Received: 31 August 2022 | Revised: 22 December 2022 | Accepted: 30 January 2023
DOI: 10.1111/avsc.12715

Applied Vegetation Science IAVS

RESEARCH ARTICLE

What explains inconsistencies in field-based ecosystem mapping?

Adam Eindride Naas[1] | Rune Halvorsen[1] | Peter Horvath[1,2] | Anders Kvalvåg Wollan[1] | Harald Bratli[1] | Katrine Brynildsrud[1] | Eirik Aasmo Finne[3] | Lasse Torben Keetz[1,2] | Eva Lieungh[1] | Christine Olson[4] | Trond Simensen[1,5] | Olav Skarpaas[1] | Hilde Riksheim Tandstad[6] | Michal Torma[1] | Espen Sommer Værland[7] | Anders Bryn[1,2,8,9]

[1]Geo-ecology Research Group, Natural History Museum, University of Oslo, Oslo, Norway
[2]Department of Geosciences, University of Oslo, Oslo, Norway
[3]Norwegian Institute for Nature Research, FRAM – High North Research Centre for Climate and the Environment, Tromsø, Norway
[4]Ecofact AS, Sandnes, Norway
[5]Norwegian Institute for Nature Research, Trondheim, Norway
[6]Sállir Natur AS, Tromsø, Norway
[7]Dokkadeltaet Nasjonale Våtmarkssenter AS, Odnes, Norway
[8]Division of Survey and Statistics, Norwegian Institute of Bioeconomy Research, Ås, Norway
[9]CBA, Faculty of Mathematics and Natural Sciences, University of Oslo, Oslo, Norway

Correspondence
Adam E. Naas, Geo-ecology Research Group, Natural History Museum, University of Oslo, Oslo, Norway.
Email: a.e.naas@nhm.uio.no

Funding information
Norges Forskningsråd; Norwegian Biodiversity Information Centre

Co-ordinating Editor: Hannes Feilhauer

Abstract

Questions: Field-based ecosystem mapping is prone to observer bias, typically resulting in a mismatch between maps made by different mappers, that is, inconsistency. Experimental studies testing the influence of site, mapping scale, and differences in experience level on inconsistency in field-based ecosystem mapping are lacking. Here, we study how inconsistencies in field-based ecosystem maps depend on these factors.

Location: Iškoras and Guollemuorsuolu, northeastern Norway, and Landsvik and Lygra, western Norway.

Methods: In a balanced experiment, four sites were field-mapped wall-to-wall to scales 1:5000 and 1:20,000 by 12 mappers, representing three experience levels. Thematic inconsistency was calculated by overlay analysis of map pairs from the same site, mapped to the same scale. We tested for significant differences between sites, scales, and experience-level groups. Principal components analysis was used in an analysis of additional map inconsistencies and their relationships with site, scale and differences in experience level and time consumption were analysed with redundancy analysis.

Results: On average, thematic inconsistency was 51%. The most important predictor for thematic inconsistency, and for all map inconsistencies, was site. Scale and its interaction with site predicted map inconsistencies, but only the latter were important for thematic inconsistency. The only experience-level group that differed significantly from the mean thematic inconsistency was that of the most experienced mappers, with nine percentage points. Experience had no significant effect on map inconsistency as a whole.

Conclusion: Thematic inconsistency was high for all but the dominant thematic units, with potentially adverse consequences for mapping ecosystems that are fragmented or have low coverage. Interactions between site and mapping system properties are considered the main reasons why no relationships between scale and thematic inconsistency were observed. More controlled experiments are needed to quantify the effect of other factors on inconsistency in field-based mapping.

KEYWORDS
classification, experience, field-based mapping, GIS, inter-observer variation, land-cover mapping, landscape metrics, ordination, scale, vegetation mapping

1 | INTRODUCTION

Accurate, consistent information on the extent and distribution of ecosystems is the foundation for a wide range of research activities and applied purposes, such as spatial planning and environmental impact assessments (Hersperger et al., 2021), conservation planning (Beier et al., 2015), ecosystem accounting (Hein et al., 2020), nature restoration (Hagen et al., 2022), and evaluation of the economic potential of natural resources (e.g. grazing capacity; Mysterud et al., 2014).

Characterisation of ecosystems useful for management purposes requires generalisation, by dividing gradual variation into non-overlapping thematic units of spatially appropriate sizes (Bailey, 1987). The thematic units typically delineate specific parts of ecological space, and/or consist of vegetation with homogeneous species composition, certain physiognomic characteristics, and/or indicator species (Moss, 2008; Halvorsen et al., 2020). Ecosystem maps are spatial representations of ecosystems adapted to a given scale (Küchler & Zonneveld, 1988) and the degree of generalisation (i.e., scale) is represented by thematic resolution (i.e., number of thematic units) and spatial resolution (e.g. the size of pixels or the minimum size of polygons).

Photographic interpretation (Ihse, 2007), pattern recognition or image classification (Phiri & Morgenroth, 2017), statistical modelling (Horvath et al., 2019; Simensen et al., 2020), and machine-learning methods (Talukdar et al., 2020) are often used for mapping at regional to landscape scales (i.e., spatial resolution of 10–1000 km^2). These methods are normally used for classifying thematic units defined by vegetation structure, such as dominance of different species groups (Franklin, 2013). Ecosystem maps can also be made on a field computer equipped with a GPS and aerial photos by delineating polygons and assigning them to thematic units. Field-based mapping is more time- and resource-consuming, but advantageous when thematic units are defined by the species composition of field-layer and bottom-layer vegetation (Ullerud et al., 2020). Field-based maps are complementary to other methods, for instance as reference data for training and/or validation of statistical models (Guisan & Zimmermann, 2000).

Measurements of errors in ecosystem mapping should ideally include comparison against an objective, independent ground truth data set to estimate the 'validity' of the produced maps. However, it is virtually impossible to establish accurate reference maps for map validation since validation points or maps are subject to the same biases, inconsistencies, and errors as the maps they are meant to validate. In other words, 'ground truths' do not necessarily represent truths in ecosystem mapping. Therefore, other measurements are needed in the assessments of map uncertainties. In such classification tasks (i.e., when correct classifications are difficult to obtain) inconsistency analyses are useful for addressing uncertainties (Kahneman et al., 2021). This applies to ecosystem mapping, where in some cases the thematic units are ambiguous (e.g. Natura 2000; European Commission, 2013), whereas in other cases, correct classification is virtually impossible to achieve with high certainty because the thematic units are defined by the species composition varying along complex environmental gradients (Halvorsen et al., 2020; i.e., truly correct classifications require longitudinal data from exhaustive species sampling and of different environmental factors).

Thematic inconsistency can be quantified as the spatial correspondence of assigned thematic units between pairs of maps from the corresponding area (Cherrill & McClean, 1995). In a broad sense, inconsistencies in ecosystem mapping result from three decision-making processes: assignment of thematic units to polygons (classification), delineation of polygons (delineation), and inclusion or exclusion of polygons (generalisation). Attempts to quantify inconsistencies resulting from classification and delineation exist (Haga et al., 2021), but generalisation has received far less attention (but see Mõisja et al., 2018). Ecosystem mapping can be treated as a decision-making process, where quantification of inconsistencies arising from classification, delineation and generalisation can inform what sort of calibration is most critical. Because thematic units often are defined by the presence/absence and abundance of species, inconsistencies have mainly been hypothesised to arise from seasonal variation of the vegetation or mappers' tendency to deviate in species detection or abundance estimates (Eriksen et al., 2018).

The wide range for thematic inconsistency of 9%–82% in field-based ecosystem maps (Kershaw & Bunce, 1998; Hearn et al., 2011) suggests that one or more factors strongly influence thematic inconsistency. Previous studies have either been conducted as repeat surveys (i.e., quality controls; e.g. Kershaw & Bunce, 1998) or as controlled experiments (e.g. Ullerud et al., 2018) but have weaknesses as they have been conducted in different years (Stevens et al., 2004), at different times of the year (Cherrill & McClean, 1995), using different classification methods (Hearn et al., 2011), without time limitation during fieldwork (Cherrill & McClean, 1999), or with insufficient sample size (Ullerud et al., 2018). Indications from these studies suggest thematic inconsistency is influenced by properties of the mapped ecosystems (Stevens et al., 2004), the scale and mapping system that is used (Ullerud et al., 2018), and the mappers' experience (Hearn et al., 2011). A better understanding of the influences of these factors is needed, and might be achieved with improved experimental designs. We aim to answer how inconsistencies in wall-to-wall field-based ecosystem maps depend on (1) site properties, (2) mapping scale and (3) differences in experience level among mappers.

2 | METHODS

2.1 | Study area

The study area consists of four sites within two regions: a continental region in northeastern Norway and an oceanic region in western Norway (Figure 1, Table 1).

