Finding the right pixel size

Tomislav Hengl*

European Commission, Directorate General JRC, Institute for Environment and Sustainability, Soil and Waste Unit, TP 280, Via E. Fermi 1, I-21020 Ispra (VA), Italy

Received 6 May 2005; received in revised form 23 November 2005; accepted 24 November 2005

Abstract

This paper discusses empirical and analytical rules to select a suitable grid resolution for output maps and based on the inherent properties of the input data. The choice of grid resolution was related with the cartographic and statistical concepts: scale, computer processing power, positional accuracy, size of delineations, inspection density, spatial autocorrelation structure and complexity of terrain. These were further related with the concepts from the general statistics and information theory such as Nyquist frequency concept from signal processing and equations to estimate the probability density function. Selection of grid resolution was demonstrated using four datasets: (1) GPS positioning data—the grid resolution was related to the area of circle described by the error radius, (2) map of agricultural plots—the grid resolution was related to the size of smallest and narrowest plots, (3) point dataset from soil mapping—the grid resolution was related to the inspection density, nugget variation and range of spatial autocorrelation and (4) contour map used for production of digital elevation model—the grid resolution was related with the spacing between the contour lines i.e. complexity of terrain. It was concluded that no ideal grid resolution exists, but rather a range of suitable resolutions. One should at least try to avoid using resolutions that do not comply with the effective scale or inherent properties of the input dataset. Three standard grid resolutions for output maps were finally recommended: (a) the coarsest legible grid resolution—this is the largest resolution that we should use in order to respect the scale of work and properties of a dataset; (b) the finest legible grid resolution—this is the smallest grid resolution that represents 95% of spatial objects or topography; and (c) recommended grid resolution—a compromise between the two. Objective procedures to derive the true optimal grid resolution that maximizes the predictive capabilities or information content of a map are further discussed. This methodology can now be integrated within a GIS package to help inexperienced users select a suitable grid resolution without doing extensive data preprocessing.

Keywords: Grid resolution; Scale; Inspection density; Point pattern analysis; Variogram; Terrain complexity

1. Introduction

A grid cell, popularly known as pixel, is the fundamental spatial entity in a raster-based GIS (Gatrell, 1991; DeMers, 2001). Although there is practically no difference between pixel and grid cell, geoinformation scientists like to emphasize that pixel is a technology and grid is a model (De By,
grid means ideal properties—orthogonal matrix, fixed resolution, which a raster image does not necessarily has to fit. For example, an aerial photo first needs to be ortho-rectified and then resampled to a regular grid to (approximately) fit the grid model (Rossiter and Hengl, 2002).

Grid cell can be also related (but should not be confused) with the support size, which is typically a fixed area or volume of the land that is being sampled. Support size can be increased by using composite samples or by averaging point-sampled values belonging to the same blocks of land. In geostatistics, one can also control the support size of the output models by averaging multiple predictions per regular blocks of land, which is known as “block kriging” (Heuvelink and Pebesma, 1999). This means that we can sample at point locations, then make predictions for blocks of 10 × 10 m. The latter often confuses GIS users because we can produce predictions at regular point locations (point kriging) and then display them using a raster map, but we can also make predictions for blocks of land (block kriging) and display them using the same raster model (Bishop et al., 2001). This distinction is especially important for the validation of the spatial prediction models because it can lead to serious misconceptions—validating a point model (support size of few centimeters) at 1 km support or vice versa can be quite discouraging (Stein et al., 2001).

Although the raster structure has a number of serious disadvantages such as of under- and oversampling in different parts of the study area and large data storage requirements, it will remain the most popular format for spatial modelling in the coming years (DeMers, 2001). What makes it especially attractive is that most of the technical characteristics are controlled by a single measure: grid resolution, expressed as ground resolution in meters. The enlargement of grid resolution leads to aggregation or upscaling and decrease of grid resolution leads to disaggregation or downscaling. As grid becomes coarser, the overall information content in the map will progressively decrease and vice versa (McBratney, 1998; Kuo et al., 1999; Stein et al., 2001). In cartography, coarser grid resolutions are connected with smaller scales and larger study areas, and finer grid resolutions are connected with larger scales and smaller study areas. The former definition often confuses non-cartographers because bigger pixel means smaller scale, which usually means larger study area (Fig. 1). Note in Fig. 1, that both aggregation and disaggregation can be done before or after geo-computation. If the

![Diagram](image-url)

Fig. 1. Upscaling and downscaling in a grid-based GIS. $S$ indicates scale: $S^-$ are smaller scales and $S^+$ are larger scales. Based on McBratney (1998).
model is linear, the two routes should yield the same results (Heuvelink and Pebesma, 1999); if not, there can be serious differences. Aggregation is fairly useful procedure to reduce the small scale variation and get better idea about the general pattern (Stein et al., 2001). In contrast, if our objective is to locate extremes (hot spots), then aggregation is something we should avoid.

