Comput Geosci (2011) 15:463–475
DOI 10.1007/s10596-010-9216-0

ORIGINAL PAPER

Multi-scale European Soil Information System (MEUSIS): a multi-scale method to derive soil indicators

Panos Panagos · Marc Van Liedekerke · Luca Montanarella

Received: 28 April 2010 / Accepted: 8 November 2010 / Published online: 22 December 2010

Abstract The Multi-scale Soil Information System (MEUSIS) can be a suitable framework for building a nested system of soil data that could facilitate interoperability through a common coordinate reference system, a unique grid coding database, a set of detailed and standardized metadata, and an open exchangeable format. In the context of INSPIRE Directive, MEUSIS may be implemented as a system facilitating the update of existing soil information and accelerating the harmonization of various soil information systems. In environmental data like the soil one, it is common to generalize accurate data obtained at the field to coarser scales using either the pedotransfer rules or knowledge of experts or even some statistical solutions which combine single values of spatially distributed data. The most common statistical process for generalization is averaging the values within the study area. In this paper, we do not present a simple averaging of numerical values without any further processed information. The upscaling process is accompanied with significant statistical analysis in order to demonstrate the method suitability. The coarser resolution nested grids cells (10×10 km) represent broad regions where the calculated soil property (e.g., organic carbon) can be accurately upscaled. Multi-scaled approaches are urgently required to integrate different disciplines (such as Statistics) and provide a meta-model platform to improve current mechanistic modeling frameworks, request new collected data, and identify critical research questions. Past papers have described in detail the upscaling methodology while our present approach is to demonstrate an important application of this methodology accompanied with statistical evidence.

Keywords Multi-scale soil information system · MEUSIS · Upscaling · Soil data · Descriptive statistics · INSPIRE

P. Panagos (✉) · M. Van Liedekerke · L. Montanarella
Joint Research Centre of the European Commission
Institute for Environment and Sustainability,
Via E. Fermi, 2749, 21027 Ispra, VA, Italy
e-mail: panos.panagos@jrc.ec.europa.eu

M. Van Liedekerke
e-mail: marc.van-liedekerke@jrc.ec.europa.eu

L. Montanarella
e-mail: luca.montanarella@jrc.ec.europa.eu

1 Introduction

During the last years the need for a coherent approach to soil protection has come on the political agenda in Europe and was therefore introduced as one of the thematic strategies to be developed within the Community's 6th Environment Action Programme. At present, the most complete source of soil information at European scale is the European Soil Database at scale 1:1,000,000. This database has been developed jointly with partners in participating countries resulting in the only harmonized coverage of digital soil information for Europe and it is available in the European Soil Portal (http://eusoils.jrc.ec.europa.eu/).

The INSPIRE (Infrastructure for Spatial Information in Europe) directive [13] has embarked on a common framework for spatial data in the European Union.

One of the INSPIRE ideas is to conduct reporting and analysis of environmental information on the basis of a harmonized and hierarchical system of grids with a common point of origin and a standardized location and size of grid cells. In other words, the INSPIRE grid system proposal could lay the basis for a Multi-scale European Soil Information System (MEUSIS), a system whereby soil data produced at a certain scale can easily be integrated or compared with soil data produced at another scale, provided that the rules for representation of the data are equal at all scales [19]. "Upscaling" is the term used to describe the process of reducing a set of values in an area down to a single value representing such area as a whole in order to allow comparison and integration to other data sets.

The major objective of the article is to present the application of upscaling approach and the MEUSIS meta-model concept. The main focus is the development of a methodology (meta-model) for upscaling data from local scale to a regional one and from local towards European Scale (Fig. 1). The meta-model process will utilize the larger-scale data incorporating geo-statistical upscaling rules that will result in aggregated datasets on soil functions. Moreover, the framework of data processing through a multi-scale grid system will be presented in order to make feasible the results of the upscaling methodology.

MEUSIS should provide data for the assessment of soil conditions at different levels of detail and make available a structure so that coherent and complementary data, available at a nested set of geographical scales, can fit together. The data providers most likely will need to process their original soil data (held in traditional vector-based soil databases) in order to fit the proposed grid system. More specific objectives of the paper are the following:

- The use of MEUSIS for accurate evaluation of degraded data and the suitable presentation to policy makers for further decision making
- The demonstration of the spatial variability importance in earth sciences such as soil science and how this issue may be tackled using descriptive statistics
- The identification of different purposes of upscaling soil data

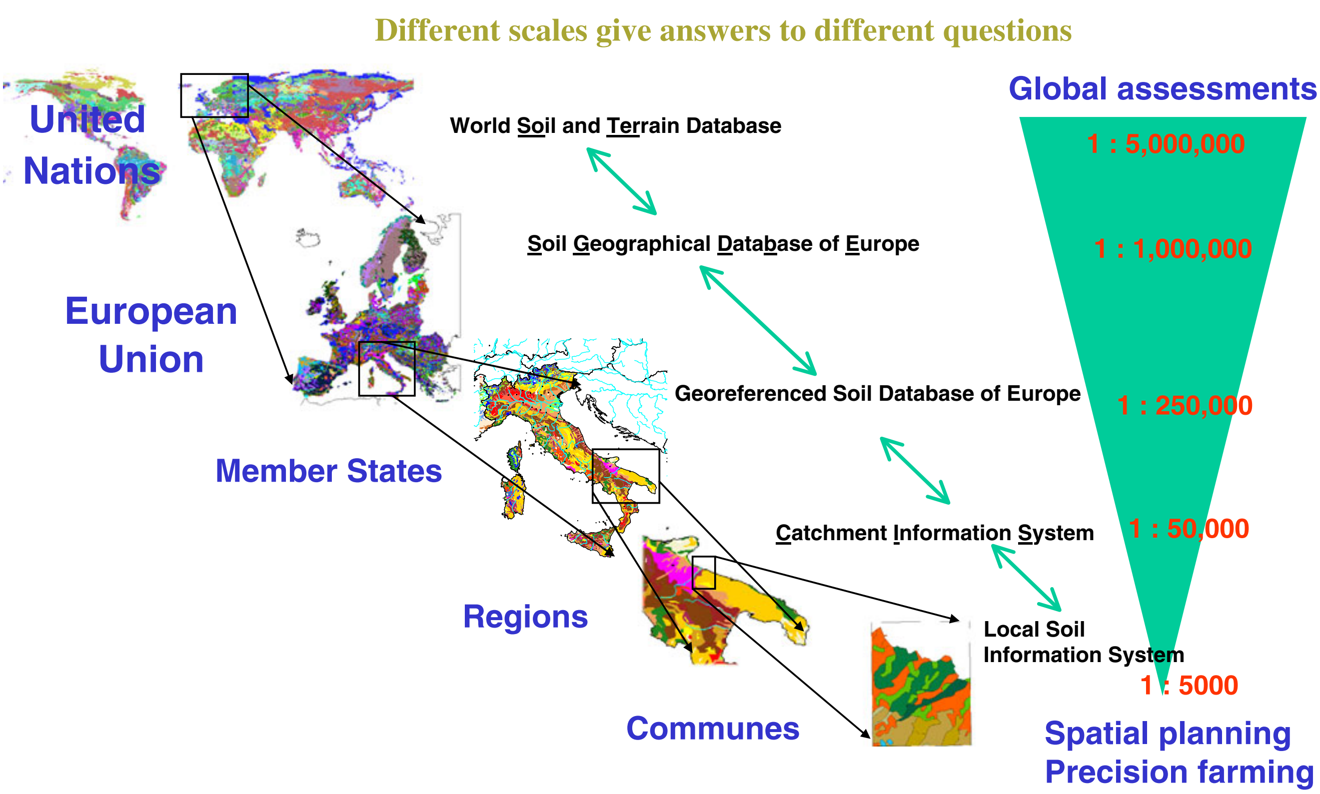

Fig. 1 Different scales of soil information

The paper introduces a multi-scale soil framework which helps the integration of research, soil science, and policy making.

