DOI 10.1111/tgis.12257

RESEARCH ARTICLE

WILEY Transactions in GIS

Completeness and classification correctness of features on topographic maps: An analysis of the estonian basic map

Kiira Mõisja[1] | Tõnu Oja[1] | Evelyn Uuemaa[1] | Jordan T. Hastings[2]

[1] Department of Geography, University of Tartu, Estonia

[2] Department of Geography, Ellison Hall, University of California, Santa Barbara, CA, USA

Correspondence
Kiira Mõisja, Vanemuise 46, Tartu 51024 Estonia.
Email: kiiram@ut.ee

Abstract

In an increasingly GIS-literate world, the availability of quality topographic maps and map databases is critical for the numerous users of spatial data. Particularly governmental agencies, first responders, and utility and transportation services, rely on the completeness and classification correctness of these maps. Estonia has systematically updated its topographic Basic Map in digital form over the past 15 years. An analysis of the Estonian production process in the period 2003-2006 provides a useful case study of both error types and error frequencies encountered in topographic mapping. Errors of completeness and classification correctness of topographic features are analyzed at two levels of specificity: in general, across all map sheets, and in detail according to the field-workers who performed the mapping. The structure of errors at the two levels was different by geometry and error types; however, both systematic and individual errors were evident. The systematic errors indicated a need for revision and improvement of the data capture specifications, which was accomplished. The individual errors were addressed by additional training for the field-workers involved.

KEYWORDS
classification correctness, completeness, error analysis, field verification, topographic mapping

1 | INTRODUCTION

Topographic maps and their accompanying vector datasets, collected by national mapping agencies (NMA), provide a reference framework for other spatial datasets and for many spatial data services (Jakobsson & Giversen, 2007) not only for the nation itself, but also regionally and globally (Jakobsson, 2012). Among the main users of the topographic data are governmental agencies, municipalities, first responders, and utility and transportation service providers (Jakobsson, 2003). The wide user-community for topographic data and the increasing adoption of GIS, requires interoperability across geographic scales and sets high expectations for data quality and also for ongoing data quality management.

 © 2016 John Wiley & Sons Ltd

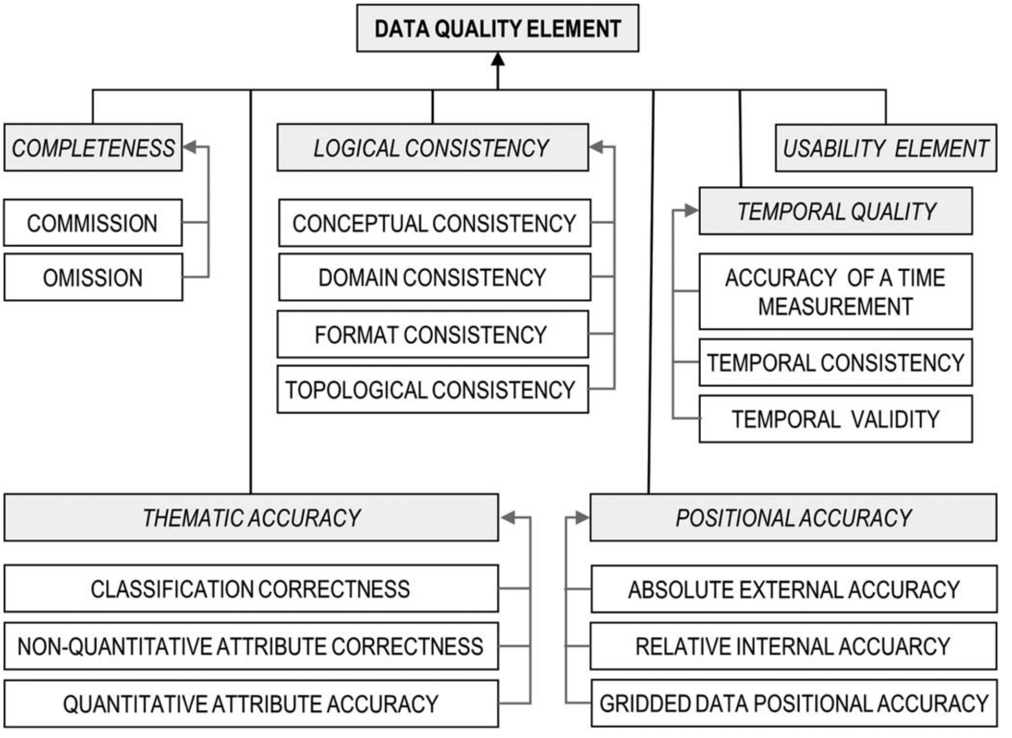

FIGURE 1 Overview of the ISO 19157:2013 data quality elements (after ISO, 2013, Figure 4)

During the last 30 years, there have been numerous discussions about describing, measuring, and reporting geospatial data quality (for example Guptill & Morrison, 1995; Veregin, 1999; Aalders, 2002; Hunter, Bregt, Heuvelink, de Bruin, & Virrantaus 2009; Devillers et al., 2010). Agreed principles of data quality have been reflected in several international standards for spatial data quality (Veršić, 2009; Kresse, Danko, & Fadaie, 2012). Specifically, ISO 19113 (2002) determined the elements of quality, ISO 19114 (2003) described the quality assessment procedure, and ISO 19138 (2006) defined the quality measures. In 2013, a new standard ISO 19157 (2013) entitled "Geographic Information – Data Quality" was published, combining all three predecessor standards and presenting a more user-centered and comprehensive view of the topic (Jakobsson, Hopfstock, Beare, & Patrucco, 2013; Leibovici, Pourabdollah, & Jackson, 2013). For this study, geospatial data quality is interpreted according to ISO 19157 (2013).

Compared with the previous standards, ISO 19157 (2013) clarifies the scope of data quality, implements "usability" as the new quality element, provides guidelines for reporting quality evaluation results, and introduces the meta-quality concept. According to the standard, data quality is expressed by 21 quality elements, which have been organized into six categories: completeness, thematic accuracy, logical consistency, temporal quality, positional accuracy, and usability (see Figure 1). Each quality element is described by a quality measure, evaluated with a selected method and the result is reported in a standalone quality report or in metadata. The quality of the data quality elements themselves is described in meta-quality.

Data quality is a concept related to uncertainty (Zhang & Goodchild, 2002; Fisher, Comber, & Wadsworth, 2006), which is endemic in all geospatial data (Goodchild, 2009). For a data producer, it is important to determine the sources of uncertainty, find the methods to measure them, and minimize them by using quality management. A main source of uncertainty is how well a feature class is defined (Fisher et al., 2006; Longley, Goodchild, Maguire, & Rhind, 2005). If the feature class is well-defined, i.e. treats clearly distinguishable geospatial objects, then uncertainty is caused by objective human errors (Fisher et al., 2006), which may be of omission or commission (Devillers & Jeansoulin, 2006; Fisher, 1999).

