4.4 Propagation of Data Uncertainty through Models

Heike Lischke

In the NFI2, several different models were used to assess how far certain forest functions, beyond timber production, were fulfilled. The forest functions included, for example, the protection against natural hazards, the preservation of the biodiversity, and the availability of recreational space. Furthermore, a model for classifying different forest types was used. The employed models contain, as does every model, different types of uncertainties. These uncertainties can be due to underlying assumptions, chosen model structures, values of the model's parameter, and only imprecisely given input values (see as an example the research of uncertainty analysis of models conducted by JOHNSON 1987 and PAHL-WOSTL *et al.* 1997). The uncertainties in the models themselves have been discussed in the appropriate chapters. This chapter tries to analyze how uncertainties of the input data can affect the outcome of the models.

These data uncertainties can be caused by systematic errors in measurements (for metric data) or in ratings of attributes (for nominal and ordinal data). Apart from this, continuous data can be randomly scattered around a mean value. Ordinal and nominal data can also have different outcomes for the same attribute when they are assessed repeatedly.

Such uncertainties of the assessed attributes, which enter the model, were determined by assessing the data a second time ("check assessment," "second survey" in Chapter 2.9) on approximately 600 sample plots in addition to the normal assessment of the NFI2 data ("first survey" in Chapter 2.9). This made it possible to determine how often the control team assessed a certain outcome of an attribute when the (first) survey team had decided upon another outcome. The results of this check assessment are presented in contingency tables (Tables 2 and 3 in Chapter 2.9). For most of the attributes, the assessments of the control team varied around the value of the first survey team (i.e., the maximum values in the contingency tables that are along the diagonal). For some of the attributes however (e.g., the attribute "stand structure," Chapter 2.9, Table 3, at the top), two outcomes were frequently mixed up (e.g., "cluster structure" and "multi-layered").

It is not possible to examine the propagation of such asymmetric uncertainties in the input data through models, which consist of many arithmetic and logical operations with the methods discussed in Chapter 4.3.2. The reason for this is because, apart from continuous data, nominal and ordinal data enter these models as well. Furthermore, the models consider many different input variables with even more case differentiations. Hence, any theoretical derivation of the model's uncertainties involves – depending on the uncertainties of the input variables – computing intense analyses and is consequently very time consuming and prone to errors, even when symbolic calculation software is used (e.g., MAPLE). The Latin-hypercube-method (JOHNSON 1987), the calculation of the model's outcome of all possible input variable combinations with the appropriate probabilities, could not be used since some input variables were continuous measurements. Furthermore, some unrealistic combinations of input variables would have occurred. The uncertainties of the results from the model were, therefore, determined using Monte-Carlo simulations (e.g., JOHNSON 1987), which are based on numerous repeated calculations with different values which were produced by a random number generator.

4.4.1 Methods

4.4.1.1 Approach

Figure 1 gives an overview of the approach of the uncertainty analysis. A detailed account can be found in the following sections. Randomly selected data from NFI sample plots were used as one group of input data. They were first assumed to be certain. From the check assessment, the uncertainty distributions of these original input data were determined. Using these distributions, several uncertain input data were generated for each original datum. The original data, as well as the associated uncertain data, were used as input data in the models. By comparing the results of the models that were simulated with the original input data and with the uncertain input data, the uncertainty distributions of the results from the models were determined. These uncertainty distributions were then used to study how data uncertainties affect the relative areas of the model results determined in the inventory.

The term “certain” refers to values that are assumed to be without any variability. “Uncertain” values are, therefore, values that include certain variability.