2 Concepts

2.1 Upscaling

Spatial data and scales in environmental research cover a wide range. Since the research community wants to develop a common understanding of the impact of management changes on coarser resolutions, there is often a difficulty because of the lack of effectively transfer data or information among scales [22]. The way that those data exchanges or model results across space scales can be effectively accomplished is critical to help everyone in understanding the methodology on how they can move information across scales.

Scale issues are considered to be important in past bibliography [2] and advantages of upscaling have been reported in hydrology [21], soil science [17], and land quality [3]. The main components in terms of changing or matching scales are upscaling and downscaling. Upscaling is the process of aggregating information collected at a large scale towards a small scale [24]. In cartography, "Upscaling" or "aggregation" are the terms used to describe the process of reducing a set of components or values in an area down to a single value representing such an area as a whole. Aggregation implies simplification, the degree of variation in the considered area is reduced and thus there is an obvious loss of information [14].

Downscaling is the process of detailing information collected at a small scale towards a larger scale [6]. In upscaling, the distinction is necessary between upscaling of processes and upscaling of data. The first concerns the existence of different processes that act at different scales. Attention focuses on the second topic which is the main theme of this paper, and addresses useful ways of upscaling for individual sets of data on indicators.

It is widely recognized that soil measurements (such as many other environmental data) cannot be scaled up directly [1]. The types of data taken at a point (1 km^2) may differ greatly from measurement made at small catchments (10 km^2) or in large areas (100 km^2). As we scale up from point to large area, we accept that there is degree of uncertainty [11].

Upscaling takes a model that is applicable at a site-specific scale (at which observations are made), and projects the outcome in a way that it is applicable at coarser scale (regional or global) [4]. One of the main problems encountered in upscaling occurs when single-site observations do not represent spatial heterogeneity of the parameters or variables at larger spatial scales. In the context of the European Soil Data Center, the whole European Soil Coverage has been described as a nested hierarchy of various grids in scales according to INSPIRE Directive.

2.2 Development of aggregation methods for the meta-model

The implementation of geo-statistical processes will be based in the hierarchical grids approach according to INSPIRE Directive and will produce regional scale datasets. Often the assumption in upscaling is that the aggregation of values to a smaller scale is linear [9] in nature or some statistical methods will make approximations to natural variability. This process results in loss of detail, and may produce erroneous relationships when the natural variability is very high. The meta-model data process objective is to minimize those errors coming from the non-linear relationships and to find out the cases where there is high natural variability.

The aggregation method should be carefully chosen in order to display the maximum spatial variation of the data in a comprehensive manner. There are many methods for upscaling and the optimal one would be the one which ensures that the new value of the whole area is the most adequate according to the objectives of the upscaling. Major objective of the article is to identify the most appropriate applicability of upscaling in a real case. Soil variables can be represented either as quantitative variables (numeric parameters) or as qualitative (classes).

For the quantitative variables like the organic carbon content, the most appropriate method for upscaling is called "zoning" in which the mean, median, or the mode of the single values of an area is used to define a larger area [16]. In soil domain, the use of mean or median statistical indicator is recommended for upscaling, instead in other cases (where critical limits or values is an important factor assessing environmental data: floods, natural hazards) the percentiles, minimum, or maximum can be more appropriate statistical indicators.

Averaging is the formal part of the coarse-graining process while upscaling involves both mathematics and the intellectual or conceptual part of the modeling process [26]. Averaging can be done, in principle, for any heterogeneous subsurface system (such as the soil one), regardless of its complexity, or the spatial variation of its properties. However, averaging itself does not necessarily result in a reduction of the information content of the system. It is the imposition of various assumptions, which are termed scaling processes in this work.

Model-based approaches [10] advocated running the model at the finer scale and then aggregating the model output at a larger scale. Soil processes models which are calibrated at a small scale (points) cannot be valid for larger-scale (catchments) applications. This implies that in case data are available in points then it is preferable to run the model at this scale and then aggregate the results to the catchment scale.

For the qualitative variables (classes) such as Dominant World Reference Base Soil Types [18], the problem is related to the semantic component of the soil types. The recommended way to perform an aggregation is firstly to reduce semantically the soil units (variability of the soil types) and then to explore the most frequent (dominant) value. The most dominant value should be accompanied with the secondary dominant value, the third dominant value and with some inter-relationships between them. Obviously, this is more complicated than the quantitative aggregation and specific upscaling rules must be defined beforehand.

2.3 Requirements for a data model of nested spatial scales

As the soil threats have been described in the proposed Soil Thematic Strategy [5], there is a need to address them and relative issues at various scales; from local/province scale, to regional/national scale, and at the end to continental/global scale. The modeling platform should be constructed in such a way that knowledge and information can be passed along the spatial scales with the minimum loss of information. Key information and data validation will be obtained from studies conducted at soil observatories covering various nested scales, from local scale to national one. The ultimate goal is to present accurate data to the European scale producing assessments for policy makers in the European Commission.

The integration of model processed data results in mapping soil functions is important in the context of the Soil Thematic Strategy [5]. The successful output of the aggregation model will be data which can be mapped in to soil threats as those described in the proposed Soil Thematic Strategy. Particular interest will be given to organic carbon decline, soil erosion, and soil compaction as they will lead to changed soil parameters. Those changes will have effects on soil functions as production of biomass and matter transformation including storing and filtering. In this way hot spots where soil functions will most likely be degraded below a level required for a sustainable land use will be identified.

This model application will improve the current status of updating the European Soil Information System and the European Soil Data Center (http://esdac.jrc.ec.europa.eu). Based on the model requirements, a data protocol will be defined for compilation, formatting, and interchange of data. The data collection will not only be useful for running the model but will indicate if data providers can easily comply with a set of requirements and thus point out the feasibility, in the Soil Thematic Strategy context, for member states authorities to collect such data.

3 Application area: Slovakia multi-scale soil information system

The paper uses the outputs of the case study implemented in Slovakia in 2006 and the resulting Slovakia Soil Database [20]. The Joint Research Centre (Ispra, Italy) has collaborated with the Soil Science and Conservation Research Institute (Bratislava) in order to develop the Slovakia Multi-scale Soil Information System. In MEUSIS, all geographical information (attributes and geometry components) are represented by the grid of regular spatial elements (pixels). The representation of various spatial resolution details is following the INSPIRE recommendations and three spatial resolution levels of geographical information have been defined for the application of MEUSIS: coarse resolution, 10 km^2; medium resolution, 5 km^2; fine resolution, 1 km^2.