If the feature class is poorly-defined, however, then uncertainty comes from vagueness or ambiguity (Fisher et al., 2006), which is more subjective. Not only natural phenomena like shorelines and forests, but also some anthropogenic phenomena like building ruins, can be vague. How many trees are required to form a forest? When does a building have only its foundation left? Ambiguity also arises when the same feature could be placed into two or more classes according to different interpretations of classification rules. For example, Comber, Fisher, and Wadsworth (2005) analyzed and graphically depicted several definitions of the term "forest" used across the world. These definitions resulted in substantially different areas of forest to be mapped. Therefore specifications need to be determined prior mapping and made clear to map- and data-producers.

Ultimately, maps need to be validated by a second-round of field-work, i.e. post-production inspection. Such field-inspection is costly and time-consuming (Devillers & Beard, 2006), and hardly exciting for most. Still, for topographic maps, the subject has been addressed by Pätynen, Kemppainen, and Ronkainen (1997), Jakobsson (2002), and briefly by Harding (2006). Allied studies treat mapping of habitat (Cherrill & McClean, 1999; Stevens, Blackstock, Howe, & Stevens, 2004). The number of papers discussing thematic accuracy and uncertainty is most developed in remote sensing, especially in land use and land cover mapping (for example Ahlqvist, 2004, 2008; Comber et al., 2005; Foody, 2002, 2008).

Hunter et al. (2009) and Devillers et al. (2010) note the persistent problem that data quality tends to be analyzed and presented at a summary, dataset level rather than at more detailed levels of granularity for individual features, attributes, etc. Earlier, following the work of several authors, Devillers and Beard (2006) introduced the *hierarchical model* with four levels of quality detail addressing in turn the overall dataset, its feature classes, individual feature instances, and their geometric primitives. At roughly the same time, Sadiq, Duckham, and Hunter (2006) pointed out the *spatial variation* in data quality that arises due to different authors' techniques for data capture, compilation, analysis, and representation. In this study, we extend the list of concerns to *human factors*: the abilities and biases of field-workers who inevitably interpret the landscape subjectively to some extent (Cherrill & McClean, 1999; Stevens et al., 2004). To understand the range of discrepancies, i.e. errors of completeness and classification correctness, in topographical digital data, it is necessary to analyze both field mapping and quality control procedures in all these ways.

This article analyzes data quality for the Estonian Basic Map at two levels of *specificity*: *in general*, where all errors are analyzed together; and *in detail*, where the same errors are analyzed according to the field-worker. We also discuss the influence these differences have on the determination of the most error-prone feature classes and the reasons for those errors.

2 | THE ESTONIAN BASIC MAP

The Estonian Basic Map (1:10 000) is a national topographic vector database in MicroStation (Bentley, 2015) .dgn format. This map comprises traditional topographic phenomena like roads, buildings, water network, land cover, and relief (see Figure 2). However, there are also several phenomena characteristic of Estonian landscapes: heaps-of-stones varying in size from a few square meters up to hundreds of square meters, scattered and sometimes isolated trees in wooded meadows, and abandoned farmland along with foundations of relict farmhouses. In total, there are 129 different feature classes defined and symbolized on the Basic Map.

The Basic Map was completed by a process of stereoplotting from aerial imagery and extensive field-work (Figure 3). The latter was needed due to the poor quality of the imagery and the lack of earlier trustworthy mapping and state registers. Moreover, topographic maps from the Soviet period were not considered reliable because of their age or large distortions (Mardiste, 2009). The Basic Map is divided into 5x5 km map sheets. The smallest area covered in field-work by a single person was one-fourth of a map sheet (6.25 km^2); this means that one sheet could be mapped by up to four different field-workers.

Original field-work on the Basic Map occupied a decade, 1997-2006. In 1999, guidelines for the digital production of the Basic Map were established, together with the quality requirements based on the then-emergent ISO 19113 (2002) standard. The mapping used in this study comes from the period 2003-2006. The Estonian Land Board

FIGURE 2 An example of Estonian Basic Map (left) along with photos of characteristic landscapes: (1) a heap-of-stones (Koov, 2007); and (2) a forest cutline (V. Uri, 2015, photograph), also located on the map by those numbers in red

performed the quality control for mapping field-work via its own field-inspection. Thus, in the current article, we distinguish between *field-work*, the mapping process of checking and amending stereoplots in the field, and *field-inspection*, the validation of field-work, also in the field. The time-difference between field-work and the field-inspection was usually about two months. However, there were still cases where the landscape had changed in that period, the typical case having to do with forest logging, where the trees were removed in the interim. Such cases were recorded by the field-inspectors and fixed in the map database, but not considered as field-workers' errors. Also, field-workers were given the opportunity to rebut field-inspectors' reviews, where appropriate. In event of doubt, the decision was made in favor of the field-worker.

Field-inspection was carried out in sections selected so that the work of as many different field-workers as possible would be inspected. Generally, instead of a rectangular area, a linear route was inspected; the inspector walked through and reviewed the mapping along the route (Figure 4). Altogether, 1,140 km of field-inspection was performed along 71 routes; the adjacent mapped area totaled approximately 126 km^2. The inspection results were documented in a detailed quality control report (Mõisja, 2003).

Following ISO 19157 (2013), the quality of field-work was assessed first by completeness and second by thematic accuracy, the latter consisting of three quality elements: classification correctness, non-quantitative attribute correctness, and quantitative attribute accuracy. (As the Basic Map is a CAD database, focused on spatial features essentially

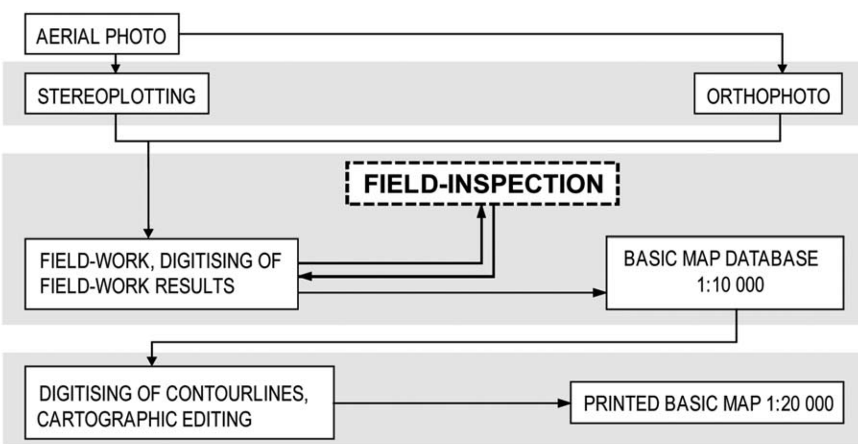

FIGURE 3 Production scheme for the Estonian Basic Map (adapted from Estonian Land Board). The focus of this study is the center text box, dash-outlined