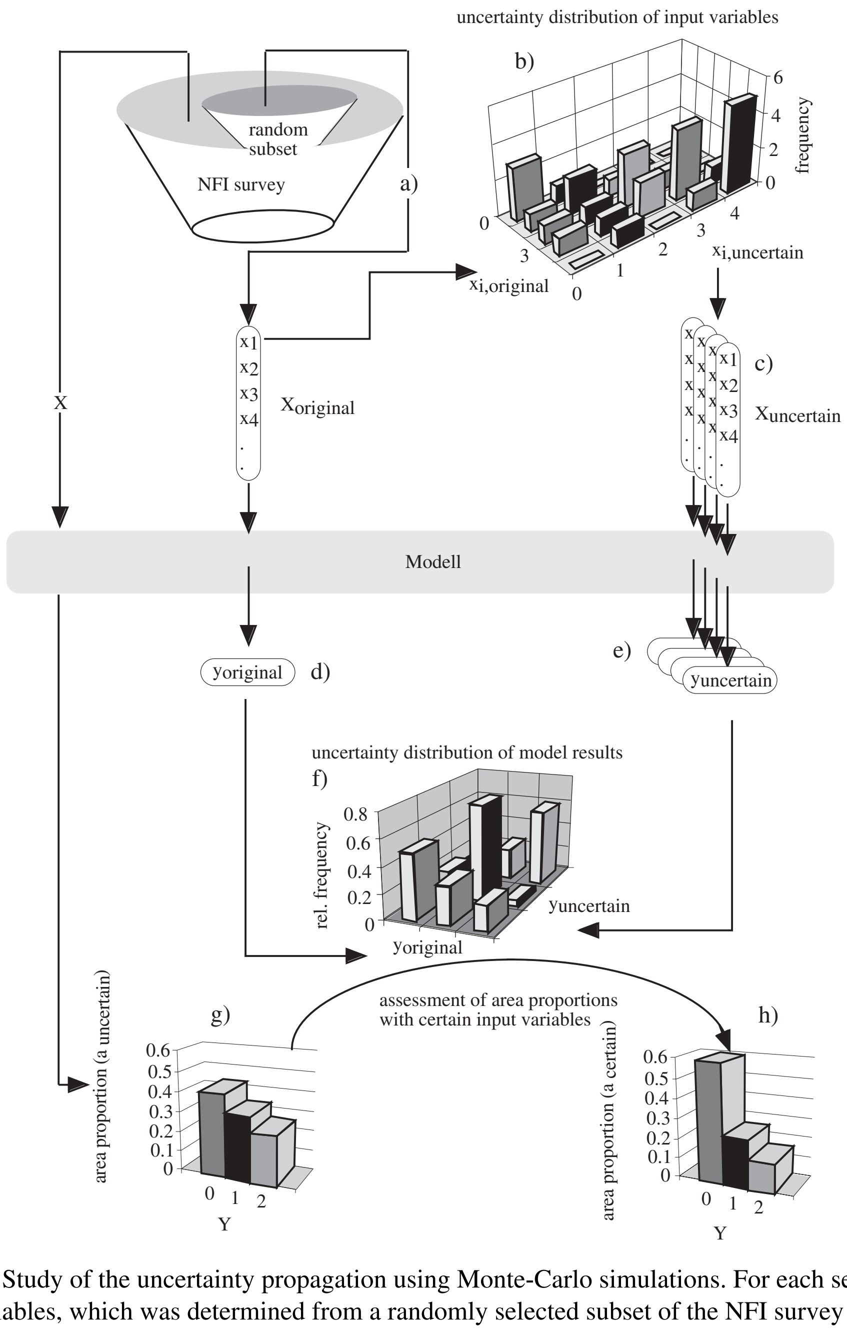


Figure 1. Study of the uncertainty propagation using Monte-Carlo simulations. For each set $X_{original}$ of input variables, which was determined from a randomly selected subset of the NFI survey (a) several sets $X_{uncertain}$ (c) are randomly selected from a given empirical uncertainty distribution (b). The model is used to calculate the results $y_{original}$ (d) and $y_{uncertain}$ (e) with the help of the input data sets $X_{original}$ and $X_{uncertain}$, respectively. By repeating this simulation many times (indicated by the different layers), the uncertainty distribution of the results (f) is determined. Together with the area proportions of the model results determined from all NFI sample plots (g), the uncertainty distribution is then used to assess what the area proportions of the results would have been if the input data had been certain (h).

4.4.1.2 Examined Models

The uncertainty propagation was studied in the models for:

1. The stability of the protection forest (protection forest model, see "Stability Standards in the Protection Forest", Chapter 3.6)
2. The ecological quality of forest areas (biotope model, version BIOLFI2, see Chapter 3.8)
3. The ecological quality of forest edges (ecotone model, version OEKOLFI2, see Chapter 3.8)
4. The recreational quality of forests (natural characteristic model, version ERHNAT3, see Chapter 3.7)
5. The determination of the forest type (EAFV 1988)

The structure of the models is described in the appropriate chapters and literature; the input variables that enter the individual models are presented in Table 1.

Table 1. Models and attributes used in the uncertainty analysis. For a precise explanation, see the model descriptions or documentation of the variables in the appendix. "Fixed" means that no check assessment was available for this attribute. It was therefore not changed. "Uncertain" means that it was changed based on the uncertainty distribution of the check assessments.

Model	Input variable	Definition	Fixed	Uncertain
Protection forest Chapter 3.6	STRUK	Stand structure		x
	VERJDG	Closure of regeneration		x
	SCHLUSSG	Crown closure		x
	KROLAE	Crown length		x
	KROFRM	Shape of crown		x
	BHD	Diameter at breast height		x
	BHOHD	Derived tree height		x
	PROB1-3	Probability of VUNIT1-3	x	
	VUNIT1-3	Most probable PNV	x	
Forest type according to (EAFV 1988)	STRUK	Stand structure		x
	EST	Stage of development		x
	WTYP	Type of forest		x
	WFRM	Origin and management type of forest		x
	NUTZKAT	Utilization category	x	
Biotope rating Chapter 3.8	STRUK	Structure		x
	EST	Stage of development		x
	SCHLUSSG	Crown closure		x
	BHDGT50	% trees with DBH > 50	x	
	BSTSGRAD	Degree of damage	x	
	WARA	Forest edge present?	x	
	BESTGRE	Stand edge		x
	LUECKEN	Type of gap		x
	STRADG	Closure of shrub species		x
	BEERDG	Closure of berries		x
	STOECKE	Stumps		x
	DUERRSTA	Standing dead trees		x
	AHAUFEN	Heaps of branches		x
	BWNATURN	Biotope rating closeness to nature	x	
	BWARTEN	Biotope rating species	x	
Ecotone value Chapter 3.8	WRARTEN	Species at forest edge	x	
	AUFBAU	Type of forest edge (vertical)		x
	MANTELBR	Width of forest edge		x
	STRABR	Width of shrub belt at forest edge		x
	KRAUTBR	Width of herbals belt at forest edge		x
	VERLAUF	Type of forest edge (horizontal)		x
	DICHTE	Density of forest edge		x
	WRUMG	Surrounding of forest edge		x
Natural characteristics Chapter 3.7	BODENVEGDG	Closure of ground vegetation		x
	STRUK	Stand structure (vertical layers)		x
	EST	Stage of development		x
	STRADG	Closure of shrub species		x
	WRUMG	Surrounding of forest edge		x
	LUECKEN	Type of gaps		x
	ERHARTEN	Species important for recreation	x	

4.4.1.3 Original Input Values

The original input values (i.e., the values of the attribute x_i that are assumed to be known as certain) were first generated by a random number generator. It was assumed that the attributes were uniformly distributed over the range of all possible NFI values and that they were independent. However, this led to many unrealistic attribute combinations and to biased result distributions. Thus, in the final simulations, the original input values were determined from a randomly selected subset of NFI samples, so that the resulting combinations were realistic.

4.4.1.4 Uncertainties of the Input Variables

The uncertainty distributions of the input variables of the models (see Table 1) were derived from the control study that was conducted between 1993 and 1995 (see Chapter 2.9). During the study, some of the sample plots were assessed twice. The results were presented in contingency tables (Chapter 2.9, Table 2 and 3) which indicated how often the control team decided upon the value $x_{i,j}$ for the attribute x_i when the survey team had chosen $x_{i,k}$.