The northeastern region is situated in the northern-boreal bioclimatic zone (Bakkestuen et al., 2008), characterised by high temperature seasonality, mean annual temperature between −1 and −2°C, and annual precipitation below 500 mm (seNorge, 2022). The western region is situated in the boreo-nemoral bioclimatic zone (Bakkestuen et al., 2008) with low temperature seasonality, mean annual temperature between 6 and 8°C, and annual precipitation in the range of 2000–3000 mm (seNorge, 2022).

Iškoras is dominated by mires, bogs, mountain birch forests and open alpine heathlands. Permafrost is appearing scattered as palsa mires. Wintertime grazing by reindeer locally reduces the lichen cover in alpine heaths and on ridges (Tømmervik et al., 2012).

Guollemuorsuolu is dominated by well-drained pine forests, alluvial forests, and flooded, gravel-dominated riverbanks. In addition, there are wetlands and areas influenced by domestic grazing and rangeland mowing. Farming alongside the Kárášjohka River was abandoned several decades ago.

Landsvik is dominated by oceanic moist forests, open moist heaths, and wetlands. The area was previously used extensively for grazing, outfield mowing and peat extraction, but these activities were abandoned decades ago. Today, dense Sitka spruce (*Picea sitchensis*) plantations are a major influence on the landscape.

Lygra is maintained as a traditional coastal heathland, burned at regular intervals and used for year-round sheep grazing and rangeland mowing (Måren & Vandvik, 2009). The heathland is dominated by heather (*Calluna vulgaris*), or, in pioneer stages after burning, by grasses. Locally, semi-natural grassland occurs in patches more intensively grazed by sheep. Wetlands dominate in depressions and mostly have clear signs of former peat extraction.

2.2 | Study design

Each site in each region was mapped by a team of 12 mappers, 10 of whom participated in mapping both regions. The number of maps

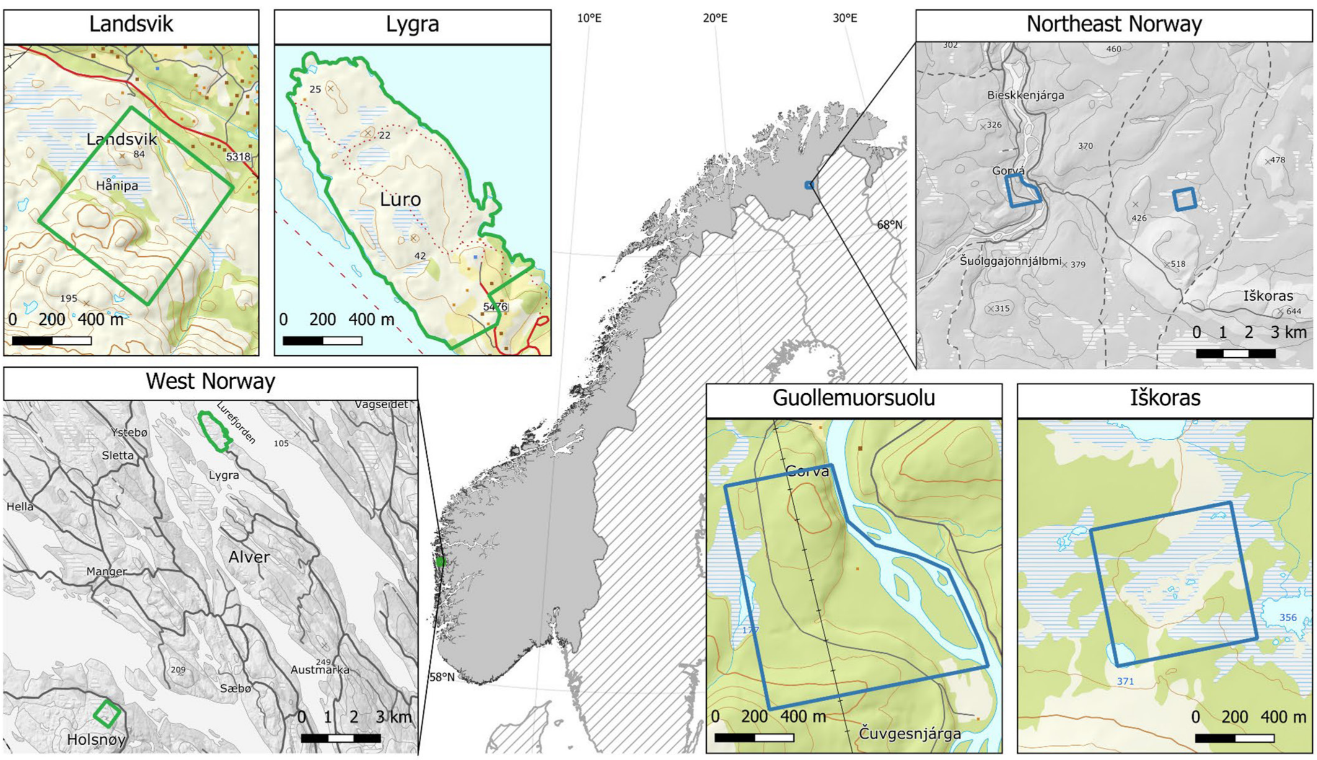

FIGURE 1 Location of the two study regions within Norway (western to the left, outlines in green; northeastern to the right, outlines in blue) with the study sites shown in insert maps. Background images were obtained from the Norwegian Mapping Authority (ETRS89, UTM zone 33 N).

TABLE 1 Basic study site information.

Region	Site	Centroid	Altitude range	Site size
Northeastern Norway	Iškoras	69.341/25.296	354–371	0.52
	Guollemuorsuolu	69.353/25.140	146–228	1.00
Western Norway	Landsvik	60.596/5.056	12–177	0.51
	Lygra	60.697/5.101	0–41	0.88

Note: Region, site name, centroid coordinates (latitude/longitude), altitude range (m a.s.l.), and size (km^2).

produced was 2 regions × 2 sites × 12 mappers = 48 maps. Each team was sorted into four groups, into which one mapper from each of three experience levels was assigned. A survey of the skills of each mapper, which was the basis for allocating them to experience levels, is given in Appendix S1. The 12 mappers in each team were randomly assigned to 12 combinations of three treatments in a 2 × 2 × 3 manner:

1. Site order:

i. First mapping of Guollemuorsuolu (northeastern region) or Landsvik (western region)

ii. First mapping of Iškoras or Lygra

2. Scale order:

i. First mapping of one site to scale 1:5000, then the other site to scale 1:20,000

ii. First mapping of one site to scale 1:20,000, then the other site to scale 1:5000

3. Experience:

i. Beginner

ii. Intermediate

iii. Experienced

The mappers had three days to complete 1:5000 maps and one day to complete 1:20,000 maps. Each of the 12 treatment combinations was replicated two times (once in each region).

2.3 | Ecosystem typology and mapping guidelines

Field-based mapping was carried out with the Nature in Norway system (NiN version 2.2; Halvorsen et al., 2020). The NiN ecosystem typology is hierarchical, with three nested levels (major-type groups, major types and minor types). Major types are defined by a set of principles, for example separating ecosystems differing in type of disturbance, or dominance of ecosystem engineering species groups (e.g. trees, helophytes). Within major types, each minor type spans an interval of standardised size along each main local complex environmental gradient (e.g. risk of drought gradient, lime richness gradient). This interval comprises a standardised amount of species compositional turnover [25% change in species composition, defined as one ecological distance unit (EDU); see Halvorsen et al., 2020]. Minor types are aggregated to mapping units adapted to mapping at different scales (e.g. 1:5000 and 1:20,000). Hereafter, 'mapping units', 'major-type units', and 'thematic units' refers to units in the original maps, in maps where the mapping units are aggregated to the major-type level, and to any unit in general respectively. All terrestrial and wetland units of the NiN system were allowed during mapping, as well as five Freshwater and Marine units. This corresponds to four major-type group units, 60 major-type units, 281 1:5000 units, and 155 1:20,000 units. The mapping followed standard NiN guidelines (see appendix S6 in Halvorsen et al., 2020), which for mapping to 1:5000 are a minimum area of 250 m^2 and a minimum polygon width of 7 m. For mapping to 1:20,000, the corresponding numbers are 2500 m^2 and 20 m respectively.

2.4 | Data collection, calibration, and material used for mapping

Mapping was conducted in August 2020 (northeastern region) and August 2021 (western region). Each year, the day before mapping was spent on calibration in the field, outside the areas to be mapped. By discussion, the aim was to achieve a common understanding of classification and delineation of dominating thematic units in each region. Thereafter, comparison of results or discussion of site-specific topics was not allowed.

