The Use of RHDI and SHDI for Measuring Sustainability in EU Regions

Antonín Hořčica[1]

Abstract: This article builds on the previous contributions presented at the INPROFORUM conferences, which focused on the application of functional time series analysis to assess sustainability through selected indicators. The aim of the present study is to compare two human development indicators applicable to the assessment of sustainability at the regional level: the Regional Human Development Index (RHDI), developed for the European Commission, and the Subnational Human Development Index (SHDI), based on the UNDP methodology and compiled by the Global Data Lab. The analysis covers EU regions at the NUTS 0, NUTS 1, and NUTS 2 levels over the period 2000–2020, corresponding to three programming periods of EU cohesion policy. Descriptive statistics and functional time series analysis were used. Although the two indices differ in their construction, they exhibit similar trends. The RHDI provides a more detailed regional perspective, whereas the SHDI allows for broader international comparability. This study is among the first to systematically compare these indicators in the context of European regions and highlights their potential for evaluating the impacts of European policies, as well as their practical application within the framework of cohesion policy.

Keywords: sustainability, human development, RHDI, SHDI, EU regions, functional time series analysis, cohesion policy

JEL Classification: C14, Q01, R11, R58

1 Introduction

The European Union (EU) has adopted the concept of sustainable development as one of the key principles of its long-term strategies. This framework, originally defined in the report Our Common Future (WCED, 1987) and further developed in the 2030 Agenda for Sustainable Development (United Nations, 2015), forms the foundation for the transformative strategic documents of the EU (European Commission, 2001), including the European Green Deal (European Commission, 2019). Thus, sustainability in the EU context is conceived as balanced economic, social, and environmental development. This integrated approach is embedded in cohesion policies, which represent the EU's primary instrument for reducing regional disparities and implementing sustainability objectives in practice.

Effective planning, implementation and evaluation of these policies require reliable and comparable indicators capable of capturing the complex state and dynamics of development at the regional level. Regions, as defined within the standardized Nomenclature of Territorial Units for Statistics (NUTS) classification, represent the crucial level for analyzing the impacts of European interventions. Analysis at the NUTS 0 (countries), NUTS 1, and NUTS 2 (major and minor regions) levels allow not only the identification of specific development challenges but also facilitate evidence-based policy design and more targeted allocation of European funds. Measurement beyond traditional economic metrics such as GDP is essential to understand genuine quality of life and sustainability.

To measure the multifaceted nature of regional development, human development indices have gained prominence. Among the most significant is the Regional Human Development Index (RHDI), developed at the request of the European Commission (Hardeman & Dijkstra, 2014) specifically for monitoring purposes within the EU using harmonized Eurostat data. The second is the Subnational Human Development Index (SHDI), produced by the Global Data Lab based on the globally recognized United Nations Development Program (UNDP) methodology that ensures high international comparability. Although both indices track the same three dimensions, health, education, and standard of living, they differ in their subindicators, data sources, and computational methodologies.

The research gap lies in the absence of a comparative study employing advanced statistical methods capable of capturing and comparing complete developmental trajectories in the time series of both indices. This article aims to fill this gap and extend the author's previous work (Hořčica, 2023, 2024) using functional data analysis (FDA) methods.

[1] University of South Bohemia in České Budějovice, Faculty of Economics, Department of Regional Management, Studentská 13, 370 05 České Budějovice, Czech Republic, ahorcica@ef.jcu.cz

2 Goals

The primary objective of this study is to systematically compare the Regional Human Development Index (RHDI) and the Subnational Human Development Index (SHDI) as tools to assess sustainability at the regional level of the EU (NUTS 0, 1, and 2). The analysis covers the period 2000-2020, encompassing three complete programming periods.

Specific objectives include analyzing and comparing methodological differences in the construction of both indices, including their components and normalization procedures. Additionally, the study aims to compare both indices using descriptive statistics and functional data analysis. Finally, the research seeks to assess the practical suitability and potential of both indices for monitoring and evaluating the impacts of European cohesion policies.