In this study, the term "uncertainty" is used for such deviations instead of the term "error." The reason for this is because the check assessment is not an assessment of the error in the strict sense, since it is not possible to determine the deviation of the survey from a fixed true value. The contingency tables reflect the variability of the assessment between teams. The true variability must be assumed to be slightly lower than the variability of the contingency tables, since the latter combines the variability of the survey team with the one from the control team. This should be taken into account for the evaluation of the results.

Neither the uncertainty distributions of the input variables nor those of the model results changed significantly if the marginal distributions of the columns in the contingency tables were used instead of the marginal distributions of the rows (e.g., Chapter 2.9, Table 3). In other words, if the survey from the control team was used as a reference value instead of the survey from the first team. The final study was conducted using the marginal distributions of the rows.

An example for the uncertainty distribution of discrete attributes (see Table 1) is given in Figure 2. The distributions for the stand structure and regeneration coverage assessed by the control team are plotted as frequency distributions of the differences to each of the possible outcomes assessed by the survey team.

The uncertainties of the continuous DBH measurement proved to be symmetric in the 1993 check assessment. Thus, it was assumed that the DBH followed a normal distribution with mean 0 and standard deviation (measurement error) 0.7 cm, as determined by the check assessment.

The tree height (BHOHD, see variable documentation, Chapter 6.1) was estimated with an empirical model that used the DBH (Chapter 3.2). Its (random) model error (i.e., the standard deviation of possible results from the model around the mean model function, which is estimated with a regression analysis) amounts to 3.8 m. This model was calibrated with the measured height against the measured DBH. The measured height, in itself, contains a random error (standard deviation) of 1.5 m. This results in an overall error (standard deviation) of

$$\sqrt{3.8^2 + 1.5^2} \cong 4.$$

The most probable potential natural vegetation unit (VUNIT1, see variable documentation, Chapter 6.1), an input variable of the protection forest model, was determined using the potential natural vegetation (PNV) models of BRZEZIECKI, KIENAST and WILDI (see also Chapter. 3.1, BRZEZIECKI *et al.* 1993; 1995; KIENAST *et al.* 1994; KIENAST *et al.* 1996). This model provides for each sample point, depending on the site conditions that exist there, the three most probable forest communities together with the probability of occurrence.

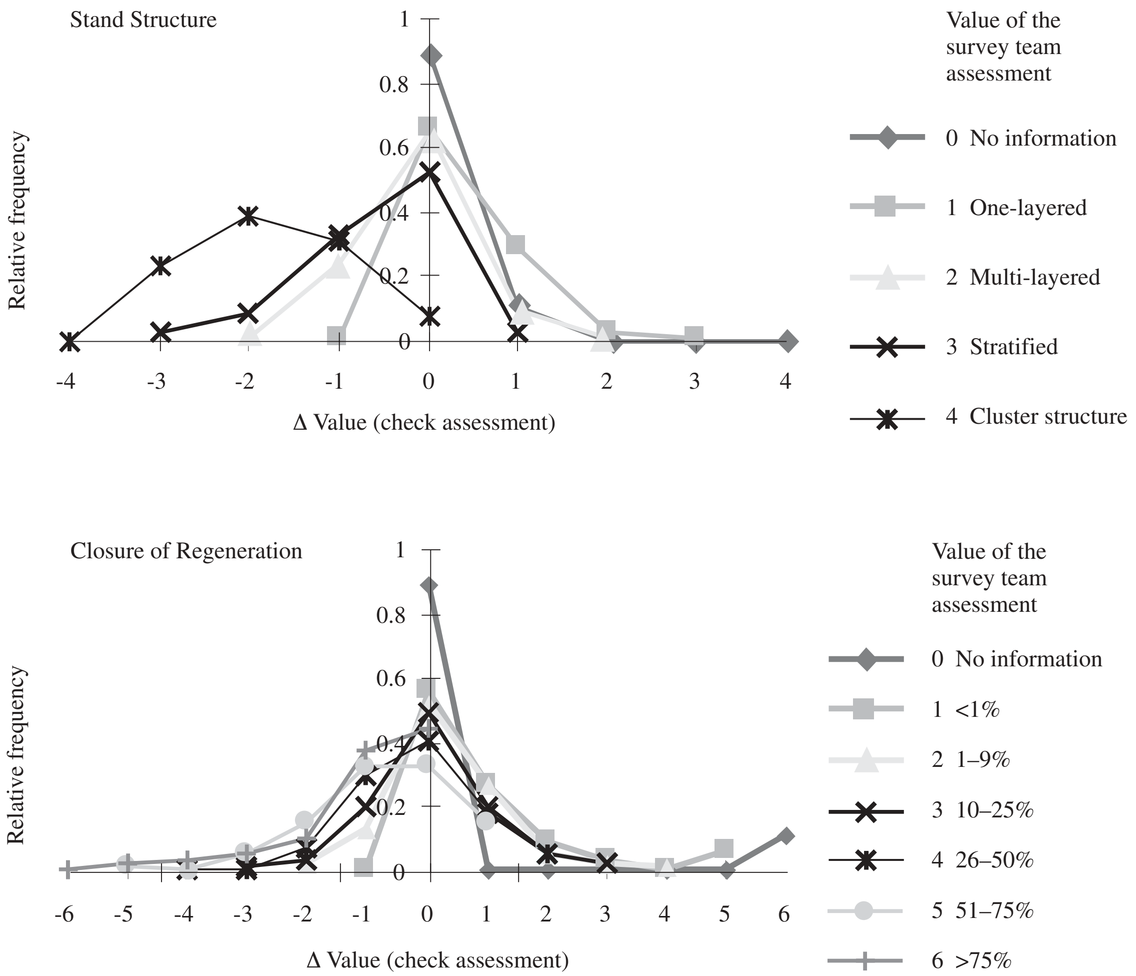


Figure 2. Uncertainty distribution of the input variables from the check assessment, using the example of stand structure and closure of regeneration. The relative frequencies of the differences between the assessment of the control team and the survey teams are presented. Each line represents an attribute value selected by the survey team. For example, if the survey team assessed the stand structure as 4 (cluster structure), the control team assessed, in 40% of the cases, the same stand with the decision 4-2 = 2 (multi-layered).

