and west in the EU; however, if the comparison line runs between north and south, this process is stagnating considerably. Many authors believe that the EU funds do not have any significant impact on the convergence process (e.g., Boldrin and Canova 2001; Dall'Erba and Le Gallo 2008; Esposti and Bussoletti 2008) or even that their economic impacts may be more significant in more developed areas (Cappelen et al. 2003).

On the other hand, there are monetary transfers, which can be included under the term "capital drain", such as company dividends or bank interests from loans to enterprises or states by institutions from the European core (Keller 2007). These are not extensively studied, although they may even exceed FDI or EU funds flows in volume, and their direction is oppositional, from the periphery to the core. For example, in Czechia in a third of the year 2015 (based on profits from foreign direct investment, income from work abroad, interest, and mandatory contributions to the common budget of the EU), the flow of transfers abroad was about 450 billion CZK (approximately 20 billion euros; 1 euro = 23 CZK), which corresponded to roughly 9% of the Czech GDP. At the same time, about 160 billion CZK (approximately 7 billion euros) flowed into Czechia from abroad (Kučera 2016), plus 93 billion CZK on average (4 billion euros) from EU funds in the period 2014–2020 (Ministry of Regional Development 2023). These disparities could have both international and intranational consequences.

It may explain a certain failure of EU regional policy in European cohesion because there are two reverse directional flows, of which EU funds play a minor role (Keller 2007).

3. Methodological framework

Unfortunately, there are no pan-European statistics to contribute to this discussion about capital drain. Therefore, we must calculate the capital drain from available indicators. The methodology is based on the paper by Kučera (2016), who assesses the drain of one country as the difference between the gross domestic product (GDP) and the gross national disposable income (GDI). By following this procedure, we obtain positive or negative capital drain values for each country, which are determined by the inflow or outflow of the mentioned transfers.

The most frequent indicator of the level of economic development is GDP, as the basic index of production performance of an economy (Rojíček et al. 2016). It can be calculated in three ways: using the production approach, the expenditure approach, or the income approach. The basic form uses the expenditure approach and is as follows:

$$GDP = C + 1 + (X - M)$$

where C represents final consumption (household and governmental institution expenditures for final consumption); I is the creation of gross capital; X is the export of goods and services, and M is the import of goods and services.

However, it has two fundamental shortcomings (Piketty 2014). Firstly, the inclusion of expenditures on the restoration of used capital (production equipment and buildings, including the restoration of infrastructure after natural and other disasters) which is necessary to avoid constant depreciation of assets, leading to a reduction in production capacity and income, but which is not income by itself. Secondly, and most importantly for the aim of this paper, GDP does not reflect interstate or interregional capital flows. For example, a country where businesses are owned by foreign owners will have lower revenues than its GDP value. On the contrary, countries with investments abroad can have significantly higher incomes than the GDP they produce on their territory (Piketty 2014). The use of the GDP indicator would thus mask a "capital drain" and increasing inequalities at the international level (Alvaredo et al. 2018). With the increasing volume of international capital and financial flows, multiplied by the free movement of labor between states, in multiple countries, the GNI development significantly differs from GDP. Therefore, the GDP index is not appropriate for measuring living standards and interstate inequalities (Alvaredo et al. 2018).

The following equation defines the transition from GDP to gross national income (GNI):

$$GNI = GDP + NY$$

where NY is the balance of initial income of residents with non-residents (Rojíček et al. 2016). Eurostat, based on the European System of Accounts (ESA) 2010 (Eurostat 2013), classifies the three types of transactions between residents and non-residents as initial income. Firstly, there are employee remunerations containing salaries and other benefits paid in cash or in kind that were awarded to individuals for their work performed for enterprises in a different location than their place of residence (workers at the borders, seasonal workers, employees of international organizations etc.). Secondly, there are employee salaries paid to non-resident workers or paid by non-resident employers. The most voluminous item of initial income is yields on investments representing income originating from the ownership of foreign financial assets and liabilities paid by the residents of one economy to the residents of a different economy. This includes interest, dividends, payments of branch profits, and direct investors' share in undivided profit of companies operating in the field of direct investments, and income allocated to insured persons under insurance systems, pension security, and standardized security schemes.

The third equation expresses the transition from GNI to GDI:

$$GNI = GDP + NCT$$

where NCT is the balance of current transfers in relation to foreign countries. Rojíček et al. (2016) state that the main types of current transfers are, besides the ordinary taxes and social allowances, also current transfers between governments or international organizations and so-called remittance, i.e. payments transferred by foreign employees to their families – residents of a given country. An important equation is the application of GDI:

$$GNI = C + S$$

where C is the final consumption and S is the gross national savings. Finally, the last equation describing the relation between GDI and GDP is:

$$GNI = GDP + NY + NCT$$

where the difference between GDI and GDP is the balance of initial income and current transfers.

Most of the data comes from the Eurostat database (2019), in particular, data concerning GDP, GDI (Non-financial transactions [nasq_10_nf_tr]), and the population size for the EU member states (Population change – Demographic balance and crude rates at national level [demo_gind]), indicating the population sizes at 1 January of the respective year. The observed period represents an interval from 1999

to 2018, which was selected in an effort to capture the period of European integration that involves the development before the widest accession in 2004, the critical financial crisis in 2008, and the post-crisis development by 2018. The period ends a year before the turbulent time affected by the COVID-19 pandemic and avoids any measures from the end of 2019 responding to the approaching virus threat as well. Croatia was omitted because it was not a full member of the EU for the majority of the observed period, and Eurostat did not have all the required information available, which is necessary for an empirical analysis. The UK is considered to be a member state because it was throughout the studying period.

For assessing international inequalities between EU member states, we used the Gini coefficient as the most frequently used tool for comparing the relative values and their regional concentration. Based on the definition, its values may range from 0 to 1, where the value 0 represents equal distribution of wealth, and 1 represents maximum inequality, where wealth is concentrated in the hands of a single individual. An average coefficient may be used for comparison of various populations, countries, or regions (Eckey and Türck 2005). Another method of measuring regional disparities is a population-weighted form of the Gini coefficient which reflects differences in the population of units. The EASYSTATS statistic tool from Novotný et al. (2014) was used for data administration and Gini coefficient calculations.

4. Observing capital drain and its impact on inequalities within the EU

The EU presents many opportunities to its member states. Apart from its broad social and political influence on them, the opportunities for these states are

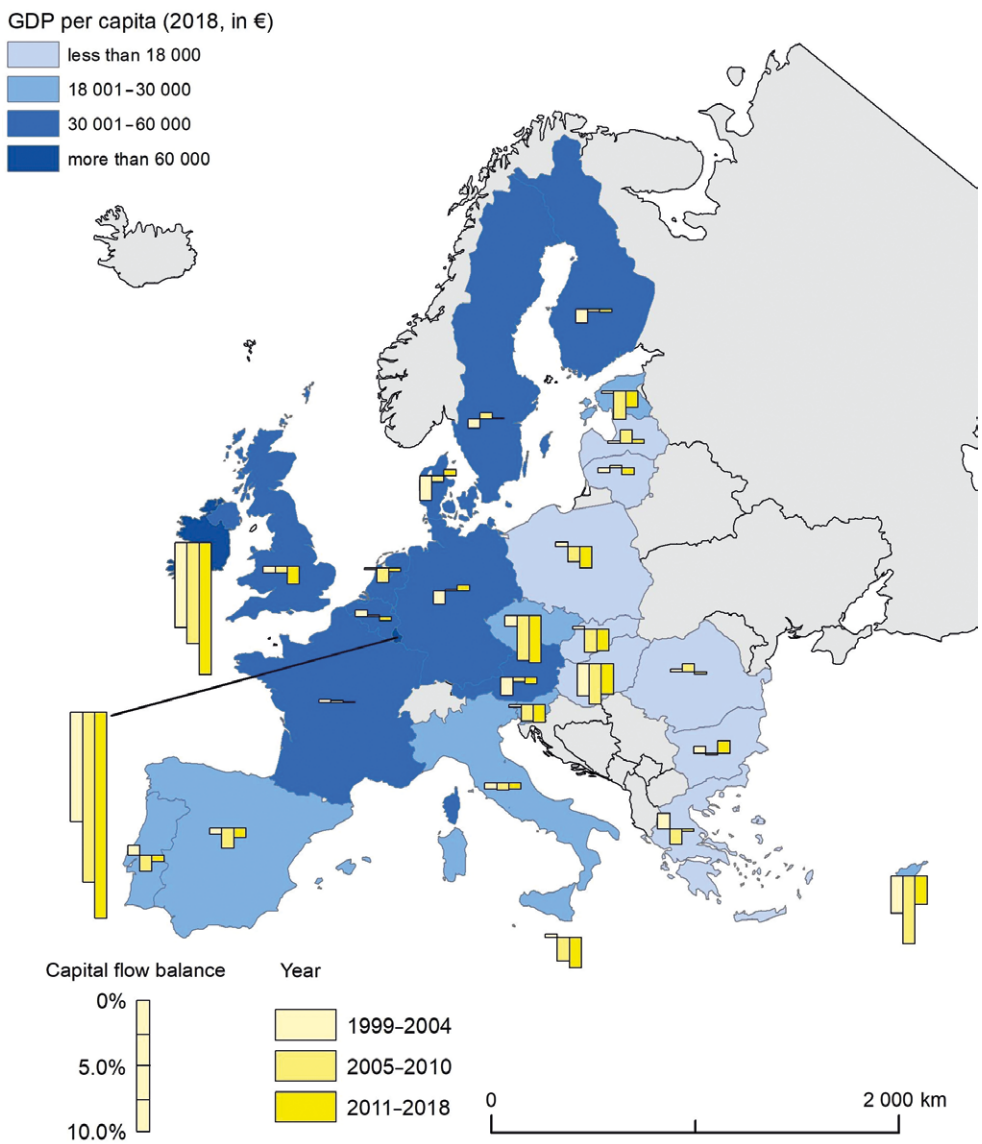


Fig. 1 GDP per capita and average yearly balance of financial transactions with non-residents (2018 in €; 1999–2018 in % of GDP).
Source: Eurostat 2019; authors' calculations.

to influence world political events and to participate intensively in the future of Europe, as well as the EU's emphasis on human rights and peaceful dispute resolution, the economic aspect mainly revolves around access to the EU market and EU funds. However, there are also some "hidden" negative consequences that could have crucial adverse impacts on economic performance, EU convergence, and the economic and political development of some countries. One such consequence is capital drain, best reflected in the difference between GDP and GDI. This difference, relative to the GDP amount, is displayed in Fig. 1 for three different periods: 1999–2004 as the period before large-scale accessions to the EU, 2005–2010 as the crisis period, and 2011–2018 as the period after the end of the global financial crisis and before the COVID-19 pandemic.

In Fig. 1, we can observe several extremes. Ireland and Luxembourg are the largest ones. These two countries are top European tax havens (Delate 2022; Hána 2022), and thanks to their favorable tax system, multinational companies may reroute their profits to these countries, resulting in their GDP being much higher than GDI. This example thus emphasizes the need for careful interpretation, as sometimes the negative difference between GDP and GDI can be caused by capital drain, where a part of GDP flows abroad (and therefore GDI is less than GDP), while other times it can be caused by a given country significantly increasing its GDP by receiving various types of capital flows from abroad. A similar example can be noted to a lesser extent in Cyprus, which is also considered one of the European tax havens (Delate 2022; Hána 2022). In contrast, this difference between the two indicators does not appear in the case of the Netherlands, although according to some sources, it also exhibits characteristics of a tax haven (Delate 2022; Hána 2022).

By 2004, we can generally say that the higher the FDI, the greater the GDP growth, and the more massive the outflow of capital in the form of dividends from the host country that reduces the value of FDI remaining in the investing economy (Kučera 2016). Except for France, Portugal, Belgium, and Greece, all EU-15 countries had negative balances in this period 1999–2004. This can again be explained in two ways. The example of Spain as a top receiver of FDI (Carbonell and Werner 2018) shows that a negative balance can be caused by the positive balance of investment flow. Basically, the same situation is typical for Central and Eastern European countries in this period (Mahutga and Bandelj 2008; Simionescu et al. 2017) which opened their economies to FDI mainly from Western European countries (which may cause negative values in Western Europe). However, it has to be noted that even Western Europe is not exempt from capital drain, and it may influence negative values in this region as well. The EU's annual balance is negative, with an average of 150 billion euros 'flowing away' from it annually (Eurostat 2019; authors' calculations). The question is, where to (there is a possible influence of the flow to non-European tax havens or to the U.S. and Eastern-Asian investments, which should be researched in more detail). Hungary's position is quite special. We could assume that its significant negative balance in this period is due to a different liberalization process in Hungary, which began before 1989 during the last decade of the Communist regime when it allowed joint ventures with foreign firms and later legalized their foreign ownership in the 1980s (Mahutga and Bandelj 2008). Similarly, its privatization process in the transition period was faster than in other countries (e.g., Bonin et al. 2005). In the studied period, therefore, there could already have been massive capital drain from Hungary.

To a large extent, the 2005–2010 period can be considered a transition stage. As it is the period after the EU accession of new member states in 2004 and the opening of the labor market (with a transition period of several years in some countries), we can observe, in addition to the already mentioned FDI and tax havens, the important role of EU funds for the resulting balance and the influence of remittances as well. Some new member states (such as Latvia, Lithuania, and Romania) experienced a high level of economic emigration, bringing back a considerable amount to the country, manifested in GDI increase. On the other hand, a capital drain in the Central European countries grew stronger. The situation since 2011 is more or less similar to the previous one. Development in Denmark is remarkable. In the first period observed, Denmark ended up with a significantly negative balance, but currently, the balance is positive. This example can demonstrate the general characteristic cycle. In the first stage, a western country is an investor whose investments do not yield any profits yet. In the second stage, the balance is settled, and in the final stage, the country reaches a positive balance of transactions with foreign countries. Czechia can be seen as an opposite example: the massive inflow of FDI from the 1990s caused the growth of negative balance, making it the economy with the most negative balance of transactions with non-residents, after omitting the tax havens' outliers.

