Tomasz Bartuś

# Geodiversity Tools

## for ArcGIS Pro

Version 0.2.0 – February 02, 2026

AGH University of Krakow

## Manual Revision History

| Version | Date | Name | Description |
|---|---|---|---|
| 0.1.0 | 06/01/2026 | Tomasz Bartuś | Initial Creation |
| 0.2.0 | 02/02/2026 | Tomasz Bartuś | A description of the Steinhaus Vertical Relief Index tool has been added. A number of details describing the operating conditions of the Geodiversity Tools have also been included. |
| | | | |
| | | | |

# Contents

# Section 1 Introduction

**Author information**:
**Author**: Tomasz Bartuś
**Affiliation**: AGH University of Krakow
**Contact**: bartus[at]agh.edu.pl

## 1.1. Purpose of the Geodiversity Tools documentation

The purpose of this documentation is to provide a clear and structured description of the Geodiversity Tools toolbox, including its conceptual assumptions, data requirements, processing logic, and implementation details. The document is intended to support correct application, interpretation, and reproducibility of geodiversity analyses performed using the toolbox. It explains how individual tools operate, how input data are processed, and how output indicators should be understood within the adopted geodiversity model.

## 1.2. Scope of the toolbox

The scope of the Geodiversity Tools toolbox includes the calculation and analysis of quantitative and qualitative geodiversity indicators using spatial data processed in ArcGIS Pro 3.x with the ArcPy geoprocessing framework. The toolbox operates on both vector and raster datasets and uses a regular analytical grid as the common spatial reference.

The tools are organized according to data geometry and processing requirements, distinguishing between raster-based tools and vector-based tools, the latter further divided into point-, line-, and polygon-based analyses. The core part of the documentation provides a standardized description of each tool, including its purpose, methodological background, required input data, generated outputs, graphical representation of the processing concept, computational algorithm, workflow diagram, and interpretation of results in the context of geodiversity assessment. The documentation is complemented by a glossary of terms and a bibliography supporting the theoretical and methodological foundations of the toolbox.

## 1.3. Intended users

The Geodiversity Toolbox for ArcGIS Pro is intended for GIS analysts, Earth science professionals, and nature conservation specialists, as well as academic teachers and students involved in spatial analysis, geodiversity assessment, and environmental research and education.

# Section 2 Conceptual framework

## 2.1. Geodiversity concept

The concept of geodiversity emerged in response to the evolution of the nature conservation paradigm, which over the last century has progressed from the protection of individual landscape elements, through species conservation, to a holistic approach that treats the environment as an interconnected system of biotic and abiotic components[1,2]. In this perspective, abiotic elements such as geological structure, landforms, soil cover, hydrological conditions, and topoclimatic variability are now recognized as the foundation of the functioning of natural systems and a key factor determining patterns of biodiversity[3–7].

Geodiversity refers to the natural variability of abiotic elements of the landscape and the processes responsible for their formation and transformation[8–13]. As a synthetic concept, it describes environmental heterogeneity in a holistic manner; however, in practice, geodiversity studies and assessments are often limited to selected environmental components, such as landform diversity (morphodiversity)[14–16] or soil diversity (pedodiversity). Regardless of whether all abiotic elements or only some of them are subject to analysis, their assessment requires the analysis of measurable environmental properties, referred to as **landscape features** (e.g. lithology, stratigraphy, slope gradient, aspect)[17].

Authors define geodiversity in various ways[8–13]; however, most emphasize the importance of the diversity of abiotic features, their spatial relationships, and the processes shaping the structure and dynamics of the environment. A common element of these definitions is the perception of geodiversity as an essential component of landscape diversity, closely linked to biodiversity[4,6,7,18–21].

Quantitative assessment of geodiversity plays an important role in geoconservation[22–30], environmental analyses[31,32], and decision-making processes[33–35]. The development of GIS methods, morphometrics, landscape ecology, and other disciplines that provide tools for landscape analysis has enabled the development of geodiversity as an independent scientific discipline. Today, it allows for objective, repeatable, and scalable assessments of the diversity of abiotic elements, as well as their integration with other environmental data[10,15,16,28,36–41]. *Geodiversity Tools* are part of the development of the geodiversity paradigm, providing methods and tools useful for the assessment of abiotic components of nature.

The possibilities for the practical application of geodiversity, including the representation of landscape elements in GIS data and their linkage with features, indices, and assessment criteria, are presented in the following subsection.

## 2.2. Components of Geodiversity Represented by GIS Data

### 2.2.1. From geodiversity concept to measurable landscape features

As mentioned above, the assessment of the diversity of abiotic landscape elements requires measurable attributes referred to as landscape features. From a practical perspective, geodiversity analysis requires the identification of these features and their representation in digital form. In subsequent stages of the analysis, they become input data for the calculation of diversity indices. The values of diversity indices describing a selected landscape feature are referred to as assessment criteria.

In the GIS environment, this process can be described as a sequence of logical stages:

**landscape element → landscape feature → index → assessment criterion**.

A **landscape element** is an abiotic component of the environment constituting a distinct spatial component of the landscape (e.g. geological structure); a **landscape feature** describes its measurable property (e.g. stratigraphy); an **index** provides a mathematical or statistical description of the diversity of that feature (e.g. $SHDI$); whereas an **assessment criterion** relates the index to a specific feature and serves as a component of the geodiversity model (e.g. $SHDI_{Strat}$).

### 2.2.2. Landscape elements and their GIS-based attributes

In geodiversity analyses conducted using GIS, the primary objects of interest are abiotic landscape elements. They are described by sets of landscape features that constitute their digital representation.

**Geological structure**

Geological structure is one of the fundamental components of geodiversity. In the GIS environment, it is most commonly represented by landscape features such as lithology or lithofacies, stratigraphy, tectonics, and the presence of geosites. These components are regionalized categorical variables and are most often described using digital data in the vector model, represented by polygons, lines, and points.

**Landforms**

Landforms, due to their continuous spatial nature and strong links with other environmental components, play a key role in quantitative geodiversity models. Consequently, they are often the subject of independent scientific studies (morphodiversity), constituting, for example, a good approximation of total geodiversity in mountainous areas[14–16]. In GIS, landforms are represented mainly by digital elevation models and their derivatives, described using primary and secondary topographic attributes. The most important analyzed landscape features include relief amplitude, slope gradient, aspect, plan and profile curvature, and various morphometric indices describing the variability of landforms. Currently, morphodiversity components are most often represented in the raster model.

**Soil cover**

Soil cover diversity (pedodiversity) is an important component of geodiversity. It reflects the long-term influence of geological, geomorphological, and hydrological processes. In GIS, soils are most commonly described using soil types and classes, usually in the form of vector data (polygons).

**Hydrology**

Hydrological elements, such as springs, streams, rivers, and lakes, are most often represented in GIS in the vector model (point, line, or polygon features). Landscape features include, among others, object types, stream length, hydrographic network density, and their spatial distribution. Hydrological features constitute an important component of geodiversity due to their role in shaping geomorphological and soil-forming processes.

**Climate**

Climate is an abiotic component of nature that, compared to those discussed above, is relatively rarely addressed in geodiversity studies. This is partly due to differences in the definitions of geodiversity and partly to the strong links between climate and other landscape features, as well as difficulties related to its modelling. Most often, studies of climatic diversity focus on small-area climates, i.e. topoclimates[42–46]. In GIS, they can be represented in both raster and vector models (polygons).

### 2.2.3. From landscape features to indicators and assessment criteria

Landscape features themselves do not constitute a measure of geodiversity. Their analysis requires the application of quantitative indices describing the degree of diversity, structure, or spatial relationships of features within defined analytical units, most commonly in the form of a regular analytical grid.

The **indices** used in geodiversity analyses depend on the data model (vector or raster) and on the geometry of the objects representing the analyzed landscape feature[15,16,38,40,47]. For data expressed as regionalized categorical variables, they include, among others, measures of the number of categories and richness (e.g. $Nc$, $Ne$) and entropy indices (e.g. $SHDI$, $Hu$); for datasets describing features expressed as regionalized continuous variables, they include statistical indices (e.g. $SD$, $SDc$). These measures are universal in nature and can be applied to various landscape features.

For practical use in geodiversity models, indices are assigned to specific landscape features, forming **partial assessment criteria**. They explicitly define which index is calculated for which feature, e.g. $SHDI_{Strat}$ – entropy of stratigraphic diversity, $Nc_{Litho}$ – number of lithological categories, or $SDc_{Aspect}$ – circular variability of slope aspect. The criterion name also includes information on the type of input data ($A$ – area, $L$

– line, $P$ – point, $R$ – raster) for which the index is calculated. This allows the distinction between criteria based on area data (e.g. $A\_Nc_{Litho}$) and point data (e.g. $P\_Nc_{Litho}$). Partial criteria constitute the output data of the Geodiversity Tools and the components of quantitative geodiversity models.

#### 2.2.4. Role of GIS-based components in quantitative geodiversity models

The defined geodiversity elements, their features, and the indices and assessment criteria assigned to them form a coherent analytical system enabling the quantitative assessment of the diversity of abiotic environmental elements. This approach allows for objective, repeatable, and scalable geodiversity analyses at different spatial scales and using diverse GIS data.

The presented conceptual structure constitutes the direct conceptual background for quantitative geodiversity models and the operation of the Geodiversity Tools. In the following chapters, the **element–feature–index–criterion** scheme is used to describe computational methods, algorithms, and the interpretation of geodiversity analysis results.

## 2.3. Geodiversity model

Geodiversity models are quantitative–qualitative frameworks used to assess the diversity of abiotic landscape elements. They link partial criteria with landscape features, landscape elements, and total geodiversity. In reference to the definition, geodiversity models are most often defined as sums of the diversity of abiotic landscape elements, with appropriate landscape features and partial criteria assigned to each such component. Such geodiversity models are referred to as Aggregating Ratings (AR)[14–16]. Some authors assign weights to selected landscape features in an expert-based manner. Models of this type can then be referred to as Weighted Linear Combination (WLC).

An important aspect of modelling is the avoidance of data redundancy resulting from the aggregation of partial criteria[16,41]. The selection of applied criteria should therefore be carefully considered. They are chosen so as to best describe different aspects of diversity, e.g. number of categories ($Nc$), number of elements ($Ne$), and entropy ($SHDI$). In practice, this means that, for example, geological structure diversity can be assessed using a combination of different criteria, such as the number of lithological units and categories, the number of stratigraphic units and categories, the number of geosite units and categories, and the length of faults, or by indices measuring entropy. It should be noted that without the use of advanced computational techniques, such as artificial neural networks[16,41], redundancy can be minimized but cannot be completely eliminated.

In the case of raster data, morphodiversity assessment is performed by combining standardized measures of variability of selected terrain features, such as elevation, slope, aspect, or slope curvatures. Such models are referred to as Raster Continuous Morphodiversity (RCM)[15]. Their application allows for an improvement in computational quality while simultaneously reducing costs.

In AR-type geodiversity models, partial criteria are aggregated to obtain a **final assessment**. To ensure equal weighting of criteria and thus enable their comparison, they are standardized. This involves transforming the range of variability of each criterion to a common range, traditionally ⟨0; 1⟩. Geodiversity Tools automatically calculate both the "raw" and standardized values of each criterion. For this purpose, they apply *Min–Max* standardization[48] (1). Its use assumes that the digital datasets describing landscape features are free of outlier values of an artefactual nature. The diversity of landscape elements, as well as total geodiversity, is calculated as the sum of standardized partial criteria.

$$x_i' = \frac{(x_i - x_{min})}{x_{max} - x_{min}} \tag{1}$$

Explanations: $x_i'$—feature value after standardization; $x_i$—feature value before standardization; $i$—$i^{th}$ statistical zone; $x_{min}$, $x_{max}$—minimum and maximum value of the feature set before standardization.

For some partial criteria, particularly those based on counting objects or the length of elements (e.g. $P\_Nc$, $P\_Ne$, $L\_Tl$), the treatment of analytical grid cells lacking input data is of key importance. Values equal to 0 and `Null` values, although they may have different semantic meanings at the stage of "raw" calculations, directly affect the range of criterion variability and the values of $x_{min}$ and $x_{max}$ during the standardization process. Consequently, the choice of how to handle cells without data may modify the distribution of standardized values and the result of the final aggregation. For this reason, **Geodiversity Tools allow the user to explicitly define the behavior of the tools in such situations**, in accordance with the principles described in Section 3.8.1.

The presented aggregation scheme makes it possible to obtain **comparable results** for different types of GIS data, across different partial criteria, and at different spatial scales.

# Section 3 Spatial analysis framework

## 3.1. Units of analysis

Assessments of the natural environment are performed using **analytical grids**, which divide the study area into smaller units within which partial criteria are calculated[49,50]. The units of analytical grids (**statistical zones**, basic fields) are the smallest elements for which partial criteria can be calculated and for which quantitative comparisons of the diversity of landscape features and elements can be made.

Depending on the method of delineation, three types of grids can be distinguished:

1. **Natural** – these include areas delineated on the basis of actual environmental characteristics, such as lithology, morphology, or hydrography. Their delimitation is based on an adopted conceptual model of the landscape. Two main approaches are recognized in the literature: the geocomplex model[51–56] originating from physical geography and the patch–corridor–matrix model[57] derived from landscape ecology. An example of natural grids are geocomplexes forming the basis of physiographic regionalization.
2. **Pseudonatural** – the boundaries of basic fields do not directly correspond to natural environmental components but partially reflect environmental divisions. An example is units based on administrative boundaries. Such analytical grids are used primarily in demographic and economic analyses.
3. **Artificial (regular)** – these are characterized by a regular shape and equal area of statistical zones. They form a regular grid of adjacent geometric figures (e.g. squares or regular hexagons) that cover the entire study area. The choice of cell shape and arrangement affects the compactness of the grid and the spatial relationships between basic fields. Squares arranged in a regular pattern or "brick-like" layouts are most commonly used.

The size of the basic field and its shape have a significant influence on the quality and comparability of the results.

## 3.2. Analytical grid / spatial units

In geodiversity analyses, the analytical grid is a tool that organizes the study space. It provides the possibility of data synthesis within basic fields, enables comparisons of index values within the adopted units, and allows for the aggregation of data derived from analyses of individual partial criteria as well as landscape features and elements.

When using natural or pseudonatural analytical grids in which individual statistical zones differ in their area, the values of geodiversity indices may not be comparable between zones. Therefore, the direct use of automatically standardized results provided by the Geodiversity Tools is not recommended. For such grids, it is recommended to first perform **normalization** by dividing the calculated values of the partial criterion by the area of the given statistical zone. Only these normalized results should then be subjected to standardization (e.g., using the *Min–Max* method). This approach helps avoid interpretative errors resulting from differences in the size of analytical grid cells and ensures the comparability of measures across the entire study area.

As already mentioned, in the case of artificial grids, basic fields may take the form of squares, hexagons, or other shapes, forming a systematic, coherent arrangement covering the entire study area. The most advantageous shapes of analytical grids depend on the spatial compactness of the units and the optimal spatial arrangement of fields relative to one another. Using, as a measure of compactness, the minimum length of unit boundaries, and as a measure of favorable arrangement, the minimum distance between the centroids of two neighboring objects, it can be stated that the most compact analytical grids are formed by structures composed of regular hexagons (with equilateral triangles being the least compact), while the most favorable spatial arrangement is provided by a structure of squares arranged in a "brick-like" pattern (with regular hexagons being the least favorable)[58]. The use of square basic fields arranged in rows and columns is characterized by relatively intermediate values of both analyzed parameters. The main advantages of such analytical grids include ease of construction and the possibility of use and comparison at any level of analysis, whereas their main drawback is most often the random spatial distribution and the crossing of natural boundaries[47].

## 3.3. Influence of spatial resolution and extent

The choice of grid cell size (spatial resolution) and the extent of the analysis is crucial for the results of geodiversity assessment. Excessively large basic fields may lead to a loss of information on local variability, whereas excessively small ones generate data redundancy and may increase the random variability of indices.

The optimal size of analytical grid cells depends on the characteristics of the population of the analyzed variables. Methods used to determine it include, among others, semivariogram analysis[59], analysis of the number of categories within grid cells[60,61], lacunarity analysis[62,63], spectral analysis[64], the paired quadrat variance method[65], and fractal methods[66]. In practice, the choice of resolution may also be arbitrary, based on the researcher's experience, experimental results, or the relevant literature[36].

One of the most commonly used methods for selecting the optimal size of statistical zones is based on the analysis of the ranges of autocorrelation of landscape features represented by regionalized continuous variables[14,67–69]. These ranges indicate the distance at which the variability of the analyzed feature ceases to be autocorrelated and becomes random. Empirical analyses show that the basic fields of analytical grids are homogeneous when their size is 3–5 times smaller than the autocorrelation ranges of these variables[70].

The extent of analytical grids should be representative of a given landscape type. It should include all typical elements of geological structure and landforms. A good practice is to define it based on the boundaries of physiographic units.

The resolution and extent of the grid affect:

- the accuracy of partial criteria estimation,
- the consistency of results at the regional scale,
- the comparability of analyses with other studies.

A well-designed analytical grid forms the foundation of quantitative geodiversity assessment, while simultaneously ensuring computational efficiency and the practical usability of the results.

## 3.4. Data representation

Geodiversity analysis is based on spatial data that enable the representation of landscape elements and their features in digital form. In the GIS environment, these data may be vector, raster, or tabular in nature, and their selection depends on the type of analyzed landscape feature, the spatial scale, and the requirements of the geodiversity model

### 3.4.1. Vector data (points, lines, polygons)

Vector data allow for precise analysis of the shape, distribution, and spatial relationships of landscape elements represented by points, lines, and polygons.

- **Points** – used to represent landscape features that are too small to be depicted as area or linear objects. These may include geological sites, geomorphological features, springs, etc. Points form the basis for deriving partial criteria based on diversity indices of point objects: *P_Nc, P_Ne*, and *Hu*.
- **Lines** – describe linear objects that are too narrow to be represented as area objects, such as faults, thrust boundaries, rivers, or streams. Indices may take into account the length, density, or spatial configuration of lines. Objects with linear geometry constitute input data, for example, for criteria based on diversity indices of linear objects, such as *L_Tl*.
- **Polygons** – are used to represent objects that are too large to be depicted using point or line geometry. They may represent, for example, lithological units, geomorphological forms, soil classes, or water bodies. The diversity of polygon objects is typically described using diversity indices of polygon objects: *A_Nc*, *A_Ne*, and *A_SHDI*.

### 3.4.2. Raster data

Raster data represent landscape features using a grid of regular cells (pixels), in which each cell is assigned an attribute value. In the context of geodiversity analyses, rasters are particularly useful for describing continuous landscape features, such as elevation and derived morphometric features (slope, aspect, curvature, TPI).

Raster data allow for straightforward calculation of partial criteria based on statistical measures of variability, such as *SD*, *SDc*, and *SDcm*. Indices that use selected pixel values, such as the Steinhaus vertical relief index (*M*), generally involve more complex algorithms and longer geoprocessing times. Raster models are well suited for computations over large areas.

### 3.4.3. Attribute and non-spatial tables

In addition to spatial data, geodiversity analysis makes use of attribute tables and tables not directly linked to object locations.

- **Attribute tables** contain information describing spatial objects. They are organized into columns and rows. Rows describe individual objects, and columns represent their attributes. In GIS systems, attribute tables are permanently linked to datasets describing the geometry of vector objects. Selecting an object on the map simultaneously highlights the corresponding row in the attribute table.
- **Non-spatial tables** describe spatial objects but are not directly linked to their geometry. They may contain descriptive attributes, classification values, or calculation results.

## 3.5. Input data and processing environment

### 3.5.1. Input data requirements

Geodiversity Tools require the provision of input data representing landscape features. The data should be stored in a project geodatabase in the form of vector feature classes or raster datasets. Together with the vector, polygon-based analytical grid class, they constitute the input data for the Geodiversity Tools. Warning: the tools do not support shapefile data.

Geodiversity Tools calculate both "raw" and standardized values of partial criteria. These are stored for each basic field of the analytical grid class as attribute values. As a result, the outputs are directly linked to the analytical units, which facilitates further aggregation into complete geodiversity models.

### 3.5.2. Geometry types

Partial criteria are based on specific geodiversity indices that have precisely defined requirements regarding the geometry of input data:

- **Point criteria** require feature classes with point geometry (e.g. geological sites, springs). `Point Z` feature classes may require conversion to `Point` feature classes.
- **Line criteria** are based on input feature classes with line geometry (e.g. faults, thrust boundaries, rivers). For tools that do not support Z geometry, `Line Z` feature classes must be converted to `Line` feature classes.
- **Polygon criteria** require polygon data (e.g. lithological units, soil classes). `Polygon Z` feature classes may require conversion to `Polygon` type.

**The analytical grid class must always have polygon geometry**, as each basic field serves as the reference unit for the calculated indices. The analytical grid feature class requires objects with `Polygon` geometry type.

### 3.5.3. Attribute requirements

Partial criteria that depend on categorical landscape features (e.g. number of categories, entropy) require a defined attribute specifying that category. Category values may be arbitrary – integers or character strings. Attributes based on other data types may not be visible through the tool's GUI.

The analytical grid class must have a **unique attribute identifying each basic field**. Typically, a primary key is used for this purpose, such as `OBJECTID` or `FID`. This enables the unambiguous linkage of index and partial-criterion calculation results to the corresponding analytical units.

### 3.5.4. Coordinate system consistency

All feature classes representing landscape features, raster datasets, and the analytical grid class should use the same **coordinate system**. This ensures correct layer overlay and spatial analysis. Inconsistencies in coordinate systems may lead to errors in data aggregation and in the calculation of geodiversity indices.

## 3.6. GIS environment

Geodiversity Tools require ArcGIS Pro 3.x and Python 3.x (ArcGIS Pro environment).

Source data should be stored in a file geodatabase.

During geoprocessing, intermediate feature classes and non-spatial tables are stored in the file geodatabase that contains the feature class (or raster) of the analyzed landscape feature or, where possible, in the in-memory workspace. These datasets are removed in the final stages of geoprocessing. In exceptional cases, they may remain undeleted and may require manual removal before the next execution of the tool.

## 3.7. Architecture of Geodiversity Tools

### 3.7.1. Tool grouping and naming convention

In Geodiversity Tools, tools are organized according to the type of data sources describing landscape features. At the highest level of the architecture, the toolset is divided into tools that require vector data and those that operate on raster data. This approach reflects the fundamental division of data models used in GIS systems and enables an unambiguous assignment of tools to the type of analyzed spatial information.

The vector data toolset is further divided into tools designed for the analysis of point, line, and polygon geometry data. This division directly results from the nature of the objects representing landscape features and from the different metrics used to describe their diversity.

Tools describing diversity based on category richness are named *Nc* (Number of Categories). Diversity based on the number of objects is calculated using the *Ne* (Number of Elements) index. Among tools employing entropy measures, the *SHDI* (Shannon Diversity Index), based on the proportions of area or length of individual categories, and the *Hu* (Unit entropy) index, based on the proportions of the number of elements belonging to particular categories, are distinguished. For linear objects, for which length is a key property, the *Tl* (Total Length) tool was developed, enabling a quantitative description of the intensity of occurrence of linear elements within the analysis units.

The applied system of tool grouping and naming conventions allows for rapid identification of the type of analyzed diversity, the metric used, and the type of input data already at the tool selection stage.

### 3.7.2. Prefixes (A, P, L, R)

All tool names are additionally prefixed with a symbol indicating the type of input data required by a given metric. The prefix A denotes tools operating on areal vector objects (polygons), P denotes tools using point geometry objects, L denotes tools intended for the analysis of linear objects, and R identifies tools operating on raster data.

The use of prefixes enables an unambiguous distinction between tools with identical metric names but employing different data models (e.g. *A_Nc* vs. *P_Nc*). This solution increases the readability of the toolbox structure, reduces the risk of incorrect tool selection, and facilitates the development of the Geodiversity Tools set through the addition of further metrics and computational variants.

## 3.8. General processing scheme

### 3.8.1. Input data → intermediate data → output data

At a high level of generality, the algorithms of all Geodiversity Tools operate according to a common processing scheme. The user begins the analysis by selecting a specific tool corresponding to a chosen partial geodiversity assessment criterion. The tool selection determines both the type of required input data and the subsequent course of the algorithm.

In the case of tools based on vector data, the user is asked to indicate the dataset representing the analyzed landscape feature. Geodiversity Tools verify the correctness of the input data selection by restricting the list of available layers exclusively to those whose geometry matches the requirements of the given tool (point, line, or polygon). Similarly, tools based on raster data analysis allow the selection of raster layers only.

If the selected tool uses categories of a landscape feature, the user additionally specifies the attribute defining these categories. In this case, the list of available attributes is automatically limited to fields storing text or integer data types, which prevents data interpretation errors.

All geoprocessing tools require the definition of an analytical grid class and an attribute that uniquely and unambiguously identifies each basic field. These units constitute the basis for aggregating the results of geodiversity index calculations.

For some tools ($P\_Nc$, $P\_Ne$, $L\_Tl$), the user additionally defines how to handle situations in which no objects representing the analyzed landscape feature occur within an analytical grid cell. This choice determines the semantic interpretation of empty grid cells—either as units with a real indicator value equal to 0 or as units with an undefined value (`NULL`), excluded from further statistical processing. This decision directly affects both the generation of intermediate data and the standardization procedure, as well as the interpretation of the final results.

During algorithm execution, intermediate data are created in the form of feature classes, rasters, or non-spatial tables. These data are temporarily stored in the project geodatabase or in the computer's working memory and are used exclusively for the purposes of the current processing. Depending on the selected mode for handling empty cells, the intermediate data may contain both zero values and `NULL` values, which affects the range of variability used in subsequent stages of the calculations.

The final outcome of Geodiversity Tools execution is the creation of new fields in the attribute table of the analytical grid class, containing the values of the calculated partial criteria. For each criterion, both the "raw" value and its standardized counterpart are stored. For some tools, the standardization process is performed in accordance with the previously defined method for handling 0 and `NULL` values, ensuring consistency between input data, intermediate data, and final results.

The names of the created fields consist of a prefix identifying the analyzed landscape feature (most often the first three letters of the dataset or category name) and the symbol of the partial criterion. For example, for lithology data, the attribute `Lit_ANc` is created (together with its corresponding alias), whereas its standardized counterpart is stored as `Lit_ANc_MM` with the alias `Std_Lit_A_Nc`.

### 3.8.2. Use of intermediate feature classes and tables

The intermediate data used in the algorithms of Geodiversity Tools include spatial feature classes, raster datasets, and non-spatial attribute tables. Their creation is necessary for operations such as data aggregation, calculation of category proportions, summation of linear feature lengths, or determination of the number of elements within analysis units.

Intermediate classes enable stepwise data processing and increase the readability and modularity of the algorithms. At the same time, their scope and number are limited to the minimum required for correct computation, in order to reduce system resource load and processing complexity.

### 3.8.3. Memory management and cleanup strategy

Memory management and handling of temporary data constitute an integral part of the Geodiversity Tools architecture. Intermediate data created during tool execution are automatically removed after the completion of computations, regardless of whether they are stored in the project geodatabase or in the computer's working memory.

This approach minimizes the risk of cluttering the working environment, reduces disk space usage, and prevents the accidental use of outdated data in subsequent analyses. The strategy also promotes analysis reproducibility and tool stability in a production environment.

## 3.9. Algorithm representation

### 3.9.1. Flow diagrams

Each tool in the Geodiversity Tools toolbox is accompanied by a step-by-step algorithm description and a flow diagram illustrating the main stages of the geoprocessing workflow. These diagrams provide a graphical representation of the tool logic and depict the sequence of operations performed on input, intermediate, and output data, including the geoprocessing tools used and the dependencies between individual analysis stages.

### 3.9.2. Purpose of flow diagrams

Algorithm descriptions and flow diagrams allow users to understand the operational principles of the Python scripts without the need for direct inspection of the source code. They provide valuable support in cases of processing errors, troubleshooting computational issues, or modifying tools to adapt them to specific input datasets.

For GIS developers, flow diagrams facilitate rapid orientation within the algorithm structure, support efficient implementation of fixes, enable functional extensions, and simplify the integration of Geodiversity Tools with other geoprocessing workflows.

### 3.9.3. Graphical conventions (shapes and colors)

The flow diagrams representing the algorithms of the Geodiversity Tools toolbox follow a consistent set of graphical conventions to clearly distinguish between object types and processing stages.

Oval shapes are used to represent input, intermediate, and output data. Data stored in the geodatabase are marked in dark green. Some objects displayed in this color have a rectangular shape, indicating non-spatial intermediate tables. Intermediate data stored exclusively in the computer's working memory are shown in light green. Rectangular objects in yellow represent geoprocessing tools and algorithmic operations. Oval and rectangular grey objects indicate code fragments that are commented out by default. Once uncommented, they may be useful during script debugging. Red quadrilateral objects illustrate conditional instructions.

These graphical conventions are intended to enhance diagram readability and facilitate interpretation of the processing workflow, regardless of the user's level of experience.

### 3.9.4. Relation between flow diagrams and script implementation

Flow diagrams constitute a simplified, conceptual representation of the algorithms implemented in Python scripts. They do not reflect syntactic details or all conditional statements present in the code, but instead highlight the key geoprocessing operations and the flow of data between successive analysis stages.

Each diagram element corresponds to specific sections of the source code, enabling a direct linkage between the graphical representation and the script implementation. This approach enhances algorithm transparency, facilitates verification, and supports further development of the Geodiversity Tools toolbox.

## 3.10. Interpretation and limitations

The results generated by the Geodiversity Tools take the form of quantitative indicators describing the variability of selected landscape features within the statistical zones of the analytical grid. These indicators do not constitute a direct measure of environmental "value"; instead, they provide information on the degree of heterogeneity of a given feature, in relation to its data type and the adopted scale of analysis. Proper interpretation of the results requires consideration of the relevant environmental context (e.g. geological, geomorphological, or hydrographic), as well as limitations arising from data resolution and data quality.