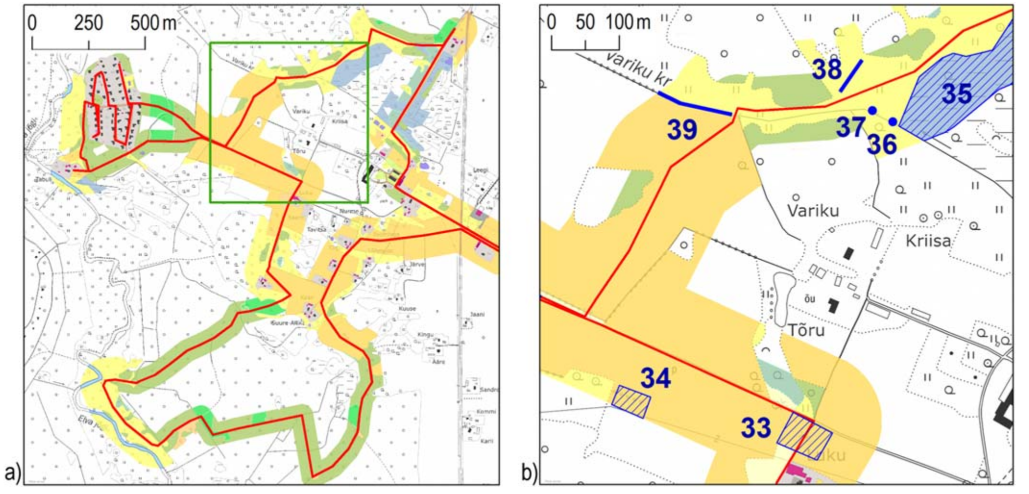

FIGURE 4 Sample section of field-inspection, at two scales. On left, an overview of the field-inspection route is marked with a red line and the inspected area with color fills: yellow and light blue polygons are landscapes with open view; green and grey polygons are landscapes with closed view. On the right, in detail, mapped and reported discrepancies are numbered in dark blue: 36, 37 refer to point features; 38, 39 to line features; 33, 34, 35 to polygon features

without attributes, the latter two elements were not considered.) Completeness was measured in terms of *omission*, where landscape entities that are covered by the map specifications do not appear as map features, and of *commission*, where map features exist that are not evident in the landscape, or if evident do not conform to the map specifications (Harding, 2006). Classification correctness meant, first and foremost, conformance of map features to entities in the landscape. Additionally, the field-inspectors observed (by eye, without direct measurement) "wrong size" and "wrong place" as indicators of positional inaccuracy. All errors were recorded as Boolean values.

3 | DATA AND METHODS

The objective of this study was to determine the prevalence of various types of errors across the Estonian Basic Map, i.e. spatially, and by mapper. We first describe our sampling framework to capture these details unambiguously, followed by the spatial database that recorded those errors for us, and finally analyses we made.

3.1 | The sampling framework

For a similar study in Finland, Jakobsson and Marttinen (2003), used square 1x1 km sampling plots. The authors recognized that defining sampling units for topographic map data is complex, because of the large difference in the number of features between densely populated squares (up to 562 features per square) and the forest squares (as few as 10 features per square).

For the Estonian Basic Map, we adopted the approach of sample *sections* formed by buffers around the field-inspectors' routes, as mentioned above (see Figure 4). A buffer width of 50 m was taken if the surrounding landscape had a "closed" viewshed (in forests, shrubbery, and residential areas) and 100 m otherwise, in "open" viewsheds. We feel that this sampling framework of buffered routes across the landscape represents the areas to be inspected better than fixed-size plots. Field-inspectors were careful to consider only the errors within the buffered sections.

3.2 | The error database

A spatial database of errors in mapping, referred to as the error database, was created by transcribing the field-inspectors' quality control reports. This database notes 3,916 errors found in 40 map sheets produced by 19 field-workers in the years 2003-2006. Ten of the field-workers had carried out 67% of all field-work.

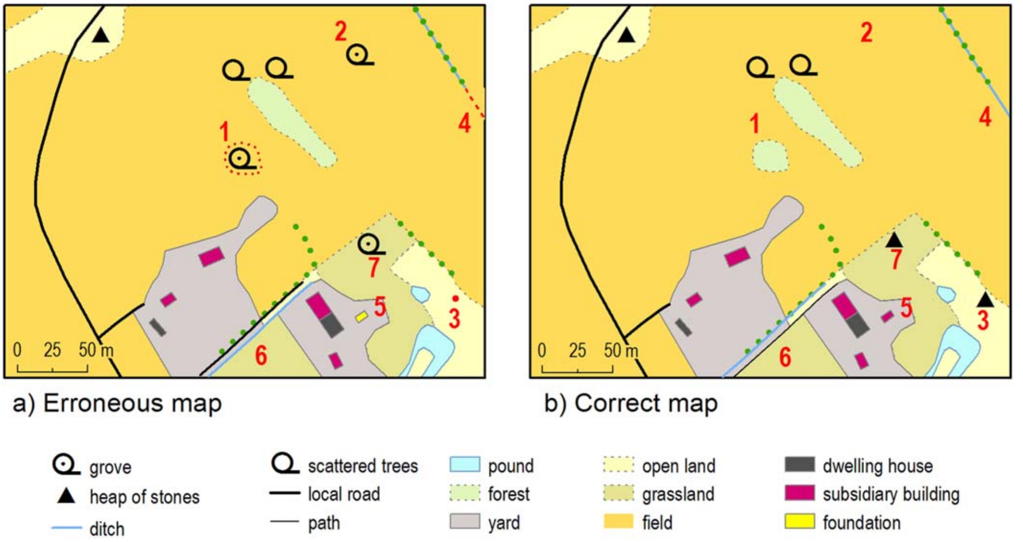

FIGURE 5 Examples of discrepancies in an erroneous map (left) and the corrected map after field-inspection (right). The erroneous "grove" (numbered 1) can be considered in three ways: a simple misclassification (forest instead of grove); a paired commission (grove) and omission (forest); or a commission (grove) and misclassification (forest instead of field)

While developing the error database, it became apparent that field-inspectors themselves had classified errors differently. For example, the misclassification of a subsidiary building as a foundation could alternatively be considered as an error of omission of a subsidiary building plus an error of commission in defining a foundation. Error classification became particularly complicated if the correct mapping required a change of geometry type, for example from a grove of trees (point feature) in a field to a small forest stand (polygon feature) of its own. The latter sort of errors might be recorded in three ways: (1) misclassification (grove instead of forest); (2) commission (grove) plus omission (forest); or (3) commission (grove) plus misclassification (forest instead of field).

