# Systematic and random variation vegetation monitoring data

Milberg, P., Bergstedt, J., Fridman, J., Odell, G & Westerberg, L.

Original publication:

Milberg, P., Bergstedt, J., Fridman, J., Odell, G & Westerberg, L., Systematic and random variation vegetation monitoring data, 2008, Journal of Vegetation Science, (19), 633-644.
http://dx.doi.org/10.3170/2008-8-18423.

*Journal of Vegetation Science 19: 633-644, 2008*
*doi: 10.3170/2008-8-18423, published online 10 April 2008*

# Observer bias and random variation in vegetation monitoring data

**Milberg, Per[1*]; Bergstedt, Johan[1,4]; Fridman, Jonas[2]; Odell, Gunnar[3] & Westerberg, Lars[1,5]**

[1]*IFM Division of Ecology, Linköping University, SE-581 83 Linköping, Sweden;*
[2]*Department of Forest Resource Management, SLU, SE-901 83 Umeå, Sweden; jonas.fridman@srh.slu.se;*
[3]*Department of Forest Soils, SLU, SE-901 83 Umeå, Sweden; E-mail gunnar.odell@sml.slu.se;*
[4]*E-mail jober@ifm.liu.se;* [5]*E-mail lawes@ifm.liu.se;*
[*]*Corresponding author; E-mail permi@ifm.liu.se*

**Abstract**

**Question:** Detecting species presence in vegetation and making visual assessment of abundances involve a certain amount of skill, and therefore subjectivity. We evaluated the magnitude of the error in data, and its consequences for evaluating temporal trends.

**Location:** Swedish forest vegetation.

**Methods:** Vegetation data were collected independently by two observers in 342 permanent 100-$m^2$ plots in mature boreal forests. Each plot was visited by one observer from a group of 36 and one of two quality assessment observers. The cover class of 29 taxa was recorded, and presence/absence for an additional 50.

**Results:** Overall, one third of each occurrence was missed by one of the two observers, but with large differences among species. There were more missed occurrences at low abundances. Species occurring at low abundance when present tended to be frequently overlooked. Variance component analyses indicated that cover data on 5 of 17 species had a significant observer bias. Observer-explained variance was < 10% in 15 of 17 species.

**Conclusion:** The substantial number of missed occurrences suggests poor power in detecting changes based on presence/absence data. The magnitude of observer bias in cover estimates was relatively small, compared with random error, and therefore potentially analytically tractable. Data in this monitoring system could be improved by a more structured working model during field work.

**Keywords**: Forest; Observer error; Permanent plot; Statistical power; Sweden.

**Nomenclature source**: Karlsson (1998).

**Abbreviation**: SK = Swedish Survey of Forest Soils and Vegetation.

## Introduction

In vegetation analysis, the investigator often subjectively estimates the abundance of individual species in the field. Most often, the data have been visual estimates of a species' cover while the alternative data collection strategy has been to record presence/absences of species in points or subplots (Kent & Coker 1992, Mueller-Dombois & Ellenberg 2002). Considering the pivotal importance of the data collected, relatively little interest has focused on the error in such data. There are several factors contributing to this error, and among the most important for trustworthiness is observer bias. Crucial for power analysis are estimates of the 'random' error, i.e. variation that cannot be explained or accounted for.

Methods based on presence/absence of species in plots describe a single, relatively uncontroversial aspect of the vegetation and therefore it is possible to evaluate its accuracy (i.e. how well a method describes the true, underlying pattern; Jonasson 1988). It is less clear what aspect of vegetation that a visual estimate of cover describes, and consequently what reference to use to evaluate its accuracy (e.g. biomass, photographs or visual estimates of many observers).

While the accuracy of cover estimates is potentially controversial, evaluating the precision in detection of species and of cover estimates is much more straightforward (i.e. the repeatability of records). Also, high precision is more important than accuracy when evaluating temporal trends in data. A number of studies have compared the scores of abundances made by two, or more, field workers when screening the same plots (e.g. Hope-Simpson 1940; Smith 1944; Sykes et al. 1983; Gotfryd & Hansell 1985; Kennedy & Addison 1987; Tonteri 1990; West & Hatton 1990; McCune et al. 1997; van Hees & Mead 2000; Scott & Hallam 2002; Brandon et al. 2003; Carlsson et al. 2005). The overall conclusion from these studies is that differences can be substantial and, therefore, that conclusions on changes over time might be suspect. If records are biased, an observed change over time might be due to systematic differences between fieldworkers (van Hees & Mead 2000). If, on the other hand, variation in field records is large or mainly random, this will result in lower statistical power than expected. Hence, it is important to know the extent of bias in data, and the size of the random error, when evaluating vegetation data and when designing monitoring programmes (Kery et al. 2006; Legg & Nagy 2006).

The present contribution aims to describe the error in vegetation data, and its random and observer bias components by using a large data set collected in mature boreal forest of Sweden. The protocol for the fieldwork, which was part of a national survey and monitoring scheme in Sweden, included scoring the cover, according to 15 classes, of 29 taxa (species or species groups). In addition, the presence/absence of selected taxa was noted (50 in our data set). We posed the following questions:

1. What is the magnitude of missed occurrences of taxa?
2. Can attributes of species explain variation in missed occurrences?
3. Do missed occurrences vary with abundance of taxa?
4. What is the magnitude of random error and observer bias components of cover estimates of taxa?

For (1) and (4), results were used to estimate minimal detectable differences for data in this monitoring programme.

## Material and Methods

### *The National Forest Inventory of Sweden*

From 1983 to 1987, the National Forest Inventory of Sweden established permanent plots with an aim to revisit them at 5-year intervals. Parallel to this, the plots were also subjected to a soil and vegetation inventory by the Swedish Survey of Forest Soils and Vegetation (SK). The Swedish University of Agricultural Sciences (SLU) executes both inventories and manages the databases. Currently, the monitoring scheme is called the Swedish Forest Soil Inventory (Anon. 2007).

For logistic reasons, the sample plots are located along the sides of a quadrat; one quadrat corresponds in central and northern Sweden to one day's work, and in southern Sweden to a half day's work by an inventory team. Quadrat size varies from 0.09 to 1.4 km$^2$ depending on location in Sweden. A quadrat has one sample plot at each corner of the quadrat and one in the middle of each side, i.e. eight in total (Lindroth 1995); the minimum distance between plots consequently varies from 600 m in northern Sweden to 200 m in the south (Hägglund 1985). Hence, there might be different degrees of spatial autocorrelation in data from the same and different quadrats. The current analyses are based on sample plot data ($N$ = 342; see details below) from 192 quadrats (84, 75, 27, 3 and 3 quadrats contributed 1, 2, 3, 4 and 5 sample plots, respectively), ignoring possible autocorrelation. The sample plots are circular, permanently marked with an aluminium pole, and vegetation is surveyed in a 100-m$^2$ plot. An inventory team consists of trained technicians, one with the task to sample soil and vegetation.

During the initial phase of SK, 1983-1987, separate inventory teams, that independently screened vegetation and some other variables, visited a selection of the plots. The purpose was to enable evaluation of the precision in the data collected, and we hereafter refer to it as the 'quality assessment inventory'. In total there were 1071 such plots (ca. 5% of all plots surveyed), but for many of them only partial data had been collected. Therefore, we used a subset consisting of 342 plots where both teams had made a full vegetation assessment (see inclusion criteria below). It is important to note that, at each sample plot, the quality assessment team had somewhat more time for vegetation analyses (perhaps 25% more, up to 30 min; Ola Löfgren pers. comm.). To reduce the cost of the quality assessment, quadrats were chosen that were easily accessed (from roads) while remote ones were generally ignored.