An additional interpretative aspect, important at the stage of result synthesis, is the treatment of missing data (`NULL` values) in partial criteria and their impact on the aggregation of landscape element geodiversity indicators and overall geodiversity.

### 3.10.1. Interpretation of indicator values

Geodiversity indicator values should be interpreted as measures of the relative variability of a given landscape feature within an individual basic unit of the analytical grid.

High indicator values generally correspond to:

- a large number of landscape feature categories (e.g. lithological units, stratigraphic units),
- a high number of spatial elements (e.g. geosites, linear features),
- strong variability of continuous quantitative attributes (e.g. elevation, slope, curvature),
- the absence of a dominant orientation in the case of directional features.

Low indicator values, in turn, indicate relative homogeneity of the analysed feature, expressed as the dominance of a single category, a small number of elements, or low variability of attribute values.

It should be emphasised that most indicators describe the structure of the feature distribution rather than its absolute values. Consequently, two areas may exhibit the same level of geodiversity despite substantial differences in the mean values of the analysed variables. Interpretation should therefore focus on the degree of heterogeneity, rather than on the intrinsic character of the feature itself.

In the context of further result aggregation, particular attention should be paid to the presence of `NULL` values in the analytical grid table. In the ArcGIS Pro environment, arithmetic operations on standardized values (e.g., sum, weighted mean) return a `NULL` result if any of the components has a `NULL` value. This means that a single missing value in one partial criterion can propagate to the synthetic result, leading to a loss of information for the entire analytical grid cell.

The interpretative properties of each indicator are described in detail in the section of this documentation dedicated to the implemented geodiversity metrics.

### 3.10.2. Scale and resolution limitations

The results of geodiversity analyses are strongly dependent on the adopted spatial scale[26,61,68,71], in particular on:

- the size of the basic units of the analytical grid,
- the resolution of the input data,
- the spatial extent of the analysis.

Changes in the size of analysis units may lead to substantial differences in indicator values, even when the source data remain unchanged. This phenomenon is known as the Modifiable Areal Unit Problem (MAUP) and represents one of the fundamental limitations of spatial analyses[72,73].

In the case of raster data, increasing spatial resolution does not necessarily result in changes in indicator values, provided that the spatial structure of the analysed feature remains stable. However, an insufficient number of observations within a basic unit may lead to unstable or incidental results, particularly for statistical indicators.

For these reasons, the choice of analysis scale should be deliberate and justified by the nature of the investigated phenomenon and the objectives of the study.

### 3.10.3. Data-quality-related constraints

The quality of results produced by the Geodiversity Tools is directly dependent on the quality of the input data. The most important constraints affecting result interpretation include:

- geometric and topological errors in vector data,
- heterogeneous resolution or positional accuracy of raster data,
- incomplete or incorrectly defined attributes describing landscape feature categories,
- differences in data acquisition methodologies.

Missing data, manifested as `NULL` values, represent a particular limitation in multi-criteria analyses. Depending on the nature of the criterion, a `NULL` value may indicate an actual absence of the analyzed feature (e.g., absence of linear objects) or a lack of information resulting from incomplete source data. Distinguishing between these situations is crucial for the correct interpretation of results and for subsequent stages of the analysis.

For this reason, selected Geodiversity Tools have been equipped with mechanisms that allow the conversion of `NULL` values into 0 values, interpreted as the actual absence of the landscape feature in the base cell. This solution prevents the undesired propagation of `NULL` values during the aggregation of partial criteria and enables the calculation of synthetic geodiversity indicators without loss of spatial information.

Diversity indicators based on variability measures can be sensitive to outliers or artifacts resulting from interpolation[59], generalization[74], or data noise.

It is therefore strongly recommended to perform a preliminary data quality assessment and a critical evaluation of results in all analyses, especially in comparative studies and decision-support applications.

# Section 4 Installation

This chapter describes the contents of the Geodiversity Tools toolbox and its integration with the ArcGIS Pro environment.

## 4.1. Download Geodiversity Tools

The latest version of the Geodiversity Tools toolbox can be downloaded from the GitHub repository at: https://github.com/tomaszbartus/GeodiversityTools.

It is recommended to download the tool as an official Release package, rather than using the “Download ZIP” option, in order to avoid automatic suffixes being added to the folder name (e.g., “`–master`”).

The downloaded archive should be extracted into any folder, preserving the original name of the main folder: `GeodiversityTools/`, outside (not within) any ArcGIS Pro projects. The installation paths of the tool as well as the paths of ArcGIS Pro projects should not be excessively long. When naming projects, the use of special characters – including spaces and diacritical marks specific to national languages – is also not recommended. Tests have shown that failure to follow the above guidelines may result in errors during script execution.

## 4.2. Struktura toolboxa Geodiversity Tools

The structure of the extracted toolbox is as follows:

```
GeodiversityTools/
│
├── Docs/
│   ├── installation.md
│   ├── usage_examples.md
│   ├── licence.md
│   └── ...
│
├── Symbology/
│   ├── Colors.md
│   ├── GeodiversityTools.stylex
│
├── TestData/
│   ├── TestData.md
│   ├── TestData.ppkx
│
├── Toolbox/
│   ├── Images/
│   │   ├── A_Nc.png
│   │   ├── A_Ne.png
│   │   ├── A_SHDI.png
│   │   └── ...
│   ├── Scripts/
│   │   ├── A_Nc.py
│   │   ├── A_Ne.py
│   │   ├── A_SHDI.py
│   │   └── ...
│   └── GeodiversityTools.atbx
│
└── README.md
```

The main `GeodiversityTools/` directory contains the file `README.md`, which provides a brief description of the toolbox.

The `Docs/` folder contains documentation files, including: `licence.md`, which provides license details; `installation.md`, which contains the key information required to run the tool; `usage_examples.md`, which presents examples of tool usage; and this document, `GeodiversityTools.pdf`.

The `Symbology/` folder contains the style file `GeodiversityTools.stylex`, which defines a proposed symbology for polygon features of the analytical grid class. Five styles are defined and can be used when creating choropleth maps of partial criteria diversity, landscape elements, or total geodiversity. They enable interval-based classification of statistical zones into five classes: very low (I), low (II), medium (III), high (IV), and very high (V) geodiversity. Additionally, the `Colors.md` file presents the symbology colors defined in the RGB color space.

The `TestData/` folder contains data that can be used for testing the Geodiversity Tools. The `TestData.ppkx` file includes a compressed ArcGIS Pro project, while `TestData.md` provides a detailed description of its contents.

The most important part of the toolbox is contained in the `Toolbox/` folder. It includes the toolbox file `GeodiversityTools.atbx` and two subfolders: `Scripts/` and `Images/`. The former contains tool scripts (`.py`), while the latter contains conceptual diagrams used in the tool help pages displayed in ArcGIS Pro.

## 4.3. Integration of the Geodiversity Tools toolbox with the ArcGIS Pro environment

To use the Geodiversity Tools in ArcGIS Pro, start the ArcGIS Pro software and open a selected project.

It is recommended that the `GeodiversityTools/` folder be located **outside** any ArcGIS Pro project directories. Placing the toolbox inside project folders may lead to workspace conflicts, file locking issues, and geoprocessing errors.

In the **Catalog pane**, right-click the **Toolboxes** tab and select **Add Toolbox** from the context menu. In the next step, browse the directory structure of your computer to locate the folder where the tool archive was extracted. Navigate to the tool folder: `...\GeodiversityTools\Toolbox\` and select the toolbox file `GeodiversityTools.atbx`.

The toolbox will now appear in your Toolboxes list and is ready to use.

### 4.3.1. Configure Script Paths (if needed)

The tools may reference external Python script files located in: `...\GeodiversityTools\Scripts\`. If the tool repository is moved to a new location, it is necessary to update the script paths.

To update the path to a .py script file for one of the tools (e.g. `A_SHDI.py`), open the toolbox file `GeodiversityTools.atbx` in the **Catalog pane**. Navigate to the folder containing the required tool (e.g. *A_SHDI*). In this case, it will be: *GeodiversityTools.atbx → Vector Metrics → Polygon Diversity*. Right-click the tool (in this case: *Shannon–Weaver index (A_SHDI)*) and select **Properties** from the context menu. Go to the **Execution** tab and update the path in the *Script File* field. Finally, confirm the changes by clicking *OK*. If necessary, save the changes using *Save*. Repeat this procedure to correct the paths for all tool script files.

### 4.3.2. Updating the Toolbox

To update the tool to a newer version, download the latest version of Geodiversity Tools from GitHub and replace the existing `GeodiversityTools/` folder with the new version, keeping its name and location outside any ArcGIS Pro project directories. Then restart ArcGIS Pro.

# Section 5 Vector Based Metrics

Vector-based partial geodiversity criteria are computed using geometry-specific landscape metrics. Accordingly, the Geodiversity Toolbox provides separate tools for point (Chapter 5.1), line (Chapter 5.2), and polygon (Chapter 5.3) feature classes.

## 5.1. Metrics for Point-Based Datasets

Geodiversity Tools provides three tools designed for landscape features represented by point geometries: Number of Point Elements ($P\_Ne$), Number of Point Categories ($P\_Nc$), and Unit Entropy ($P\_Hu$).

## P_Ne Number of Point Elements

| | |
|---|---|
| Short Description: | The tool calculates the number of elements of a selected landscape feature (point feature class) within each polygon of the analytical grid. |
| Version: | 0.1.1 |
| Author: | Tomasz Bartuś |
| Date: | 2026-01-26 |

### Method Description

Number of Point Elements ($P\_Ne$) is a basic measure of landscape diversity, designed for the analysis and assessment of classes of features described by categorical, regionalized variables and represented by point geometry objects. The metric determines how many individual elements (points) of a given landscape feature occur within the boundaries of each analytical grid cell (2).

$$P\_Ne_k = \sum_{i=1}^{n} e_i \quad (2)$$

Explanations: $P\_Ne_k$—the number of point elements in the $k$-th cell of the analytical grid; $n$—the total number of point elements in the $k$-th cell; $e_i$—a single object with point geometry in the $k$-th cell.

**Notes on missing data**:

In the case of an analytical grid containing base cells without points, the user can choose how to handle missing data during the standardization of $P\_Ne$. Two tool behavior options are available:

- `Replace NULL with 0 (MIN = 0, MAX from Ne)` – missing values (grid cells without points) are replaced with 0, and standardization is performed using a minimum value of `MIN` = 0 and a maximum value `MAX` determined from the observed number of points in the grid cells.
- `Keep NULL (MIN/MAX from observed Ne only)` – missing values remain as `NULL`, and standardization is based on the actual range of observed $Ne$ values, with empty cells not affecting the determination of `MIN` and `MAX`.

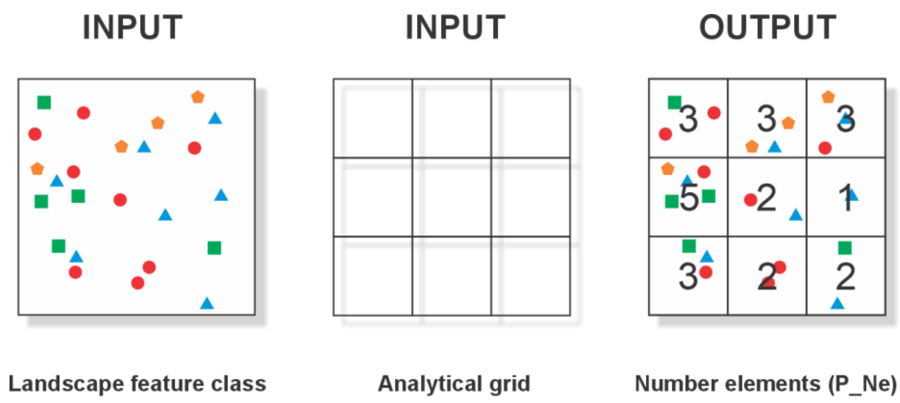
***Input data***

| Parameter | Type | Description |
|---|---|---|
| `landscape_fl` | Point feature layer | A point feature class representing the landscape feature to be analyzed. The number of its elements will be counted within each cell of the analytical grid. |
| `grid_fl` | Polygon feature layer | A polygon feature class representing the analytical grid. For each grid cell (statistical zone), the partial geodiversity criterion is calculated based on the number of elements of the selected landscape feature. |
| `grid_id_field` | Field | An attribute of the analytical grid that uniquely identifies each grid cell (statistical zone) in which geodiversity indices will be calculated. |
| `null_handling_mode` | String / Choice | Specifies how analytical grid cells without geosite points are handled during the standardization of $P\_Ne$. Available options:<br>• `Replace NULL with 0 (MIN = 0, MAX from Ne)` – missing $Ne$ values are replaced with 0;<br>• `Keep NULL (MIN/MAX from observed Ne only)` – missing `Ne` values remain as `NULL` and do not affect the standardization range. |

**Output data**

| Index | Type | Description |
|---|---|---|
| `{prefix}_PNe` | Output field | The number of elements (points) of the analyzed landscape feature within each polygon of the analytical grid. |
| `{prefix}_PNe_MM` | Output field | The standardized number of elements of the analyzed landscape feature within each polygon of the analytical grid. |

## Conceptual Diagram

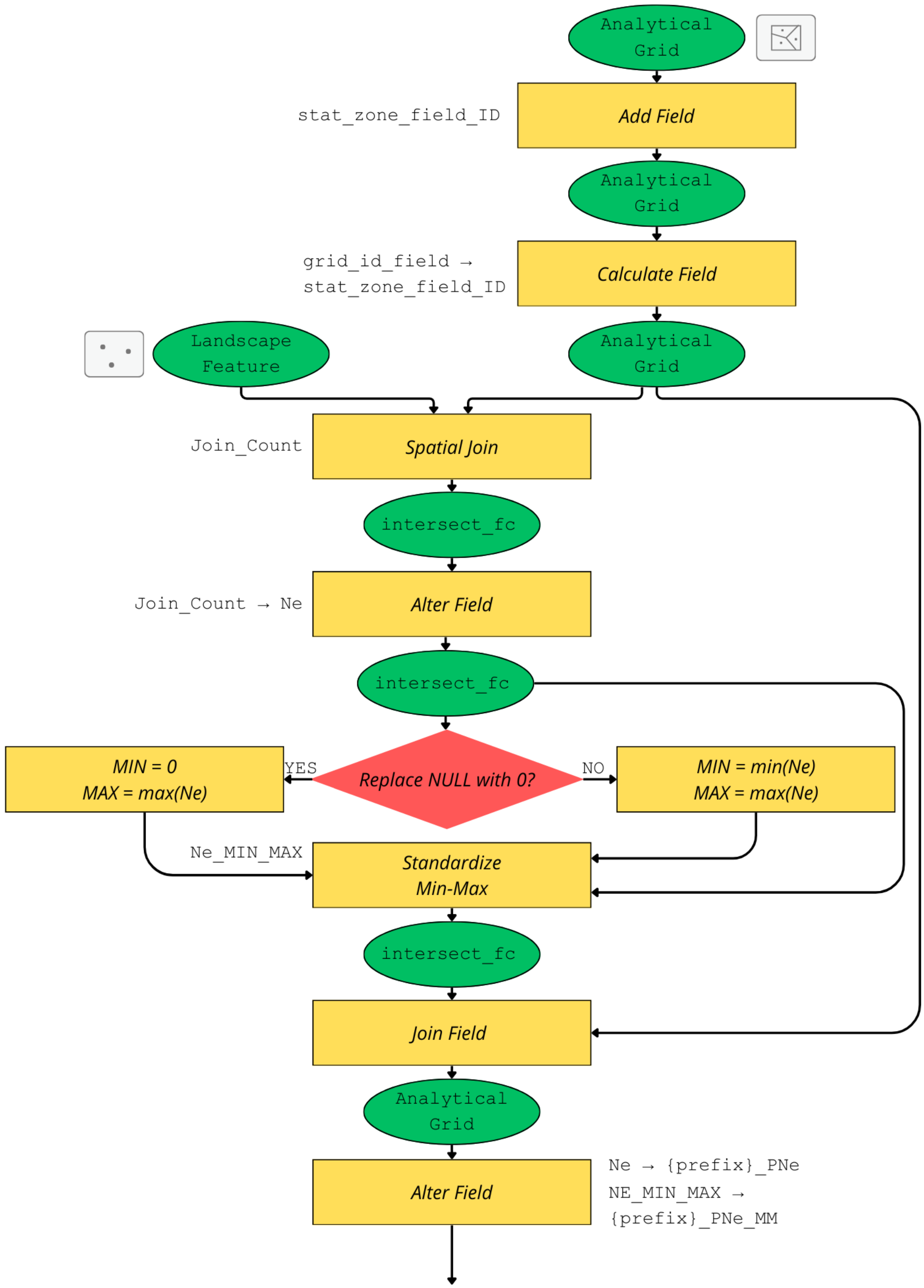


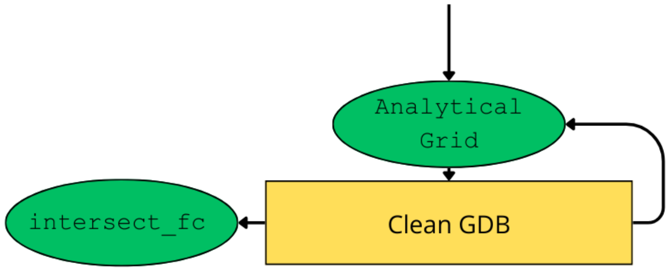

## Algorithm

| | |
|---|---|
| Step 1: | Retrieves input parameters and initializes variables; removes locks from input data; deletes remaining intermediate datasets in the geodatabase (`intersect_fc`); checks whether the analytical grid table already contains fields intended for results. The missing data handling mode (`null_handling_mode`) is specified, determining whether missing values will be replaced with 0 or remain as `NULL`. |
| Step 2: | Creation of an auxiliary statistical zone identifier by copying the primary key of the analytical grid. First, a field named `stat_zone_field_ID` of type `Long` is added. Next, using the *Calculate Field* function, the values of the primary key field (`grid_id_field`) are assigned to `stat_zone_field_ID`. This operation ensures the uniqueness of statistical zone identifiers during table join operations and prevents conflicts with system-defined fields. |
| Step 3: | Perform a *Spatial Join* between the analytical grid feature class (`grid_fl`) and the point feature class representing geosites (`landscape_fl`). The function counts the number of points located within each grid cell. The result is saved as the `intersect_fc` feature class, containing the field `Join_Count`. The tool uses the `JOIN_ONE_TO_ONE` option and the `INTERSECT` spatial relationship. |
| Step 4: | Rename the `Join_Count` field to `Ne` using the *Alter Field* function, where `Ne` represents the number of geosites within the statistical zone. |
| Step 5: | Missing data handling: when the `Keep NULL (MIN/MAX from observed Ne only)` option is selected, 0 values in the Ne field are converted to `NULL` so that grid cells without points remain as missing data. |
| Step 6: | Calculates the minimum and maximum `Ne` values in the intermediate class (`intersect_fc`) taking into account the missing data handling mode. If all `Ne` values are identical (`MIN` = `MAX`), standardization is skipped, and the fields in the created attribute `Ne_MIN_MAX` are assigned a value of 0. Otherwise, classic Min–`Max` standardization is performed, assigning the fields of the `Ne_MIN_MAX` attribute standardized values in the range ⟨0; 1⟩. `NULL` values are preserved in `Ne_MIN_MAX` if the user has selected the option to keep `NULL`s. |

| | |
|---|---|
| Step 7: | Before joining the results to the analytical grid table, the script checks for the existence of the `Ne` and `Ne_MIN_MAX` fields in `grid_fc` and deletes them if they already exist. Then, using the *Join Field* function, the `Ne` and `Ne_MIN_MAX` fields from the `intersect_fc` class are joined to the analytical grid table based on `stat_zone_field_ID`. After the join operation is completed, these values are copied into the analytical grid table. |
| Step 8: | Rename the copied fields to match the tool's naming convention: field `Ne` → `{prefix}_PNe` and field `Ne_MIN_MAX` → `{prefix}_PNe_MM`. Assigning appropriate aliases to the fields. |
| Step 9: | Delete the intermediate feature class `intersect_fc` from the geodatabase, remove the temporary `stat_zone_field_ID` field from the grid feature class table, and optimize the database using the *Compact* function. |

## Flow Diagram






## Relevance to Geodiversity Assessment

- *P_Ne* is a simple diversity metric, describing diversity based on the number of point geometry objects representing a selected landscape feature.
- The metric reflects the density and intensity of occurrence of point-based landscape features within the base cells of the analytical grid.
- *P_Ne* values increase with a greater number of points located within the statistical zone.
- Low *P_Ne* values indicate a small number of point objects, which may suggest landscape homogeneity, lack of diagnostic features, or low intensity of geological or geomorphological processes.
- High *P_Ne* values indicate a large number of points, which may result from highly diverse landforms, high landscape fragmentation, local structural discontinuities, or intensive geodynamic processes.
- Absence of diversity (homogeneity) in the point-based approach means that no points representing the analyzed feature occur within the grid cell.
- Unlike area-based metrics (e.g., *SHDI*), *P_Ne* does not consider the size of the objects, only their number. Like its polygon-based counterpart *A_Ne*, it measures purely the elemental component of diversity.
- Two analytical grid cells may receive identical *P_Ne* values even if the points represent features of different origin, geomorphological significance, or size – the metric is sensitive to object accumulation but does not distinguish their category or spatial significance.
- If no point objects (geosites) occur within an analytical grid cell, the *P_Ne* indicator value is determined according to the user-selected option for handling empty grid cells. Depending on the chosen variant:
    - the empty cell may be assigned *P_Ne* = 0, indicating a complete absence of geosites in the analytical unit,
    - or *P_Ne* remains `NULL`, indicating missing information and excluding the unit from further statistical analysis.
- Standardized values (*Std_P_Ne*) are also determined differently depending on the user's decision.
- If the "`Replace NULL with 0 (MIN = 0, MAX from Ne)`" option is selected, empty cells are treated as units with *P_Ne* = 0 and are included in the standardization process, where:
    - the minimum standardization value is fixed at `MIN` = 0,
    - the maximum value `MAX` is determined from the actually observed *P_Ne* values,
    - empty cells are assigned a standardized value *Std_P_Ne* = 0.
- If the "`Keep NULL (MIN/MAX from observed Ne only)`" option is selected, empty cells remain as `NULL` and do not affect the determination of the `MIN`–`MAX` range used for standardization. *Min–Max* standardization is performed only on grid cells containing at least one point object.

- If all analytical grid cells contain at least one point object ($P\_Ne \geq 1$), *Min–Max* standardization is performed based on the actually observed $P\_Ne$ values, maintaining a consistent interpretation: each cell with at least one geosite receives a standardized value $Std\_P\_Ne > 0$.
- If all analytical grid cells have identical $P\_Ne$ values (i.e., no variation in the number of point objects across the entire area), the standardized partial criterion $Std\_P\_Ne$ is assigned a value of 0 for all cells, regardless of the selected `NULL` handling option, reflecting the lack of spatial contrast for this criterion.

# *P_Nc Number of Point Categories*

| | |
|---|---|
| Short Description: | The tool calculates the number of point categories (e.g., geosites) of a selected landscape feature within each polygon of the analytical grid. |
| Version: | 0.1.1 |
| Author: | Tomasz Bartuś |
| Date: | 2026-01-26 |

## *Method Description*

Number of Point Categories ($P\_Nc$) is a simple measure of landscape diversity, designed for the analysis and assessment of features described by categorical, regionalized variables and represented by point feature classes. The metric determines how many categories (classes) of a given landscape feature occur within the boundaries of each analytical grid cell (3).

$$P\_Nc_k = \sum_{i=1}^{n} c_i \qquad (3)$$

Explanations: $P\_Nc_k$—the number of categories of point elements in the $k$-th cell of the analytical grid; $n$—the total number of point elements in the $k$-th cell; $c_i$—the category of point geometry objects in the $k$-th cell.

**Notes on missing data**:

In the case of analytical grid cells without point elements, the user can choose how to handle missing data during the standardization of P_Nc. Two tool behavior options are available:

- Replace `NULL with 0 (MIN = 0, MAX from Nc)` – missing values are replaced with 0, and standardization is performed using a minimum value of point objects in the statistical zones `MIN` = 0 and a maximum value `MAX` determined from the observed variability of $Nc$.
- `Keep NULL (MIN/MAX from observed Nc only)` – missing values remain as `NULL`, and standardization is based on the actual range of observed $Nc$ values, with empty cells not affecting the determination of `MIN` and `MAX`.

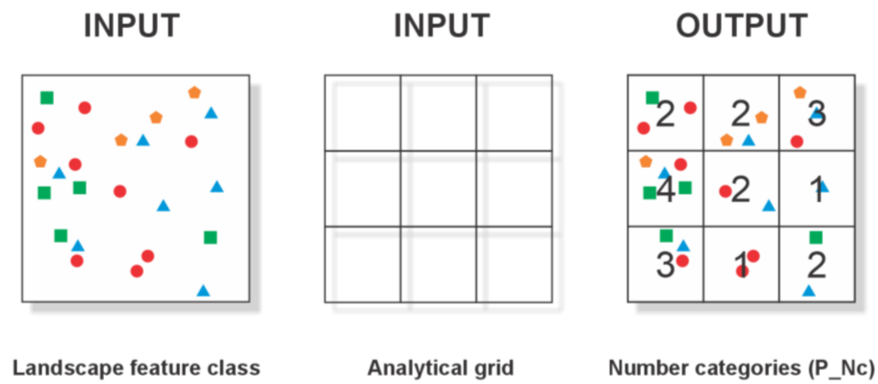
***Input data***

| Parameter | Type | Description |
|---|---|---|
| `landscape_fl` | Point feature layer | A point feature class representing the landscape feature to be analyzed. Its categories will be counted within each cell of the analytical grid. |
| `landscape_attr` | Field | The attribute field in the input landscape feature that defines its categories (classes), which will be counted within each cell of the analytical grid. |
| `grid_fl` | Polygon feature layer | A polygon feature class representing the analytical grid. For each cell of the analytical grid (statistical zone), the partial geodiversity criterion is calculated based on the number of categories of the selected landscape feature. |
| `grid_id_field` | Field | An attribute of the analytical grid that uniquely identifies each cell of the analytical grid (statistical zone) in which geodiversity indices will be calculated. |
| `null_handling_mode` | String / Choice | Specifies how analytical grid cells without point objects are handled during the standardization of $P\_Nc$. Available options:<br>• `Replace NULL with 0 (MIN = 0, MAX from Nc)` – missing $P\_Nc$ values are replaced with 0;<br>• `Keep NULL (MIN/MAX from observed Nc only)` – missing $P\_Nc$ values remain as `NULL` and do not affect the standardization range. |

**Output data**

| Index | Type | Description |
|---|---|---|
| `{prefix}_PNc` | Output field | The number of categories (types of point geometry objects) of the analyzed landscape feature within each polygon of the analytical grid. |
| `{prefix}_PNc_MM` | Output field | The standardized number of categories of the analyzed landscape feature within each polygon of the analytical grid. |