These objectives lead to the following research questions: (1) Are the temporal and spatial trends of the RHDI and SHDI indices consistent across EU regions? (2) What are the specific advantages and limitations of each index for assessing regional sustainability, and which is more appropriate for evaluating EU cohesion policy objectives?

To address these questions, three working hypotheses have been formulated: H1: The RHDI and SHDI indices exhibit a high mutual statistical correlation across EU regions, confirming their general consistency as tools for measuring human development. H2: RHDI better captures specific regional disparities within EU member states than SHDI due to its use of indicators directly linked to European policy objectives and Eurostat data. H3: Although the temporal trajectories of both indices during 2000-2020 show overall convergence tendencies, functional analysis will reveal significant differences in the dynamics and volatility of EU development that traditional methods would not detect.

3 Methods

3.1 Geographical aspects of sustainability and regional division of the EU

For the purposes of regional sustainability analysis in the European Union, a unified statistical classification of territorial units has been developed, known as the Nomenclature of Territorial Units for Statistics (NUTS). This hierarchical structure, established by Eurostat, divides the territory of member states into three main levels (NUTS 1, 2, and 3). It is designed to enable comparability of statistical data across EU countries and over time. In the context of European strategies, cohesion policies most frequently utilizes the NUTS 2 level, which represents the key administrative unit for allocating financial resources from EU funds and for assessing territorial disparities.

Given the different geographic coverage of the indices examined, data harmonization was necessary for this study. While RHDI is based primarily on data at the NUTS 2 level with the possibility of aggregation to higher NUTS 1 and NUTS 0 levels, SHDI operates with a combination of NUTS 1 and NUTS 2 levels depending on the size of individual states and the availability of underlying data. Therefore, only those regions for which values of both indicators were available for the entire 2000-2020 observation period were included in the final analysis. A more detailed analysis of data availability and comparability will be presented in the results section of this article.

3.2 Comparison of the construction of the SHDI and RHDI indices

Both regional indices examined in this study are based on the theoretical foundation of the Human Development Index (HDI), which has been developed by the United Nations Development Program since 1990 (UNDP, 2019). HDI combines three fundamental dimensions, health, education, and standard of living, which are aggregated using the geometric mean after normalization.

The SHDI, managed by the Global Data Lab, represents a direct extension of this global methodology to the subnational level (Smits & Permanyer, 2019). It maintains an identical structure to the HDI, which ensures its primary advantage: global comparability (Global Data Lab, n.d.). For normalization, it uses the fixed goalposts method.

For the health dimension, life expectancy at birth has a minimum value of 20 years and a maximum of 85 years. For the education dimension, expected years of schooling and mean years of schooling have minimum values of 0 and maximum values of 18 and 15 years, respectively. For the standard of living dimension, measured as GNI per capita, the minimum is set at $100 and the maximum at $75,000 in purchasing power parity (PPP) terms.

Table 1 Comparison of the calculation of the SHDI and RHDI Indices

Calculation Ph.	Criterion	SHDI	RHDI
1. Normalization	Establishing min-max values	Fixed, global boundaries (established by UNDP)	Dynamic, relative boundaries (from EU regional data) and five-year projection
	Basic formula	*actual value - min / max - min*	
	Application of basic formula	Natural logarithm is applied to the income dimension	Formula is inverted for indicators with negative trends
2. Averaging	Type of mean	Arithmetic mean	
	Application of mean	Only for education dimension (average of two indicators)	For all three dimensions, always averaging two indicators.
3. Final Calculation	Aggregation method	Geometric mean (of dimension indices): $SHDI\ nebo\ RHDI\ =\ \sqrt[3]{D_{health} \cdot D_{education} \cdot D_{income}}$	

Source: Author's elaboration based on UNDP (2019), Global Data Lab (n.d.), Hardeman & Dijkstra (2014).

In contrast to the global approach of SHDI, the RHDI was developed specifically for the European Union to better reflect its policy priorities and data availability (Hardeman & Dijkstra, 2014). The methodology, further elaborated by the Joint Research Centre (JRC), differs in two key aspects.