Each mapper constructed one ecosystem map of each designated area, with wall-to-wall coverage, by delineating polygons on a field computer equipped with a built-in GPS receiver and a QGIS with forms specifically adapted for the study (see https://github.com/geco-nhm/NiN_QGIS_3.x). The application provided an overview map (N50; scale 1:50,000) and orthophoto imagery at different resolutions and time points. Digital elevation models (DEM) and digital terrain models (DTM) were only available for the western region (see Appendix S2 for sources of background layers). All NiN system documentation (Bratli et al., 2017, 2019) was available.

2.5 | Quality control and rasterisation

Technical errors (topology errors etc.) were identified and manually corrected with the core plugin *Topology Checker* in QGIS version 3.20 before analyses in R version 4.1.1. Packages used for uploading shape files and rasterisation were *rgdal* and *raster*, while landscape indices were generated with *landscapemetrics*. Zero-skewness transformation of variables and multivariate analyses were performed with *e1071* and *vegan* respectively. The data set consisted of 48 maps, each rasterised to 1 m × 1 m pixels. For each pixel, one categorical variable was recorded, providing the identity of the mapping unit that covered the largest fraction of the pixel.

2.6 | Map-inconsistency indices

For each map, the area assigned to each mapping unit and all landscape indices from the R package *landscapemetrics* at the landscape level were calculated. To be comparable across sites only indices independent of site size (e.g. patch density) were included in further analyses. The density of polygons below the minimum mapping area in the study area (except polygons cut by the landscape edges) for each scale (1:5000: 250 m^2, 1:20,000: 2500 m^2) were counted in vector-format maps. For each pair of 'comparable' maps (i.e., all map pairs from the same site mapped to the same scale; $15 \times 4 \times 2 = 120$), inconsistency was calculated for all landscape indices and density of polygons below minimum mapping area as the absolute value of the difference between the values for a given index. In addition, area inconsistency was calculated as the sum of absolute differences in area assigned to each mapping unit divided by the size of the study area. Each number was then divided by two to avoid double counting.

For each 'comparable' map pair, a confusion matrix was generated with cells providing the number of pixels assigned to each combination of mapping units in corresponding pixels. Thematic inconsistency M was quantified as

$$M = 1 - \frac{\sum_{i=1}^{d} x_i}{\sum_{j=1}^{e} y_j}$$

where x is the ith diagonal element in a matrix with d diagonal elements, and y is the jth element in the same matrix, with e elements in total. Thematic inconsistency was calculated for original maps (mapping-unit inconsistency), maps with mapping units aggregated to major-type units (major-type-unit inconsistency), and in two non-overlapping parts of Guollemuorsuolu (Figure 2). This subdivision was motivated by indications of strong variation in mapping-unit inconsistency within this site in a pilot study (see Appendix S3).

Thematic inconsistency was also calculated for individual mapping units across all 'comparable' maps. First, all confusion matrices were added separately for each combination of site and scale. Then each 'summary matrix' was made symmetric by adding its transpose and dividing the diagonal by two to avoid double counting. Next, each element was standardised by division with the respective column sum. An inconsistency matrix was finally obtained by subtraction of each element from one. All thematic inconsistency values were converted from proportions to percentages.

Mapping-unit pairs, regardless of belonging to the same or different major types or different major-type groups, were characterised by the ecological distance (ED) separating them (see Eriksen et al., 2018; Haga et al., 2021). The data files and R scripts for generating the ED matrix, providing the ED between all combinations of mapping units are available on GitHub (https://github.com/geco-nhm/NiN_ecological_distance). Residual ED between comparable maps was calculated by multiplication of the confusion matrix with the ED

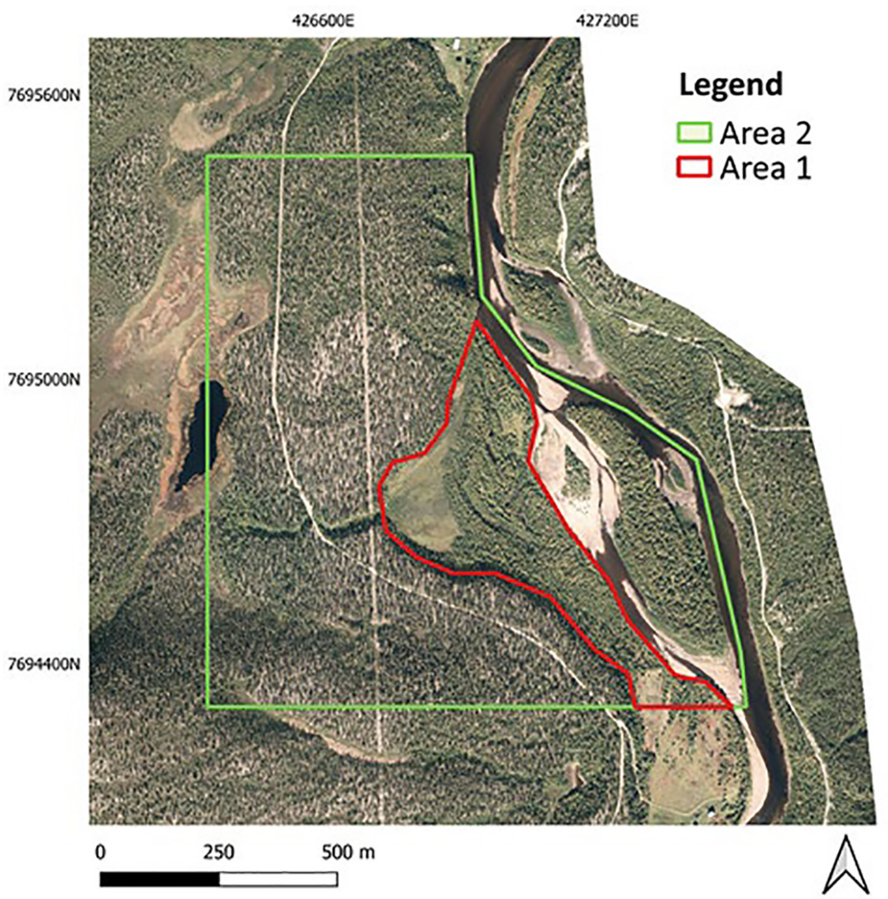

FIGURE 2 Subdivision of the Guollemuorsuolu site into non-overlapping areas for analysis of within-site variation in thematic inconsistency. The background image was obtained from the Norwegian Mapping Authority (ETRS89, UTM zone 33 N).

matrix and division by the number of inconsistent pixels between the two corresponding maps.

The map-inconsistency indices were assigned to five groups. One addresses the overall spatial inconsistency (thematic inconsistency and ED) while the others address inconsistency in composition (area and number of ecosystem units), spatial configuration of the landscape (e.g. degree of fragmentation, clumpiness or connectivity), shape of patches, and violation of mapping guidelines (i.e., minimum mapping area and width of polygons). See Appendix S4 for a description of map-inconsistency indices and their grouping.

2.7 | Multivariate analyses

Principal components analysis (PCA; Legendre & Legendre, 2012) was carried out on a matrix with pairwise comparisons of 120 maps as rows and 24 map indices as variables in columns. Each element in the matrix corresponds to the value of one inconsistency index calculated for one map pair. Before multivariate analyses, all variables were zero-skewness were transformed in accordance with Økland et al. (2001) to remove heteroscedasticity and level out differences in the variation among variables. Kendall's rank correlation coefficients (Kendall, 1938) were calculated between all variable pairs to identify sets of highly intercorrelated variables ($|\tau| > 0.7$) and one in each set was retained for further analyses. Additional variables were excluded due to many zero observations or because they were considered strongly similar to another variable (e.g. although not highly correlated, mutual information resembles relative mutual information). Global non-metric multidimensional scaling (GNMDS; Minchin, 1987) with four axes was used as complementary ordination method in parallel with PCA, and Kendall's rank correlation coefficients between GNMDS and PCA axes were calculated to identify 'true' gradients in the data, in accordance with van Son and Halvorsen (2014). Kendall's rank correlation coefficients were also calculated between all variables and the principal components to identify variables most strongly related to each axis.

Redundancy analysis (RDA; Rao, 1964) was performed on the same matrix of inconsistencies as used for PCA, with seven experience indices as constraining variables, generated from a questionnaire on mapping experience answered by the mappers after mapping was completed. The experience indices address knowledge of species (plants, mosses, lichens and fungi), experience with wall-to-wall mapping, relevant experience adapted to the respective site and scale, etc (see descriptions and questionnaire in Appendix S1). A matrix of pairwise comparisons between the mappers was constructed, with rows corresponding to the rows in the matrix used for the PCA and differences with respect to the experience indices and time consumption as columns. The elements of this matrix provide information about differences in experience or time consumption between mapper pairs. The variables were zero-skewness-transformed and one in each set of highly intercorrelated variables ($|\tau| > 0.7$) was retained. Together with site, scale and the difference in experience level (factor variable with six levels corresponding to the experience levels of a mapper pair), these variables were added to the constrained ordination model in a forward selection manner. The significant variable with the highest explanatory power quantified with the *F*-test was added as condition variable in subsequent cycles, and other variables were added until none significantly explained residual variation. *p*-Values were calculated under the null hypothesis that the constraining variable did not explain more variation than *s* random permutations of the variable ($\alpha \leq 0.0001$, $s = 9999$). Interactions were tested once all main effects were included in the model.

