Consistency of defoliation data of the national training courses for the forest condition survey in Germany from 1992 to 2012

Nadine Eickenscheidt · Nicole Wellbrock

Received: 10 January 2013 / Accepted: 29 July 2013 / Published online: 18 August 2013
© The Author(s) 2013. This article is published with open access at Springerlink.com

Abstract The consistency of visual assessment of tree defoliation, which represents the most widely used indicator for tree condition, has frequently been in the focus of scientific criticism. Thus, the objective of the present study was to examine the consistency of the defoliation data from the annual national training courses for the forest condition survey in Germany from 1992 to 2012. Defoliation assessments were carried out in stands of beech (Fagus sylvatica), oak (Quercus robur and Quercus petraea), Norway spruce (Picea abies), and pine (Pinus sylvestris). Among the observer teams, the absolute deviation from the observer mean of all years was ±4.4 % defoliation and the standard deviation of defoliation was ±5.5 %. On average, 94 % of the assessments were located within the ±10 % interval of deviation from the mean. Tree species-specific differences did not occur when all years were considered. A trend towards increasing consistency was observed from 1992 to 2012, in particular for oak and spruce. The deviation of defoliation assessments depended non-linearly on the level of defoliation with highest deviations at intermediate defoliations. In spite of high correlations and agreements among observers, systematic errors were determined in nearly every year. However, within-observer variances were higher than between-observer variances. The present study applied a three-way evaluation approach for the assessment of consistency and demonstrated that the visual defoliation assessment at the national training courses in general produced consistent data within Germany from 1992 to 2012.

Keywords Forest condition survey · Defoliation · Observer error · Harmonisation · Systematic error · Quality assurance

Introduction

In Europe, the issue of forest decline emerged as the major environmental concern of the 1980s (e.g. Innes 1993). As a consequence, the ‘International Co-operative Programme on Assessment and Monitoring of Air Pollution Effects on Forests’ (ICP Forests) was established in 1985 by the Economic Commission for Europe of the United Nations under the Convention on Long-Range Transboundary Air Pollution (Innes 1993). Within the framework of the ICP Forests Programme, efforts were made to widely harmonise and standardise methods for forest monitoring throughout Europe. The methods were recorded in the ICP Forests manual (UNECE 2010) that was first published in 1985 and has continuously been subject to updates since its publication.

The forest condition survey represents an essential part of the forest monitoring, which was described in Part IV of the ICP Forests manual (Eichhorn et al. 2010) and became mandatory throughout the European Union in 1987 (Redfern and Boswell 2004; Solberg and Strand 1999). The survey on forest condition has been conducted annually on the systematic wide-scale monitoring plots (Level I), which were established wherever
forest coincided with a 16 × 16 km grid over Europe, as well as on plots of the intensive monitoring programme (Level II; since the 1990s) (Eichhorn et al. 2010; Ferretti et al. 1999; Innes 1993). The forest condition survey according to ICP Forests, has been taken place annually in Western Germany since 1984 and in the whole Federal Republic of Germany since 1990 (BMELV 2012). The 16 German federal states (including three city states) are responsible for the field assessment of the forest condition survey and publish the results for the federal states annually. Within the federal states grid, densifications are common. The evaluation of the 16 × 16 km grid data for the whole of Germany (415 plots in 2012; ranging from 4 plots per state (Berlin) to 96 plots (Bavaria), and no plots for the city states Bremen and Hamburg) is carried out by the Thünen Institute of Forest Ecosystems and the results are finally published by the Federal Ministry for Food, Agriculture and Consumer Protection, which represents the National Focal Centre of Germany (BMELV 2012; Eichhorn et al. 2010).

The forest condition survey is based on the assessment of defoliation (e.g. Arbeitsgruppe AG Diagnose und Klassifizierung der neuartigen Waldschäden 1984; Eichhorn et al. 2010; Innes et al. 1993), which is the most widely used indicator for assessment of tree condition (Ferretti 1997; Ghosh et al. 1995; Innes et al. 1994). Defoliation is assessed visually using binoculars and a scoring system of 5 % classes (e.g. Durrant and Boswell 2002; Eichhorn et al. 2010). Consistency and reproducibility of defoliation data has frequently been in the focus of scientific criticism due to subjectivity in the visual assessment (Ferretti 1998; Innes 1988; Innes 1993; Schöpfer 1985a). Consistency of observations, however, is of major importance for spatial and temporal data evaluations. Several studies conducted in different European countries during the 1980s and beginning 1990s revealed a poor level of reproducibility and significant variations among observer teams (training courses) and between observer and control teams (field check) (Ferretti et al. 1999; Ghosh et al. 1995; Innes et al. 1993; Kandler and Innes 1995; Köhl et al. 2000; Solberg and Strand 1999). In addition, Klap et al. (2000) found the factor ‘country’ as the most statistically significant predictor for European defoliation data whereas Seidling (2004) detected the factor ‘federal states’ as the most significant predictor for defoliation data of the 89 German Level II plots. Despite criticism, an objective and feasible alternative method for the determination of defoliation is not yet available. The image analysis system CROCO has, however, been recommended for verification of visual assessments (Mizoue 2002; Nakajima et al. 2011). Hence, main emphasis has been placed on the quality assurance (QA) programme to improve and to document the consistency and reproducibility of visually assessed defoliation data (e.g. Ferretti et al. 1999). Regular training courses and further QA procedures (e.g. field checks) are believed to remove a great amount of subjectivity and variation among individual observers (Innes 1993; Köhl 1991; Schöpfer 1985b). For instance, a rapid and steady improvement of defoliation data consistency in Italy following the adoption of the QA programme was reported by Bussotti et al. (2009).

Thus, the present study investigated the defoliation data from the annual training courses in Germany from 1992 to 2012. The aim was (1) to evaluate the consistency of defoliation assessments over time and (2) to determine possible tree species-specific differences in the consistency.

Materials and methods

Procedure of the national training courses

The national training course has taken place annually in June since 1984 before the training courses of the federal states, which in turn have taken place immediately before the field surveys in July and August. The four most frequent tree species of Germany are investigated, namely, beech (Fagus sylvatica), oak (Quercus robur and Quercus petraea), Norway spruce (Picea abies), and pine (Pinus sylvestris). The participating observer teams consist of at least one representative who conducts the forest condition survey or who is responsible for the training course in its respective federal state. The training courses aim at eliminating differences in the assessments of defoliation among the observer teams of different federal states in order to obtain consistent data in terms of spatial and temporal comparability within Germany during the forest condition survey. Therefore, interfering factors such as the social position of trees, tree species, and visibility of the crown are kept constant during the course (Köhl 1993). Easily visible trees are generally chosen, although good visibility does usually not reflect the real condition during the forest
condition survey. Moreover, the assessment occurs from a fixed observation point. Before the assessment, five trees of each species are jointly assessed and discussed. Subsequently, ten trees are independently assessed by the observer teams (first round). The results are recorded and distinct discrepancies among the teams are discussed following the first round. Finally, a second round where another ten trees are independently assessed is carried out. Discrepancies among the teams are discussed again. Over the years, the procedure of the training courses has changed. The second round including the discussion in between has been obligatory since 2011 but had already been performed in earlier years, e.g. in 2008 for beech and oak. In general, ten to twenty trees were assessed per tree species; however, the quantity of assessed trees ranged from ten to one hundred (Table 1). During the first years, assessment was not necessarily performed from the same observation point. In 1997, the guidelines for assessment was not necessarily performed from the same observation point. In 1997, the guidelines for assessment were introduced (internal guideline of the AG Dauerbeobachtungsflächen – Waldschäden published in Dammann et al. 2001) as well as the reference book for defoliation ‘Waldbäume – Bilderserien zur Einschätzung von Kronenverlichtungen bei Waldbäumen’ (Arbeitsgruppe AG Diagnose und Klassifizierung der neuartigen Waldschäden 1984; Meining et al. 2007) were introduced. The age of trees was not recorded before 2011, but was over 60 years for most trees. Furthermore, the observer teams of the different federal states have partly changed over time; however, this was not recorded. In some cases, two separate persons representing two different federal states started as a joint team (Berlin and Brandenburg in most years, Bremen and Hamburg in earlier years, and Baden-Württemberg and Bavaria in 2003). The number of persons per team varied as well and was not noted before 2011. Generally, two persons built a team, but sometimes up to five persons or, in some cases, an individual person represented one federal state. Additionally, experience in defoliation assessment varied among the observers from persons being part of the courses since the beginning in the 1980s to inexperienced beginners. The annual training courses have usually taken place in one federal state for two consecutive years (Table 1). Sites are selected so as to reflect a large distribution of the defoliation levels, but the focus is on trees with intermediate defoliation because it has been reported that the defoliation assessment is more difficult than that of healthy or heavily damaged trees (e.g. Solberg and Strand 1999). For 1993, no data were available for oak due to strong pest infestation and for 1996 and 1997, data were unrecoverable.