The selection of the spatial units (1, 5, and 10 km^2) is proposed for the development of Spatial Data Infrastructures (SDIs) compliant with INSPIRE Directive. Alternative spatial unit size can impose problems in the interoperability of the output datasets since they may not fit to standard SDIs requirements. Having built the Slovakia Soil Database in this way, a system of spatially nested hierarchical grids has been developed.

In Fig. 2, the three illustrated spatial resolutions correspond to the three levels of data collection:

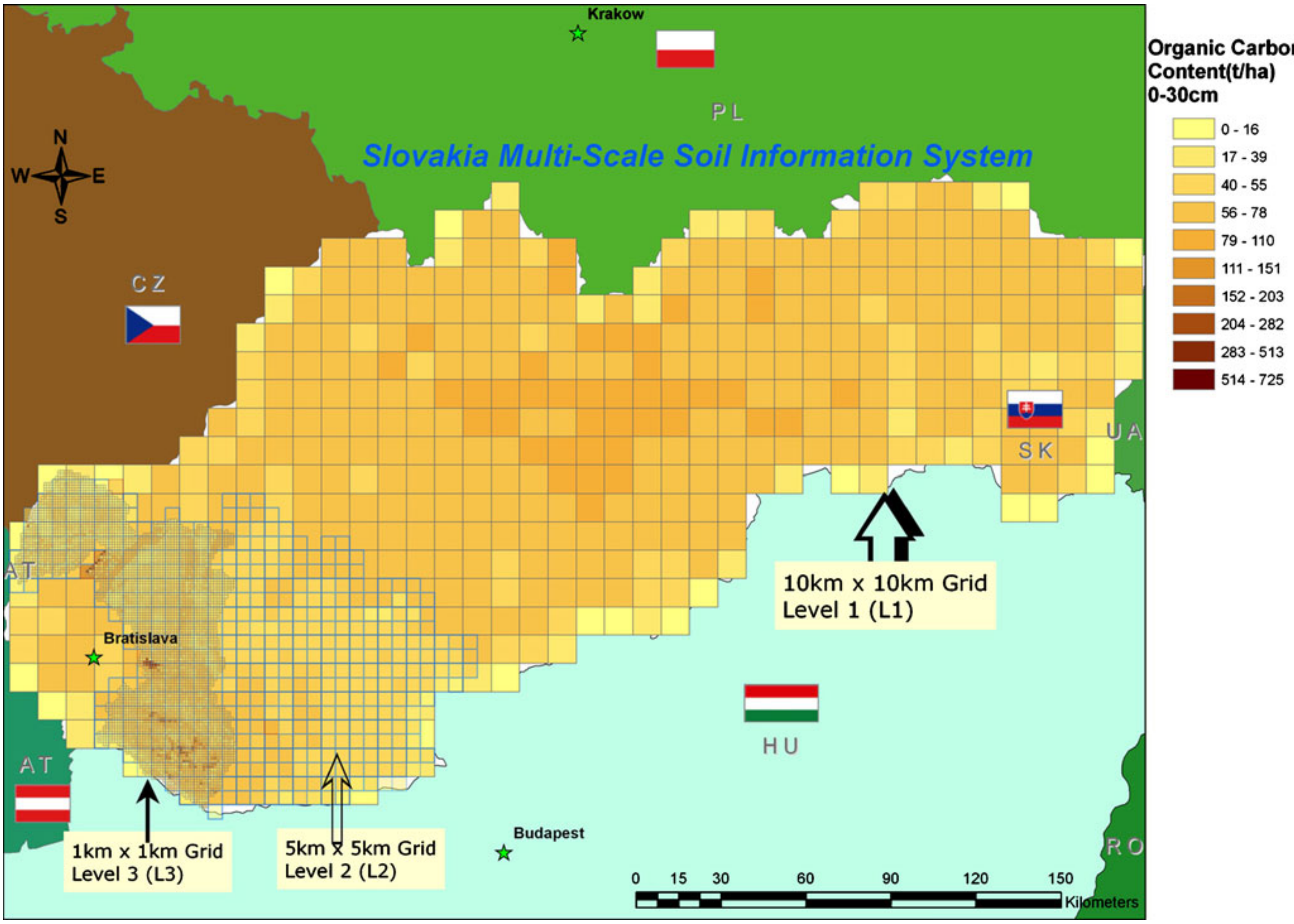

Fig. 2 Slovakia MEUSIS

- Level 1 (L1): 509 cells of 10 × 10 km covering an area of 49,026 km^2 (whole Slovakian surface).
- Level 2 (L2): 475 cells of 5 × 5 km covering an area of 10,489 km^2 (districts of Trnava and Nitra).
- Level 3 (L3): 4,409 cells of 1 × 1 km covering an area of 4,409 km^2 (only district of Trnava).

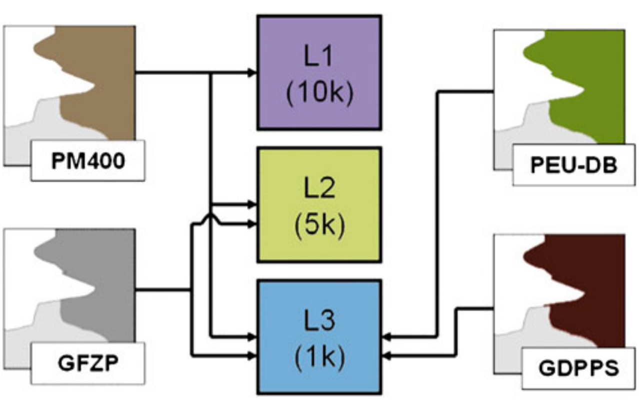

Fig. 3 Input data allocation scheme

3.1 Data sources

In order to fill in the database in the three resolutions, the data provider has used the following information sources:

- Digital soil map of Slovakia at scale 1:400,000 (PM400)
- Digital geo-referenced database of soil ecological units at scale 1:5,000 (PEU-DB)
- Set of digital regional maps of geo-factors of landscapes—soils at scale 1:50,000 (GFZP)
- Georeferenced database of agricultural soils in Slovakia (GDPPS)

In Fig. 3, the allocation scheme per level (L1, L2, and L3) is presented. Some basic allocation rules have been defined in order to complete the data input according to the most accurate pixel scale. So, each pixel should be filled with the most detailed information source, but in case of lack of information the next more general information source provides the data value.

4 Application of upscaling methodology

In order to describe the upscaling methodology, an input data field such as the organic carbon (OC) content in the surface horizon 0–30 cm (quantitative variable) of the Slovakia Soil Database will be used. The organic carbon is measured as tons/hectare according to the following equation:

$$\mathrm{OC}\,(\mathrm{t/ha}) = \mathrm{Cox} \times \mathrm{BD} \times d,$$

where

Cox (%) is the average content of organic carbon for topsoil/subsoil,
BD (g/cm^3) is the average soil bulk density for topsoil/subsoil, d (cm) is the volume of topsoil/subsoil.