**Fig. 1.** Map of Sweden indicating the location of the 342 forest vegetation plots.

### Inclusion criteria

The field protocol first involved the decision whether a plot should be divided into subplots. As such a division might be different for different observers, no cases where the standard and/or quality assessment team had divided plots were included. Selecting only undivided plots for the analysis may exclude the more heterogeneous plots, e.g. vegetation in, or close to, edges in the landscape.

As to the vegetation inventory there was also a decision whether a section of the plot should be omitted from assessment. Reasons for partial omission were, in most cases, the area occupied by the standing stems of timber (normally < 0.4 m$^2$), boulders, rock outcrops and a few cases when the area had been affected by soil scarification etc. (e.g. Lundmark et al. 1985). Vegetation was then assessed on the remaining area. Naturally, observers differed somewhat in whether, and how much, they excluded. To avoid cases where the area under consideration was substantially different, we excluded 33 plots where the difference in area was >10 m$^2$. This limit is arbitrary, but the procedure reflects what users of data from this monitoring system are likely to do to reduce a possible bias in data.

Application of the above criteria resulted in a selection of 342 plots, which are spread over the entire country (Fig. 1). Data from these plots had been assembled by 36 observers in the standard inventories, and each observer had, on average, visited 9.5 plots (range 1-27; median 5.5). Only two persons were involved in the quality assessment inventory, visiting 212 and 130 plots, respectively. The median time period between the two inventories was 7 days (74% of the plots within 14 days); with a single exception, the standard inventory teams were always first on the plot. Most vegetation types involved are resilient to trampling, with only small phenological changes over the season.

### Variables in vegetation records

The objective of the SK was originally to improve the forest site classification system (Hägglund & Lundmark 1981), which is reflected in the pre-determined list of almost 100 species and groups of species subjected to the inventory (Anon. 1983-2002). Presence/absence data were collected but for some species cover was also estimated. In total, 79 species (Table 1) and groups of species (Table 2) had been recorded in the plots used in the current study. Of these, 50 were only recorded as presence/absence (Tables 1 and 2). In some cases, we also used two 'negative' variables, indicating the lack of vegetation cover by vascular plants, moss or lichen (Table 3).

Cover was visually estimated on a non-linear, 15-point scale: < 0.1, 0.1-1, 1-3, 3-6, 6-9, 9-12, 12-15, 15-20, 20-25, 25-30, 30-40, 40-50, 50-60, 60-80 and 80-100% cover. Throughout this paper, we used the class, or number of steps between classes, and consider it as an interval variable.

Some species were not included in the first year or years, leading to 260, 168, 174 or 82 observation plots for them instead of 342.

### Statistical analysis

#### Magnitude of missed occurrences

The purpose of the analysis of missed occurrences was to reveal taxa prone to under- or over-estimation. 'Missed occurrences' of a taxon was calculated as the percentage of plots where only one observer recorded it out of the total number of plots where at least one observer had noted it. As this ratio is sensitive to small numbers, it was only calculated for taxa where total number of observation > 25. This ratio does not account for cases where both observers missed a taxa that was actually present, nor where misidentification meant the scoring of a non-existing taxa. Therefore, we believe that our estimates of 'missed occurrences' are more likely to be 'underestimates' than 'overestimates'.

The number of missed occurrences of a taxon is strongly affected by its frequency in a data set. To remove this effect, we fitted a second degree polynomial (1) to data on taxa. Hence we assume that when frequency is approaching its minimum (total absence) or maximum (presence in all samples), the number of missed occurrences is low, while it would be largest at intermediate frequencies.

$$y = a * (x - 0.5)^2 + b \qquad (1)$$

where $y$ is the arcsine-transformed square root of the proportion of plots where a taxon was recorded by one observer, $x$ is the proportion of plots where a taxon was recorded by at least one observer, $a$ and $b$ are constants. The residuals of the 21 most abundant species (excluding groups of species) were used in one-way General Linear Model ANOVA using Statistica 7 software (Anon. 2004). The explanatory variables were life form (three different: geophytes, hemicryptophytes and woody chamaephytes; according to Ellenberg et al. 1992) and plant height (continuous; according to Lid 1985). In a preliminary analysis, we had shown residuals to be independent of $x$, i.e. the proportion of plots where a taxon was recorded by at least one observer (correlation: $R = 0.050$, $P = 0.830$).

When evaluating the change in presence/absence data over time, the McNemar test (McNemar 1947) might be used (Elzinga et al. 1998; Pollock 2006), as

**Table 1**. Presence of species (or groups of closely related species) as noted independently by two observers in sample plots (100-m$^2$). *N* = frequency; 'Both' is the number of plots where the species was recorded by both observers; 'One' is the number of plots where the species was recorded by only one of the observers. 'K-S test' reports the outcome of a Kolmogorov-Smirnov test comparing frequency distributions of plots where only one, and where two observers noted a species. Residuals are from function fitted in Fig. 3. Life form and heights are, in most cases, according to Ellenberg et al. (1992) and Lid (1985), respectively.