Observer errors

Potential errors that may occur during surveys are sampling and non-sampling errors (Kish 1995). Sampling errors can be attributed to the fact that the sample does not represent the entirety of the population and can therefore be reduced by increasing the sample size provided that a probabilistic sampling design is adopted (Kish 1995; Schöpfer 1985b). Non-sampling errors include the observer error, which is the result of the visual assessment of defoliation (Kish 1995; Schöpfer 1985b) and the error which is addressed in the present study. The observer error has to be considered additionally to the sampling error when regarding errors of defoliation assessments during the forest condition survey (e.g. Gertner and Köhl 1995). According to Cochran (1977) and Kish (1995), the total error, which is usually given as mean-squared error (MS), is defined as the sum of variable (random) error and bias (systematic error):

\[ MS = total\ error = (\text{variable\ error})^2 + (\text{bias})^2 \] (1)

Precision refers to the random error (size of deviation of the estimated value from the sample mean value) whereas accuracy refers to the total error including the bias (size of deviation of the estimated value from the true mean value) (Cochran 1977; Kish 1995). Therefore, an estimate may be precise but biased. The observer error is defined as difference between the true value of defoliation and the assessed value (Gertner and Köhl 1995; Wulff 2002). Hence, random as well as systematic components can be included in the observer error (Köhl et al. 2000). Systematic errors can be ascribed to several sources of error such as different definitions of the assessable crown, weather conditions, flowering of pine trees, as well as the observer's style of assessment (Schöpfer 1985b; Solberg and Strand 1999). Calculation of the observer error is not trivial, since the true value of defoliation is unknown. In studies like the present one, it is not possible to estimate the true errors and the accuracy of defoliation assessments. When it is assumed that the arithmetic mean of defoliation assessments of all observer teams regarding an individual tree is the unbiased true value, then the mean absolute deviation of defoliation assessments represents an estimator for the observer error.
In the present study, the consistency of defoliation data was evaluated. Therefore, deviations, correlations, and agreements among observer teams (federal states) were examined and variance components were estimated. The defoliation data are scored in 5% classes and thus are pseudo-continuous.

The absolute deviation of defoliation assessments between the observer teams and the arithmetic mean of all observer teams was calculated as well as the standard deviation of defoliation assessments, which corresponded to the standard deviation of the deviation. At site level (one tree species per year), the number of degrees of freedom was corrected because the assessments of different teams on an individual tree were not independent.

A two-way analysis of variance (ANOVA) was performed in order to estimate the variance components as well as systematic and random errors (Wulff 2002). The observer teams were assumed to be a random selection of a population of possible observers, some of which will later carry out the forest condition survey (Bravo and Potvin 1991; Wulff 2002). Additionally, the precision of assessment was assumed to be the same for all k observer teams, i.e. $\sigma_i^2 = \ldots = \sigma_k^2 = \sigma_E^2$ (Köhl 1993; Wulff 2002).

### Table 1 Year, federal state, location, number of assessed trees (trees), and number of participating observer teams (observers) at the national training courses given for the main tree species

| Year | State | Location     | Beech Trees | Beech Observers | Oak Trees | Oak Observers | Spruce Trees | Spruce Observers | Pine Trees | Pine Observers |
|------|-------|--------------|-------------|-----------------|-----------|---------------|--------------|------------------|------------|----------------|
| 1992 | NW    | Altenbeken   | 10          | 15              | 10        | 15            | 10           | 15               | 10         | 15             |
| 1993 | TH    | Schmiedefeld | 10          | 14              | –          | –             | 10           | 13               | 10         | 14             |
| 1994 | BY    | Freising     | 10          | 15              | 15         | 15            | 15           | 13               | 10         | 15             |
| 1995 | BY    | Freising     | 20          | 16              | 20         | 16            | 20           | 16               | 20         | 16             |
| 1996 | BB    | Potsdam      | –           | –               | –          | –             | –            | –                | –          | –              |
| 1997 | BB    | Potsdam      | –           | –               | –          | –             | –            | –                | –          | –              |
| 1998 | SH    | Pronstorf    | 10          | 16              | 10         | 16            | 10           | 16               | 10         | 15             |
| 1999 | SN    | Graupa       | 20          | 15              | 20         | 15            | 20           | 16               | 20         | 16             |
| 2000 | SN    | Graupa       | 20          | 16              | 20         | 15            | 20           | 15               | 20         | 15             |
| 2001 | SH    | Pronstorf    | 90          | 17              | 10         | 17            | 10           | 16               | 10         | 16             |
| 2002 | HE    | Hann. Münden | 95          | 14              | 10         | 13            | 10           | 14               | 10         | 14             |
| 2003 | HE    | Hann. Münden | 20          | 13              | 18         | 13            | 100          | 13               | 15         | 13             |
| 2004 | RP    | Trippstadt   | 17          | 15              | 15         | 15            | 10           | 15               | 16         | 15             |
| 2005 | RP    | Trippstadt   | 15          | 15              | 15         | 15            | 15           | 15               | 10         | 15             |
| 2006 | MV    | Dümmer       | 15          | 13              | 15         | 12            | 15           | 13               | 30         | 13             |
| 2007 | MV    | Dümmer       | 15          | 14              | 15         | 14            | 15           | 14               | 20         | 14             |
| 2008 | BY    | Freising     | 20          | 13              | 20         | 13            | 10           | 13               | 10         | 13             |
| 2009 | BY    | Freising     | 10          | 12              | 10         | 12            | 10           | 12               | 10         | 11             |
| 2010 | SL    | Homburg      | 10          | 12              | 10         | 12            | 10           | 12               | 10         | 12             |
| 2011 | SL    | Bexbach      | 20          | 10              | 20         | 10            | 20           | 10               | 20         | 10             |
| 2012 | HE    | Witzenhausen | 20          | 10              | 20         | 10            | 20           | 10               | 20         | 10             |

For 1993, no data were available for oak due to strong pest infestation, and for 1996 and 1997, data were unrecoverable. The observer teams generally represent the federal states, but Saxony-Anhalt started with two teams from 1993 to 2005, Berlin and Brandenburg started as joint team in most years, and in 2006, a cross-federal states institution was established comprising the states Hesse, Lower Saxony, and Saxony-Anhalt as well as Schleswig-Holstein (since 2011). The institution started with one to three observer teams.