Many researchers investigated the effects of grid resolution on the accuracy of their models. Applications range from mapping of soil properties (Florinsky and Kuryakova, 2000), modelling of surface runoff (Kuo et al., 1999; Molnr and Julien, 2000), sea currents (Davies et al., 2000) or modelling of meteorological data (McQueen et al., 1995; Noda and Niino, 2003). In the case of terrain data, aggregation of grid resolutions will seriously deteriorate accuracy of terrain parameters (Weihua and Montgomery, 1994; Dietrich et al., 1995; Thompson et al., 2001). Wilson et al. (2000) demonstrated the impact the grid resolution makes on hydrological analysis: as the grid resolution increased from 30 to 200 m the flow-path length and specific catchment area maps changed drastically. The grid resolution can also be crucial for the accuracy of the simulation model such as surface runoff or erosion models (Sánchez Rojas, 2002). Kienzle (2004) gives a systematic overview of effects of various grid resolutions on the reliability of terrain parameters suggesting finer grid resolutions from 5–20 m. Liang et al. (2004), on the other hand, observed impacts of different spatial resolutions on modelling surface runoff and concluded that the resolution needs to be improved only to a critical level after which the model will not necessarily perform better. Weihua and Montgomery (1994) also got a substantial improvement with 10 m grid resolution over 30 and 90 m data, but 2 or 4 m data gave only marginal improvements. Florinsky and Kuryakova (2000) focused specifically on the importance of grid resolution of terrain parameters on the efficiency of spatial prediction of soil variables. They plotted correlation coefficients versus different grid resolutions and looked for the grid size with most powerful prediction efficiency. Bishop et al. (2001) suggested use of the Shannon’s information criterion to select the optimal block size for block kriging. All these experiments clearly prove two things: (1) grid resolution plays an important role for the efficiency of the mapping and (2) its selection can be optimized, to a certain level, to satisfy both processing capabilities and representation of spatial variability.

Although much has been published on the effect of grid resolution on the accuracy of spatial modelling, choice of grid resolution is seldom based on the inherent spatial variability of the input data (Vieux and Needham, 1993; Bishop et al., 2001). In fact, in most GIS projects, grid resolution is selected without any scientific justification. In the ESRI’s package ArcGIS, for example, the default output cell size is suggested by the system using some trivial rule: in the case the point data is being interpolated in Spatial Analyst, the system will take the shortest side of the study area and divide it by 250 to estimate the cell size (ESRI, 2002). Obviously, such pragmatic rules do not have a sound scientific background.

This motivated me to produce methodological guides to select a suitable grid resolution for output maps based on the inherent properties of the input data. I tried to relate the choice of grid resolution to measurable cartographic and statistical concepts such as: scale, processing power, positional accuracy, inspection density, spatial dependence structure and complexity of terrain. I will first recommend some general rules of thumb to select the grid resolution and then demonstrate how to select a legible grid resolution given the real datasets.

2. Methods

2.1. Grid resolution and cartographic concepts

Although we live in a digital era where we do not necessarily work with hard copy maps, spatial resolution and extent are still strongly related with the traditional cartographic concepts (Quattrochi and Goodchild, 1997; Goodchild, 2001). For example, in traditional soil cartography the scale of an existing map is commonly assessed by estimating either the maximum location accuracy (MLA) or average size area (ASA) of the polygons on the ground (Rossiter, 2003). These cartographic definitions can also be used to estimate the suitable grid resolution for a given mapping scale. As a rule of thumb, Rossiter (2003) suggests that four grid cells can be considered equivalent to the minimum legible delineation (MLD), which is the smallest size area that we map. According to the definition of Vink (1975) the MLD is 0.25 cm$^2$ on the map, so the suitable grid resolution can estimated based on
the scale number \((SN)\):

\[
p \leq \sqrt{\frac{MLD}{4}} = \sqrt{\frac{SN^2 \cdot 0.000025}{2}} = SN \cdot 0.0025
\]  

(1)

where \(p\) is the grid (pixel) size and MLD is the minimum legible delineation area on the ground in \(m^2\). This means that for a 1:50K scale, MLD is 6.25 ha and suitable grid resolution is 125 m, which seems fairly coarse. Somewhat larger grid resolutions from 0.5 to 3 mm on the map have been also recommended by Valenzuela and Baumgardner (1990).