First, it employs six sub-indicators: for health, these are life expectancy at birth and infant mortality rate; for education, the NEET rate (Not in Education, Employment, or Training) and the proportion of the population with tertiary education; and for income, net adjusted household disposable income and employment rate. Second, rather than using fixed global values for normalization, it employs a dynamic min-max method where boundaries are determined directly from the European regional dataset, extended with a five-year linear projection. The methodology also distinguishes between indicators with positive and negative trends. An overview of the computational differences is presented in Table 1.

3.3 Data Sources, time frame, and regional classification

Data for this study were obtained from two primary sources: the SHDI database version 7.0, managed by the Global Data Lab, and RHDI, for which calculations utilized data from the Eurostat database. This dataset builds on the author's previous work, in which data gaps were identified and addressed (Hořčica, 2024). The analysis covers the period 2000-2020 and the EU regions at NUTS levels 0, 1, and 2.

A critical methodological step involved the processing and harmonization of both datasets to create a final analytical dataset where each unit has valid values for both indices at the same NUTS level. The process included addressing data gaps, where missing values in the RHDI time series were imputed using multiple imputation methods (Amelia algorithm) following Hořčica (2024). In rare cases where data for a particular unit were completely missing, they were supplemented from alternative official sources (e.g., UNDP, CEIC).

During the harmonization of the SHDI database, when the NUTS 2 level was not used, a targeted regional selection was necessary. Additionally, the historical evolution of the NUTS nomenclature was taken into account. Regions that could not be reliably harmonized across both datasets were excluded from the final analysis.

3.4 Statistical analysis methods and R software

All statistical processing and data visualization were performed in the R programming environment (R Core Team, 2023), using tidyverse packages for data processing (Wickham et al., 2019). The analysis was carried out in two phases.

The first phase comprised descriptive statistical analysis (calculation of means and standard deviations) and correlation analysis (Pearson and Spearman coefficients) for the basic assessment of the properties of the data throughout the programming periods and verification of hypothesis H1 regarding the mutual consistency of the indices.

The core of the analytical approach is functional data analysis (FDA), which, unlike classical methods, treats the entire temporal evolution of a region as a single functional unit, a curve, enabling a more robust comparison of trends and dynamics (Hořčica, 2023). For FDA implementation, the GET package (Global Envelope Tests) in R was employed (Mrkvička, 2020). Two key techniques were used: functional boxplots (fBoxplot()) to visualize central trends, variability, and outlying trajectories, and the global envelope test for formal statistical verification of the hypothesis regarding differences between RHDI and SHDI trajectories (Myllymäki et al., 2017). These methods enabled visual and statistical comparison of the two indices.

4 Research results

4.1 Data availability and harmonization

The first analytical step involved creating a harmonized dataset for the SHDI and RHDI indices, ensuring that each analyzed territorial unit had valid values for both indices at the same NUTS level, primarily NUTS 2.

Table 2 Regional inclusion in analysis according to the SHDI and RHDI data harmonization

NUTS Level	Countries	SHDI and RHDI Harmonization
NUTS 2	Belgium, Bulgaria, Czech Republic, Denmark, Italy, Netherlands, Portugal, Austria, Romania, Greece, Slovakia, Spain, Sweden	Fully harmonized and complete NUTS 2 regions
Incomplete NUTS 2	Finland, France, Croatia, Lithuania	For Croatia, older NUTS 2 nomenclature used; for RHDI, regions FI20 and FRY5 excluded; for Lithuania, only 1 NUTS 2 region used
NUTS 0 = NUTS 2	Estonia, Cyprus, Luxembourg, Latvia, Malta	Countries without regional structure; NUTS 0 data used; SHDI and RHDI data supplemented from external sources
NUTS 2 + NUTS 1	Poland, Hungary	In SHDI one region corresponds to NUTS 1 level
NUTS 1	Germany	NUTS 1 level used for both indices
Excluded	Ireland, Slovenia	SHDI available only at incompatible NUTS 3 level

Source: Author's elaboration

Analysis of the source data revealed specific issues that required adjustments, as summarized in Table 2. Therefore, the resulting analytical dataset contains 206 comparable NUTS 1 and NUTS 2 regions. This comprises 18 NUTS 1 regions (16 German, 1 Hungarian and 1 Polish), 5 national regions that correspond to NUTS 2, and 183 NUTS 2 regions. This structure forms the basis for all subsequent statistical tests and visualizations. The complete dataset containing the RHDI and SHDI data is available upon request.