This model also contains uncertainties which are, for the most part, due to the insufficient quality and quantity of the underlying data in some regions. In order to estimate the influence of such uncertainties of the simulated PNV on the protection forest simulation, the probability $P_{vunitMod}$ was introduced for the confidence of the PNV model results. With this probability $P_{vunitMod}$, one of the three most probable forest communities was selected. Within these three forest communities the probabilities calculated by the PNV model were used for the selection. With the probability $1\text{-}P_{vunitMod}$, however, *any* forest community was randomly selected. During a sensitivity analysis, the $P_{vunitMod}$ varied between 0.5 and 1.

For variables where no check assessments were conducted, the values assessed in the NFI were used (i.e., they were assumed to be certain) ("fixed" in Table 1).

4.4.1.5 Monte-Carlo Simulation

The four examined models were programmed in Modula-2. From 600 randomly selected NFI sample plots, the outcome was determined and used as the input variable. For each set of "original values" $X_{original}$, 100 additional sets with "uncertain values" $X_{uncertain}$ were selected with the help of a random number generator from the empirical distribution of the check assessment.

The random selection from an empirically determined discrete distribution of the attribute *m* with the probability density *f(m)* and the cumulative distribution function *F(m)* is illustrated in Figure 3. Using a random number generator, a value *u* is selected from a [0,1] uniform distribution (in the example 0.55) and determined in which interval of the cumulative distribution function *F(m) u* falls. The corresponding *m*-value (in example 2) is then selected.

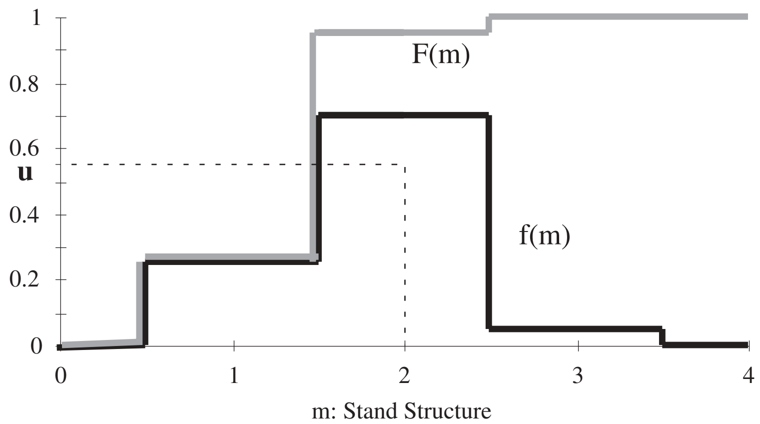


Figure 3. Random selection from a discrete, empirical distribution *p*. The example here is the empirical distribution f(stand structure) when the original value of the stand structure is 2 (multi-layered) (see Figure 2). The cumulative distribution function *F* is calculated from the empirically determined density function *p*. Using a random number generator, a number *u* is randomly selected from a uniform [0,1] distribution (here: 0.55). The next step is to determine in which interval of *F u* falls. The stand structure associated with this interval (here: 2) is selected.

The models were applied using the original ($X_{original}$) input variables and the uncertain ($X_{uncertain}$) input variables associated with them. This led to original result values and to 100 uncertain result values $y_{uncertain}$ for each $y_{original}$. It was then determined how frequently each $y_{uncertain}$ was obtained for a certain $y_{original}$ (see example in Table 2).

At the end of the simulation, the relative frequencies of the uncertain results per original result value were determined. This resulted in the matrix *U* of the uncertainty distribution (Table 3). The relative frequencies in the rows of *U* indicate for a certain original value how frequently the same or other (uncertain) values resulted. In this form they can be used for a-priori uncertainty estimations; in other words, they indicate how a (hypothetical) true result, which is based on certain input data, would be distorted by the uncertainty of the input data.

Table 2. Absolute frequencies of the model results from the Monte-Carlo simulation with the example of the protection forest model. Presented are the proportions of fulfilled criteria for a sufficient stability of the protection forest, which were calculated with original data and with data determined with the uncertainty distribution.

$y_{unsicher}$ / $y_{original}$	alle	≥2/3, <1	>1/3, <2/3	≤1/3	Σ
alle	92	59	19	1	**171**
≥2/3, <1	42	118	54	1	**215**
>1/3, <2/3	19	71	112	5	**207**
≤1/3	1	1	3	2	**7**
Σ	**154**	**248**	**189**	**9**	**600**

Table 3. Uncertainty distributions (matrix *U*) for the model's results, referring to results calculated with the original data for the example of Table 2.

$y_{unsicher}$ / $y_{original}$	alle	≥2/3, <1	>1/3, <2/3	≤1/3	∑
alle	0.538	0.345	0.111	0.006	**1**
≥2/3, <1	0.195	0.549	0.251	0.005	**1**
>1/3, <2/3	0.092	0.343	0.541	0.024	**1**
≤1/3	0.143	0.143	0.429	0.286	**1**
∑	**0.968**	**1.380**	**1.332**	**0.320**	**4**

Table 4. *A-posteriori* probability distribution (Matrix U^{-1}) for the model results of the example in Table 2 (see equation 2).

	alle	≥2/3, <1	>1/3, <2/3	≤1/3	∑
alle	2.445	-1.740	0.346	-0.051	**1**
≥2/3, <1	-0.977	3.283	-1.390	0.084	**1**
>1/3, <2/3	0.255	-1.877	2.837	-0.215	**1**
≤1/3	-1.116	2.045	-3.734	3.805	**1**
∑	**0.607**	**1.710**	**-1.941**	**3.624**	**4**

4.4.1.6 Uncertainties of the Area Proportions

The analysis of the normally assessed NFI data that used different models showed the area in which a certain result category was obtained; separately for each larger region and for all of Switzerland. According to the protection forest model, approximately 38% of the area in protection forest regions (Chapter 3.6) meet all of the criteria for sufficient stability of the protection forest; 27% meet between one and two-thirds of the criteria; and only 2% meet less than one-third of the criteria.