**Abbreviations:** BY Bavaria, BB Brandenburg, HE Hesse, MV Mecklenburg-Western Pomerania, NW North Rhine-Westphalia, RP Rheinland-Palatinate, SH Schleswig-Holstein, SL Saarland, SN Saxony, and TH Thuringia.
The general model was described by Bravo and Potvin (1991) and Wulff (2002):

\[ X_{ij} = \mu + S_i + O_j + E_{ij}, \]

where \( X_{ij} \) is the assessment by observer team \( j \) on tree \( i \), \( \mu \) is the grand mean of all estimations, \( S_i \) is the tree effect, \( O_j \) is the observer bias for observer team \( j \), and \( E_{ij} \) is the random error of the assessment by observer team \( j \) on tree \( i \). The bias is assumed to be the same on all kinds of trees. The term \( \mu + S \) denotes the true value. In practice, \( \mu + S \) will be the arithmetic mean value of all assessments. The term \( O + E \) denotes the error term, which is differentiated in systematic and random error components. The estimations for the corresponding variance (\( s^2 \)) components are (Bravo and Potvin 1991; Wulff 2002):

\[ s^2_E = MS_E \]

\[ s^2_O = \frac{MS_O - MS_E}{n_i}, \]

where \( n \) is the number of trees. It was tested if observer teams differed significantly from the average defoliation assessment (\( H_0: \sigma_o^2=0 \)).

Correlations between the observer teams were examined using the Pearson correlation coefficient \( r \) and agreements among observers were examined using the intraclass correlation coefficient (ICC) (Bravo and Potvin 1991):

\[ ICC = \frac{n(MS_S - MS_E)}{nMS_S + kMS_O + (nk-n-k)MS_E}, \]

where \( k \) is the number of observer teams.

Furthermore, analysis of variance combined with Tukey’s HSD test was used for multiple comparisons of means. In the case of non-normality of data, the non-parametric Kruskal–Wallis H test combined with the multiple comparison test by Castellan and Siegel (1988) were used. The non-parametric Mann–Whitney U test was used for comparison of two means. In addition, simple linear regressions were performed. In the case of heteroscedasticity of variances the generalised least squares method with a variance function was applied. Normality of residuals and homogeneity of variances were tested prior to all statistical analyses. Statistical significance was stated at \( P \leq 0.05 \). The whiskers of the boxplots correspond to 1.5 times the interquartile distance. Evaluation and visualisation of data were performed using R 2.15.1 (R Development Core Team 2012).

**Results**

The absolute deviation from the arithmetic mean at site level averaged 4.4 % defoliation (3.2–7.1 %) from 1992 to 2012, and no significant differences were observed among the tree species. The absolute deviations of oak \( (P=0.015) \) and spruce \( (P=0.017) \) decreased from 1992 to 2012 (Fig. 1). In 2011 and 2012, for example, the mean absolute deviations for all tree species were 3.5 and 3.6 %, respectively. The highest absolute deviations were observed in 2008 (7.1 %) and 2009 (5.8 %) at pine sites. On average, 93.4 % (spruce) to 95.2 % (oak) of the assessments from 1992 to 2012 were located within the ±10 % interval of deviation from the mean (Fig. 2). Assessments outside this interval were detected at all sites in every year. The proportion of assessments within the interval increased over the years for oak \( (P=0.001) \) and spruce \( (P=0.02) \). In 2011, 94.0 (beech) to 99.5 % (spruce) and in 2012, 96.0 (pine) to 99.0 % (oak) of the assessments lay within this interval. Moreover, 72.0 (beech) to 86.6 % (pine) of the assessments were located within the ±5 % interval of deviation in 2011 and 70.5 (pine) to 84.0 % (oak) of assessments in 2012. The lowest proportion of assessments within the ±5 % interval (48.5 and 56.4 %) and ±10 % interval of deviation (78.5 and 85.5 %) were observed at the pine sites in 2008 and 2009. The introduction of two rounds during the training courses had no significant effect on the error of assessment in beech, oak, and spruce stands considering the years 2008 (only beech and oak), 2011, and 2012 (Fig. 3). The range of deviation even tended to increase from the first to the second round. For pine, the absolute deviation, however, decreased from the first to the second round \( (P=0.001) \) (Fig. 3). This decrease occurred in 2011 as well as in 2012.

The maximal positive and negative deviations at individual trees were 34.7 and −47.8 % defoliation both of which were observed at the pine sites in 1995. The maximal deviations at individual oak (−25.0 %) and spruce (32.3 %) trees were also found in 1995 whereas the maximal deviation of beech trees (30.7 %) was observed in 2004. Deviations greater
than ±20 % were rarely observed and accounted for less than or equal to 1 % of the total assessments of one site except for the pine sites in 1995 and 2008 and the oak site in 1995 where these deviations accounted for up to 3 % of the total assessments.

The average standard deviation at site level was 5.5 % defoliation ranging from 3.9 to 8.6 %. As with the absolute deviation, no significant differences occurred among tree species and the standard deviation of oak ($P<0.005$) and spruce ($P<0.026$) decreased from 1992 to 2012 (Fig. 4). Mean standard deviations for all tree species in 2011 and 2012 were 4.4 and 4.5 % defoliation, respectively.

Systematic errors in assessed defoliations that were attributed to the observer teams were detected at all tree species and in nearly every year (Table 2). However, few consistent temporal or species-specific patterns regarding systematic errors among observer teams were determined. The variance among observers of oak decreased from 1992 to 2012 ($P=0.036$, $R^2=0.26$). The variance among observers (systematic error) was approximately one fifth of the variance of the random error regarding every year between 1992 and 2012 (Table 2). The average ICC was 0.83 and ranged from 0.52 (pine 1994, spruce 2001) to 0.97 (pine 2005, oak 2012) (Table 2). No temporal or species-specific trends were observed. In 2011 and 2012, the mean ICC were 0.89 and 0.91 and the Pearson correlation coefficients $r$ were 0.91 (0.71–0.98) and 0.93 (0.73–0.99), respectively. For both years, ICC and $r$ displayed the highest values for oak trees. The Pearson correlations were significant among all combinations of observer teams ($P<0.001$).

The four tree species differed in the frequency of the level of defoliation (Fig. 5). Beech trees showed the highest mean defoliation with 37 % whereas pine trees showed the lowest mean defoliation with 25 %. The frequency distributions of defoliation of all tree species displayed a distinct skewness to the right and defoliations more than 60 % were rarely present. The absolute deviation from the mean significantly and non-linearly depended on the defoliation level ($P>0.001$) (Fig. 6), being highest at intermediate levels of defoliation.