## *Conceptual Diagram*

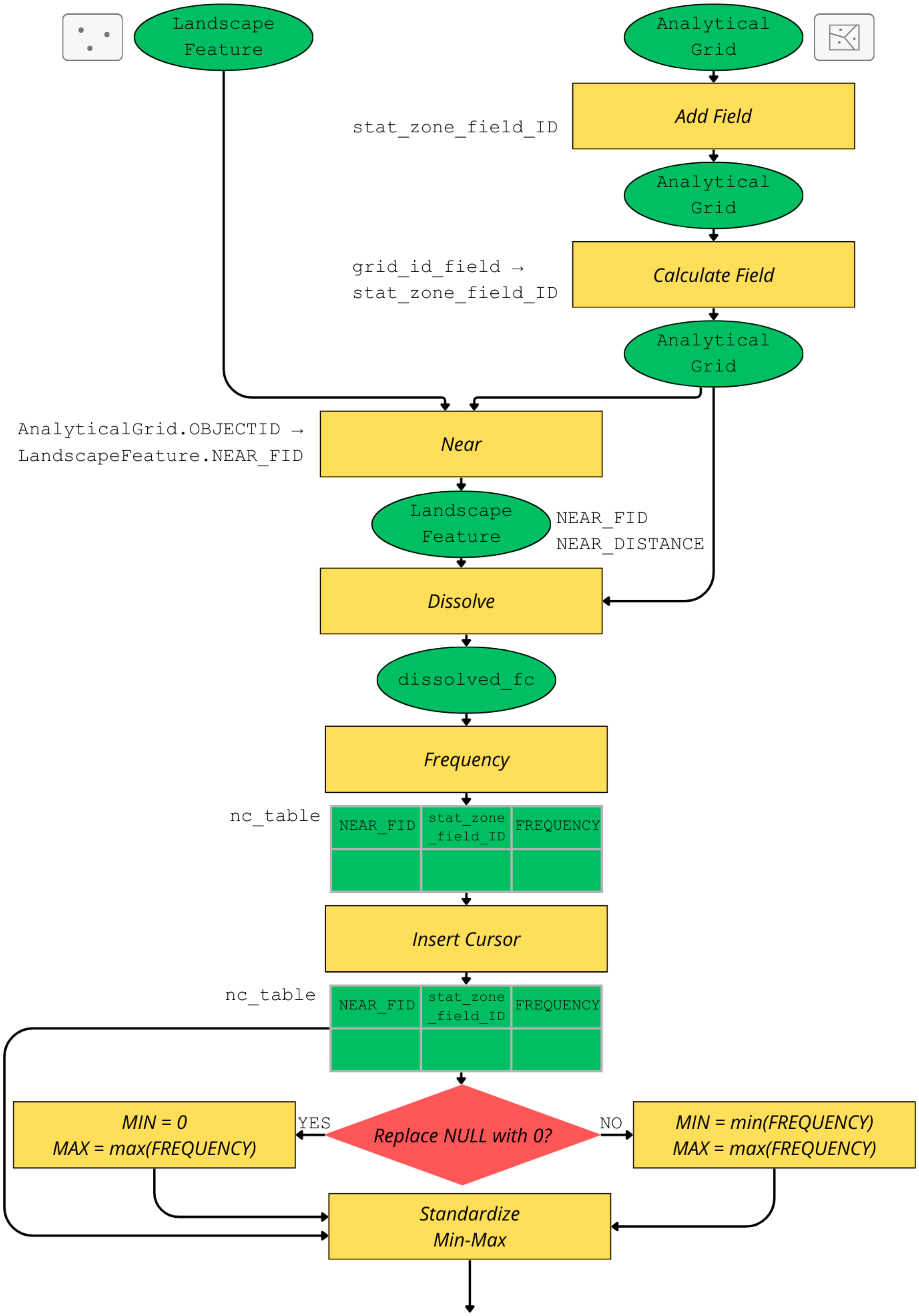


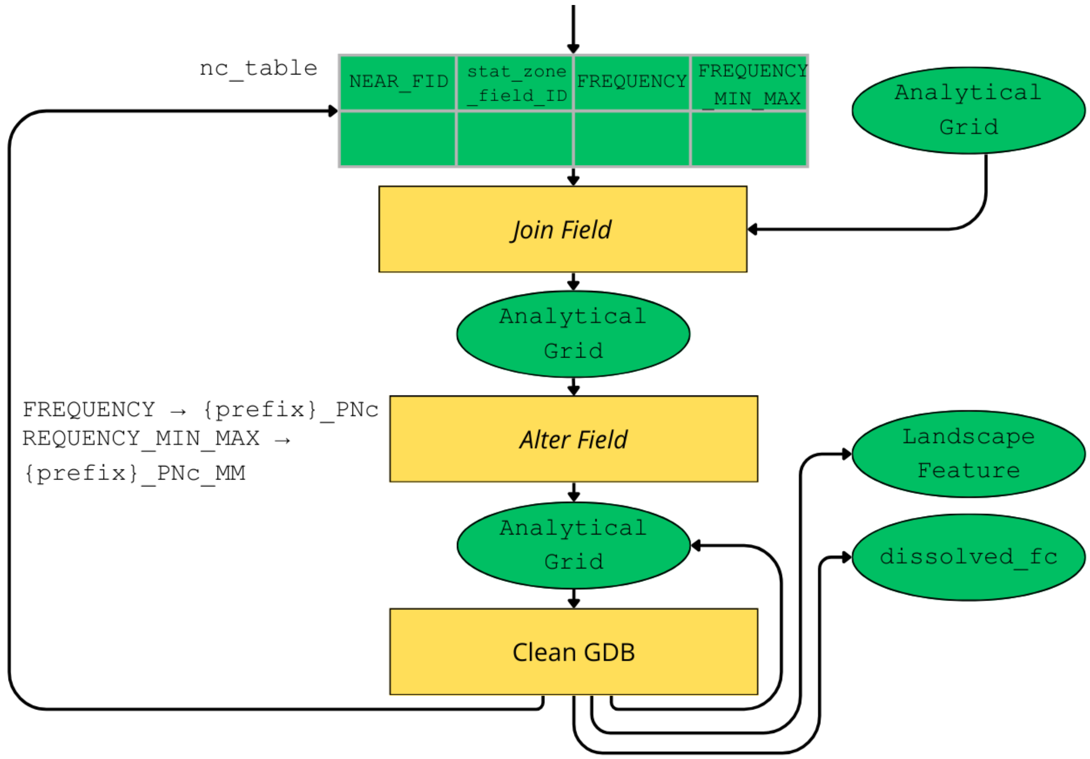

## *Algorithm*

| | |
|---|---|
| Step 1: | Retrieval of input parameters and initialization of variables; removal of data locks; verification of the existence of objects in the project geodatabase with names reserved for intermediate feature classes and non-spatial tables; and validation that the attribute table of the analytical grid feature class does not already contain fields with names intended for the output results. Additionally, the missing data handling mode (`null_handling_mode`) is specified, determining whether missing values will be replaced with 0 or remain as `NULL`. |
| Step 2: | Creation of an auxiliary statistical zone identifier by copying the primary key of the analytical grid. First, a field named `stat_zone_field_ID` of type `Long` is added. Next, using the *Calculate Field* function, the values of the primary key field (`grid_id_field`) are assigned to `stat_zone_field_ID`. This operation ensures the uniqueness of statistical zone identifiers during table join operations and prevents conflicts with system-defined fields. |
| Step 3: | The *Near* function assigns to each point (geosite) the identifier (e.g. `OBJECTID`) of the analytical grid cell in which it is located. This value is stored in the `NEAR_FID` attribute of the point feature class. This step enables the geometric association of points with the analytical grid. |
| Step 4: | Within each analytical grid cell, the *Dissolve* function merges points of the same category into multipoint features. A new feature class `dissolved_fc` is created, where multipoint objects are described by the attributes: `landscape_attr` – defining the landscape feature category, and `NEAR_FID` – describing the statistical zones. |
| Step 5: | The *Frequency* function counts the number of point feature categories for each grid cell. Aggregation is based on the `NEAR_FID` attribute. A non-spatial table `nc_table` is created, in which the `FREQUENCY` field contains the number of unique point categories for each cell of the analytical grid. |
| Step 6: | Analytical grid cells that contain no point objects are supplemented in the `nc_table` with rows where `FREQUENCY` = 0 (*Insert Cursor* function). Additionally, depending on the user-selected `null_handling_mode` option, `NULL` values may either be left as missing data or replaced with 0. This ensures a clear interpretation of the $P\_Nc$ indicator as a complete absence of point category diversity in these cells. |

| | |
|---|---|
| Step 7: | *Min–Max* standardization of the `FREQUENCY` field into `FREQUENCY_MIN_MAX` takes into account the user-selected `null_handling_mode` option:<br>• If `Replace NULL with 0 (MIN = 0, MAX from Nc)` is selected, the `MIN` value for standardization is fixed at 0, and `MAX` is determined from the actually observed `FREQUENCY` values. Cells without point objects are assigned a standardized value $Std\_P\_Nc = 0$, and any occurrence of at least one point results in a value > 0.<br>• If `Keep NULL (MIN/MAX from observed Nc only)` is selected, `NULL` values remain in the table, `MIN` and `MAX` are determined from observed $Nc$ values (excluding `NULLs`), and cells without points remain as `NULL`. If all grid cells have identical $P\_Nc$ values, all statistical zones are assigned a value of 0. |
| Step 8: | In the `nc_table`, an attribute `FREQUENCY_MIN_MAX` is created. If `MIN = MAX`, standardization is skipped and all `FREQUENCY_MIN_MAX` fields are assigned a value of 0. Otherwise, classic *Min–Max* standardization is performed using the *Standardize* function, assigning `FREQUENCY_MIN_MAX` fields values in the range ⟨0; 1⟩. |
| Step 9: | Before performing the table join operation, the script verifies whether the analytical grid attribute table contains any existing fields intended for storing the results and removes them if present. Next, using the *Join Field* function, the results of the analysis stored in the non-spatial table `nc_table` (fields `FREQUENCY` and `FREQUENCY_MIN_MAX`) are joined to the analytical grid attribute table (`grid_fl`) based on the key pair `stat_zone_field_ID` and `NEAR_FID`. After the join operation is completed, the values are copied into the analytical grid table. |
| Step 10: | Using the *Alter Field* function, the script assigns tool-compliant names to the copied fields: `FREQUENCY` is renamed to `{prefix}_PNc`, and `FREQUENCY_MIN_MAX` to `{prefix}_PNc_MM`. At the same time, readable field aliases are defined. |
| Step 11: | Remove the auxiliary fields `NEAR_FID` and `NEAR_DIST` from the point feature class of landscape features (`landscape_fl`). Remove from the geodatabase the intermediate feature class `dissolved_fc`, the non-spatial table `nc_table`, and the temporary field `stat_zone_field_ID` from the analytical grid table. Then, optimize the geodatabase using the *Compact* function. |

## Flow Diagram






## *Relevance to Geodiversity Assessment*

- *P_Nc* is a simple diversity metric, describing diversity based on the number of unique point categories (types of geosites, classes of geomorphological features, categories of point-based geological observations, etc.) occurring within each cell of the analytical grid.
- The metric reflects the typological diversity of point-based landscape features – its values increase with a greater number of different point categories located within the analytical cell.
- Low *P_Nc* values indicate the presence of points belonging to a small number of categories, which may suggest homogeneity of geological and/or geomorphological processes, uniform types of outcrops, or dominance of a single type of geosite.
- Absence of diversity (category homogeneity) means that at most one category of point landscape elements occurs within the grid cell – even if the number of points is large, the lack of multiple types results in a low metric value.
- High *P_Nc* values indicate the presence of numerous, diverse types of geosites, which may result from complex geological structures, the co-occurrence of different geomorphological processes, or a rich mosaic of forms and outcrops of varied origin.
- Unlike point abundance metrics (e.g., *P_Ne*), *P_Nc* does not reflect object density but only the diversity of their categories. Similarly to categorical metrics known from polygon-based analyses (e.g., *A_Nc*), *P_Nc* measures the qualitative component of geodiversity for point geometry objects.
- Two statistical zones may receive identical *P_Nc* values even if the total number of points in these cells is very different – the metric is sensitive to type diversity but does not distinguish occurrence intensity, concentration, or internal spatial structure of points.
- If no point objects (geosites) occur within an analytical grid cell, the *P_Nc* indicator value is determined according to the user-selected option for handling empty grid cells. Depending on the chosen variant:
    - the empty cell may be assigned *P_Nc* = 0, indicating a complete absence of point category diversity in the analytical unit,

  - or $P\_Nc$ remains `NULL`, indicating missing information and excluding the unit from further statistical analysis.
- Standardized values ($Std\_P\_Nc$) are also determined differently depending on the user's decision.
- If the "`Replace NULL with 0 (MIN = 0, MAX from Nc)`" option is selected, empty cells are treated as units with $P\_Nc$ = 0 and are included in the standardization process, where:
  - the minimum standardization value is fixed at `MIN` = 0,
  - the maximum value `MAX` is determined from the actually observed $P\_Nc$ values,
  - empty cells are assigned a standardized value $Std\_P\_Nc$ = 0.
- If the "`Keep NULL (MIN/MAX from observed Nc only)`" option is selected, empty cells remain as `NULL` and do not affect the determination of the `MIN`–`MAX` range used for standardization. *Min–Max* standardization is performed only on grid cells containing at least one point category.
- If all analytical grid cells contain at least one point category ($P\_Nc \geq 1$), *Min–Max* standardization is performed based on the actually observed $P\_Nc$ values, maintaining a consistent interpretation: each cell with at least one point category receives a standardized value $Std\_P\_Nc$ > 0.
- If all analytical grid cells have identical $P\_Nc$ values (i.e., no variation in the number of point categories across the entire area), the standardized partial criterion $Std\_P\_Nc$ is assigned a value of 0 for all cells, regardless of the selected `NULL` handling option, reflecting the lack of spatial contrast for this criterion.

# *P_Hu Unit entropy of Point Elements*


| | |
|---|---|
| Short Description: | The tool calculates the unit-based entropy ($P\_Hu$) of a selected point-based landscape feature within each polygon of the analytical grid. |
| Version: | 0.1.1 |
| Author: | Tomasz Bartuś |
| Date: | 2026-01-26 |

## *Method Description*

Unit Entropy ($P\_Hu$) is a measure of disorder and internal compositional diversity of categorical, regionalized variables expressed through point geometry object classes. It is based on the classical Shannon entropy, calculated for the probability distribution of point categories $\{q_c\}$ using the natural logarithm. The tool calculates $P\_Hu$ for each polygon of the analytical grid by analyzing the evenness of the distribution of point feature classes.

The category proportions are calculated using Equation 4.

$$q_{ck} = \frac{Nc_k}{Ne_k} \tag{4}$$

Explanations: $q_{ck}$—the proportion (probability) of occurrence of category $c$ in the $k$-th cell of the analytical grid, with $\sum_{c=1}^{n} q_{ck} = 1$; $Nc_k$—the number of points of category $c$ in the $k$-th cell of the analytical grid; $Ne_k$—the total number of points in the $k$-th cell of the analytical grid.

The $P\_Hu$ value for a given cell of the analytical grid is calculated as the sum of the entropy components for all categories occurring in that cell (5).

$$P\_Hu_k = -\sum_{c=1}^{n} q_{ck} \times ln\, q_{ck} \tag{5}$$

Explanations: $P\_Hu_k$—the unit entropy of point elements in the $k$-th cell of the analytical grid.

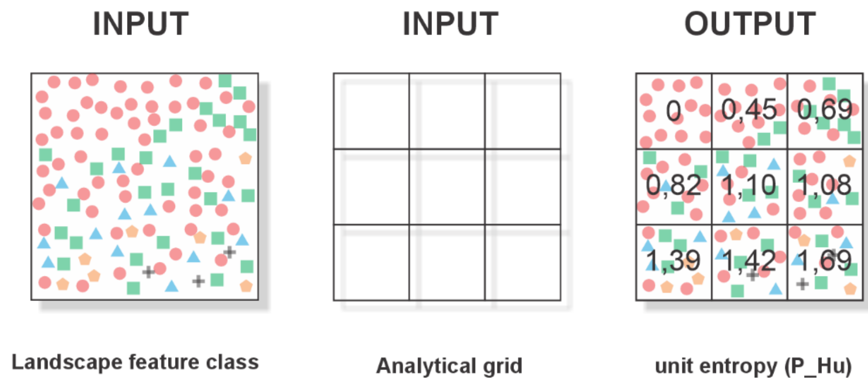
***Input data***

| Parameter | Type | Description |
|---|---|---|
| `landscape_fl` | Point feature layer | A point feature class representing the landscape feature to be analyzed, for which the unit entropy ($P\_Hu$) will be calculated within each cell of the analytical grid. |
| `landscape_attr` | Field | The field in the input landscape feature that defines the categories (classes) of the analyzed landscape feature, which will be considered within each cell of the analytical grid. |
| `grid_fl` | Polygon feature layer | A polygon feature class representing the analytical grid, where for each cell (statistical zone) the partial geodiversity criterion is calculated based on the unit entropy ($P\_Hu$) of the selected landscape feature. |
| `grid_id_field` | Field | An attribute of the analytical grid that uniquely identifies each cell (statistical zone) in which geodiversity indices will be calculated. |

**Output data**

| Index | Type | Description |
|---|---|---|
| `{prefix}_PHu` | Output field | The unit entropy ($P\_Hu$) of the analyzed landscape feature within each polygon of the analytical grid. |
| `{prefix}_PHu_MM` | Output field | The standardized unit entropy of the analyzed landscape feature within each polygon of the analytical grid. |

## *Conceptual Diagram*

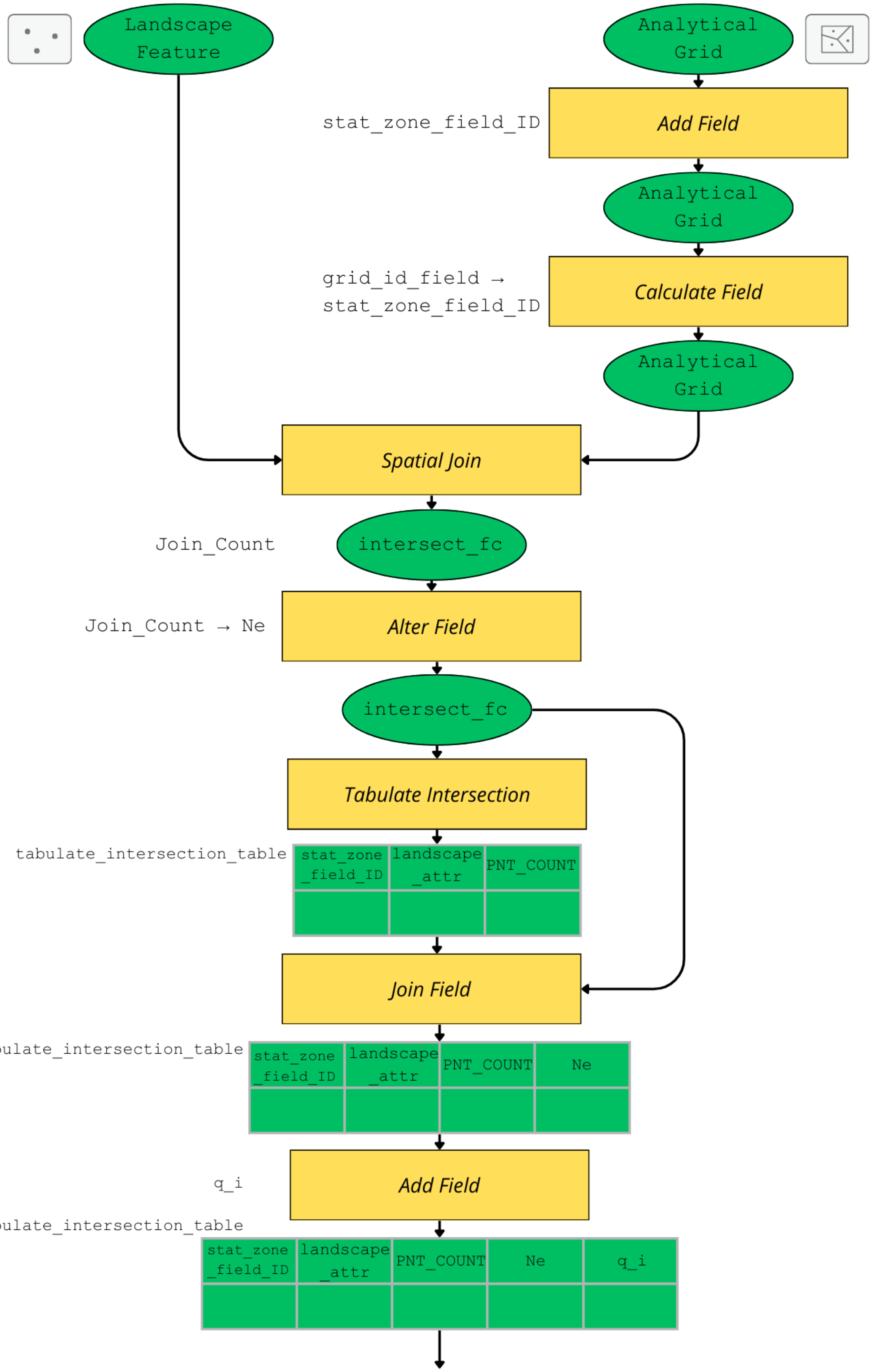


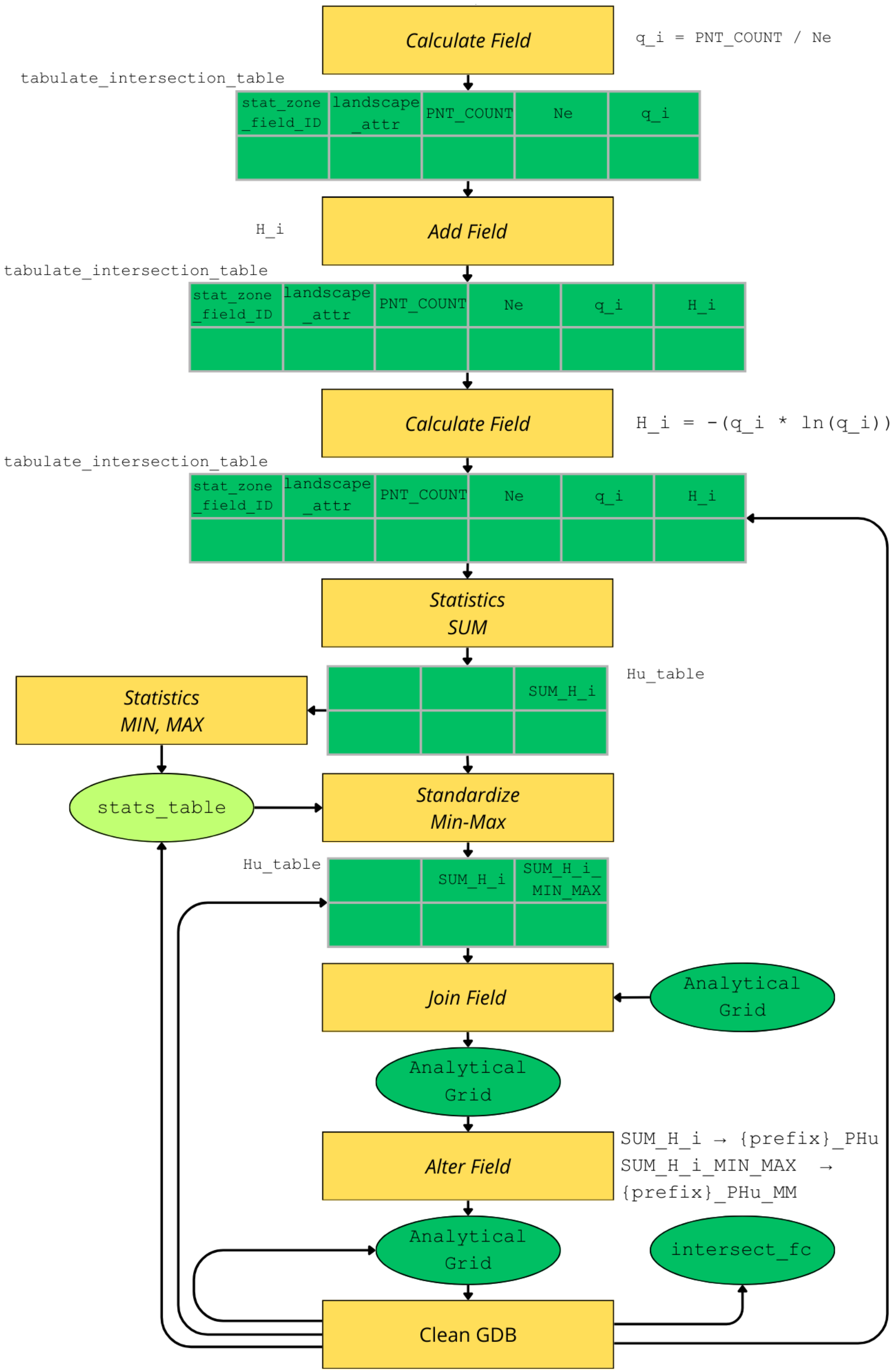

## *Algorithm*

| | |
|---|---|
| Step 1: | Retrieval of input parameters and initialization of variables; removal of data locks; verification of the existence of objects in the project geodatabase with names reserved for intermediate feature classes and non-spatial tables; and validation that the attribute table of the analytical grid feature class does not already contain fields with names intended for the output results. |
| Step 2: | Creation of an auxiliary statistical zone identifier by copying the primary key of the analytical grid. First, a field named `stat_zone_field_ID` of type `Long` is added. Next, using the *Calculate Field* function, the values of the primary key field (`grid_id_field`) are assigned to `stat_zone_field_ID`. This operation ensures the uniqueness of statistical zone identifiers during table join operations and prevents conflicts with system-defined fields. |
| Step 3: | Perform a *Spatial Join* in which each cell of the analytical grid (`grid_fl`) is assigned the number of point feature objects representing the selected landscape feature (`landscape_fl`) located within it. An intermediate feature class `intersect_fc` is created, containing the attribute `Join_Count` representing the number of points in the given grid cell. The tool uses the `JOIN_ONE_TO_ONE` option and the `INTERSECT` spatial relationship. |
| Step 4: | Rename the `Join_Count` attribute to `Ne` using the *Alter Field* function. |
| Step 5: | Apply the *Tabulate Intersection* function to create a non-spatial table `tabulate_intersection_table`, in which, for each cell of the analytical grid, the following are listed: the statistical zone identifier (`stat_zone_field_ID`), the category of point landscape objects (`landscape_attr`), and the number of points in that category (`PNT_COUNT`). For each grid cell and each category occurring within it, information on the number of objects is included. |
| Step 6: | Using the *Join Field* function, join the attribute table of the intermediate feature class `intersect_fc` – specifically the `Ne` attribute – to the table `tabulate_intersection_table`. In `tabulate_intersection_table`, a new attribute `q_i` is created, whose values are then calculated using the formula `q_i` = `PNT_COUNT` / `Ne` with a condition that prevents division by zero. |

| | |
|---|---|
| Step 7: | Add a field `H_i` to the table `tabulate_intersection_table` using the *Add Field* function. Using *Calculate Field*, calculate the entropy component for each point category within the statistical zone with the formula $H\_i = -(q\_i \times ln(q\_i))$. Categories with `q_i = 0` are assigned `H_i = 0`. |
| Step 8: | For each grid cell, the *Statistics* function sums the `H_i` values calculated for individual categories. A non-spatial table `Hu_table` is created, in which the field `SUM_H_i` contains the unit entropy (`P_Hu`) value for each statistical zone (grouping by `stat_zone_field_ID`). |
| Step 9: | Using the *Statistics* function, the temporary table `stats_table` created in memory (`memory`) is used to calculate the minimum and maximum of `SUM_H_i`. If `MIN = MAX`, standardization is skipped, and all records in Hu_table are assigned a value of 0. Otherwise (if `MIN ≠ MAX`), the *Standardize Field* function with the *Min–Max* method is applied to create the attribute `SUM_H_i_MIN_MAX` in `Hu_table`, containing standardized values in the range ⟨0; 1⟩. |
| Step 10: | Before the next algorithm step, the script checks for the existence of the `SUM_H_i` and `SUM_H_i_MIN_MAX` fields in the analytical grid attribute table (`grid_fl`) and removes them if they are present. Then, using the *Join Field* function, the `SUM_H_i` and `SUM_H_i_MIN_MAX` fields from the non-spatial table `Hu_table` are joined to the analytical grid attribute table (`grid_fl`). After the join operation is completed, these values are copied to the grid feature class. |
| Step 11: | The *Alter Field* function is used to rename the copied attributes according to the tool's naming convention: `SUM_H_i` → `{prefix}_PHu` and `SUM_H_i_MIN_MAX` → `{prefix}_PHu_MM` , and assigns them appropriate aliases. |
| Step 12: | Using the *Delete* function, the intermediate feature class `intersect_fc` and the non-spatial tables `tabulate_intersection_table` and `Hu_table` are removed, as well as the temporary table `stat_table` stored in memory and the temporary field `stat_zone_field_ID` from the analytical grid feature class. Then, the geodatabase is optimized using the *Compact* function. |

## *Flow Diagram*




Calculate Field
q_i = PNT_COUNT / Ne
tabulate_intersection_table
stat_zone_field_ID
landscape_attr
PNT_COUNT
Ne
q_i
H_i
Add Field
tabulate_intersection_table
stat_zone_field_ID
landscape_attr
PNT_COUNT
Ne
q_i
H_i
Calculate Field
H_i = -(q_i * ln(q_i))
tabulate_intersection_table
stat_zone_field_ID
landscape_attr
PNT_COUNT
Ne
q_i
H_i
Statistics SUM
SUM_H_i
Hu_table
Statistics MIN, MAX
stats_table
Standardize Min-Max
Hu_table
SUM_H_i
SUM_H_i_MIN_MAX
Join Field
Analytical Grid
Analytical Grid
Alter Field
SUM_H_i → {prefix}_PHu
SUM_H_i_MIN_MAX → {prefix}_PHu_MM
Analytical Grid
intersect_fc
Clean GDB

## *Relevance to Geodiversity Assessment*

- $P\_Hu$ measures the diversity of point feature categories within a base cell of the analytical grid.
- The values of $P\_Hu$ depend on: the number of categories ($Nc_k$), the proportions of each class ($q_1$, $q_2$, …, $q_c$), and the evenness of their distribution. They do not depend on the absolute number of points in the cell (as long as the category proportions remain unchanged), the cell area, or the spatial density of points.
- Unlike the classical Shannon–Weaver Diversity Index ($SHDI$), $P\_Hu$ does not account for the area of features, making it suitable for analyzing the diversity of point objects – such as geosites, geomorphological forms, or field observations.
- $P\_Hu$ increases with greater evenness of category proportions and decreases when one category dominates or strongly exceeds the others.
- High $P\_Hu$ values indicate high heterogeneity of categories and a complex structure of point landscape elements, often characteristic of areas with complex geological structure or multiprocess genesis.
- Low $P\_Hu$ values suggest dominance of a single category and low internal diversity of point landscape features.
- Homogeneity ($P\_Hu$ = 0) is defined as the occurrence of only a single category of point features within the cell.
- The absence of point objects within a base cell results in `NULL` values for both $P\_Hu$ and $Std\_P\_Hu$. Such values prevent the aggregation of partial criteria at the level of landscape element diversity and overall geodiversity (in ArcGIS Pro, the sum of partial criterion values (`Float` type) and `NULL` returns `NULL`).
- Cells with identical category proportions ($q_1$ … $q_c$) have the same $P\_Hu$ values, regardless of the total number of points ($Ne$). Changing the number of points does not affect the index if the proportional structure remains the same.
- For the same number of categories ($Nc$), $P\_Hu$ values may vary slightly depending on the number of points ($Ne$). In smaller samples, entropy is more sensitive to random distributions of points among categories.
- $P\_Hu$ values range from ⟨0; 1⟩, so the index does not require additional standardization.

## 5.2. Metrics for Line-Based Datasets

Geodiversity Tools provides a single diversity metric for landscape features represented by line geometry objects – Total Length of Line Elements ($L\_Tl$).

# *L_Tl Total Length of Line Elements*


| | |
|---|---|
| Short Description: | The tool calculates the total length of elements of a selected landscape feature (line feature class) within each polygon of the analytical grid. |
| Version: | 0.1.1 |
| Author: | Tomasz Bartuś |
| Date: | 2026-01-26 |

### *Method Description*

Total Length of Line Elements ($L\_Tl$) is a simple diversity metric designed for the analysis and assessment of landscape features described by categorical, regionalized variables and represented by line geometry feature classes. The metric evaluates the diversity of a selected feature based on the total length of line objects occurring within the boundaries of each cell of the analytical grid (6).

$$L\_Tl_k = \sum_{i=1}^{n} l_i \qquad (6)$$

Explanations: $L\_Tl_k$—the total length of line objects of the selected landscape feature in the $k$-th cell of the analytical grid; $n$—the total number of line elements in the $k$-th cell of the analytical grid; $l_i$—the length of the $i$-th line geometry object in the $k$-th statistical zone.