| | Acronym* | Sample plots | *N* | Both | One | Missed occurrences 'One'/N (%) | K-S test probability | Residual | Life form** | Height (cm) |
|---|---|---|---|---|---|---|---|---|---|---|
| **Cover records** | | | | | | | | | | |
| *Vaccinium myrtillus* | BLAB | 342 | 326 | 311 | 15 | 5 | < 0.001 | 0.029 | Z | 30 |
| *Vaccinium vitis-idaea* | LING | 342 | 298 | 265 | 33 | 11 | < 0.001 | 0.005 | Z | 10 |
| *Polytrichum commune* | VBM | 342 | 126 | 68 | 58 | 46 | < 0.001 | 0.135 | B | |
| *Calluna vulgaris & Erica tetralix* | LJUN | 342 | 123 | 94 | 29 | 24 | < 0.001 | – 0.051 | Z;Z | - |
| *Empetrum nigrum* | KRAK | 342 | 111 | 85 | 26 | 23 | < 0.001 | – 0.049 | Z | 10 |
| *Vaccinium uliginosum* | ODON | 342 | 74 | 59 | 15 | 20 | < 0.001 | – 0.056 | Z | 40 |
| *Epilobium angustifolium* | MJOO | 342 | 50 | 30 | 20 | 40 | NS | 0.066 | H | 100 |
| *Pteridium aquilinum* | ORNB | 342 | 45 | 38 | 7 | 16 | NS | – 0.061 | G | 100 |
| *Andromeda polifolia, Vaccinium oxycoccos & microcarpum* | ROTR | 342 | 29 | 17 | 12 | 41 | < 0.005 | 0.056 | Z;Z | |
| *Rhododendron tomentosum* | SKVA | 342 | 28 | 25 | 3 | 11 | NS | – 0.067 | Z | 60 |
| *Rubus chamaemorus* | HJOR | 342 | 27 | 23 | 4 | 15 | < 0.025 | – 0.045 | H | 15 |
| *Sphagnum fuscum* | RVM | 342 | 7 | 0 | 7 | | NA | 0.078 | B | |
| *Arctostaphylos uva-ursi* | MJOL | 342 | 4 | 1 | 3 | | NS | 0.024 | Z | 5 |
| **Presence/absence records** | | | | | | | | | | |
| *Maianthemum bifolium* | EKORRBA | 342 | 121 | 86 | 35 | 29 | | – 0.005 | G | 10 |
| *Linnaea borealis* | LINNEA | 260 | 95 | 70 | 25 | 26 | | – 0.027 | Z | 10 |
| *Oxalis acetosella* | HARSYRA | 342 | 79 | 63 | 16 | 20 | | – 0.059 | G/H | 10 |
| *Potentilla erecta* | BLODROT | 342 | 61 | 37 | 24 | 39 | | 0.067 | H | 15 |
| *Gymnocarpium dryopteris* | EKBRAKE | 342 | 58 | 48 | 10 | 17 | | – 0.064 | G | 20 |
| *Anemone nemorosa* | VITSIPP | 342 | 52 | 38 | 14 | 27 | | – 0.000 | G | 15 |
| *Rubus saxatilis* | STENBAR | 342 | 40 | 28 | 12 | 30 | | 0.017 | H | 20 |
| *Geranium sylvaticum* | SKOGSNA | 342 | 39 | 34 | 5 | 13 | | – 0.070 | H | 40 |
| *Trientalis europaea* | SKOGSST | 82 | 35 | 21 | 14 | 40 | | 0.092 | G | 15 |
| *Fragaria* spp. | SMULTRO | 342 | 31 | 18 | 13 | 42 | | 0.061 | H | 10 |
| *Rubus idaeus* | HALLON | 260 | 30 | 18 | 12 | 40 | | 0.060 | | 100 |
| *Solidago virgaurea* | GULLRIS | 82 | 21 | 13 | 8 | | | 0.064 | H | 40 |
| *Filipendula ulmaria* | ALGORT | 342 | 19 | 16 | 3 | | | – 0.034 | | |
| *Equisetum palustre* | KARRFRA | 342 | 16 | 9 | 7 | | | 0.042 | | |
| *Carex digitata* | VISPSTA | 260 | 16 | 4 | 12 | | | 0.130 | | |
| *Anemone hepatica* | BLASIPP | 342 | 14 | 10 | 4 | | | 0.003 | | |
| *Phegopteris connectilis* | HULTBRA | 342 | 14 | 6 | 8 | | | 0.064 | | |
| *Geum rivale* | HUMLEBL | 342 | 11 | 6 | 5 | | | 0.032 | | |
| *Convallaria majalis* | LILJEKO | 260 | 11 | 10 | 1 | | | – 0.055 | | |
| *Cornus suecica* | HONSBAR | 342 | 9 | 8 | 1 | | | – 0.043 | | |
| *Rumex acetosa* | ANGSSYR | 342 | 6 | 1 | 5 | | | 0.052 | | |
| *Cirsium helenioides* | BORSTTI | 342 | 6 | 3 | 3 | | | 0.016 | | |
| *Cicerbita alpina* | TORTA | 342 | 6 | 6 | 0 | | | – 0.090 | | |
| *Crepis paludosa* | KARRFIB | 342 | 5 | 1 | 4 | | | 0.039 | | |
| *Circium palustre* | KARRTIS | 342 | 5 | 3 | 2 | | | – 0.001 | | |
| *Mycelis muralis* | SKOGSSA | 342 | 5 | 5 | 0 | | | – 0.086 | | |
| *Angelica sylvestris* | STRATTA | 342 | 5 | 4 | 1 | | | – 0.028 | | |
| *Alchemilla* spp. | DAGGKAP | 342 | 3 | 1 | 2 | | | 0.007 | | |
| *Paris quadrifolia* | ORMBAR | 342 | 3 | 2 | 1 | | | – 0.019 | | |
| *Parnassia palustris* | SLATTER | 342 | 3 | 2 | 1 | | | – 0.019 | | |
| *Lathyrus vernus* | VARART | 342 | 3 | 0 | 3 | | | 0.028 | | |
| *Anemone ranunculoides* | GULSIPP | 342 | 2 | 0 | 2 | | | 0.011 | | |
| *Trollius europaeus* | SMORBOL | 342 | 2 | 1 | 1 | | | – 0.015 | | |
| *Selaginella selaginoides* | DVARGLU | 342 | 1 | 1 | 0 | | | – 0.069 | | |
| *Milium effusum* | HASSLEB | 260 | 1 | 0 | 1 | | | – 0.003 | | |
| *Aegopodium podagraria* | KIRSKAL | 342 | 1 | 1 | 0 | | | – 0.069 | | |
| *Platanthera bifolia & chlorantha* | NATTVIO | 260 | 1 | 0 | 1 | | | – 0.003 | | |
| *Aconitum lycoctonum* | NORDISK | 342 | 1 | 0 | 1 | | | – 0.011 | | |
| *Moneses uniflora* | OGONPYR | 260 | 1 | 0 | 1 | | | – 0.003 | | |
| *Daphne mezereum* | TIBAST | 260 | 1 | 1 | 0 | | | – 0.071 | | |

*Acronym used in data base and its documentation; **Life form: G = geophyte; H = hemicryptophyte; Z = woody chamaephyte; B = bryophyte.

it compares the proportions of presence and absence at two points in time. A null hypothesis would state that there is no difference in the proportions of losses ($P_{10}$) and gains ($P_{01}$) of a particular species. We calculated minimal detectable change in proportion ($P_{10} - P_{01}$) for a different number of sample plots, assuming $\alpha = 0.05$, $\beta = 0.1$ (i.e. power = 0.9), using the software StudySize 1.09 (CreoStat HB, Gothenburg, Sweden).

*Do missed occurrences vary with abundance of taxa?*

The question whether occurrences are more likely to be overlooked when they are sparse was addressed by calculating missed occurrences per abundance class (as recorded by the observer that found it). As a comparison, the total number of records per abundance class was also calculated, and a Kolmogorov-Smirnov test was conducted to test the null hypothesis that the two frequency distributions come from the same population.

Another question is whether species that in general are more abundant when present would be easier to detect than less abundant species. This was evaluated using linear regression with missed occurrences of species as a function of the average cover when present. The average cover was based on all records, i.e. duplicating the number of records from plots with two observers. There were 11 species in this analysis (excluding two rare species and groups of species).

*Error in cover estimates*

We calculated the average differences in the pairs of cover estimates between observers, and the SD of these differences, based on all plots where at least one observer had noted a taxon. This estimate of differences gives an indication of the magnitude of variation in data to be used to assess changes over time, i.e. incorporating both missed occurrences and variation in cover assessment. We also calculated SD based only on the plots where both observers had noted a taxon, hence estimating the magnitude of differences in visual estimates of cover ($SD_{cover}$).

An evaluation of temporal changes in data from the current monitoring scheme might involve the difference between two points in time (paired *t*-test). A null hypothesis would state that the average difference is zero. Assuming estimated Standard Deviation SD in this study can be used, we calculated minimal detectable difference ($\delta$) for $N$ using $\alpha = 0.05$, $\beta = 0.1$ (i.e. power 0.9), with the software StudySize 1.09 (CreoStat HB, Gothenburg, Sweden).

To formally compare the magnitude of observer bias in cover estimates, compared with random error, we extracted variance components for individual taxa. We considered the identity of the conventional observer as a random factor and the two quality assessment observers as fixed covariables, hence conducting a mixed model ANOVA (using Statistica 7 software; Anon. 2004). Analyses were conducted for 17 taxa that were relatively common First, data were only included from a conventional observer if he/she had recorded the taxon in at least five plots. Second, the taxon was included only if there were data from at least seven conventional observers. Corresponding analyses were also conducted for the variables ‘lack of mosses and lichens’ and ‘lack of vascular plants’.

## Results

The plots had on average 9.1 taxa (SD 3.69; range 1-26). The two quality assessment observers generally recorded somewhat lower cover of taxa (Table 3) but on average an additional 0.79 taxa per plot than the conventional observers.