![Fig. 1 Mean absolute deviation of defoliation assessments from the arithmetic mean given for the beech sites (top, left side), the oak sites (top, right side), the spruce sites (bottom, left side), and the pine sites (bottom, right side) for the years 1992 to 2012. The lines and parameters of significant (oak, spruce) and insignificant (beech, pine) linear regressions are indicated](image-url)
Discussion

Temporal changes in the observer error

The mean absolute deviation of 4.4 % defoliation, which may be used as estimate for the observer error, assuming that the arithmetic mean of all assessments of an individual tree represents the unbiased true value, as well as the mean standard deviation of 5.5 % defoliation were comparable or lower than deviations reported in the literature. Additionally, both measures displayed a decreasing trend from 1992 to 2012 for the four tree species, which was statistically significant for oak and spruce sites. In comparison, Schöpfer (1985b) estimated a higher observer error of ±8 % defoliation at the national training courses in 1983 and 1984. The mean deviation that was calculated among three European countries at a training course in 1989 was higher as well (deviation between two countries) (Innes et al. 1993). Standard deviations of single observations of control surveys (field checks) in Sweden amounted to 4.7–12.6 % defoliation between 1995 and 1999 (Wulff 2002). Solberg and Strand (1999) estimated a standard deviation of 10 % for single trees and of 5 % for plot means from field checks in Norway between 1990 and 1995. The reported standard deviations are similar but
not directly comparable to our results since the reported ones were derived from pairwise assessments. In general, a deviation of ±10 % defoliation from a reference value is an acceptable limit of deviation (e.g. Innes 1988; Köhl 1991). In the present study, 94 % of the assessments were located within these limits and the proportion of assessments within the limits displayed an increasing trend, which was significant for oak and spruce. In fact, pronounced deviations (more than ±20 % defoliation) from the mean were not observed during the last three years. Proportions determined during the national training courses in Italy and the South European training courses between 1996 and 2004 were slightly lower ranging from 80 to 90 %, and a marginal increase in the proportion was observed over time (Bussotti et al. 2009). The proportions calculated in the present study are high in comparison to those reported in the literature from the second assessments by a control team during field checks. Innes et al. (1994) reported from a field check in Switzerland in 1993 that the quality limits had to be broadened to ±15 % defoliation to achieve an acceptable result with more than 90 % of the assessments lying within this interval. Innes (1993), Ferretti et al. (1999), and Solberg and Strand (1999) came to similar results during field checks in Great Britain (1988), Italy (1996), and Norway (1990–95), respectively. In contrast, Bussotti et al. (2009) found that more than or equal to 90 % of the assessments during the field control in Italy from 1999 to 2004 fell between ±10 % with respect to the reference team. In recent times, the prescribed aim (data quality limit) of ICP Forests for field checks during the forest condition survey was set to at least more than or equal to 70 % of assessments that must not deviate more than ±10 % from one another (Eichhorn et al. 2010). The results at hand suggest to introduce a data quality limit for national and international training courses requiring that 90 % of the assessments have to range within ±10 %
from the reference value (here arithmetic mean of all assessments on an individual tree).

The observed trend towards more consistent assessments among observer teams may be ascribed to several changes over time. In particular, the introduction of the guidelines for assessment (internal guideline of the AG Dauerbeobachtungsflächen – Waldschäden published in Dammann et al. 2001) and the defoliation reference book in 1997 (Arbeitsgemeinschaft AG Dauerbeobachtungsflächen der Länder und des Bundes 1997; Meining et al. 2007) may have improved the assessments, although no abrupt improvement occurred following the introduction. Additionally, the definition of the observation point for assessment probably led to greater consistency. A positive effect of the introduction of two rounds was shown for pine. Maintenance of the achieved consistency during the forest condition survey, however, is of great importance. The positive effect was not observed for beech, oak, and spruce where the discussion in between possibly led to an increased uncertainty in the assessment. Since mandatory second rounds have been introduced only recently, no final conclusion can be made whether this implementation represents an improvement for harmonisation.

Explanation for outliers

Occasionally, deviations of more than ±20 % defoliation were determined at individual trees of all tree species but the cases where deviations of more than ±20 % accounted for slightly more than 1 % of total assessments occurred only at three sites and thus played a negligible role during the training courses. The most pronounced deviations and the highest proportion of deviations outside the ±10 % interval of deviation were observed at pine trees in 1995, 2008, and 2009. Innes et al. (1993) reported similar high deviations of ±45 % defoliation at individual trees, which were observed at a training course among three European countries in 1989. Innes et al. (1993) and Wulff (2002) also mentioned difficulties in the defoliation assessment of pine trees. However, despite difficulties in 1995, 2008, and 2009, pine trees did not differ from the other species over time in the present study. In these years, the training courses were carried out...
| Year | Tree   | Source of variance | df | SS      | MS   | $s^2_O$ | F ratio | P value | ICC  |
|------|--------|--------------------|----|---------|------|---------|---------|---------|------|
| 2012 | Beech  | Tree               | 19 | 58,357  | 3,071| 1.98    | 3.37    | 0.0008  | 0.94 |
|      |        | Observer           | 9  | 506     | 56   | 1.98    | 3.37    | 0.0000  | 0.97 |
|      |        | Error              | 171| 2,851   | 17   | 1.98    | 3.37    | 0.0000  | 0.97 |
|      | Oak    | Tree               | 19 | 111,053 | 5,845| 1.96    | 3.60    | 0.0004  | 0.97 |
|      |        | Observer           | 9  | 488     | 54   | 1.96    | 3.60    | 0.0004  | 0.97 |
|      |        | Error              | 171| 2,577   | 15   | 1.96    | 3.60    | 0.0004  | 0.97 |
|      | Spruce | Tree               | 19 | 35,297  | 1,858| 3.79    | 4.88    | 0.0000  | 0.89 |
|      |        | Observer           | 9  | 857     | 95   | 3.79    | 4.88    | 0.0000  | 0.89 |
|      |        | Error              | 171| 3,338   | 20   | 3.79    | 4.88    | 0.0000  | 0.89 |
|      | Pine   | Tree               | 19 | 30,366  | 1,598| 5.22    | 5.53    | 0.0000  | 0.85 |
|      |        | Observer           | 9  | 1,148   | 128  | 5.22    | 5.53    | 0.0000  | 0.85 |
|      |        | Error              | 171| 3,945   | 23   | 5.22    | 5.53    | 0.0000  | 0.85 |
| 2011 | Beech  | Tree               | 19 | 37,154  | 1,956| −0.18   | 0.88    | 0.5440  | 0.87 |
|      |        | Observer           | 9  | 236     | 26   | −0.18   | 0.88    | 0.5440  | 0.87 |
|      |        | Error              | 171| 5,082   | 30   | −0.18   | 0.88    | 0.5440  | 0.87 |
|      | Oak    | Tree               | 19 | 50,997  | 2,684| 1.73    | 3.17    | 0.0014  | 0.94 |
|      |        | Observer           | 9  | 454     | 51   | 1.73    | 3.17    | 0.0014  | 0.94 |
|      |        | Error              | 171| 2,720   | 16   | 1.73    | 3.17    | 0.0014  | 0.94 |
|      | Spruce | Tree               | 19 | 23,198  | 1,221| 1.09    | 2.29    | 0.0186  | 0.87 |
|      |        | Observer           | 9  | 348     | 39   | 1.09    | 2.29    | 0.0186  | 0.87 |
|      |        | Error              | 171| 2,882   | 17   | 1.09    | 2.29    | 0.0186  | 0.87 |
|      | Pine   | Tree               | 19 | 26,592  | 1,400| 2.07    | 3.92    | 0.0002  | 0.89 |
|      |        | Observer           | 9  | 501     | 56   | 2.07    | 3.92    | 0.0002  | 0.89 |
|      |        | Error              | 171| 2,431   | 14   | 2.07    | 3.92    | 0.0002  | 0.89 |
| 2010 | Beech  | Tree               | 9  | 29,872  | 3,319| 6.17    | 5.01    | 0.0000  | 0.93 |
|      |        | Observer           | 11 | 848     | 77   | 6.17    | 5.01    | 0.0000  | 0.93 |
|      |        | Error              | 99 | 1,523   | 15   | 6.17    | 5.01    | 0.0000  | 0.93 |
|      | Oak    | Tree               | 9  | 47,573  | 5,286| 2.93    | 1.99    | 0.0368  | 0.93 |
|      |        | Observer           | 11 | 646     | 59   | 2.93    | 1.99    | 0.0368  | 0.93 |
|      |        | Error              | 99 | 2,915   | 29   | 2.93    | 1.99    | 0.0368  | 0.93 |
|      | Spruce | Tree               | 9  | 20,032  | 2,226| −0.83   | 0.63    | 0.8020  | 0.90 |
|      |        | Observer           | 11 | 154     | 14   | −0.83   | 0.63    | 0.8020  | 0.90 |
|      |        | Error              | 99 | 2,212   | 22   | −0.83   | 0.63    | 0.8020  | 0.90 |
|      | Pine   | Tree               | 9  | 18,229  | 2,026| 6.00    | 4.52    | 0.0000  | 0.88 |
|      |        | Observer           | 11 | 847     | 77   | 6.00    | 4.52    | 0.0000  | 0.88 |
|      |        | Error              | 99 | 1,688   | 17   | 6.00    | 4.52    | 0.0000  | 0.88 |
| 2009 | Beech  | Tree               | 9  | 20,096  | 2,233| 12.45   | 5.25    | 0.0000  | 0.81 |
|      |        | Observer           | 11 | 1,692   | 154  | 12.45   | 5.25    | 0.0000  | 0.81 |
|      |        | Error              | 99 | 2,901   | 29   | 12.45   | 5.25    | 0.0000  | 0.81 |
|      | Oak    | Tree               | 9  | 18,551  | 2,061| 4.33    | 3.40    | 0.0005  | 0.88 |
|      |        | Observer           | 11 | 674     | 61   | 4.33    | 3.40    | 0.0005  | 0.88 |
|      |        | Error              | 99 | 1,784   | 18   | 4.33    | 3.40    | 0.0005  | 0.88 |
|      | Spruce | Tree               | 9  | 52,544  | 5,838| 4.33    | 3.40    | 0.0005  | 0.88 |
Table 2 (continued)