4.2 Comparison of indices using descriptive statistics

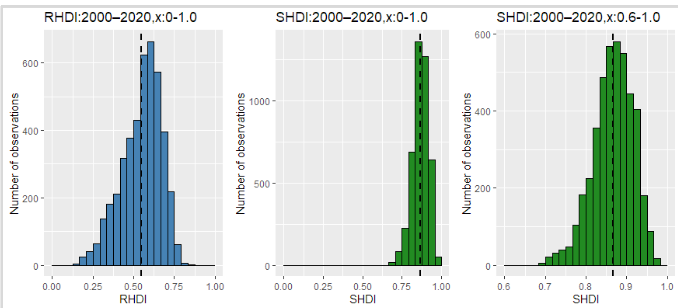
Descriptive analysis was first conducted to assess the fundamental properties of both indices and their mutual relationship. Basic summary statistics are presented in Table 3. The two indices operate on different scales. SHDI achieves a mean value of 0.867 with a low standard deviation of 0.047, indicating that values are clustered in the upper portion of the global spectrum. In contrast, RHDI has a mean value of 0.551 but more than doubles the standard deviation of 0.114, as illustrated in the histograms in Figure 1.

Table 3 Basic statistical characteristics of RHDI and SHDI (2000-2020)

Index	Minimum	Mean	Std. Deviation	Maximum
RHDI	0.215	0.551	0.114	0.778
SHDI	0.721	0.867	0.047	0.951

Source: Author's elaboration

Figure 1 Distribution histograms of the RHDI and SHDI values for the period 2000-2020



Source: Author's elaboration

From a temporal perspective, analysis of data divided according to three seven-year programming periods revealed two key trends. First, the mean values of both indicators systematically increase over time, demonstrating improvement in human development levels across the EU. Second, both indices exhibit a continuous decrease in standard deviation over time (see Table 4). This phenomenon, known as sigma-convergence, confirms that inequalities in human development levels among EU member states are diminishing.

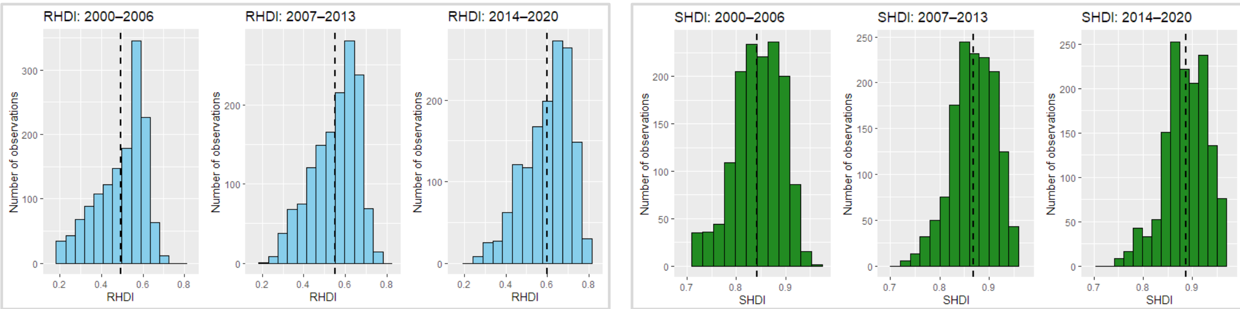
Table 4 Evolution of RHDI and SHDI across programming periods

Index	Period	Mean	Standard Deviation
RHDI	2000–2006	0.486	0.113
	2007–2013	0.553	0.098
	2014–2020	0.615	0.090
SHDI	2000–2006	0.839	0.049
	2007–2013	0.871	0.038
	2014–2020	0.892	0.037

Source: Author's elaboration

Histograms for individual periods (see Figure 2) confirm the first trend, but the convergence is not visually conclusive.