4.1 Economic inequalities measured on the basis of GDP

First, let us briefly examine the level and development of economic inequalities based on GDP, which is a macroeconomic indicator that does not reflect a capital drain. Fig. 2 shows data from 1999 to 2018, with the Gini index starting at 0.40, indicating very high inequalities in the observed area. A lower level of the population-weighted form of the Gini coefficient can be explained as follows. The new member states (referring to states that joined the EU since 2004) represent only about 20% of the total population of

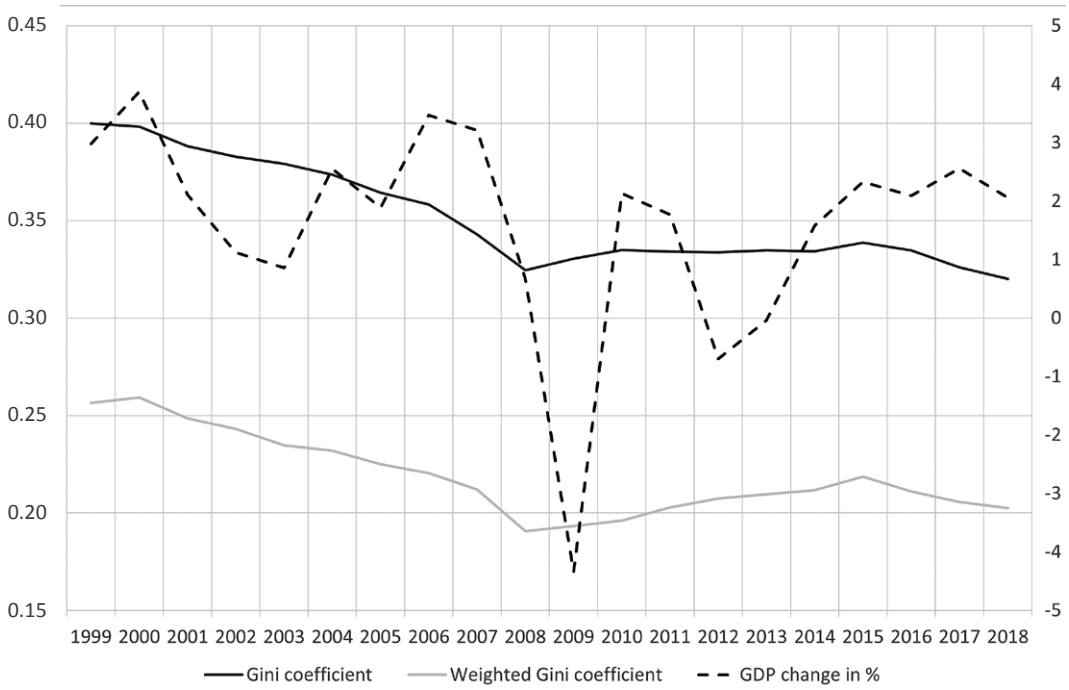


Fig. 2 Gini coefficient of GDP per capita for EU-27 (1999–2018).
Source: Eurostat 2019; authors' calculations.

the EU-27 (Eurostat 2019; author's calculations), which reduces the weight of the overall polarization between the old and new EU members significantly.

The first stage, by 2004, is defined as the EU accession preparation stage of the candidate countries. We can observe a sharp decline in inequalities, explained by the massive inflow of capital into Central and Eastern Europe and the introduction of more advanced manufacturing technologies in these countries. In the second stage, from 2004 to 2008, GDP continued to grow, and the Gini index dropped from 0.37 in 2004 to 0.33 in 2008, which can be considered a great success of the EU in reducing disparities.

The third stage, from 2008 to 2015, influenced by the global economic crisis (2008–2009), saw slightly growing inequalities. It is not necessarily just the weak impact of the EU cohesion policy; we must consider the fact that periods of crisis always have a negative impact on increasing inequalities (Novotný 2006; Goda 2018). The Greek rescue package and the establishment of the European Stability Mechanism (ESM) by the Eurozone member states might have been important factor as well. The duration of the rescue process is similar to the period during which we noticed the growth of inequalities. The drop in Greece's GDP, gradually drifting away from the average, was increasing the overall dispersion of values, which consequently has an impact on statistical indicators of inequalities. We must also consider the problems of other countries, such as Spain, Portugal, Italy, and Ireland. We can see, mainly from the population-weighted form of the inequality coefficient, how much Spain and Italy, countries with large populations, dropped by 100% of the EU average in that period, which again increases the dispersion of values and the aggregate coefficient. In the case of Ireland, its drop in GDP between 2008 and 2014 can be seen as quite the opposite because its development actually appears to be approaching the average, reducing the dispersion of GDP values. For the record, Eurostat indicates that Ireland dropped from 148% of the EU average in 2007 to 129% in 2009. In 2018, it grew again to 181%, this time resulting in divergence compared to other countries. Considering that many countries experienced a drop below the average, we have to ask who was growing above the EU average in the last period. It was Germany, which grew from 117% of the average in 2007 to 126% in 2015, Denmark, which grew from 123% to 128% in 2014, and Hungary, which grew towards the EU average between 2008 and 2015 by 8 percentage points. Recession in the Baltic states, prolonged stagnation in Slovenia, and Czechia during the post-crisis years resulted in divergent movements of European economies.

The last stage, from 2015 to 2018, is characterized by a reduction in overall disparities. GDP growth values exceed 2% again, which has a positive impact on the overall reduction of disparities and overall convergence. We must note that the EU-27 did not manage to reach the inequality values before the global economic crisis until 2018, when the observed coefficients already show slightly lower values than in 2008. However, based on the authors' calculations,

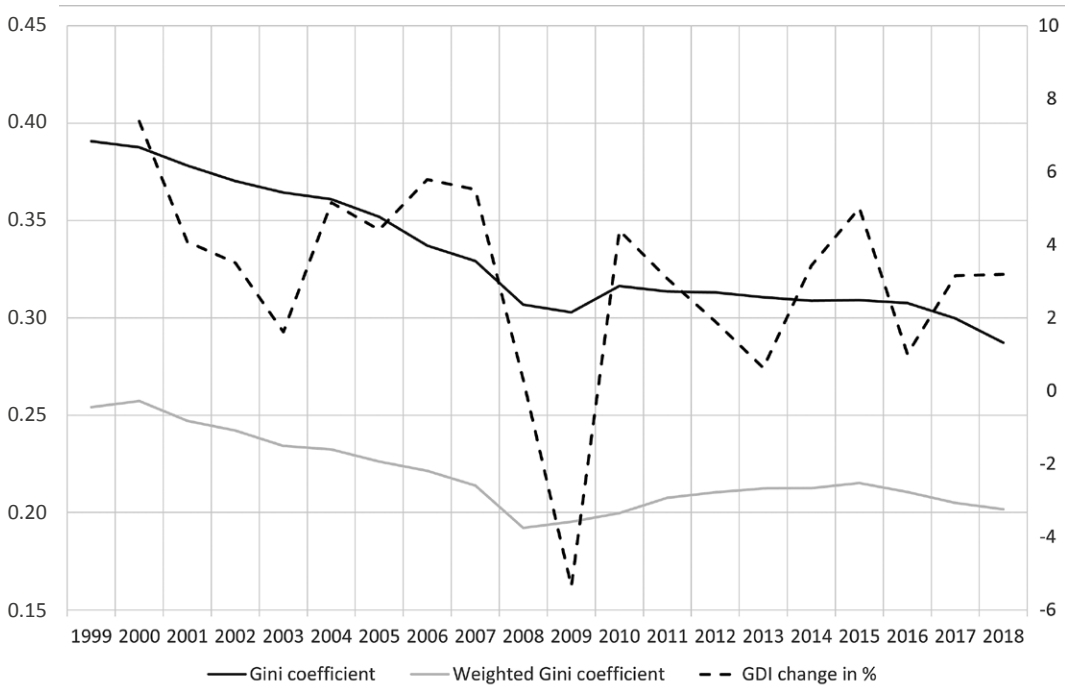


Fig. 3 Gini coefficient of GDI per capita for EU-27 (1999–2018).
Source: Eurostat 2019; authors' calculations.

the Gini index for the former EU-15 in 2018 was 0.24, compared with a mere 0.18 in 1999.

4.2 Economic inequalities measured on the basis of GDI

Inequalities measured based on GDI led to an interesting observation and surprising results at first glance. When comparing the charts in Fig. 2 and 3, it is possible to conclude that the level of disparities measured by GDI is slightly lower than when using GDP. However, the explanation is clear. Upon closer examination of the outliers, Ireland and Luxembourg, we can see their significant decrease (Ireland falls by 20%, and Luxembourg by 30%), which, in turn, results in the reduction of dispersion from the average and a decrease in variability rates. A similar situation is observed at the opposite extreme in the second period, in the case of Romania, and in the third period, in the case of Bulgaria, with positive balances of transactions with non-residents, which again leads to a reduction of dispersion from the average compared to the variability measured from GDP. After disregarding the outliers, Ireland and Luxembourg, we obtain an opposite result. However, the difference in the Gini coefficient of GDI is higher by only 3 thousandths compared to GDP. The development trends in inequalities measured by GDI are almost identical to those measured by GDP. In the observed period from 1999 to 2018, convergence between EU member states was revealed.

However, the aggregate index of inequalities within the entire EU does not indicate the development of individual countries or regions. Fig. 4 illustrates how the GDI value has changed in relation to the EU average over the three periods from Fig. 1. In simple terms, a reduction of European disparities would occur if richer countries achieved values below 100, and poorer countries achieved values above 100. Richer countries, in most cases, move slightly below the value of 100. But their distance from the EU average is not so crucial as to claim that it contributes to reducing disparities in the EU. Moreover, in several cases, there is an increase (above 100) in the distance from the EU average in some richer countries (e.g., Belgium, Germany, Netherlands, and Austria). Conversely, there are a few positive cases of poorer countries approaching the EU average (e.g., high values of Bulgaria, Romania, Lithuania, and Latvia), but it slows down over time. It corresponds to the sharp reduction of inequalities in 1999–2008 in Fig. 3 and could be the effect of incoming FDI in the first period and the subsequent capital drain, which has slowed the divergence with the core region. In several cases, the values of these countries are around 100 or even below it (e.g., Hungary in 2005–2010 and Poland in 1999–2004).

In the discussion of the world system theory, we came across a difference between the old and new member states. The development of new member states, which we can describe as peripheries, and their convergence with the old and economically more advanced member states from the European core (Storper 2018; Pavlínek 2022a) needs to be considered as one of the priorities of the European cohesion policy, as stated in one of the specific goals stated already in the Treaty of Rome. In Fig. 5, we see

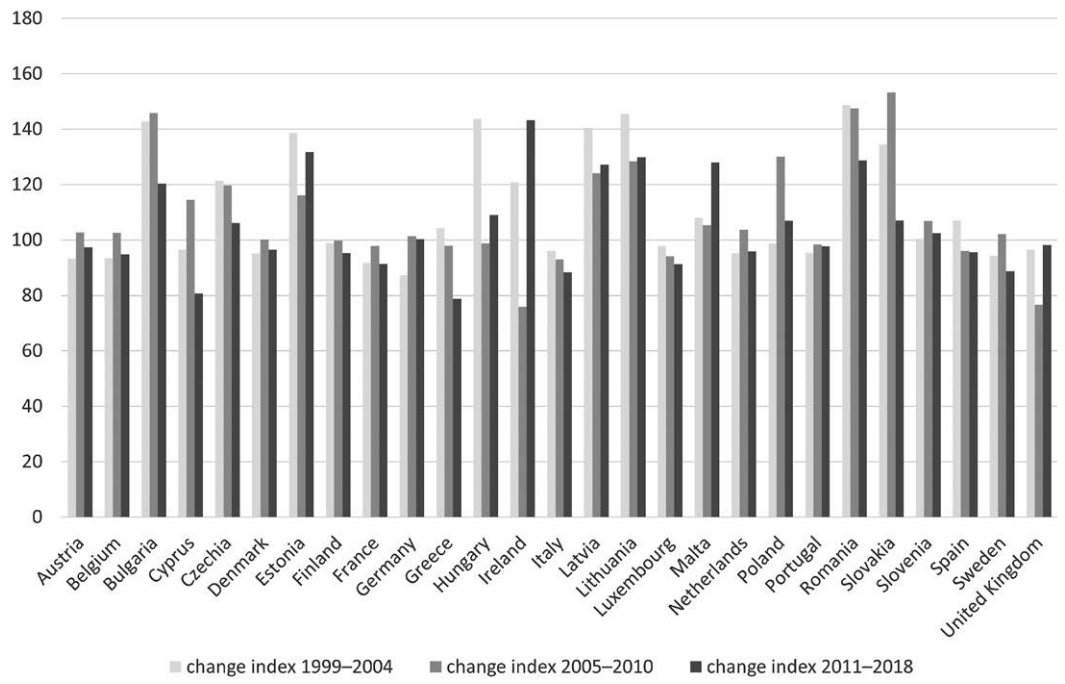


Fig. 4 Change index of the annual GDI development of individual states compared to the annual European Union average (1999–2018).
Source: Eurostat 2019; authors' calculations.

a continuous increase in the share of new member states in the total GDP and GDI of the EU. Again, we can confirm that there is a certain degree of convergence between the old and new member states. Nevertheless, the share of new member countries still does not correspond to the share of their population, which we could then call a sign of equalization of inequalities within the EU. If the inhabitants of these countries have fewer wealth resources than the rest of the EU, we can still consider them the poorer part of the EU. Moreover, we can see a difference between the share of GDP and GDI as well. A significant part of the produced capital is drained from these countries, reducing the disposable wealth made by their labor.

