Methodology: Building the Geoeconomic Interconnectivity Index

This document describes the methodology underlying the Geoeconomic Interconnectivity Index by Bertelsmann Stiftung, wiiw and ECIPE. For further details see geoii.eu.

The Geoeconomic Interconnectivity Index (GEOII) is based on a rigorous and transparent methodology for measuring economic, financial, and policy interlinkages between 21 EU neighbouring countries and four major global powers — the EU, United States, China, and Russia — from 2010 to 2023. This section outlines how the index was constructed, including data sources, indicator selection, treatment of missing data and outliers, normalisation, and aggregation techniques.

1. Data collection of indicators

The data for each indicator of the Geoeconomic Interconnectivity Index (GEOII) – henceforth referred to as "the index", as opposed to the sub-indices – were collected from the original source, as indicated in the indicator list. The method by which the raw data were subsequently computed into the indicators of bilateral interconnectivity for the dataset, as well as the formula used, can also be found in the table.

The final dataset covers 41 bilateral interconnectivity indicators between 21 EU neighbouring countries and four powers – the EU, the US, China and Russia – from 2010 to 2023. The indicators were organised into three categories (or "sub-indices"), each of which measures a distinct yet complementary dimension of economic interconnectivity (i.e. trade, finance and policy) based on their relevance to economic and policy linkages:

- The trade sub-index captures the intensity, quality and specialisation of bilateral trade flows, including goods, services and sector-specific interactions (e.g. intellectual property rights, or IPRs), which collectively represent the commerce dimension of economic interdependence.
- The finance sub-index measures cross-border bilateral financial flows, investment linkages and external debt relationships, which collectively represent the capital-based dimension of economic interdependence.
- The policy sub-index reflects the bilateral alignment of institutional, regulatory and policy framework that govern bilateral relationships. Policies act as an enabler of (or barrier to) interconnectivity, shaping the conditions under which trade and finance operate.

Indicators for each sub-index were chosen to represent both the input side (i.e. policy) and the output side (i.e. trade and finance) of interconnectivity. This ensures that the index captures both (a) the existence or outcomes of bilateral economic links and (b) their contributions to interconnectivity between countries as set by policies (e.g. policy regulations regarding trade and finance).

Sector-specific indicators, including those related to information and communication technology (ICT) services, green goods and critical raw materials, were included to account for the importance of specialised industries in forging interconnectivity. These industries are central to global economic relations and reflect the broader trends of digitalisation, sustainability and strategic resource dependence.

The EU neighbours include Albania, Algeria, Armenia, Azerbaijan, Belarus, Bosnia and Herzegovina, Egypt, Georgia, Israel, Jordan, Kosovo, Lebanon, Moldova, Montenegro, Morocco, North Macedonia, Palestine, Serbia, Tunisia, Türkiye and Ukraine. The countries are grouped into four regional categories which are "EU Neighbours East", "EU Neighbours South", "Türkiye" and "Western Balkans". Although it is a single country, Türkiye is categorised as a separate region due to its economic importance and size. The main report provides the rationale for the country selection. Libya and Syria, initially part of the selection, were excluded due to data limitations. Note that the EU's geographic neighbourhood includes territories with disputed political status, such as Kosovo and the Palestinian Territories. The authors and their respective institutions do not take a position on the status of these territories. The term "countries" is used throughout the project for the sake of convenience.

The methodological approach taken for constructing the GEOII, including the three sub-indices, aligns with best practices as outlined in the Joint Research Centre (JRC) and Organisation for Economic Co-operation and Development (OECD) Handbook on Composite Indicators (OECD & JRC 2008). The GEOII therefore builds on the established frameworks for composite indicators designed to measure multidimensional phenomena.

2. Correlations of indicators and the handling of missing data

To ensure that the indicators appropriately contribute to each sub-index, the correlations between them were assessed. This check aimed to identify potential redundancies or overlapping information by flagging correlations above 0.95. If any indicator was found to be highly correlated, weights would need to be applied. However, no indicators were found to exceed this correlation threshold.

The next step was to address missing values. This process of handling missing values in the data is essential to ensure that the index remains robust and representative. The choice of imputation methods was guided by the length and nature of the data gaps and was manually checked for each indicator.

For short gaps of one to two consecutive years, linear interpolation was used to fill in the missing values based on the trends observed in the adjacent years. This method is appropriate when the indicator shows a smooth and predictable pattern. For data that changed gradually over time, forward or backward filling was applied using recent data points to fill the gaps.

Special consideration was given to ensure that no data was inferred in situations where external factors (e.g. conflicts) might have significantly disrupted established bilateral relationships. Even if trends appeared stable, imputations were avoided when the underlying dynamics could plausibly have changed due to political or economic shocks. This was done to prevent introducing misleading assumptions into the index.