Significant differences ($\alpha \leq 0.05$) between the mean mapping-unit inconsistency and sites, mappers, and experience level of each mapper pair was tested by unpaired two-tailed *t*-tests with unequal variances (Legendre & Legendre, 2012). The same test was used to analyse differences between scales, averaged over all sites and for each site.

3 | RESULTS

Averaged over sites and scales, area inconsistency was $36 \pm 12\%$ (mean $\pm$ SD), and $22 \pm 14\%$ when maps were aggregated to the major-type level. Averaged over sites and scales, mapping-unit inconsistency was $51 \pm 13\%$, while major-type-unit inconsistency was $35 \pm 14\%$ (Table 2).

Mapping-unit inconsistency varied among sites, being highest at Iškoras ($61 \pm 11\%$, *t*-test: *p*-value = 0.0001), followed by Landsvik ($55 \pm 10\%$, *t*-test: *p*-value = 0.13), Guollemuorsuolu ($45 \pm 13\%$, *t*-test: *p*-value = 0.04) and Lygra ($44 \pm 7\%$, *t*-test: *p*-value < 0.0001). Averaged over scales, mapping-unit inconsistency was $80 \pm 14\%$ in Area 1 and $39 \pm 15\%$ in Area 2 at Guollemuorsuolu subsites. Residual ED between maps varied among sites and increased in the order: Iškoras (2.2 ± 0.3 EDU), Guollemuorsuolu (2.6 ± 0.7 EDU), Lygra (2.8 ± 0.4 EDU) and Landsvik (3.3 ± 0.2 EDU).

At each site and scale, between zero and four thematic units had lower than 50% thematic inconsistency (Table 3). These were typically dominant thematic units, such as Forest at Guollemuorsuolu and Coastal heath at Lygra. Compared to the mapping units they contain aggregation to major-type units reduced thematic inconsistency, especially at Iškoras and Guollemuorsuolu for Forest and Open fen. Averaged over all sites and scales, the median thematic inconsistency for individual units was 98.9% and 92.5% at the mapping unit and the major-type unit levels respectively.

Mapping-unit inconsistency did not differ significantly between scales averaged over all sites, but differed significantly between scales at Guollemuorsuolu (1:5000; $53 \pm 9\%$, 1:20,000; $38 \pm 13\%$, *t*-test of difference: *p*-value = 0.0008) and Landsvik (1:5000; $50 \pm 9\%$, 1:20,000; $59 \pm 9\%$, *t*-test of difference: *p*-value = 0.007). Aggregation of maps to the major-type level reduced thematic inconsistency with 14, 25, 5, and 10 percentage points at Iškoras, Guollemuorsuolu, Landsvik, and Lygra respectively. In all cases, 1:5000 maps were less inconsistent than 1:20,000 maps when all maps were aggregated to the major-type level.

TABLE 2 Mean and standard deviation of thematic inconsistency (%) for mapping units and major-type units at each combination of scale and site ($n = 15$ in each group), and averaged over scales ($n = 30$), sites ($n = 60$) and all pairwise comparisons ($n = 120$).

	Original scale	Iškoras	Guollemuorsuolu	Landsvik	Lygra	Averaged over sites
Mapping units	1:5000	60 ± 10	53 ± 9	50 ± 9	46 ± 4	52 ± 10
	1:20,000	62 ± 13	38 ± 13	59 ± 9	41 ± 9	50 ± 15
	Averaged over scales	61 ± 11	45 ± 13	55 ± 10	44 ± 7	51 ± 13
Major-type units	1:5000	29 ± 7	19 ± 4	43 ± 9	32 ± 6	31 ± 11
	1:20,000	44 ± 9	20 ± 5	57 ± 11	36 ± 10	39 ± 16
	Averaged over scales	37 ± 11	20 ± 4	50 ± 12	34 ± 8	35 ± 14

The standard deviation in average mapping-unit inconsistency for individual mappers was 5%. Two mappers had significantly lower mapping-unit inconsistency than the mean (mappers 10 and 11), whereas one had significantly higher (mapper 14). The lowest mapping-unit inconsistency was found between experienced mappers (42 ± 11%, *t*-test: *p*-value = 0.05) while the highest was found between beginners and mappers with intermediate experience (55 ± 13%, *t*-test: *p*-value = 0.17). All other combinations of experience-level groups had mapping-unit inconsistencies between 50% and 51% (see full results for thematic inconsistency in Appendix S5).

3.1 | Multivariate analyses

The first three principal components (PC) explained 19.4%, 11.8% and 10.2% of the variation in the data set. The correlation between PC1 and GNMDS1 was $|\tau| = 0.82$, between PC2 and GNMDS3 $|\tau| = 0.55$, and between PC3 and GNMDS2 $|\tau| = 0.70$. The PCA plot for PC1 and PC2 revealed a gradient from left to right in Figure 3, running from Landsvik via Iškoras to Guollemuorsuolu and Lygra, with only Guollemuorsuolu separated from the other sites along PC2. The two scales were not separated along PC1, although they were separated to variable extents within sites. Along PC2, and particularly PC3, the two scales were separated. PC1 was most strongly correlated with indices of overall spatial inconsistency and inconsistency in landscape composition [highest correlation with ED ($|\tau| = 0.65$) and major-type-unit inconsistency ($|\tau| = 0.62$)]. PC2 was most strongly correlated with inconsistency indices of landscape spatial configuration and violation of mapping guidelines [inconsistency in patch density, mean radius of gyration, density of polygons below minimum mapping area, edge density, relative mutual information, and standard deviation of radius of gyration ($|\tau| = 0.30$–0.54)]. PC3 was correlated at $|\tau| = 0.30$–0.54 with inconsistency indices of landscape spatial configuration (inconsistency in largest patch index, effective mesh size, connectance, interspersion and juxtaposition, and patch density) (see full results in Appendix S6).

Site, scale and their interaction were selected as constraining variables during forward selection with RDA (*F*-test: *p*-value < 0.0001). The fraction of the total variation explained in RDA was 17.9% for sites, 5.4% for scales and 6.3% for their interaction.

4 | DISCUSSION

Area inconsistency was 36% for mapping units and 22% for major-type units, while thematic inconsistency for the corresponding hierarchical levels was 51% and 35%. These results are likely to be considered unacceptable if the maps' spatial properties are intended for practical use, for example for management of specific areas. Moreover, they show that wall-to-wall maps have less uncertainty if area coverage rather than exact spatial distributions of ecosystems is to be reported. In surveys with the purpose of obtaining area-representative statistics, ecosystems are normally classified in systematically placed points (e.g. Bryn et al., 2018; Breidenbach et al., 2020; d'Andrimont et al., 2020). The area inconsistency values reported here are in the same range as those reported for point sampling with the same mapping system (43% for mapping units and 19% for major-type units respectively) by Eriksen et al. (2018). Thus, area statistics for systematic ecosystem accounting and other uses can be obtained by similar reliability from wall-to-wall maps as from point-mapping methods.

Previous studies have reported widely different results for thematic inconsistency in field-based maps. Thematic inconsistency values in the low range (9%–17%) are typically found when intensive calibration of mappers from the same organisation is combined with low thematic resolution (Barr et al., 1993; Kershaw & Bunce, 1998; Stevens et al., 2004), while the converse (non-calibrated mappers from different organisations combined with high thematic resolution) characterises high thematic inconsistency (66%–82%) reports (Cherrill & McClean, 1999; Hearn et al., 2011). All these studies used experienced mappers. Our study reveals thematic inconsistency values intermediate between these extremes (ranging from 19% to 62%), varying among mappers, sites and aggregation levels of mapping units. This makes sense, since our calibration effort was low to moderate, our mappers were mainly from the same organisation and the thematic and spatial resolutions were high to very high (depending on the map scale and generalisation level of the type hierarchy).

4.1 | Sites

Mapping-unit inconsistency ranged between 44% and 61% across sites, the highest observed at Iškoras. Averaged over scales, 21% of

TABLE 3 Mean and standard deviation of thematic inconsistency (%) for all mapping units and major-type units with thematic inconsistency lower than 50%.