The grid resolution can also be related to the MLA, which commonly ranges from 0.25 mm to maximum of 0.1 mm on the map (Vink, 1975). This gives the smallest legible resolutions

\[
p \geq SN \cdot MLA = SN \cdot 0.00025 \cdot (0.0001)
\]  

(2)

So for 1:50K scale, the smallest legible grid resolution is 12.5 m (5 m). Resolutions finer than 5 m truly do not make sense as it will be hard to visualize or print them at this scale of work. Following notions in microscopy, and the Nyquist frequency concept from signal processing (Shannon, 1949), which states that the original signal can be reconstructed if sampling frequency is two times the original frequency, McBratney et al. (2003, Table 1), suggested that there should be at least \(2 \times 2\) pixels to represent smallest rounded objects of interest and at least two pixels to represent the width of elongated objects. The smallest objects are typically of size \(1 \times 1\) mm on the map, so that the grid resolution can be determined using the \(p = 0.5\) mm rule (Fig. 2a). The former can be used as the universal rule of thumb to relate scale with grid resolution.

2.2. Grid resolution and computer processing power

The grid resolution can also be related with the size of area and processing power of our computer. Although we might insist on using the finest grid resolution possible, the calculation time will increase exponentially (cubically) with the total number of pixels in a map (McBratney, 1998). This means that grid resolution needs to match capabilities of our computer and time given to complete a GIS project. Following the popular Moore’s law, Lagacherie and McBratney (2005) discussed the relationship between the grid resolution and processing capabilities of standard desktop PC’s and discovered the following rough relationship between the log of the image size and the current year (Fig. 3a):

\[
\log_{10}(m) = 0.14 \cdot (Y - 1955)
\]  

(3)

where \(\log_{10}(m)\) is the logarithm of the image size in pixels and \(Y\) is the current year. According to this formula, the standard image size in year 2005 would be about \(10^7\) pixels, which makes an image of 3000 \(\times\) 3000 pixels.

The grid resolution can also be estimated by dividing the size of the study area by the number of pixels that computer can handle. For example, if the size of area is about 100 000 km\(^2\) and computer can handle \(10^7\) pixels, we should probably work with resolutions of 100 or more meters. Based on this principle, we can also derive the coarsest global standard resolution. This is the resolution of an

---

Fig. 2. Popular cartographic rules to select grid resolution: (a) relationship between grid resolution and cartographic scale; (b) relationship between grid resolution and density of observation points for soil mapping applications.
image covering the whole Globe. The surface of Earth is about $5.10 \times 10^8 \text{ km}^2$ (Yoder, 1995), which means that the coarsest global standard resolution in the year 2005 is about 7 km. If the Eq. (3) is correct, the coarsest global standard resolution in year 2040 will be about 25 m (Fig. 3b). At that time the computers will be so powerful that they will be able to handle images of million by million pixels! These are of course rather simple models and real figures might differ from application to application.

2.3. Grid resolution and GPS positioning

If the scale of a project is unknown or non-standard, we can assess it by analyzing the mapping methodology. For example, the choice of the grid resolution can be related with the positional accuracy of our field positioning method. This is especially important for integration of GPS with remote sensing images and aerial photos where we should be certain that our GPS reading will (most probably) fall inside the right pixel. To ensure this, we first need to estimate the confidence radius of the positioning method using some control points. The confidence radius is the radius of a circle where we expect the most likely ($P = 95\%$) of the points to appear (Arnaud and Flori, 1998). It is normally evaluated using the error vector ($r_E$), which is as a difference between the measured ($X_{GPS}, Y_{GPS}$) and the true location of the control point ($X_T, Y_T$):

$$r_E = \sqrt{(X_{GPS} - X_T)^2 + (Y_{GPS} - Y_T)^2}$$

As a rule of thumb, one should select the grid resolution so that the area of circle described by the error radius is equal or smaller than the area of pixel (Fig. 4b):

$$p^2 \geq r_E^2 \cdot \pi$$

where $r_E^2$ is the average error radius and $p$ is the grid resolution. The recommended grid resolution is the one where most of the points ($95\%$) would fall within the pixel

$$p = \sqrt{r_E^2(P=95\%)} \cdot \pi \approx 1.8 \cdot r_E(P=95\%)$$

where $r_E(P=95\%)$ is the 95% probability error radius or confidence radius derived from the cumulative distribution of the error vectors from a set of measurements (Fig. 4a). Garmin (www.garmin.com), for example, claims that their handheld receivers (selective availability turned off) achieve horizontal error of not more than 15 m in 95% of time. This would be then compatible with a grid resolution of 27 m.