Assessing changes of soil carbon due to land management practices, land use and climate change plays a crucial role in the fields of climate change mitigation (soil carbon sequestration), land–atmosphere interactions, and soil fertility. Soil organic carbon is an important soil component as it influences soil structure and aggregation, soil moisture conditions, soil nutrient status, and soil biota, and hence influences ecosystem functioning [15].

4.1 Upscaling methodology

Aggregating functions in the upscaling methodology and spatial data process will be done using ArcGIS software. In order to prove the utility and applicability of this nested approach, the technical steps of the upscaling process from 5 to 10 km^2 are described in the next flowchart (Fig. 4). Find below the aspects that determine the successful results of this operation:

- The Aggregation Technique establishes how the value of each output cell will be determined. In this aggregation exercise, the statistical function "Mean" is used as the most appropriate technique.
- The Factor by which to multiply the cell size of the input raster in order to obtain the desired resolution. Since the desired resolution is 10 and the input resolution is 5, then the cell factor is 2 as it should satisfy the following equation: Output cell size = input cell size × cell factor

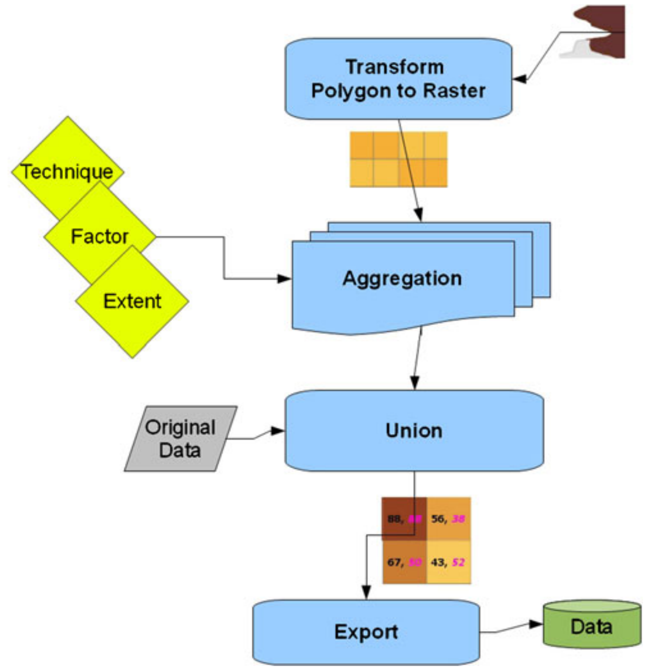

Fig. 4 Technical steps in the upscaling process

- The Extent of the output area should be the same as the Extent of original data. Each output cell should fit with the original cell of the same size in order to allow further process.

4.2 Upscaling results from 5 × 5 km grid towards the 10 × 10 km grid

Using the above described upscaling methodology, the 475 cells of 5 km^2 yield 134 cells of 10 km^2. As a rule described in the INSPIRE principles, 4 cells of 5 × 5 km grid cell are requested in order to upscale their value to a cell of 10 × 10 km using the statistical mean process. In this case, the additional cells of the upscaled data are due to the fact that less than 4 may be upscaled to 1 cell of coarser resolution (cases near the borders, upper left, or down right corners).

In Fig. 5, the reader may observe both the upscaled data as a result of the above described process and the original data (transparency has been applied). The upscaled data of 10 km^2 scale are located in the center of the cell and have a blue font over the source data of 5 km^2 scale in light grey font. The use of transparency allows the reader to justify the correspondence between source and resulting data. In the majority of the cases the upscaled data is the mean of four input data values of 5 × 5 km cells(except the cases near borders where less than four cells can be upscaled to 1).

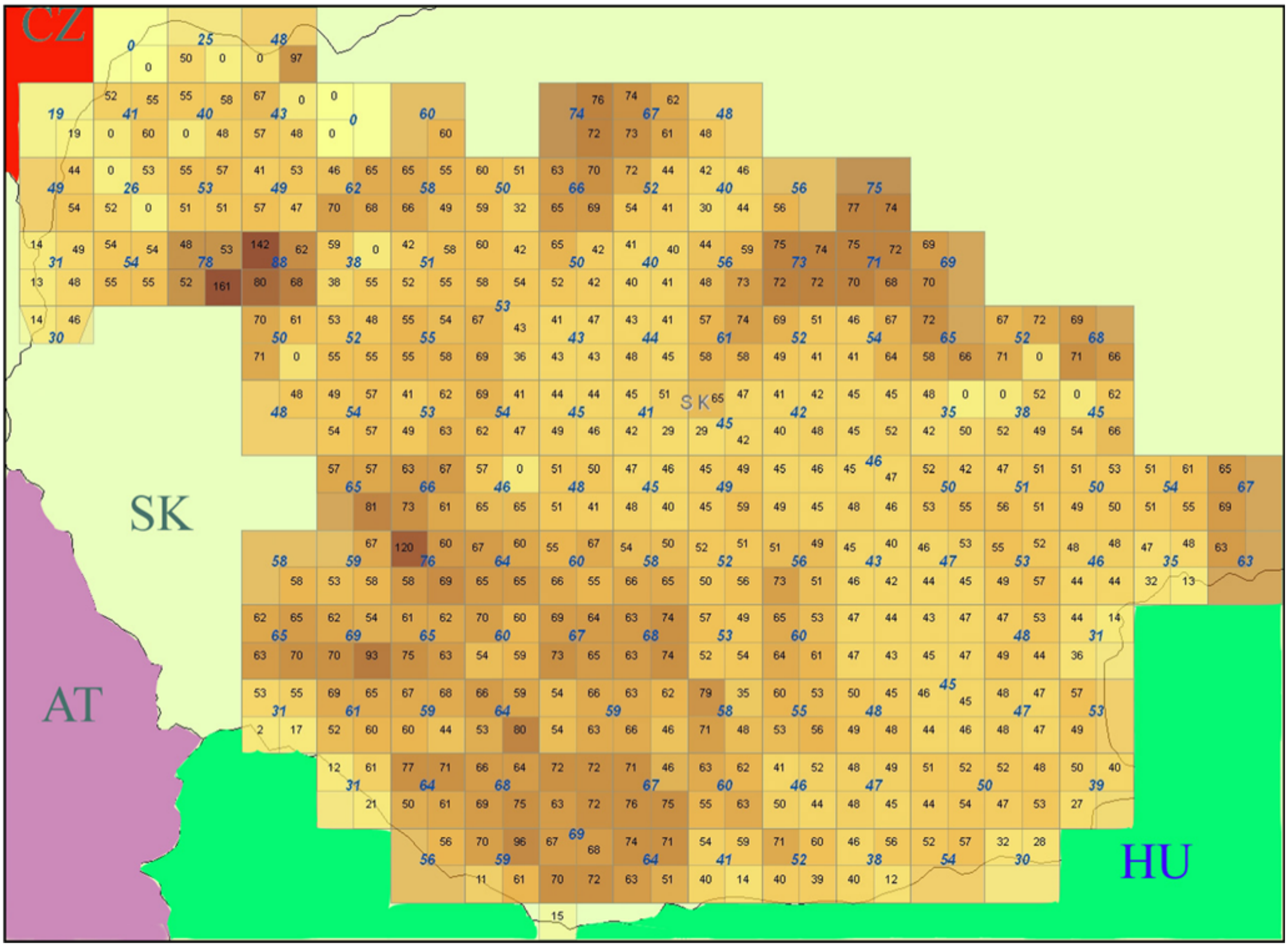

Fig. 5 Upscaled results from 5 km towards 10 km

Next recommended GIS operation is to union the cells that fit both in the original and in the upscaled data (result, 131 cells). The graphical representation of this comparison may be viewed in Fig. 6 as both values can be visualized for comparison. In each of the 131 cells, the reader may view both the original value (first one in black font) and the upscaled one (second one in red italic font). All the steps in the upscaling methodology are documented with the attached figures for the better conceptual understanding of the methodology.