| Year | Tree        | Source of variance | df | SS   | MS  | $s^2_O$ | F ratio | P value | ICC  |
|------|-------------|--------------------|----|------|-----|---------|---------|---------|------|
| 2008 | Beech       | Observer           | 11 | 552  | 50  | 1.71    | 1.52    | 0.1380  | 0.93 |
|      |             | Error              | 99 | 3,279| 33  |         |         |         |      |
|      | Pine        | Tree               | 9  | 7,984| 887 |         |         |         |      |
|      |             | Observer           | 10 | 1,882| 188 | 15.53   | 5.72    | 0.0000  | 0.62 |
|      |             | Error              | 90 | 2,964| 33  |         |         |         |      |
|      | 2008        | Oak                | 9  | 7,984| 887 |         |         |         |      |
|      |             | Observer           | 12 | 1,335| 111 | 4.52    | 5.36    | 0.0000  | 0.78 |
|      |             | Error              | 228| 5,653| 25  |         |         |         |      |
|      | 2008        | Spruce             | 9  | 7,873| 875 |         |         |         |      |
|      |             | Observer           | 12 | 1,939| 162 | 13.10   | 5.28    | 0.0000  | 0.60 |
|      |             | Error              | 108| 3,307| 31  |         |         |         |      |
|      | 2007        | Beech              | 14 | 29,775| 2,127|         |         |         |      |
|      |             | Observer           | 13 | 829  | 64  | 3.09    | 3.67    | 0.0000  | 0.88 |
|      |             | Error              | 182| 3,159| 17  |         |         |         |      |
|      | 2007        | Oak                | 14 | 11,890| 849 |         |         |         |      |
|      |             | Observer           | 13 | 935  | 72  | 3.42    | 3.50    | 0.0001  | 0.71 |
|      |             | Error              | 182| 3,737| 21  |         |         |         |      |
|      | 2007        | Spruce             | 14 | 17,305| 1,236|         |         |         |      |
|      |             | Observer           | 13 | 1,021| 79  | 3.83    | 3.73    | 0.0000  | 0.78 |
|      |             | Error              | 182| 3,835| 21  |         |         |         |      |
|      | 2006        | Beech              | 14 | 29,528| 2,109|         |         |         |      |
|      |             | Observer           | 12 | 1,822| 152 | 8.69    | 7.07    | 0.0000  | 0.84 |
|      |             | Error              | 168| 3,609| 22  |         |         |         |      |
|      | 2006        | Oak                | 14 | 33,038| 2,360|         |         |         |      |
|      |             | Observer           | 11 | 1,014| 92  | 4.44    | 3.60    | 0.0002  | 0.87 |
|      |             | Error              | 154| 3,942| 26  |         |         |         |      |
|      | 2006        | Spruce             | 14 | 13,660| 976 |         |         |         |      |
|      |             | Observer           | 12 | 1,604| 134 | 7.00    | 4.67    | 0.0000  | 0.67 |
|      |             | Error              | 168| 4,804| 29  |         |         |         |      |
|      | 2005        | Beech              | 14 | 54,490| 3,892|         |         |         |      |
|      |             | Observer           | 14 | 2,116| 151 | 7.94    | 4.71    | 0.0000  | 0.87 |
|      |             | Error              | 196| 6,287| 32  |         |         |         |      |
| Year  | Tree    | Source of variance | df  | SS     | MS  | $\sigma^2_O$ | F ratio | P value | ICC |
|-------|---------|--------------------|-----|--------|-----|-------------|---------|---------|-----|
| 2004  | Beech   | Tree               | 16  | 55,109 | 3,444 | 5.57 | 3.72 | 0.0000 | 0.85 |
|       |         | Observer           | 14  | 1,812  | 129  | 3.18 | 2.28 | 0.0083 | 0.91 |
|       |         | Error              | 224 | 7,785  | 35   | 3.09 | 2.48 | 0.0050 | 0.94 |
|       | Spruce  | Tree               | 9   | 60,938 | 6,771 | 105  | 6.34 | 0.0001 | 0.97 |
|       |         | Observer           | 14  | 1,703  | 122  | 3.09 | 2.48 | 0.0050 | 0.94 |
|       |         | Error              | 126 | 3,680  | 29   | 3.09 | 2.48 | 0.0050 | 0.94 |
|       | Pine    | Tree               | 15  | 89,670 | 5,978 | 5.03 | 3.68 | 0.0000 | 0.92 |
|       |         | Observer           | 14  | 1,546  | 110  | 4.87 | 5.03 | 0.0000 | 0.92 |
|       |         | Error              | 210 | 6,294  | 30   | 4.87 | 5.03 | 0.0000 | 0.92 |
| 2003  | Beech   | Tree               | 19  | 109,002| 5,737 | 9.25 | 4.17 | 0.0000 | 0.92 |
|       |         | Observer           | 12  | 959    | 80   | 3.19 | 4.95 | 0.0000 | 0.96 |
|       |         | Error              | 228 | 3,679  | 16   | 3.19 | 4.95 | 0.0000 | 0.96 |
|       | Spruce  | Tree               | 17  | 58,100 | 3,418 | 3.42 | 4.09 | 0.0000 | 0.92 |
|       |         | Observer           | 12  | 977    | 81   | 3.42 | 4.09 | 0.0000 | 0.92 |
|       |         | Error              | 204 | 4,061  | 20   | 3.42 | 4.09 | 0.0000 | 0.92 |
|       | Pine    | Tree               | 15  | 45,055 | 3,218 | 4.87 | 5.03 | 0.0000 | 0.91 |
|       |         | Observer           | 12  | 1,095  | 91   | 4.87 | 5.03 | 0.0000 | 0.91 |
|       |         | Error              | 168 | 3,055  | 18   | 4.87 | 5.03 | 0.0000 | 0.91 |
| 2002  | Beech   | Tree               | 94  | 241,307| 2,567 | 15.657 | 13.05 | 45.63 | 0.0000 | 0.78 |
|       |         | Observer           | 13  | 5,927  | 456  | 4.52 | 17.37 | 0.0000 | 0.85 |
|       |         | Error              | 1,222 | 32,085 | 26   | 4.52 | 17.37 | 0.0000 | 0.85 |
|       | Oak     | Tree               | 9   | 45,882 | 5,098 | 4.87 | 5.03 | 0.0000 | 0.91 |
|       |         | Observer           | 12  | 525    | 44   | 2.56 | 2.42 | 0.0082 | 0.95 |
|       |         | Error              | 108 | 1,956  | 18   | 2.56 | 2.42 | 0.0082 | 0.95 |
|       | Spruce  | Tree               | 9   | 7,155  | 795  | 7.35 | 4.52 | 0.0000 | 0.66 |
|       |         | Observer           | 13  | 1,227  | 94   | 7.35 | 4.52 | 0.0000 | 0.66 |
|       |         | Error              | 117 | 2,442  | 21   | 7.35 | 4.52 | 0.0000 | 0.66 |
|       | Pine    | Tree               | 9   | 47,666 | 5,296 | 3.09 | 2.48 | 0.0050 | 0.94 |
|       |         | Observer           | 13  | 672    | 52   | 3.09 | 2.48 | 0.0050 | 0.94 |
### Table 2 (continued)