2.4. Grid resolution and remote sensing systems

As with raster-based GIS, grid resolution controls many aspects of remote sensing systems used for mapping. Characteristics of objects are scale dependent and their total number, area, average size or perimeter of objects being mapped will differ for different grid resolution (Lillesand and Kiefer, 2000, pp. 598–603). In fact, one will never be able to absolutely determine how many islands or lakes are present.
in the world and what is their perimeter. Although we cannot actually change the grid resolution in remote sensing images, we should at least be certain that the images are adequate for our mapping application. We can do that by inspecting the size of smallest spatial objects that are being mapped and then select the appropriate grid resolution, i.e. remote imaging system. Again, there should be at least four pixels to represent smallest objects and at least two pixels to represent the narrowest objects, which can be expressed mathematically as

\[
p \leq \begin{cases} \frac{\sqrt{a_{MLD}}}{4} & \text{if } S < 3 \\ \frac{w_{MLD}}{2} & \text{if } S > 3 \end{cases}
\]

(7)

where \(a_{MLD}\) is the area of the smallest objects, \(w_{MLD}\) is the width of the narrowest objects and \(S\) is the shape complexity index derived as the perimeter to boundary ratio:

\[
S = \frac{P}{2 \cdot r \cdot \pi}, \quad r = \sqrt{\frac{a}{\pi}}
\]

(8)

where \(P\) is the perimeter of polygon, \(a\) is the area of polygon and \(r\) is the radius of circle with the same surface area (Hole, 1953). Arbitrary value of 3 is used to differentiate between compact and narrow/long polygons. This means that we can first sample size of spatial objects (e.g. crop plots, water bodies, forest patches, roads) in part of the study area and then decide on the suitable sensor. To be more certain, one can plot the histogram with cumulative distributions and derive the 5% probability area, i.e. 5% probability width of the reference objects (Garbrecht and Martz, 1994). For example, if the objective of our mapping project are agricultural plots, and if the smallest plots are about 1 ha, the grid resolution should be at least 50 m, which means that we should use Landsat or similar imagery. If the smallest plots are about 0.05 ha (corresponds to the scale 1:5000), we will need images with resolutions of 10 m and finer.

In addition to the size of smallest objects, one also needs to take into account how contrasting is an object of interest when compared with its surroundings. The local contrast between adjacent objects can be assessed by statistically comparing the reflectance values between the target object \(r_T\) and surrounding object \(r_S\). Obviously, the higher the difference, the less strict we have to be about the grid resolution. For example, consider the centerline road markers of size of 20 cm. Such markers will be visible even on the 60 cm resolution QuickBird images because of high contrast with the surrounding objects. The impact of local contrast in the case of aggregation is illustrated in Fig. 5b: if the contrast is fairly low, the resampling will totally diminish point or line object so they cannot be distinguished from the surroundings.

2.5. Grid resolution and point samples

In many mapping projects, a map is made out of the point samples collected in the field and then used
to make predictions. To be consistent, every mapping project should have approximately an equal density of samples per area, also called inspection density. Obviously, the denser the observation points, the larger the scale of mapping. A cartographic rule, used for example in soil mapping, is that there should be at least one (ideally four) observation per 1 cm² of the map (Avery, 1987, Table 1). This principle can be used to estimate the effective scale of a data set consisting of sampled points only. For example, 10 observations per km² corresponds to the scale of about 1:500. The same principle can be also expressed mathematically

$$ SN = \sqrt{4 \cdot \frac{A}{N} \times 10^2} \ldots SN = \sqrt{\frac{A}{N} \times 10^2} $$  \hspace{1cm} (9)

where $A$ is the surface of the study area in m² and $N$ is the total number of observations. Remember from Eq. (1) that the scale number can be used to estimate the grid resolution. If we take the intermediate number of 2.5 observations per cm² and combine it with the $p = 0.5$ mm on the map rule of thumb, with a bit of reduction, we finally get a simple formula

$$ p = 0.0791 \cdot \sqrt{\frac{A}{N}} $$  \hspace{1cm} (10)

So, for example, if we deal with 100 samples and the size of the area is 10 km², a recommended grid resolution would be 25 m (see also Fig. 2b), which can also be expressed as 160 pixels per point sample.