For consistency in our subsequent analyses, all errors were recorded according to a common classification system applying the following rules, as exemplified in Figure 5:

- Point features could have all error types (see Figure 5, errors 2, 3, and 7).
- Line features could have all error types. Where line lengths were either shorter or longer than they should have been, the error was noted as either omission or commission, respectively, not as a wrong size (Figure 5, error 4). In addition, where nearby parallel line features had swapped places with each other (Figure 5, error 6), a classification error was noted, not a wrong placement.
- Polygon features that participated in a full tessellation (no holes or overlap) could only be misclassified. However, small polygon features (Figure 5, error 5) that were recognized as point features during field-inspection could have all error types, as could short linear features (culverts, bridges).
- Finally, when the geometry type changed from point to polygon or from line to polygon (Figure 5, error 1), the point or line was recorded as an error of commission and polygon as a misclassification, not an omission.

3.3 | Error analysis

The mapping errors noted by the field-inspectors were analyzed at two levels of specificity: first *in general*, considering the whole set across all maps, and second *in detail*, according to the originating field-worker. Analyses were made with regard to: (1) the type of errors (e.g. completeness, misclassification); (2) the geometry of erroneous features (point, line, or polygon); and (3) the specific feature classes involved. Finally, the most misclassified feature classes were determined.

TABLE 1 Data quality measures used in analysis

Analysis	Measure	Definition
Type of errors (quality elements)	error count	number of incorrect items
Geometry of errors (point, line, polygon)	error count	number of incorrect items
Feature classes of errors	error sum	total number of incorrect point items, length of incorrect line items, or area of incorrect polygons
Most misclassified feature classes	error sum	total number of incorrect point items, length of incorrect line items, or area of incorrect polygons

ISO 19157 (2013) specifies numerous measures for each quality element. In this article, only error count, error sum, and error rate are considered (Table 1). *Error count* indicates the total number of errors irrespective of feature geometry, thus lumping them all together. *Error sum* is calculated differently for each geometry type: the number of incorrect point features, the aggregate length of incorrect line features, and the aggregate area of incorrect polygon features. The counts and sums characterize the overall impact of errors differently, although they have subjective interpretations. For example, the error database included three missing paths with a total length of 300 m and one superfluous ditch with a length of 500 m. For some data users, three missing paths (count) might represent a conspicuously larger problem as the number of features is higher; whereas for others, the single fictitious ditch (sum) may have a higher impact. Finally, *error rate* expresses as a percentage the number erroneous features with respect to the total number of correct features that should have been present. Error rate is thus normalized and useful for a comparative analysis.

We used error count to rank field-workers' performance according to error type and geometry type of erroneous features. Five error types were distinguished: omission, commission, misclassification, wrong size, and wrong place (Table 2).

The most error-prone feature classes were determined by using scatter plots, which are remarkably helpful but sadly uncommon in data quality analysis. These plots visually elucidate the relation between error and landscape, and are made separately for each feature geometry. The feature classes that appear in the upper left quadrant on the scatter plots were considered critical. Across field-workers, the distribution of errors are described by coefficient of variation, i.e. the ratio of the standard deviation to the mean, a unitless quantity.

An error matrix (also called a contingency or confusion matrix) (Congalton & Green, 1993; Foody, 2002) was used to study the misclassifications in the feature database. This matrix was generated separately for each geometry type, using error sum as the quality measure.

4 | RESULTS

The mapping errors were analyzed in three ways – by type (Figure 6), by geometry (Figure 7), and by feature class (Figure 8, Table 3) – at both levels of specificity. As expected, the analysis by feature class and field-mapper was both most time-consuming and most instructive.

4.1 | Types of errors

In general, the largest share of all errors came from omission (48%) and misclassification (33%) (Figure 6, column Total). Among point features, omission accounted for 59% of all errors, and for line features 54% (Table 2). According to our protocol (Section 3.1), polygon features could only be misclassified.

In detail, Figure 6 shows that the predominant error among field-workers was also omission. Four field-workers had misclassification as their most common error; one field-worker had almost equal rates of omission (44%) and misclassification (42%).

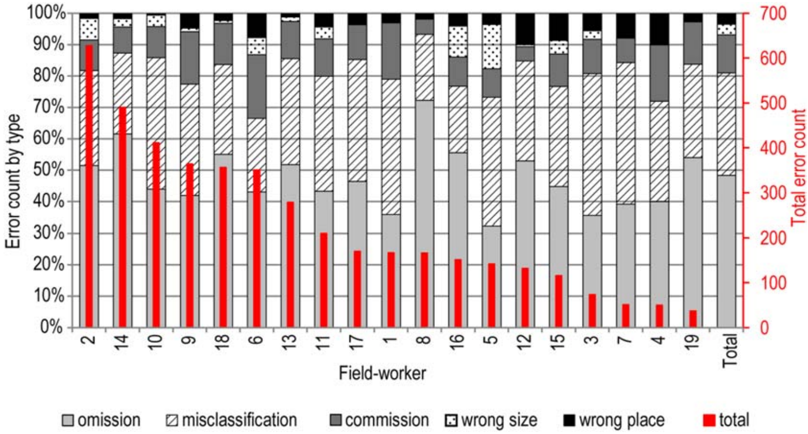

FIGURE 6 The variability of errors by type among field-workers. The ordinate (X-axis) ranks field-workers by their decreasing share of errors overall (red stripe), measured by Y-axis on right. The Y-axis on left shows the distribution of errors by type for each field-worker

4.2 | Geometry of errors

In general, the majority of errors involved lines (46%) and points (40%); polygon errors were much fewer (14%) (Figure 7, column Total). The share of features by geometry type in the sampled sections was more nearly uniform: lines (47%), points (30%), and polygons (23%).

In detail, much larger variability appeared. For example, over two-thirds of all errors made by one field-worker related to lines; whereas for another field-worker only about one-sixth did so. For most field-workers, polygon errors were least prevalent, less than 20%. The results presented in Figure 7 suggest three groups of field-workers: (1) those who make a disproportionate number of errors regarding line features (42%); (2) those who make a disproportionate number of errors regarding point features (37%); and (3) those whose errors are in similar proportion to the database overall (accomplished by 21% of field-workers).

Considering geometry type and error type together, three equal groups can be distinguished: (1) six field-workers with omissions comprising over 50% of all error both in line and point features; (2) another six field-workers with omissions exceeding 50% only for point features; and (3) the final six having omission exceeding 50% only for lines. By contrast, only one field-worker consistently misclassified features of all geometries.