**Notes on missing data**:

In the case of analytical grid cells containing no linear objects (e.g., faults), the user can choose how to handle missing data during the calculation and standardization of the $L\_Tl$ indicator. Two tool behavior options are available:

1. `Replace NULL with 0 (MIN = 0, MAX from L_Tl)` – missing values (grid cells without lines) are replaced with 0. Standardization is performed using a minimum value `MIN` = 0 and a maximum value `MAX` determined from the observed line lengths in the grid cells.
2. `Keep NULL (MIN/MAX from observed L_Tl only)` – missing values remain as `NULL`, and standardization is based on the actual range of observed $L\_Tl$ values, with empty cells not affecting the determination of `MIN` and `MAX`.

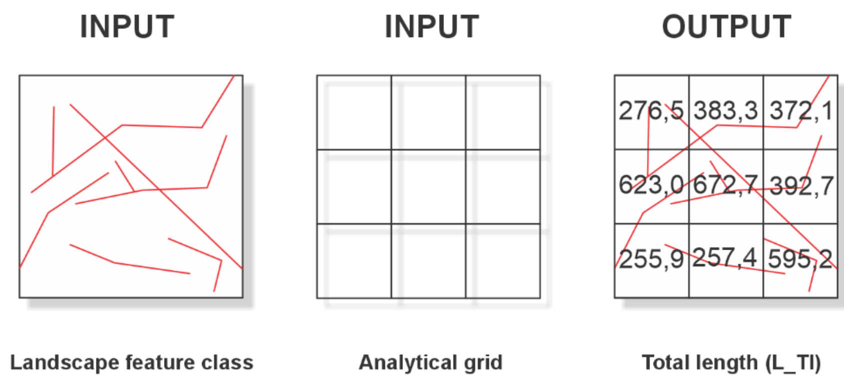
***Input data***

| Parameter | Type | Description |
|---|---|---|
| `landscape_fl` | Line feature layer | A line feature class representing the landscape feature to be analyzed. The lengths of its features are aggregated within each polygon of the analytical grid. |
| `grid_fl` | Polygon feature layer | A polygon feature class representing the analytical grid. For each grid cell (statistical zone), the partial geodiversity criterion is calculated based on the total length of the selected line features within that cell. |
| `grid_id_field` | Field | An attribute of the analytical grid that uniquely identifies each grid cell (statistical zone) in which the geodiversity indices are calculated. |
| `null_handling_mode` | String / Choice | Specifies how analytical grid cells without linear objects are handled during the calculation and standardization of $L\_Tl$. Available options:<br>• `Replace NULL with 0 (MIN = 0, MAX from L_Tl)` – missing $L\_Tl$ values are replaced with 0, and standardization is performed using a minimum value `MIN` = 0 and a maximum value `MAX` determined from the observed line lengths.<br>• `Keep NULL (MIN/MAX from observed L_Tl only)` – missing values remain as `NULL` and do not affect the determination of `MIN`/`MAX` during standardization. |

**Output data**

| Index | Type | Description |
|---|---|---|
| `{prefix}_LTl` | Output field | The total length of all linear features of the analyzed landscape feature within each polygon of the analytical grid. |
| `{prefix}_LTl_MM` | Output field | The standardized total length of all linear features of the analyzed landscape feature within each polygon of the analytical grid. |

## *Conceptual Diagram*

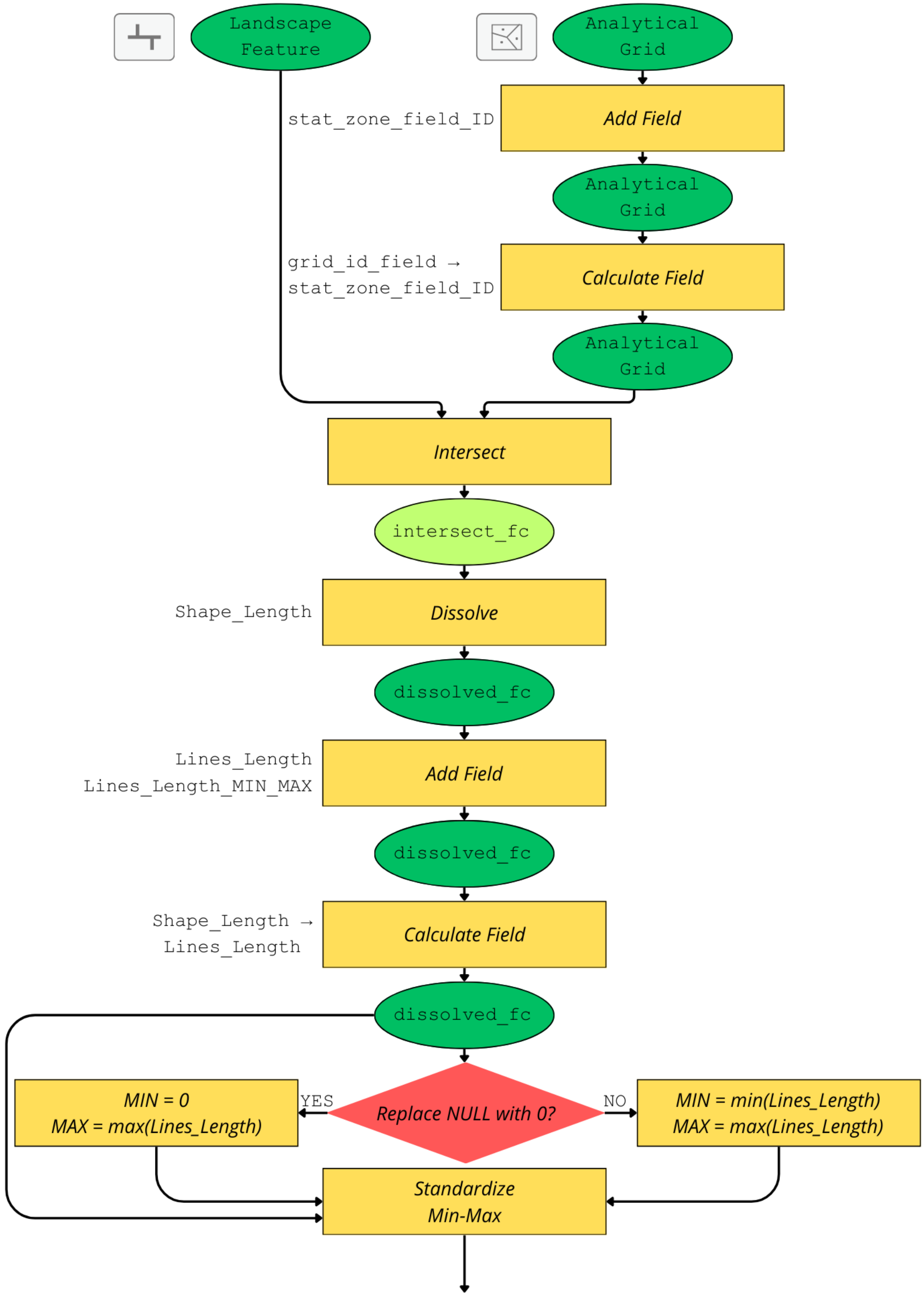


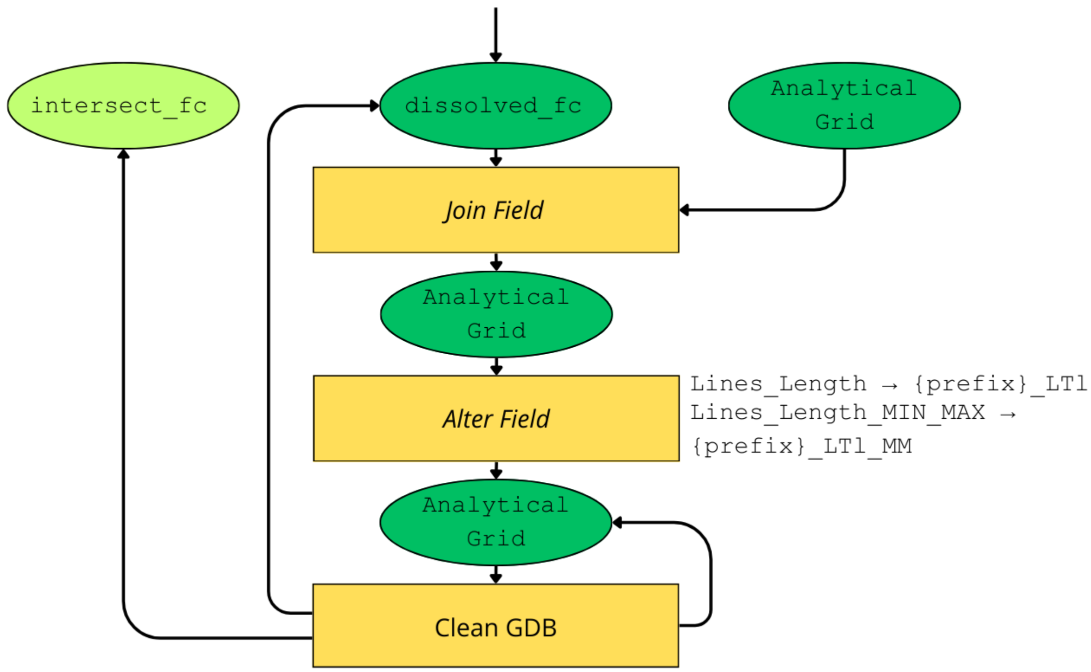

## *Algorithm*

| | |
|---|---|
| Step 1: | Retrieval of input parameters and initialization of variables; removal of data locks; verification of the existence of objects in the project geodatabase with names reserved for intermediate feature classes and non-spatial tables; and validation that the attribute table of the analytical grid feature class does not already contain fields with names intended for the output results. Additionally, the missing data handling mode (`null_handling_mode`) is specified, determining whether missing values will be replaced with 0 or remain as `NULL`. |
| Step 2: | Creation of an auxiliary statistical zone identifier by copying the primary key of the analytical grid. First, a field named `stat_zone_field_ID` of type `Long` is added. Next, using the *Calculate Field* function, the values of the primary key field (`grid_id_field`) are assigned to `stat_zone_field_ID`. This operation ensures the uniqueness of statistical zone identifiers during table join operations and prevents conflicts with system-defined fields. |
| Step 3: | The *Intersect* function intersects the line geometry feature class representing the selected landscape feature (`landscape_fl`) with the polygon feature class of the analytical grid (`grid_fl`). As a result, an intermediate feature class `intersect_fc` is created in the RAM memory, containing line segments bounded by the boundaries of individual statistical zones. Each line segment has the attribute `stat_zone_field_ID`, identifying the grid cell in which it is located. |
| Step 4: | Using the *Dissolve* function, line segments are aggregated within individual cells of the analytical grid. An intermediate feature class `dissolved_fc` is created, in which all line segments belonging to the same grid cell (grouped by `stat_zone_field_ID`) are merged, and their total length is stored in the `Shape_Length` field. As a result, each record in `dissolved_fc` represents the total length of line objects within a given grid cell. The `dissolved_fc` class is not created in memory because it will be used in the *Join Field* function. |
| Step 5: | In the `dissolved_fc` table, the `Lines_Length` attribute field is created using the *Add Field* function, and the values from the `Shape_Length` attribute are copied into it using *Calculate Field*. This operation is necessary because the system field `Shape_Length` cannot be used directly in the table join process. In the same table and using the same function, the `Lines_Length_MIN_MAX` field is created to store values standardized to the range ⟨0; 1⟩. |

| | |
|---|---|
| Step 6: | Manual *Min–Max* standardization is performed on the `Lines_Length` field, without using the *Standardize* function. If the user has selected Replace `NULL` with 0, missing values are replaced with 0 prior to standardization, and values in the `Lines_Length_MIN_MAX` field are calculated using `MIN` = 0 and `MAX` equal to the maximum observed line length. Otherwise, missing values remain as `NULL`, and `MIN` and `MAX` are determined from the actually observed values. |
| Step 7: | Prior to the next algorithm step, the script checks for the existence of the `Lines_Length` and `Lines_Length_MIN_MAX` fields in the analytical grid feature class attribute table and removes them if they are present. Then, using the *Join Field* function, the attribute table of the `dissolved_fc` feature class is joined to the analytical grid attribute table. After the join operation is completed, the values stored in the `Lines_Length` and `Lines_Length_MIN_MAX` fields are copied to the grid feature class. If `null_handling_mode = Replace NULL` with 0, `NULL` values in the target fields are replaced with 0 after the *Join Field* operation. |
| Step 8: | The *Alter Field* function renames the copied fields according to the tool's naming convention. The `Lines_Length` field is renamed to `{prefix}_LTl`, and the `Lines_Length_MIN_MAX` field is renamed to `{prefix}_LTl_MM`. Appropriate aliases are assigned to the attribute names. |
| Step 9: | Delete the intermediate datasets (`intersect_fc` and `dissolved_fc`) and the temporary field `stat_zone_field_ID` from the analytical grid table using the *Delete* function, and then optimize the input geodatabase using the *Compact* function. |

## Flow Diagram






## *Relevance to Geodiversity Assessment*

- *L_TL* is a basic quantitative metric for characterizing linear landscape elements, describing the total length of all line objects (e.g., faults, streams, lithostratigraphic boundaries, morphologic edges) located within each unit of the analytical grid.
- The metric reflects the intensity of occurrence and degree of development of linear landscape structures – its values increase with greater total length of objects intersecting the grid cell, regardless of their number.
- Low *L_TL* values indicate limited development of linear structures, which may suggest simpler geological structure, a poorly developed drainage network, a limited number of erosional edges, or low morphotectonic diversity.
- High *L_TL* values indicate strong development of linear elements, which may result, among others, from complex tectonics, an extensive drainage network, numerous lithological boundaries, or intensive morphogenetic processes leading to the formation of numerous edges and transition zones.
- The absence of linear elements in a base cell results in *L_Tl* = 0, indicating homogeneity of the cell with respect to linear features and a complete lack of contribution from this criterion to the spatial diversity of the landscape.
- Unlike category diversity metrics (e.g., *A_Nc*, *P_Nc*), *L_TL* measures a quantitative rather than qualitative component – it describes “how much line” occurs within the cell, not “which types” of lines are present.
- Unlike object count metrics (e.g., *P_Ne*, *L_Ne*), *L_TL* is not sensitive to the number of segments but to their total length – a cell containing a single long dislocation may have a higher diversity value than a cell with many short faults.
- Two analytical grid cells may receive identical *L_TL* values despite having very different numbers of line objects – the metric does not distinguish whether the total length is distributed among many short segments or concentrated in a single long element.
- Unlike relative metrics such as line density, *L_TL* is an absolute measure that does not account for the area of the statistical zone or the relationship between length and its size. As a result,

*L_TL* values are strongly dependent on the size of grid units, which may hinder direct comparisons between cells of different sizes or between grids of different resolutions. Therefore, to ensure comparability of results in inter-area analyses and in studies combining grids of different geometry and spatial scales, the *L_TL* metric requires standardization.

- If no linear objects occur within an analytical grid cell, the *L_Tl* indicator value is determined according to the user-selected option for handling empty grid cells. Depending on the chosen variant:
    - the empty cell may be assigned *L_Tl* = 0, indicating a complete absence of linear landscape elements in the analytical unit,
    - or *L_Tl* remains `NULL`, indicating missing information and excluding the unit from further statistical analysis.
- Standardized values (*Std_L_Tl*) are also determined differently depending on the user's decision.
- *L_Tl* values are standardized using the *Min–Max* method based on the total range of linear feature lengths within the analytical grid cells.
- If the "`Replace NULL with 0 (MIN = 0, MAX from L_Tl)`" option is selected, empty cells are treated as units with *L_Tl* = 0 and are included in the standardization process, where:
    - the minimum standardization value is fixed at `MIN` = 0,
    - the maximum value `MAX` is determined from the actually observed *L_Tl* values in cells containing linear objects,
    - empty cells are assigned a standardized value *Std_L_Tl* = 0.
- If the "`Keep NULL (MIN/MAX from observed L_Tl only)`" option is selected, empty cells remain as `NULL` and do not affect the determination of the `MIN`–`MAX` range used for standardization. *Min–Max* standardization is performed only on grid cells containing at least one linear object.
- If all analytical grid cells contain at least one linear object (*L_Tl* > 0), *Min–Max* standardization is performed based on the actually observed *L_Tl* values, maintaining a consistent interpretation: each cell with linear objects receives a standardized value *Std_L_Tl* > 0.
- If all analytical grid cells have identical *L_Tl* values (i.e., no variation in total linear feature length across the entire area), the standardized partial criterion *Std_L_Tl* is assigned a value of 0 for all cells, regardless of the selected `NULL` handling option, reflecting the lack of spatial contrast for this criterion.

## 5.3. Metrics for Polygon-Based Datasets

Geodiversity Tools provides three diversity metrics for landscape features represented by polygon geometry objects: Number of Polygon Elements ($A\_Ne$), Number of Polygon Categories ($A\_Nc$), and the Shannon–Weaver Diversity Index ($A\_SHDI$).

# *A_Ne Number of Polygon Elements*

| | |
|---|---|
| Short Description: | The tool calculates, for a selected landscape feature (polygon feature class), the number of its elements within each polygon of the analytical grid. |
| Version: | 0.1.1 |
| Author: | Tomasz Bartuś |
| Date: | 2026-01-26 |

### *Method Description*

Number of Polygon Elements ($A\_Ne$) is a basic measure of landscape diversity, designed for the analysis and assessment of features described by categorical, regionalized variables and represented by polygon geometry objects. The metric determines how many individual elements (polygons) of a given feature occur within the boundaries of each cell of the analytical grid (7).

$$A\_Ne_k = \sum_{i=1}^{n} e_i \qquad (7)$$

Explanations: $A\_Ne_k$—the number of polygon elements in cell $k$ of the analytical grid; $n$—the total number of polygon elements in cell $k$ of the analytical grid; $e_i$—a single polygon geometry object in cell $k$ of the analytical grid.

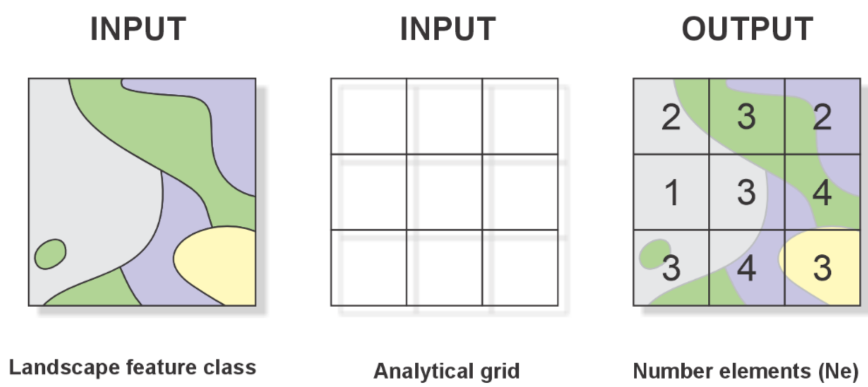
### *Input data*

| Parameter | Type | Description |
|---|---|---|
| `landscape_fl` | Polygon feature layer | A polygon feature class representing the landscape feature to be analyzed. The number of its features is counted within each polygon of the analytical grid. |
| `grid_fl` | Polygon feature layer | A polygon feature class representing the analytical grid. For each grid cell (statistical zone), the partial geodiversity criterion is calculated based on the number of features of the selected landscape feature. |
| `grid_id_field` | Field | An attribute of the analytical grid that uniquely identifies each grid cell (statistical zone) in which geodiversity indices are calculated. |

**Output data**

| Index | Type | Description |
|---|---|---|
| `{prefix}_ANe` | Output field | The number of elements (individual polygons) of the analyzed landscape feature within each polygon of the analytical grid. |
| `{prefix}_ANe_MM` | Output field | The standardized number of elements of the analyzed landscape feature within each polygon of the analytical grid. |

***Conceptual Diagram***

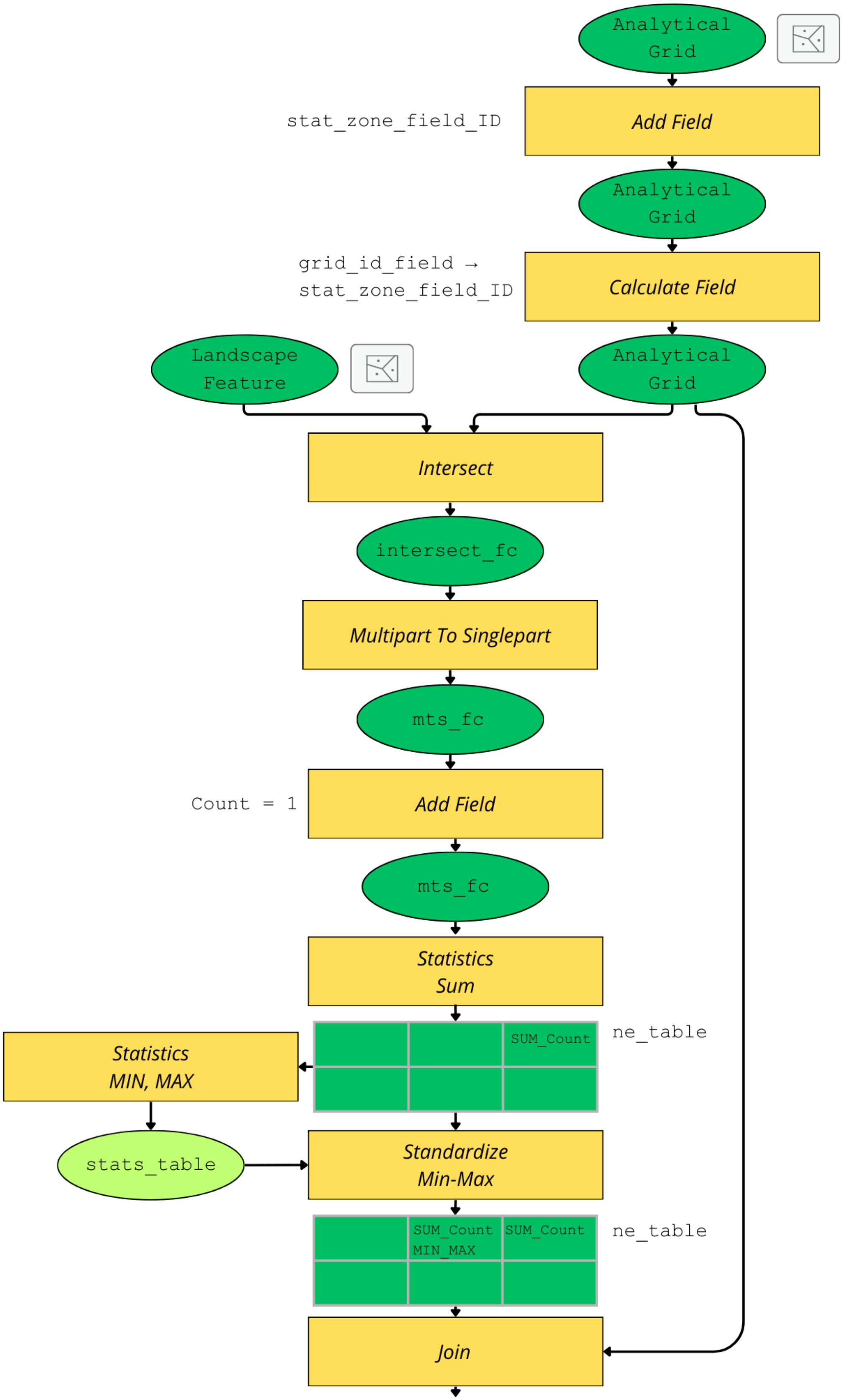


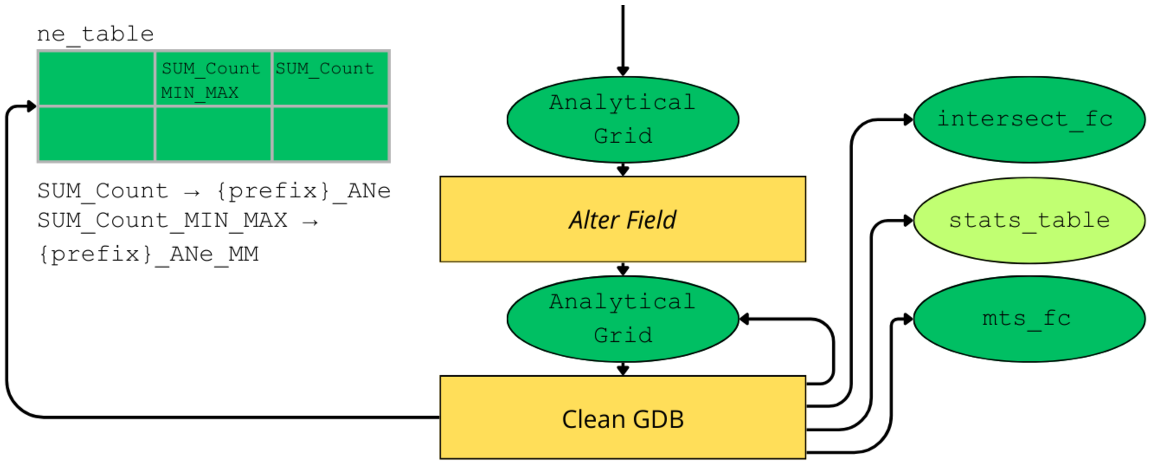

***Algorithm***

| | |
|---|---|
| Step 1: | Retrieval of input parameters and initialization of variables; removal of data locks; verification of the existence of objects in the project geodatabase with names reserved for intermediate feature classes and non-spatial tables; and validation that the attribute table of the analytical grid feature class does not already contain fields with names intended for the output results. |
| Step 2: | Creation of an auxiliary statistical zone identifier by copying the primary key of the analytical grid. First, a field named `stat_zone_field_ID` of type `Long` is added. Next, using the *Calculate Field* function, the values of the primary key field (`grid_id_field`) are assigned to `stat_zone_field_ID`. This operation ensures the uniqueness of statistical zone identifiers during table join operations and prevents conflicts with system-defined fields. |
| Step 3: | Intersect the selected landscape feature layer (`landscape_fl`) with the analytical grid feature layer (`grid_fl`). The *Intersect* function creates the intermediate feature class `intersect_fc`, containing the `stat_zone_field_ID` attribute used for grouping polygons. |
| Step 4: | Convert multipart polygons within successive cells of the analytical grid into singlepart polygons using the *Multipart To Singlepart* function. An intermediate feature class `mts_fc` is created. This operation ensures that each individual landscape patch is counted as a separate element (`Ne`) within a grid cell. |
| Step 5: | Add a new field `Count` to the attribute table of `mts_fc` and assign the value `Count = 1` to all objects in the table. |

| | |
|---|---|
| Step 6: | Using the *Statistics* function, calculate the total number of polygon geometry elements within each cell of the analytical grid. The summation result is stored in the `SUM_Count` field of the non-spatial table `ne_table`. The statistics are grouped based on the `stat_zone_field_ID` field. |
| Step 7: | The *Standardize* function (*Min–Max* method) is used to standardize the calculation results. Before standardization, the *Statistics* function calculates the minimum and maximum values of the `SUM_Count` attribute. The results are stored in the temporary statistics table stats_table created in memory (`memory`). In the `ne_table`, the `SUM_Count_MIN_MAX` field is created. If there is no data variability (`MIN = MAX`), standardization is not performed and all records are assigned a value of 0. Otherwise, the *Standardize* function calculates *Min–Max* standardized values of the `SUM_Count` attribute in the `SUM_Count_MIN_MAX` field. |
| Step 8: | The non-spatial table `ne_table` is joined to the attribute table of the analytical grid feature class (`grid_fl`) using the *Join* function. The join key is the `stat_zone_field_ID` field. |
| Step 9: | Before the next algorithm step, the script checks for the existence of the `SUM_Count` and `SUM_Count_MIN_MAX` fields in the analytical grid attribute table and removes them if present. Then, the values stored in the `SUM_Count` and `SUM_Count_MIN_MAX` attributes are copied to the analytical grid attribute table. |
| Step 10: | The copied attributes are renamed to `{prefix}_ANe` and `{prefix}_ANe_MM`, respectively, and are assigned aliases derived from the name of the landscape feature layer. |
| Step 11: | The tool removes intermediate feature classes and tables, the temporary `stat_zone_field_ID` field from the analytical grid table, as well as other temporary objects stored in memory from the geodatabase, and then performs database optimization. |

## Flow Diagram






## Relevance to Geodiversity Assessment

- *A_Ne* is a simple diversity metric that describes diversity based on the number of elements (polygons) of a selected landscape feature occurring within each unit of the analytical grid.
- The metric reflects the degree of spatial fragmentation of the analyzed landscape feature – its values increase with a greater number of distinct objects (e.g., many small lithological units, geomorphological forms, etc.).
- Low *A_Ne* values indicate spatial homogeneity within statistical zones (a small number of elements), typical of compact, non-fragmented landscape features.
- Absence of diversity (homogeneity) is defined as the occurrence of only a single polygon of the analyzed landscape feature within one cell of the analytical grid.
- High *A_Ne* values indicate the presence of numerous, fragmented polygons, which may result from a high degree of geological complexity or relief diversity in the studied area.
- Unlike area-based metrics (e.g., *SHDI*), *A_Ne* does not consider polygon size, only their number – thus measuring exclusively the structural component of diversity.
- Two grid cells may receive identical *A_Ne* values even when the areas of the objects differ, as long as the number of elements is the same – this makes the metric sensitive to mosaic patterns but not to differences in areal proportions.