Grid resolution can also be related to geometry of point patterns, i.e. distance between the sampled points (Boots and Getis, 1988). Following, again, the Nyquist frequency concept (Shannon, 1949) the grid resolution should be at most half the average spacing between the closest point pairs

$$ p \leq \frac{\bar{h}_{ij}}{2} $$  \hspace{1cm} (11)

where $\bar{h}_{ij}$ is the average distance between two closest point pairs also called mean shortest distance. In the case of regular point samples, the formula simplifies to

$$ p = 0.5 \cdot \sqrt{\frac{A}{N}} $$  \hspace{1cm} (12)

or 4 pixels per point sample. Note that the difference between the factors in Eqs. (10) and (12) is fairly big. You should keep in mind that Eq. (12) is valid only if we are dealing with (spatially) absolutely regular point samples. If the point samples show more random or clustered distribution of points, we need to be somewhat more strict about the grid resolution. If we are dealing with random point samples, then the average spacing between the closest point pairs is approximately half the spacing between closest point pairs in regular point samples (Fig. 6b).
is because random sampling has equal probability of producing totally clustered and totally regular samples (Fig. 6c). See also the supplementary materials (http://hengl.pfos.hr/PIXEL/) for further argumentation. The factor then modifies to

\[ p = 0.25 \cdot \sqrt{\frac{A}{N}} \]  

We can be even more strict and look for a grid resolution where \( >95\% \) of closest point pairs do not fall into the same pixel. This can be done using some well known point pattern analysis algorithms such as the one described in Boots and Getis (1988) and Rowlingson and Diggle (1993). We first need to derive the probability of finding the first point pair at different distances and for a given data set. We can then select the 5\% probability smallest distance (see Fig. 6c)

\[ p \geq h_{ij}(p=5\%) \]  

In addition, we can inspect the spatial correlation structure of the point data set and use this information to select the grid resolution. Here the key is to estimate the range of spatial dependence. Obviously, a variable that is spatially auto-correlated at shorter distances would require finer grid resolution and vice versa. Selection of grid spacing can be related with estimation of the optimal bin size that can estimate the probability density function, used for example in statistics to display histograms (Izenman, 1991). This gives a rather simple formula

\[ p = h_R \cdot m^{-\frac{1}{4}} \]  

After we have determined the range of spatial dependence \( (h_R) \), we need to count the number of point pairs \( (m) \) within that range and then derive a suitable grid resolution that can be used to represent the spatial dependence structure.

Another useful thing of estimating the spatial dependence structure is that we get an estimate of the nugget variance, i.e. the variation that happens at zero distances. This is a part of variation that we would typically like to remove or diminish, which can be achieved by using for example block kriging (Heuvelink and Pebesma, 1999; Bishop et al., 2001). Because the nugget variance refers to microscale variation (in fact, infinitely small distances), any block size with positive area would average out short-range variation. To be on the safe side, we can increase the size of the block \( (B) \) depending on the amount of nugget variation. A satisfactory block size can be estimated by comparing the derived estimation variance with the original variation in the data, so-called normalized estimation error

\[ \sigma_{E\%} = \frac{\sigma_E}{s_z} \]  

where \( \sigma_E \) is the estimation error (precision) and \( s_z \) is the total variation in the spatial data. This ratio is equivalent to the coefficient of multiple determination \( (R^2) \) used in regression analysis to show how much of the variation has been explained by a model. If \( \sigma_{E\%} \leq 40\% \) this means that the prediction model has explained 85\% or more of the variation in the data. In contrary, when the \( \sigma_{E\%} \) gets close to 100\%, this means that the prediction model is completely weak \( (R^2 \approx 0) \). In practice, we can run block-kriging predictions for different block sizes and then select the block size for which the normalized estimation error is good enough to map the whole area (e.g. \( \sigma_{E\%} = 40\% \)). Similarly,
there is no gain in using larger block sizes if there is no nugget variation or if already 95% of variation is explained by the prediction model ($\sigma_{E^2} = 23\%$, $R^2 = 0.95$).

2.6. Grid resolution and terrain analysis

The key problem when selecting a grid resolution for terrain analysis is that there can be a significant difference between the surface elevation on a coarser grid versus the actual topography, meaning that some peaks and channels might disappear in a raster DEM. In general, an increase in the detail in the DEM will also mean more accurate terrain parameters. This increase, however, depends on the general variability of the landscape. For example, a generally simple and smooth landscape might not need a fine resolution DEM. Even more so, if the grid resolution is too fine it might introduce local artefacts or slow down computation of terrain parameters. Obviously, we need a grid resolution that optimally reflects the variability of the elevation surface and is able to represent the majority of geomorphic features (Borkowski and Meier, 1994; Kienzle, 2004).