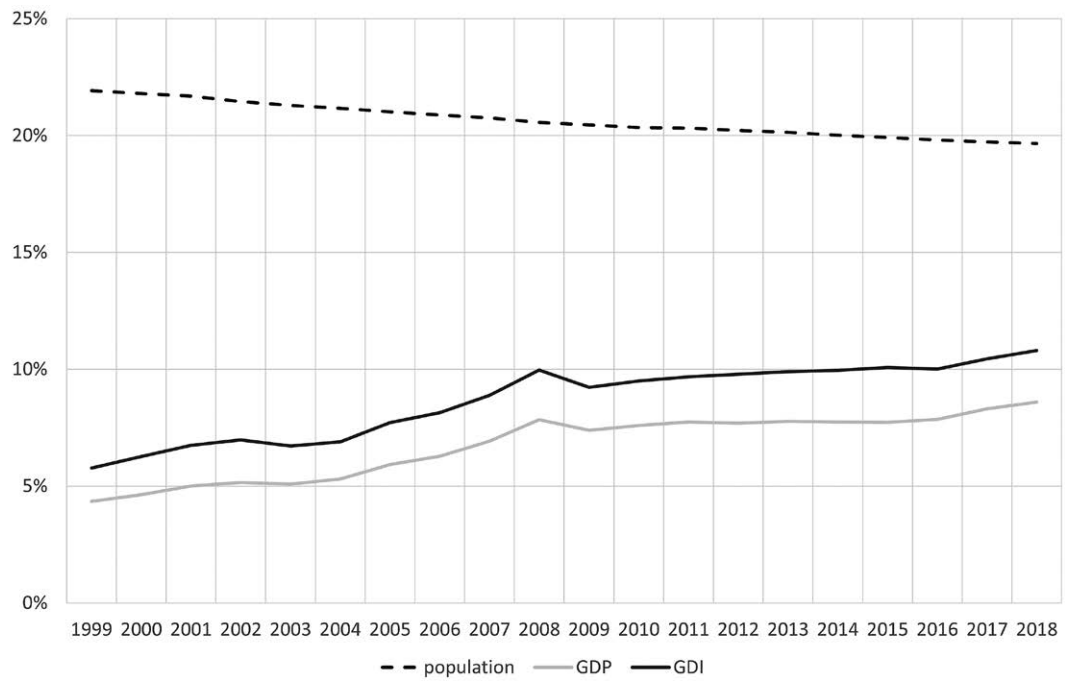


Fig. 5 Shares of new EU member states in selected attributes (1999–2018).
Source: Eurostat 2019; authors' calculations.

5. Discussion about cohesive and political consequences

The available literature offers several interpretations of inequalities in the EU territory. However, almost none of them also evaluate the hidden capital drain that is not shown in the value of the commonly used GDP. Surprisingly, when using GDI, which accounts for the capital drain that, according to world systems theory, flows from the periphery to the core (Myrdal 1957; Holubec 2009; Sorinel 2010; Wallerstein 2011), inequalities in the EU are smaller than when using GDP. However, it is necessary to add that this is primarily due to the distorting extremes of Ireland and Luxembourg, the countries with the highest GDP in the EU and the biggest difference between GDI and GDP. It highlights the fact that both countries are tax havens, which fundamentally affects the level of GDP (due to profit shifting by corporations) and the subsequent comparison. If we disregard these outliers, we get an opposite result of the Gini coefficient of GDI, which is, however, only slightly higher compared to GDP. At the same time, we can see a considerable reduction in inequalities both in the case of counting GDP and GDI. Although this stopped in 2008, and since then there has been a slight increase or stagnation caused by the crisis (a common characteristic of crises, as see in Novotný 2006; Goda 2018), in the last years of the observed period, the values of the Gini coefficient have fallen again. However, this decrease in inequalities is not at the same pace as before 2008.

However, this generalized view cannot reveal all the details associated with capital drain between countries or EU regions. Many countries from the European core and periphery have GDIs that approach the European average from both sides. In light of this, we can conclude that some degree of economic convergence among member states is occurring in the EU. However, there are a few countries within the European core that are extending their lead over the EU average in GDI (Belgium, Germany, Netherlands, Austria). Conversely, there are a few countries that are moving away from the mean (Hungary, Poland, but only in certain periods). Based on the difference of GDP and GDI shares in the new EU members states (accessed since 2004), we may see that capital drain reduces the statistics of their economic performance. Therefore, we can say that capital drain does not have a major impact at the pan-European level but has a negative impact on some new member states that have less available capital or wealth than they have generated through their own labor (Kučera 2016).

There are important cohesive and political implications of this capital drain. If we compare periods before and after 2004 when the new member states accessed the EU, the convergence was only slightly slower, and GDI growth was only slightly lower in the first period. This calls into question the impact of EU membership in reducing inequalities, which should be one of its essential political tasks according to the founding treaties. Is there this trend, or is the EU riddled with capital drain that disrupts any political efforts to support the peripheries with a more difficult position in the common market? It is necessary to note that the effort to invest and the use of profit is quite natural in capitalism and not a negative phenomenon in principle. However, in the global economy and international relations, leading firms from the world's economic core control value creation and capital flows regardless of any other interests unless government policy sets out clear regulations (Henderson et al. 2002; Gereffi 2005; Pavlínek and Ženka 2011). Thanks to the hiddenness of capital drain, it can exceed a tolerable level and can significantly disrupt the positions and well-being of some countries when they cannot use the wealth they have created on their territory to a large extent. However, the crisis period from 2008 to 2015 is very risky when evaluating the efficiency of the cohesion policy. The growth of inequalities might be seen as its failure, but there could also be the potential positive impact on the moderation of European inequalities growth during the crisis. Thus, the question remains about how EU inequalities will continue to change after the end of the series of crises associated with the COVID-19 pandemic and the Russian invasion of Ukraine.

Evaluating the capital drain, one of the highest levels of negative balance for Czechia requires consideration. The 7% outflow of produced assets is significantly above the European average. Therefore, it is necessary to resolve this situation (e.g., by reforming the Czech tax system or introducing a common EU tax system) and open a political discussion about this phenomenon nationally and in the EU as well (Chmelař et al. 2016). The reason is crucial. The reduction of disparities under these conditions might not be regarded as positive by voters in the new member states. These states can reach the average income of the EU in 40–80 years; it is not so hopeful an outlook, which may lead to a wave of populist approaches promising a faster convergence process using less acceptable interventions (Mihaljek 2018). In more general terms, strongly unequal societies tend to disintegrate democracy by voting authoritarian leaders and to become unstable societies (Tridico 2018), which is now threatening Europe (Weeks 2018; Tismaneanu 2019), mainly the Central European region (Plenta 2020; Mravcová and Havlík 2022).

6. Conclusion

The aim of this paper was to analyze and discuss capital drain in the EU and its cohesive and political impact. Once capital drain is taken into account, the level of inequality is slightly lower. This suggests that the inflow or outflow of capital in the form of dividends does not significantly contribute to increasing

disparities within the EU. However, it does have a crucial impact on some states in the EU's periphery, from which a relatively large amount of capital is drained. This may lead to negative political consequences, including rising instability and the emergence of populist political parties that promise radical interventions. Such a political situation is particularly prevalent in Central European countries, which serve as the primary source of capital drain. This article seeks to highlight the necessity of discussing the current situation at both the national and EU levels, with a focus on tax system reforms and the potential introduction of a common European tax policy aimed at curbing this capital drain. By doing so, it may be possible to meet the expectations of new member states and the lofty objectives set forth upon the founding of the EU with the signing of the Treaty of Rome in 1957.

Acknowledgements

This work was supported by the Technology Agency of the Czech Republic under Grant TL02000195.

References

Alfaro, L., Chanda, A., Kalemli-Ozcan, S., Sayek, S. (2004): FDI and economic growth: the role of local financial markets. Journal of International Economics 64(1), 89–112, https://doi.org/10.1016/S0022-1996(03)00081-3.

Alvaredo, F., Chancel, L., Piketty, T., Saez, E., Zucman, G. (2018): World Inequality Report 2018. Cambridge, MA and London, England: Harvard University Press, https://doi.org/10.4159/9780674984769.

AP News (2019): Poorer EU members oppose cuts in long-term budget. AP News, 5 November 2019, https://apnews.com/679ba30dd20440bfa066d7108a2f009f (accessed on 30 January 2020).

Becker, S. O., Egger, P. H., Von Ehrlich, M. (2010): Going NUTS: The effect of EU Structural Funds on regional performance. Journal of Public Economics 94(9–10), 578–590, https://doi.org/10.1016/j.jpubeco.2010.06.006.

Bečicová, I., Blažek, J. (2015): Is there a credit-gap in a periphery? The perception of this problem by small entrepreneurs. Journal of Rural Studies 42, 11–20, https://doi.org/10.1016/j.jrurstud.2015.09.006.

Boldrin, M., Canova, F. (2001): Inequality and convergence in Europe's regions: reconsidering European regional policies. Economic Policy 16(32), 206–253, https://doi.org/10.1111/1468-0327.00074.

Bonin, J. P., Hasan, I., Wachtel, P. (2005): Privatization matters: Bank efficiency in transition countries. Journal of Banking & Finance 29(8–9), 2155–2178, https://doi.org/10.1016/j.jbankfin.2005.03.012.

Borensztein, E., DeGregorio, J., Lee, J. W. (1998): How does foreign direct investment affect economic growth? Journal of International Economics 45(1), 115–135, https://doi.org/10.1016/S0022-1996(97)00033-0.

Cappelen, A., Castellacci, F., Fagerberg, J., Verspagen, B. (2003): The impact of EU regional support on growth and convergence in the European Union. JCMS: Journal of Common Market Studies 41(4), 621–644, https://doi.org/10.1111/1468-5965.00438.

Carbonell, J. B., Werner, R. A. (2018): Does Foreign Direct Investment Generate Economic Growth? A New Empirical Approach Applied to Spain. Economic Geography 94(4), 425-456, https://doi.org/10.1080/00130095.2017.1393312.

Chmelař, A., Pícl, M., Bittner, J., Volčík, S., Nechuta, A. (2016): Analýza odlivu zisků: Důsledky pro českou ekonomiku a návrhy opatření [Analysis of the outflow of profits: Consequences for the Czech economy and proposals for measures]. Úřad vlády České republiky, Praha.

Coe, N. M., Hess, M., Yeung, H. W., Dicken, P., Henderson, J. (2004): 'Globalizing' Regional Development: A Global Production Networks Perspective. Transactions of the Institute of British Geographers 29(4), 468–484, https://doi.org/10.1111/j.0020-2754.2004.00142.x.

Cuaresma, J. C., Ritzberger-Grünwald, D., Silgoner, M. A. (2008): Growth, convergence and EU membership. Applied Economics 40(5), 643–656, https://doi.org/10.1080/00036840600749524.

Dall'Erba, S., Fang, F. (2017): Meta-analysis of the impact of European Union Structural Funds on regional growth. Regional Studies 51(6), 822–832, https://doi.org/10.1080/00343404.2015.1100285.

Dall'Erba, S., Le Gallo, J. (2008): Regional convergence and the impact of European structural funds over 1989–1999: A spatial econometric analysis. Papers in Regional Science 87(2), 219–244, https://doi.org/10.1111/j.1435-5957.2008.00184.x.

Delatte, A.-L., Guillin, A., Vicard, V. (2022): Grey zones in global finance: The distorted geography of cross-border investments. Journal of International Money and Finance 120(102540), https://doi.org/10.1016/j.jimonfin.2021.102540.

Eckey, H. F., Türck, M. (2005): Convergence of EU-Regions. A Literature Report. Investigaciones Regionales 10(1), 5–32.

Ederveen, S., Groot, H. L., Nahuis, R. (2006): Fertile soil for structural funds? A panel data analysis of the conditional effectiveness of European cohesion policy. Kyklos 59(1), 17–42, https://doi.org/10.1111/j.1467-6435.2006.00318.x.

Esposti, R., Bussoletti, S. (2008): Impact of Objective 1 funds on regional growth convergence in the European Union: a panel-data approach. Regional Studies 42(2), 159–173, https://doi.org/10.1080/00343400601142753.

European Commission (2014): An introduction to EU Cohesion Policy 2014–2020, https://ec.europa.eu/regional_policy/sources/docgener/informat/basic/basic_2014_en.pdf (accessed on 16 September 2019).

Eurostat (2013): European system of accounts ESA 2010, https://ec.europa.eu/eurostat/documents/3859598/5925693/KS-02-13-269-EN.PDF/44cd9d01-bc64-40e5-bd40-d17df0c69334 (accessed on 16 September 2019).

Eurostat (2019): Eurostat database, https://ec.europa.eu/eurostat/data/database (accessed on 9 May 2019).

Fiala, P., Krutílek, O., Pitrová, M. (2018): Evropská unie [European Union]. Centrum pro studium demokracie a kultury, Brno.

Gereffi, G. (2005): The Global Economy: Organization, Governance, and Development. In: Smelser, N. J., Swedborg, R. (Eds.): The Handbook of Economic Sociology, Second Edition. Princeton University Press, Princeton, 160–182, https://doi.org/10.1515/9781400835584.160.

Goda, T. (2018): The global concentration of wealth. Cambridge Journal of Economics 42(1), 95–115, https://doi.org/10.1093/cje/bex020.

Goulet, R. (2011): Cohesion Policy 2014-2020. Investing in growth and jobs. European Commission, Directorate-General for Regional Policy, Brussels.

Grela, M., Majchrowska, A., Michałek, T., Mućk, J., Stążka-Gawrysiak, A., Tchorek, G., Wagner, M. (2017): Is Central and Eastern Europe converging towards the EU-15? Narodowy Bank Polski, Warszawa.

Gros, D. (2018): Convergence in the European Union: Inside and outside the euro. Informal meeting of Economic and Financial Affairs Ministers, Sofia, Bulgaria, 27–28 April 2018.

Hakenes, H., Schnabel, I. (2010): The threat of capital drain: A rationale for regional public banks? Journal of Institutional and Theoretical Economics JITE 166(4), 662–689, https://doi.org/10.2139/ssrn.906993.

Hána, D. (2022): Daňové ráje v globálním geografickém pohledu [Tax havens from a global geographical perspective]. Geografické rozhledy 31(4), 32–33.

Hána, D., Hellebrandová, L. (2018): Spatial and Sectoral Differentiation of Support to Innovative Companies from EU Funds in Czechia. European Planning Studies 26(8), 1598–1615, https://doi.org/10.1080/09654313.2018.1485135.

Harvey, D. (2010): The Enigma of Capital and the Crises of Capitalism. Oxford University Press, New York.

Henderson, J., Dicken, P., Hess, M., Coe, N., Yeung, H. W.-C. (2002): Global production networks and the analysis of economic development. Review of International Political Economy 9(3), 436–464, https://www.jstor.org/stable/4177430.