# A_Nc Number of Polygon Categories

| | |
|---|---|
| Short Description: | The tool calculates the number of categories of a selected polygon-based landscape feature within each polygon of the analytical grid. |
| Version: | 0.1.1 |
| Author: | Tomasz Bartuś |
| Date: | 2026-01-26 |

## Method Description

The Number of Polygon Categories ($A\_Nc$) is a simple measure of landscape diversity, intended for the analysis and assessment of features described by regionalized discrete variables, represented by classes of polygon features. The metric determines the number of categories (classes) of a given landscape feature occurring within the boundaries of each analytical grid cell (8).

$$A\_Nc_k = \sum_{i=1}^{n} C_i \tag{8}$$

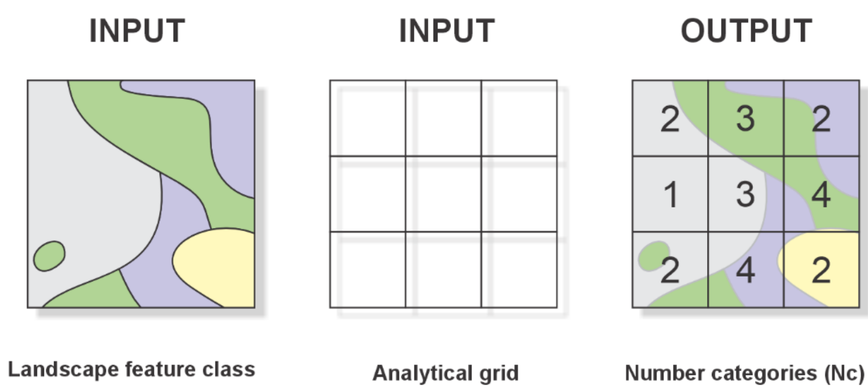
Explanations: $A\_Nc_k$—the number of polygon categories in the $k$-th elementary grid cell; $n$—the total number of polygon elements in the $k$-th elementary grid cell; $c_i$—a unique polygon category in the $k$-th elementary grid cell.

## Input data

| Parameter | Type | Description |
|---|---|---|
| `landscape_fl` | Polygon feature layer | A polygon feature class representing the landscape feature to be analyzed. Its categories are counted within each polygon of the analytical grid. |
| `landscape_attr` | Field | The attribute field of the input landscape feature that defines its categories (classes), which are counted within each grid cell. |
| `grid_fl` | Polygon feature layer | A polygon feature class representing the analytical grid. For each grid cell (statistical zone), the partial geodiversity criterion is calculated based on the number of categories of the selected landscape feature. |
| `grid_id_field` | Field | An attribute of the analytical grid that uniquely identifies each grid cell (statistical zone) in which geodiversity indices are calculated. |

| Output data | | |
|---|---|---|
| **Index** | **Type** | **Description** |
| `{prefix}_ANc` | Output field | The number of categories (polygon types) of the analyzed landscape feature within each polygon of the analytical grid. |
| `{prefix}_ANc_MM` | Output field | The standardized number of categories of the analyzed landscape feature within each polygon of the analytical grid. |

## *Conceptual Diagram*

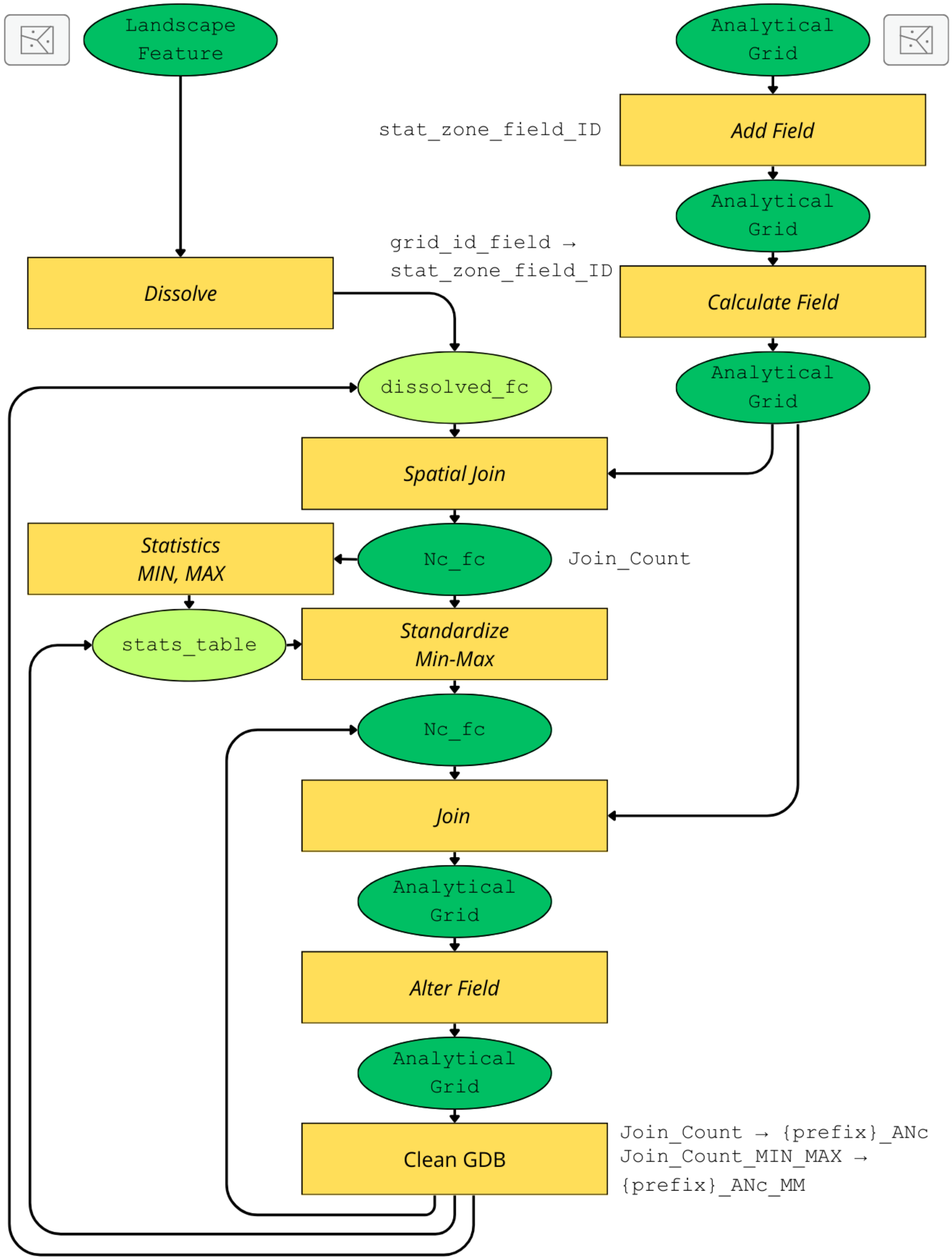



## *Algorithm*

| | |
|---|---|
| Step 1: | Retrieval of input parameters and initialization of variables; removal of data locks; verification of the existence of objects in the project geodatabase with names reserved for intermediate feature classes and non-spatial tables; and validation that the attribute table of the analytical grid feature class does not already contain fields with names intended for the output results. |
| Step 2: | Creation of an auxiliary statistical zone identifier by copying the primary key of the analytical grid. First, a field named `stat_zone_field_ID` of type `Long` is added. Next, using the *Calculate Field* function, the values of the primary key field (`grid_id_field`) are assigned to `stat_zone_field_ID`. This operation ensures the uniqueness of statistical zone identifiers during table join operations and prevents conflicts with system-defined fields. |
| Step 3: | Using the *Dissolve* function, all objects of the polygon feature class representing the selected landscape feature are merged into multipart polygons. The dissolve key is the input attribute that divides this feature into $n$ categories. The operation is performed using an intermediate feature class `Dissolved_fc` created in operational memory (`memory`), without saving a permanent feature class in the geodatabase. |
| Step 4: | The polygon analytical grid feature class is joined with the created multipart polygon feature class using the *Spatial Join* tool. An intermediate polygon feature class `Nc_fc` is created, in which the values of the `Join_Count` attribute are automatically calculated. For each elementary grid cell, this attribute stores the number of multipart polygons (categories of the analysed feature) occurring within it. The `INTERSECT` spatial relationship and the `JOIN_ONE_TO_ONE` option |

| | |
|---|---|
| | are applied. |
| Step 5: | The *Statistics* function calculates the minimum and maximum values of the `Join_Count` attribute. The results are saved to a temporary table `stat_table` stored in operational memory (`memory`). If there is no variation in the data (`MIN = MAX`), standardisation is not performed and a value of 0 is assigned to all records of the `Nc_fc` table. If `MIN ≠ MAX`, the *Standardize* function performs *Min–Max* standardisation of the `Join_Count` attribute. The results are written to the `Nc_fc` table in the `Join_Count_MIN_MAX` attribute field. |
| Step 6: | The attribute table of the `Nc_fc` feature class is joined to the attribute table of the analytical grid feature class using the *Join* function. The relationship is built based on the temporary unique identifier field `stat_zone_field_ID`. |
| Step 7: | Before the next algorithm step, the script checks for the existence of the `Join_Count` and `Join_Count_MIN_MAX` fields in the analytical grid attribute table and removes them if present. Then, the values stored in the `Join_Count` and `Join_Count_MIN_MAX` attributes are copied to the analytical grid attribute table. |
| Step 8: | The copied attribute names are changed to `{prefix}_ANc` and `{prefix}_ANc_MM`, respectively, using the *Alter Field* function, which also assigns them clear, user-friendly aliases. |
| Step 9: | Removal of intermediate feature classes, the temporary `stat_zone_field_ID` field from the analytical grid table, and other temporary objects stored in operational memory from the geodatabase, followed by database compaction (optimization). |

## Flow Diagram



### Relevance to Geodiversity Assessment

- *A_Nc* is a basic, simple measure of categorical diversity, describing the number of unique types of polygonal landscape elements (e.g. lithological units or types of landforms) occurring within the statistical zones of the analytical grid.
- The index reflects the typological diversity of areal landscape features – its values increase with an increasing number of different polygon categories present within a given grid cell.
- Low *A_Nc* values indicate a limited number of categories, which may suggest geological, geomorphological, or pedological homogeneity.
- High *A_Nc* values indicate the presence of many polygon classes, which may result from complex geological structure, landscape mosaicity, or high diversity of geomorphological forms.
- The absence of diversity (*A_Nc* = 1) means that only one category of polygonal landscape elements occurs within a given grid cell – even if the cell is intersected by numerous polygons of the same class, the index remains low because it refers exclusively to the number of types, not to their area or geometry.
- Unlike areal measures (e.g. *A_SHDI*), *A_Nc* does not account for area proportions, degree of fragmentation, or spatial relationships. The index measures exclusively the qualitative component of diversity – the number of categories, not their size, proportions, or mutual spatial arrangement.
- Two statistical zones may obtain identical *A_Nc* values even if the grid cells differ significantly in terms of the areas of individual landscape classes – the index is sensitive only to the number of categories, not to their share, balance, or degree of dominance.
- Because *A_Nc* is an absolute measure, comparisons across different spatial scales require standardisation of the results.

# A_SHDI Shannon-Weaver Diversity Index

| | |
|---|---|
| Short Description: | The tool calculates the Shannon–Weaver diversity index of a selected polygon-based landscape feature within each polygon of the analytical grid. |
| Version: | 0.1.1 |
| Author: | Tomasz Bartuś |
| Date: | 2026-01-26 |

### Method Description

The Shannon-Weaver Diversity Index (*A_SHDI*) is a classical measure of compositional diversity used in the analysis of discrete regionalized variables expressed as polygon feature classes[75]. Often, it is simply referred to as *entropy*. The index quantifies the level of disorder and the degree of evenness in the area shares of polygon categories within the cells of the analytical grid. Its value depends on the proportional areas of individual categories, thus reflecting the structural complexity and heterogeneity of the landscape.

For each grid polygon, the tool calculates the proportional area of each category (9).

$$q_{ck} = \frac{Ac_k}{Ae_k} \tag{9}$$

Explanations: $q_{ck}$—the share (probability) of occurrence of category $c$ in the $k$-th elementary field of the analytical grid, with $\sum_{c=1}^{n} q_{ck} = 1$; $Ac_k$—the area of polygons of category *c* in the $k$-th elementary field of the analytical grid; $Ae_k$—the total area of all polygons in the $k$-th elementary field of the analytical grid.

The *A_SHDI* index is the negative sum of these proportions multiplied by their logarithm (10).

$$A\_SHDI_k = -\sum_{c=1}^{n} q_{ck} \times ln\, q_{ck} \tag{10}$$

Explanation: $A\_SHDI_k$—entropy of polygon elements in the $k$-th basic field of the analytical grid.

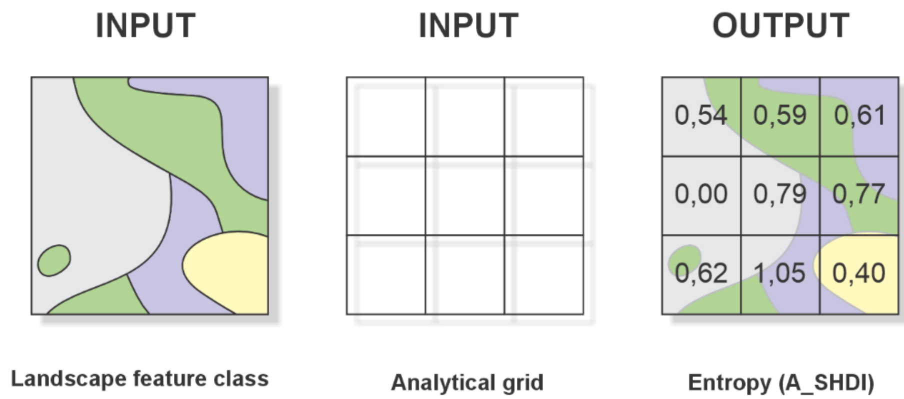
**Input data**

| Parameter | Type | Description |
|---|---|---|
| `landscape_fl` | Polygon feature layer | A polygon feature class representing the landscape feature to be analyzed, for which the *SHDI* is calculated within each polygon of the analytical grid. |
| `landscape_attr` | Field | The attribute field of the input landscape feature that defines its categories (classes), which are used to calculate the *SHDI* within each grid cell. |
| `grid_fl` | Polygon feature layer | A polygon feature class representing the analytical grid. For each grid cell (statistical zone), the partial geodiversity criterion is calculated based on the *SHDI* of the selected landscape feature. |
| `grid_id_field` | Field | An attribute of the analytical grid that uniquely identifies each grid cell (statistical zone) in which geodiversity indices are calculated. |

**Output data**

| Index | Type | Description |
|---|---|---|
| `{prefix}_ASHDI` | Output field | The Shannon–Weaver diversity index (*SHDI*) of the analyzed landscape feature within each polygon of the analytical grid. |
| `{prefix}_SHDIMM` | Output field | The standardized Shannon–Weaver diversity index (SHDI) of the analyzed landscape feature within each polygon of the analytical grid. |

## *Conceptual Diagram*

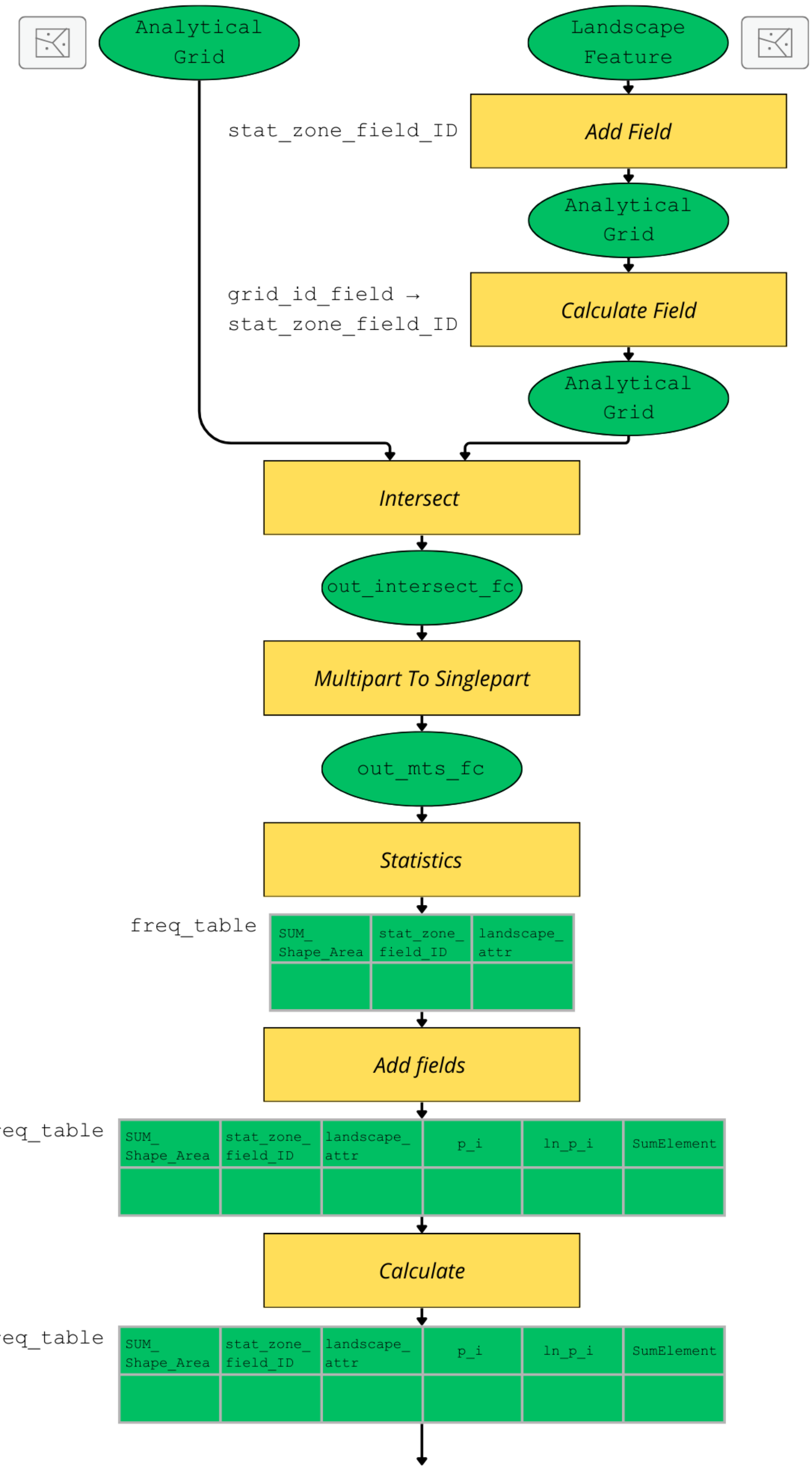


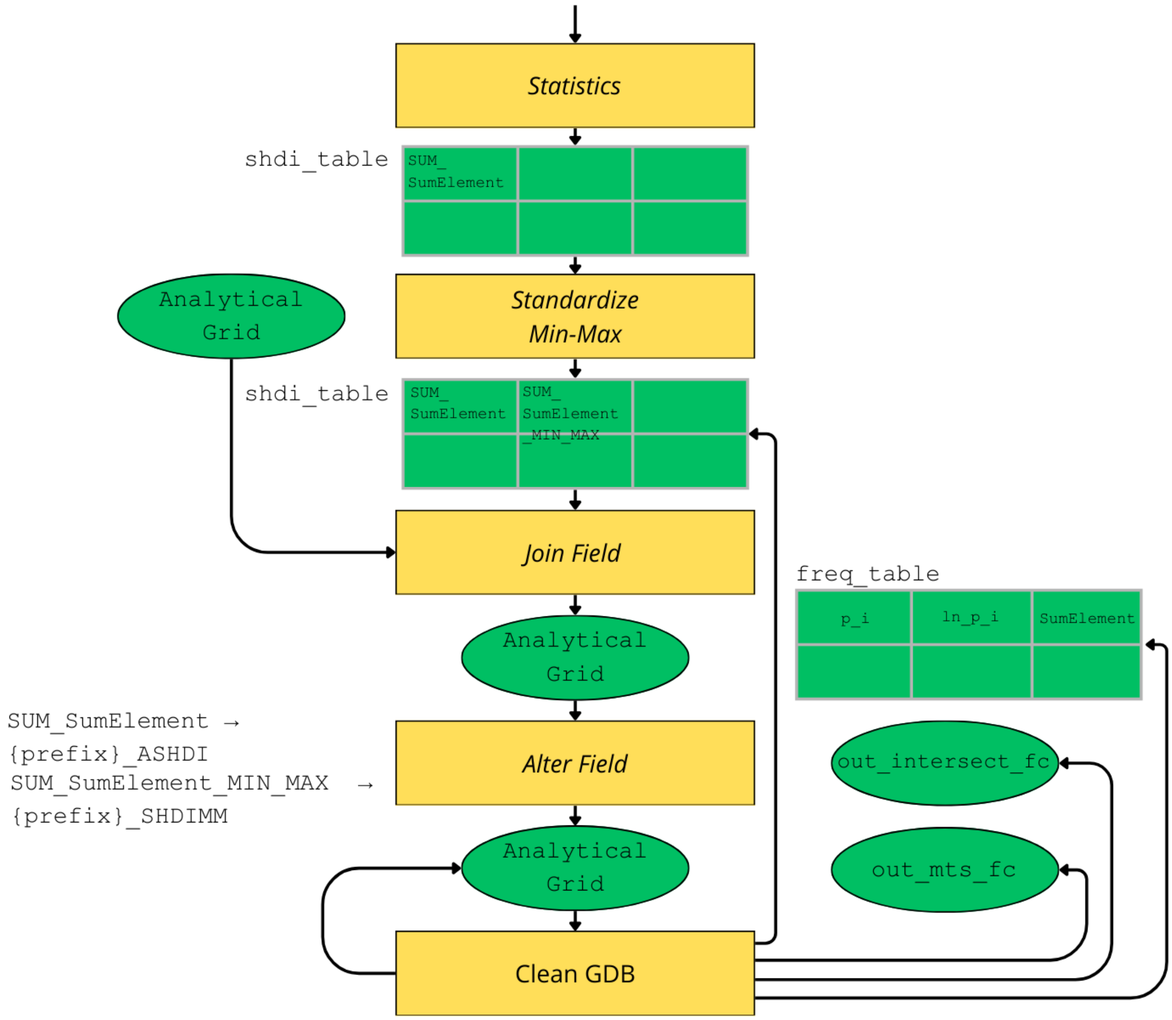

## *Algorithm*

| | |
|---|---|
| Step 1: | Retrieval of input parameters and initialization of variables; removal of data locks; verification of the existence of objects in the project geodatabase with names reserved for intermediate feature classes and non-spatial tables; and validation that the attribute table of the analytical grid feature class does not already contain fields with names intended for the output results. |
| Step 2: | Creation of an auxiliary statistical zone identifier by copying the primary key of the analytical grid. First, a field named `stat_zone_field_ID` of type `Long` is added. Next, using the *Calculate Field* function, the values of the primary key field (`grid_id_field`) are assigned to `stat_zone_field_ID`. This operation ensures the uniqueness of statistical zone identifiers during table join operations and prevents conflicts with system-defined fields. |
| Step 3: | Intersection of polygons representing the analysed landscape feature with polygons of the analytical grid. The *Intersect* function creates an intermediate feature class `out_intersect_fc`. Each fragment of a landscape feature polygon is assigned to the corresponding grid cell. In this process, the objects inherit the unique statistical zone identifier `stat_zone_field_ID`. |
| Step 4: | Conversion of the resulting multipart polygons into singlepart polygons using the *Multipart To Singlepart* function. An intermediate feature class `out_mts_fc` is created. This operation is essential for correctly recalculating the actual geometry of patch fragments after they have been intersected by the grid boundaries. |
| Step 5: | Calculation of area by category and grid unit. The *Statistics* function creates an intermediate non-spatial table `freq_table` containing the summed areas (`SUM_Shape_Area`) for each combination of grid cell identifier (`stat_zone_field_ID`) and landscape feature category (`landscape_attr`). |
| Step 6: | Addition of auxiliary fields to the table: `p_i` (area proportion of a category within a grid cell), `ln_p_i` (natural logarithm of the proportion), and `SumElement` (Shannon summation element: $-p_i \ln p_i$). These fields are defined as `FLOAT` type to preserve calculation precision. |
| Step 7: | For each record in the `freq_table`, the values of the fields `p_i`, `ln_p_i` (if `p_i > 0`), and `SumElement` (0 for `p_i = 0`) are calculated. The calculations are performed using an `Update Cursor`, supported by a Python dictionary |

| | |
|---|---|
| | (`total_area`) that aggregates the total area of all categories for each grid cell. |
| Step 8: | The *Statistics* function creates an intermediate non-spatial table `shdi_table` containing the field `SUM_SumElement`, representing the *SHDI* value for each grid unit. The grouping is performed based on the `stat_zone_field_ID` field. |
| Step 9: | Before performing the standardization, the script automatically removes any existing fields with the same names from the analytical grid's attribute table to prevent schema conflicts. Then, the *Standardize* function performs *Min–Max* standardization, creating the `SUM_SumElement_MIN_MAX` field. |
| Step 10: | The *Join Field* function joins the fields `SUM_SumElement` and `SUM_SumElement_MIN_MAX` to the attribute table of the analytical grid. The relationship is established based on the `stat_zone_field_ID` field. |
| Step 11: | The *Alter Field* function renames the resulting fields in the analytical grid attribute table to `{prefix}_ASHDI` and `{prefix}_SHDIMM`, respectively, and assigns them readable aliases. |
| Step 11: | Removal of intermediate feature classes and tables (`out_intersect_fc`, `out_mts_fc`, `freq_table`, `shdi_table`), deletion of the temporary `stat_zone_field_ID` field from the analytical grid attribute table, followed by cache cleaning and geodatabase compaction. |

## *Flow Diagram*






## Relevance to Geodiversity Assessment

- $A\_SHDI$ measures the compositional diversity of the landscape based on the area proportions of polygon feature categories.
- The values of $A\_SHDI$ depend on the number of categories ($Nc_k$), their area proportions, and the degree of evenness, but do not depend on the absolute size of the basic field, as they are calculated from the sum of relative proportions.
- The range of $A\_SHDI$ values extends from 0 to ln($Nc_k$).
- The index increases with greater evenness of category area proportions and reaches its maximum when all classes have similar shares.
- High $A\_SHDI$ values indicate high landscape heterogeneity resulting from the presence of many categories with comparable area proportions. This often reflects a complex geomorphological, lithological, or land-use structure.
- Low $A\_SHDI$ values indicate a small number of categories, dominance of a single category, or strong disproportions in area between categories, which within a basic field point to a more homogeneous spatial structure.
- Homogeneity ($A\_SHDI$ = 0) is defined as the occurrence of only one polygon category within a field ($q_c$ = 1).

- Two fields with identical category area proportions ($q_c$) have the same $A\_SHDI$ values, regardless of the absolute areas of these categories. As long as the proportions remain unchanged, increasing or decreasing the total field size does not affect the index value.
- For comparisons between different categorical variables, different analytical grids, or regions, the use of standardization (e.g. the *Min–Max* method) is recommended.

# Section 6 Raster Based Metrics

Geodiversity Tools provides three tools designed for landscape features represented by raster datasets: Raster Standard Deviation (*R_SD*), Circular Standard Deviation (*R_SDc*) and Steinhaus Vertical Relief Index (*R_M*).

## R_SD Raster Standard Deviation

| | |
|---|---|
| Short Description: | The tool calculates, for a selected landscape feature (raster dataset), the standard deviation of its pixel values within each polygon of the analytical grid. |
| Version: | 0.1.1 |
| Author: | Tomasz Bartuś |
| Date: | 2026-01-26 |

### Method Description

Raster Standard Deviation (*R_SD*) is the most important measure used to assess the diversity of landscape features represented by continuous regionalized variables. The index defines the degree of variability of pixel values in raster images[15]. Calculations are performed in accordance with the classical definition of the population standard deviation (11). It measures the mean deviation of raster pixel values from the local arithmetic mean. In order to aggregate deviations both below and above the mean, the differences $(x_i - \bar{x})$are squared, and to express the final results in the same units as the analysed feature, they are subsequently square-rooted. This facilitates interpretation and comparison between different cells of the analytical grid. Calculations are performed separately for each basic cell of the analytical grid.

*R_SD* is the most universal diversity measure for continuous data based on statistical measures of variability. Unlike the *Range* statistic (*min–max*), *R_SD* accounts for all pixel values rather than only extreme values, and therefore better reflects actual spatial heterogeneity. In contrast to *Variance*, *R_SD* is expressed in the same units as the analysed feature (e.g. metres, percentages), which significantly facilitates geographical interpretation. Compared with the *Coefficient of Variation* (*CV*), *R_SD* does not require preliminary assumptions and ensures interpretative correctness. Although *CV* has the advantage of being a relative measure (unitless results), its application is valid only when a positive relationship exists between means and standard deviations in the analysed population. As demonstrated by Mastej & Bartuś[15], the use of *CV* in diversity analyses when this condition is not met leads to serious distortions of results, making correct assessment impossible.