4.3 | Feature classes of errors

Of the 104 feature classes that appeared in the field-inspection sample sections, 20 were absolutely error-free; these were features that were clearly recognizable in stereo images (lake, railway, radio-tower, high-voltage power line), or that were corroborated by other reliable databases (first- and second-class roads), or that appeared so infrequently (ruins of windmill, light tower) that field-workers had time to study them.

TABLE 2 Distribution of errors (count) by geometry and error

	Omission	Commission	Misclassification	Wrong size	Wrong place	TOTAL Errors	Total Features
Point	922	207	333	38	57	1,557	16,383
Line	968	265	399	96	76	1,804	26,255
Polygon			555			555	12,393
TOTAL	1,890	472	1,287	134	133	3,916	55,031

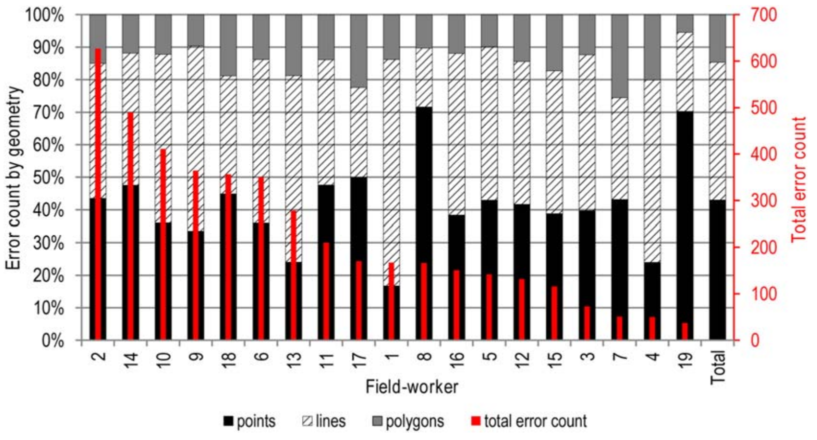

FIGURE 7 The variability of errors by geometry among field-workers. The ordinate (X-axis) ranks field-workers by their decreasing share errors overall (red stripe), measured by Y-axis on right. The Y-axis on left shows the distribution of errors by geometry for each field-worker

In general, the feature classes with many errors compared to their abundance appear in the upper left quadrant of the scatter plots (Figure 8). Five point, three line, and two polygon features are highlighted in the critical quadrant.

In detail by field-workers, the variability of error-prone feature classes in the scatter plots was much higher (Table 3). Two feature classes in particular (footpath and heap-of-stones) occurred in the critical quadrant for at least half of the field-workers. Altogether, eight feature classes were error-prone for about one-third of the field-workers. One feature class (forest) appeared only in these detailed analyses. Conversely, another feature class (narrow ditch <2 m wide), which was clearly problematic in general, only caused errors for four field-workers.

Four feature classes (grasslands, narrow ditches, culverts and open space) had the highest variation among field-workers (Table 3). More than half the errors regarding grasslands and narrow ditches were made by just two field-workers; regarding culverts and open space, most of the mistakes were made by three field-workers. The most uniformly problematic feature classes were foundation, footpath, and grove of trees.

More than 60% of the point errors were omissions. The highest share of omissions (92%) applied to heap-of-stones. By contrast, the highest share of commissions occurred for scattered trees (27%). The highest share of misclassification involved grove-of-trees (43%). These facts are summarized by confusion matrices, per feature class (Figure 9).

In general, the most misclassified point features involved two feature pairs: deciduous grove vs. deciduous tree and dwelling house vs. subsidiary building (Figure 9 left). The two building types were equally confused. Deciduous

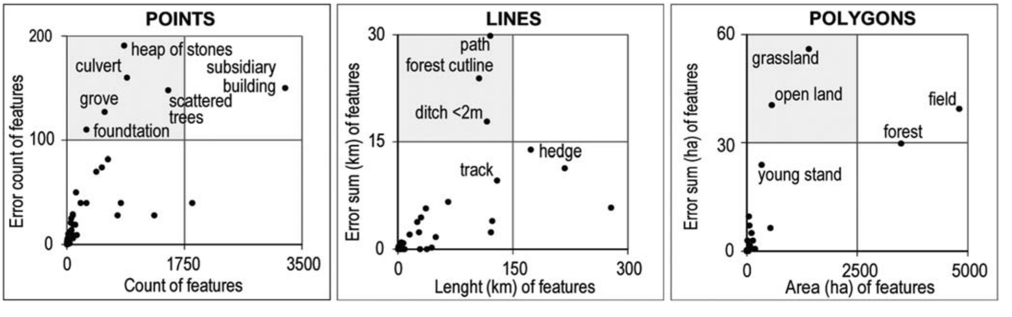

FIGURE 8 Quartile-quartile plots of errors in the three feature geometries against their summary measures, with individual feature classes labeled. The feature classes having relatively few errors appear close to the horizontal axis. The features placed clearly above the diagonal (upper left quadrant) may be considered more problematic: the frequency of this particular feature class in the landscape is relatively low but the number of errors is high compared to other feature classes

TABLE 3 Distribution of errors (count, m, m^2) for selected feature classes among field-workers. In the stub, the parenthesized number tabulates field-workers who had features represented in the critical quadrants of scatter plots (Figure 8). In the body, the measures of errors made by those field-workers are shown in italics

Feature class	Field-worker 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	StDev	Mean	CV
POINTS																						
heap-of-stones (9)	*8*	*27*	10		*9*	8	*3*	*29*	8	*7*	*15*	4	*9*	3	*6*	1	3	10	*3*	7.5	9	0.83
foundation (8)	1	11	*3*	1		4	*5*	9		*8*	4	1	*6*	3		*10*	*8*	*12*	*2*	3.6	6	0.66
scattered trees (7)	1	*26*			1	5	*3*	12	12	*13*	8	3	*8*	3		3	*13*	*19*	*3*	6.9	8	0.83
grove (7)	4	19		*3*	*6*	9		*16*		*11*	*11*	3	2	3	1	5	*8*	8	*2*	5.1	7	0.73
culvert (7)		*19*			2	*29*			*33*	*16*	5	*12*	4	2			1	*14*	*3*	10.5	12	0.90
LINES																						
path (14)	*1,714*	*2,982*	930	*610*	164	1,566	272	*862*	*2,718*	*2,940*	*1,974*	*415*	1,755	372	*642*	*976*	*1,404*	*2,862*	*182*	958	1,334	0.72
forest cutline (7)	*2,232*	2,398	256	28	94	1,300	464		*2,127*	*3,071*	*1,779*	177	*3,634*	*1,769*		254	*2,298*	1,030		1,112	1,432	0.78
ditch <2m (4)	420	1,358		*399*		*6,018*		213	1,332	1,233	683	212	*2,739*	489		89		*1,652*		1,542	1,295	1.19
POLYGONS																						
open space (8)	77	19,673	*5,625*		1,883	*36,397*		*9,090*	*52,727*	10,335	*9,907*	*51,027*	*101,136*	*16,023*			12,681	38,027		26,745	26,043	1.03
grassland (6)	4,046	48,698	3,343		3,206	*18,654*	*6,596*		7,982	*9,941*	5,703	4,056	*29,277*	*11,701*	*2,807*	2,270	*127,208*	*191,245*		51,582	29,796	1.73
forest (6)	*34,240*	47,763	*13,161*	1,783	2,230	16,065	10,037	*11,866*	19,592	*21,368*	2,788	3,968	36,064	6,107	*3,364*	*8,065*	7,782	17,805		12,730.2	14,669	0.87