Hlaváček, P., Bal-Domanska, B. (2016): Impact of Foreign Direct Investment on Economic Growth in Central and Eastern European Countries. Inzinerine Ekonomika-Engineering Economics 27(3), 294–303, https://doi.org/10.5755/j01.ee.27.3.3914.

Holubec, S. (2009): Sociologie světových systémů: hegemonie, centra, periferie [Sociology of world systems: hegemony, centers, periphery]. Sociologické nakladatelství (SLON), Praha.

Javorcik, B. S. (2004): Does foreign direct investment increase the productivity of domestic firms? In search of spillovers through backward linkages. American Economic Review 94(3), 605–627, https://doi.org/10.1257/0002828041464605.

Jensen, N. M. (2003): Democratic governance and multinational corporations: Political regimes and inflows of foreign direct investment. International Organization 57(3), 587–616, https://doi.org/10.1017/S0020818303573040.

Keller, J. (2017): Evropské rozpory ve světle migrace [European contradictions in the light of migration]. Sociologické nakladatelství (SLON), Praha.

Keynes, J. M. (2018): The General Theory of Employment, Interest and Money. Palgrave Macmillan, Cambridge, https://doi.org/10.1007/978-3-319-70344-2.

Kučera, L. (2016): Odliv peněz z ČR je jeden z nejsilnějších v Unii [The outflow of money from the Czech Republic is one of the strongest in the Union]. Statistika & my 2, https://www.statistikaamy.cz/2016/02/odliv-penez-z-cr-je-jeden-z-nejsilnejsich-v-unii/ (accessed on 23 March 2019).

Kyriacou, A. P., Roca-Sagalés, O. (2012): The impact of EU structural funds on regional disparities within member states. Environment and Planning C: Government and Policy 30(2), 267–281, https://doi.org/10.1068/c11140r.

Li, Q., Resnick, A. (2003): Reversal of fortunes: Democratic institutions and foreign direct investment inflows to developing countries. International Organization 57(1), 175–211, https://doi.org/10.1017/S0020818303571077.

Lolos, S. E. G. (2009): The effect of EU structural funds on regional growth: assessing the evidence from Greece, 1990–2005. Economic Change and Restructuring 42(3), 211–228, https://doi.org/10.1007/s10644-009-9070-z.

Magrini, S. (1999): The evolution of income disparities among the regions of the European Union. Regional Science and Urban Economics 29(2), 257–281, https://doi.org/10.1016/S0166-0462(98)00039-8.

Mahutga, M. C., Bandelj, N. (2008): Foreign Investment and Income Inequality. The Natural Experiment of Central and Eastern Europe. International Journal of Comparative Sociology 49(6), 429–454, https://doi.org/10.1177/0020715208097788.

Mihaljek, D. (2018): Convergence in Central and Eastern Europe: Can All Get to EU Average? Comparative Economic Studies 60(2), 217–229, https://doi.org/10.1057/s41294-018-0063-7.

Ministry of Regional Development (2023): Drawing in the 2014–2020 period, https://dotaceeu.cz/en/statistiky-a-analyzy/cerpani-v-obdobi-2014-2020 (accessed on 29 May 2023).

Molle, W. (2017): The economics of European integration: theory, practice, policy. Routledge, London, https://doi.org/10.4324/9781315240121.

Mravcová, H., Havlík, V. (2022): Pragmatism and support for the EU in Slovakia's politics. East European Politics 38(1), 123–143, https://doi.org/10.1080/21599165.2021.1929187.

Myrdal, G. (1957): Economic Theory and Under-developed Regions. Gerald Duckwords, London.

Nerudová, D., Dobranschi, M., Solilová, V., Litzman, M. (2023): Onshore and offshore profit shifting and tax revenue losses in the European Union. Economic Modelling (early access), https://doi.org/10.1016/j.econmod.2022.106111.

Novotný, J. (2006): Negativní vlivy společensko-ekonomických nerovností a mechanismy jejich regulace: argumenty z rozvojových zemí [The Negative Aspects of Socio-economic Inequalities and Mechanisms of their Regulation: Some Arguments from Developing Countries]. Ekonomický časopis 54(7), 709–724.

Novotný, J., Nosek, V., Jelínek, K. (2014): EasyStat, http://web.natur.cuni.cz/~pepino/EasyStat.zip (accessed on 16 September 2019).

Pavlínek, P. (2022a): Relative positions of countries in the core-periphery structure of the European automotive

industry. European Urban and Regional Studies 29(1), 59–84, https://doi.org/10.1177/09697764211021882.
Pavlínek, P. (2022b): Revisiting economic geography and foreign direct investment in less developed regions. Geography Compass 16(4), 1–21, https://doi.org/10.1111/gec3.12617.
Pavlínek, P., Ženka, J. (2011): Upgrading in the automotive industry: firm-level evidence from Central Europe. Journal of Economic Geography 11(3), 559–586, https://doi.org/10.1093/jeg/lbq023.
Piketty, T. (2014): Capital in the Twenty-First Century. Harvard University Press, Cambridge, https://doi.org/10.4159/9780674369542.
Piketty, T., Saez, E. (2014): Inequality in the long run. Science 344(6186), 838–843, https://doi.org/10.1126/science.1251936.
Plenta, P. (2020): Conspiracy theories as a political instrument: utilization of anti-Soros narratives in Central Europe. Contemporary Politics 26(5), 512–530, https://doi.org/10.1080/13569775.2020.1781332.
Puigcerver-Peñalver, M. C. (2007): The impact of structural funds policy on European regions' growth. A theoretical and empirical approach. The European Journal of Comparative Economics 4(2), 179–208.
Rojíček, M., Spěváček, V., Vejmělek, J., Zamrazilová, E., Žďárek, V. (2016): Makroekonomická analýza – teorie a praxe [Macroeconomic analysis – theory and practice]. Grada, Praha.
Sala-i-Martin, X. X. (2002): The Disturbing "Rise" of Global Income Inequality. National Bureau of Economic Research, Cambridge, https://doi.org/10.3386/w8904.
Shahbaz, M., Nasir, M. A., Roubaud, D. (2018): Environmental degradation in France: The effects of FDI, financial development, and energy innovations. Energy Economics 74, 843–857, https://doi.org/10.1016/j.eneco.2018.07.020.
Simionescu, M., Lazányi, K., Sopková, G., Dobeš, K., Balcerzak, A. P. (2017): Determinants of Economic Growth in V4 Countries and Romania. Journal of Competitiveness 9(1), 103–116, https://doi.org/10.7441/joc.2017.01.07.
Smith, N. (2008): Uneven Development: Nature, Capital, and the Production of Space. University of Georgia Press, Athens, https://doi.org/10.1353/book11443.
Sorinel, C. (2010): Immanuel Wallerstein's World System Theory. Annals of Faculty of Economics 1(2), 220–224, http://anale.steconomiceuoradea.ro/volume/2010/n2/031.pdf (accessed on 30 January 2020).
Storper, M. (2018): Separate worlds? Explaining the current wave of regional economic polarization. Journal of Economic Geography 18(2), 247–270, https://doi.org/10.1093/jeg/lby011.
Tismaneanu, V. (2019): What Went Wrong and Why? Nationalism versus Democracy in Eastern and Western Europe. Contemporary European History 28(1), 69–72, https://doi.org/10.1017/S096077731800084X.
Treaty establishing the European Economic Community (Treaty of Rome, 25 Mar 1957).
Tridico, P. (2018): The determinants of income inequality in OECD countries. Cambridge Journal of Economics 42(4), 1009–1042, https://doi.org/10.1093/cje/bex069.
Wallerstein, I. (2011): The Modern World-System I: Capitalist Agriculture and the Origins of the European World-Economy in the Sixteenth Century. University of California Press, San Francisco, https://doi.org/10.1525/9780520948570.
Weeks, J. (2018): Free Markets and the Decline of Democracy. Review of Radical Political Economics 50(4), 637–648, https://doi.org/10.1177/0486613418772167.

Improving vegetation spatial distribution mapping in arid and on coastal dune systems using GPR in Tottori Prefecture (Japan)

Christopher Gomez[1,2], Jiaqi Liu[3], Jing Wu[4], Frans Persendt[5], Balazs Bradak[6], Yousefi Saleh[7], Danang Sri Hadmoko[2,*]

[1] Kobe University, Faculty of Oceanology, Laboratory of Sediment Hazards and Disaster Risk, Japan
[2] Universitas Gadjah Mada, Department of Geography, PSBA Laboratory, Yogyakarta, Indonesia
[3] Tottori University, Dry Land Research Centre, Japan
[4] Tsukuba Meteorological Research Institute, Japan
[5] University of Namibia, Namibia
[6] Kobe University, Faculty of Oceanology, Laboratory of Extraterrestrial Oceanography, Japan
[7] Soil Conservation and Watershed Management Research Department, Chaharmahal and Bakhtiari Agricultural and Natural Resources Research and Education Center, Iran
[*] Corresponding author: hadmoko@ugm.ac.id

ABSTRACT

In this article, desertification and dune progression over vegetation was quantified using remote sensing data. However, vegetation buried under sand blowout could not be counted using this method. Therefore, to estimate the extent of buried vegetation, a GPR campaign was conducted over the coastal sand-dune of Tottori Prefecture (Japan) in combination with a high-resolution topographic UAV-based survey of the topography. The results show that buried vegetation exists underneath sand-blowout, especially near the dune ridges, and can extend from 20 to 30 meters further than the estimate based on airborne remote sensing. Furthermore, the presence of palaeo-vegetation in palaeodune layers also provides the information on the long-term evolution of sand dunes, which can be used to reconstruct Quaternary coastal environments.

KEYWORDS

coastal dune; ground penetrating radar; buried vegetation; vegetation mapping

Received: 27 March 2023
Accepted: 11 December 2023
Published online: 18 December 2023

Gomez, C., Liu, J., Wu, J., Persendt, F., Bradak, B., Saleh, Y., Hadmoko, D. S. (2023): Improving vegetation spatial distribution mapping in arid and on coastal dune systems using GPR in Tottori Prefecture (Japan). AUC Geographica 58(2), 238–249
https://doi.org/10.14712/23361980.2023.18
© 2023 The Authors. This is an open-access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0).

1. Introduction

Desertification is one of the challenging results of climate-change and anthropogenic land-degradation (Kasas 1995; Stringer 2008). It threatens biodiversity (Musila et al. 2001), modify climates at the regional scale and can even destabilize food production and economic chains.

The progression of sand against vegetation is particularly acute on coastal sand dune, where urbanization, recreational and touristic activities, erosion from climate change, from anthropogenic origin, or from extreme events (Opelt and Berg 2004; Lee et al. 2006; Lavigne et al. 2009) are all adding supplementary pressure to the bio-geomorphologic system. Despites negative impacts from a human ethics and economic perspective, aeolian sand movement on coastal dunes is a natural process that responds to climate fluctuations, and without human intervention, vegetation is also controlling the erosion from waves runup (Feagin et al. 2019), resulting in complex systems, with complex effects on vegetation. Indeed, the variable geomorphologic units can generate localized micro-climate increasing the species richness and diversity (Ranwell 1972; Kenworthy 1990) and contributing to the diversity generated by nutrients, pH, earthiness and rate of change (Yu and Rhew 2007). Dune mobility is thus displacing vegetation, while creating local diversity.

From a management perspective, however, dune mobility has been proven to negatively impact economic activities (Maun and Baye 1989, Martinez et al. 1997; Levin et al. 2006), so that the local communities and the authorities have attempted to fix the fluctuations of this frontline using vegetation planting (Menashe 1998; Van Loon-Steensma and Schelfhout 2017; UNCDD 2022) and a combination of imported soil and plants (Hong and Lee 2016). This management relies on a knowledge of: (a) the spatial distribution of vegetation, which is assessed by remote sensing (Laporte-Fauret et al. 2020) to extract notably the NDVI index (Levin et al. 2006); and (b) the dynamic processes from flume and field experiments, which have led to the generation of "factor of safety". It expresses the role of the plant in increasing the stability of a sand dune against wave runup for instance (1 and 2 in Laporte-Fauret et al. 2020):

$$\vartheta = -\frac{31.809}{[-31.809 + (0.034 \times a)]} \quad (1)$$

$$\vartheta = -\frac{48.039}{[-48.039 + (0.027 \times a) + (0.005 \times b) + (8.2 \times c)]} \quad (2)$$

where a is the above ground biomass, b is the number of leaves per square meter and c is the average stem circumference in cm (cf. Feagin et al. (2019) for further details on the data and calculation). In both empirical equations, the emphasis has been given to the aerial part of the vegetation, but using four (4) different species (Sesuvium portulacastrum, Panicum amarum, Ipomoea pescapre and Spatina patens), however Feagin et al. (2019) have also argued that the part below ground have an impact on wave erosion. Extending this idea, buried vegetation and root systems certainly have an effect on the erodibility of the coastal sand dunes, would it be from the action of waves or wind-blows. Consequently, understanding the distribution of buried vegetation and roots in the subsurface of coastal sand dunes is essential to understand the balance at the frontlines between vegetation and sand surface, and for any accurate estimation of sand-dunes biomass.

For this purpose, one method that has proven to be particularly effective in surveying the subsurface of coastal and other sand dunes is Ground Penetrating Radar (GPR). GPR is an electromagnetic method that relies on the change of velocity and reflected energy from linear and punctual objects in the ground, with change in the dielectric permittivity being the main reason for those variations. By recording series of single radar traces, juxtaposed one to another, it is then possible to obtain an "image" of the subsurface. This technology began to emerge in the 1960s and was well explained by the end of the 20th Century (Davis and Annan 1989; Conyers and Goodman 1997; Reynolds 1997). For coastal dunes, originally, digging deep trenches in unconsolidated sediments was lined with difficulties, and GPR started to reveal the complex internal structure of coastal dunes over long-transects in otherwise inaccessible settings (Neal and Roberts 2001), as well as dunes in other environments – e.g. in Antarctica (Bristow et al. 2010a).