*Magnitude of missed occurrences*

Missed occurrences (Tables 1 and 2) made up, on average, 26% (SD = 12.3) for species and slightly more for groups of species (34%, SD = 20.5). Missed occurrences ranged from 3% (‘mosses of mesic ground’; Table 2) to 74% (‘tall-grown sedges of wet ground’; Table 2). Using these estimates of missed occurrences, we calculated the minimal detectable difference within this monitoring system. For example, 11% of the occurrences of *Vaccinium vitis-idea* were missed (Table 1) which corresponds to ($P_{01} + P_{10}$) in Fig. 2. This suggests that the minimal detectable difference comparing two points in time and having 100 plots would be ± 9% (Fig. 2). A similar estimate for an often missed species such as *Epilobium angustifolium* (40%; Table 1) would be ± 21%.

The residuals from the function fitted to data in Fig. 3 indicated that *Polytrichum commune* and *Trientalis europaea* were likely to be overlooked (largest positive residuals in Table 1). In contrast, negative residuals were smaller and *Gymnocarpium dryopteris, Geranium sylvaticum, Oxalis acetosella* and *Vaccinium uliginosum* seemed easy to detect (Table 1).

There was no relationship between the residuals of missed occurrences and life-form ($F_{(2,17)} = 0.751$; $P = 0.487$) or height of the species ($F_{(1,17)} = 0.094$; $P = 0.763$).

**Table 2**. Presence of groups of species was noted independently by two observers in sample plots (100-m$^2$). $N$ = frequency; 'Both' is the number of plots where the taxon was recorded by both observers; 'One' is the number of plots where the taxon was recorded by only one of the observers. 'K-S test' reports the outcome of a Kolmogorov-Smirnov test comparing frequency distributions of plots where only one, and where two observers noted a species. Residuals were calculated from the function in Fig. 3.

| | Acronym* | Sample plots | $N$ | Both | One | K-S test probability | Missed occurrences : 'One'/$N$ (%) | Residual |
|---|---|---|---|---|---|---|---|---|
| **Cover records** | | | | | | | | |
| Mosses of mesic ground | FMM | 342 | 339 | 329 | 10 | <0.001 | 2.9 | 0.055 |
| Other herbs | OVOR | 342 | 279 | 239 | 40 | <0.001 | 14.3 | -0.013 |
| Narrow-leaved grasses | SMGR | 342 | 264 | 223 | 41 | <0.001 | 15.5 | -0.032 |
| Other species | OVAR | 342 | 182 | 115 | 67 | <0.001 | 36.8 | 0.079 |
| Low-grown sedges of non-moist ground | EFLH | 342 | 178 | 111 | 67 | NS | 37.6 | 0.084 |
| Other mosses of wet ground | OSM | 342 | 148 | 55 | 93 | <0.005 | 62.8 | 0.293 |
| Broad-leaved grasses | BRGR | 342 | 139 | 94 | 45 | <0.001 | 32.4 | 0.025 |
| *Cladina* spp., *Cladonia* spp. & *Stereocaulon* spp. | CPL | 342 | 137 | 104 | 33 | <0.001 | 24.1 | -0.048 |
| Low-grown herbs | LAGO | 342 | 133 | 115 | 18 | <0.001 | 13.5 | -0.148 |
| *Sphagnum* spp. (excl *S. fuscum*) | OVM | 342 | 131 | 108 | 23 | <0.001 | 17.6 | -0.108 |
| *Equisetum sylvaticum, Menyanthes trifoliata & Carex globularis* | SVAK | 342 | 85 | 62 | 23 | NS | 27.1 | -0.012 |
| *Lycopodiaceae* | LUMME | 342 | 82 | 57 | 25 | <0.001 | 30.4 | 0.013 |
| Tall-grown herbs | HOGO | 342 | 80 | 59 | 21 | <0.005 | 26.2 | -0.016 |
| Other lichens | OVL | 342 | 78 | 35 | 43 | NS | 55.1 | 0.168 |
| Low-grown sedges of wet ground | FFLH | 342 | 72 | 19 | 53 | NS | 73.6 | 0.265 |
| Tall-grown sedges of wet ground | FFHH | 342 | 43 | 11 | 32 | <0.001 | 74.4 | 0.202 |
| **Presence/absence records** | | | | | | | | |
| Tall-grown or low-grown herbs | HOLAF | 342 | 135 | 122 | 13 | | 9.6 | -0.193 |
| *Viola* spp. | VIOL | 342 | 56 | 39 | 17 | | 30.4 | 0.017 |
| *Melampyrum* spp. | KOVA | 82 | 53 | 34 | 19 | | 35.8 | 0.094 |
| 'False' tall-grown ferns | FHORM | 168 | 29 | 16 | 13 | | 44.8 | 0.095 |
| *Veronica* spp. | VERO | 342 | 28 | 11 | 17 | | 60.7 | 0.116 |
| Fam. *Orchidaceae* | ORKI | 342 | 22 | 5 | 17 | | | 0.139 |
| *Ranunculus* spp.** | SMORBLO | 342 | 21 | 10 | 11 | | | 0.075 |
| *Hieracium* spp. | FIBB | 82 | 14 | 4 | 10 | | | 0.226 |
| 'Genuine' tall-grown ferns | AHORM | 174 | 10 | 4 | 6 | | | 0.090 |
| *Galium* spp. | MARA | 82 | 10 | 5 | 5 | | | 0.103 |

*Acronym used in data base and its documentation; **Only 10-50 cm tall, yellow-flowering species.

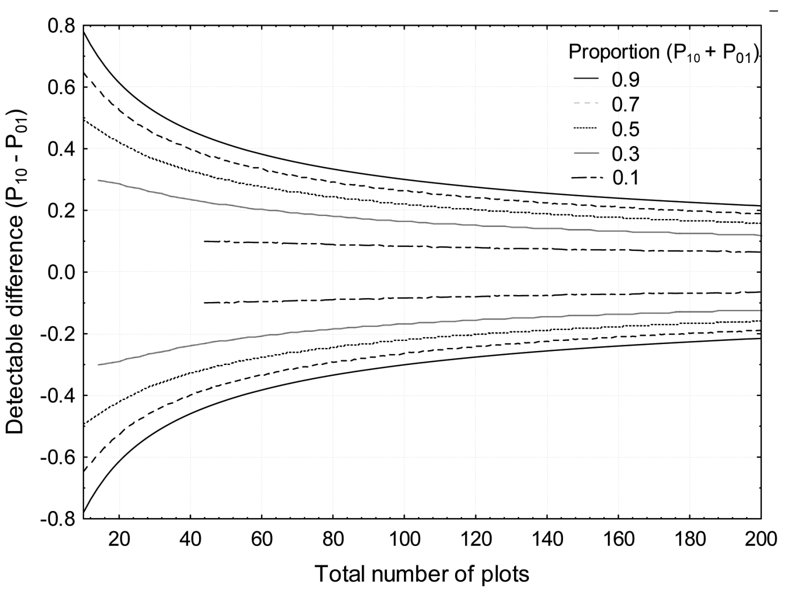

**Fig. 2.** Minimal detectable difference in proportions $P_{10}$ and $P_{01}$ in two-sided McNemar test as a function of sample size and the proportion ($P_{10}$+$P_{01}$), using $\alpha = 0.05$, Power ($1-\beta$) = 0.9, and $H_0$: mean = 0. For an evaluation of temporal trends, $P_{10}$ may represent the proportion of plots where a species was recorded only on the first occasion, $P_{01}$ proportion of plots where a species was recorded only on the second occasion, while ($P_{10}$+$P_{01}$) would be calculated from data given in Table 1 or 2 ('Missed occurrences').