| Year | Tree  | Source of variance | df  | SS     | MS     | $s_{o}^{2}$ | F ratio | P value | ICC  |
|------|-------|--------------------|-----|--------|--------|-------------|---------|---------|------|
| 2001 | Beech | Error              | 117 | 2,437  | 21     |             |          |         |      |
|      |       | Tree               | 89  | 183,824| 2,065  |             |          |         |      |
|      |       | Observer           | 16  | 14,313 | 895    | 9.58        | 27.84   | 0.0000  | 0.74 |
|      |       | Error              | 1,424| 45,751 | 32     |             |          |         |      |
|      | Oak   | Tree               | 9   | 23,414 | 2,602  |             |          |         |      |
|      |       | Observer           | 16  | 1,415  | 88     | 6.91        | 4.56    | 0.0000  | 0.85 |
|      |       | Error              | 144 | 2,791  | 19     |             |          |         |      |
|      | Spruce| Tree               | 9   | 6,226  | 692    |             |          |         |      |
|      |       | Observer           | 15  | 1,464  | 98     | 6.54        | 3.03    | 0.0003  | 0.52 |
|      |       | Error              | 135 | 4,354  | 32     |             |          |         |      |
|      | Pine  | Tree               | 9   | 15,091 | 1,677  |             |          |         |      |
|      |       | Observer           | 15  | 934    | 62     | 3.51        | 2.29    | 0.0063  | 0.77 |
|      |       | Error              | 135 | 3,669  | 27     |             |          |         |      |
| 2000 | Beech | Tree               | 19  | 69,070 | 3,635  |             |          |         |      |
|      |       | Observer           | 15  | 1,972  | 131    | 5.05        | 4.32    | 0.0000  | 0.86 |
|      |       | Error              | 285 | 8,676  | 30     |             |          |         |      |
|      | Oak   | Tree               | 19  | 184,565| 9,714  |             |          |         |      |
|      |       | Observer           | 14  | 1,100  | 79     | 2.49        | 2.74    | 0.0009  | 0.95 |
|      |       | Error              | 266 | 7,634  | 29     |             |          |         |      |
|      | Spruce| Tree               | 19  | 21,455 | 1,129  |             |          |         |      |
|      |       | Observer           | 14  | 1,575  | 113    | 4.69        | 6.04    | 0.0000  | 0.76 |
|      |       | Error              | 266 | 4,955  | 19     |             |          |         |      |
|      | Pine  | Tree               | 19  | 30,674 | 1,614  |             |          |         |      |
|      |       | Observer           | 14  | 1,587  | 113    | 4.62        | 5.40    | 0.0000  | 0.81 |
|      |       | Error              | 266 | 5,589  | 21     |             |          |         |      |
| 1999 | Beech | Tree               | 19  | 56,770 | 2,988  |             |          |         |      |
|      |       | Observer           | 14  | 1,611  | 115    | 4.02        | 3.33    | 0.0001  | 0.84 |
|      |       | Error              | 266 | 9,202  | 35     |             |          |         |      |
|      | Oak   | Tree               | 19  | 178,749| 9,408  |             |          |         |      |
|      |       | Observer           | 14  | 952    | 68     | 2.37        | 3.29    | 0.0001  | 0.96 |
|      |       | Error              | 266 | 5,491  | 21     |             |          |         |      |
|      | Spruce| Tree               | 19  | 20,805 | 1,095  |             |          |         |      |
|      |       | Observer           | 15  | 1,141  | 76     | 2.85        | 4.00    | 0.0000  | 0.75 |
|      |       | Error              | 285 | 5,417  | 19     |             |          |         |      |
|      | Pine  | Tree               | 19  | 37,445 | 1,971  |             |          |         |      |
|      |       | Observer           | 15  | 1,044  | 70     | 2.53        | 3.67    | 0.0000  | 0.85 |
|      |       | Error              | 285 | 5,402  | 19     |             |          |         |      |
| 1998 | Beech | Tree               | 9   | 44,195 | 4,911  |             |          |         |      |
|      |       | Observer           | 15  | 1,312  | 87     | 6.68        | 4.23    | 0.0000  | 0.92 |
|      |       | Error              | 135 | 2,792  | 21     |             |          |         |      |
|      | Oak   | Tree               | 9   | 81,787 | 9,087  |             |          |         |      |
|      |       | Observer           | 15  | 1,809  | 121    | 9.34        | 4.43    | 0.0000  | 0.94 |
|      |       | Error              | 135 | 3,678  | 27     |             |          |         |      |
|      | Spruce| Tree               | 9   | 24,765 | 2,752  |             |          |         |      |
Table 2 (continued)