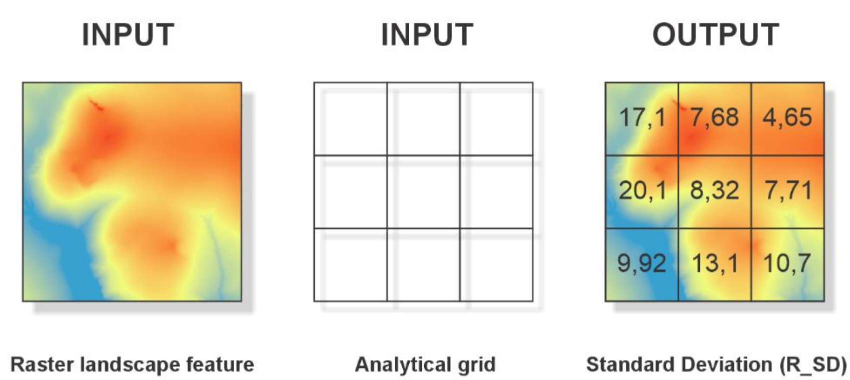
$$SD = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n}} \qquad (11)$$

Explanations: $n$—the number of pixels within a statistical zone of the analytical grid; $x_i$—the value of the landscape feature represented by the $i$-th pixel of the analysed basic grid cell; $\bar{x}$—the mean value of the landscape feature within the analysed basic grid cell.

| *Input data* | | |
|---|---|---|
| **Parameter** | **Type** | **Description** |
| `landscape_fl` | Input raster dataset | A raster dataset representing the landscape feature to be analyzed. The standard deviation of its pixel values is calculated within each polygon of the analytical grid. |
| `grid_fl` | Polygon feature layer | A polygon feature class representing the analytical grid. For each grid cell (statistical zone), the partial geodiversity criterion is calculated based on the standard deviation of the selected raster landscape dataset. |
| `grid_id_field` | Field | An attribute of the analytical grid that uniquely identifies each grid cell (statistical zone) in which geodiversity indices are calculated. |

| **Output data** | | |
|---|---|---|
| **Index** | **Type** | **Description** |
| `{prefix}_RSD` | Output field | The standard deviation of the analyzed landscape feature within each polygon of the analytical grid. |
| `{prefix}_RSD_MM` | Output field | The standardized standard deviation of the analyzed landscape feature within each polygon of the analytical grid. |

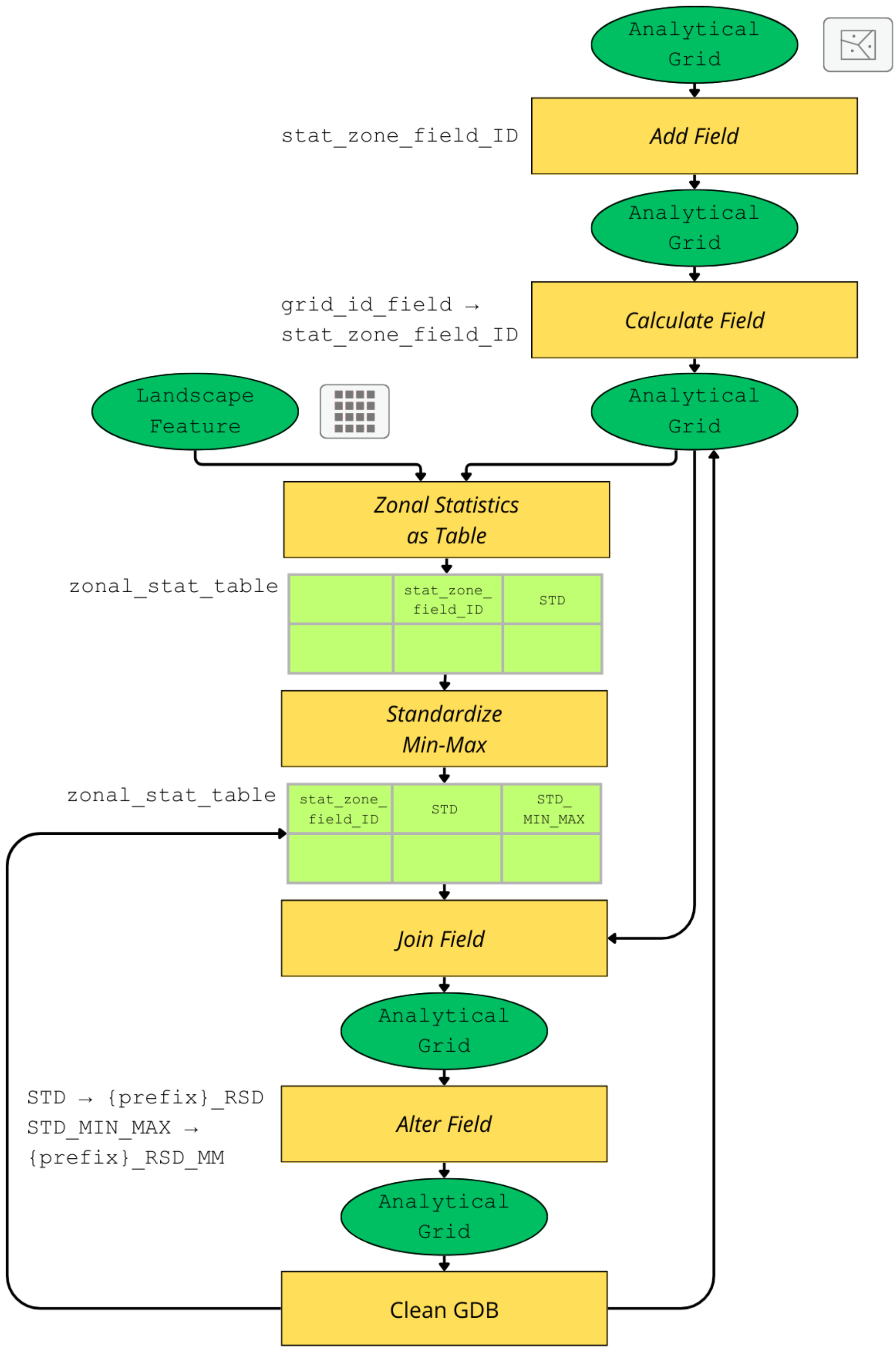
## *Conceptual Diagram*




| *Algorithm* | |
|---|---|
| Step 1: | Retrieval of input parameters and initialization of variables; removal of data locks; verification of the existence of objects in the project geodatabase with names reserved for intermediate feature classes and non-spatial tables; and validation that the attribute table of the analytical grid feature class does not already contain fields with names intended for the output results. |

| | |
|---|---|
| Step 2: | Creation of an auxiliary statistical zone identifier by copying the primary key of the analytical grid. First, a field named `stat_zone_field_ID` of type `Long` is added. Next, using the *Calculate Field* function, the values of the primary key field (`grid_id_field`) are assigned to `stat_zone_field_ID`. This operation ensures the uniqueness of statistical zone identifiers during table join operations and prevents conflicts with system-defined fields. |
| Step 3: | The *Zonal Statistics as Table* function calculates the standard deviation (`STD`) of raster pixel values of the landscape feature (`landscape_ras`) within the boundaries of each analytical grid polygon. The result is written to an intermediate non-spatial table `zonal_stat_table`, in which each record corresponds to one statistical zone (`stat_zone_field_ID`) and contains the field `STD`. This table is created in the system memory (RAM). |
| Step 4: | The *Standardize Field* function performs *Min–Max* standardization of the `STD` field, creating a new field `STD_MIN_MAX` with values rescaled to the range ⟨0; 1⟩. If it is detected that the `STD` values are identical across all statistical zones (no diversity), the script automatically assigns a value of 0 to all records, preventing division-by-zero errors. |
| Step 5: | Before the next join operation, the script checks for and removes any existing result fields named `STD` and `STD_MIN_MAX` in the attribute table of the analytical grid. The Join Field function then copies both result values – `STD` and `STD_MIN_MAX` – to the grid table. Each statistical zone thus receives two indicators: the “raw” and the rescaled one. |
| Step 6: | The *Alter Field* function renames the `STD` and `STD_MIN_MAX` fields to the final indicator names, respectively: `{prefix}_RSD` and `{prefix}_RSD_MM`. Additionally, appropriate aliases are assigned to the fields. |
| Step 7: | Removal of the unnecessary field `stat_zone_field_ID` from the analytical grid table and the `zonal_stat_table`, clearing the geoprocessing cache, and compacting the geodatabase. |

## *Flow Diagram*



### Relevance to Geodiversity Assessment

- *R_SD* measures the variability of continuous raster values within a basic cell of the analytical grid by quantifying the degree to which individual pixels deviate from the local arithmetic mean. It is the fundamental and most widely used measure of variability for continuous spatial data.
- *R_SD* values depend exclusively on the dispersion of pixel values, not on their absolute magnitudes. Two cells with identical distributions of relative values (deviations from the mean), but different absolute levels, may have the same *R_SD*.
- High *R_SD* values indicate strong internal variability of the feature (e.g., large elevation differences).
- Low *R_SD* values indicate internal homogeneity of the cell – minimal differences between pixels and a lack of local variability.
- `R_SD = 0` denotes absolute homogeneity of the cell, in which all pixels have identical values.
- *R_SD* values do not depend on the number of pixels within a basic cell, provided that the distribution of their values remains the same. Increasing raster resolution (a larger number of pixels) does not affect the result if the structure of variability is unchanged.
- *R_SD* is sensitive to outliers, which makes it an indicator that emphasizes local anomalies.
- The statistic requires that the number of pixels within a basic cell is not less than 30.
- *R_SD* is expressed in the units of the analyzed feature (e.g., meters, millimeters, percentages), which facilitates interpretation and comparison between cells.
- *R_SD* is not suitable for assessing the diversity of circular data. As reported by Mastej & Bartuś[15], there is a certain exception to this rule. If the classical standard deviation of an angular variable is less than 25°, the standard (linear) standardized deviation may be used instead of the circular one. In such cases, the maximum relative error of partial diversity estimation does not exceed 5%.
- *R_SD* does not require assumptions of normality. It is scalable and, after standardization (e.g., using the *Min–Max* method), can be applied in criterion-based or area-based comparisons.

# R_SDc Circular Standard Deviation

| | |
|---|---|
| Short Description: | The tool calculates the circular standard deviation of a selected landscape feature (raster layer) within each polygon of the analytical grid. |
| Version: | 0.1.1 |
| Author: | Tomasz Bartuś |
| Date: | 2026-01-26 |

### Method Description

Circular Standard Deviation (*R_SDc*) is a specialized measure of variability for landscape features described as continuous angular variables, i.e. phenomena of a periodic nature represented on a circle. This applies, among others, to attributes such as aspect, fault orientations, or palaeotransport directions in basins, where values of 0° and 360° denote the same direction. In such cases, classical statistical methods based on real-number arithmetic are inadequate, as they do not account for the “wrapping” of the angular scale.

$R\_SDc$ is the circular counterpart of the classical standard deviation ($R\_SD$) and is used to quantify the degree of directional variability of the analysed landscape feature within each cell of the analytical grid.

Its computation is based on the principles of circular statistics. The elementary angular values represented by raster pixels can be conceptualized as a set of unit vectors (of length equal to 1), anchored at the origin of the coordinate system and inclined at an angle $\theta_i$ relative to the north direction.

A B

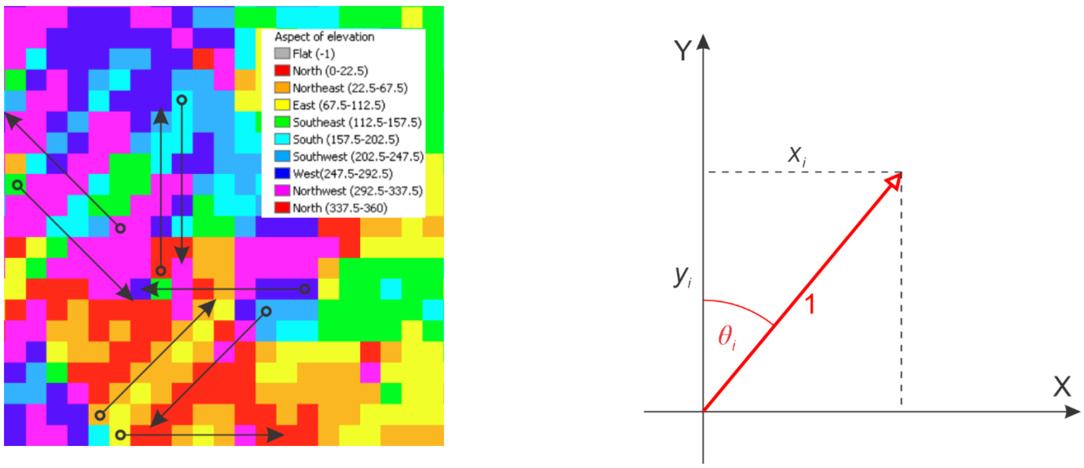



Figure 1. Fragment of the raster dataset of the *Aspect* landscape feature, with unit vectors indicating the slope aspect directions of selected pixels (A), and a unit vector of an example pixel inclined to the OY axis (north direction) by an angle $\theta_i$ (B).

The angular values are first transformed into sine (12) and cosine (13) components, which makes it possible to geometrically average the directions.

$$sin\varphi_i = \frac{x_i}{\sqrt{x_i^2 + y_i^2}} = \frac{x_i}{1} = x_i \quad (12)$$

$$cos\varphi_i = \frac{y_i}{\sqrt{x_i^2 + y_i^2}} = \frac{y_i}{1} = y_i \quad (13)$$

Explanations: $\varphi_i$—slope angle of the unit vector of the $i$-th pixel of the raster image; $x_i$—sine component of the unit vector corresponding to the angular value of the $i$-th pixel; $y_i$—cosine component of the unit vector corresponding to the angular value of the $i$-th pixel.

Next, the so-called circular resultant vector $\vec{R}$ is calculated as the sum of all unit vectors (14, 15).

$$x_r = \sum_{i=1}^{n} \sin(\theta_i) \quad (14)$$

$$y_r = \sum_{i=1}^{n} \cos(\theta_i) \quad (15)$$

Explanations: $x_r$—sine component of the resultant vector; $y_r$—cosine component of the resultant vector; $n$—number of pixels in the statistical zone.

The length of the vector $\vec{R}$ can be calculated using the Pythagorean theorem (16)

$$R = \sqrt{x_r^2 + y_r^2} \tag{16}$$

The mean (resultant) direction of the resultant vector – $\bar{\theta}_c$– is then the direction of the hypotenuse of the triangle formed by $x_r$ and $y_r$ [15,76–78] and is calculated using formula (17).

$$\bar{\theta}_c = \begin{bmatrix} arctg\dfrac{x_r}{y_r} \;\; for\; x_r > 0 \; and \;\; y_r > 0 \\ arctg\dfrac{x_r}{y_r} + 180° \; for \;\; y_r < 0 \\ arctg\dfrac{x_r}{y_r} + 360° \; for \; x_r < 0 \; and \;\; y_r > 0 \end{bmatrix} \tag{17}$$

The length of the vector $\vec{R}$ reflects the degree of directional concentration.

The case of maximum concentration of unit vectors around the mean angular direction $\bar{\theta}_c$ (no variability) occurs when the angles of all unit vectors are identical (all vectors are parallel to each other). In this case, $R$ would equal n (since these are unit vectors), and the concentration of the resultant vector $\bar{R}$ would be 1 [79].

Conversely, in the case of minimal concentration, when all vectors represent maximally different angles, the mean angle $\bar{\theta}_c$ is undefined, and $\bar{R} = 0$. Therefore, the magnitude of $\bar{R}$ expresses the concentration – the longer the vector $\vec{R}$ (the greater $\bar{R}$), the higher the concentration of unit vectors around the resultant direction $\bar{\theta}_c$.

In practice, the concentration of the resultant vector $\bar{R}$ is calculated by dividing its length $R$ by the number of pixels in the individual base fields of the analytical grid (18).

$$\bar{R} = \frac{R}{n} \tag{18}$$

This makes the result independent of the sometimes varying number of pixels in the base fields and standardizes it, as $\bar{R}$ ranges between 0 and 1

Angular variance is a quantity inversely related to concentration. It is calculated using the formula: $SD_c^2 = 1 - \bar{R}$. The circular standard deviation $SDc$ is calculated, among others, according to Batschelet [76] (19).

$$SD_c = \sqrt{2}\sqrt{1 - \bar{R}} = \sqrt{2(1 - \bar{R})} \tag{19}$$

If the angles are expressed in degrees, an additional factor of $\frac{180°}{\pi}$ is required in the formulas. The final form of the formula is presented in (20).

$$SD_c = \frac{180°}{\pi}\sqrt{2(1 - \bar{R})} = \frac{180°}{\pi}\sqrt{2\left(1 - \frac{\sqrt{[\sum_{i=1}^{n}\sin(\theta_i)]^2 + [\sum_{i=1}^{n}\cos(\theta_i)]^2}}{n}\right)} \tag{20}$$

The resulting measure $R\_SDc$ is calculated based on the classical definition of circular standard deviation, whose interpretation is analogous to variance and standard deviation known from linear data, but adapted to the nature of angular variables. High values of $R\_SDc$ indicate high directional heterogeneity (strong dispersion of raster pixel orientations), whereas low values indicate high directional agreement and uniformity.

Similar to *R_SD*, this index is calculated for each base field of the analytical grid, providing a map-based representation of directional variability in regionalized raster data. *R_SDc* allows for correct interpretation and comparison of landscape features of a periodic nature, avoiding errors that would result from treating angular values as ordinary real numbers.

***Input data***

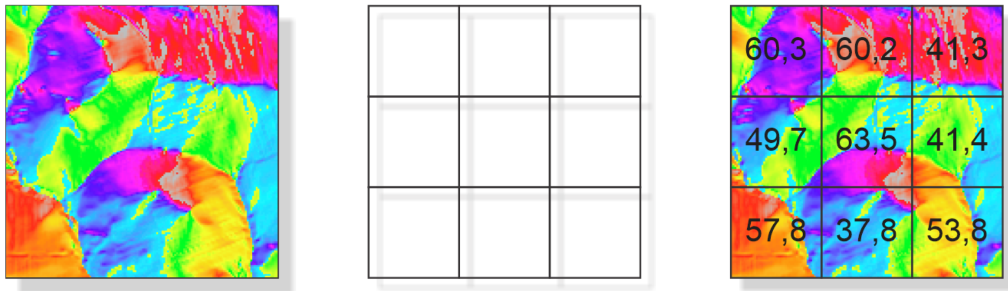
| Parameter | Type | Description |
|---|---|---|
| `landscape_raster` | Raster layer | A raster layer representing the landscape feature to be analyzed. Its circular standard deviation (*SDc*) is calculated within each polygon of the analytical grid. |
| `grid_fl` | Polygon feature layer | A polygon feature class representing the analytical grid. For each grid cell (statistical zone), the partial geodiversity criterion is calculated based on the *SDc* of the selected landscape feature. |
| `grid_id_field` | Field | An attribute of the analytical grid that uniquely identifies each grid cell (statistical zone) in which geodiversity indices are calculated. |

**Output data**

| Index | Type | Description |
|---|---|---|
| `{prefix}_RSDc` | Output field | The circular standard deviation (*SDc*) of the analyzed landscape feature within each polygon of the analytical grid. |
| `{prefix}_RSDcMM` | Output field | The standardized circular standard deviation (*SDc*) of the analyzed landscape feature within each polygon of the analytical grid. |

## *Conceptual Diagram*

**INPUT** | **INPUT** | **OUTPUT**



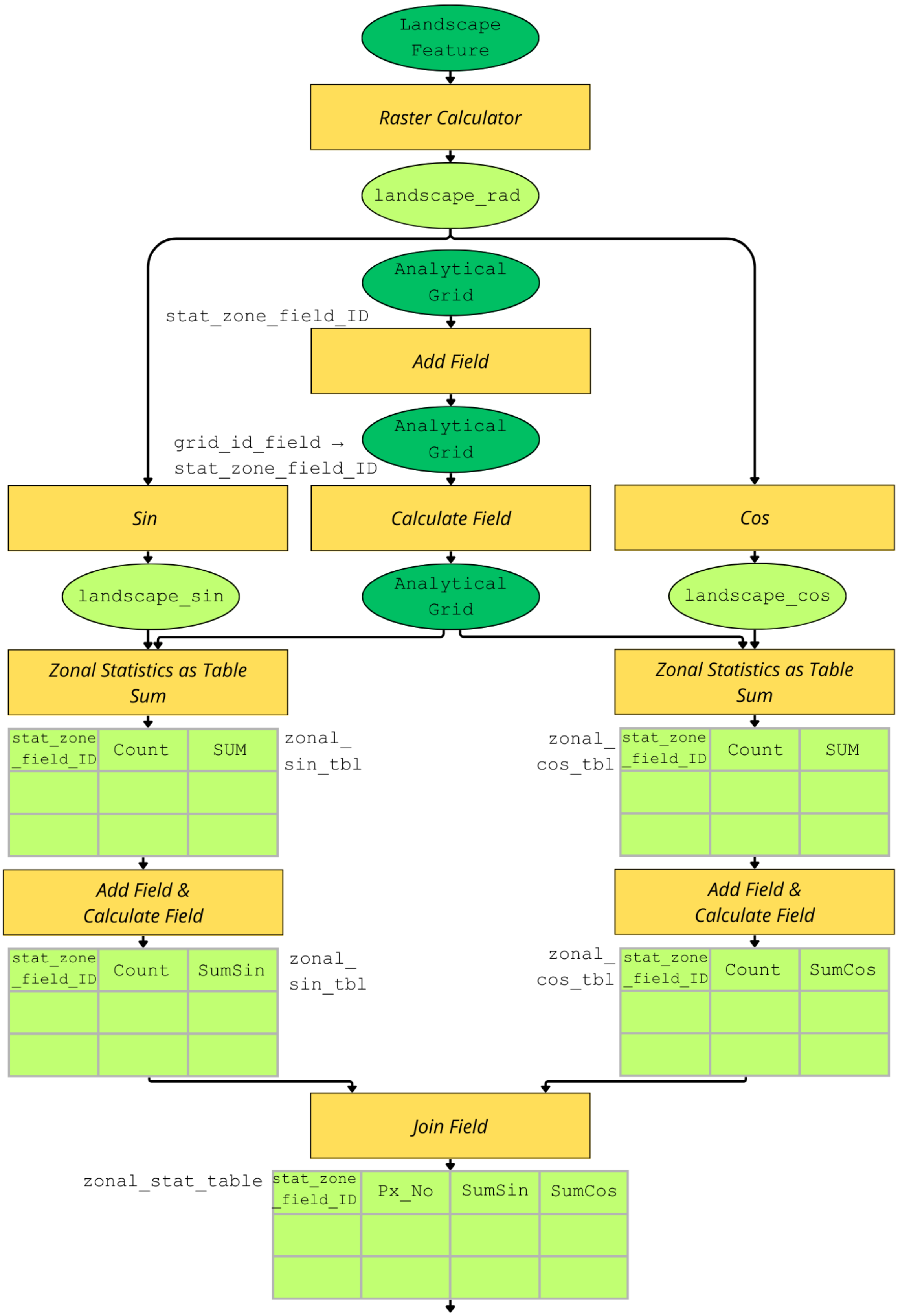

**Raster landscape feature** | **Analytical grid** | **Circular Standard Deviation (R_SDc)**

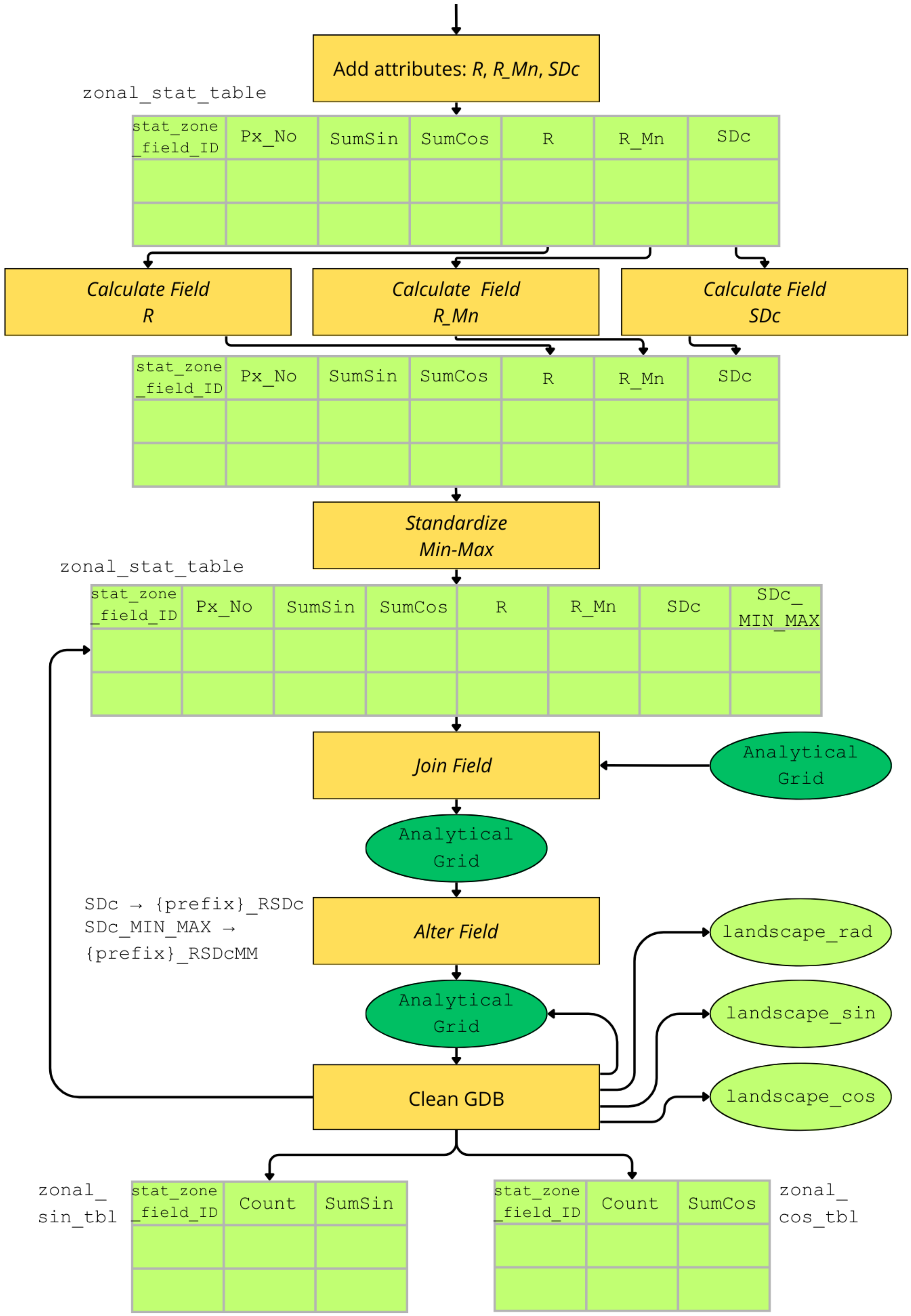
## *Algorithm*

| | |
|---|---|
| Step 1: | Retrieval of input parameters and initialization of variables; removal of data locks; verification of the existence of objects in the project geodatabase with names reserved for intermediate feature classes and non-spatial tables; and validation that the attribute table of the analytical grid feature class does not already contain fields with names intended for the output results. |
| Step 2: | Creation of an auxiliary statistical zone identifier by copying the primary key of the analytical grid. First, a field named `stat_zone_field_ID` of type `Long` is added. Next, using the *Calculate Field* function, the values of the primary key field (`grid_id_field`) are assigned to `stat_zone_field_ID`. This operation ensures the uniqueness of statistical zone identifiers during table join operations and prevents conflicts with system-defined fields. |
| Step 3: | ArcGIS Pro requires trigonometric operations to be performed in radians; therefore, the angular values of the selected landscape feature are converted from degrees to radians. A raster `landscape_rad` is created, in which the value of each pixel is calculated as: `landscape_ras / (180 / math.pi)`. For performance optimization, this raster is stored in the system memory (RAM). |
| Step 4: | Based on the raster in radians, two intermediate raster datasets are created using the *Sin* and *Cos* functions: one containing the sine values for each pixel – `landscape_sin`, and one containing the cosine values for each pixel – `landscape_cos`. |
| Step 5: | Using two calls to the *Zonal Statistics As Table* function, two non-spatial tables are generated: `zonal_sin_tbl` and `zonal_cos_tbl`. In the first table, the sums of sines (attribute `SumSin`) are calculated for each statistical zone of the analytical grid `grid_fl`, and in the second table, the sums of cosines (attribute `SumCos`). Additionally, the number of raster pixels covered by the statistical zones of the analytical grid is calculated (`Count` field). |
| Step 6: | Based on the attribute table of the analytical grid (`grid_fl`), an empty non-spatial table `zonal_stat_table` is generated, containing only the statistical |

| | |
|---|---|
| | zone identifier (`stat_zone_field_ID`). In the following steps, all numerical data required for calculating *SDc* will be collected in this table. This operation is performed using the *Copy Rows* function on one of the source tables. |
| Step 7: | Using the *Join Field* function, the tables `zonal_sin_tbl` and `zonal_cos_tbl` are joined to `zonal_stat_table`, and the attributes `SumSin`, `SumCos`, and `Count` are copied from them. |
| Step 8: | In the table `zonal_stat_table`, the attribute `Count` (number of the raster cells within a given grid field) is renamed to `Px_No`. |
| Step 9: | Attributes are added to `zonal_stat_table`: `R` – length of the resultant vector, `R_Mn` – concentration of the resultant vector, and `SDc` – circular standard deviation. |
| Step 10: | Calculation of the resultant vector length `R` (see 16). Calculations are performed ensuring conversion of the fields `SumSin` and `SumCos` to native Python types (`float64`) to avoid geodatabase data type errors and handle any possible null values (`None`). |
| Step 11: | Calculation of the concentration of the resultant vector `R_Mn` (see 18). |
| Step 12: | Calculation of the circular standard deviation `SDc` (see 19 and 20). The result is expressed in degrees. |
| Step 13: | The *Standardize* function (*Min-Max* method) standardizes the calculation results. A field `SDc_MIN_MAX` is created in `zonal_stat_table`. Due to in-memory processing, the script uses a manual procedure to calculate extremes and perform standardization. |
| Step 14: | Before performing the join, the script automatically checks for and removes any existing result fields named `SDc` and `SDc_MIN_MAX` in the attribute table of the analytical grid. Using the *Join Field* function, zonal_stat_table is then joined with the grid's attribute table, and the fields `SDc` and `SDc_MIN_MAX` are appended and copied. |
| Step 15: | Field names are changed to match the tool's naming convention: `SDc` → `{prefix}_RSDc` and `SDc_MIN_MAX` → `{prefix}_RSDcMM`, using the *Alter Field* function. Appropriate aliases are also assigned to the fields. |
| Step 16: | Auxiliary rasters (`landscape_rad`, `landscape_sin`, `landscape_cos`), non-spatial tables (`zonal_sin_tbl`, `zonal_cos_tbl`, `zonal_stat_table` and `stats_table`), and the temporary field `stat_zone_field_ID` from the analytical grid table are deleted, and the geodatabase is optimized using the *Compact* function. |

## *Flow Diagram*




Add attributes: R, R_Mn, SDc
zonal_stat_table
stat_zone_field_ID
Px_No
SumSin
SumCos
R
R_Mn
SDc
Calculate Field R
Calculate Field R_Mn
Calculate Field SDc
stat_zone_field_ID
Px_No
SumSin
SumCos
R
R_Mn
SDc
Standardize Min-Max
zonal_stat_table
stat_zone_field_ID
Px_No
SumSin
SumCos
R
R_Mn
SDc
SDc_MIN_MAX
Join Field
Analytical Grid
Analytical Grid
SDc → {prefix}_RSDc
SDc_MIN_MAX → {prefix}_RSDcMM
Alter Field
landscape_rad
Analytical Grid
landscape_sin
Clean GDB
landscape_cos
zonal_sin_tbl
stat_zone_field_ID
Count
SumSin
stat_zone_field_ID
Count
SumCos
zonal_cos_tbl

### *Relevance to Geodiversity Assessment*

- *R_SDc* measures the degree of variation in angular (circular) values of regionalized continuous variables within a base field of the analytical grid. It indicates how strongly directions (e.g., aspect, flow, orientation of linear structures) are dispersed around the dominant orientation defined by the resultant vector.
- The index is based on trigonometric analysis (sums of sines and cosines), which ensures correct handling of directional data, where values of 0° and 360° are identical and may create apparent discontinuities.
- *R_SDc* depends solely on angular dispersion, not on absolute values. Two sets of directions can have the same *R_SDc* even if their circular means differ, provided the level of dispersion around that mean is identical.
- High *R_SDc* values indicate high directional variability within a statistical zone, i.e., lack of a dominant orientation (e.g., diverse slope aspects, complex arrangements of geological structures).
- Low *R_SDc* values indicate directional uniformity, i.e., strong dominance of a single direction (e.g., uniform aspect, regular orientation of structures).
- *R_SDc* = 0 represents absolute angular uniformity, meaning all pixels have identical direction.
- The value of *R_SDc* does not depend on the number of pixels, as long as the distribution of directions remains unchanged. Higher raster resolution (more pixels) does not affect the result if the directional structure is the same.
- *R_SDc* is sensitive to outliers, especially rare but different directions, which can significantly increase the dispersion level.
- Circular statistics methods recommend that the number of angular observations should be at least 30, which increases the stability of the estimation of the resultant direction and dependent indices (including *SDc*).
- *R_SDc* is expressed in degrees, facilitating interpretation in the context of spatial data.
- Unlike classical standard deviation, *R_SDc* is the only correct measure of variability for directional data that “wraps around” at a full 360° rotation.
- The index does not require assumptions of normality, as circular statistics are based on geometric and trigonometric relationships rather than classical variance theory.
- After standardization (e.g., *Min–Max* method), *R_SDc* can be directly used in comparative analyses, assessment of complex geodiversity, and multi-criteria analyses.