Comparing trenches visuals with GPR data, strong reflectors in sand dunes have been linked to humic horizons (Neal and Roberts 2001; Gomez et al. 2010) and other paleosols buried during dune movement (Buynevich et al. 2007). Similarly, reflectors of lower amplitude and inclined at angles related to the friction angle of dry sand (between 30 and 37 degrees) and typical of the sand dune processes, have also been identified in most of the research on sand dune internal structure (e.g. Kain et al. 2014). This relation between GPR reflectors and sedimentary units has then allowed for large-scale GPR surveys in the Namib Desert (Chandlers et al. 2022), providing then extensive images of the subsurface structure.

In combination with dating from Quartz optically stimulated luminescence, the different layers identified with GPR could then be used to define the age of different paleosurfaces in sand dunes from Soulac to Biarritz in the Southwest France for instance (Bertran et al. 2020), showing that aridification periods also coincided with high-wind speeds during the last 25ka to 14ka. Such methods have also been used and extended to other environments to make inferences on extra-terrestrial bodies (e.g. Bristow et al. 2010b).

It then demonstrates the importance of sand dunes to assess aeolian activity (Anthony et al. 2010), how sea-level changes influenced the coasts (Skornik et al. 2008) as well how sand dunes can be used to asses past-meteorological variations (e.g. Sommerville et al. 2007).

Because sand dunes are mostly made of homogeneous well-sorted grain-sizes (from the GPR perspective), they are well indicated for GPR studies (Braton and Montagu 2004) and the search for buried vegetation and roots. When tree roots have a pluricentimetre diameter, it is even possible to measure the diameter from GPR, as it has been shown from 500 MHz – or higher frequency – antenna tests (Barton and Montaagu 2004). In other words, the relative homogeneity of the sand blowouts makes

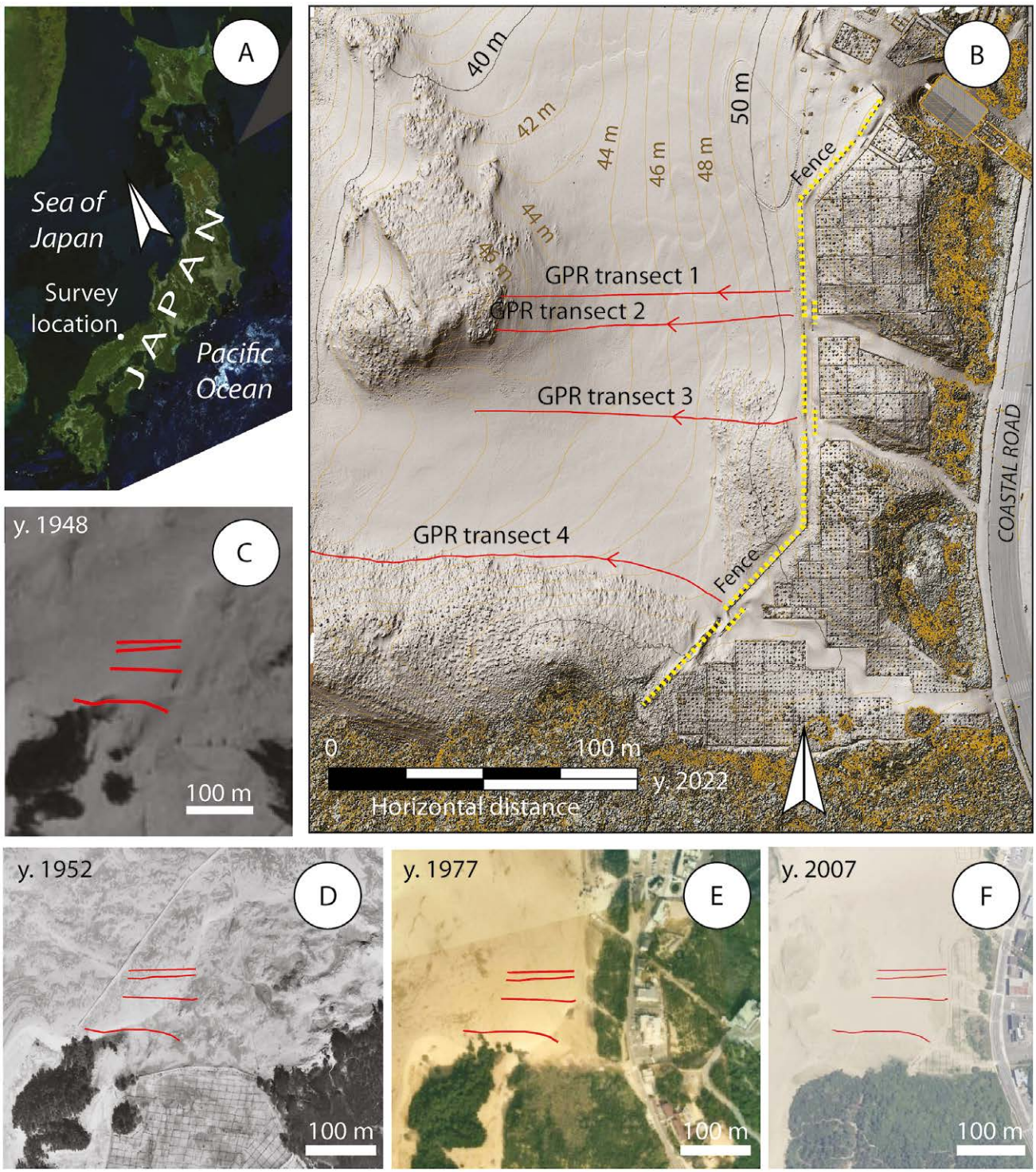


Fig. 1 Study location on (A) the West Coast of Japan, where (B) 4 GPR transects were acquired in an East-West direction. The location during the historical period shows variations in the sand dune land cover, although in (C) 1948, (D) 1952, (E) 1977 and (F) 2007, no major vegetation encroachment is recorded in the area. Only transect 4 has been in vegetated areas (1948 and 1952). The yellow-dotted line marks the fence acting as a sand-barrier.

the identification of punctual objects, such as buried anthropogenic elements and vegetation and roots, easier than in environments that are block-rich and where the material is poorly sorted (e.g. Gomez and Lavigne 2010). And, even in well-sorted environments, the variation in grain-sizes can generate artefact-limits, which can be difficult to identify (Gomez et al. 2008, 2009). Furthermore, to the contrary of sand dunes, moisture-rich silty and clay-rich soils can limit the contrasts between the substratum and the tree roots, leading to further difficulties in interpreting the results (Butnor et al. 2001). Coastal sand dunes are thus an ideal environment to study buried vegetation and roots using GPR, and thus retrieve the presence of otherwise invisible biomass and vegetation.

The present study is thus aiming to image the subsurface in the vicinity of low-grass vegetation in order to assess whether vegetation extends further underneath the sand, remaining invisible to other airborne remote sensing techniques. These techniques are essential, because, by law, one is not authorized to dig a trench or a hole in the Tottori Sand Dune over environmental concerns.

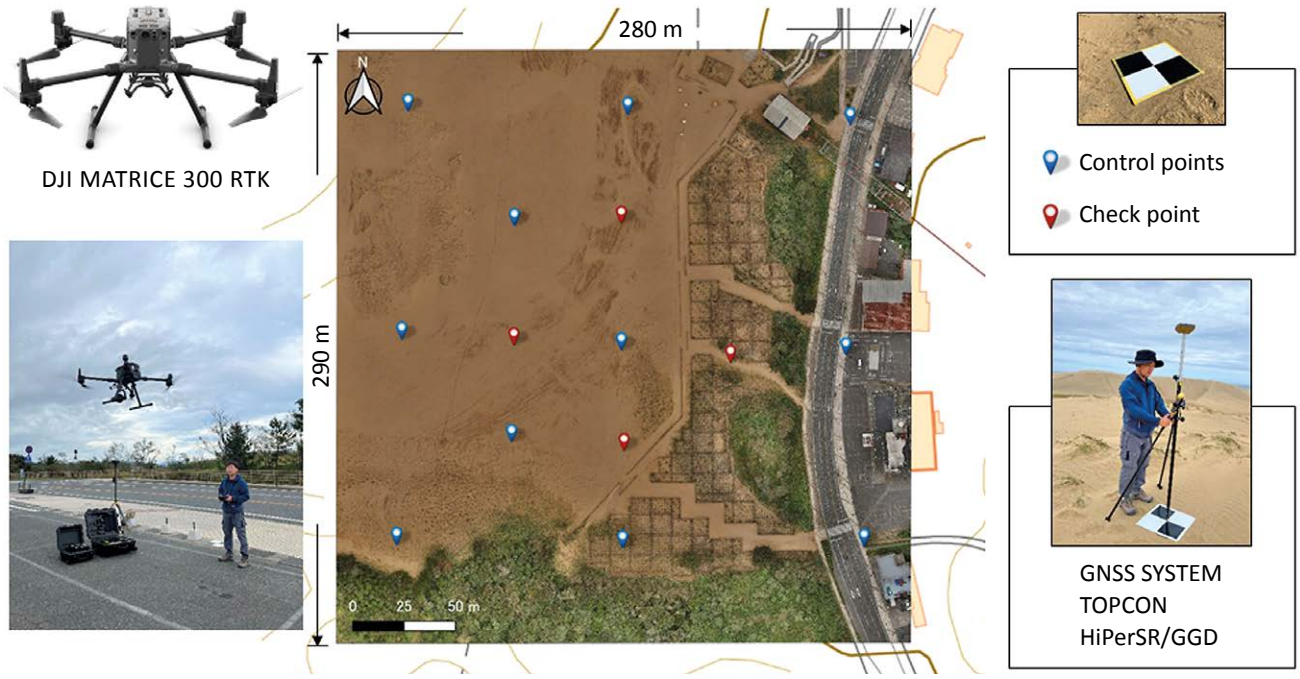
2. Study area

The present research was conducted at the Tottori Sand dune in Tottori Prefecture, Japan (Fig. 1).

The coastal sand dune is located on the West Coast of Japan and fully ex-tend along a 16 km long band, which is about 1 km in width. The sand dune has been dated from tephrochronology to be 50 ka to 70 ka (Naruse 1989), with this sand-dominated feature trapped between the sea and other coastal plain deposits (Murayama et al. 1963; Tamura et al. 2010). Although the Tottori prefecture is not an arid area per se with annual rainfall ranging between 1300 mm/year to 2000 mm/year in the last 20 years (JMA 2023).

In the present survey, we focused on the area named Hamasaka (Tamura et al. 2011), at the landward-edge of the dune, where the dune is active and sand fencing has been installed. Fencing is necessary because the main wind directions oscillate between WNW and NNW at velocities >8 m/s 3.5% of the time, and >12 m/s about 1% of the time (Tamura et al. 2011). Using GPR, Tamura et al. (2010, 2011a, 2011b) have imaged the Tottori sand dune along two parallel transects of 600 m and 1,200 m, which provide an overall overview of the dune structure. But there has been no localized study working on the vegetation, nor has there been detailed imaging of the first 2 m, as previous studies have concentrated on the general structure.

On the sand-dunes of Tottori prefecture, a total of 132 vascular plant species have been found, among which 4 are endemic species, distribution of which is



Total error	Long. error (m)	Lat. error (m)	Alt. error (m)
Control points	0.025	0.025	0.003
Check point	0.025	0.031	0.031

Fig. 2 Topographic data acquisition along a 280 m × 290 m square using the DJI MATRICE 300RTK UAV in combination with Ground Control Points and Check Points recorded by GNSS commercialized by TOPCON.

linked to the dunes' dynamics and the anthropogenic activities (Iwasato and Nagamatsu 2018).

3. Methodology

3.1 Topographic data and orthophotographs

As sand dunes change topography and surface vary rapidly, notably under the influence of wind blowout, the topography and the imagery was acquired within a few hours of the GPR data, to assure that the topography is conform to the GPR transect, and also because high-resolution imagery allows to confirm the length of the transects recorded by the GPR.

The UAV used for the survey is a DJI® Matrice 300 RTK (Fig. 2), which was flown using automatic flight. From the UAV a set of photographs was taken using a Zenmuse P1 camera at an altitude of 30 m, with a flight speed of 3 m/s, and with an overlap ratio of 80% laterally and in the direction of the flight.

In combination with the Ground Control Points recorded using a GNSS (Global Navigation Satellite System) commercialized by TOPCON® (Fig. 2), the photographs were processed in the SfM-MVS (Structure from Motion-Multiple View Stereophotogrammetry) processing software Agisoft® Metashape generating an orthophotograph at a resolution of 0.01 m/pixel (for detailed explanation on the methods, see Gomez 2022), and an altitude error at the check points of 0.031 m, which is below error that would influence the result of the GPR results.

3.2 Ground Penetrating Radar data acquisition and processing

The internal architecture of the sand dune in the survey area was imaged using a Ground Penetrating Radar (GPR) Mala® ProEx mounted with a shielded 800 MHz antenna, from which distance was measured with a coding wheel. Four transects of length 93 m to 178 m were recorded starting landward, and directed towards the sea in an East–West direction (Fig. 1). The accuracy of the GPR signal is related to the vertical resolution (R) of the collected data (3), which has been calculated to be equivalent to the radial resolution, which can be approximated to be 1/4 of the wavelength (Reynolds 1997), and which can be determined as the velocity (V) divided by the nominal frequency of the antenna (f):

$$R = \frac{1}{4}\frac{V}{f} \tag{3}$$

In such a way, the ratio of velocity over frequency for a 800 MHz antenna in free-space (velocity 299,792,458 m/s) is 37.47 cm, and it becomes 18.75 cm (velocity 150,000,000 m/s), so that the maximum accuracy that is attainable with the present dataset is 4.6875 cm. In other words, the 800 MHz antenna will not be able to image single roots or single stems in vegetation, but it will provide an image when there are "clots" of vegetation and roots and soil, or wetter sand that can create a local reflector.

The collected GPR data were then processed following a 7 steps procedure (standard procedure found for instance in (Ettinger et al. 2014; Gomez et al. 2008, 2009; Gomez and Lavigne 2010): (1) the meantime zero correction, so that the 0 corresponds to the topographic surface; (2) DEWOW function application to limit the effects of the surface "ringing effect"; (3) Gain correction compensation with AGC Gain correction, with an empirical value of 1.1 dB/m and a maximum amplitude of 10k dB/m to avoid over-saturation; (4) based on the slopes of the hyperbolae, the average velocity was set to 1.5 m/ns (typical of dry unsaturated and dry sands); (5) Migration of the radargram for velocity, using a unit-velocity field and migration for topography based on the topographic data extracted from the topographic data acquired by UAV.