These numbers seem to contradict the column margins in Table 2. However, it is important to note that the 600 NFI sample plots, which were used to generate the original data for this analysis, originate from all over Switzerland, while the analysis of the protection forest model only refers to those regions in which the forest fulfilled a certain protective function. Despite this, the uncertainty matrices that were determined using these data are still generally valid, since they only contain the proportion of the uncertain results (i.e., the column margins are not important).

Under the unrealistic assumption that these area proportions are the result of the model's analysis using "certain" input variables, it is possible to find out, with the help of the uncertainty distribution matrix *U* determined in Chapter 4.4.1.5, how these uncertainties affect the area proportions in the result categories. In order to accomplish this, the areas associated with the result categories (vector $a_{certain}$) are newly distributed according to the matrix *U* into the category.

$$a_{uncertain} = \begin{pmatrix} a_{uncertain,1} \\ \vdots \\ a_{uncertain,n} \end{pmatrix} = \begin{pmatrix} u_{1,1} & \cdots & u_{1,n} \\ \vdots & \ddots & \vdots \\ u_{n,1} & \cdots & u_{n,n} \end{pmatrix} \cdot \begin{pmatrix} a_{certain,1} \\ \vdots \\ a_{certain,n} \end{pmatrix} = U \cdot a_{certain} \quad (1)$$

Example 1: In the example of the protection forest with *U* from Table 3, 38% of the area with the category "all stability criteria fulfilled" is newly distributed with 0.345*38%=13.11% in the category "more than 2/3 of the stability criteria fulfilled;" with 0.111*38%=4.218% in the category "between 1/3 and 2/3 of of the stability criteria fulfilled;" and with 0.006*38%=0.228% in the category "less than 1/3 of the stability criteria fulfilled." Only 0.538*38%=20.44% remain in the category "all stability criteria fulfilled."
The new area proportions result then by newly distributing the area proportions of all categories and adding them up.

The NFI surveys are subject to the variability of the survey teams; thus, the input variable and the results from the models, as well as their area proportions, must be considered uncertain (i.e., $a_{uncertain}$ denotes the uncertain, but known, results of the survey). Therefore, it is of interest to estimate in retrospect (*a posteriori*), how the results and their area proportions $a_{certain}$ would have been if the results had been certain (i.e., if the survey teams would have assessed the attributes in the same way). For this (1) is solved for $a_{certain}$:

$$\begin{aligned} & U \cdot a_{certain} = a_{uncertain} \\ \Leftrightarrow\ & a_{certain} = U^{-1} \cdot a_{uncertain} \end{aligned} \tag{2}$$

For the example of the protection forest model U^{-1}, the inverse matrix of U is given in Table 4.

The uncertainty distribution of the results from the models and the *a-posteriori* area proportions calculated with (2) are presented in the following.

4.4.2 Results

The results of the uncertainty analysis are given in Figures 4 through 11. The first figure for each model shows the uncertainty distribution of the results from the models as the difference to the results of the original values.

> Example 2: If in Figure 6, the biotope rating of original data takes on the value 4 (category "high," a curve with solid squares), the biotope ratings for 60 % of the uncertain data also have the value 4, i.e., Δbiotope rating = 0; in 40% of the cases they have the value 3 ("tends to be high"), i.e., Δ biotope rating = –1. Values of 2 and 1 (category "tends to be low" and "low") do not exist at all, i.e., Δ biotope ration = –2 and –3.

The second figure shows the assessed area proportions together with the a-posteriori area proportions.

Table 7 shows for all models the mean value of the attribute category that was determined with the area proportion – with and without variability.

4.4.2.1 Protection Forest Model

Figure 4 describes the uncertainty distributions of the proportions, in which the stability criteria of the protection forest are fulfilled, for three different values of P_{vunitMod} (i.e., the confidence of the PNV model). The difference between the three simulations is small, which indicates that the uncertainty of the PNV model has only a very small influence to the overall result. If less than a third of all stability criteria are fulfilled by simulations that have "certain" original values (category 4, curves with solid squares), then the uncertainty distribution is shifted strongly in the direction of the higher proportions (smaller categories). For the other values that were generated with the original data (other curves), the value itself was selected with a probability of about 0.6; whereas, the neighboring values were selected with a probability of 0.2 to 0.4. The area proportion (Figure 5) for "less than two-thirds of the stability criteria fulfilled" increased slightly. The area proportion for "all of the stability criteria fulfilled" strongly increased at the expense of the area proportion, for "more than two-thirds of the stability criteria fulfilled."