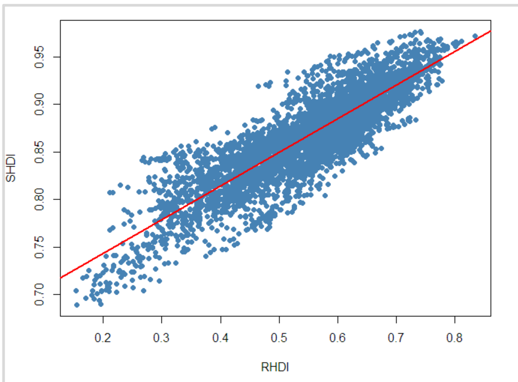
Figure 2 Evolution of RHDI and SHDI distributions across programming periods



Source: Author's elaboration

To examine the relationship between both indices, a linear regression model was constructed with SHDI as the dependent variable and RHDI as the independent variable. The results confirmed a very strong and statistically highly significant positive linear relationship, further supporting the findings of the correlation analysis. The visual representation of this relationship is evident in Figure 3.

Figure 3 Linear regression model of the relationship between RHDI and SHDI



Source: Author's elaboration

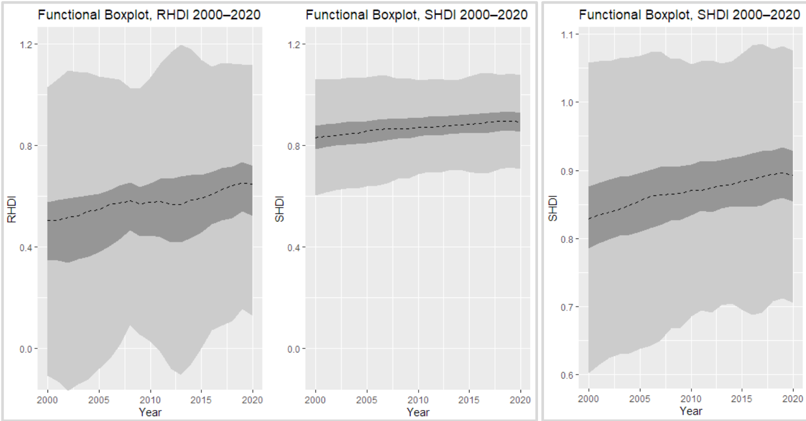
A key finding is the Multiple R-squared value, which reached 0.7554. This indicates that changes in RHDI values can explain approximately 75.5% of the variance in SHDI values, confirming their mutual dependence.

Analysis of the model coefficients yielded the following equation: SHDI = 0.672 + 0.354 * RHDI. The general validity of the model is also highly significant, as confirmed by the F-test value (1.335e+04) and the virtually zero p-value ($p < 2.2e\text{-}16$). These findings quantitatively demonstrate that despite methodological differences, both indices are strongly interconnected and capture very similar underlying dynamics of regional human development.

4.3 Analysis using functional data analysis

Figure 4 presents functional boxplots for RHDI and SHDI at the regional level for the period 2000-2020, displayed without outlier regions for clarity of the main trends. The dashed line indicates the median trend, the dark gray band represents 50% of the most common trajectories (analogous to the interquartile range), and the light gray band shows the total range.

Figure 4 Functional boxplots of RHDI and SHDI at the regional level (2000-2020)