4. Results

The radargram extracted along transect 1 (Fig. 3) is topographically lower seaward. The internal structure of the radargram is composed of sets of 30 to 35 degrees (Fig. 3-B) with two areas of ~30 m leeward (Fig. 3-C), and ~25 m in the upper area, where a < 1 m thick sets of layers are displaying series of units' sub-parallel to the present topography. There is a sharp contact between these sets of units near the surface and the underlying units at a steeper angle. The two units near the surface (a) and (b) (Fig. 3-B), are the latest deposits, and underneath (b) two generations of accumulations separated by 'reactivation zone' (c and d) serve as a base to the most recent accumulations (b). Underneath the inclined layers, there is a buried unit (f) that resembles the b, c, d units.

Both zooms C and D (Fig. 3) are also displaying a high-concentration of hyperbolae (the sets of black and white arrows). The black-arrows are for hyperbolae are generated by objects located at the surface. In the majority of cases, the hyperbolae are aligned in sets of single layers, instead of being mixed at different depths. The hyperbolae have all very close slope angles in both the rising and falling limbs, across the radargram confirming the similarity of the material.

At both the beginning and the end of the transect, the concentration of hyperbolae is linked to the presence of punctual objects, which are related to the dune vegetation, and invisible from the surface.

The second transect, (Fig. 4) is parallel to transect 1 (Fig. 3) and it presents similarities with the later. The roots or buried aerial part of the vegetation creates a large number of hyperbolae between 72 m and the end of the transects. The hyperbolae are again concentrated in one single layer (Fig. 4-B)

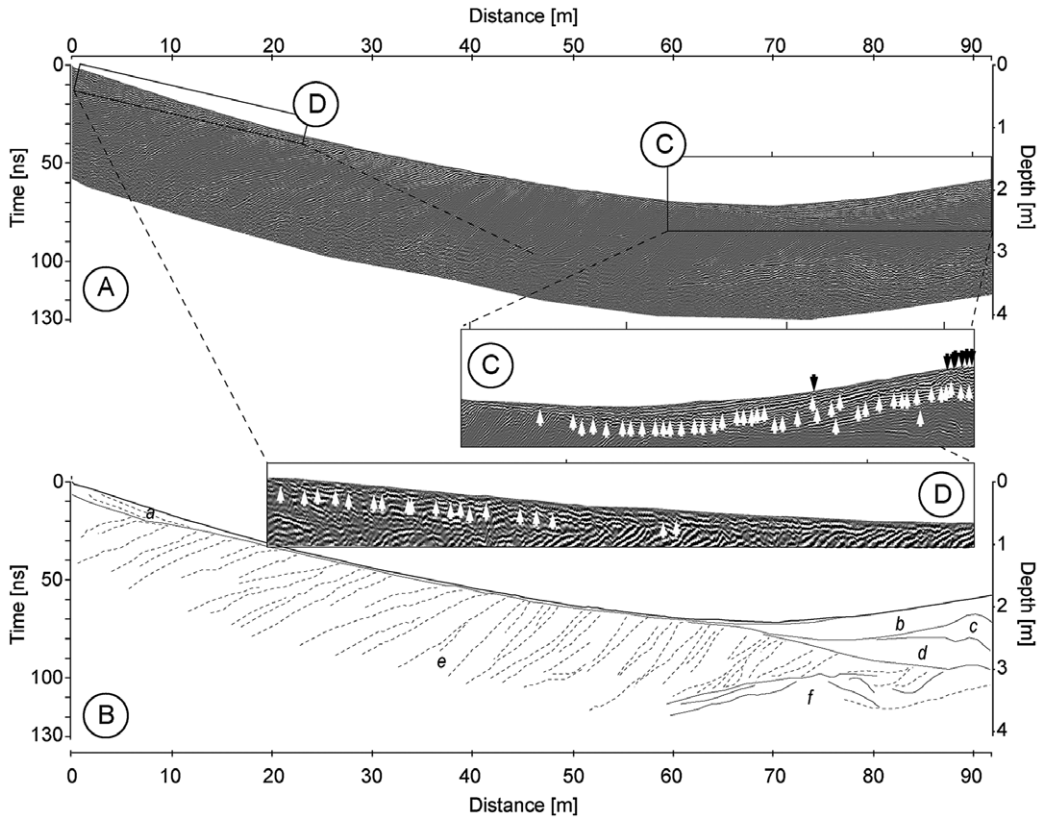


Fig. 3 The radargram acquired at transect 1 and the explanation of both the major units and the distribution of the punctual objects in the subsurface noted as black and white arrows in the zoom; please note that these are not exhaustive lists and other hyperbolae exists, although their interpretation was uncertain (High-resolution files available upon request).

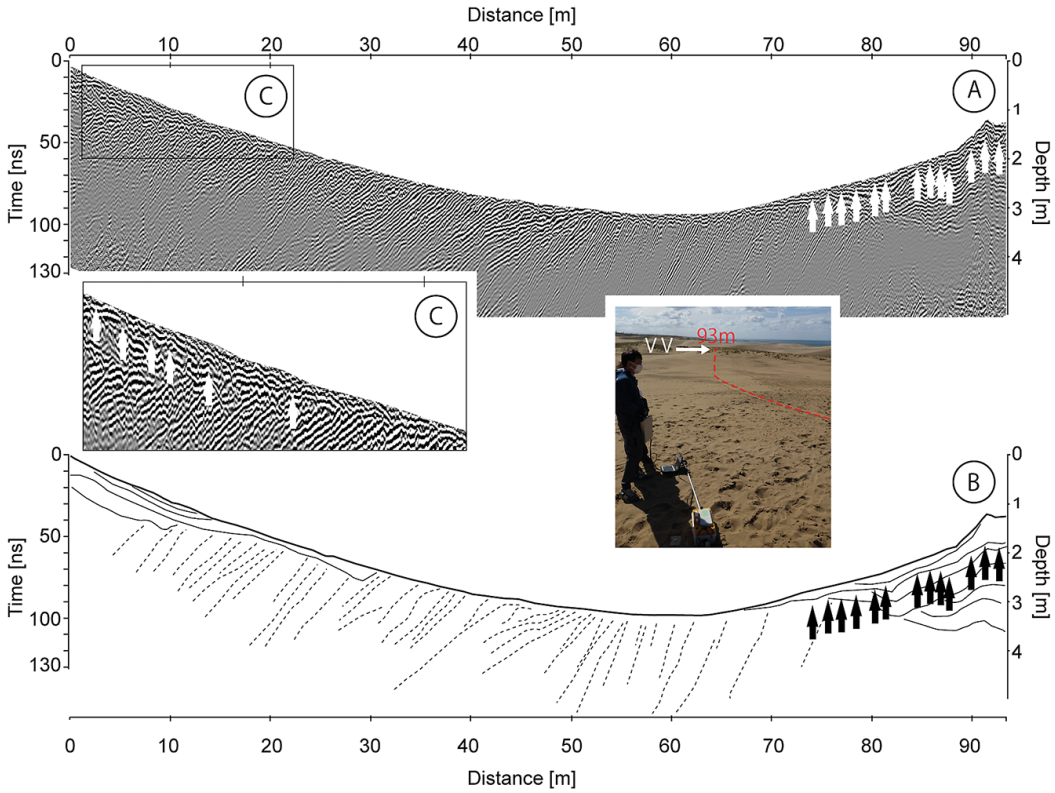


Fig. 4 Transect 2 divided between the radargrams (A), the limits between the different units (B) and a Zoom showing a potential blowout sets of layers (C). A photograph also shows the place where the GPR was dragged (High-resolution files available upon request).

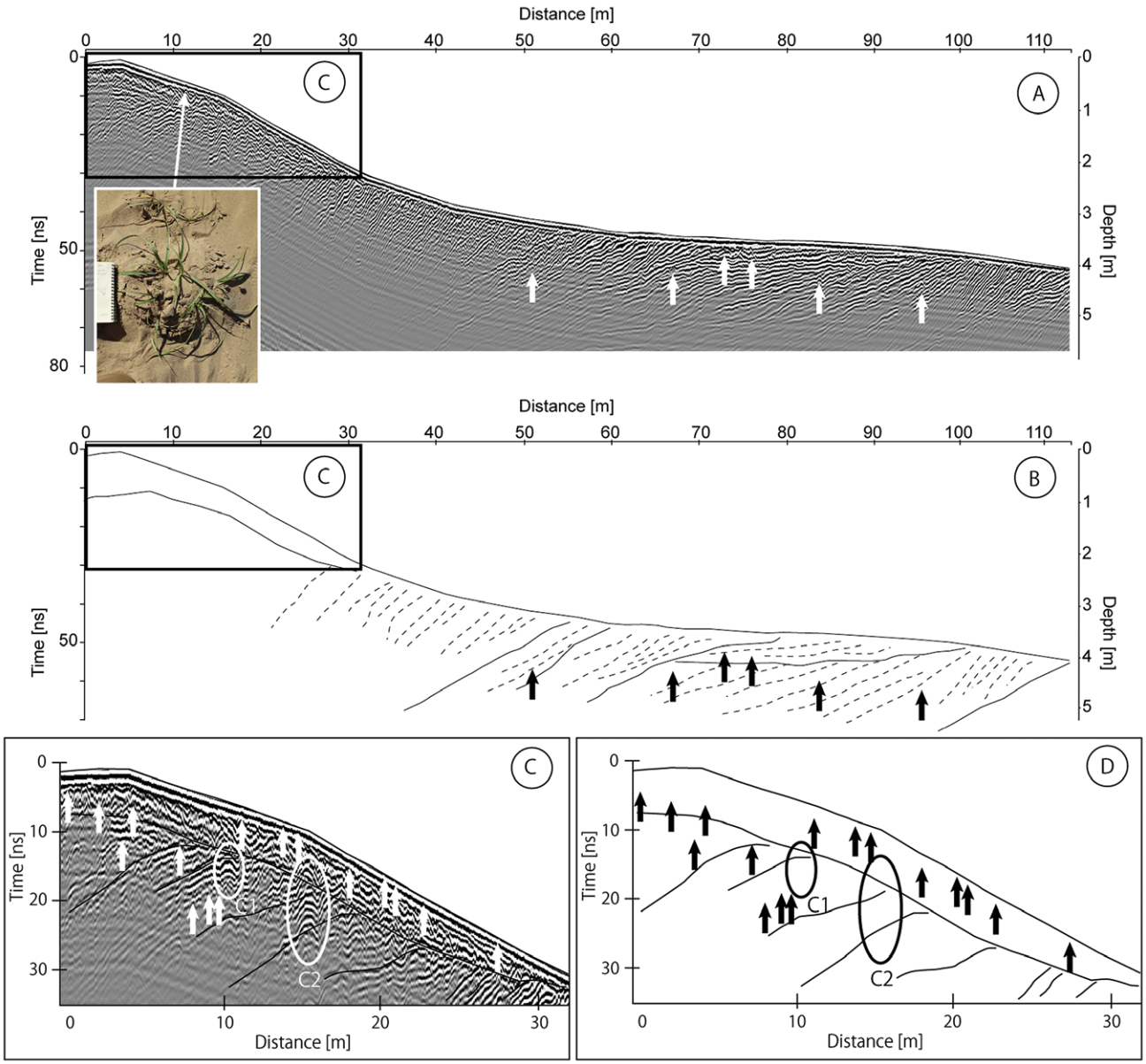


Fig. 5 Transect 3 acquired by GPR (A) and of which the major units have been extracted (B). The first 30 m shows a complex set of layers and hyperbolae (C), where two hyperbolae (C1 and C2) also suggest the presence of foreign material (D) (High-resolution files available upon request).

although vegetation only appears towards the end of the transect near the ridge at 90 m distance (VV on the photograph at the end of the transect on Fig. 4). At the beginning of the transect, there is also, as in transect 1, a series of subhorizontal layers that forms a discontinuity with the slanted layers underneath. With the white arrows, the clearly visible hyperbolae have been identified, although there are others, of which only a small portion is visible (and thus were not marked). At the centre of the transect, between 45 and 50 m, there is one hyperbolae that was not present in transect 1. The signal amplitude at the reflector is stronger than for other hyperbolae and may be due to a foreign object.

The third transect (Fig. 5) is decreasing seaward and not climbing up towards the end as with transects 1 and 2. At this location the first 30 m display a similar pattern as in transects 1 and 2. A sub-horizontal layer is forming a discontinuity with the underlying layers, but this time the layer a slightly thinner (~80 cm) and it is very rich in hyperbolae (Fig. 5-A and B). The hyperbolae are not confined to this layer, but they are also to be found in the slanted layers underneath. (Fig. 5-C). Hyperbolae that have a tip reaching the surface have been identified to be related to vegetation at the surface (cf. photograph in Fig. 5), and among the hyperbolae, there are two hyperbolae with strong reflectors (C1 and C2), with vertical repeats suggesting that material like iron composes the reflectors showing anthropogenic impacts.

The last recorded transect (Transect 4: Fig. 6) is the longest, with 178 m, and it shows further complexity compared to the other transects. If the subsurface is dominated by sub-parallel units at a slanted angle of

30 to 35 degrees (like the other transects), there is a buried dune crest with a change of dipping angles in the layers (Fig. 6-B and C and Zoom F), as well as on the opposite a set of concave units (Fig. 6-G). On Fig. 6, the entire transects presents hyperbolae in the immediate subsurface (comparing with the previous transect on Fig. 5, one can observe that in transect 4 the first few centimetres are discontinuous in-stead of being dominated by a "uniform" surface and few-first decimetres).

Underneath this layer, there is a set of hyperbolae in the first 30 m, corresponding to the palaeo-crest of a dune, as well as along layers with an apparent dip (not in 3D) towards the sea. These series of hyperbolae have been further identified with red arrows (Fig. 6). These groups of hyperbolae are mostly linked to reflectors of higher amplitudes, rather than the layers defined by sets of low amplitude signal. At the end of the transect (from 140 m) an area that is flatter, topographically, is defined by a group of "plate-like" layers, with reflectors hyperbolae (Fig. 6-G).