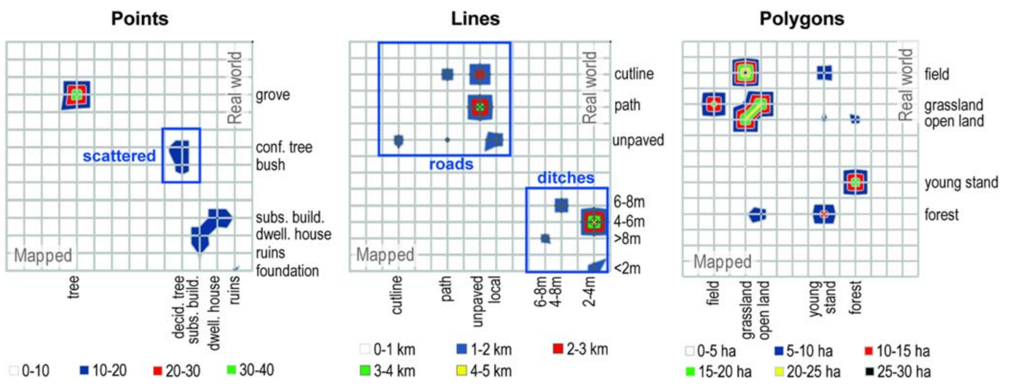

FIGURE 9 Selections from confusion matrices for the most misclassified feature classes, by geometry. The horizontal-axis shows features presented on the map, the vertical-axis features occurring in the real world

groves were sometimes considered deciduous trees but not *vice versa*, the reason being that in the original stereoplotting both trees and groves were marked as trees. For field-workers this situation was confusing, as they could not keep track of whether the particular tree was already classified or not.

Among line features, footpath caused 40% of all misclassifications and ditch caused 32% (Figure 9 center). Footpaths and forest cutlines were classified as tracks, i.e. unpaved road, a higher class, in 25% of the cases. For *erroneous* ditches, a related problem in misclassification was the determination of their width class: in 20% of cases the class was overestimated and in 80% of cases it was underestimated.

Among polygon features, 43% of all misclassifications involved the three most common classes: arable field, grassland, and open space (Figure 9 right). Predominantly, an arable field was misclassified as grassland. Another pair of misclassifications had to do with (mature) forest vs. young forest classes, as the latter also was used for clear-cuts.

In detail by field-workers, similar misclassification rates appeared for point and line features as in general. However, differences appeared in polygon features. The most frequent mix-ups again involved three feature classes: forest, grassland, and open space. Forest vs. grassland and forest vs. open space misclassifications were made by 19 and 14 field-workers, respectively. The total area of misclassifications was small, which is why it did not appear in the general analysis.

5 | DISCUSSION

Quality control (QC) analysis, as done above, is one of the constituents of data quality management (Jakobsson, 2003). The results of QC serve as a basis for the improvement of data production (Dassonville, Vauglin, Jakobsson, & Luzet, 2002; Jakobsson & Marttinen, 2003). Detailed QC is important for quality management as the variability across individuals can be high.

In our study, the analyses in general and in detail, by error type and feature geometry, demonstrated considerable differences across the Estonian Basic Map; this is of concern for both map users and map producers. Based only on the general analysis, we might conclude that point features and line features were equally problematic. Their errors were mainly caused by omission and misclassification. However, the detailed analysis revealed variability between field-workers. For example, one field-worker mainly omitted point features, whereas another field-worker misclassified lines. Even the analysis by error type and feature geometry did not sufficiently explain the variability of errors across the map. Very likely the number of errors in mapping does not depend so much on the geometry but rather on the specific feature class.

To determine the *sources* of error, further analysis by feature classes, as recommended Hunter et al. (2009) and Devillers et al. (2010), was needed. At both levels of specificity, our analyses highlighted several common critical feature classes and associated errors, which can be divided into three groups: (1) features that were frequently omitted, e.g. heap-of-stones, foundation, culvert, path, forest cutline; (2) features that tended to be committed in excess, e.g. scattered trees; and (3) features that were mostly misclassified, e.g. deciduous grove and open space.

The omission of features might have occurred for several reasons. Regarding heap-of-stones and footpath, the reason is clearly insufficient definition of the feature classes involved. Estonia has many heaps-of-stones of different sizes and shapes and also numerous footpaths of different widths and qualities, which are not always obvious in forested areas. In these cases, the high number of errors relates to the field-workers' inability to decide which class to use based on the existing definitions, or whether to map the feature at all. For culverts, relict foundations, and forest cutlines the problem may also be visibility. Foundations are often overgrown by vegetation limiting their detection in the landscape. Cutlines may be similarly obscured by new growth which makes them difficult to detect. Analogous studies performed in Finland (Pätynen et al., 1997; Jakobsson, 2002) showed that the biggest difficulties with omission and commission together related to buildings, streams <2 m in width, light-traffic routes, as well as some types of agricultural lands.

With regard to misclassification, the problematic features were again buildings and agricultural lands. The direct comparison with the Finnish study is complicated as the feature classes (for example, types of agricultural land) are different. However, it is evident that similar feature classes – light-traffic routes in Finland and footpaths or forest cutlines in Estonia – are the critical omissions in both cases. The most misclassified features were those "neighbouring" each other (like agricultural field – grassland – open space); and those rapidly changing in time (forest – clearcut – young forest). Similar issues are well-known in habitat mapping (Cherrill and McClean, 1999; Stevens et al., 2004). The abundant commission of scattered trees in this study is closely related to the misclassification of forest overall.