Source: Author's elaboration

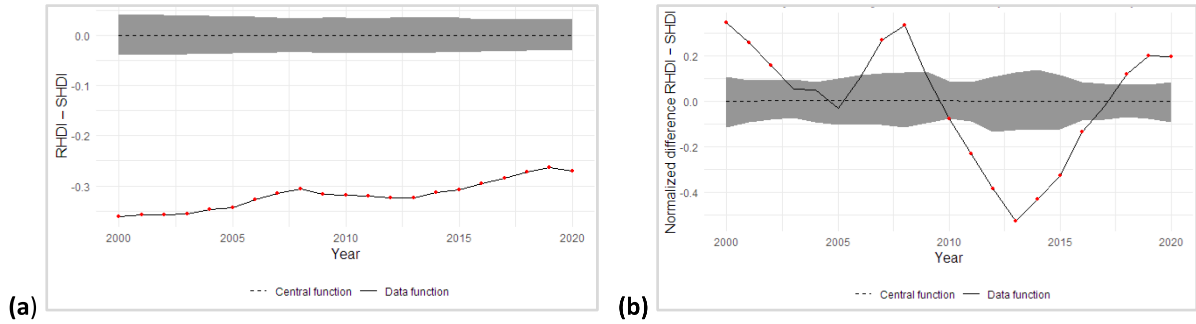
Visual comparison yields consistent conclusions confirming the descriptive statistics: the rising median indicates improvement, and the narrowing bands suggest convergence. The graphs simultaneously reveal different sensitivity and scale: The RHDI covers a broad range and better illustrates internal differences, while the standard SHDI (center) has values compressed in the upper part of the global scale. The added SHDI box

plot with an adjusted scale (right) allows for a better visual comparison of the curve shapes, showing that, despite differing levels, the dynamics and trends of both indices are very similar. However, for a detailed measurement of disparities within the EU, RHDI remains the more suitable tool due to its construction and scale.

To formally statistically verify the differences between the indices, two global envelope tests from the GET package were performed. The first test (Figure 5a) compared the difference in the mean values of RHDI and SHDI, while the second (Figure 5b) compared the similarity of their trajectory shapes after data normalization (removing differences in level and scale).

The first test confirmed that the difference between the mean RHDI and SHDI values is highly statistically significant (p-value = 0.002). As shown in Figure 5a, the actual difference curve lies entirely outside the gray "envelope of randomness." This formally proves that both indices produce demonstrably different results, which is a logical consequence of their different construction and purpose: SHDI for global comparison, RHDI for detailed analysis within the EU. Notable is the slightly convergent trend of the difference curve, attributable to a "ceiling effect" in SHDI for the most developed countries.

Figure 5 Global envelope test of differences (a) and similarity of curve shapes (b)



Source: Author's elaboration

The second test, which focuses on the similarity of curve shapes after normalization, did not reveal a statistically significant difference in overall shape (p-value = 0.14). However, as evident from Figure 5b, although the difference curve on average moves within the "envelope of randomness," it locally deviates outside it during several periods (e.g., around 2004, 2009-2015, and after 2018).

This indicates that while the overall developmental dynamics are statistically indistinguishable, these local deviations suggest that the sensitivity of both indices to specific events or year-on-year changes may differ. Therefore, the indices may respond differently to economic shocks or policy impacts in specific years. This result confirms overall trend agreement (supporting H1) while simultaneously highlighting the differences in dynamics that reflect the different construction and sensitivity of both indicators.

4.4 Evaluation of working hypotheses and answers to research questions

Based on the analysis conducted, all three established hypotheses were confirmed:

Hypothesis H1 was confirmed. The values of the RHDI and SHDI index exhibit very strong statistical correlation, as demonstrated by the analysis of the average trends and the linear regression model (4.2). Additionally, the curve similarity test confirmed that their developmental trajectories are statistically indistinguishable shapes (4.3). This demonstrates that, despite different constructions, both indices measure very similar fundamental concepts and capture the same human development dynamics.

Hypothesis H2 was confirmed. Descriptive statistics (4.2) and visualizations using histograms and functional boxplots (Section 4.3) showed that RHDI has significantly higher value dispersion than SHDI. This confirms that RHDI is a more sensitive tool for capturing and measuring regional disparities within the European Union.

Hypothesis H3 was confirmed. Time series analysis of both indices (4.2) revealed a continuous decrease in standard deviation, providing direct evidence of an ongoing convergence process (sigma-convergence) among EU regions.

These findings directly address the research questions posed in the introduction. The analysis revealed that RHDI and SHDI indices are highly consistent in terms of temporal trends and dynamics, yet differ in their scale, sensitivity, and spatial differentiation. While SHDI, through its global methodology, provides valuable context for international comparison, its ability to distinguish differences among highly developed EU regions is limited. Conversely, RHDI proved to be a more suitable tool for assessing regional sustainability in the European context, primarily due to its higher sensitivity and use of sub-indicators directly aligned with European policy objectives.