| Year | Tree     | Source of variance | $df$ | SS   | MS   | $s_{O}^2$ | F ratio | $P$ value | ICC  |
|------|----------|--------------------|------|------|------|----------|---------|-----------|------|
|      |          |                    |      |      |      |          |         |           |      |
| 1995 | Beech    | Tree               | 19   | 45,508 | 2,395 |          |         |           |      |
|      |          | Observer           | 15   | 687   | 46   | 1.14     | 1.99    | 0.0159    | 0.86 |
|      |          | Error              | 285  | 6,560 | 23   |          |         |           |      |
| 1994 | Beech    | Tree               | 9    | 15,217 | 1,691 |          |         |           |      |
|      |          | Observer           | 14   | 1,307 | 93   | 5.83     | 2.66    | 0.0020    | 0.73 |
|      |          | Error              | 126  | 4,423 | 35   |          |         |           |      |
| 1993 | Beech    | Tree               | 9    | 38,176 | 4,242 |          |         |           |      |
|      |          | Observer           | 13   | 1,689 | 130  | 9.31     | 3.52    | 0.0001    | 0.87 |
|      |          | Error              | 117  | 4,314 | 37   |          |         |           |      |
| 1992 | Beech    | Tree               | 9    | 14,323 | 1,591 |          |         |           |      |
|      |          | Observer           | 14   | 726   | 52   | 3.12     | 2.51    | 0.0035    | 0.81 |
|      |          | Error              | 126  | 2,607 | 21   |          |         |           |      |
|      | Oak      | Tree               | 9    | 65,228 | 7,248 |          |         |           |      |
|      |          | Observer           | 14   | 923   | 66   | 4.00     | 2.54    | 0.0031    | 0.94 |
|      |          | Error              | 126  | 3,270 | 26   |          |         |           |      |
out in Bavaria. According to the participants, the selected pine trees showed an uncommon type of growth for pine trees in Germany. It was not possible to classify the defoliation using the reference book due to the special type of growth. However, these errors are negligible for the forest condition surveys where this growth type is very rare.

Common reasons for frequent occurrence of deviations of more than ±10% on individual trees were difficulties in setting the assessable crown or anomalies such as crown damages, substitution crowns, and uncommon growth types (the information was derived from participants but was not recorded in the past). In spite of high consistencies among the observer teams.

Table 2 (continued)

| Year | Tree | Source of variance | df | SS | MS | $s^2_o$ | F ratio | $P$ value | ICC |
|------|------|-------------------|----|----|----|--------|--------|---------|-----|
| 1992 | Spruce | Tree | 9  | 48,061 | 5,340 |        |        |        |     |
|      |       | Observer         | 14 | 1,232 | 88  | 5.35   | 2.55   | 0.0030  | 0.90 |
|      |       | Error            | 126| 4,344 | 34  |        |        |         |     |
| 1993 | Pine  | Tree             | 9  | 39,984 | 4,443 |        |        |        |     |
|      |       | Observer         | 14 | 1,292 | 92  | 6.91   | 3.98   | 0.0000  | 0.91 |
|      |       | Error            | 126| 2,921 | 23  |        |        |         |     |

Negative variances should be interpreted as zero. For 1993, no data were available for oak due to strong pest infestation and for 1996 and 1997, data were unrecoverable.

Fig. 5 Frequency distribution of defoliation levels for beech (top, left side), oak (top, right side), spruce (bottom, left side), and pine (bottom, right side) including all data from 1992 to 2012. The number of assessed trees is indicated by ‘n’.
in 2011 and 2012, the assessments on beech and pine displayed higher deviations compared to the other species in both years. The tree species showed an extraordinary fructification in 2011, which in case of beech was additionally accompanied by notably small-sized leaves, which caused the comparable high uncertainties in 2011. However, a systematic influence of fructification on defoliation assessment could not be observed between 1992 and 2012. In 2012, the actual needle set of the pine trees was still lacking, as the training course took place in early June. The absence of the actual needle set increased the level of error because the observers had to imagine the needle set. This error, however, is unimportant for the forest condition survey, which takes place later in the year when the actual needles are developed.

Systematic and random errors

Systematic errors among observer teams were found in nearly every year for the four tree species and have frequently been reported from studies investigating the quality of defoliation assessment (Ghosh et al. 1995; Innes 1993; Mues and Seidling 2003; Solberg and Strand 1999; Wulff 2002). Systematic errors resulting from different weather conditions could be excluded because all assessments took place at the same time and were additionally conducted from the same observation point. Systematic temporal or species-specific patterns regarding significant differences among observer teams (federal states) were hardly observed due probably to changes in the assembling of observer teams over time. However, in several years, two of the federal states showed systematic deviations for one tree species each. In one case, the corresponding tree species is rarely represented in the respective federal state and plays no role for the forest condition survey. The federal state that deviated the most over the years and for the four tree species was meanwhile taken over by a cross-federal states institution, which is now responsible for the forest condition survey. Under the cross-federal states institution, the survey is conducted by the same teams as before but systematic errors have not been observed so far.

Between-observer variances were lower than variances of the random error (within-observer variances). The variances were comparable to variances given for observer teams at the national training courses in Sweden in 1995–1999 (Wulff 2002). In 2011 and 2012, variances were in general low as compared to earlier years. The fructification and small-sized leaves of beech in 2011 resulted in comparably high uncertainties in the assessments due to a relatively high random error whereas a systematic error was not observed.
In contrast, the absence of the actual needle set at pine trees in 2012 resulted in a comparably high systematic error, which however was not relevant to the survey as already mentioned.

In spite of significant systematic errors, correlations and agreements among the observer teams were extraordinarily high. Agreements were slightly higher for deciduous trees than for coniferous trees. To evaluate whether assessments of one year and one species were consistent, we applied a three-way evaluation approach. Assessments were considered inconsistent if (1) the mean absolute deviation from the mean was more than ±5 %, (2) less than 90 % of the assessment lay within the ±10 % interval of deviation from the mean, and (3) significant systematic differences among observer teams existed. The present study demonstrated that defoliation assessments at the national training courses were consistent with exception of the assessments at the beech site in 2005, at the spruce site in 1993, and at the pine sites in 1995, 2008, and 2009. The inconsistency at the pine sites in Bavaria can be attributed to the growth type (see above) whereas reasons for the inconsistency at the beech and spruce sites could not be reconstructed. Results from the international cross-comparison courses in Europe in 2001 and 2002 also indicated consistent assessments among teams from one country (Mues and Seidling 2003). Despite consistent defoliation assessments among the German federal states, temporal and spatial evaluations of defoliation data from the forest condition survey should focus on pronounced alterations due to between-observer variance and, particularly, due to often considerable within-observer variance. Solberg and Strand (1999) as well as Wulff (2002) made similar statements for the Norwegian and Swedish forest condition survey.

Conclusions

The present study demonstrated that the visual defoliation assessment produced consistent results within Germany at the national training courses for the forest condition survey from 1992 to 2012. Significant tree species-specific differences in the deviations were not observed but the assessment of deciduous trees tended to be slightly more consistent than the assessment of coniferous trees. In large part, pronounced deviations that were observed at the courses (e.g. assessments on pine trees in 1995, 2008, and 2009) are probably of little relevance for the national forest condition survey. Assuming similar assessment behaviours of the observers, a similar distribution of defoliation levels, and that the mean value of all observations on an individual tree represents the unbiased true value, then an observer error of one ±5 % class (absolute deviation of ±4.4 % defoliation from the mean) has to be considered in addition to the sampling error during the forest condition survey. The true bias could, however, not be calculated in the present study. In order to ensure that assessment behaviours do not drift apart subsequent to the training course and that the results can be generalised to the forest condition survey, an intercomparison course during the forest condition survey at which the federal states have to assess trees without previous consultation may be an appropriate measure. In addition, considerable temporal and spatial alterations should be the main focus of interest of the national defoliation assessments rather than short-term trends due to random and systematic observer errors.
Acknowledgments We thank the responsible institutions of the federal states for providing the data of the training courses. We further acknowledge the discussions and the effort of the Arbeitskreis Kronenzustand as well as of the participants of the national training courses. We particularly like to thank I. Dammann (NW-FVA) and H.W. Schröck (FAWF) for information and discussion.