The features that did not appear problematic in the general analysis but where many field-workers made errors that were obvious in detail needs additional discussion. With regard to the forest feature class, for example, the reason is the large number of small size misclassified forest polygons found in the sampled sections. The high coefficient of variation (0.87) for this class also indicates uniform distribution of errors among the field-workers, who frequently mixed-up forest with grassland and open space. This surprising result can be explained by the location of the errors, which were mainly located along the forest edge where neighbouring grassland with scattered trees or open space with scattered trees made the border of forest vague. Exactly where such grassland or open space with scattered trees turned into forest and *vice versa* varied by field-worker. In both cases scattered trees changed into commission of forest.

Conversely, some features that appeared frequently misclassified in the general analysis, could be traced to only a few field-workers in the detailed analyses, and thus were considered a matter of individual field-workers' skills. In this study, the feature classes ditch < 2 m and grassland were problems because of errors made by just two field-workers, who produced more than half of all such errors.

Often it is assumed that geospatial data collected by the same processes have similar quality (Frank et al., 2004). However, our study as well habitat mapping studies (Cherrill & McClean, 1999; Stevens et al., 2004) has demonstrated that this assumption does not hold in field mapping. Different field-workers may have remarkable differences in skill and hence in the quality of data they produce. We determined early in our study that to decide whether errors were systematic or not, detailed analysis by field-worker at the feature class level was needed; this is also advised by Devillers and Beard (2006).

The problematic feature classes with more-or-less even distribution of the errors across field-workers were considered to indicate the need for revision and improvement of the data capture specification (Harding, 2006). Where uneven distributions occurred, additional training for some field-workers was indicated. In our opinion, the high variability among field-workers needs further research.

Also, the impact of landscape diversity on the quality of mapping field-work should be addressed. Several studies on accuracy of land cover data derived from automated classified satellite images (van Oort et al., 2004; Smith, Stehman, Wickham, & Yang, 2003) have found that classification accuracy is positively correlated with some landscape diversity indicators.

6 | CONCLUSIONS

This research is based on an empirical database of mapping errors detected for the Estonian Basic Map found during routine field-inspection. QC analyses were carried out at two levels of specificity: (1) a general level, where all errors were considered together; and (2) a more detailed level where errors were broken-out by field-worker.

Quality analyses by error type (omission, commission, and misclassification), by geometry type (point, line, and polygon), and by feature classes showed important differences at the general vs. detailed levels, although there were commonalities, too. Most problems stemmed from insufficient definition of the desired feature classes (poor mapping specifications) and inadequate training of field-workers (individual differences). A surprise was that such an "obvious" feature class as forest could be problematic; this had to do with the features' essential fuzziness (Burrough & Frank, 1996) and temporality (Langran, 1993), which was not sufficiently realized at the outset of mapping. Other, error-prone feature classes at both levels of analysis were forest cutline, scattered trees and deciduous grove, footpath, small culvert, relict foundation, and the quintessential Estonian heap-of-stones.

ACKNOWLEDGMENTS

The field inspections were supported by the Estonian Land Board. The three Estonian authors are affiliated with the institutional grant IUT 2-16 funded by the Estonian Ministry of Education and Research and by the Marie Skłodowska-Curie Actions individual fellowships offered by the Horizon 2020 Programme under REA grant agreement number 660391. The fourth author was previously involved in analogous mapping projects during his tenure as the Chief Cartographer for the State of Nevada, USA. We also wish to thank the reviewers for constructive and helpful comments and suggestions on the manuscript.

DISCLAIMER

The map errors analyzed in this study were corrected and they did not affect the quality of the Estonian Basic Map delivered to users.

REFERENCES

Aalders, H. J. G. L. (2002). The registration of quality in a GIS. In W. Shi, P. F. Fisher, & M. F. Goodchild (Eds.), *Spatial data quality*, (pp. 186–199). London, UK: Taylor & Francis.

Ahlqvist, O. (2004). A parameterized representation of uncertain conceptual spaces. *Transactions in GIS*, *8*, 493–514.

Ahlqvist, O. (2008). In search of classification that supports the dynamics of science: The FAO Land Cover Classification System and proposed modifications. *Environment & Planning B*, *35*, 169–186.

Bentley. (2015). MicroStation. Retrieved from http://www.bentley.com/engb/Products/MicroStation/

Burrough, P. A., & Frank, A. (Eds.) (1996). *Geographic objects with indeterminate boundaries*. London, UK: Taylor & Francis.

Cherrill, A., & McClean, C. (1999). Between-observer variation in the application of a standard method of habitat mapping by environmental consultants in the UK. *Journal of Applied Ecology*, *36*, 989–1008.

Comber, A., Fisher, P. F., & Wadsworth, R. (2005). What is land cover? *Environment & Planning B*, *32*, 199–209.

Congalton, R. G., & Green, K. (1993). A practical look at the sources of confusion in error matrix generation. *Photogrammetric Engineering & Remote Sensing*, *59*, 641–644.

Dassonville, L., Vauglin, F., Jakobsson, A., & Luzet, C. (2002). Quality management, data quality and users, metadata for geographical information. In W. Shi, P. F. Fisher, & M. F. Goodchild (Eds.), *Spatial data quality* (pp. 202–215). London, UK: Taylor & Francis.

Devillers, R., & Beard, K. (2006). Communication and use of spatial data quality information in GIS. In R. Devillers & R. Jeansoulin (Eds.), *Fundamentals of spatial data quality* (pp. 237–253). London, UK: ISTE.

Devillers, R., & Jeansoulin, R. (2006). Spatial data quality: Concepts. In R. Devillers & R. Jeansoulin (Eds.), *Fundamentals of spatial data quality* (pp. 31–42). London, UK: ISTE.

Devillers, R., Stein, A. R., Bedard, Y., Chrisman, N., Fisher, P., & Shi, W. (2010). Thirty years of research on spatial data quality: Achievements, failures and opportunities. *Transactions in GIS*, *14*, 387–400.

Estonian Land Board. (2016). *Estonian basic map*. Retrieved from http://geoportaal.maaamet.ee/docs/pohikaart/tootmis-skeem.gif.

Fisher, P. F. (1999). Models of uncertainty in spatial data. In P. A. Longley, M. F., Goodchild, D. J., Maguire, & D. W. Rhind (Eds.), *Geographical information systems* (pp. 191–205). New York, NY: John Wiley & Sons.

Fisher, P. F., Comber, A., & Wadsworth, R. (2006). Approaches to uncertainty in spatial data. In R. Devillers & R. Jeansoulin (Eds.), *Fundamentals of spatial data quality* (pp. 43–59). London, UK: ISTE.

Foody, G. M. (2002). Status of land cover classification accuracy assessment. *Remote Sensing of Environment*, *80*, 185–201.

Foody, G. M. (2008). Harshness in image classification accuracy assessment. *International Journal of Remote Sensing*, *29*, 3137–3158.