The present analysis may be applied also in order to identify cases where the data provider has previously performed the "tricky" operation well-known as downscaling. Acting as an organization which collects data in Pan-European level in smaller scales, we can apply data quality mechanisms checking if the data providers have contributed with their original data or they have manipulated their data by distributing their smaller scale data to larger scale ones (downscaling). It is fairly easy for the data manager to conclude this test by viewing the following data pattern image.

As a next step, the scatter diagram (Fig. 7) has been implemented for the better data relationship pattern identification. In this diagram, the reader may compare the correspondence between the original 10 × 10 km values (X-axis) and the upscaled data (Y-axis). It is obvious that there is a linear relationship between the two datasets as there is a major concentration of data values near a line. A more detailed statistical analysis will follow in order to take into account all the statistical descriptors and make comprehensible the conclusions of this comparison.

The proposed approach is not just a theoretical one, but proved to be applied in specific cases proposing significant results. In the past, there were many theoretical references to an ideal MEUSIS as a nested system of hierarchical grids while in this paper, we describe the results of the applied MEUSIS methodology using

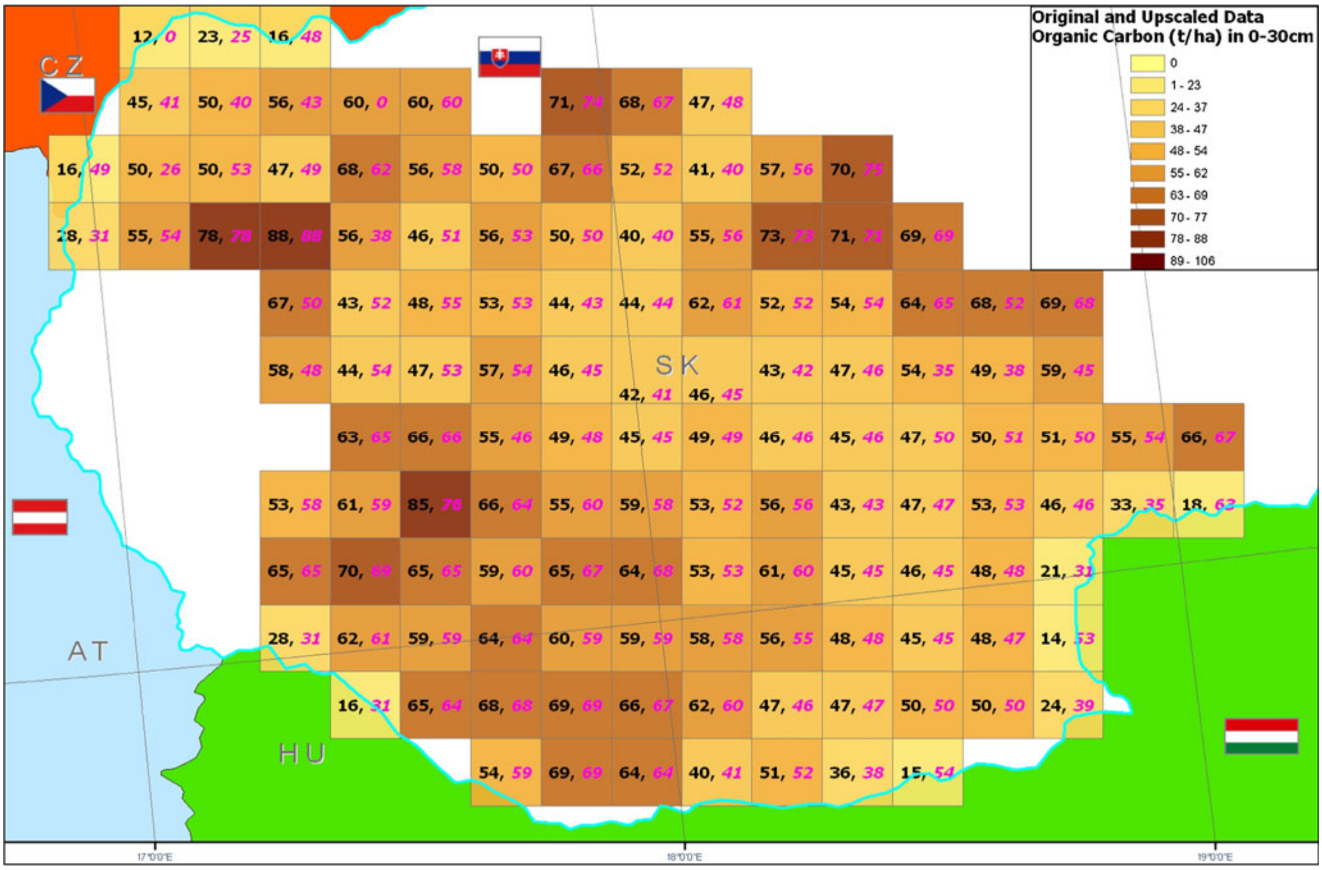

Fig. 6 Comparison of original and upscaled data

Fig. 7 Scatter diagram of the original and upscaled data

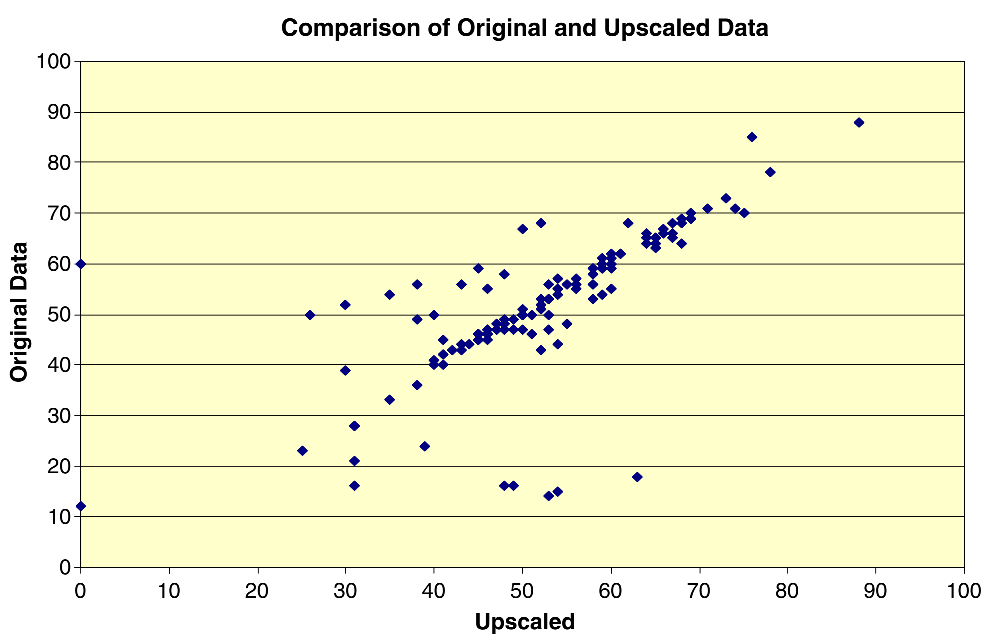

Table 1 Descriptive statistics of the upscaling process from 5 km^2 towards 10 km^2