5. Discussion

The GPR investigation of a 280 m × 290 m area of the sand dune has shown that the subsurface is dominated in its lower part by a slanted series of layers with an angle be-tween 30 and 37 degrees (one will note that they just seem steeper on the figures due to the scale relation between vertical and horizontal values). These layers reach the surface except at the end on

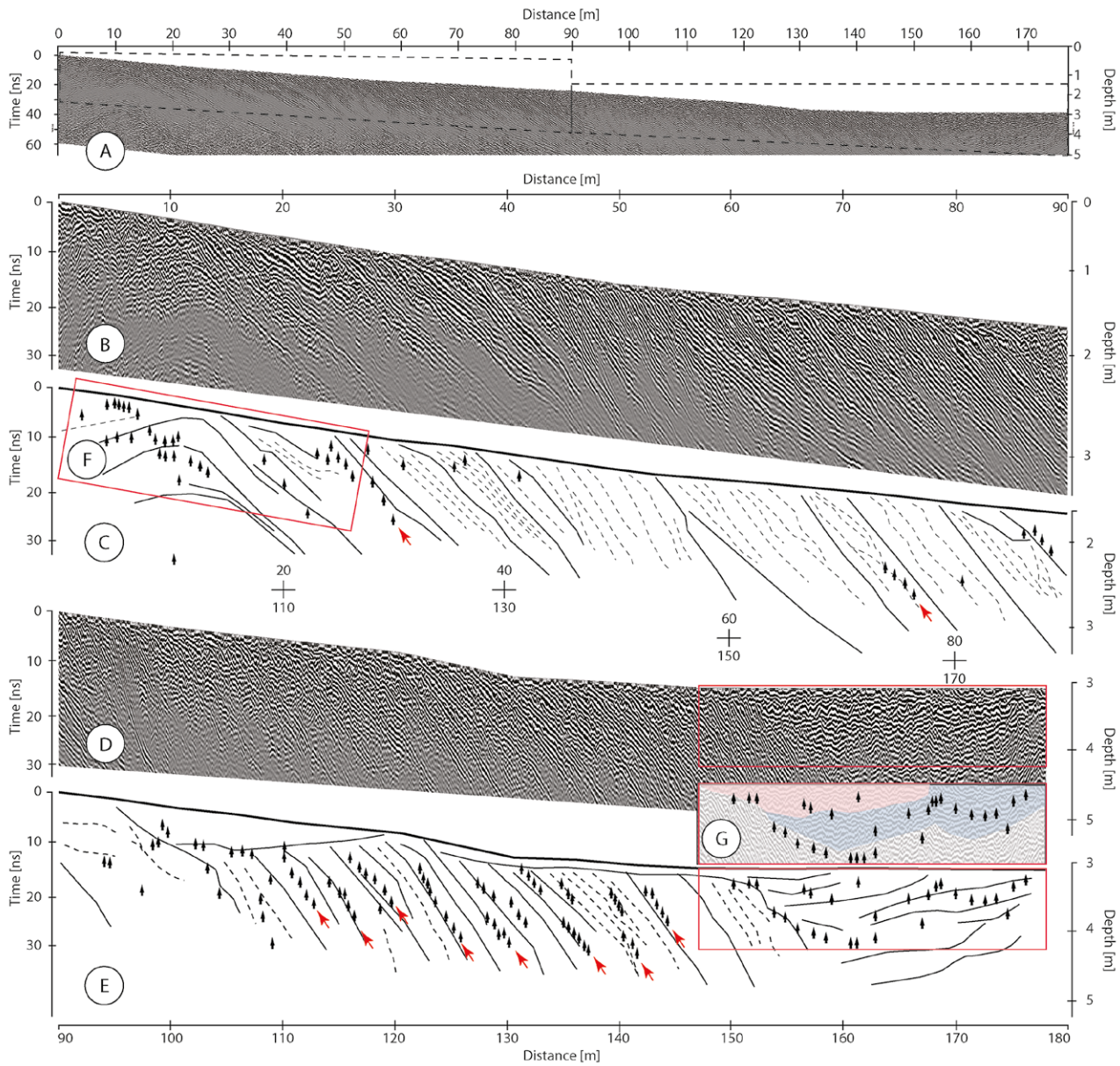


Fig. 6 Transect 4 in its full length (A), and divided between the first 90 m (B and C) and the last 87 m (D and E). The zoom F and G show two opposite layering patterns, with convex and concave shapes. The hyperbolae are shown with black arrows, and the alignments along slanted layers are further demonstrated using red arrows (High-resolution files available upon request).

the topographic highs and the low-topographic (Transect 4). At these locations, other layers are placed over these slanted layers series to form discontinuities. The layering of these top layers is subparallel to the surface, or showing a "plate" shape, as though infilling a topographic hole. As the dune is composed of rather homogeneous sand-grains, punctual objects creating hyperbolae are either due to objects introduced by humans (e.g. in transect 4 the two ringing hyperbolae that could typically be concrete blocks with steel inside), or due to vegetation roots or buried vegetation (at 40–50 m above sea level, it is too high to be driftwood or objects brought by tidal activities). The near-surface hyperbolae and the hyperbolae aligned in layers can thus be confidently attributed to buried vegetation or roots. This interpretation is further consistent with wind blowout locations near ridges or in front of the sand-fences, and for the buried alignments, it shows times of "lower" activity, when the dune is moving more slowly and the vegetation has sufficient time to encroach on the now buried surfaces.

Compared with the existing GPR imaging of the Tottori sand dune (Tamura et al. 2010, 2011a, 2011b), the presence of the slanted layers in units linked to reactivation period is consistent with the finding of the present study, where layers are included in units separated by higher-amplitude data. Furthermore, and although the transects were not taken in the same location as the three transects that are presented in the three papers of Tamura et al. (2010, 2011a, 2011b), most of the units can be attributed to present wind activity, with the layers the most inland corresponding to dates 200 to 300 years old and later. In the lower section of transect 4 however, the buried dune ridge could be linked to the Pleistocene Dune, based on the position proposed by Tamura et al. (2010).

5.1 Sand dune structure

The subsurface units with a slanted angle of 30 to 35 degrees is typical of sand dunes and coastal sand dunes and have been found to form the main

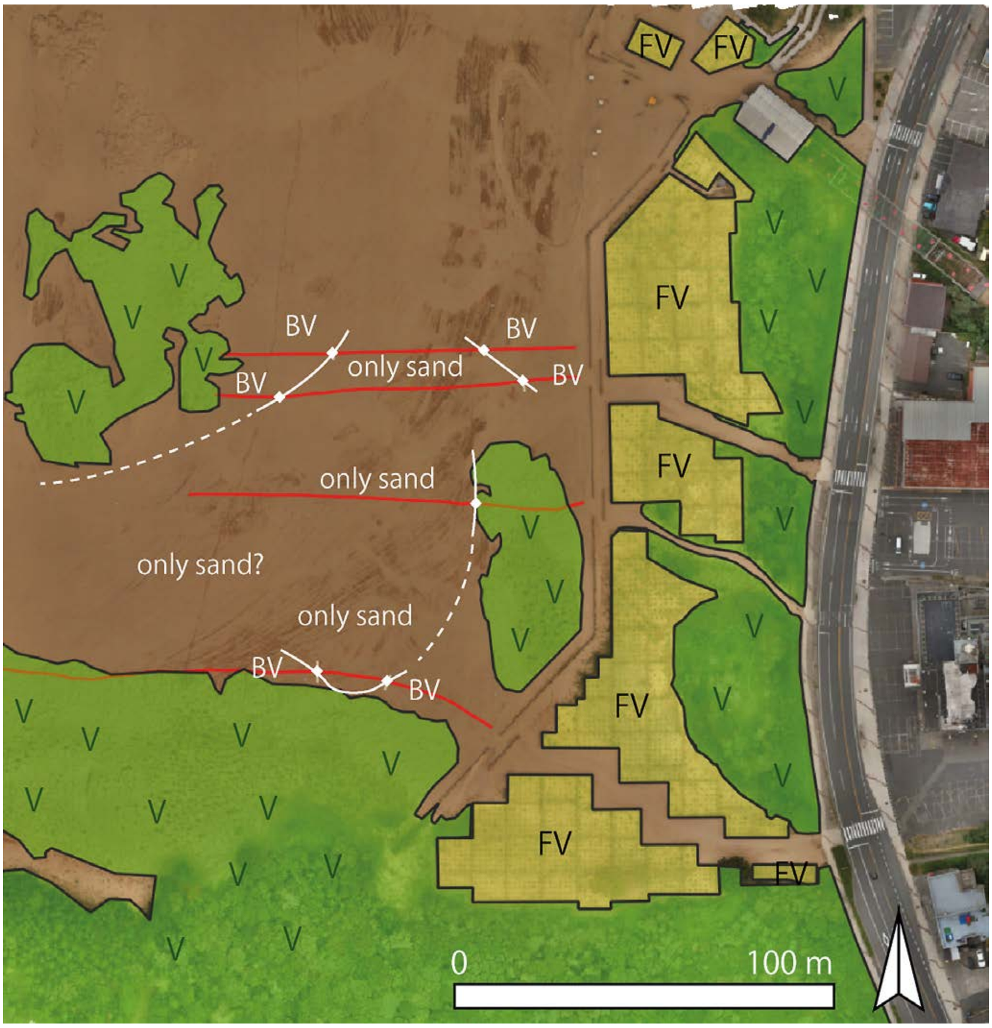


Fig. 7 Distribution of the ground-cover and potential limits of recently buried vegetation (in green with V: vegetation; in yellow with FV: Fenced vegetation; BV: Burried vegetation; the solid white line shows the interpolation of the limit of the near subsurface vegetation; and the dotted line, further interpretation of these limits.

structure of parabolic dunes on the French coast of Acquitaine (Bertran et al. 2020), and also in dunes in other environments, like Antarctica (Bristow et al. 2010b) and the Namib desert (Chandlers et al. 2022). The sand-dunes of the Namib desert, like in Tottori, are also displaying other units towards the top with also truncated units and "superposition surfaces (cf. in Chandlers et al. 2022 Fig. 7 and 8). These types of structures are thus not limited to coastal environment, but typical of wind-deflation and the results of the present study correspond well with the description and interpretation of the radargram from barchane dunes for instance (Fu et al. 2019). These authors have thus related the structure seen on transect 4 in Zoom G to be "trough cross-strata radar facies", with reflections at 28–32 degrees, values that are slightly lower than the 30–35 degrees found in the present study. This discrepancy can be due to the type of sand and its angularity, modifying the repose angle or friction angle of the sand (e.g. Souza Junior et al. 2020), and it can also be associated with the fact that in the present study, we are only examining a system that is inherited and being eroded through with new depositions and movements, meaning that the internal trough cross-strata are not directly related to the present topography.

5.2 Sand dune vegetation and buried vegetation

The interpretation of buried vegetation (whether roots or aerial components of the plant) was found in the upper layers subparallel to the slope, on the stoss side of the dune, which is the location where plants tend to develop first as it is potentially more stable than the steep leeside (facing land in the case of Tottori), as it can be seen from the survey of vegetation establishment in the coastal sand dunes of Israel between 1965 and 1999 (Kutiel et al. 2004). Once the stoss side is more stable, then the lee-side also becomes stable and then vegetation develops.

Buynevich et al. (2017) have used the same GPR and the same antenna in the coastal sand dunes of Lithuania and the Bahamas to image vegetation and roots on dunes. Using the 800 MHz antenna, they have notably imaged windblown buried vegetation as well, proving that the method used in the present survey is replicable, and that this vegetation, once buried indeed generate hyperbolae from punctual objects (e.g. the radargrams of figures 4 and 5 in Buynevich et al. 2017).

5.3 Vegetation estimation and the need for subsurface data

Finally, the findings from GPR show that vegetation and vegetation-roots should ex-tend away from the areas where vegetation is visible at the surface (Fig. 7).

Buynevitch et al. (2017) have studied in areas where vegetation is more mature with tree stands and is visible at the surface, but for lower vegetation (grass, etc.) they may disappear fully under blowout and not be counted during remote-sensing assessment of desertification for instance. The results of this study have relevance to hydrological modelling in semi-arid areas as well, especially pertaining to vegetation estimation which plays an important in flood and water resource management in these scarce areas. This is especially true to rainfall-runoff models such as the Namrom and NamPit (Namibia) which rely on vegetation cover estimates of previous year to determine water flux and storage (Hughes and Metzler 1998).

Although the present research is not sufficient to do so, it appears essential to multi-ply GPR assessment of the dune subsurface, in order to better assess the biomass, present on the sand dune. Especially, when vegetation recovered by blowout is not killed in the process, the later could "re-sprawl". Unexpectedly, such research topic has been largely overseen, because data are mostly generated from remote-sensing, and this should become a new research direction.

6. Conclusion

Combining UAV-based SfM-MVS topography and GPR, the now inactive dune sys-tem (most probably Pleistocene) is being eroded and recovered by wind blowout, which are periodically covered by vegetation. This process seems to have occurred during the palaeodune development, showing periods of stability and periods of "stronger winds". Finally, buried vegetation in the near subsurface can extend 20 to 30 m from an area where vegetation is visible from the surface, and it may be necessary to rethink the limits of vegetation, if one wants to consider more than the mapping of the subaerial organs of plants.

References

Annan, A. P. (2005): Ground penetrating radar in near surface geophysics. In D. K. Butler (Eds): Near-Surface Geophysics. Investigations in Geophysics 13, 357–438, https://doi.org/10.1190/1.9781560801719.ch11.

Anthony, E. J., Mrani-Aloui, M., Hequette, A. (2010): Shoreface sand supply and mid- to late Holocene Aeolian dune formation on the storm-dominated macrotidal coast of the southern North Sea. Marine Geology 276(1–4), 100–104, https://doi.org/10.1016/j.margeo.2010.07.006.

Barton, C. V. M., Montagu, K. D. (20004): Detection of tree roots and determination of root diameters by ground-penetrating radar under optimal conditions. Tree Physiology 24(12), 1323–1331, https://doi.org/10.1093/treephys/24.12.1323.