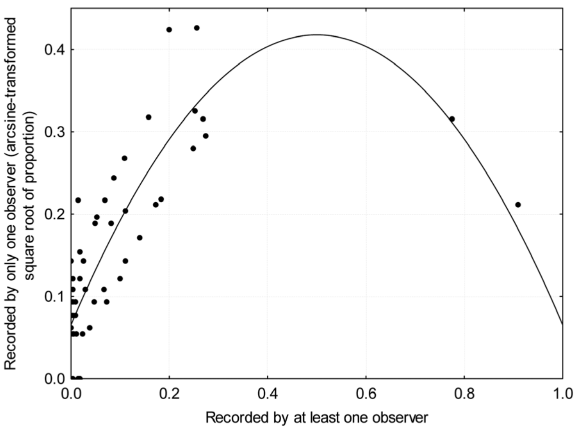

**Fig. 3.** Relationship between the proportion of cases where a taxon was recorded by both observers (abscissa) and the proportion (arcsine-transformed square root) of cases where the taxon was recorded by only one of the two observers. There were 342 plots in total. $y = -1.41*(x - 0.5)^2 + 0.418$.

### *Do missed occurrences vary with abundance of taxa?*

Cases where a taxon was recorded by only one observer were generally over-represented at low abundances: 78% of those cases were in cover class 1 and 2, while the corresponding value was 45% for cases where both observers had recorded a taxon in cover class 1 and 2. For 20 taxa, the Kolmogorov-Smirnov tests indicated a significant difference in frequency distribution of records noted by one observer and records noted by both observers (Tables 1 and 2). Exceptions were a few rare species (for which this test has low power), and *Epilobium angustifolium* (Fig. 4), *Pteridium aquilinum, Rhododendron tomentosum* (Table 1), and the species groups 'low-grown sedges of non-moist ground', '*Equisetum sylvaticum, Menyanthes trifoliata* & *Carex globularis*', 'other lichens' and 'low-grown sedges of wet ground' (Table 2). Six examples of frequency distributions are shown in Fig. 4.

When comparing species, those that were abundant when present were more seldom missed (Fig. 5), but the relatively large spread of data points and the outlying *Polytrichum commune* indicate a complex relationship.

### *Error in cover estimates*

Average differences in estimated cover class were generally relatively close to zero (Table 3). Of special interest were the estimates of SD (Table 3), that can be used in power analyses (Fig. 6).

$SD_{cover}$, i.e. considering only plots where both observers had scored a taxon thereby highlighting the variation in perception of cover, was generally smaller than SD (Table 3) although there were also cases where $SD_{cover}$, was marginally larger than SD (*Vaccinium uliginosum*, and the groups '*Cladina* spp., *Cladonia* spp. & *Stereocaulon* spp.' and 'other herbs').

In total, five of the 17 taxa tested scored a significant observer effect in the variance component analyses: *Vaccinium myrtillus*, 'other herbs', 'mosses of mesic ground', '*Cladina* spp., *Cladonia* spp. & *Stereocaulon* spp.' and 'other mosses of wet ground' (Table 3). The amounts of variation explained by observer was < 5%, 5-10% and >10% in nine, six and two cases, respectively (Table 3). The two 'negative' variables, recording the

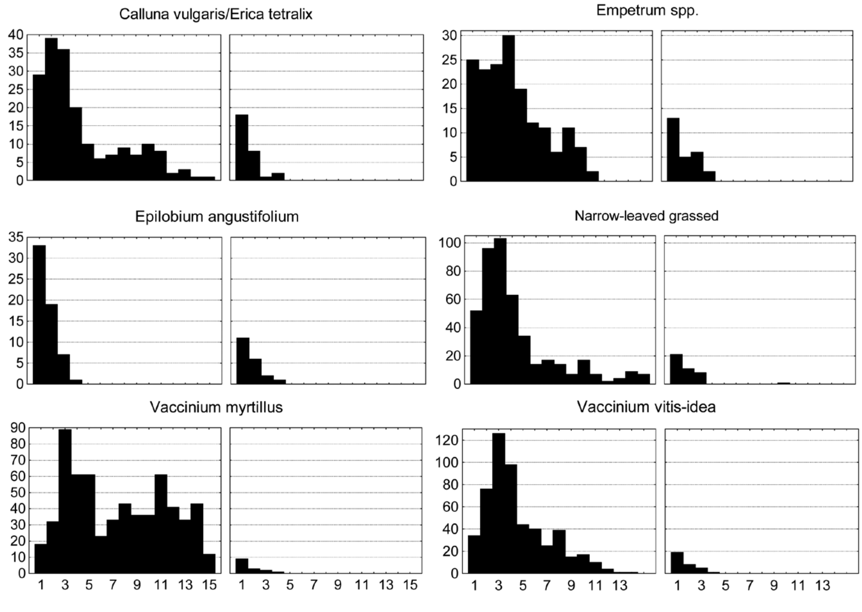

**Fig. 4.** Number of cases per abundance class (1-15) for selected taxa. Right graphs show cases where only one of the observers had noted the taxon's presence, while left graphs show data on cases where both observers had noted the taxon (based on all data, i.e. 2*$N$). Kolmogorov-Smirnov tests (Table 1) indicated that the distribution of the two data types were statistically significant for all displayed examples except *Epilobium angustifolium.*

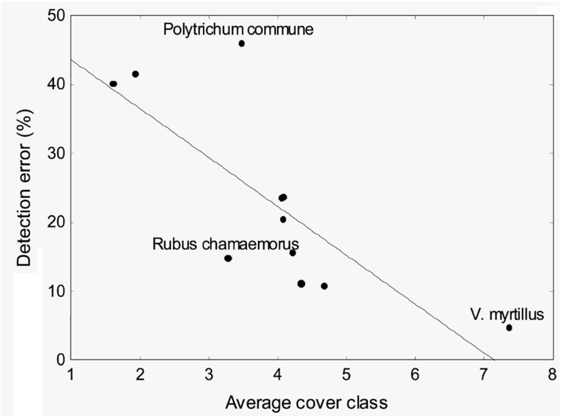

**Fig. 5.** Missed occurrences of a species as a function of average cover-class when present. The relationship ($y = 50.7 - 7.095*x$) was significant (linear regression: $F_{(1,9)} = 13.2$, $P = 0.0054$).

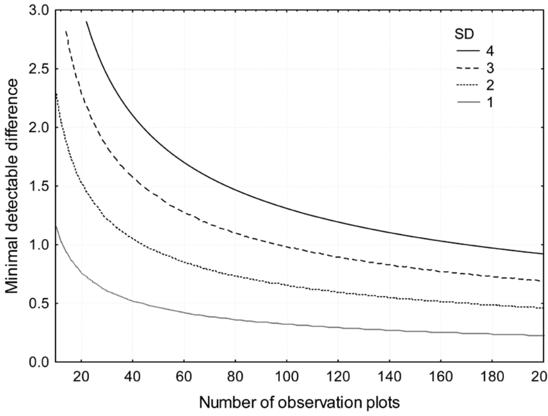

**Fig. 6.** Minimal detectable difference in two-sided paired *t*-test as a function of sample size and SD, using $\alpha = 0.05$, Power $(1 - \beta) = 0.9$, and $H_0$: mean = 0.

'lack of cover of mosses and lichens' and 'lack of cover of vascular plants', had both highly significant observer effects; the latter had the highest observer explained variance recorded in this study (16%).

Ignoring possible bias and assuming that the error is purely random, we can use the data in Table 3 to estimate the minimal detectable difference within this monitoring system. For example, *Vaccinium vitis-idaea* had SD 1.92 (Table 3). This suggests that the minimal detectable difference, when having 100 plots where *V. vitis-idaea* was present on at least one of the two points in time, would be 0.63 cover units (Fig. 6).