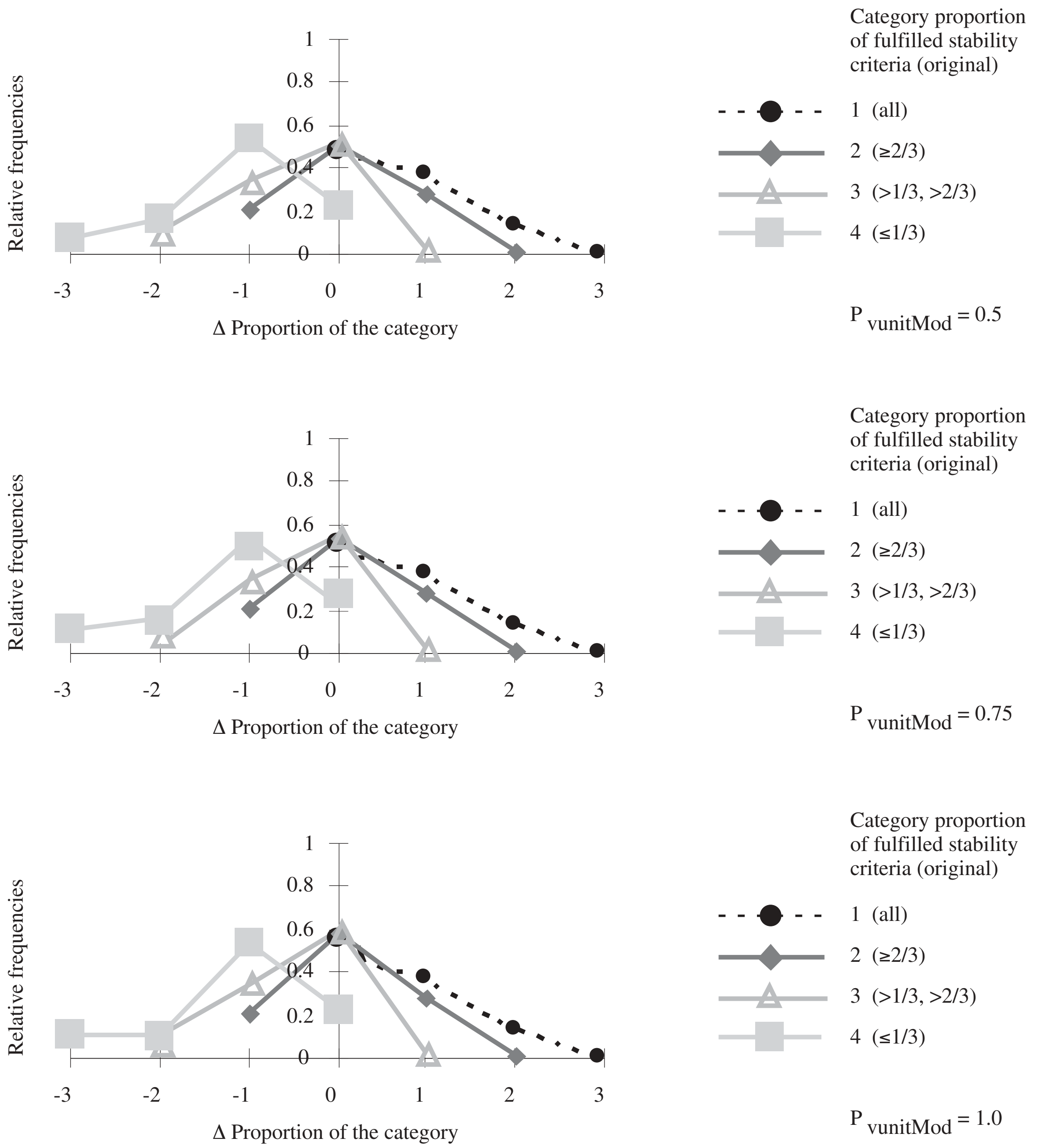


Figure 4. Uncertainty distribution for the proportion of fulfilled stability criteria according to the protection forest model. (The proportions were given in discrete categories). Plotted are the relative frequencies of the deviations between the proportions that were calculated from input variables of randomly selected NFI sample plots (“original”), and the proportions calculated with “uncertain” values of the input variables. The “uncertain” input was determined by selecting from the corresponding uncertainty distribution (e.g., Figure 2, Table 2 and 3 in Chapter 2.9). Each curve represents a proportion that is based on original data. The x-axis represents the differences of the “uncertain” proportions to the original proportions. $P_{vuniModt}$: Confidence of the PNV model.

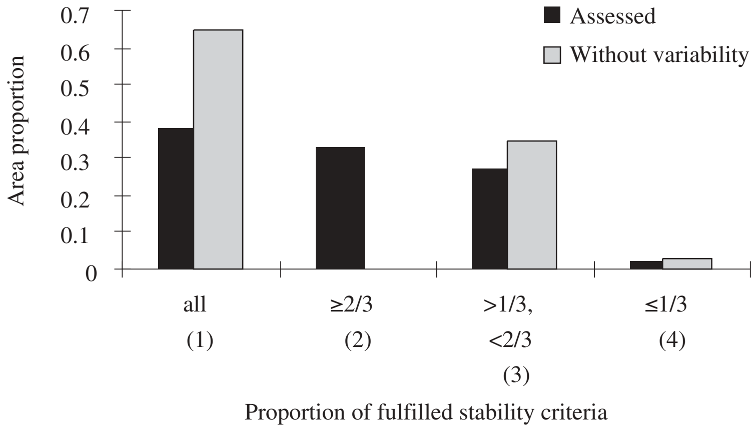


Figure 5. Area proportions for different proportions of fulfilled stability criteria in all of Switzerland. The area proportions were determined (a) with the protection forest model with assessed (i.e., variable, uncertain) data and (b) corrected with the inverse uncertainty distribution. The corrected distribution corresponds to the distribution which would occur if the data did not have any variability, i.e., would have been certain ($P_{vunitMod}$=0.75).

4.4.2.2 Biotope Model

Figure 6 shows a low sensitivity of the biotope rating model with respect to the uncertainty of the input variables, which is also reflected in the small change of the area proportion as seen in Figure 7.