A suitable grid resolution can be derived for given sampled elevations (e.g. contours) and based on the complexity of (sampled) terrain. Imagine a one-dimensional topography with a number of inflection points (Fig. 7). We can again connect the problem of the grid resolution with the Nyquist–Shannon sampling theorem (Shannon, 1949; Tempfli, 1999). In our one-dimensional example above, terrain is the signal and its frequency is determined by the density of inflection points. Hence, the grid resolution should be at least half the average spacing between the inflection points:

$$ p \leq \frac{l}{2 \cdot m(\delta z)} $$ (17)

where $l$ is the length of a transect and $m(\delta z)$ is the number of inflection points observed. We can also be more strict and look for the 5% probability spacing between the inflection points. In the example in Fig. 7 there are 20 inflection points and average spacing between them is 0.8 m. Hence, a grid resolution of at least 0.4 m is recommended. Cumulative distribution of distances between the inflection points shows that 5% threshold of the smallest spacing between the inflection points is 0.2 m. Hence, grid resolution of at least 0.4 m and at most 0.2 m is recommended.

Fig. 7. Schematic example showing effect of grid resolution on representation of topography: too coarse grid resolution ($p = 2.5$ m) will misrepresent topography; whereas finer grid resolution ($p = 0.5$ m) will be more effective in representing all peaks and channels.
In the 2D case, the suitable grid resolution can be estimated from the total length of contours. Here the contours present mapped changes of fixed elevations. We do not actually have a map of inflection points, but can approximate them using the contour map. The suitable grid resolution is then

\[ p = \frac{A}{2 \cdot \sum l} \]  

where \( A \) is the total size of the study area and \( \sum l \) is the total cumulative length of all digitized contours. A more strict approach is to evaluate the density of contour lines in an area and derive the 5% probability smallest width of contours.

3. Case studies

3.1. Example 1: GPS positioning

In this example, I will first demonstrate how to select a grid resolution based on the evaluation of the GPS positioning method. In this case, 100 positioning fixes were recorded using the single-fix GPS positioning method (Arnaud and Flori, 1998) at the control point with a known location. The fluctuation of the GPS readings can be seen in Fig. 8a. The errors ranged from 0.7 to 23.9 m, average error was 8.5 m with a standard distribution of 5.2 m. The error vectors seem to follow the log-normal distribution (Fig. 8b). The theoretical log-normal distribution gave the 95% probability radius of 19.1 m, while the experimental distribution shows a somewhat higher value (20.4). Eq. (6) gives us a suitable grid resolution of 34.4 m. If this grid resolution is selected, most (95%) of GPS fixes will fall within the right pixels. This number for example corresponds to the resolution of the Landsat imagery. A more accurate positioning method would be needed to locate points within finer grid resolutions. For example, August et al. (1994) showed that using the averaging of multiple GPS fixes, the 95% radius of a standard GPS method can be decreased to up to 2.5 times by averaging 300 replicates. If we would like to use a GPS positioning with grid resolutions of about 15 m, then we would need to use GPS positioning with averaging (5 min per point). Higher positional accuracy (5–20 times) can be achieved by using differential correction, which can improve accuracy to less than three meters on average. Such accuracy would be compatible with grid resolutions within the range 2–10 m (SPOT or IKONOS imagery).

3.2. Example 2: monitoring agricultural plots

I will now demonstrate how to select a remote sensing sensor based on the size of agricultural plots (Fig. 9a). The polygon map consists of 121 polygons in total. The smallest polygon is 0.005 ha, the
biggest is 6.903 ha, average size of polygons is 0.824 ha with standard deviation of 1.005 ha. The polygons were then separated into two groups according to the shape complexity index. In this case only six polygons classified for narrow polygons. For each of these, an average width has been estimated by taking regular measurements (10 per polygon). A histogram of areas of compact polygons and a cumulative theoretical log-normal distribution can be seen in Fig. 9b.

I further on derived the 5% inverse cumulative distribution value assuming the log-normal distribution. I got 0.046 ha, which means that the pixel size should be about 20 m. The coarsest legible grid resolution for this data set \( (P = 50\%) \) would be 70 m (\( A = 0.5 \) ha). If resolutions coarser than 70 m are used to monitor agriculture for this area, then in more than 50% of the areas there will be less than four pixels per polygon. Note that in this case we are not using the true smallest polygon size but a somewhat higher figure (0.046 ha) because the smallest value (0.005 ha) is not representative. Further inspection of the widths showed that the average width of the narrow polygons is about 16 m, which gives a somewhat more strict grid resolution of about 8 m. However, the narrow polygons occupy only 0.9% percent of the total study area, so we do not have to be as strict. Finally, I would recommend that satellite imagery in a range from 10 to 70 m can be used to monitor agriculture in this study area.