Statistic	Upscaled data	Original data
Sum	6,825	6,819
Count (cells)	131	131
Mean	52.10	52.05
Maximum	88	88
Minimum	0	12
Median	52.00	53.00
Standard deviation	13.27	14.29
Mode	53.00	50.00
Coefficient of kurtosis	2.60	1.05
Coefficient of skewness	−0.76	−0.72
Coefficient of correlation	0.693	

both GIS operations and statistical analysis (descriptors, scatter diagram) in datasets originated from a European Country.

In Table 1, a more detailed statistical analysis is presented and some descriptive statistics are shown in order to allow the better understanding of the two data series. Commenting the results of this analysis, we notice the following remarks:

- The mean of both data sets are very close and this result may be explained in the following ways:
  - either the data sources for both the 10 km^2 values and the 5 km^2 values are the same
  - or the use of abovementioned upscaling method is producing satisfactory results as the upscaled data are close to the original data
- The median which represents the value for which half of the data are below this value is very close for both data distributions. Moreover, both median values are close to the mean values which can be considered a positive fact [8]
- The mode value which represents the most frequent value in data series is also quite close for both data distributions
- Taking into account the three abovementioned measures of tendency (mean, median, and mode), we conclude that there are no extreme values that can affect the data distribution. Having in mind all those statistical indicators, we can certify that there is a small-medium variability regarding the organic carbon content in this scale (5 × 5 km).
- The standard deviation measures the dispersion of the data around the mean. Regarding the prediction of intervals, it is observed that the distribution of the upscaled data tends to be normal and as a consequence, we may use the standard normal distribution in order to estimate with 95% probability the range in which a random upscaled data value X will be included according to the Eq. 1:

$$P(-1.96\sigma \leq \mathrm{X} - \mu \leq 1.96\sigma) = 0.95 \tag{1}$$

Where μ and σ are the mean and the standard deviation. Resolving the Eq. 1 for the upscaled data distribution, we determine the range of possible values for the parameter organic carbon content

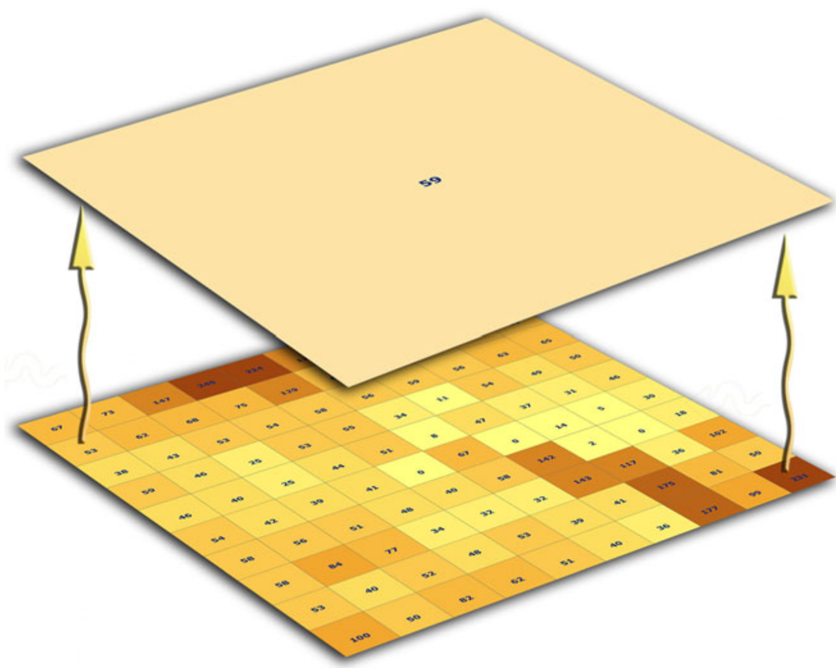

Fig. 8 Upscaling process from 1 × 1 km towards 10 × 10 km grid

0–30 cm will vary between 26.10 and 78.10 with a probability of 95%.

- The coefficient of kurtosis determines the kurtosis of the data distributions and in both cases this indicator is less than 3 which means that the both distributions are platykurtic (flat shape). In the upscaled data, the coefficient of kurtosis is close to 3 (the distribution tends to be normal; mesokurtic)
- The coefficient of skewness in both cases is negative which means that the left tail of the distribution curve is longer (flatter).
- One of the most important indicators is the coefficient of correlation which determine how strong the relationship between the two data distributions. The value 0.693 determines a quite strong relationship between the two data distributions (it is also obvious from the scatter diagram in Fig. 7).

In the abovementioned statistical analysis, we could continue by performing a regression analysis but it is not the main objective of this paper to reach such a conclusion.

4.3 Upscaling results from 1 × 1 km towards the 10 × 10 km

Using the upscaling methodology, we will go a step further and an analysis of upscaling from 1 × 1 km towards