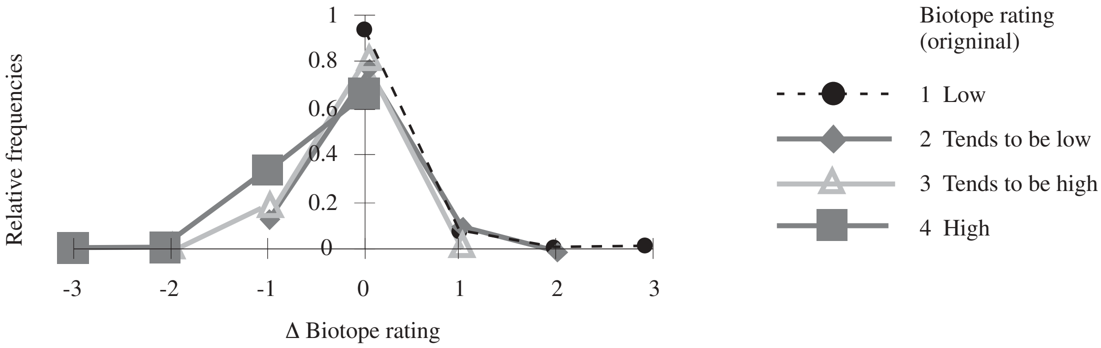


Figure 6. Relative frequencies of the deviations of the biotope rating according to the model BIOLFI2. Plotted are the relative frequencies of the deviations between the biotope ratings that were calculated from input variables of randomly selected NFI sample plots (“original”), and the biotope ratings calculated with “uncertain” values of the input variables. The “uncertain” input was determined by selecting from the corresponding uncertainty distribution (e.g., Table 2 and 3 in Chapter 2.9). Each curve represents a biotope rating that is based on original data. The x-axis represents the differences of the “uncertain” to the “original” biotope rating.

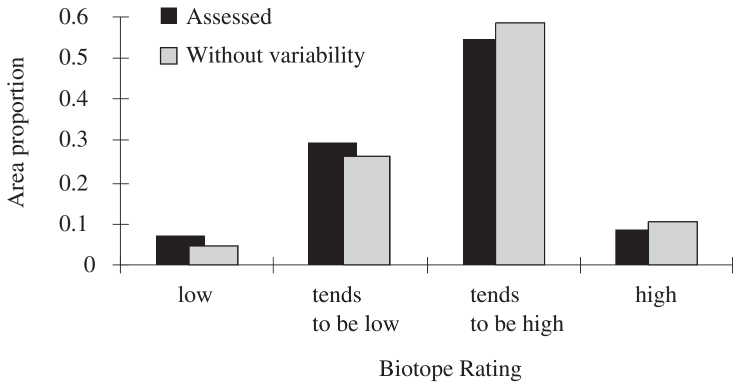


Figure 7. Proportions of the area representing different biotope ratings in all of Switzerland. The area proportions were determined (a) with the biotope rating model with assessed (i.e., variable, uncertain) data and (b) corrected with the inverse uncertainty distribution. The corrected distribution corresponds to the distribution, which would occur if the data did not have any variability, i.e., would have been certain.

4.4.2.3 Ecotone Model

The ecotone model reacted with slightly more sensitivity to the uncertainty of the input variables (Figure 8). The area proportions (Figure 9) were shifted from the lower to the higher ecotone values.

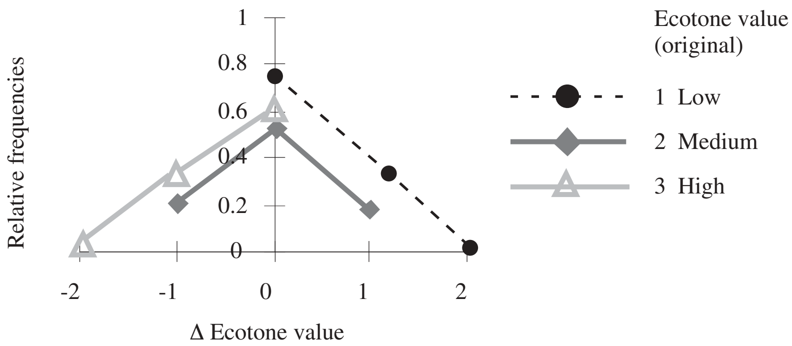
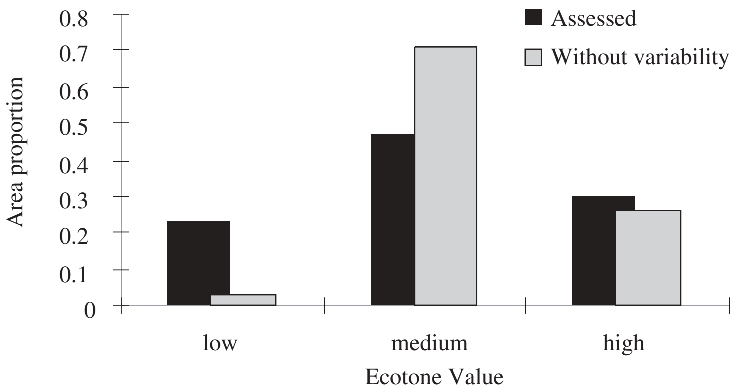


Figure 8. Relative frequencies for the deviations of the ecotone values according to the model OekoLFI2. Plotted are the relative frequencies of the deviations between the ecotone values that were calculated from input variables of randomly selected NFI sample plots ("original"), and the ecotone values calculated with "uncertain" values of the input variables, which were determined by selecting the input variables from the corresponding uncertainty distribution (e.g., Table 2 and 3 in Chapter 2.9). Each curve represents an ecotone value that is based on original data. The x-axis represents the differences of the "uncertain" to the "original" ecotone values.

Figure 9. Proportion of the area representing different ecotone values in all of Switzerland. The area proportions were determined with the ecotone model with assessed (i.e., variable, uncertain) data and (b) corrected with the inverse uncertainty distribution. The corrected distribution corresponds to the distribution, which would occur if the data did not have any variability, i.e., would have been certain.



4.4.2.4 Natural Characteristics

The uncertainty distribution for the natural characteristics' model (Figure 10) was shifted from "low" natural characteristics to "tends to be low" natural characteristics. The area proportion (Figure 11) increased for the categories "high" and "tends to be low" at the expense of the category "low" and "tends to be high."