# R_M Steinhaus Vertical Relief Index

| | |
|---|---|
| Short Description: | The tool calculates, for a selected landscape feature (raster layer), the Steinhaus Vertical Relief Index within each polygon of an analytical grid. |
| Version: | 0.1.0 |
| Author: | Tomasz Bartuś |
| Date: | 2026-01-28 |

## *Method Description*

The Steinhaus vertical relief index[80] ($R\_M$) is a metric designed for the analysis of continuous regionalized variables. It describes the degree of morphological variability of a studied surface by assessing the variation of the analyzed attribute along vertical profiles. In geomorphometric studies, this index is most commonly used to analyze terrain morphology based on elevation data.

The variability of the studied surface along a profile line can be represented as a function (21).

$$z = f(x) \tag{21}$$

Explanations: $x$—horizontal distance from the beginning of the profile; $z$—value of the analyzed attribute (e.g., elevation).

The course of this function is irregular and consists of a sequence of local minima and maxima ($z_0$, $z_1$,…, $z_n$), which in mathematics are referred to as function fluctuations. The total amplitude of these fluctuations can be expressed as the sum of the absolute differences between consecutive extrema (22).

$$\sum_{i=0}^{n} |\Delta z_i| \tag{22}$$

Steinhaus proposed a measure of the intensity of these fluctuations in the form of the $\mu$ index, defined as the sum of the square roots of the absolute differences in elevation between consecutive local minima and maxima (23).

$$\mu = \sqrt{|z_1 - z_0|} + \sqrt{|z_2 - z_1|} + \ldots + \sqrt{|z_n - z_{n-1}|} = \sum_{i=1}^{n} \sqrt{|\Delta z_i|} \tag{23}$$

Explanations: $z_0$–$z_n$—the values of consecutive local minima and maxima of the analyzed parameter.

In computational practice, to obtain a representative measure of terrain morphology for an elementary area, the μ index is determined along several vertical profiles. In the $R\_M$ tool, the analysis is performed separately for each cell of the analytical grid. Through the centroid of each cell, two orthogonal vertical profiles are drawn: one along the north–south (N–S) axis and the other along the west–east (W–E) axis.

Along each profile, the local minima and maxima of the $Z$ attribute are identified. Additionally, to maintain the correct definition of the profile within the boundaries of the cell, values at the intersections of the profile line with the cell boundary are also included in the set of profile points. Based on this, separate values of the μ index, $\mu_{NS}$ and $\mu_{WE}$, are calculated.

Finally, for each cell, the arithmetic mean of these values is determined (24).

$$\bar{\mu} = \frac{\mu_{NS} + \mu_{WE}}{2} \tag{24}$$

The final form of Steinhaus' vertical relief index, denoted as $M$, includes normalization with respect to the area of the analyzed cell (25).

$$M = \frac{\bar{\mu}}{\sqrt{A}} \; [-] \tag{25}$$

Explanations: $M$—Steinhaus vertical relief index; $\bar{\mu}$—mean sum of the square roots of elevation differences between consecutive extrema; $A$—area of the analyzed cell.

Normalization by the square root of the area allows comparison of the index values between cells of different sizes and ensures its spatial scalability.

As reported by Wieczorek & Żyszkowska[81], studies by Szczepankiewicz[82], and Gadzojannis & Plewniak[83] have shown that the Steinhaus index effectively reflects terrain variability.

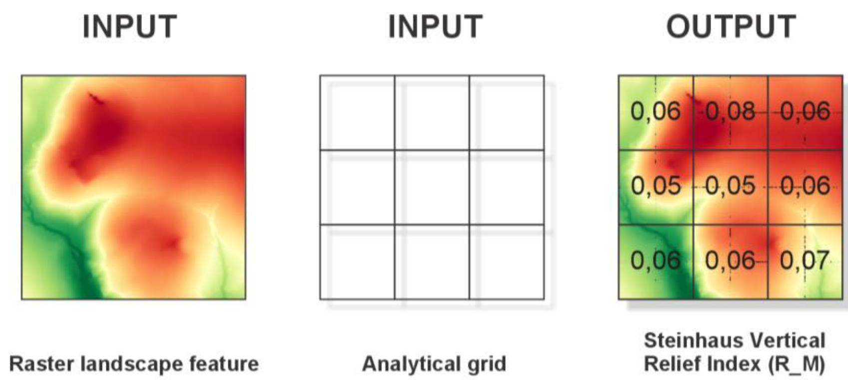
***Input data***

| Parameter | Type | Description |
|---|---|---|
| `landscape_raster` | Input raster dataset | A raster layer representing the landscape feature to be analyzed. Its Steinhaus vertical relief index ($R\_M$) will be counted within each cell of the analytical grid. |
| `grid_fl` | Polygon feature layer | A polygon feature class representing the analytical grid. For each grid cell (statistical zone), the partial geodiversity criterion is calculated based on the $R\_M$ of the selected landscape feature. |
| `grid_id_field` | Field | An attribute of the analytical grid that uniquely identifies each grid cell (statistical zone) in which geodiversity indices will be calculated. |

**Output data**

| Index | Type | Description |
|---|---|---|
| `{prefix}_RM` | Output field | Vertical relief index of the analyzed landscape feature within each polygon of the analytical grid. |
| `{prefix}_RM_MM` | Output field | Standardized vertical relief index of the analyzed landscape feature within each polygon of the analytical grid. |

## Conceptual Diagram




## Algorithm

**Step 1:** Retrieve the tool input parameters (landscape feature raster, analytical grid polygon feature class, and identifier field), initialize the workspace and project metadata. The script verifies the data coordinate system, sets the overwrite mode, and disables inheritance of `Z` and `M` coordinates in the created feature classes.

**Step 2:** Validate the input data, including:

- checking that the analytical grid is not a shapefile (the tool requires geodatabase feature classes),
- recalculating the spatial extent of the analytical grid,
- verifying that the raster and the analytical grid overlap.

If these conditions are not met, the calculation is terminated.

**Step 3:** Check for conflicts in the names of output fields in the analytical grid attribute table. The script ensures that fields designated for the $R\_M$ index and its standardized form do not already exist in the data, preventing overwriting of results.

**Step 4:** Remove any locks from the landscape feature raster and the analytical grid feature class to allow safe data modifications in subsequent steps.

**Step 5:** Create a temporary statistical zones identifier by adding a `StatZoneID` field to the analytical grid attribute table and copying the primary key values into it. This ensures unique identification of grid cells for table and spatial operations.

**Step 6:** Create an in-memory helper table, `centroids_table`, containing the coordinates of the centroids of the analytical grid cells. For each cell, the `X` and `Y` centroid coordinates are stored along with the `StatZoneID`.

**Step 7:** Based on the raster extent and centroid locations, generate a global set of elevation profiles:

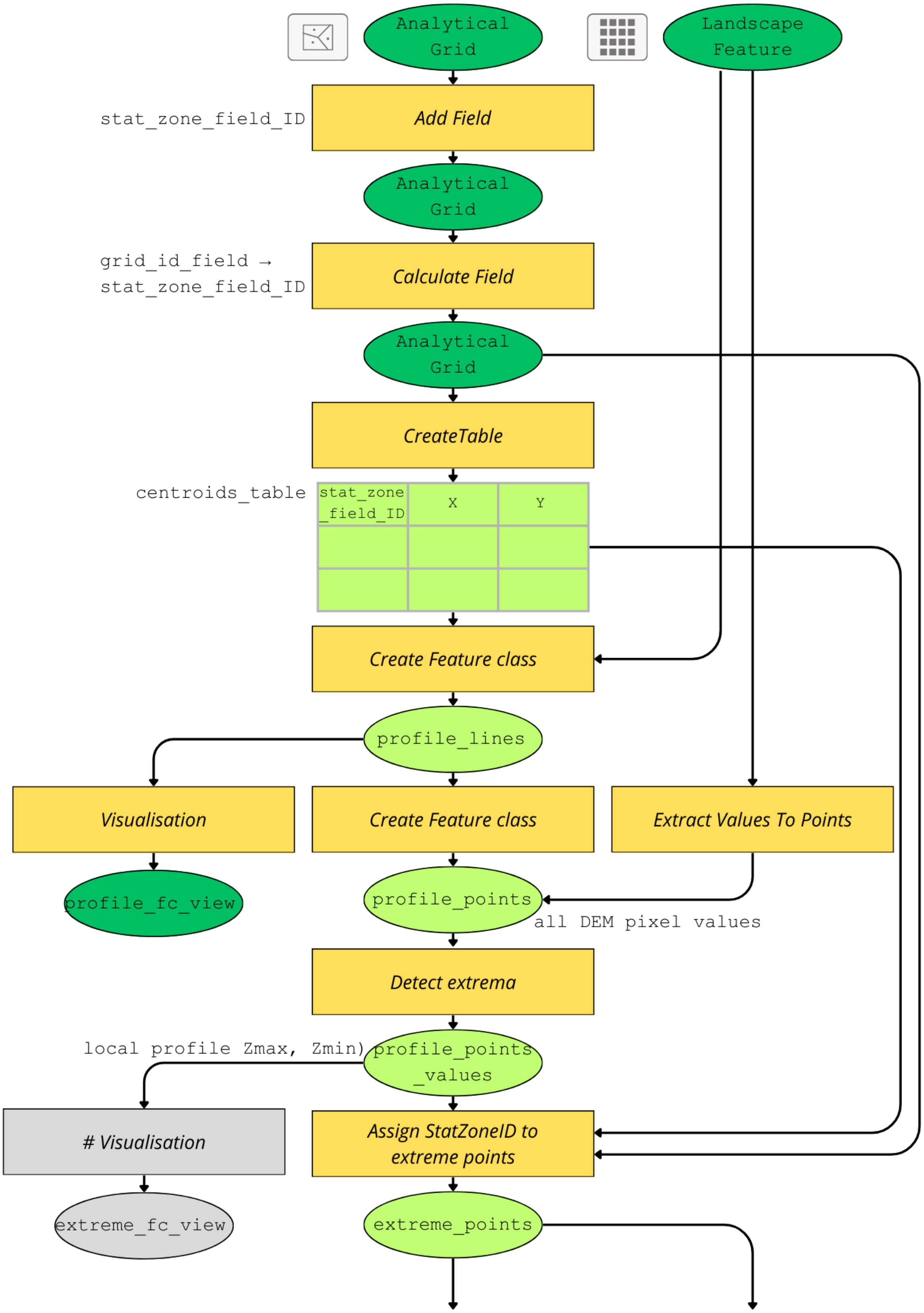
- west-east (W–E) profiles,
- north–south (N–S) profiles.

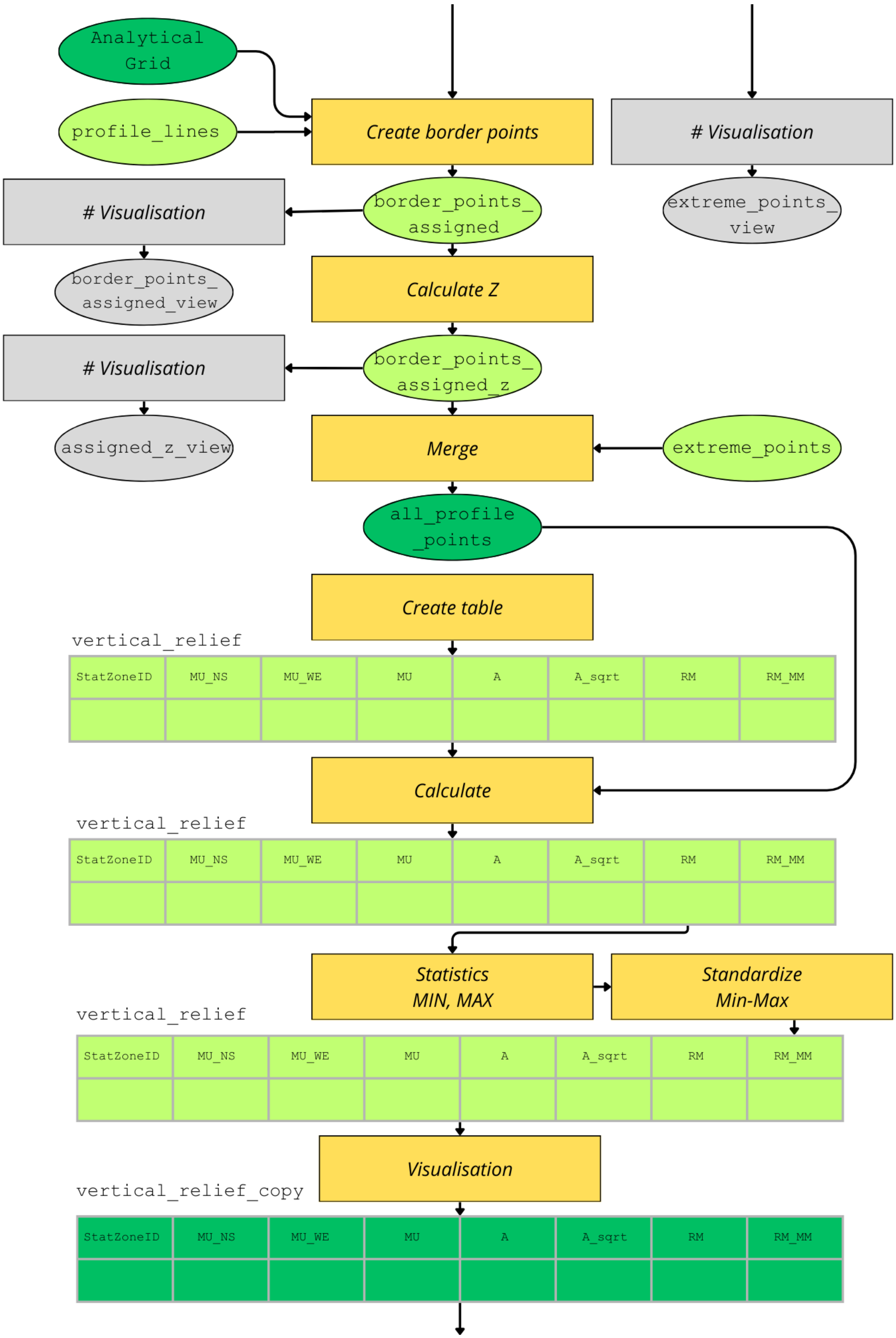
Profiles pass through unique centroid coordinates and cover the entire raster extent. Profile lines are stored as in-memory polyline features.

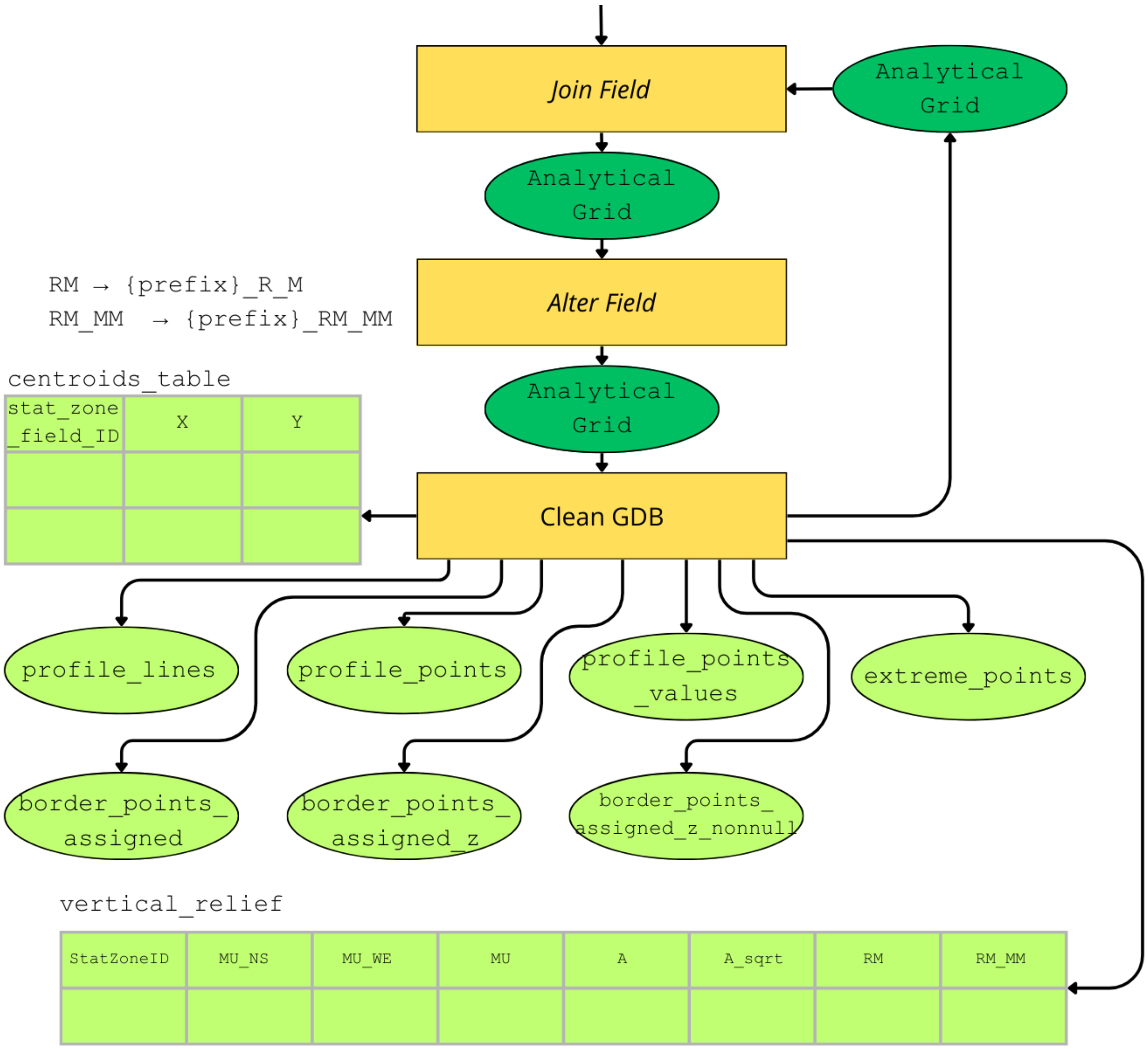
**Step 8:** Along each profile, generate a regular set of sampling points at a resolution corresponding to the raster cell size. Raster elevation values are interpolated to

| | |
|---|---|
| | these points using the *Extract Values to Points* function. |
| Step 9: | Analyze the elevation values along each profile to detect local extrema (maxima and minima). Additionally, include the start and end points of the profile in the set of extrema. Results are stored in memory as an in-memory point feature class called `extreme_points`. |
| Step 10: | Using a *Spatial Join* and additional validation against centroid axes (with half-pixel tolerance), assign the extrema points to the corresponding analytical grid cells. Geometric validation prevents incorrect assignment of points lying on the boundaries between adjacent cells. |
| Step 11: | Generate boundary points for the profiles of each cell, corresponding to intersections of the W–E and N–S profiles with the polygon boundaries. These points represent the entry and exit locations of the profiles for each cell. |
| Step 12: | Extract raster elevation values for the boundary points. Points with null values (`NULL`) are automatically removed to ensure correctness of subsequent calculations. |
| Step 13: | Merge the extreme points and boundary points into a single point feature class containing complete information on the elevation profile for each analytical grid cell. |
| Step 14: | Create an in-memory non-spatial table, `vertical_relief`, for calculating the vertical relief index. The table contains fields for the cell identifier (`StatZoneID`), $\mu\_NS$ (`MU_NS`) and $\mu\_WE$ (`MU_WE`) components, mean $\bar{\mu}$ (`MU`), cell area parameters (`A`, `A_sqrt`), and Steinhaus vertical relief index values (`RM` and `RM_MM`). |
| Step 15: | Load all profile points into a Python dictionary in RAM to avoid repeated database queries. For each zone (`StatZoneID`), sort points by coordinates (`X` for W–E profiles, `Y` for N–S profiles), then calculate `MU_NS` and `MU_WE`. Next, compute the mean value `MU` (23). |
| Step 16: | Read cell areas (`A`) from the analytical grid attribute table. For each cell, calculate the square root of the area (`A_sqrt`) as a normalization factor for the index. |
| Step 17: | Calculate the Steinhaus vertical relief index `R_M` (25). Store the results in the `RM` field of the `vertical_relief` table. |
| Step 18: | Standardize `R_M` using the *Min–Max* method. The script manually determines the minimum and maximum `RM` values of the population, then calculates the normalized form of the index, `RM_MM`. |
| Step 19: | Join the calculation results to the input analytical grid using a table join operation, then rename temporary fields to the final names according to the project naming conventions. |
| Step 20: | Finally, remove the temporary `StatZoneID` field from the analytical grid, clear the workspace cache, and terminate the script. |

## *Flow Diagram*




Analytical Grid
profile_lines
Create border points
# Visualisation
border_points_assigned
extreme_points_view
# Visualisation
border_points_assigned_view
Calculate Z
border_points_assigned_z
# Visualisation
assigned_z_view
Merge
extreme_points
all_profile_points
Create table
vertical_relief
StatZoneID
MU_NS
MU_WE
MU
A
A_sqrt
RM
RM_MM
Calculate
vertical_relief
StatZoneID
MU_NS
MU_WE
MU
A
A_sqrt
RM
RM_MM
Statistics MIN, MAX
Standardize Min-Max
vertical_relief
StatZoneID
MU_NS
MU_WE
MU
A
A_sqrt
RM
RM_MM
Visualisation
vertical_relief_copy
StatZoneID
MU_NS
MU_WE
MU
A
A_sqrt
RM
RM_MM




## Relevance to Geodiversity Assessment

- The Steinhaus vertical relief index (*R_M*) is conceptually simple but computationally demanding geomorphometric measure designed for regionally continuous variables. While it can be applied to any 3D variable, it was originally developed and is most widely used for describing the variability of terrain energy. Therefore, in the following description, it will be referred to in the context of terrain morphology.
- Unlike indices based on the analysis of the variability of all pixels within an entire area (e.g., *SD* or *SDc*), the values of the Steinhaus vertical relief index depend on the variability of terrain elevation along orthogonal profiles intersecting the analyzed surface. The index describes the intensity of elevation changes along these profiles, taking into account both their amplitude and frequency.
- Calculating the index does not require measuring horizontal distances, but only the analysis of elevation points encountered along the profile lines. This makes the index robust against errors from horizontal projection, focusing exclusively on the vertical aspect of geodiversity.
- In known applications, including the Geodiversity Tools, the morphological variability of each analyzed area is assessed along two perpendicular profiles passing through the centroid of each analytical grid cell.

- The index values range from 0—for theoretically perfectly flat areas—to infinity for areas with maximally dynamic terrain relief.
- High $R\_M$ values indicate high heterogeneity and complex vertical structure of the landscape, which is characteristic of areas with multiprocess genesis. Lower values correspond to areas with leveled elevation, reflecting low morphodiversity.
- The $R\_M$ index is additive. If calculations are based on the same set of geometric profiles (N–S and W–E) passing through the centroids of statistical zones, the sum of the Steinhaus vertical relief index values calculated for disjoint subareas equals the index value calculated for the entire area. This means that $R\_M$ does not depend on the size of the analytical cell or the number of profile points, as long as the same profile definition and the same resolution of the elevation raster from which Z values are sampled are maintained.

# Acknowledgements

This study was supported by the grant from AGH University in Krakow 16.16.140.315.

# Section 7 Glossary

*The glossary provides a set of fundamental concepts used in geodiversity assessments and the GIS tools developed within the project. The definitions encompass geomorphological and geological terminology, as well as terms related to spatial data processing, landscape structure, and quantitative metrics. The objective of the glossary is to standardize terminology, facilitate the interpretation of results, and ensure terminological consistency across documentation, source code, and scientific publications. Entries are cross-referenced, allowing for efficient navigation between related concepts and enhancing the understanding of the broader context of landscape and geoconservation analyses.*

**Analytical grid** – a polygon ➤feature class covering the study area and dividing it into *n* disjoint features—basic units (➤statistical zones)—within which ➤geodiversity assessments are performed. The grid may be based on natural, pseudo-natural, or artificial boundaries. The analytical grid enables the standardization of the ➤spatial scale of geodiversity analyses.

**AR model** (Aggregating Ratings model)[16] – the most widely used quantitative ➤geodiversity model, in which the diversity of ➤partial criteria—calculated for individual ➤landscape features and ➤landscape elements—is aggregated within the ➤statistical zones of the ➤analytical grid.

**Attribute** – in ➤GIS, descriptive information assigned to features, most commonly within ➤vector datasets of a ➤feature class or ➤non-spatial tables. Attributes are stored in an ➤attribute table, typically linked to the geometry of ➤vector datasets, and can encompass both ➤qualitative and ➤quantitative spatial variables. In ➤geodiversity analyses, attributes enable the calculation of ➤geodiversity indexes as well as the classification and aggregation of data within the ➤statistical zones of an ➤analytical grid.

**Attribute table** – a tabular data structure describing a set of geographic features, in which rows correspond to individual features, while columns represent their ➤attributes. Attribute tables refer to both ➤vector and ➤raster datasets. In ➤GIS, attribute tables are most commonly linked to ➤feature classes, and their fields can be used for querying, filtering, analyzing, and symbolizing spatial features.

**Categorical spatial variables** – a ➤spatial variable whose values describe spatial objects or their qualitative characteristics (e.g., rock types, land cover classes, categories of morphological landforms) (➤qualitative spatial variable). Such data are discrete in nature and cannot be ordered along a natural numerical scale. They are widely used in landscape mosaic and ➤geodiversity analyses.

**Centroid** – in ➤GIS, a point representing the geometric center of a spatial object, such as a polygon, line, or a set of points.

**Circular statistics** – a branch of statistics concerned with the analysis of directional or angular ➤quantitative spatial variables that are cyclical in nature (e.g., dip azimuths, wind directions, etc.). In the context of landscape and ➤geodiversity analyses, circular statistics allow for the description of the directional variability of features such as aspect and for the calculation of aggregate measures, such as mean directions or standard deviations, within the ➤statistical zones of the ➤analytical grid. Typical measures include the mean resultant vector, circular variance, and circular standard deviation.

**Continuous spatial variable** – a ➤spatial variable whose values can change fluently across space (➤quantitative spatial variable) and can take any numerical value within a defined range (e.g., elevation above sea level, slope, temperature). These variables describe spatial processes or properties that do not occur as clearly separated categories.

**Diversity index** – (metric, indicator) a ➤landscape index used for the quantitative assessment of the diversity of ➤landscape features. In ➤geodiversity analyses, it is referred to as a ➤geodiversity index.

**Entropy** – (in ➤geodiversity) a measure of disorder, complexity, or spatial heterogeneity of observed ➤landscape units. It describes the degree of differentiation between object types (e.g., lithological classes, geomorphological landforms, ➤geosites) and the evenness of their proportions within the analyzed ➤statistical zone. Higher entropy indicates greater diversity and a more balanced distribution of categories, whereas lower entropy points to the dominance of single elements or a simplified landscape structure. Entropy is based on the probability distribution of class proportions and typically utilizes the natural logarithm, ensuring comparability of results between regions and across different ➤spatial scales.