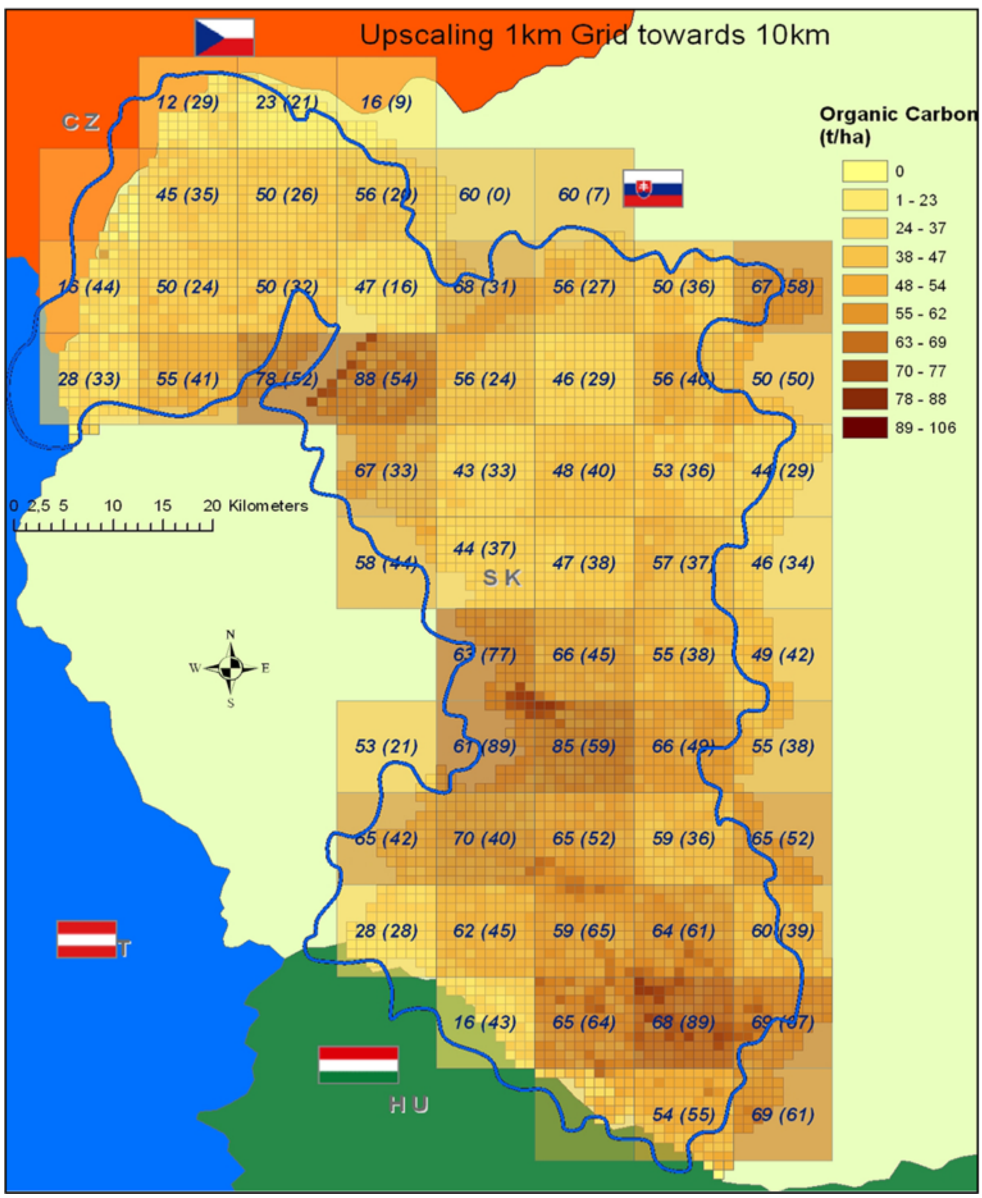

Fig. 9 Comparison of original and upscaled data (1 km^2 towards 10 km^2)

the 10 × 10 km will be presented below. In order to update one cell of 10 × 10 km, it is requested 100 cells of 1 × 1 km (Fig. 8). In the Slovakia Soil Database, there are available 4,409 cells which are upscaled to 61 cells of 10 km (there are cases where less than 100 cells may be upscaled to 1 cell of the upper scales).

The methodology applied is the one which has been described above and the results are shown in Fig. 9. After executing the Union operation (joining two datasets), the 61 common cells (original 10 × 10 km and upscaled ones) are compared in Fig. 9 where the reader may view both data values. The blue outline reflects the borders of the area where data in scale of 1 km^2 have been provided.

Proceeding with the statistical analysis, the statistical descriptors are compared in the Table 2:

- There is a slightly significant difference between the two means (40.20 and 53.64) which may be explained as the upscaled data tend to have lower values than the original ones due to high dispersion of original data.
- Regarding the median and the mode, there is even a larger difference between the two datasets due to the trend of the upscaled data to have lower values.
- The coefficient of skewness is positive in the upscaled data (right tails is longer) and negative for the original data (left tail is longer).
- The value range of upscaled data can be estimated with probability 95% and as a result a random upscaled value may vary between 6.02 and 74.38 according to the Eq. 1.
- The coefficient of correlation has a value of 0.449 which express a medium relationship (neither strong, nor weak) between the two data distributions.
- Comparing the upscaling process from 1 km^2 towards 10 km^2 with the one from 5 km^2, we have noticed some differences. The results produced in the case of 1 km^2 upscaling can be considered satisfactory as it is more possible to have a better result in a certain area of 10 km^2 by upscaling 100 data values of 1 km than having one random value in this large area. It can be considered that the average of 100 samples is more accurate than the random sample in a 10 km^2 area.

Table 2 Descriptive statistics of the upscaling process from 1 km^2 towards 10 km^2

Statistic	Upscaled data	Original data
Sum	2,452	3,272
Count (cells)	61	61
Mean	40.20	53.64
Maximum	89.00	88.00
Minimum	0.00	12.00
Median	38.00	56.00
Standard deviation	17.44	15.70
Mode	28.00	50.00
Coefficient of kurtosis	1.03	1.10
Coefficient of skewness	0.58	−0.81
Coefficient of correlation	0.449	

5 MEUSIS responses to limitations of current soil data availability

The data availability problem rises since the soil observations are sparse and are made on relatively small geometric support but extensive resources are required for making larger regional assessment. Due to high sampling and laboratory analysis costs [7], we base our input data in current samples or expert judgments which are representative of soil mapping units. MEUSIS can be applied to cases where based on few representative data, we have to respond in fast and efficient way to policy makers demands and to data requests coming from the other scientific communities where soil data are needed as input for modeling purposes. This capability will contribute significantly to the policy and stakeholder-oriented information, for example, assessing soil indicators that can be included in eco-services for EU citizens helping to improve their quality of life for environment and security [12].

MEUSIS may receive critics by traditional soil scientists since there are cases where spatial soil diversity and landscape dynamics [25] cannot be represented in a satisfactory way by such an information system. The way proposed to face the problem of huge natural variability is to estimate the "Descriptive Statistical Indicators" as they propose a satisfactory overview of the data variability and estimate the uncertainty of upscaled data. The statistical analysis supports the reader to better recognize the natural variability of soils across the landscapes. The numerical processed results are not only accurate in mathematical terms but represent somehow the reality in the ground since the reader takes into account also the other statistical descriptors (standard deviation, coefficient of correlation).

INSPIRE will enable the sharing of environmental spatial information among public sector organizations and better facilitate public access to spatial information across Europe. This ambitious objective will be achieved through the establishment of integrated spatial information services, based upon a distributed network of databases, linked by common standards and protocols to ensure compatibility. MEUSIS approach faces the problems connected with incompatibility, low accessibility, data gaps, duplication of collection, varying standards, and low level of coordination [23] in the following ways:

- MEUSIS reference grids may be proposed as a standard for soil data collection in various scales. The soil data should be collected once and maintained at the level where this can be done most effectively. MEUSIS promotes the maintenance of the original soil data in local/regional level while it facilitates data sharing to other policy makers institutions in different levels (region, country, and European Union).
- The MEUSIS proposed spatial units which are INSPIRE compliant allow the soil data to be interoperable with other spatial data coming from different sources and used for modeling purposes. In such way, soil indicators such as soil organic carbon can be interoperable and used easily by scientific communities for which there is an increasing demand of soil data (climate change, biodiversity, energy, food, and water).
- Soil data management and coordination is facilitated with the use of harmonized grids since MEUSIS proposes a common standard with specific rules for data collection. The common standards include specifications for grid system, co-ordinate transformation, labeling, spatial resolutions, data exchange format, and metadata information.