### 3.3. Example 3: point data interpolation

In this example, I will use the Wespe point data previously used in numerous soil mapping applications (De Gruijter et al., 1997). The dataset consist of 552 profile observations where various soil variables have been described. The target variable is the membership value to enk earth soil type to analyze the spatial dependence structure. The values range from 0 to 1, with an average of 0.232 and a standard deviation of 0.322. The total size of the area is 12.1 km², which gives a sampling density of about 45 observations per km², which corresponds to the scale of 1:25k, i.e. grid resolution of 12 m (Eq. (10)). If we inspect the spreading of the points, we see that the average spacing between the closest point pairs is about 120 m, which is fairly close to the regular point sampling (for this data set—148 m). The cumulative distribution showed that 95% of points are at distances of 5 and more meters. This means that the legible grid resolutions are between 5 and 150 m (Eq. (12)).
Automated fitting of the variogram using a global model further gave a nugget parameter \( (C_0) \) of 0.042, a sill \((C_0 + C_1)\) of 0.097 and a range parameter \((R)\) of 175.2 m, which means that the variable is correlated up to the distance of about 525 m \((h_R)\). There are 11 807 point pairs within this range, which finally means that the optimal lag/grid size would be about 23 m. The pattern analysis of the point data set further shows that there is clear regularity in the point geometry: most of the distances are grouped at 180 m. The final interpolated map in resolution of 10 m can be seen in Fig. 10d. If we apply block-kriging, the nugget variation, which seems to be significant in this case \((C_0 = 0.042)\), would be diminished and the model would explain 71\% of variation at the support size of 10 m and 78\% at the support size of 60 m. Note that although further increasing of the block size would result in somewhat better normalized estimation error, on the other hand, we would unnecessarily lose a level of detail by increasing the block size above 120 m because more than 50\% point pairs would be within the same blocks. Also note that this sampling density corresponds to the scale of 1:25K, which means that the support size should not exceed 25 m (Fig. 11).

### 3.4. Example 4: contour data for terrain modelling

In this case study, I will demonstrate how a grid resolution can be selected from a map of contours, i.e. a dataset consisting of lines digitized from a topo-map. The study area is described in detail in Hengl et al. (2004). Contour lines were extracted

---

**Fig. 10.** Selection of grid resolution based on point pattern analysis: (a) a set of 552 soil profile observations; (b) probability of finding one point in neighborhood and graph of distances to closest point; (c) variogram and parameters fitted automatically in gstat (http://www.gstat.org); (d) interpolated map using ordinary kriging at grid resolution of 10 m.
from the 1:50K topo-map (Fig. 12a), with the contour interval of 10 m and supplementary 5 m contours in areas of low relief. The total area is 13.69 km² and elevations range from 80 to 240 m. There were 127.6 km of contour lines in total, which means that the average spacing between the contours is 107 m. The grid resolution should be at least 53.5 m to present the most of the mapped changes in relief. I then derived the distance from the contours map using the 5 m grid and displayed the histogram of the distances (Fig. 12b) to derive the 5% probability distance. Absolutely shortest distance between the contours is 7 m, and the 5% probability distance is 12.0 m. Finally, I can conclude that the legible resolution for this data set is within the range 12.0–53.5 m. Finer resolutions than 12 m are unnecessary for the given complexity of terrain. Note that selection of the most suitable grid resolution based on the contour maps is scale dependant. For the contour lines digitized from the 1:5K topo maps (Fig. 12c), the average spacing between the lines is 26.6 m and the 5% probability distance is 1.6 m. This means that, at 1:5K scale, the recommended resolutions are between 1.6 and 13.3 m.

4. Discussion and conclusions

There are three important concepts brought out in this paper that need to be further emphasized. First, principles from the general statistics and information theory, such as Nyquist frequency concept from signal processing and equations to estimate the probability density function, can be closely related to selection of the grid resolution. Second, there are three standard grid resolutions that can be derived for each input data:

- **Coarsest legible grid resolution**—this is the largest resolution that we should use given a specific scale, positional accuracy, size of objects or/and complexity of terrain. Using resolutions coarsest than the coarsest legible resolution means that we are either not respecting the scale of work,
positional accuracy, inspection density, size of objects being mapped or the complexity of terrain.

- **Finest legible grid resolution**—this is the smallest grid resolution that represents the most (95%) of spatial objects or topography. This is the finest meaningful resolution which corresponds to the concept of the maximum location accuracy. Resolutions finer than the finest legible resolution are probably just waste of memory.

- **Recommended grid resolution**—this is a compromising resolution, usually set as the intermediate number between the coarsest and finest resolutions.