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

References

Arbeitsgemeinschaft (AG) Dauerbeobachtungsflächen der Länder und des Bundes. (1997). Waldbäume - Bilderserien zur Einschätzung von Kronenverlichtungen bei Waldbäumen. (pp. 155). Kassel, Germany: Verlag M. FASTE.

Arbeitsgruppe (AG) Diagnose und Klassifizierung der neuartigen Waldschäden. (1984). Diagnose und Klassifizierung der neuartigen Waldschäden. (pp. 24). AFZ-Farbbildheft. München, Germany: BLV-Verl.-Ges.

BMELV. (2012). Ergebnisse der Waldzustanderhebung 2011. Berlin, Germany: BMELV.

Bravo, G., & Potvin, L. (1991). Estimating the reliability of continuous measures with Cronbach's alpha or the intraclass correlation coefficient: toward the integration of two traditions. Journal of Clinical Epidemiology, 44(4/5), 381–390.

Bussotti, F., Cozzi, A., Cenni, E., Bettini, D., Sarti, C., & Ferretti, M. (2009). Quality Assurance and measurement errors in monitoring tree crown conditions in Italy. Journal of Environmental Monitoring, 11(4), 769–773.

Castellan, N., & Siegel, S. (1988). Nonparametric statistics for the behavioural sciences. New York: MacGraw Hill Int.

Cochran, W. G. (1977). Sampling techniques (pp. 428). New York: Wiley.

Dammann, I., Herrmann, T., Körver, F., Schröck, H. W., & Ziegler, C. (2001). Dauerbeobachtungsflächen Waldschäden im level II-Programm – Methoden und Ergebnisse der Kronenansprache seit 1983. (pp. 85). Bonn, Germany: BMVEL.

Durrrant, D., & Boswell, R. (2002). Comparison of crown density assessments on trees within the stand and on ridge edges within the forest. Forest Ecology and Management, 157, 1–6.

Eichhorn, J., Roskams, P., Ferretti, M., Mues, V., Szepesi, A., & Durrant, D. (2010). Visual assessment of crown condition and damaging agents. In UNECE (Ed.), Manual on methods and criteria for harmonized sampling, assessment, monitoring and analysis of the effects of air pollution on forests (pp. 49). Hamburg, Germany: UNECE, ICP Forests Programme Co-ordinating Centre.

Ferretti, M. (1997). Forest health assessment and monitoring - Issues for consideration. Environmental Monitoring and Assessment, 48, 45–72.

Ferretti, M. (1998). Potential and limitation of visual indices of tree condition. Chemosphere, 36(4–5), 1031–1036.

Ferretti, M., Bussotti, F., Cenni, E., & Cozzi, A. (1999). Implementation of quality assurance procedures in the Italian programs of forest condition monitoring. Water, Air, and Soil Pollution, 116, 371–376.

Gertner, G., & Köhl, M. (1995). Correlated observer errors and their effects on survey estimates of needle-leaf loss. Forest Science, 41(4), 758–776.

Ghosh, S., Innes, J. L., & Hoffmann, C. (1995). Observer variation as a source of error in assessments of crown condition through time. Forest Science, 41(2), 235–254.

Innes, J. L. (1988). Forest health surveys: problems in assessing observer objectivity. Canadian Journal of Forest Research, 18, 560–565.

Innes, J. L. (1993). Forest health: its assessment and status. Wallingford: Can International.

Innes, J. L., Landmann, G., & Mettendorf, B. (1993). Consistency of observations of forest tree defoliation in three European countries. Environmental Monitoring and Assessment, 25, 29–40.

Innes, J. L., Böhm, J. P., Bucher, J. B., Dobbertin, M., Jansen, E., Kull, P., et al. (1994). Sanasilva-Bericht 1993: Der Zustand des Schweizer Waldes. Berichte der Eidgenössischen Forschungsanstalt für Wald, Schnee und Landschaft (Vol. 339, pp. 60). Birmensdorf, Switzerland: Eidgenössischen Forschungsanstalt für Wald, Schnee und Landschaft.

Kandler, O., & Innes, J. L. (1995). Air pollution and forest decline in central Europe. Environmental Pollution, 90(2), 171–180.

Kish, L. (1995). Biases and nonsampling errors, in: survey sampling (pp. 509–573). New York: Wiley.

Klap, J. M., Oude Voshaar, J. H., De Vries, W., & Erisman, J. W. (2000). Effects of environmental stress on forest crown condition in Europe. Part IV: statistical analysis of relationships. Water, Air, and Soil Pollution, 119, 387–420.

Köhl, M. (1991). Waldschadensentwicklungen: mögliche Ursachen der Variation der Nadel–Blattverlustschätzung zwischen Beobachtern und Folgerungen für die Kontrollaufnahmen. Allgemeine Forst- und Jagdzeitung, 162(11/12), 210–221.

Köhl, M. (1993). Quantifizierung der Beobachterfehler bei der Nadel-/Blattverlustschätzung. Allgemeine Forst- und Jagdzeitung, 164(5), 83–92.

Köhl, M., Traub, B., & Päivinen, R. (2000). Harmonisation and standardisation in multi-national environmental statistics - mission impossible? Environmental Monitoring and Assessment, 63, 361–380.

Meining, S., Bauer, A., Dammann, I., Gawehn, P., Schröck, H. W., & Wendland, J. (2007). Waldbäume - Bilderserien zur Einschätzung von Kronenverlichtungen bei Waldbäumen. (pp. 128). Kassel, Germany: Verlag M. FASTE.

Mizoue, N. (2002). CROCO: semi-automatic image analysis system for crown condition assessment in forest health monitoring. Journal of Forest Planning, 8(1), 17–24.

Mues, V., & Seidling, W. (2003). Evaluation of the international cross-calibration courses 2001 and 2002. (pp. 78). Hamburg, Germany: Institute for World Forestry, Federal Research Centre for Forestry and Forest Products.

Nakajima, H., Kume, A., Ishida, M., Ohiyama, T., & Mizoue, N. (2011). Evaluation of estimates of crown condition in forest monitoring: comparison between visual estimation and automated crown image analysis. Annals of Forest Science, 68(8), 1333–1340.

R Development Core Team. (2012). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org. Accessed 2 July 2012.
Redfern, D. B., & Boswell, R. C. (2004). Assessment of crown condition in forest trees: comparison of methods, sources of variation and observer bias. Forest Ecology and Management, 188, 149–160.

Schöpfer, W. (1985a). Das Schulungs- und Kontrollsystem der terrestrischen Waldschadensinventuren. AFZ, 40(50), 1353–1357.

Schöpfer, W. (1985b). Zur Genauigkeit terrestrischer Waldschadensinventuren. Der Forst- und Holzwirt, 8, 221–224.

Seidling, W. (2004). Crown condition within integrated evaluations of Level II monitoring data at the German level. European Journal of Forest Research, 123, 63–74.

Solberg, S., & Strand, L. (1999). Crown density assessments, control surveys and reproducibility. Environmental Monitoring and Assessment, 56, 75–86.

UNECE. (2010). Manual on methods and criteria for harmonized sampling, assessment, monitoring and analysis of the effects of air pollution on forests. Hamburg, Germany: UNECE, ICP Forests Programme Co-ordinating Centre.

Wulff, S. (2002). The accuracy of forest damage assessments - experiences from Sweden. Environmental Monitoring and Assessment, 74, 295–309.