**Feature class** – the fundamental data unit in ➤vector datasets of ➤GIS systems, containing a homogeneous collection of objects of the same type, with the same geometry (point, line, or polygon), described by the same set of ➤attributes, and referring to the same ➤spatial scale. In Geodiversity Tools, feature classes serve as input datasets of ➤landscape features used for calculating ➤landscape indexes and ➤geodiversity indexes, as well as intermediate datasets created during ➤geoprocessing.

**Geoconservation** – a branch of nature conservation concerned with the identification, inventory, documentation, protection, and rational use of abiotic nature elements (➤geosites). It encompasses the protection of geological, geomorphological, pedological, and hydrological landforms and processes, as well as features of the highest scientific, landscape, educational, and cultural value. The assessment of ➤geodiversity using ➤geodiversity indexes is one of the primary tools supporting geoconservation.

**Geodiversity** – the variety of abiotic elements of the Earth's environment (geosphere), encompassing the diversity of geological structure, geomorphological landforms, soils, hydrology, and climatic conditions. It includes both the variety of types (compositional diversity) and their distribution, configuration, and spatial relationships. Geodiversity influences landscape functioning, biodiversity, the geotourism potential of a region, and other factors. In ➤GIS analyses, the quantitative assessment of geodiversity is enabled by ➤landscape metrics, including dedicated ➤geodiversity indexes.

**Geodiversity index** – a ➤landscape metric used for the quantitative assessment of ➤geodiversity based on ➤vector or ➤raster datasets. ➤Partial criteria are defined based on geodiversity indexes and ➤landscape features. Quantitative geodiversity analysis enables the identification of biodiversity potential and the comparison of ➤analytical grid units across various ➤spatial scales.

**Geodiversity model** – a quantitative model used to assess ➤geodiversity, linking the diversity of ➤partial criteria with the variations of ➤landscape features, ➤landscape elements, and total geodiversity.

**Geographic Information System (GIS)** – an integrated computer system designed for acquiring, storing, managing, processing, analyzing, visualizing, and interpreting spatial data. GIS allows for the linking of geographic information with its descriptive ➤attributes, supporting landscape analyses, ➤spatial modeling, and the assessment of environmental phenomena. In the context of ➤geodiversity, GIS enables the creation of ➤analytical grids, the calculation of ➤geodiversity indexes, and the visualization and comparison of ➤landscape features across various ➤spatial scales.

**Geosite** – a location or an object of abiotic nature characterized by scientific, educational, cultural, or aesthetic value. Geosites are key components of ➤geodiversity and are frequently subject to documentation, valuation, and protection within the framework of ➤geoconservation. They may serve as reference points for field research, tourism, or the monitoring of natural processes.

**Geoprocessing** – a set of analytical and transformational operations applied to spatial data within ➤GIS systems, encompassing activities such as selection, editing, format conversion, spatial analysis, modeling, and process automation. The objective of geoprocessing is to transform input data (➤vector, ➤raster, or ➤nonspatial table datasets) into new geographic information, models, and cartographic products that enable interpretation, analysis, and spatial decision support.

**GIS** – ➤Geographic Information System.

**Landscape elements** – constituent components of the landscape, including geological structure, landform, hydrosphere, soils, climate, vegetation, and fauna, which represent the direct subject of physical-geographical research. They function in parallel, interpenetrating and interacting with one another, forming the vertical structure of the landscape. Each of the landscape elements can be described by a number of ➤landscape features. The abiotic elements of the landscape are the subject of ➤geodiversity studies.

**Landscape feature** – a measurable characteristic or attribute describing ➤landscape elements (e.g., lithology, stratigraphy, slope, aspect). It can be represented using ➤vector or ➤raster datasets. In ➤geodiversity analysis, it is used for the quantitative assessment of the diversity of a selected ➤landscape element.

**Landscape metrics (index)** – quantitative measures originating from landscape ecology that describe the structure (composition) and spatial organization (configuration) of ➤landscape units. The main groups of metrics include: area and edge, shape, core area, contrast, aggregation, and ➤diversity indexes. The latter are utilized in ➤geodiversity assessment, where they are adapted to analyze the differentiation of abiotic ➤landscape elements.

**Linear statistics** – (also referred to as Euclidean or Standard statistics) statistics applied to ➤quantitative spatial variables of a linear, non-cyclical nature (in contrast to ➤circular statistics). In the context of ➤geodiversity and ➤GIS analyses, these include, among others, the range and standard deviation of ➤landscape features such as pixel values (elevation), line lengths, or polygon areas. Linear statistics are used to assess the diversity and quantitative characteristics of ➤landscape features within ➤statistical zones.

**Landscape unit** – the fundamental unit for describing landscape structure and configuration, representing a single spatial object. In ➤vector datasets, it can take the form of a point, line, or polygon. In ➤geodiversity analyses, the properties of landscape units are aggregated within an ➤analytical grid to calculate the ➤partial criteria of ➤geodiversity.

**Morphodiversity** – in ➤geodiversity analysis, the diversity of landforms (a ➤landscape element). Typically described using ➤landscape features such as: relative relief (denivelation), slope, aspect, profile and plan curvature, among others. Some of the ➤spatial criteria describing them are angular in nature and require the use of ➤geodiversity indexes that employ ➤circular statistics.

**Non-spatial table** – in ➤GIS – an ➤attribute table that lacks geometric information. It can store descriptive information, but because it does not contain a geographic component, its records cannot be rendered on a map.

**Normalization** – in ➤GIS – a ➤geoprocessing operation consisting in converting ➤quantitative spatial variables or ➤landscape metrics values into relative values by relating them to the size of the spatial unit within which they were calculated. It is applied in analyses based on ➤analytical grids with natural or pseudonatural ➤statistical zones in order to ensure the comparability of results between units that differ in area, shape, or spatial extent. Normalization removes the influence of the geometry of reference units on indicator values. It is not equivalent to ➤standardization, which concerns the unification of value scales.

**Partial criterion** – a quantitative metric of a ➤landscape feature, developed based on ➤landscape metrics and ➤geodiversity indexes, describing a specific aspect of the compositional or structural diversity of the landscape (e.g., *Lithology_Nc*, *Aspect_SDc*). Within ➤geodiversity models, partial criteria are used to calculate the diversity of ➤landscape features, ➤landscape elements, and total ➤geodiversity.

**Quantitative spatial variable** – a ➤spatial variable expressed in numerical form, allowing for statistical analysis, calculations, and the comparison of results (e.g., area, elevation, number of elements, slope). Quantitative variables can be both ➤continuous and ➤categorical spatial variables, provided that the values represent measurable magnitudes.

**Qualitative spatial variable** – a ➤spatial variable describing the properties of ➤landscape units or areas in a categorical, non-numerically measurable way. The values of a qualitative variable represent classes, types, or categories, such as lithology types or soil types.

**Raster dataset** – a digital spatial dataset based on the raster model, consisting of a matrix of cells (pixels) organized into rows and columns, where each cell is assigned a value representing specific information. Rasters can store both ➤quantitative spatial variables (e.g., elevation, slope, temperature values) and ➤categorical spatial variables (e.g., soil type, land cover class). In ➤geodiversity analyses, raster datasets enable the calculation of ➤geodiversity indexes within the ➤statistical zones of the ➤analytical grid.

**RCM model** – Raster Continuous Morphodiversity model[15] – a morphodiversity model utilizing ➤quantitative spatial variables in the form of ➤continuous spatial variables and ➤raster datasets, as well as ➤partial criteria that employ ➤circular statistics.

**Script** – in ➤GIS systems, a standalone executable file (in this project, written in Python) used for the automation of ➤geoprocessing. Scripts can be embedded in ➤toolboxes, and within them, in ➤toolsets. They typically retrieve input data and parameters, are designed to process these data, and generate results, such as ➤geodiversity index values within an ➤analytical grid.

**Spatial autocorrelation** – a statistical phenomenon described in 1970 by Waldo Tobler's First Law of Geography, stating that the similarity between geographic objects depends on their distance from one another. It posits that near locations (areas, units) have more similar characteristics than those that are farther apart.

**Spatial modeling** – a methodology or set of analytical procedures used to examine and describe the spatial relationships between geographic phenomena. It integrates ➤GIS and remote sensing techniques, enabling the creation of models that reflect the spatial and temporal differentiation of ➤landscape features. Models built in this manner allow for the prediction of future landscape configurations under varying environmental conditions.

**Spatial scale** – the specification of the level of detail or the geographic extent within which spatial analyses are conducted. Spatial scale can refer to the size of the study area, the data resolution (e.g., pixel size of ➤raster datasets), or the size and shape of the ➤statistical zones within an ➤analytical grid. In ➤geodiversity analyses, the appropriate selection of spatial scale is crucial for the comparability of results, the interpretation of ➤geodiversity indexes, and the identification of landscape structures and patterns.

**Spatial variable** – a characteristic or property assigned to a specific location within geographic space. The values of this variable are linked to a location and may exhibit differentiation depending on position. Spatial variables are the foundation of ➤GIS analysis, ➤spatial modeling, and ➤geodiversity research, as they enable the description and comparison of environmental properties across different parts of the study area.

**Standardization** – the process of transforming numerical values, such as ➤geodiversity indexes, so that they are comparable with one another, regardless of their original units, ranges, or ➤spatial scale. In ➤geodiversity analyses, standardization enables the comparison of ➤partial criteria calculated from different types of ➤vector and ➤raster datasets and facilitates their interpretation and integration within a chosen ➤geodiversity model. One of the most commonly used standardization methods is Min–Max, which transforms the values of a given index into a range of ⟨0; 1⟩.

**Statistical zone** (pole podstawowe, komórka siatki) – podstawowa jednostka obliczeniowa ➤analytical grid, w której dokonywane są oceny krajobrazowe. W analizach ➤geodiversity, za pomocą ➤geodiversity indexes obliczane są ➤partial criteria. Ich agregacja w ramach ➤AR models umożliwia syntezę informacji w ramach ➤landscape elements oraz georóżnnorodności całkowitej.

**Toolbox** – in ➤GIS – a collection of thematically related ➤scripts and ➤geoprocessing tools that enable reproducible analysis and consistent implementation through graphical user interfaces. A toolbox organizes script tools into ➤toolsets.

**Toolset** – in ➤GIS, a subset of a ➤toolbox that groups tools by theme or function (e.g., ➤geodiversity indexes). A toolset facilitates the organization, navigation, and logical linking of tools within a single ➤toolbox.

**Vector datasets** – data that describe geographic objects using points, lines, and polygons. In ➤GIS, the vector data model is particularly useful for storing information about objects with discrete boundaries (➤qualitative spatial variables). Each vector object, in addition to its geometry, can contain descriptions in the form of ➤attributes, organized into ➤attribute tables.

# Section 8 References


1. Neef E. *Die Theoretischen Grundlagen Der Landschaftslehre*. VEB Hermann Haack; 1967.
2. Leser H, Nagel P. Landscape diversity — a holistic approach. In: Barthlott W, Winiger M, Biedinger N, eds. *Biodiversity: A Challenge for Development Research and Policy*. Springer Berlin Heidelberg; 2001:129-143. doi:10.1007/978-3-662-06071-1_9
3. Batlle M, Ramón J, Hoek (van der) Y. Plant community associations with morpho-topographic, geological and land use attributes in a semi-deciduous tropical forest of the Dominican Republic. *Neotrop Biodivers*. 2021;7(1):465-475. doi:10.1080/23766808.2021.1987769
4. Tukiainen H, Toivanen M, Maliniemi T. Geodiversity and Biodiversity. *Geological Society, London, Special Publications*. 2023;530(1):31-47. doi:10.1144/SP530-2022-107
5. Tukiainen H, Bailey JJ, Field R, Kangas K, Hjort J. Combining geodiversity with climate and topography to account for threatened species richness. *Conservation Biology*. 2017;31(2):364-375. doi:10.1111/cobi.12799
6. Antonelli A, Kissling WD, Flantua SGA, et al. Geological and climatic influences on mountain biodiversity. *Nat Geosci*. 2018;11(10):718-725. doi:10.1038/s41561-018-0236-z
7. Semeniuk V. The linkage between biodiversity and geodiversity. In: Eberhard R, ed. *Pattern & Processes: Towards a Regional Approach to National Estate Assessment of Geodiversity*. Technical Series 2. Australian Heritage Commission & Environment Forest Taskforce, Environment; 1997:51-58.
8. Stanley M. Geodiversity. *Earth Heritage*. 2000;14:15-18. Accessed July 12, 2023. http://www.earthheritage.org.uk/wp/wp-content/uploads/2018/03/EH14-2000.pdf
9. Stanley M. Geodiversity: Our foundation. *Geology Today*. 2003;19(3):104-107.
10. Gray M. *Geodiversity: Valuing and Conserving Abiotic Nature*. 2nd Edition. Wiley-Blackwell; 2013.
11. Gray M. Geodiversity: the origin and evolution of a paradigm. *Geological Society, London, Special Publications*. 2008;300(1):31-36. doi:10.1144/SP300.4
12. Brilha J, Gray M, Pereira DI, Pereira P. Geodiversity: An integrative review as a contribution to the sustainable management of the whole of nature. *Environ Sci Policy*. 2018;86:19-28. doi:10.1016/j.envsci.2018.05.001
13. Gray M. Geodiversity and the ecosystem approach. *Parks Stewardship Forum*. 2022;38(1):39-45. doi:10.5070/P538156117
14. Bartuś T, Mastej W. Morphodiversity as a Tool in Geoconservation: A Case Study in a Mountain Area (Pieniny Mts, Poland). *Sustainability*. 2023;15(14):11357. doi:10.3390/SU151411357
15. Bartuś T, Mastej W. HOW to use continuous variables in geodiversity assessments – RASTER continuous morphodiversity model. *Environmental Modelling & Software*. 2025;193(0):106597. doi:10.1016/j.envsoft.2025.106597
16. Mastej W, Bartuś T. Supervised classification of morphodiversity using artificial neural networks on the example of the Pieniny Mts (Poland). *Catena (Amst)*. 2024;242:108086. doi:10.1016/j.catena.2024.108086
17. Sołowiej D. *Podstawy Metodyki Oceny Środowiska Przyrodniczego Człowieka*. Wydawnictwo Naukowe Uniwersytetu im. Adama Mickiewicza; 1992.

18. Jedicke E. Biodiversität, Geodiversität, Ökodiversität. Kriterien zur Analyse der Landschaftsstruktur – ein konzeptioneller Diskussionsbeitrag. *Naturschutz und Landschaftsplanung*. 2001;33(2/3):59-68. Accessed July 11, 2023. https://www.researchgate.net/profile/Eckhard_Jedicke/publication/279620421_Biodiversity_Geodiversity_Ecodiversity_Criteria_for_the_analysis_of_the_landscape_structure_-_A_conceptual_contribution_to_the_discussion/links/56e18fac08ae40dc0abf5117/Biodiversity-Geodiversity-Ecodiversity-Criteria-for-the-analysis-of-the-landscape-structure-A-conceptual-contribution-to-the-discussion.pdf

19. Crofts R. Linking geoconservation with biodiversity conservation in protected areas. *International Journal of Geoheritage and Parks*. 2019;7(4):211-217. doi:10.1016/j.ijgeop.2019.12.002

20. Alahuhta J, Toivanen M, Hjort J. Geodiversity – biodiversity relationship needs more empirical evidence. *Nat Ecol Evol*. 2019;4(1):2-3. doi:10.1038/s41559-019-1051-7

21. Gordon JE, Bailey JJ, Larwood JG. Conserving nature's stage provides a foundation for safeguarding both geodiversity and biodiversity in protected and conserved areas. *Parks Stewardship Forum*. 2022;38(1):46-55. doi:10.5070/P538156118

22. Sharples C. *A Methodology for the Identification of Significant Landforms and Geological Sites for Geoconservation Purposes*. 1993. Accessed June 20, 2023. https://www.researchgate.net/profile/Chris-Sharples-2/publication/266617978_A_Methodology_for_the_Identification_of_Significant_Landforms_and_Geological_Sites_for_Geoconservation_Purposes/links/5435db450cf2dc341db2d2a3/A-Methodology-for-the-Identification-of-Significant-Landforms-and-Geological-Sites-for-Geoconservation-Purposes.pdf

23. Sharples C. *Concepts and Principles of Geoconservation*. Park and Wildlife Service, Department of Environment and Land Management; 1998. Accessed July 12, 2023. https://www.researchgate.net/publication/266021113_Concepts_and_principles_of_geoconservation

24. Sharples C. *Concepts and Principles of Geoconservation*. Tasmanian Parks & Wildlife Service; 2002. Accessed June 12, 2023. https://www.researchgate.net/publication/266021113_Concepts_and_principles_of_geoconservation

25. Cleal CJ, Thomas BA, Bevins RE, Wimbledon WAP. GEOSITES - an international geoconservation initiative. *Geology Today*. 1999;15(2):64-68. doi:10.1046/j.1365-2451.1999.1502006.x

26. Brocx M, Semeniuk V. Geoheritage and geoconservation - History, definition, scope and scale. *J R Soc West Aust*. 2007;90(2):53-87. Accessed July 11, 2023. https://rswa.org.au/publications/Journal/90(2)/volume90part253-87l.pdf

27. Wimbledon WAP. Geoheritage in Europe and its conservation. *Episodes*. 2013;36(1):68-68. doi:10.18814/epiiugs/2013/v36i1/010

28. Reynard E, Brilha J. *Geoheritage: Assessment, Protection, and Management*. Elsevier; 2018.

29. Crofts R, Gordon JE, Brilha J, et al. *Guidelines for Geoconservation in Protected and Conserved Areas*. (Groves C, ed.). IUCN, International Union for Conservation of Nature; 2020. doi:10.2305/IUCN.CH.2020.PAG.31.en

30. Pescatore E, Bentivenga M, Giano SI. Geoheritage and Geoconservation: Some Remarks and Considerations. *Sustainability*. 2023;15(7):5823. doi:10.3390/su15075823

31. Alkhasawneh MSh, Ngah UK, Tay LT, Isa NAM. Determination of importance for comprehensive topographic factors on landslide hazard mapping using artificial neural network. *Environ Earth Sci*. 2014;72(3):787-799. doi:10.1007/s12665-013-3003-x

32. Chen Y, Li N, Chang M, et al. Spatial Heterogeneity of Dominant Controlling Factors for Seismic Landslide Susceptibility Zones: A Case Study of the Barkam Earthquake. *Transactions in GIS*. 2025;29(6). doi:10.1111/tgis.70116

33. Fox N, Graham LJ, Eigenbrod F, Bullock JM, Parks KE. Incorporating geodiversity in ecosystem service decisions. *Ecosystems and People*. 2020;16(1):151-159. doi:10.1080/26395916.2020.1758214

34. Comer PJ, Pressey RL, Hunter ML, et al. Incorporating geodiversity into conservation decisions. *Conservation Biology*. 2015;29(3):692-701. doi:10.1111/cobi.12508

35. Barrow CJ. *Environmental Management: Principles and Practice*. Routledge; 1999. Accessed June 13, 2023. https://uomustansiriyah.edu.iq/media/lectures/5/5_2020_03_04!03_12_11_PM.pdf

36. Kot R. *Georóżnorodność – Problem Jej Oceny i Zastosowania w Ochronie i Kształtowaniu Środowiska Na Przykładzie Fordońskiego Odcinka Doliny Dolnej Wisły i Jej Otoczenia*. Towarzystwo Naukowe w Toruniu, Uniwersytet Mikołaja Kopernika; 2006.

37. Fassoulas C, Mouriki D, Dimitriou-Nikolakis P, Iliopoulos G. Quantitative Assessment of Geotopes as an Effective Tool for Geoheritage Management. *Geoheritage*. 2012;4(3):177-193. doi:10.1007/s12371-011-0046-9

38. Najwer A, Zwoliński Z. Semantyka i metodyka oceny georóżnorodności – przegląd i propozycja badawcza. *Landform Analysis*. 2014;26:115-127. doi:10.12657/landfana.026.011

39. Brilha J. Inventory and Quantitative Assessment of Geosites and Geodiversity Sites: a Review. *Geoheritage*. 2016;8(2):119-134. doi:10.1007/s12371-014-0139-3

40. Zwoliński Z, Najwer A, Giardino M. Methods for Assessing Geodiversity. In: Reynard Emmanuel, Brilha J, eds. *Geoheritage. Assessment, Protection, and Management*. Elsevier; 2018:27-52. doi:10.1016/B978-0-12-809531-7.00002-2

41. Bartuś T. *Struktura i Różnorodność Abiotycznych Komponentów Krajobrazu w Ocenie i Delimitacji Obszarów Chronionych Na Przykładzie Ojcowskiego Parku Narodowego i Jego Otoczenia*. Wydawnictwa AGH; 2020. Accessed June 12, 2023. https://open.icm.edu.pl/bitstreams/0207012c-2c1f-410b-aa9f-b75bde9eaade/download

42. Paszyński J. Metody sporządzania map topoklimatycznych. In: Kluge M, ed. *Metody Opracowań Topoklimatycznych. Dokumentacja Geograficzna, 3*. Instytut Geografii i Przestrzennego Zagospodarowania PAN; 1980:13-28.

43. Bartuś T. Model zmienności topoklimatycznej rejonu Ojcowskiego Parku Narodowego. *Prądnik Prace i Materiały Muzeum W Szafera*. 2014;24:25-46. Accessed December 23, 2025. https://open.icm.edu.pl/items/468e3012-a826-48a5-bcf0-21b8871dfd72

44. Paszyński J, Miara K, Skoczek J. Wymiana energii między atmosferą a podłożem jako podstawa kartowania topoklimatycznego. In: *Wymiana Energii Między Atmosferą a Podłożem Jako Podstawa Kartowania Topoklimatycznego. Dokumentacja Geograficzna, 14*. Instytut Geografii i Przestrzennego Zagospodarowania PAN; 1999.

45. Paszyński J. Wymiana energii na powierzchni czynnej jako podstawa klasyfikacji topoklimatycznej. *Acta Agrophysica*. 2004;3(2):351-358.

46. Kicińska B, Olszewski K, Żmudzka E. Uwagi o wykorzystaniu klasyfikacji J. Paszyńskiego do kartowania topoklimatycznego (z doświadczeń Zakładu Klimatologii Uniwersytetu Warszawskiego). In: Kuchcik M, ed. *Współczesne Badania Topoklimatyczne. Dokumentacja Geograficzna, 23*. Instytut Geografii i Przestrzennego Zagospodarowania; 2001:143-151.

47. Kot R, Leśniak K. Ocena georóżnorodności za pomocą miar krajobrazowych – podstawowe trudności metodyczne. *Przegląd Geograficzny*. 2006;78(1):25-45.

48. Han J, Kamber M, Pei J. *Data Mining : Concepts and Techniques*. 3rd ed. Elsevier / Morgan Kaufmann; 2012. Accessed December 22, 2025. https://ia800702.us.archive.org/7/items/datamining_201811/DS-book%20u5.pdf

49. Pietrzak M. *Syntezy Krajobrazowe – Założenia, Problemy, Zastosowania*. Bogucki Wydawnictwo Naukowe; 1998.

50. Bartkowski T. *Zastosowania Geografii Fizycznej*. Państwowe Wydawnictwo Naukowe; 1986.

51. Bartsch H. Arbeitsmetoden in der Landschäftsökologie. In: Heyer E, ed. *Arbeitsmethoden in Der Physischen Geographie*. Volk und Wissen; 1968:34-46.

52. Kondracki J. *Podstawy Regionalizacji Fizycznogeograficznej*. Wyd. 2. PWN; 1976.

53. Kondracki J, Richling A. Próba uporządkowania terminologii w zakresie kompleksowej geografii fizycznej. *Przegląd Geograficzny*. 1983;55(1):201-217.

54. Balon J. Unifikacja typów geokompleksów w skali kraju podstawą waloryzacji krajobrazu. In: Kistowski M, Korwel-Lejkowska B, eds. *Waloryzacja Środowiska Przyrodniczego w Planowaniu Przestrzennym*. Fundacja Rozwoju Uniwersytetu Gdańskiego; 2007:25-33.

55. Balon J. O dwóch szkołach wyróżniania I klasyfikacji geokompleksów. In: Ziaja W, Jodłowski M, eds. *Struktura Środowiska Przyrodniczego a Fizjonomia Krajobrazu*. Instytut Geografii i Gospodarki Przestrzennej, Uniwersytet Jagieloński; 2014:53-61.

56. Balon J. Organizacja krajobrazu w skali globalnej, regionalnej i lokalnej. In: Beczkiewicz M, ed. *Akademia Talentów Przyrodniczych. Podwyższanie Jakości Kształcenia Kompetencji Naukowych i Przyrodniczych w Szkołach Ponadgimnazjalnych w Ramach Przedmiotu Ekologia Krajobrazu*. Instytut Środowiska Rolniczego i Leśnego Polskiej Akademii Nauk; 2018:1-14.

57. Forman RTT, Godron M. *Landscape Ecology*. Wiley and Sons; 1991.

58. Parysek JJ. *Modele Klasyfikacji w Geografii*. Vol 31. Seria Geografia. Wydawnictwo Naukowe Uniwersytetu Adama Mickiewicza; 1982.

59. Matheron G. Principles of geostatistics. *Economic Geology*. 1963;58:1246-1266.

60. Eiden G, Kayadjanian M,, Vidalm Claude. *From Land Cover to Landscape Diversity in the European Union*. 2000.

61. Bartuś T. Basic unit size in the analysis of the distribution of spatial landscape elements on the basis of the lithostratigraphic geodiversity of the Ojców National Park (Poland). *Geology, Geophysics & Environment*. 2017;43(2):95. doi:10.7494/geol.2017.43.2.95

62. Plotnick RE, Gardner RH, O'Neill R V. Lacunarity indices as measures of landscape texture. *Landsc Ecol*. 1993;8(3):201-211. doi:10.1007/BF00125351

63. Plotnick RE, Gardner RH, Hargrove WW, Prestegaard K, Perlmutter M. Lacunarity analysis: a general technique for the analysis of spatial patterns. *Physical Review*. 1996;53:1-8.

64. Legendre P, Fortin MJ. Spatial pattern and ecological analysis. *Vegetatio*. 1989;80:107-138.

65. Guo Q, Kelly M. Interpretation of scale in paired quadrat variance methods. *Journal of Vegetation Science*. 2004;15(6):763-770. doi:10.1111/j.1654-1103.2004.tb02319.x

66. Krummel JR, Gardner RH, Sugihara G, O'Neill RV, Coleman PR. Landscape Patterns in a Disturbed Environment. *Oikos*. 1987;48(3):321. doi:10.2307/3565520

67. Rossi RE, Mulla DJ, Journel AG, Franz EH. Geostatistical Tools for Modeling and Interpreting Ecological Spatial Dependence. *Ecol Monogr*. 1992;62(2):277-314. doi:10.2307/2937096

68. Meisel JE, Turner MG. Scale detection in real and artificial landscapes using semivariance analysis. *Landsc Ecol*. 1998;13(6):347-362. doi:10.1023/A:1008065627847

69. Radeloff VC, Miller TF, He HS, Mladenoff DJ. Periodicity in spatial data and geostatistical models: autocorrelation between patches. *Ecography*. 2000;23(1):81-91. doi:10.1111/j.1600-0587.2000.tb00263.x

70. Suchożebrski J. The Size of the Basic Unit in Geographical Analysis. *Miscellanea Geographica*. 2004;11(1):151-160. doi:10.2478/mgrsd-2004-0017

71. Turner MG, O'Neill R V., Gardner RH, Milne BT. Effects of changing spatial scale on the analysis of landscape pattern. *Landsc Ecol*. 1989;3(3-4):153-162. doi:10.1007/BF00131534

72. Gehlke CE, Biehl K. Certain effects of grouping upon the size of the correlation coefficient in census tract material. *J Am Stat Assoc*. 1934;29:169-170.

73. Openshaw S. *The Modifiable Areal Unit Problem. CATMOG 38*. GeoBooks; 1984.

74. Bartuś T. Raster images generalization in the context of research on the structure of landscape and geodiversity. *Geology, Geophysics & Environment*. 2014;40(3):271. doi:10.7494/geol.2014.40.3.271

75. Shannon CE, Weaver W. *The Mathematical Theory of Communication*. University of Illinois Press; 1949.

76. Batschelet E. *Statistical Methods for the Analysis of Problems in Animal Orientation and Certain Biological Rhythms*. American Institute of Biological Sciences; 1965.

77. Mahan RP. *Circular Statistical Methods: Applications in Spatial and Temporal Performance Analysis*. U.S. Army Research Institute for the Behavioral and Social Sciences; 1991. Accessed July 2, 2024. https://apps.dtic.mil/sti/tr/pdf/ADA240751.pdf

78. Swan ARH, Sandilands M, McCabe P. *Introduction to Geological Data Analysis*. Blackwell Science Ltd.; 1995.

79. Agterberg FP. *Geomathematics. Mathematical Background and Geo-Science Applications*. Elsevier Scientific Pub. Co.; 1974.

80. Steinhaus H. O wskaźniku ukształcenia pionowego. *Przegląd Geograficzny*. 1947;21(1-2):113-115.

81. Wieczorek M, Żyszkowska W. Geomorfometria – parametry morfometryczne w charakterystyce rzeźby terenu. *Polski Przegląd Kartograficzny*. 2011;43(2):130-144. Accessed December 23, 2025. https://yadda.icm.edu.pl/baztech/element/bwmeta1.element.baztech-article-BAR9-0010-0080/c/wieczorek_zyszkowska_geomorfometria_2_2011.pdf

82. Szczepankiewicz S. Intensywność urzeźbienia krajobrazu okolic Wałbrzycha. *Prace Wrocławskiego Tow Naukowego" ser B*. 1948;8:1-23.

83. Gadzojannis A, Plewniak W. Mapa intensywności urzeźbienia Sudetów Kłodzkich. *Acta Universitatis Wratislaviensis Prace Inst Geogr, ser A*. 1974;236:79-83.