Third, the choice of the grid resolution needs to be considered in relation to a number of inherent properties of a dataset or study area. The parameters and summary equations to derive coarsest and finest grid resolutions discussed in this paper are given in Table 1. These formulas can now be integrated within a GIS package to help inexperienced surveyors derive grid maps without doing extensive data preprocessing. A simple calculator can be found via the website listed at the beginning of the paper.

From all discussed statistical approaches to determine the true optimal pixel size, two seems to be most promising. The first is to select the pixel size that yields best predictive properties. For example, if we wish to use terrain parameters to predict soil properties over entire study area, we can test terrain parameters derived at various resolutions and then select the one that offers the highest correlation coefficient with the target variable (Florinsky and Kuryakova, 2000). A problem of this approach is...
### Table 1
Summary equations to select grid resolution: \( SN \) is scale factor, \( r_E \) is positioning error, \( \bar{r}_E \) is average positioning error, \( \bar{a} \) is average size of delineations, \( a_{MLD} \) is area of the minimum legible delineation, \( w_{MLD} \) is width of narrowest legible delineation, \( A \) is surface of study area, \( N \) is number of sampled points in study area, \( h_j \) is spacing between closest point pairs, \( h_R \) is average spacing between closest points, \( h_E \) is range of spatial dependence, \( m \) is number of point pairs within range of spatial dependence, and \( \bar{f} \) is total length of contours.

<table>
<thead>
<tr>
<th>Aspect</th>
<th>Coarsest legible resolution</th>
<th>Finest legible resolution</th>
<th>Recommended compromise</th>
</tr>
</thead>
<tbody>
<tr>
<td>Working scale</td>
<td>( \leq SN \cdot 0.0025 )</td>
<td>( \geq SN \cdot 0.0001 )</td>
<td>( = SN \cdot 0.0005 )</td>
</tr>
<tr>
<td>GPS positioning error</td>
<td>( \leq 1.8 \cdot r_E(P=99%) )</td>
<td>( \geq \bar{r}_E \cdot \sqrt{\pi} )</td>
<td>( = 1.8 \cdot r_E(P=99%) )</td>
</tr>
<tr>
<td>Size of reference objects</td>
<td>( \leq \frac{\bar{a}}{2} )</td>
<td>( \geq 2 \cdot a_{MLD} )</td>
<td>( = \frac{\bar{a}}{2} )</td>
</tr>
<tr>
<td>Inspection density</td>
<td>( \leq 0.1 \cdot \sqrt{\frac{2}{\pi}} )</td>
<td>( \geq 0.05 \cdot \sqrt{\frac{2}{\pi}} )</td>
<td>( = 0.0791 \cdot \sqrt{\frac{2}{\pi}} )</td>
</tr>
<tr>
<td>Distance between points</td>
<td>( \leq \frac{h_j}{4} )</td>
<td>( \geq h_R(P=5%) )</td>
<td>( = 0.25(0.5) \cdot \sqrt{\frac{2}{\pi}} )</td>
</tr>
<tr>
<td>Spatial dependence structure</td>
<td>( \leq \frac{h_j}{4} )</td>
<td>( \geq h_R(P=5%) )</td>
<td>( = h_R \cdot m^{-1} )</td>
</tr>
<tr>
<td>Complexity of terrain</td>
<td>( \leq \frac{\bar{f}}{4} )</td>
<td>( \geq \frac{w_{MLD}}{2} )</td>
<td>( = \frac{\bar{f}}{4} )</td>
</tr>
</tbody>
</table>

that it can be fairly time consuming to test all possible combinations of different grid resolutions. In addition, the prediction power versus grid resolution graph might give a set of different peaks for different target variables, so that we cannot select a single ‘optimal’ grid resolution. Finally, this grid resolution is then valid only for this study area and its effects might be different outside the area.

The second analytical approach to select the resolution is to derive information content for different block size values and then select the one that offers the richest amount of information per unit area (Bishop et al., 2001). In this case, only a single combination of the prediction model and block size should yield the highest information content. Such principles still need to be refined and tested in various case study to see if they really lead to optimal grid resolutions with maximum information content for a given level of detail.

Selection of the right pixel size will remain an issue that is relative to application type and project objectives. Moreover, standard grid resolutions (20–200 m in most cases) with which we work today, will soon shift toward finer and finer, which means that we need to consider grid resolution in time context also. No absolute ideal pixel size exist, that is for sure. One should at least try to avoid using resolutions that do not comply with the inherent properties of the input datasets.

### References


Sánchez Rojas, R., 2002. GIS-based upland erosion modeling, geovisualization and grid size effects on erosion simulations with CASC2D-SED. Ph.D. Dissertation, Colorado State University, Fort Collins, CO.