Note that data coverage for 2023 is lower than in the previous years, primarily due to delays in reporting and the availability of data (or lack thereof) for several indicators. However, the imputation methods described above were carefully applied to maintain the reliability of the index. While this may introduce a small degree of uncertainty, the impact on overall trends is expected to be minimal.

Additionally, alternative approaches were tested for outlier treatment, which could particularly affect indicators with missing values (see below). Ultimately, the most conservative method was chosen to minimise the effect of missing data for 2023. Users should be mindful of this context, however, when examining data for the most recent year.

3. Outlier detection and treatment

The next step is to detect and address outliers, which is crucial for maintaining the accuracy of the index. Unchecked extreme values can distort results, potentially leading to biased conclusions and reducing the reliability of comparisons. By addressing outliers, the index can better capture trends and relationships without being skewed by irregularities.

Outliers were initially identified using a standard deviation threshold of three standard deviations from the mean. Data points exceeding this threshold were flagged as potential outliers and reviewed. This approach was complemented by an analysis of skewness and kurtosis to better understand the distribution of each indicator. Indicators exhibiting high skewness (greater than 2.5) or kurtosis (greater than 5) were subject to closer scrutiny to confirm the presence of outliers. While the standard deviation threshold was the primary detection method, the skewness and kurtosis metrics were also considered to flag potential outliers. When multiple methods flagged outliers, the most conservative approach was applied to ensure robust results.

The method ultimately chosen for addressing outliers was winsorisation, which adjusts extreme values to lie within specific percentiles. Two so-called winsorisation rates were applied, tailored to the degree of skewness and kurtosis observed for each indicator. For highly skewed indicators, values were adjusted to lie within the 1st and

99^{th} percentiles. For moderately skewed data, values were adjusted to the 5^{th} and 95^{th} percentiles.

Formally, the adjusted values X to lie within the specified percentiles can be expressed as:

$$X_{winsorized} = \begin{cases} P_{min} \ if \ X < P_{min} \\ P_{max} \ if \ X > P_{max} \\ X \quad otherwise \end{cases}$$

where P_{min} and P_{max} are the lower and upper percentile thresholds, which vary depending on the skewness and kurtosis of the indicator: the 1^{st} and 99^{th} percentiles for highly skewed indicators, and the 5^{th} and 95^{th} percentiles for moderately skewed indicators.

The rationale for using different winsorisation rates is to balance preserving meaningful variations in the data with minimising the impact of extreme values. A stricter threshold is applied to heavily skewed indicators to prevent large deviations from distorting the overall index scores.

Note that logarithmic transformations are often used to handle outliers, but we chose not to apply them. This approach is most effective for data with large variations or a multiplicative pattern. However, many indicators in our dataset have a value of 0 or contain very small values, which cannot be transformed using a logarithmic scale. Additionally, applying a log transformation could obscure meaningful differences in the data that do not follow an exponential pattern, making interpretation more difficult. For these reasons, we opted for winsorisation, which better preserves the integrity of the data while effectively managing extreme values.

4. Normalisation

To make the indicators comparable across different measurement units and scales, all were normalised to a common scale ranging between 0 and 100 using a so-called min-max normalisation. This means that for each indicator and year, the minimum and maximum values were identified and rescaled, so that eventually the final values range from 0 (least interconnected) to 100 (most interdependent).

More formally, this normalisation for all indicators for each year can be expressed as follows:

$$X' = \frac{X - min(X)}{max(X) - min(X)} \times 100$$

where X' is the normalised value, and *min(X)* and *max(X)* are the minimum and maximum values, respectively, of the indicator for each year. Note that for indicators where higher values indicated poorer performance (e.g. tariffs and other trade restrictions), a simple inversion was applied to ensure that higher scores consistently reflected better performance.

$$X'_{inverted} = 100 - X'$$

5. Index aggregation

The final step is the aggregation procedure. Before aggregating the final index, the scores of the sub-indexes are computed to represent the performance across each bilateral relationship for the three dimensions of interconnectivity. This was achieved by taking the unweighted arithmetic mean of the normalised values for all indicators within each sub-index for a given neighbour-to-power relationship and year. By using an unweighted arithmetic mean, we assume that all indicators within a sub-index contribute equally to measuring its respective dimension of interconnectivity. This approach aligns with those of widely recognised indices (e.g. the Human Development Index, the Environmental Performance Index and the Global Innovation Index), which often employ equal weighting at the sub-index level. In the absence of empirical or theoretical justification for prioritising specific indicators, equal weighting and arithmetic mean ensure neutrality.

The resulting sub-index scores were then standardised to ensure comparability across years and partners. Standardisation involves subtracting the mean from each sub-index score for a given year and then dividing the result by the standard deviation. To enhance interpretability, the standardised scores for each sub-index and year were again rescaled to a normalised scale ranging between 0 and 100, as discussed above. This step ensures that all subindex scores are aligned to the same range, promoting consistency across the three dimensions.