## Discussion

This study showed that the magnitude of observer bias in plant cover estimates is most often <10% of the total variation when accounting for between-plot differences (Table 3). This study also showed that error, both random error and observer bias, is species and cover dependent (Tables 1-3, Figs. 3-5). It is worth pointing out that the large plot size (100 m$^2$), the species-poor vegetation (average of 9 taxa per plot), the use of species groups (Table 2), and the use of 15 cover classes in our study, might limit the transferability of the results. For example, there are probably more missed occurrences with increasing plot size and with increasing species richness.

### *Error in detection of taxa and in cover estimates*

Most estimates of error in vegetation data previously reported are the sum of random error and observer bias and are directly comparable to the presented estimates of missed occurrences (Table 1 and 2), SD and SD$_{\text{cover}}$ (Table 3).

### *Presence/absence data*

In our study, on average, 26 and 34% of the occurrences were missed for species and species groups, respectively, both with considerable variation (Tables 1 and 2). Both accidentally missing a species and misidentification could have contributed to our variable 'missed occurrences'. We believe, however, that the influence of misidentification is low as the list of taxa was determined beforehand based on, among other things, the ease of identification. Part of the variation in missed occurrences among taxa depended on the total frequency of taxa (Fig. 3, error in absolute number presumably largest when 50%), a variable that is at least partly scale (plot size) dependent. Another explanatory variable for missed occurrences is the abundance of a species when present (Fig. 4): scant occurrences are more likely missed (see also Lepš & Hadincová 1992; Kercher et al. 2003; Vittoz & Guisan 2007). Furthermore, the overall average abundance of a species could partly explain the missed occurrences (Fig. 5) with ubiquitous species less likely to be missed. The two latter explanatory variables might give some guidance regarding data quality and training priorities.

There is some indication that missed occurrences increase with plot size. Ringvall et al. (2005), using numerous observers, partly from the same monitoring system as we used, reported average agreements between surveyors of 87 and 89% in forest vegetation for plot size of 0.01 and 0.33 m$^2$, respectively. This is better than in our study (Tables 1 and 2) and that of Archaux et al. (2006), both using large plots in forest vegetation. Large plots probably give more room for manifestation of search and detection skills among observers.

The model used to describe missed occurrences as a function of frequency (Eq. 1) might not be quite realistic

**Table 3**. Species and groups of species whose cover was estimated independently by two observers in 342 sample plots (100-m$^2$). 'Average difference' (Quality Assessment observer minus Conventional Observer) and SD is based on the taxon's abundance (one of 15 cover classes) when recorded by at least one observer. In contrast, $SD_{cover}$ is based only on plots where both observers had noted a taxon, hence showing variation in visual perception of cover. $N_{CO}$ and $N_{plots}$ is number of conventional observers and plots, respectively, used when calculating 'Observer-explained variance' (and corresponding $P$-value) in a variance component analysis.

| | Acronym* | $N$ | Average difference in cover | SD | $SD_{cover}$ | $N_{CO}$; $N_{plots}$ | Observer-explained variance (%) | $P$ |
|---|---|---|---|---|---|---|---|---|
| **Species** | | | | | | | | |
| *Vaccinium myrtillus* | BLAB | 326 | -0.061 | 1.96 | 1.96 | 20; 284 | 7.3 | 0.0058 |
| *Vaccinium vitis-idaea* | LING | 298 | 0.473 | 1.92 | 1.91 | 19; 256 | 0.0 | 0.726 |
| *Polytrichum commune* | VBM | 126 | -0.421 | 2.03 | 1.87 | 10; 91 | 7.5 | 0.099 |
| *Calluna vulgaris & Erica tretralix* | LJUN | 123 | -0.081 | 1.45 | 1.33 | 11; 87 | 2.1 | 0.330 |
| *Empetrum nigrum* | KRAK | 111 | 0.279 | 1.95 | 1.82 | 7; 79 | 1.5 | 0.337 |
| *Vaccinium uliginosum* | ODON | 74 | 0.324 | 2.01 | 2.09 | | | |
| *Epilobium angustifolium* | MJOO | 50 | 0.180 | 1.30 | 0.73 | | | |
| *Pteridium aquilinum* | ORNB | 45 | -0.044 | 2.41 | 1.75 | | | |
| *Andromeda polifolia, Vaccinium oxycoccos & microcarpum* | ROTR | 29 | -0.310 | 1.14 | 1.12 | | | |
| *Rhododendron tomentosum* | SKVA | 28 | -0.071 | 1.54 | 1.44 | | | |
| *Rubus chamaemorus* | HJOR | 27 | 0.296 | 1.35 | 1.34 | | | |
| *Sphagnum fuscum* | RVM | 7 | 0.714 | 7.27 | | | | |
| *Arctostaphylos uva-ursi* | MJOL | 4 | -0.500 | 2.38 | | | | |
| **Groups of species** | | | | | | | | |
| Mosses of mesic ground | FMM | 339 | -0.614 | 2.71 | 2.67 | 21; 299 | 7.5 | 0.0040 |
| Other herbs | OVOR | 279 | -0.025 | 1.63 | 1.65 | 15; 221 | 6.2 | 0.023 |
| Narrow-leaved grasses | SMGR | 264 | -0.129 | 1.80 | 1.66 | 20; 232 | 1.3 | 0.310 |
| Other species | OVAR | 182 | -0.511 | 1.38 | 1.19 | 14; 140 | 5.2 | 0.116 |
| Low-grown sedges of non-moist ground | EFLH | 178 | -0.247 | 1.28 | 0.84 | 12; 128 | 1.0 | 0.362 |
| Other mosses of wet ground | OSM | 148 | -0.520 | 2.72 | 1.88 | 14; 118 | 12.5 | 0.016 |
| Broad-leaved grasses | BRGR | 139 | 0.014 | 2.05 | 1.79 | 11; 92 | 5.0 | 0.187 |
| *Cladina* spp., *Cladonia* spp. and *Stereocaulon* spp. | CPL | 137 | 0.073 | 1.69 | 1.73 | 11; 105 | 15.8 | 0.0055 |
| Low-grown herbs | LAGO | 133 | 0.203 | 1.88 | 1.81 | 10; 85 | 0.0 | 0.953 |
| *Sphagnum* spp. (excl. *S. fuscum*) | OVM | 131 | -0.557 | 3.15 | 2.78 | 10; 85 | 0.0 | 0.826 |
| *Equisetum sylvaticum, Menyanthes trifoliata & Carex globularis* | SVAK | 85 | -0.094 | 2.57 | 1.90 | 9; 66 | 3.9 | 0.265 |
| *Lycopodiaceae* | LUMME | 82 | -0.183 | 0.89 | 0.80 | | | |
| Tall-grown herbs | HOGO | 80 | 0.162 | 1.63 | 1.46 | | | |
| Other lichens | OVL | 78 | -0.256 | 1.62 | 0.89 | 8; 57 | 0.0 | 0.510 |
| Low-grown sedges of wet ground | FFLH | 72 | -0.167 | 3.41 | 2.08 | | | |
| Tall-grown sedges of wet ground | FFHH | 43 | -0.884 | 3.31 | 2.90 | | | |
| **Lack of vegetation** | | | | | | | | |
| Area not covered with mosses or lichens | BSA | 342 | 0.439 | 3.04 | - | 21; 302 | 9.3 | 0.00078 |
| Area not covered with vascular plants | FSAK | 342 | -0.868 | 2.92 | - | 21; 302 | 16.0 | 0.000001 |

*Acronym used in data base and its documentation.

and a skewed curve may better describe the data. For example, we would expect the risk of not detecting a rarely encountered species to be larger than missing (to note) the presence of a ubiquitous one (Fig. 5). Nevertheless, we expect the general conclusions on magnitude of missed occurrence to hold, namely that there are species that are harder to detect than equally abundant species (e.g. *Polytrichum commune*, residual 0.135, and *Empetrum nigrum*/*hermaphroditum*, – 0.049; Table 1; cf. Brandon et al. 2003; van Hees & Mead 2000). This illustrates that species and cover interact, not through the simple species' attributes tested here (height and life form), but through some other observer bias component (Fig. 5). Taxa with large residuals in this analysis could be targets for improving precision (see below).