5 Conclusion

The objective of this study was to systematically compare the Regional Human Development Index (RHDI) and the Subnational Human Development Index (SHDI) as tools for measuring sustainability at the regional level within the European Union. The analysis, based on descriptive statistics and functional data analysis methods, confirmed that despite methodological differences, both indices capture similar macro-level trends, yet they differ in their sensitivity and intended purpose.

The findings suggest that the two indices are not mutually exclusive, but rather complementary. For policymakers, it is recommended to employ the RHDI as a tool for detailed monitoring and evaluation of the impacts of cohesion policy programs, while the SHDI can serve as a supplementary indicator for global benchmarking of the EU’s performance.

Several avenues for future research are proposed. It would be beneficial to apply comparative functional analysis at a more granular regional level (NUTS 2 and NUTS 3) once data availability permits. Furthermore, investigating the causes of local variations in the dynamics of both indices, identified through the shape similarity test and linking them to specific economic shocks or policy interventions would provide valuable insights. A key step towards a more comprehensive measurement of sustainability would also involve extending the analysis to include the environmental dimension, which is currently absent from both human development indices.

References

European Commission. (2001). *A sustainable Europe for a better world: A European Union strategy for sustainable development.* Publications Office of the European Union. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52001DC0264

European Commission. (2019). *Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee, and the Committee of the Regions: The European Green Deal.* Publications Office of the European Union. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52019DC0640

Global Data Lab. (n.d.). *Subnational Human Development Index (SHDI) Database (Version v7.0*). https://globaldatalab.org/shdi/

Hardeman, S., & Dijkstra, L. (2014). *The EU Regional Human Development Index*. Luxembourg: Publications Office of the European Union. https://publications.jrc.ec.europa.eu/repository/bitstream/JRC90538/online%20version%20a4.pdf

Hořčica, A. (2023). *Functional analysis in R: A new approach to sustainability indicators in the context of regional development*. Inproforum, 17(1), 29-35. https://doi.org/10.32725/978-80-7694-053-6.04

Hořčica, A. (2024). *Multiple imputation in R: A method to bridging data gaps in regional sustainability analysis*. In Proceedings of the 18th International Scientific Conference INPROFORUM: Entrepreneurship for Sustainability (pp. 8–16). https://doi.org/10.32725/978-80-7694-102-1.02

Mrkvička, P. (2020). *GET: Global Envelopes and Tests*. R package version 0.4-7. https://cran.r-project.org/package=GET

Mrkvička, P., Myllymäki, M., Jílek, M., & Hahn, U. (2021). *Multiple Monte Carlo testing, with applications in spatial point processes*. Statistics and Computing, 31(4), 44. https://arxiv.org/abs/1506.01646

Myllymäki, M., Mrkvička, T., Grabarnik, P., Seijo, H., & Hahn, U. (2017). *Global envelope tests for spatial processes*. Journal of the Royal Statistical Society Series B: Statistical Methodology, 79(2), 381–404. https://doi.org/10.1111/rssb.12172

Permanyer, I., & Smits, J. (2020). *Inequality in human development across the globe*. Population and Development Review, 46(3), 583–601. https://doi.org/10.1111/padr.12343

R Core Team. (2023). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.r-project.org/

Smits, J., & Permanyer, I. (2019). *The subnational human development database*. Scientific Data, 6(1). https://doi.org/10.1038/sdata.2019.38

UNDP. United Nations Development Programme. (2019). *Human Development Reports*. Technical Notes [Online]. http://hdr.undp.org/en/content/human-development-index-hdi

United Nations. (2015). *Transforming our World: The 2030 Agenda for Sustainable Development*. New York: United Nations. https://sustainabledevelopment.un.org/post2015/transformingourworld/publication

WCED. (1987). *Report of the World Commission on Environment and Development: Our Common Future*. http://www.un-documents.net/our-common-future.pdf

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., ... Yutani, H. (2019). *Welcome to the tidyverse*. Journal of Open Source Software, 4(43), 1686. https://doi.org/10.21105/joss.01686