Frank, A. U., Grum, E., & Vasseur, B. (2004). Procedure to select the best dataset for a task. In M. J. Egenhofer, H. J. Miller, & C. Freksa (Eds.), *Proceedings of the third international conference on geographic information science* (pp. 81–93). Berlin, Germany: Springer.

Gervais, M., Bedard, Y., Levesque, M., Bernier, E., & Devillers, R. (2009). Data quality issues and geographic knowledge discovery. In H. J. Miller & J. Han (Eds.), *Geographic data mining and knowledge discovery* (pp. 99–115). Boca Raton, FL: CRC Press.

Goodchild, M. F. (2009). The quality of geospatial context. In K. Rothermel, D. Fritsch, W. Blochinger, & F. Dürr (Eds.), *Quality of context: First international workshop* (pp. 15–24). Berlin, Germany: Springer.

Guptill, S. C., & Morrison, J. L. (Eds.) (1995). *Elements of spatial data quality*. Amsterdam, the Netherlands: Elsevier.

Harding, J. (2006). Vector data quality: A data provider's perspective. In R. Devillers & R. Jeansoulin (Eds.), *Fundamentals of spatial data quality* (pp. 141–159). London, ISTE.

Hunter, G. J., Bregt, A. K., Heuvelink, G. B. M., de Bruin, S., & Virrantaus, K. (2009). Spatial data quality: Problems and prospects. In G. Navratil (Ed.), *Research trends in geographic information science* (pp. 101–121). Berlin, Germany: Springer Lecture Notes in Geoinformation & Cartography.

ISO. (2002). *ISO 19113:2002 Geographic Information – Quality Principles*. Geneva, Switzerland: International Organization for Standardization.

ISO. (2003). *ISO 19114:2003 Geographic Information – Quality Evaluation Procedures*. Geneva, Switzerland: International Organization for Standardization.

ISO. (2006). *ISO/TS 19138:2006 Geographic Information – Data Quality Measures*. Geneva, Switzerland: International Organization for Standardization.

ISO. (2013). *ISO 19157:2013 Geographic Information – Data Quality*. Geneva, Switzerland: International Organization for Standardization.

Jakobsson, A. (2002). Data quality and data quality management: Examples of quality management in European national mapping agencies. In W. Shi, P. F. Fisher, & M. F. Goodchild (Eds.), *Spatial data quality* (pp. 216–229). London, UK: Taylor & Francis.

Jakobsson, A. (2003). Framework and requirements for management of topographic data in Europe. In *Proceedings of the 9th Scandinavian Research Conference on Geographical Information Science* (pp. 91-102). Espoo, Finland.

Jakobsson, A. (2012). Introducing a new paradigm for provision of European Reference Geo-Information: Case study of the European Location Framework Concept. In M. Jobst (Ed.), *Service oriented mapping 2012* (pp. 51–62). Vienna, Austria: JOBST Media.

Jakobsson, A., & Giversen, J. (Eds.) (2007). Guidelines for implementing the ISO 19100 geographic information quality standards in national mapping and cadastral agencies. Retrieved from http://www.eurogeographics.org

Jakobsson, A., & Marttinen, J. (2003). Data quality management of reference datasets: Present practice in European National Mapping Agencies and a proposal for a new approach. In *Proceedings of the 21st International Cartographic Conference* (pp. 1748-1756). Durban, South Africa: ICA.

Jakobsson, A., Hopfstock, A., Beare, M., & Patrucco, R. (2013). Quality management of reference geo-information. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*, *40-2/W1*, 127–132.

Koov, H. (2007). Photo of heap of stones. Retrieved from http://www.monument.ee/vihulavald/palmse-naljakangrud

Kresse, W., Danko, D. M., & Fadaie, K. (2012). Standardization. In W. Kresse & D. M. Danko (Eds.), *Springer handbook of geographical information* (pp. 393–566). Berlin, Germany: Springer.

Langran, G. (1993). *Time in geographic information systems*. London, UK: Taylor & Francis.

Leibovici, D. G., Pourabdollah, A., & Jackson, M. J. (2013). Which spatial data quality can be metapropagated? *Journal of Spatial Science*, *58*, 3–14.

Longley, P. A., Goodchild, M. F., Maguire, D. J., & Rhind, D. W. (2005). *Geographic information systems and science* (2nd Ed.). Chichester, UK: John Wiley & Sons.

Mardiste, H. (2009). Consequences of the Soviet map secrecy to national cartography in Estonia. In D. Unverhau (Ed.), *Geheimhaltung und Staatssicherheit. Zur Kartographie des Kaltes Krieges. Archiv zur DDR-Staatssicherheit*. (Vol. *9.1*, pp.107–118). Munster, Germany: LIT Verlag.

Mõisja, K. (2003). Estonian Basic Map and its quality management. *Transactions of the Estonian Agricultural University*, *216*, 135–142.

Pätynen, V., Kemppainen, I., & Ronkainen, R. (1997). Testing for completeness and thematic accuracy of the national topographic data system in Finland. In *Proceedings of the 18th International Cartographic Conference* (pp. 1360-1367). Stockholm, Sweden: ICA.

Pierkot, C., Zimányi, E., Lin, Y., & Libourel, T. (2011). Advocacy for external quality in GIS. In C. Claramunt, S. Levashkin, & M. Bertolotto (Eds.), *Geospatial semantics* (pp. 151–165). Berlin, Germany: Springer-Verlag.

Sadiq, M. Z., Duckham, M., & Hunter, G. J. (2006). Modeling spatial variation in data quality using linear referencing. In *Proceedings of the 7th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Science* (pp. 225-235). Lisbon, Portugal: ISARA.

Smith, J. H., Stehman, S. V., Wickham, J. D., & Yang, L. (2003). Effects of landscape characteristics on land-cover class accuracy. *Remote Sensing of Environment*, *84*, 342–349.

Stevens, J. P., Blackstock, T. H., Howe, E. A., & Stevens, D. P. (2004). Repeatability of Phase 1 habitat survey. *Journal of Environmental Management*, *73*, 53–59.

van Oort, P. A. J., Bergt, A. K., de Bruin, S., de Wit, A. J., & Stein, A. (2004). Spatial variability in classification accuracy of agricultural crops in the Dutch national land-cover database. *International Journal of Geographical Information Science*, *18*, 611–626.

Veregin, H. (1999). Data quality parameters. In P. A. Longley, M. F. Goodchild, D. J. Maguire, & D. W. Rhind (Eds.), *Geographical information systems* (pp. 177–189). New York, NY: John Wiley & Sons.

Veršič, A. (2009). *Spatial data quality control process based on ISO 19113 and ISO 19114*. Paper presented at the 24th International Cartographic Conference, Santiago, Chile.

Zhang, J., & Goodchild, M. F. (2002). *Uncertainty in geographical information*. London, UK: Taylor & Francis.