The next step is to select an appropriate method for combining the three sub-indices into a single final composite indicator of interconnectivity. Several approaches are available for this task, including both arithmetic (unweighted) and weighted methods, each with its own set of trade-offs. The chosen method should align with the goal of capturing the strength of the interconnected relationships between the neighbours and the powers, while also supporting meaningful policy discussions.

For this reason, the geometric mean was selected as the primary aggregation method, as it effectively captures the multiplicative effects of interconnectivity. The geometric mean is calculated by multiplying the values and then taking the root of the total number of the data values involved. More formally, the geometric mean can be expressed as follows:

$$Index_{i,k} = \left(\prod_{j=1}^{3} X'_{i,j,k}\right)^{1/3}$$

where $X'_{i,j}$ is the normalised score for subindex j for country i and power k.

The advantage of using the geometric mean is that it places greater emphasis on the multidimensional relationships between each country pair (i.e. neighbour-power pair). Unlike the arithmetic mean, which can bias the final index by a strong performance in a single sub-index, the geometric mean rewards a balanced performance across all sub-indices, making it particularly suited for an index that measures interconnectivity. For instance, a country with consistent scores of 70 in all sub-indices will score higher in the final index than a country scoring 90 in the trade and finance sub-indices but only 30 in the policy sub-index. Conceptually, this approach reflects the idea that true interconnectivity between countries requires consistent engagement across trade, finance and policy, rather than excelling in just one area, and it thereby offers a more holistic view of interdependency. The JRC-OECD Handbook supports the use of the geometric mean for indices that seek to capture multidimensional phenomena, as it prevents a single dimension from dominating the overall score (OECD & JRC 2008).

By penalising unevenness in the scores, the geometric mean emphasises the complementarity between the dimensions. It helps identify areas for improvement and prevents countries with imbalances in their bilateral relationships (e.g. high trade integration but poor policy coordination) being overrepresented in the final index score. This approach ensures that the index accurately captures both the depth and breadth of a country's relationships.

Furthermore, the final GEOII is provided at different levels to provide insights across various dimensions of interconnectivity.

- Country-level aggregation: Scores were calculated for each EU neighbouring country-power pair across all years.
- Group-level aggregation: Scores were calculated for each regional grouping – namely, EU Neighbours East, EU Neighbours South, Türkiye and the Western Balkans – by aggregating bilateral interconnectivity scores across neighbours. Additionally, each country within the region was weighted in the aggregation process based on its population, ensuring that each neighbouring country contributes to the regional scores in proportion to its market size.
- Power-level aggregation: Scores were calculated for each power – the EU, the US, China and Russia – by aggregating bilateral interconnectivity scores across neighbours. This aggregation was also weighted by each neighbour's population, ensuring that the contributions of each neighbouring country to the overall score for each power reflects its relative market sizes.

6. Sensitivity analysis

To assess the GEOII's robustness, we conducted a sensitivity analysis comparing index scores using different outlier detection methods (standard deviation vs. skewness/kurtosis) and aggregation techniques (geometric vs. arithmetic mean). The results show high correlations across all methods, with Pearson coefficients ranging from 0.97 to 0.99 for the overall index and the individual sub-indices. This suggests that the choice of method and technique has minimal impact, reinforcing the GEOII's robustness.

Despite some differences in country and region rankings, the analysis shows consistent results across the different methods. Rankings using geometric mean with the two outlier detection methods are nearly identical, with a Spearman correlation of 0.998. Comparisons between geometric and arithmetic means show slightly lower correlations (0.873 and 0.866, respectively). These findings confirm that while the aggregation method choice introduces some degree of variation, the overall rankings remain highly consistent, ensuring the index's robustness.

Monte Carlo simulations, which are computer-based experiments used to test how results might change when random variations are introduced, show that the rankings for the GEOII remain very stable. This even holds when small random changes are added to the data. This stability holds true no matter which method is used to combine (or "aggregate") the data – whether it is the geometric or the arithmetic mean.

The average rankings produced by these two methods are almost identical, meaning the choice of method does not significantly affect the final ranking outcomes. Additionally, the way rankings shift up or down across the different simulations is very similar for both methods, demonstrating that both approaches are equally reliable when faced with minor inconsistencies in the input data. Finally, the range of rankings (i.e. the highest and lowest positions a country or region might achieve) stays consistent, showing no sign that one method unfairly impacts the relative standings of countries or regions.

Overall, the sensitivity analysis demonstrates that the GEOII is robust to changes in the outlier detection methods and aggregation techniques. The chosen specification, which is the geometric mean with standard deviation outlier detection, provides a balanced and theoretically sound approach to measuring economic interconnectivity.