	Iškoras		Guollemuorsuolu		Landsvik		Lygra	
Mapping units	**1:5000**	**1:20,000**	**1:5000**	**1:20,000**	**1:5000**	**1:20,000**	**1:5000**	**1:20,000**
Lichen and heath forest	-	-	-	29 ± 20	-	-	-	-
Lime-poor subxeric coastal heath	-	-	-	-	-	-	36 ± 5	-
Lime-poor submesic to subxeric coastal heath	-	-	-	-	-	-	-	42 ± 4
Agriculturally improved pasture	-	-	-	-	-	-	-	39 ± 14
Major-type units								
Open shallow-soil ground	-	-	-	-	44 ± 4*	-	-	-
Forest	24 ± 3	34 ± 12	12 ± 3	12 ± 4	-	-	-	-
Open alluvial sediment	-	-	41 ± 12	48 ± 34*	-	-	-	-
Coastal heath	-	-	-	-	-	-	31 ± 6	42 ± 14
Agriculturally improved grassland	-	-	-	-	-	-	-	33 ± 11
Open fen	43 ± 12	-	41 ± 9	48 ± 31	-	-	-	-
Tree plantation	-	-	-	-	-	-	21 ± 9*	16 ± 4*
Freshwater bottom systems	15 ± 8*	-	27 ± 10*	34 ± 13*	24 ± 4*	-	-	-
Marine seabed systems	-	-	-	-	-	-	26 ± 4*	25 ± 4*

Note: Major-type unit inconsistencies with equal thematic inconsistency for mapping units are denoted by an asterisk.

the mapping-unit inconsistency at this site was attributed to confusion between Wetland and Terrestrial systems (see Appendix S5). This is not surprising, given that the discussion about criteria for separating these major-type groups remained inconclusive and the areas in question were therefore mapped as Mountain heath by some, and Bog or Open fen by others. The higher mapping-unit inconsistency at Iškoras is most likely a result of disagreement among mappers, unsettled during calibration. This inability to converge on a classification consensus during discussion in the field illustrates the challenges with obtaining accurate classification with high certainty. Classification requires information on for instance land-use history, tree cover, peat production, and the presence and abundance of species, meaning that accurate classification with complete certainty is virtually impossible in practice due to the lack of information. Consistency is more achievable than accuracy because it only requires that the information available to mappers is interpreted equally. In a review article, Cherrill (2013) points to the much higher thematic inconsistency of non-calibrated mappers than of calibrated mappers. This suggests that proper calibration has the potential to be a major inconsistency-reduction measure, to avoid mapper-specific interpretation and practical use of mapping guidelines. Our experience suggests that a calibration leader with the authority to settle disagreements is required. In large mapping projects, calibration of leaders will be necessary to avoid development of different 'schools'.

The lowest mapping-unit inconsistency was observed at Lygra, which is largely dominated by one thematic unit, Lime-poor coastal heath. A likely explanation is that the homogeneity of the area entails fewer borders and thus less potential for delineation and generalisation inconsistencies. As a result, thematic inconsistency propagating from delineation and generalisation inconsistencies could be reduced at Lygra due to dominance of one thematic unit. The maps from Guollemuorsuolu illustrate the same point. Area 1 was fragmented, with a mire and interspersed forested areas subjected to flooding disturbance and former farming activities of varying intensities. In contrast, most of Area 2 was dominated by xeric pine forest with clear borders to surrounding ecosystems. These results also accord with results of thematic inconsistency for individual units, which indicates its clear relationship to dominance and distinctiveness of borders. This is exemplified by the lower thematic inconsistency for Forest at Guollemuorsuolu, Tree plantation and Agriculturally improved grassland at Lygra (a small patch in the middle of the open coastal heath landscape), in addition to Freshwater and Marine bottom system units. Another consequence is that thematic inconsistency was extremely high for virtually all rare thematic units. Previous studies have traditionally related high thematic inconsistency to specific thematic units (Stevens et al., 2004; Ullerud et al., 2018; Haga et al., 2021), although with contrasting results for which ecosystems are the most 'difficult' to map. Our results suggest that dominance of individual thematic units and the distinctiveness of borders are also important for thematic inconsistency. The reasons for thematic inconsistency are therefore probably more context-dependent and not as strongly related to specific thematic units as previous studies suggest.

Relative to the other sites, the residual ED at Iškoras and Guollemuorsuolu was low (i.e., mapping units separated by few EDUs was relatively frequently confused). Several studies report confusion of neighbouring or ecologically similar thematic units as the most

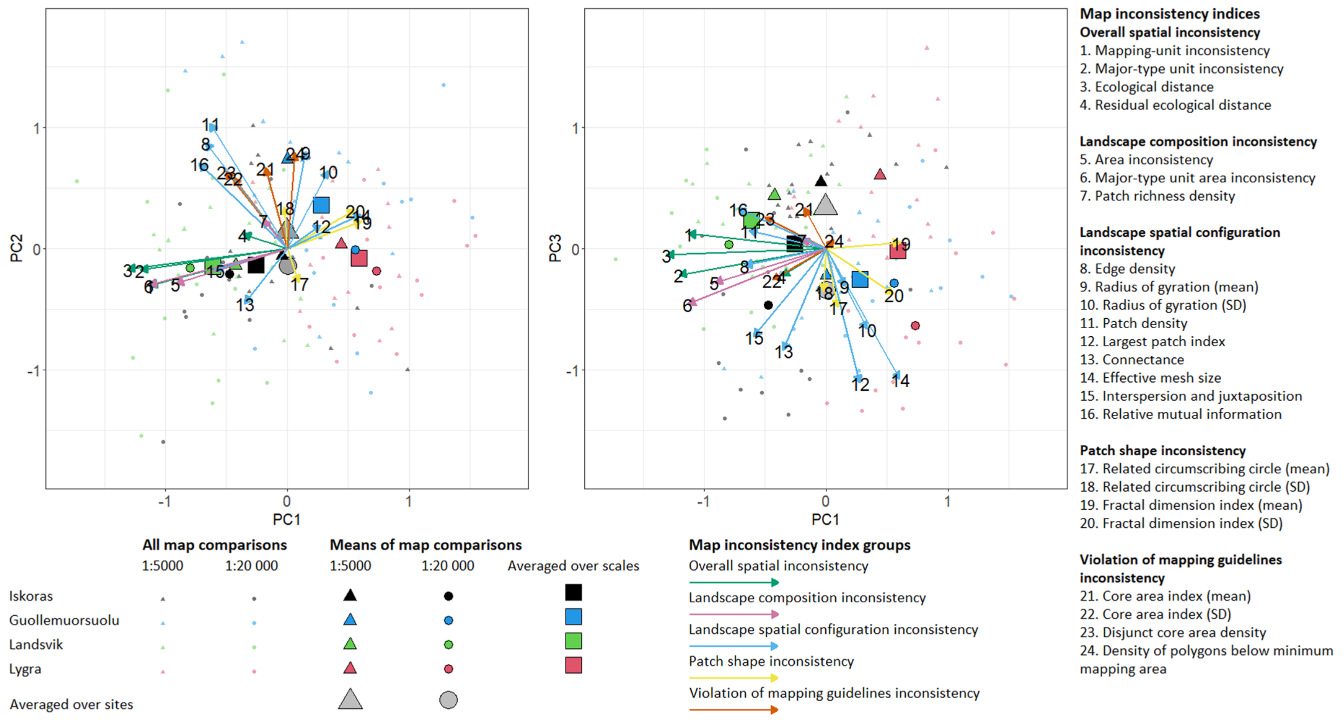

FIGURE 3 Principal components analysis of pairwise map comparisons (PC1 and PC2 to the left; PC1 and PC3 to the right). Arrows point in the direction of maximum increase of inconsistencies for a given map index (see Appendix S4 for descriptions of these).

common type of thematic inconsistency (Cherrill & McClean, 1999; Hearn et al., 2011; Ullerud et al., 2018; Haga et al., 2021), but the prevalence of such inconsistencies also depend on other factors such as site, scale, mapper and interactions between these. This can be exemplified by the most dominating (and most frequently confused) mapping unit at Landsvik, Lime-poor submesic to subxeric forest. This mapping unit was most often confused with Lime-poor submesic boreal heath and Tree plantation. The two mapping-unit pairs represent separations of four and six EDUs respectively, but in reality, the confused areas were similar, both ecologically and in appearance due to their successional state at the time of mapping. In such situations, distinguishing thematic units at lower hierarchical levels, separated by few EDUs, might be as difficult as distinguishing those at higher hierarchical levels. Thus, this interaction between site properties (i.e., variation in succession) and mapping system structure (i.e., aggregation of mapping units) may have affected our results, resulting in the higher residual ED at Landsvik and Lygra.