### *Cover data: random error and observer bias*

The calculated SD of cover estimates seemed relatively low (Table 3), but as the data are from a cover class scale, they are difficult to compare with published reports. Sykes et al. (1983) reported SD values between 7 and 22% (seven taxa, three plot sizes) of cover estimates (20 cover classes) in British woodlands.

There are few previous estimates of observer bias in vegetation studies (van Hees & Mead 2000; Nagy et al. 2002; Ringvall et al. 2005). Hence, the relatively small portion of total variation (ca. 10%) that comes from systematic differences between observers when assessing cover of taxa is noteworthy (Tables 3 and 4). Nevertheless, some previous reports have failed to detect substantial observer bias: Bråkenhielm & Liu (1995) concluded, from forest vegetation monitoring plots, that inter- and intra-observer error was nearly the same and Lepš & Hadincová (1992) failed to detect significant observer effects in cover estimates (7 cover classes). In contrast, Sykes et al. (1983) reported consistent inter-observer differences of cover estimates (20 cover classes).

Our results should generally strengthen the reliability of conclusions drawn from vegetation data comprised of cover estimates, assuming that we consider observer bias < 10% to be acceptable.

The two quality assessors scored more taxa but recorded lower cover than the conventional observers; this may reflect both the greater skills and the slightly longer times the former could spend per plot. Overall, however, differences in cover estimates were close to 0 (Table 3).

### *Remedies for improved precision*

There are a number of reasons for why there is such substantial disagreement when doing repeated surveys of the same plot. 1. Re-locating the plot in the field and edge-effect – small differences in placement of a point or frame will mean that slightly different areas are covered contributing to random error. 2. Skills in finding what is in the plot vary, contributing to possible bias. 3. There is also a random component, simply suggesting that even the most skilled field worker, working without time constraints, still will miss some occurrences (Ringvall et al. 2005; Archaux et al. 2006) and exhibit some variation in cover estimates. 4. There might also be disagreement about the identity of species detected (Gray & Azuma 2005; in our study, species difficult to identify had generally been excluded from the short list, but *Polytrichum commune* and *Sphagnum fuscum* were notable exceptions). 5. Identification skills might also add bias in data. In our study as well as that of Gray & Azuma (2005), the field workers were left with some room for determining the taxonomic level given (e.g. genus or species level).

Training, and/or experience, is likely to increase the number of species detected (e.g. Table 3; Stapanian et al. 1997) while calibration of cover estimates of target species can improve both the accuracy and precision (Gallegos 2005).

A commendable strategy is to use several observers to generate the data and thereby increase accuracy by reducing variance and influence of bias in the pooled data (Kirby et al. 1986; Nilsson 1992; Klimeš 2003; Bergfur et al. 2004; Vittoz & Guisan 2007). However, this is not common practice, nor is it likely to become so because of the high cost for field work.

Our analyses point to the vital importance of the instructions in and internal consistency of the field protocol. In the current data set, differences between the paired samples have been exaggerated as an observer might, depending on skills and confidence, allocate a moss to either *Sphagnum fuscum* or *Sphagnum* spp. (in total more than 50 species in Sweden) and either *Polytrichum commune* or 'other mosses' (including the other nine species of *Polytrichum* occurring in Sweden). The large residuals of *Polytrichum commune* and *Sphagnum fuscum* (Table 1) must be seen in this light. So, it seems best, from the perspective of the usefulness of the data, to let the taxonomic ambition be determined at the desk and not in the field.

Data in monitoring could probably be improved by a more structured working model during field work (e.g. clearer guidelines for how to search plots, how much time to spend, and marking not only the presence but also absence of species). When comparing species, those that were abundant when present had less missed occurrences (Fig. 5), which also gives some guidance when targeting species for training.

As scant occurrences were more likely to be missed (Fig. 4), it is tempting to play with the idea of ignoring such occurrences. Hence, could a minimum level of abundance for inclusion improve data quality, by eliminating lots of false zeros? Or would the new detection limit exaggerate differences, which are potentially false, when data show a transition from below to above the threshold? As a large part of the missed occurrences are infrequent species, a simpler version of slimming the data, preferentially excluding poor-quality parts, might be to eliminate data on some species.

The monitoring scheme from which the current data emerged has, over time, launched a number of improvements (Ståhl & Odell 1998). Most notably, species groups are now more well-defined and the cover classes have been abandoned in favour of recording an 'exact' percentage cover. Especially the latter has simplified analyses. On the other hand, any changes in a running monitoring program create problems for evaluation of temporal trends (Ståhl & Odell 1998).

Two of the 'worst' variables, from the point of view of observer bias, were 'area not covered with mosses or lichens' and 'area not covered with vascular plants' (Table 3). It would be prudent to reconsider their future in the system.

*The power of monitoring: precision vs. sample size*

Ignoring possible observer bias, we can use information in Tables 1-3 to estimate an appropriate minimal detectable difference within this monitoring system (Figs. 2 and 6). Such information is crucial for the overall trustworthiness of a monitoring scheme, as well as for any study of temporal, or spatial, trends based on its data (Strayer 1999, Maxwell & Jennings 2005, Legg & Nagy 2006). A challenge, however, is how to implement variation estimates (Table 3) and calculated power analysis (Figs. 2 and 6) into evaluations of data from a monitoring scheme. Another challenge is to set up reasonable data and measurement quality objectives, weighing costs against data quality (e.g. Stribling et al. 2003) for presence/absence and cover data.

**Acknowledgements.** We are deeply grateful to all those persons involved in the vegetation inventory. Financial support was partly provided by the Swedish Environmental Protection Agency (PM, LW) and Stiftelsen Oscar & Lili Lamms Minne (JB). We appreciate comments from Anders Glimskär, Ola Löfgren, Mike Palmer, Göran Ståhl and referees.

## References

Anon. 1983-2002. *Ståndortskarteringens databas, SKBAS*. http://www-sml.slu.se/sk/skbas.pdf (accessed October 2007).

Anon. 2004. *Statistica (data analysis software system), version 7*. Statsoft Inc., Tulsa, OK, US.

Anon. 2007. *Markinventeringen, RIS-MI*. Swedish Forest Soil Inventory, http://www-markinventeringen.slu.se. (accessed October 2007).

Archaux, F., Gosselin, F., Berges, L. & Chevalier, R. 2006. Effects of sampling time, species richness and observer on the exhaustiveness of plant censuses. *Journal of Vegetation Science* 17: 299-306.

Bergfur, J., Carlsson, A.L.M. & Milberg, P. 2004. Phenological changes within a growth season in two semi-natural pastures in southern Sweden. *Annales Botanici Fennici* 41: 15-25.

Bråkenhielm, S. & Liu, Q. 1995. Comparison of field methods in vegetation monitoring. *Water, Air and Soil Pollution* 79: 75-87.

Brandon, A., Spyreas, G., Molano-Flores, B., Carroll, C. & Ellis, J. 2003. Can volunteers provide reliable data for forest vegetation surveys? *Natural Areas Journal* 23: 254-261.

Carlsson, A.L.M., Bergfur, J. & Milberg, P. 2005. Comparison of data from two vegetation monitoring methods in semi-natural grasslands. *Environmental Monitoring and Assessment* 100: 235-248.