6 Conclusions and outlook

- The MEUSIS nested grids approach can be proposed as a solution in many cases where the data owner does not allow the distribution/publication of detailed data but is willing to distribute generalized data (in coarser resolution). MEUSIS can be considered a valuable methodology which contributes to the generalization (without losing the real values) of very detailed data and may allow the scientific community to access valuable information without having any copyright problems.
- Upscaling has a serious drawback in case the source dataset in the larger scale has high spatial variability. This has been shown in the upscaling process from 1 × 1 km towards the 10 × 10 km. In a latter stage, we should investigate the combination of statistical descriptors in the upscaling process in order to better calibrate this model process.
- Since national or regional soil information are organized in a complex databases with considerable limitations (high variety in data availability, representation differences, various scales, different attribute precision, and data inconsistency), the MEUSIS methodology may be proposed as a pipeline in order to upscale soil information from this detailed level towards European level. MEUSIS may overcome all the abovementioned limitations and problems. This implies that high amount of information at national and regional level can be used in order to improve the European scale datasets.
- Upscaling has a real positive effect whenever the visual representation of very fine resolution maps is difficult both for viewing purposes and calculation procedures (example of 20 × 20 m European Soil Sealing Map). MEUSIS can be proposed as a good data management practice in different decision making levels (e.g., the technical officer of a region would need different scale of information about soil organic carbon decline than the policy officer in the European Commission).
- MEUSIS approach is significantly different than conventional polygon maps as it is pixel-based. The output results may be easily manipulated, imported in models, and used by other earth and social sciences. The methodology can be used also for other environmental indicators (water quality, air pollution, etc.) which have small or medium spatial variability.
- MEUSIS contributes significantly to the recent policy developments (INSPIRE, Soil Thematic Strategy) and overcomes obstacles of the limited data availability.

Acknowledgements The data for the application of MEUSIS methodology were provided by the Soil Science and Conservation Institute in Slovakia.

References

1. Beven, K.J.: Changing ideas in hydrology—the case of physically-based models. J. Hydrol. **105**(1–2), 157–172 (1989)
2. Bierkens, M.F.P., Finke, P.A., De Willigen, P.: Upscaling and Downscaling. Methods for Environmental Research. Kluwer Academic Publishers, Dordrecht (2000)
3. Bouma, J., Hoosbeek, M.R.: Obtaining soil and land quality indicators using research chains and geostatistical methods. Nutr. Cycl. Agroecosyst. **50**, 35–50 (1998)
4. Camarero, L., Garcia-Pausas, J., Huguet, C.: A method for upscaling soil parameters for use in a dynamic modelling assessment of water quality in the Pyrenees. Sci. Total Environ. **407**(5), 1701–1714 (2009)
5. CEC: Soil Thematic Strategy (COM(2006) 232) of the European Commission (2006) http://eusoils.jrc.ec.europa.eu/

ESDB_Archive/Policies/Directive/com_2006_0231_en.pdf, Accessed 05/10/2010
6. Cheng, Q.: Modeling local scaling properties for multiscale mapping. Vadose Zone J. **7**(2), 525–532 (2008)
7. Christy, C.D.: Real-time measurement of soil attributes using on-the-go near infrared reflectance spectroscopy. Comput. Electron. Agric. **61**(1), 10–19 (2008)
8. Dikmen, O., Akin, L., Alpaydin, E.: Estimating distributions in genetic algorithms. comput. Inf. Sci. **2869/2003**, 521–528 (2003)
9. Ewert, F., van Keulen, H., van Ittersum M.K., Giller, K.E., Leffelaar, P.A., Roetter, R.P.: Multi-scale analysis and modelling of natural resource management options. In: Proceedings of 3rd Biennial meeting of the International Environmental Modelling and Software Society (2006)
10. Heuvelink, G.B.M., Pebesma, E.J.: Spatial aggregation and soil process modelling. Geoderma **89**, 47–65 (1999)
11. Hewett, C., Quinn, P., Heathwaite, A., Doyle, A., Burke, S., Whitehead, P., Lerner, D.: A multi-scale framework for strategic management of diffuse pollution. Environ. Model. Softw. **24**, 74–85 (2009)
12. GMES, Global Monitoring for Environment and Security. http://www.gmes.info/ Accessed 05/10/2010
13. INSPIRE: Directive 2007/2/EC of the European Parliament and of the Council (14 March 2007): establishing an Infrastructure for Spatial Information in the European Community (2007), http://inspire.jrc.ec.europa.eu/ Accessed 05/10/2010
14. Kokkonen, T., Koivusalo, H., Laurén, A., Penttinen, S., Starr, M., Kellomäki, S., Finér, L.: Implications of processing spatial data from a forested catchment for a hillslope hydrological model. Ecol. Model. **199**(4), 393–408 (2006)
15. Lal, R.: Soil Carbon sequestration impacts on global climate change and food Security. Science **304**, 1623–1627 (2004)
16. Lembo, A.J., Lew, M.Y., Laba, M., Baveye, P.: Use of spatial SQL to assess the practical significance of the modifiable areal unit problem. Comput. Geosci. **32**, 270–274 (2006)
17. McBratney, A.B.: Some considerations on methods for spatially aggregating and disaggregating soil information. Nutr. Cycl. Agroecosyst. **50**, 51–62 (1998)
18. Nachtergaele, F.: The "soils" to be classified in the world reference base for soil resources. Eurasian Soil Sci. **38**(SUPPL. 1), S13–S19 (2005)
19. Panagos, P., Van Liedekerke, M., Lado Rodriguez, L., Montanarella, L.: MEUSIS: multi-scale european soil information system. GEOconnexion International Magazine **7**(2), 39–41 (2008)
20. Skalský, R.: Multiscale European soil information system—pilot project for Slovakia. In: Proceedings No. 28, 2006 Soil Science and Conservation Research Institute, Bratislava, pp 89–98 (2006)
21. Stewart, J.B., Engman, E.T., Feddes, R.A., Kerr, Y.H.: Scaling up in hydrology using remote sensing: summary of a Workshop. Int. J. Remote Sens. **19**(1), 181–194 (1998)
22. Thwaites, R.N., Slater, B.K.: Soil-landscape resource assessment for plantations—a conceptual framework towards an explicit multi-scale approach. For. Ecol. Manag. **138**(1–3), 123–138 (2000)
23. Tuchyna, M.: Establishment of spatial data infrastructure within the environmental sector in Slovak Republic. Environ. Model. Softw. **21**(11), 1572–1578 (2006)
24. Van Bodegom, P.M., Verburg, P.H., Stein, A., Adiningsih, S., Denier van der Gon, H.A.C.: Effects of interpolation and data resolution on methane emission estimates from rice paddies. Environ. Ecol. Stat. **9**(1), 5–26 (2002)
25. Wilding, L.P., Lin, H.: Advancing the frontiers of soil science towards a geoscience. Geoderma **131**(3–4), 257–274 (2005)
26. Wood, B.: The role of scaling laws in upscaling. Adv. Water Resour. **32**(5), 723–736 (2008)