4.2 | Scales

The reduction of thematic inconsistency when mapping units were aggregated to higher hierarchical levels was 25 percentage points at Guollemuorsuolu, and well below that at the other sites (e.g. five percentage points at Landsvik). Lower thematic resolution entails lower thematic inconsistency, but the magnitude of this effect depends on site-specific properties such as which mapping units are present and how they are aggregated into higher hierarchical units. The likely stronger effect of lowering thematic resolution at Guollemuorsuolu might explain the higher reduction in thematic inconsistency with coarser map scale at Guollemuorsuolu relative to the other sites. The significantly lower mapping-unit inconsistency in 1:20,000 maps observed at Guollemuorsuolu agrees with findings in previous studies, showing increasing thematic inconsistency with finer map scales (i.e., thematic resolution and spatial resolution) (Ullerud et al., 2018).

In contrast to these findings, mapping-unit inconsistency was significantly higher in 1:20,000 than 1:5000 maps at Landsvik. Moreover, aggregated 1:5000 maps were always less inconsistent than 1:20,000 maps, when all maps were aggregated to the major-type unit level. These differences may be attributed to differences in spatial resolution during field mapping. A possible explanation of why thematic inconsistency decreases with higher spatial resolution is that problem solving and mapping-related decisions at finer spatial resolutions induce higher precision.

Several properties of our study setup may have influenced these results. Our mappers had more experience with 1:5000 compared to 1:20,000 mapping, and all but two mappers had equal or more years of experience with 1:5000 mapping. Moreover, NiN species lists, convenient for classifying mapping units, only exist for 1:5000 units. University courses and Norwegian Environmental Agency's (2022) mapping manual used for mapping projects in Norway are also mainly oriented towards mapping at this scale. Time constraints may also apply. The mappers had three days to finish 1:5000 maps and one day to finish 1:20,000 maps for the same area. More time available for mapping allows for denser walking patterns, less reliance on aerial photo interpretation, more time spent on decision-making in the field, etc. Alternatively, higher thematic inconsistency in aggregated 1:20,000 maps might be caused by variation between mappers. Because a different combination of mappers mapped the two scales at each site, it is difficult to separate the effects of these factors.

4.3 | Differences in mapping experience

We found a difference in mapping-unit inconsistency between the group of most experienced mappers and the mean of all mappers of nine percentage points. This accords with the result of Hearn et al. (2011), who found correlation between the average thematic inconsistency of each mapper and years of experience. In our analysis, we instead related pairwise inconsistencies with differences in experience. In our opinion, this is more sensible as inconsistency is a measure of differences. This rationalises for using mapper differences according to variables instead of the raw variables themselves as explanatory variables. Furthermore, it allows the separation of within-group and between-group variation.

For effective organisation of calibration programmes, experience must be decomposed into specific traits of relevance to inconsistency reduction. A common view among environmental managers and ecologists is that lack of experience is important for thematic inconsistency because they think beginners overlook indicator species (Cherrill, 2016). In our study, specific mappers deviated significantly from the mean mapping-unit inconsistency, but still we detected no relationships between mapping-unit inconsistency and differences in knowledge of species or other components of mapping experience (e.g. experience with wall-to-wall mapping, or aerial photo interpretation). Other mapper-dependent factors, such as background or personality traits, may be more important but were not analysed. Eriksen et al. (2018) found a positive relationship between correct assignment of thematic units to points and knowledge of species, mapping units and major-type units. However, wall-to-wall mapping is more complex than point mapping since it requires skills in delineation and generalisation, in addition to classification. Possibly, no simple component of mapping experience will explain a large fraction of variation in map inconsistencies. Field-based wall-to-wall ecosystem mapping is complex and requires a combination of skills (e.g. botanical, geological, cartographic skills and their interactions). In addition, there are difficulties with quantifying all relevant skills.

4.4 | Multivariate analyses

Our results demonstrate the utility of multivariate analyses for disentangling complex relationships behind variation in map inconsistency, as maps can be characterised by multiple indices. PC1 was interpreted as a gradient in overall spatial inconsistency and inconsistency in landscape composition. The positive correlations between indices from these groups are not surprising as area inconsistencies are directly affected by thematic inconsistencies, and thus, reducing thematic inconsistencies improves the reliability of both the spatial and the compositional information in the maps.

PC2 was strongest correlated with inconsistencies in violation of mapping guidelines and landscape spatial configuration, the latter group also correlated strongest with PC3. Information on spatial configuration of the landscape is often required for assessment of landscape function (Frank et al., 2012). For instance, effective mesh size quantifies the degree of fragmentation, which is used to assess the provision of ecosystem services, such as biodiversity, habitat provision, and recreation potential (Dobbs et al., 2014), as well as land-use and land-cover changes (Uuemaa et al., 2013). Most of the inconsistency indices of landscape spatial configuration and violation of mapping guidelines are indicators of the degree of map detail suggesting that mappers have different tendencies to generalise during mapping. In a previous study, classification and delineation were found to contribute equally to errors in field-based ecosystem maps (Haga et al., 2021). To our knowledge, no studies have attempted to quantify the relative contribution of generalisation errors. Our results suggest generalisation errors may account for considerable uncertainty in field-based ecosystem mapping, and that these errors may reduce the reliability of analyses using landscape spatial configuration indices.

Visual inspection of the PCA plot indicates the presence of systematic differences in map inconsistencies across sites, scales and between scales within sites. This was confirmed by RDA results, identifying site, scale, and their interaction as the most important predictors of map inconsistencies. Scales were only separated along PC2 and PC3, indicating that the effect of scale across sites most likely resulted from inconsistencies in landscape spatial configuration and violation of mapping guidelines.

4.5 | Recommendations and implications for research and applications

A take-home message from this study is that map inconsistencies often are highly context-dependent, resulting in our inability to separate the effects of scale differences from other effects on mapping-unit inconsistency (e.g. the combination of mappers). These context dependencies should be taken into consideration in the planning of future experiments. Our results from PCA showed that considerable variation among maps was manifested in the inconsistency indices for landscape spatial configuration. Quantification of these inconsistencies and identification of their relationships to other factors is outside the scope of this study. However, spatial configuration is often taken into account when ecosystem maps are used in practice, and this aspect should therefore be investigated further.

Our findings confirm and strengthen those reported in previous studies substantiating that high inconsistency is typical of contemporary field-based ecosystem mapping. As areas with high inconsistency might require more calibration for accurate mapping, knowledge of relations between inconsistency and specific landscape patterns and/or thematic units should guide calibration. Errors in ecosystem maps may lead to inaccurate assessments of ecological site value, erroneous planning and management decisions, negative consequences for biodiversity and economic losses for parties involved (Cherrill, 2016). Therefore, every organisation delivering and/or ordering field-based ecosystem mapping services should perform and/or request quality controls (e.g. by giving mappers slightly overlapping mapping areas intended for map-inconsistency analysis) and carry out a cost–benefit analysis of inconsistency reduction. Organisations finding reduction of map inconsistencies to be cost-effective will benefit from availability of an inconsistency reduction protocol for field-based ecosystem mapping. We are unaware of any published protocols made specifically for that purpose, although many mapping systems are equipped with general guidelines (e.g. NiN; Bryn & Naas, 2021 or Phase 1; JNCC, 2004).

More research should be devoted to how field-based mapping can be complemented with other methods (e.g. remote sensing). For instance, some areas may be mapped accurately with minimal mapping effort if field-based and remote-sensing methods are combined, allowing more strategic resource allocation (Stenzel et al., 2014; Neumann, 2020). Remote-sensing methods may also be used in synergy with field-based methods to map gradual transitions between ecosystems, which is difficult or costly to map with field-based methods alone (Raab et al., 2018). Since transitions in nature are gradual, this would give more realistic maps.

Our results on inconsistency provide indications of inaccuracy, which is usually quantified as the average difference of maps from a reference map considered to be 'true' (Haga et al., 2021). Using inconsistency as a performance measure avoids complications with establishing the 'truth'. No gold standard for making accurate maps exists, although independent field-based mapping followed by a consensus process to agree on a reference map has been used previously (Haga et al., 2021). Such consensus maps should be produced by groups of mappers with diverse backgrounds, as field-based ecosystem mapping requires a comprehensive and diverse skill set. A comparison of reference maps made by independent groups may reveal if these are subject to similar magnitudes of inconsistency as the maps in our study.

Conservationists, modellers, and other end users should be critical of data sources and careful when using data collected without consideration of inconsistencies and that are not quality checked. Particularly, this applies to data used as 'ground truth', for example for decision-making or as evaluation data for distribution models. Results obtained from these methods (e.g. reported model accuracies, estimated land cover and fragmentation changes) should be interpreted conservatively and/or uncertainties propagating from the 'ground truth' should be taken into account in the analyses. This applies particularly to analyses performed in fragmented and rare ecosystems. Further research should test if and how inaccuracies in maps propagate to other maps using field-based maps in one or more methodological steps, and how inaccuracies affect nature management (e.g. area conservation, spatial planning, temporal change detection; see Cherrill, 2016).