Bertran, P., Andrieux, E., Bateman, M. D., Fuchs, M., Klinge, M., Marembert, F. (2020): Mapping and chronology of coversands and dunes from the Aquitaine

southwest France. Aeolian Research 47: 100628, https://doi.org/10.1016/j.aeolia.2020.100628.
Bristow, C. S., Jol, H. M., Augustinus, P., Wallis, I. (2010a): Slipfaceless 'whaleback' dunes in a polar desert, Victoria Valley, Antarctica, Insights from ground penetrating radar. Geomorphology 114(3), 361–372, https://doi.org/10.1016/j.geomorph.2009.08.001.
Bristow, C. S., Augustinus, P. C., Wallis, I. C., Jol, H. M., Rhodes, E. J. (2010b): Investigation of the age and migration of reversing dunes in Antarctica using GPR and OSL, with implications for GPR on Mars. Earth and Planetary Sciences Letters 289(1–2), 30–42, https://doi.org/10.1016/j.epsl.2009.10.026.
Butnor, J. R., Doolittle, J. A., Kress, L., Cohen, S., Johnsen, K. H. (2001): Use of ground-penetrating radar to study tree roots in the southeastern United States. Tree Physiology 21(17), 1269–2001, https://doi.org/10.1093/treephys/21.17.1269.
Buynevich, I. V., Bitinas, A., Pupienis, D. (2007): Lithological anomalies in a relict coastal dune: Geophysical and palaeoenvironmental markers. Geophysical Research Letters 34(9): L09707, 1–5, https://doi.org/10.1029/2007GL029767.
Buynevich, I. V., Savarese, M., Allen Curran, H., Bitinas, A., Glumac, B., Pupienis, D., Kopcznski, K., Dobrotin, N., Gnivecki, P., Park Boush, L., Damusyte, A. (2017): Sand incursion into temperate (Lithuania) and tropical (the Bahamas) maritime vegetation: georadar visualization of target-rich Aeolian lithosomes. Estuarine Coastal Shelf Sciences 195, 69–75, https://doi.org/10.1016/j.ecss.2017.02.011.
Chandlers, C. L., Rabedaugh, J., McBride, J. H., Morris, T. H., Narteau, C., Arnold, K., Lorenz, R. D., Barnes, J. W., Hayes, A., Rodriguez, S., Rittenour, T. (2022): Near-surface structure of a large linear dune and an associated crossing dune of the northern Namib sand Sea from Ground Penetrating Radar: Implications for the history of large linear dunes on Earth and Titan. Aeolian Research 57:100813, https://doi.org/10.1016/j.aeolia.2022.100813.
Davis, J. L., Annan, A. P. (1989): Ground-penetrating radar for high-resolution mapping of soil and rock stratigraphy. Geophysical Prospecting 37(5), 531–551, https://doi.org/10.1111/j.1365-2478.1989.tb02221.x.
Conyers, L. B., Goodman, D. (1997): Ground-penetrating radar: an introduction for archaeologists. Walnut, Creek, CA: AltaMira Press.
Ettinger, S., Manville, V., Kruse, S., Paris, R. (2014): GPR-derived architecture of a lahar-generated fan at Cotopaxi volcano, Ecuador. Geomorphology 213, 225–239, https://doi.org/10.1016/j.geomorph.2014.01.013.
Feagin, R. A., Furman, M., Salgado, K., Martinez, M. L., Innocenti, R. A., Eubanks, K., Figlus, J., Huff, T. P., Sigren, J., Silva, R. (2019): The role of beach and sand dune vegetation in mediating wave run up erosion. Estuarine Coast Shelf Sciences 219, 97–106, https://doi.org/10.1016/j.ecss.2019.01.018.
Fu, T., Wu, Y., Tan, L., Li, D., Wen, Y. (2019): Imaging the structure and reconstructing the development of a barchan dune using ground-penetrating radar. Geomorphology 341, 192–202, https://doi.org/10.1016/j.geomorph.2019.05.014.
Gomez, C., Lavigne, F., Lespinasse, N., Hadmoko, D. S., Wassmer, P. (2008): Longitudinal structure of pyroclastic-flow deposits, re-vealed by GPR survey at Merapi Volcano, Java, Indonesia. Journal of Volcanology Geothermal Research 176(4), 439–447, https://doi.org/10.1016/j.jvolgeores.2008.04.012.
Gomez, C., Lavigne, F., Hadmoko, D.S., Lespinasse, N., Wassmer, P. (2009): Block-and-ash flow deposition: A conceptual model from a GPR survey on pyroclastic-flow deposits at Merapi Volcano, Indonesia. Geomorphology 110(3–4), 118–127, https://doi.org/10.1016/j.geomorph.2009.03.024.
Gomez, C., Lavigne, F. (2010): Transverse architecture of lahar terraces, inferred from radargrams: preliminary results from Semeru Volcano, Indonesia. Earth Surface Processes and Landforms 35(9), 1116–1121, https://doi.org/10.1002/esp.2016.
Gomez, C., Lavigne, F., Lespinasse, N. (2010): L'apport du Radar Geologique pour l'etude des impacts geomorphologiques du tsuna-mi du 26 Decembre 2004, In F. Lavigne, R. Paris (Eds.): Le Tsunami du 26 Decembre 2004. Publications de la Sorbonne, Paris, France. 127–136, https://doi.org/10.4000/books.psorbonne.3804.
Gomez, C. (2022): Point-cloud Technologies for Geomorphologists: from data acquisition to processing, Springer, Heidelberg, Germany, https://doi.org/10.1007/978-3-031-10975-1.
Hong, S. H., Lee, E. Y. (2016): Restoration of eroded coastal sand dunes using plant and soil-conditioner mixture. International Biodeterioration and Biodegradation 113, 161–168, https://doi.org/10.1016/j.ibiod.2016.04.021.
Hughes, D. A., Metzler, M. (1998): Assessment of three monthly rainfall-runoff models for estimating the water resource yield of semiarid catchments in Namibia, Hydrological Sciences Journal 43(2), 283–297, https://doi.org/10.1080/02626669809492122.
JMA (2023) Annual Rainfall at the coast of Tottori. Data from the Japanese Meteorological Agency (in Japanese), https://www.data.jma.go.jp/obd/stats/etrn/view/annually_a.php?prec_no=69&block_no=1519&year=&month=&day=&view=p5 (accessed on 9 December 2023).
Kain, C., Gomez, C., Wassmer, P., Lavigne, F., Hart, D. (2014): Truncated dunes as evidence of the 2004 tsunami in North Sumatra and environmental recovery post-tsunami. New-Zealand Geographer 70(3), 165–178, https://doi.org/10.1111/nzg.12052.
Kassas, M. (1995): Desertification: a general review. Journal of Arid Environment 30(2), 115–128, https://doi.org/10.1016/S0140-1963(05)80063-1.
Kenworthy, J. B. Botanical surveys. (1990): In W. Ritchie; L. Kingham (Eds.): The St Fergus Coastal Environment: The physical and Biological Characteristics. C.E.M.P. Aberdeen University Research and Industrial Services, Aberdeen, United Kingdom.
Kutiel, P., Cohen, O., Shoshany, M., Shub, M. (2004): Vegetation establishment on the southern Israeli coastal sand dunes between the years 1965 and 1999. Landscape and Urban Planning 67(1–4), 141–156, https://doi.org/10.1016/S0169-2046(03)00035-5.
Laporte-Fauret, Q., Lubac, B., Castelle, B., Michalet, R., Marieu, V., Bombrun, L., Launeau, P., Giraud, M., Normandin, C., Rosebery, D. (2020): Classification of Atlantic Coastal Sand Dune Vegetation Using In Situ,

UAV, and Airborne Hyperspectral Data. Remote Sensing 12(14): 2222, https://doi.org/10.3390/rs12142222.

Lavigne, F., Paris, R., Grancher, D., Wassmer, P., Brunstein, D., Vautier, F., Leone, F., Flohic, F., De Coster, B., Gunawan, T., Gomez, C., Setiawan, A., Cahyadi, R., Fachrizal. (2009): Reconstruction of Tsunami Inland Propagation on December 26, 2004 in Banda Aceh, Indonesia, through Field Investigations. Pure and Applied Geophysics 166, 259–281, https://doi.org/10.1007/s00024-008-0431-8.

Lee, M. S., Do, J. O., Park, M. S., Jung, S., Lee, K. H., Bae, K. S., Park, S. J., Kim, S. B. (2006): Dominance of Lysobacter sp. In the rhizo-sphere of two coastal sand dune plant species, Calystegia soldanella and Elymus mollis. Antonie Van Leuwenhoek International Journal 90¸ 19–70, https://doi.org/10.1007/s10482-006-9056-z.

Levin, N., Kidron, G. J., Ben-Dor, E. (2006): The spatial and temporal variability of sand erosion across a stabilizing coastal dune field. Sedimentology 53, 697–715, https://doi.org/10.1111/j.1365-3091.2006.00787.x.

Martinez, M. L., Moreno-Casasola, P., Vazquez, G. (1997): Effects of disturbance by sand movement and inundation by water on tropical dune vegetation dynamics. Canadian Journal of Botany 75(11), 2005–2014, https://doi.org/10.1139/b97-912.

Maun, M. A., Baye, P. R. (1989): The ecology of Ammophila breviligulata Fern on coastal dune ecosystem. CRC Critical Review of Aquatic Sciences 1, 661–681.

Menashe, E. (1998): Vegetation and Erosion: A Literature Review. Greenbelt Consulting. 10p. Retrieved from http://www.greenbeltconsulting.com/assets/pdfs/VegetationAndErosion.pdf (accessed on 27 November 2022).

Murayama, M., Isshiki, N., Sakamoto, T. (1963): Tottori hokubu, Tottori nanbu 1:50,000 Geological map. Chishitsu, 1–8 (in Japanese with English abstract).

Musila, W. M., Kinyamario, J. L., Jungerius, P. D. (2001): Vegetation dynamics of coastal sand dunes near Malindi, Kenya. Afrrican Journal of Ecology 39(2), 170–177, https://doi.org/10.1046/j.1365-2028.2001.00294.x.

Iwasato, M., Nagamatsu, D. (2018): Plant species diversity and habitat conditions in a protected large coastal dune area of Western Japan. Landscape and Ecological Engineering 14, 99–113, https://doi.org/10.1007/s11355-017-0334-x.

Naruse, T. (1989): Coastal Sand Dunes in Japan. Geographical Review of Japan 62A, 129–144 (in Japanese), https://doi.org/10.4157/grj1984a.62.2_129.

Neal, A., Roberts, C.L. (2001): Internal structure of a through blowout, determined from migrated ground-penetrating radar profiles. Sedimentology 48(4), 791–810, https://doi.org/10.1046/j.1365-3091.2001.00382.x.

Opelt, L., Berg, G. (2004): Diversity and antagonistic potential of bacteria associated with bryophytes from nutrient poor habitats of Baltic Sea coast. Applied Environmental Microbiology 70(11), 6569–6579, https://doi.org/10.1128/AEM.70.11.6569-6579.2004.

Ranwell, D. S. (1972): Ecology of Salt Marshes and Sand Dunes. Chapman & Hall, London, United Kingdom.

Reynolds, J. M. (1997): An introduction to Applied and Environmental Geophysics. Wiley, Chichester, United Kingdom.

Skornik, K., Gehrels, W. R., Murray, A. S. (2008): Aeolian sand movement and relative sea-level rise in Ho Bugt, western Denmark, during the 'Little Ice Age'. The Holocene 18(6), 951–965, https://doi.org/10.1177/0959683608091800.

Sommerville, A. A., Hansom, J. D., Housley, R. A., Sanderson, D. C. W. (2007): Optically stimulated luminescence (OSL) dating of coastal Aeolian sand accumulation in Sanday, Orkney Islands, Scotland. Holocene 17(5), 627–237, https://doi.org/10.1177/095968360707898.

Souza Junior, P. L., Santos, Junior, O. F., Fontoura, T. B., Freitas Neto, O. (2020): Drained and undrained behavior of an Aeolian sand from Natal, Brazil. Soil Rocks Technical Note 43(2), 1–8, https://doi.org/10.28927/SR.432263.

Stringer, L. C. (2008): Reviewing the International Year of Deserts and Desertification 2006: What contribution towards combating global desertification and implementing the United Nations Convention to Combat Desertification? Journal of Arid Environment 72(11), 2065–2074, https://doi.org/10.1016/j.jaridenv.2008.06.010.

Tamura, T., Kodama, Y., Saitoh, Y., Watanabe, K., Yamaguchi, N., Matsumoto, D. (2010): Ground-penetrating radar profile of the Tottori coastal dunes. Quaternary Research of Japan 49–6, 357–367 (in Japanese), https://doi.org/10.4116/jaqua.49.357.

Tamura, T., Bateman, M. D., Kodama, Y., Saitoh, Y., Watanabe, K., Yamaguchi., Matsumoto, D. (2011a): Building of shore-oblique transverse dune ridges revealed by ground-penetrating radar and optical dating over the last 500 years on Tottori coast, Japan Sea. Geomorphology 132(3–4), 153–166, https://doi.org/10.1016/j.geomorph.2011.05.005.

Tamura, T., Kodama, Y., Bateman, M.D., Saitoh, Y., Watanabe K., Matsumoto, D., Yamaguchi, N. (2011b): Coastal barrier dune con-struction during sea-level highstands in MIS 3 and 5a on Tottori coast-line, Japan. Palaeogeography Palaeoclimatology Palaeocology 308(1–2), 492–501, https://doi.org/10.1016/j.palaeo.2011.05.054.

UNCCD (United Nations Convention to Combat Desertification). (2022): Great Green Wall Initiative. UNCCD, Bonn, Germany, https://www.unccd.int/our-work/ggwi (accessed on 27 November 2022).

Van Loon-Steensma, J. M., Schelfhout, H. A. (2017): Wide green dikes: a sustainable option with benefits for both nature and land-scape values. Land Use Policy 63, 528–538, https://doi.org/10.1016/j.landusepol.2017.02.002.

Yu, K. B., Rhew, H. (2007): Coastal Dunes on the West Coast of Korea; a Geomorphological Perspective. Seoul National University Press, Seoul, Korea.