

REAL-LIFE EXAMPLES FROM JORDAN

Spatial Data Quality Degradation

The competitiveness of an organisation is negatively affected by poor data quality, as introduced through the multitude of transformations and transfers carried out on original data. One risk involves uncontrolled exchange of data between organisations and departments within them. The author uses real-life examples from Jordan to demonstrate the impact of resolution, vector-to-raster conversion, scale, generalisation, classification of remotely sensed images and file exchange upon data quality.

Data acquisition is the most important step in any GIS and Geomatics project. All results are influenced by the quality of the original data and subsequent steps involving transformation, transfer and exchange. Principal spatial data sources are satellite imagery, aerial photographs, field surveying and scanned maps and documents. Original data cannot normally be used directly in GIS projects. For example, satellite images have to be corrected for atmosphere and Earth-curvature effects, among others. And geo-referencing has to be applied to bring them into a preferred reference system. All these transformations introduce error and degradation of quality.

Image Resolution

Figure 1 shows two satellite images of different resolutions: on the right is part of a Landsat image of 30m resolution; the left image stems from an Ikonos image of 1m resolution. Both were taken in the same year. Much information present in the left-hand image is absent in the right-hand one. The number of mixed pixels (mixels) in the low-resolution image causes degradation in shape and in a number of recognisable features. In the high-resolution image one can identify three pools, whereas in the low-resolution image only one is visible. Different resolutions will also result in differing values for areas, as shown in Table 1. The expansion in the urban area of Greater Amman, and the runoffs in the Walla and Habisse, were computed from 30m and 15m-Landsat data and from a digital vector map. The computations summarised in Table 1 demonstrate how coarse resolutions introduce overestimation of area size.

Vector to Raster

Vector-to-raster (V-to-R) transformation also introduces degradation, depending on resolution: i.e. size of raster cells. A point in the vector database, for example, becomes a square of the size of one pixel or raster cell. V-to-R transformation can be carried out either by central-point or by dominant-unit rasterisation. To demonstrate the effects on shape and size, we in Figure 2 transformed a vector polygon (A) to raster (B) and back again to raster mode (C). The resulting polygon overlaid on the original (see D) shows how severe degradation can be.

Scale and Detail

A map feature must be visible and easily identifiable under normal conditions; that is, at a distance of 30cm from the eyes and under normal light conditions. The number of features that can be visualised is proportional to scale. Not only number, but also the shape of features depends on scale. Smaller details will appear only when the scale becomes large enough. Whether certain details will be represented or not depends on the rules of legibility, which include:

- visual acuity of differentiation: in general, the factors which determine the recording of an image by eye are colour, lighting conditions and size of objects
- visual acuity of alignment: ability of the eye to see whether two lines are aligned to each other
- parting threshold: the eye requires a minimal space between two features in order to distinguish them as separate. This space depends on the thickness of the lines with which the features are drawn; for thick lines the parting threshold is 0.15mm
- differential threshold: ability of the eye to record differences in size of features.

The minimal dimensions that the eye can record without ambiguity, under natural conditions, are shown in Table 2. When the dimension of a map feature is smaller than the minimal dimension this feature cannot be represented at its true size and has therefore to be represented by a conventional sign. For example, a building of 25 x 25 metres can be represented at scale on maps of scales better than 1:50,000. But it cannot be represented in its true size at a scale of, for example, 1:500,000, because then the map dimensions of this building would become 0.05 x 0.05mm.

Generalisation

Figure 3 shows a curved line represented at different scales. Details are lost when represented at medium-scale (B), and become a straight line when represented at small-scale (C). Suppose this curved line is a segment of a road; then the length would vary according to scale and the expected error could exceed 100%. Figure 4 shows the effect of three types of generalisation of a group of islands: elimination of the small islands (top), regrouping the small islands into one island (middle) and regrouping the small islands to become part of the large island (bottom). All three generalisations show large degradation in area, shape and number of features.

Classification

When classifying remotely sensed images, degradation may be introduced resulting from:

- limited number of classes: only a limited number of classes can be selected, while in reality the number of classes may be much larger; features which would not belong to any selected class will be erroneously allocated to one of the selected classes
- percentage of pixels correctly classified is generally less than 100%
- mixed pixels: a pixel cannot belong to more than one class at a time; a mixed pixel will belong to one class only, leading to area and shape errors
- unclassified pixels: pixels that are not allocated to any of the selected classes.

Further, geo-referencing introduces degradation: the Root Mean Square (RMS) error is rarely equal to zero, and re-sampling causes spectral degradation. The likelihood of degradation increases with increasing transfer of data between organisations; for example, when the format is changed during transfer from vector to image format, such as Tiff or JPG (Figure 5).

Concluding Remarks

Organisations must standardise data formats to avoid multiple transfer of formats, some of which may be of low quality. A history file must accompany data so that some estimation may be made of the quality of the data, as good-quality data can be never obtained from poor-quality data.

Further Reading

- Burough, P., A., 1986, Principles of Geographical Information Systems for Land Resources Assessment.
- Comber A.J., Fisher P., and Wadsworth R., 2004, Integrating land cover data with different ontology: identifying change from inconsistency, International Journal of Geographical Information Science 18 (7), pp 691-708.

<https://www.gim-international.com/content/article/spatial-data-quality-degradation>