5 | CONCLUSIONS

Of the factors addressed in our study, among-site differences explain most variation, which probably is related to dominance of thematic units and variation in the distinctiveness of borders between them. Scale only influences mapping-unit inconsistency in a site-dependent manner, partly because of interactions between properties of the sites and mapping system. Except for the most experienced group of mappers, who have significantly lower mapping-unit inconsistency than the mean, differences in experience have no documented effect.

Fine-scale ecosystem maps depict the spatial distribution of ecosystems in detail. This makes them highly valuable for use in local decision-making processes, although the spatial uncertainty in wall-to-wall field-based ecosystem maps often may be considered unacceptable for such purposes. Reports of low thematic inconsistency do exist (e.g. Stevens et al., 2004), which open the possibility of factors explaining large fractions of variation in inconsistency. More controlled experiments are needed to disentangle and gain more certainty on the effect of these factors on inconsistency in field-based ecosystem mapping.

AUTHOR CONTRIBUTIONS

Anders Bryn and Peter Horvath developed the idea and study design, while Rune Halvorsen also contributed with the latter. Anders Bryn, Anders Kvalvåg Wollan and Peter Horvath organised the fieldwork. Anders Bryn, Adam Eindride Naas, Michal Torma and Rune Halvorsen developed the questionnaire. Quality control of the maps was conducted by Adam Eindride Naas and Peter Horvath. Adam Eindride Naas conducted the analyses, with advice from Anders Bryn, Peter Horvath and Rune Halvorsen. He also wrote the paper, with contributions from Anders Bryn, Eva Lieungh, Espen Sommer Værland, Peter Horvath, Rune Halvorsen, and Trond Simensen. All authors contributed with fieldwork, data collection, discussed the results and commented on the manuscript.

ACKNOWLEDGEMENTS

Anne-Barbi Nilsen is acknowledged for providing field computers for mapping. The study was supported financially by the Norwegian biodiversity information Centre.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Adam Eindride Naas https://orcid.org/0000-0001-9736-9020
Rune Halvorsen https://orcid.org/0000-0002-6859-7726
Peter Horvath https://orcid.org/0000-0002-6017-5385
Anders Kvalvåg Wollan https://orcid.org/0000-0003-3947-9402
Harald Bratli https://orcid.org/0000-0002-0193-1259
Katrine Brynildsrud https://orcid.org/0000-0003-4310-2608

Eirik Aasmo Finne https://orcid.org/0000-0002-8559-3655
Lasse Torben Keetz https://orcid.org/0000-0002-8040-4122
Eva Lieungh https://orcid.org/0000-0003-4009-944X
Christine Olson https://orcid.org/0000-0002-4679-6281
Trond Simensen https://orcid.org/0000-0002-7022-1349
Olav Skarpaas https://orcid.org/0000-0001-9727-1672
Hilde Riksheim Tandstad https://orcid.org/0000-0003-3218-6500
Espen Sommer Værland https://orcid.org/0000-0002-9424-1950
Anders Bryn https://orcid.org/0000-0003-4712-8266

REFERENCES

Bailey, R.G. (1987) Suggested hierarchy of criteria for multi-scale ecosystem mapping. *Landscape and Urban Planning*, 14, 313–319. Available from: https://doi.org/10.1016/0169-2046(87)90042-9

Bakkestuen, V., Erikstad, L. & Halvorsen, R. (2008) Step-less models for regional environmental variation in Norway. *Journal of Biogeography*, 35, 1906–1922. Available from: https://doi.org/10.1111/j.1365-2699.2008.01941.x

Barr, C.J., Bunce, R.G.H., Clarke, R.T., Fuller, R.M., Furse, M.T., Gillespie, M.K. et al. (1993) *Countryside survey 1990 Main report*. London: Department of the Environment.

Beier, P., Hunter, M.L. & Anderson, M. (2015) Introduction. *Conservation Biology*, 29(3), 613–617. Available from: https://doi.org/10.1111/cobi.12511

Bratli, H., Halvorsen, R., Bryn, A., Arnesen, G., Bendiksen, E., Jordal, J. et al. (2019) *Beskrivelse av kartleggingsenheter i målestokk 1:5000 etter NiN versjon 2.2.0*. Trondheim: Norwegian Biodiversity Information Centre.

Bratli, H., Halvorsen, R., Bryn, A., Bendiksen, E., Jordal, J., Svalheim, E. et al. (2017) *Dokumentasjon av NiN versjon 2.1 tilrettelagt for praktisk naturkartlegging i målestokk 1:5000. Natur i Norge, Artikkel 8*. Trondheim: Norwegian Biodiversity Information Centre.

Breidenbach, J., Granhus, A., Hylen, G., Eriksen, R. & Astrup, R. (2020) A century of National Forest Inventory in Norway – Informing past, present, and future decisions. *Forest Ecosystems*, 7, 1–19. Available from: https://doi.org/10.1186/s40663-020-00261-0

Bryn, A. & Naas, A.E. (2021) *Feltveileder for kartlegging av terrestrisk naturvariasjon etter NiN (2.3) – Tilpasset målestokk 1:5 000 og 1:20 000, utgave 2 oppdatert mars 2022, kartleggingsveileder nr 2*. Trondheim: Norwegian Biodiversity Information Centre.

Bryn, A., Strand, G.-H., Angeloff, M. & Rekdal, Y. (2018) Land cover in Norway based on an area frame survey of vegetation types. *Norwegian Journal of Geography*, 72, 131–145. Available from: https://doi.org/10.1080/00291951.2018.1468356

Cherrill, A. (2013) Repeatability of vegetation mapping using phase 1 and NVC approaches: Implications for professional practice and surveyors' training requirements. *In Practice*, 81, 41–45.

Cherrill, A. (2016) Inter-observer variation in habitat survey data: Investigating the consequences for professional practice. *Journal of Environmental Planning and Management*, 59, 1813–1832. Available from: https://doi.org/10.1080/09640568.2015.1090961

Cherrill, A. & McClean, C. (1995) An investigation of uncertainty in field habitat mapping and the implications for detecting land cover change. *Landscape Ecology*, 10, 5–21. Available from: https://doi.org/10.1007/BF00158550

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. Available from: https://doi.org/10.1046/j.1365-2664.1999.00458.x

d'Andrimont, R., Yordanov, M., Martinez-Sanchez, L., Eiselt, B., Palmieri, A., Dominici, P. et al. (2020) Harmonised LUCAS in-situ land cover and use database for field surveys from 2006 to 2018 in the European Union. *Scientific Data*, 7, 1–15. Available from: https://doi.org/10.1038/s41597-020-00675-z

Dobbs, C., Kendal, D. & Nitschke, C.R. (2014) Multiple ecosystem services and disservices of the urban forest establishing their connections with landscape structure and sociodemographics. *Ecological Indicators*, 43, 44–55. Available from: https://doi.org/10.1016/j.ecolind.2014.02.007

Eriksen, E.L., Ullerud, H.A., Halvorsen, R., Aune, S., Bratli, H., Horvath, P. et al. (2018) Point of view: Error estimation in field assignment of land-cover types. *Phytocoenologia*, 49, 135–148. Available from: https://doi.org/10.1127/phyto/2018/0293

European Commission. (2013) *Interpretation manual of European Union habitats, version EUR 28*. European Commission, DG-ENV. Available from: http://ec.europa.eu/environment/nature/legislation/habitatsdirective/docs/Int_Manual_EU28.pdf [Accessed 16 February 2023]

Frank, S., Fürst, C., Koschke, L. & Makeschin, F. (2012) A contribution towards a transfer of the ecosystem service concept to landscape planning using landscape metrics. *Ecological Indicators*, 21, 30–38. Available from: https://doi.org/10.1016/j.ecolind.2011.04.027

Franklin, J. (2013) Mapping vegetation from landscape to regional scales. In: van der Maarel, E. & Franklin, J. (Eds.) *Vegetation ecology*, 2nd edition. Chichester: Wiley-Blackwell, pp. 486–508.