Ellenberg, H., Weber, H.E., Düll, R., Wirth, V., Werner, W. & Paulißen, D. 1992. *Zeigerwerte von Pflanzen in Mitteleuropa*. 2nd. ed. *Scripta Geobotanica* 18: 1-258.

Elzinga, C.L., Valzer, D.W. & Willoughby, J.W. 1998. *Measuring and monitoring plant populations*. Bureau of Land Management, Denver, Colorado, USA. BLM/RS/ST-98/005+1730. http://www.blm.gov/nstc/library/pdf/MeasAndMon.pdf (accessed October 2007).

Gallegos, Å. 2005. *Design and evaluation of a computer aided calibration program for visual estimation of vegetation cover*. SLU, Institutionen för skoglig resurshushållning och geomatik, Arbetsrapport 142. www.resgeom.slu.se/resana/NILS/Publikationer/arb_rapp_142.pdf (accessed October 2007).

Gotfryd, A. & Hansell, R.I.C. 1985. The impact of observer bias on multivariate analyses of vegetation structure. *Oikos* 45: 223-234.

Gray, A.N. & Azuma, D.L. 2005. Repeatability and implementation of a forest vegetation indicator. *Ecological Indicators* 5: 57-71.

Hägglund, B. 1985. *En ny svensk riksskogstaxering*. Report 37, Department of Forest Survey, Swedish University of Agricultural Sciences, Umeå, SE. (In Swedish with English summary.)

Hägglund, B. & Lundmark, J.-E. 1981. *Handledning i bonitering med skogshögskolans boniteringssystem*. 3 Vols. Skogsstyrelsen, Jönköping, SE. (In Swedish.)

Hope-Simpson, J.F. 1940. On the errors in the ordinary use of subjective frequency estimations in grassland. *Journal of Ecology* 28: 193-209.

Jonasson, S. 1988. Evaluation of the point intercept method for the estimation of plant biomass. *Oikos* 52: 101-106.

Karlsson, T. 1998. Förteckning över svenska kärlväxter. *Svensk Botanisk Tidskrift* 91: 241-560 (In Swedish with English summary.)

Kennedy, K.A. & Addison, P.A. 1987. Some considerations for the use of visual estimates of plant cover in biomonitoring. *Journal of Ecology* 75: 151-157.

Kent, M. & Coker, P. 1992. *Vegetation description and analysis*. A practical approach. John Wiley and Sons, Chichester, UK.

Kercher, S.M., Frieswyk, C.B. & Zedler, J.B. 2003. Effects of sampling teams and estimation methods on the assessment of plant cover. *Journal of Vegetation Science* 14: 899-906.

Kéry, M., Spillmann, J.H., Truong, C. & Holderegger, R. 2006. How biased are estimates of extinction probability in revisitation studies? *Journal of Ecology* 94: 980-986.

Kirby, K.J., Bines, T., Burn, A., Mackintosh, J., Pitkin, P. & Smith, I. 1986. Seasonal and observer differences in vascular plant records from British woodlands. *Journal of Ecology* 74: 123-131.

Klimeš, L. 2003. Scale-dependent variation in visual estimates of grassland plant cover. *Journal of Vegetation Science* 14: 815-821.

Legg, C.J. & Nagy, L 2006. Why most conservation monitoring is, but need not be, a waste of time. *Journal of Environmental Management* 78: 194-199.

Lepš, J. & Hadincová, V. 1992. How reliable are our vegetation analyses? *Journal of Vegetation Science* 3: 119-124.

Lid, J. 1985. *Norsk, svensk, finsk flora*. Det Norske Samlaget, Oslo, NO.

Lindroth, S. (ed.) 1995. *Skog och mark i Sverige. Fakta från Riksskogstaxeringen*. Sveriges lantbruksuniversitet, Kommentus förlag, Stockholm, SE. (In Swedish.)

Lundmark, J.-E., Odell, G. & Carlsson, E. 1985. *Fältinstruktion för ståndortskarteringen av permanenta provytor vid riksskogstaxeringen*. Sveriges Lantbruksuniversitet, Institution för Skoglig Marklära, Uppsala, SE.

Maxwell, D. & Jennings, S. 2005. Power of monitoring programmes to detect decline and recovery of rare and vulnerable fish. *Journal of Applied Ecology* 42: 25-37.

McCune, B., Dey, J.P., Peck, J.E., Cassel, D., Heiman, K., Will-Wolf, S. & Neitlich, P.N. 1997. Repeatability of community data: species richness versus gradient scores in large-scale lichen studies. *Bryologist* 100: 40-46.

McNemar, Q. 1947. Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika* 12: 153-157.

Mueller-Dombois, D. & Ellenberg, H. 2002. *Aims and methods of vegetation ecology*. Blackburn Press, Caldwell, NJ, US.

Nagy, L., Nagy, J., Legg, C.J., Sales, D.I. & Horsefield, D. 2002. Monotoring vegetation change caused by trampling: a study in the Cairngorms, Scotland. *Botanical Journal of Scotland* 54: 191-207.

Nilsson, C. 1992. Increasing the reliability of vegetation analyses by using a team of two investigators. *Journal of Vegetation Science* 3: 565.

Pollock, J.F. 2006. Detecting population declines over large areas with presence-absence, time-to-encounter, and count survey methods. *Conservation Biology* 20: 882–892.

Ringvall, A., Petersson, H., Ståhl, G. & Lämås, T. 2005. Surveyor consistency in presence/absence sampling for monitoring vegetation in a boreal forest. *Forest Ecology and Management* 212: 109-117.

Scott, W.A. & Hallam, C.J. 2002. Assessing species misidentification rates through quality assurance of vegetation monitoring. *Plant Ecology* 165: 101-115.

Smith, A.D. 1944. A study of the reliability of range vegetation estimates. *Ecology* 25: 441-448.

Ståhl, G. & Odell, G. 1998. *Vegetationsförändringar i svensk skogsmark mellan 1980- och 90-talet: en studie grundad på ståndortskarteringen*. SLU, Institutionen för skoglig resurshushållning och geomatik, Umeå, Arbetsrapport 37. http://www-sml.slu.se/sk/vegfor98.pdf (accessed October 2007).

Stapanian, M.A., Cline, S.P. & Cassell, D.L. 1997. Evaluation of a measurement method for forest vegetation in a large-scale ecological survey. *Environmental Monitoring and Assessment* 45: 237-257.

Strayer, D.L. 1999. Statistical power of presences-absence data to detect population declines. *Conservation Biology* 13: 1034-1038.

Stribling, J.B., Moulton II, S.R. & Lester, G.T. 2003. Determining the quality of taxonomic data. *Journal of the North American Benthological Society* 22: 621-631.

Sykes, J.M., Horrill, A.D. & Mountford, M.D. 1983. Use of visual cover assessments as quantitative estimators of some British woodland taxa. *Journal of Ecology* 71: 437-450.

Tonteri, T. 1990. Inter-observer variation in forest vegetation cover assessments. *Silva Fennica* 24: 189-196.

van Hees, W.W.S. & Mead, B.R. 2000. Ocular estimates of understory vegetation structure in a closed *Picea glauca*/*Betula papyrifera* forest. *Journal of Vegetation Science* 11: 195-200.

Vittoz, P. & Guisan, A. 2007. How reliable is the monitoring of permanent vegetation plots? A test with multiple observers *Journal of Vegetation Science* 18: 413-422.

West, N.E. & Hatton, T.J. 1990. Relative influence of observer error and plot randomization on detection of vegetation change. *Coenoses* 5: 45-49.

Received 17 November 2006;
Accepted 16 October 2007;
Co-ordinating Editor: A. Chiarucci.