Guisan, A. & Zimmermann, N.E. (2000) Predictive habitat distribution models in ecology. *Ecological Modelling*, 135, 147–186. Available from: https://doi.org/10.1016/S0304-3800(00)00354-9

Haga, H.E.E.S., Nilsen, A.B., Ullerud, H.A. & Bryn, A. (2021) Quantification of accuracy in field-based land cover maps: A new method to separate different components. *Applied Vegetation Science*, 24, e12578. Available from: https://doi.org/10.1111/avsc.12578

Hagen, D., Evju, M., Henriksen, P.S., Solli, S., Erikstad, L. & Bartlett, J. (2022) From military training area to National Park over 20 years: Indicators for outcome evaluation in a large-scale restoration project in alpine Norway. *Journal for Nature Conservation*, 66, 126125. Available from: https://doi.org/10.1016/j.jnc.2021.126125

Halvorsen, R., Skarpaas, O., Bryn, A., Bratli, H., Erikstad, L., Simensen, T. et al. (2020) Towards a systematics of ecodiversity: The EcoSyst framework. *Global Ecology and Biogeography*, 29, 1887–1906. Available from: https://doi.org/10.1111/geb.13164

Hearn, S., Healey, J., McDonald, M., Turner, A., Wong, J. & Stewart, G. (2011) The repeatability of vegetation classification and mapping. *Journal of Environmental Management*, 92, 1174–1184. Available from: https://doi.org/10.1016/j.jenvman.2010.11.021

Hein, L., Bagstad, K.J., Obst, C., Edens, B., Schenau, S., Castillo, G. et al. (2020) Progress in natural capital accounting for ecosystems. *Science*, 367, 514–515. Available from: https://doi.org/10.1126/science.aaz8901

Hersperger, A.M., Grădinaru, S.R., Daunt, A.B.P., Imhof, C.S. & Fan, P. (2021) Landscape ecological concepts in planning: Review of recent developments. *Landscape Ecology*, 36, 2329–2345. Available from: https://doi.org/10.1007/s10980-021-01193-y

Horvath, P., Halvorsen, R., Stordal, F., Tallaksen, L.M., Tang, H. & Bryn, A. (2019) Distribution modelling of vegetation types based on area frame survey data. *Applied Vegetation Science*, 22, 547–560. Available from: https://doi.org/10.1111/avsc.12451

Ihse, M. (2007) Colour infrared aerial photography as a tool for vegetation mapping and change detection in environmental studies of Nordic ecosystems: A review. *Norwegian Journal of Geography*, 61, 170–191. Available from: https://doi.org/10.1080/00291950701709317

JNCC. (2004) *Handbook for phase 1 habitat survey – A technique for environmental audit*. Peterborough: Joint Nature Conservation Committee.

Kahneman, D., Sibony, O. & Sunstein, C.R. (Eds.). (2021) *Noise: A flaw in human judgement*. Great Britain: William Collins.

Kendall, M.G. (1938) A new measure of rank correlation. *Biometrika*, 30, 81–93. Available from: https://doi.org/10.2307/2332226
Kershaw, D.C. & Bunce, R.G.H. (1998) *Northern Ireland countryside survey 2000: Quality assurance*. Grange-over-Sands: Centre for Ecology and Hydrology.
Küchler, A.W. & Zonneveld, I.S. (Eds.). (1988) *Vegetation mapping*. Dordrecht: Kluwer Academic Publishers.
Legendre, P. & Legendre, L. (Eds.). (2012) *Numerical ecology*. Amsterdam: Elsevier.
Måren, I.E. & Vandvik, V. (2009) Fire and regeneration: The role of seed banks in the dynamics of northern heathlands. *Journal of Vegetation Science*, 20, 871–888. Available from: https://doi.org/10.1111/j.1654-1103.2009.01091.x
Minchin, P.R. (1987) An evaluation of the relative robustness of techniques for ecological ordination. *Vegetatio*, 69, 89–107. Available from: https://doi.org/10.1007/BF00038690
Mõisja, K., Uuemaa, E. & Oja, T. (2018) The implications of field worker characteristics and landscape heterogeneity for classification correctness and the completeness of topographical mapping. *International Journal of Geo-Information*, 7, 205. Available from: https://doi.org/10.3390/ijgi7060205
Moss, D. (2008) *EUNIS habitat classification – A guide for users*. Copenhagen: European Environment Agency.
Mysterud, A., Rekdal, Y., Loe, L.E., Angeloff, M., Mobæk, R., Holand, Ø. et al. (2014) Evaluation of landscape-level grazing capacity for domestic sheep in alpine rangelands. *Rangeland Ecology & Management*, 67, 132–144. Available from: https://doi.org/10.2111/REM-D-12-00177.1
Neumann, C. (2020) Habitat sampler—A sampling algorithm for habitat type delineation in remote sensing imagery. *Diversity and Distributions*, 26, 1752–1766. Available from: https://doi.org/10.1111/ddi.13165
Norwegian Environmental Agency. (2022) *Mapping manual – Mapping of terrestrial ecosystem types following NiN2*. Trondheim: Norwegian Environmental Agency.
Økland, R.H., Økland, T. & Rydgren, K. (2001) Vegetation-environment relationships of boreal spruce swamp forests in Østmarka nature reserve, SE Norway. *Sommerfeltia*, 29, 1–190. Available from: https://doi.org/10.2478/som-2001-0001
Phiri, D. & Morgenroth, J. (2017) Developments in Landsat land cover classification methods: A review. *Remote Sensing*, 9, 967. Available from: https://doi.org/10.3390/rs9090967
Raab, C., Stroh, H., Tonn, B., Meißner, M., Rohwer, N., Balkenhol, N. et al. (2018) Mapping semi-natural grassland communities using multi-temporal RapidEye remote sensing data. *International Journal of Remote Sensing*, 39, 5638–5659. Available from: https://doi.org/10.1080/01431161.2018.1504344
Rao, C.R. (1964) The use and interpretation of principal component analysis in applied research. *Sankhyā: The Indian Journal of Statistics, A*, 26, 329–359.
seNorge. (2022) Klima. Available at http://www.senorge.no/?p=klima [Accessed 12 January 2022]
Simensen, T., Horvath, P., Vollering, J., Erikstad, L., Halvorsen, R. & Bryn, A. (2020) Composite landscape predictors improve distribution models of ecosystem types. *Diversity and Distributions*, 26, 928–943. Available from: https://doi.org/10.1111/ddi.13060
Stenzel, S., Feilhauer, H., Mack, B., Metz, A. & Schmidtlein, S. (2014) Remote sensing of scattered Natura 2000 habitats using a one-class classifier. *International Journal of Applied Earth Observation and Geoinformation*, 33, 211–217. Available from: https://doi.org/10.1016/j.jag.2014.05.012
Stevens, J., Blackstock, T., Howe, E. & Stevens, D. (2004) Repeatability of phase 1 habitat survey. *Journal of Environmental Management*, 73, 53–59. Available from: https://doi.org/10.1016/j.jenvman.2004.05.009
Talukdar, S., Singha, P., Mahato, S., Pal, S., Liou, Y.-A. & Rahman, A. (2020) Land-use land-cover classification by machine learning classifiers for satellite observations – A review. *Remote Sensing*, 12, 1135. Available from: https://doi.org/10.3390/rs12071135
Tømmervik, H., Bjerke, J.W., Gaare, E., Johansen, B. & Thannheiser, D. (2012) Rapid recovery of recently overexploited winter grazing pastures for reindeer in northern Norway. *Fungal Ecology*, 5, 3–15. Available from: https://doi.org/10.1016/j.funeco.2011.08.002
Ullerud, H.A., Bryn, A., Halvorsen, R. & Hemsing, L.Ø. (2018) Consistency in land-cover mapping: Influence of field workers, spatial scale and classification system. *Applied Vegetation Science*, 21, 278–288. Available from: https://doi.org/10.1111/avsc.12368
Ullerud, H.A., Bryn, A. & Skånes, H. (2020) Bridging theory and implementation – Testing an abstract classification system for practical mapping by field survey and 3D aerial photographic interpretation. *Norwegian Journal of Geography*, 73, 301–317. Available from: https://doi.org/10.1080/00291951.2020.1717595
Uuemaa, E., Mander, Ü. & Marja, R. (2013) Trends in the use of landscape spatial metrics as landscape indicators: A review. *Ecological Indicators*, 28, 100–106. Available from: https://doi.org/10.1016/j.ecolind.2012.07.018
van Son, T.C. & Halvorsen, R. (2014) Multiple parallel ordinations: The importance of choice of ordination method and weighting of species abundance data. *Sommerfeltia*, 37, 1–37. Available from: https://doi.org/10.2478/som-2014-0001

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

Appendix S1. Questionnaire and indices for experience
Appendix S2. Sources of background layers
Appendix S3. Heat maps and ecological distance maps
Appendix S4. Map indices
Appendix S5. Thematic inconsistency results
Appendix S6. Multivariate analyses results

How to cite this article: Naas, A.E., Halvorsen, R., Horvath, P., Wollan, A.K., Bratli, H., Brynildsrud, K. et al. (2023) What explains inconsistencies in field-based ecosystem mapping? *Applied Vegetation Science*, 26, e12715. Available from: https://doi.org/10.1111/avsc.12715