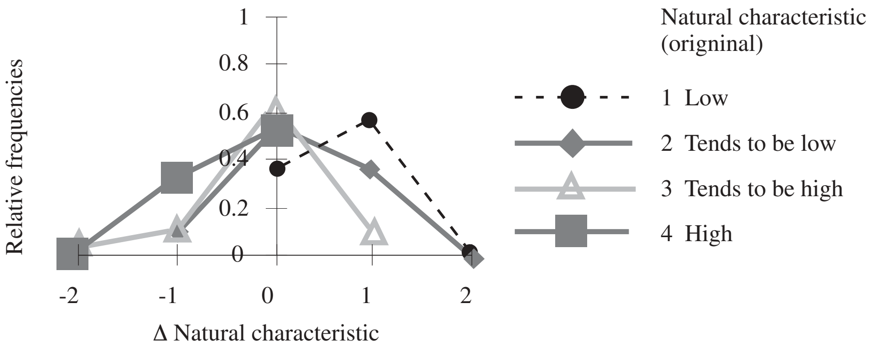


Figure 10. Relative frequency for the deviations of the natural characteristics according to the model NATUERH3. Plotted are the relative frequencies of the deviations between the natural characteristics that were calculated from input variables of randomly selected NFI sample plots (“original”), and the natural characteristics calculated with “uncertain” values of the input variables, which were determined by selecting the input variables from the corresponding uncertainty distribution (e.g., Table 2 and 3 in Chapter 2.9). Each curve represents a natural characteristic that is based on original data. The x-axis represents the differences of the “uncertain” to the “original” natural characteristics.

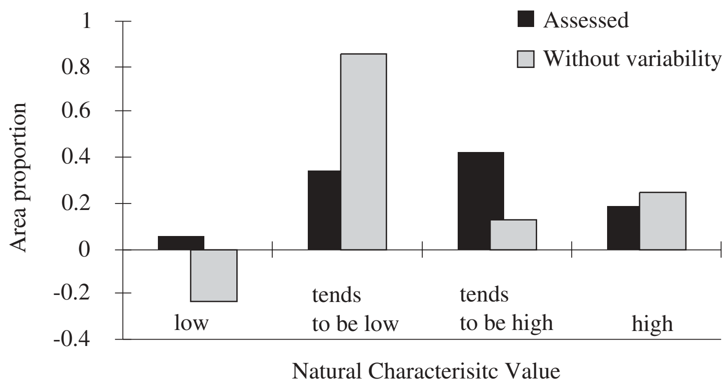


Figure 11. Proportion of the area representing different natural characteristics in all of Switzerland (a) with assessed (i.e., variable, uncertain) data and (b) corrected with the inverse uncertainty distribution. The corrected distribution corresponds to the distribution, which would occur if the data did not have any variability, i.e., would have been certain.

4.4.2.5 Forest Type Model (According to Report 305)

Table 5 demonstrates that, due to the uncertainties in the input variables, some systematic uncertainties in the determination of the forest types arise. Some of the forest types (namely 6, 10, 11, 14, and 16) are frequently classified as medium timber (type 15). Plantations are not clearly recognized and are wrongly identified as uniform high forests. The types coppice forest and coppice with standards are being confused or are classified as pole wood. Plenter type high forests have a high chance of being classified as irregular high forests or as young or medium timber. The classification of young, medium, and old timber is not well defined. In the area proportions (Figure 12) the uncertainties affect the transition from pole wood to timber in particular. For certain input data, area proportions of pole wood would decrease at the expense of the young and medium timber. Especially drastic is the influence of the uncertainties for irregular high forests. If “certain” input data would be used, the area proportion of this forest would completely change to the plenter high forest.

Table 5. Relative proportion of forest types determined from the uncertain input data (rows) per forest type determined from the original input data (columns). Proportions over 0.3 (in this example only on the diagonal) have a black background; proportions between 0.2 and 0.3 are framed bold; and proportions between 0.1 and 0.2 are framed. The codes are defined in Table 6.

Unsicher Original	0	3	5	6	7	8	9	10	11	12	13	14	15	16	17
0	0.90	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.01	0.00	0.01	0.01	0.05
3	0.01	0.99	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
5	0.01	0.00	0.99	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
6	0.01	0.00	0.00	0.48	0.00	0.01	0.01	0.04	0.08	0.03	0.07	0.05	0.12	0.09	0.01
7	0.01	0.00	0.00	0.04	0.51	0.00	0.00	0.02	0.19	0.02	0.04	0.06	0.06	0.05	0.01
8	0.00	0.00	0.00	0.04	0.00	0.40	**0.21**	0.01	0.02	0.03	0.25	0.02	0.00	0.01	0.01
9	0.01	0.00	0.00	0.03	0.00	0.19	0.38	0.01	0.07	0.02	0.15	0.07	0.04	0.02	0.00
10	0.01	0.00	0.00	0.03	0.01	0.00	0.01	**0.25**	**0.20**	0.03	0.08	0.11	0.16	0.09	0.03
11	0.02	0.00	0.00	0.03	0.01	0.00	0.01	0.05	0.42	0.04	0.07	0.09	0.17	0.09	0.01
12	0.02	0.00	0.00	0.03	0.01	0.00	0.01	0.01	0.06	0.73	0.04	0.02	0.02	0.02	0.05
13	0.02	0.00	0.00	0.03	0.01	0.00	0.01	0.01	0.05	0.03	0.74	0.06	0.01	0.01	0.02
14	0.02	0.00	0.00	0.03	0.01	0.00	0.01	0.01	0.06	0.02	0.17	0.50	0.15	0.02	0.01
15	0.02	0.00	0.00	0.03	0.01	0.00	0.01	0.02	0.11	0.00	0.01	0.12	0.53	0.14	0.01
16	0.01	0.00	0.00	0.03	0.01	0.00	0.01	0.01	0.05	0.04	0.01	0.00	0.19	0.63	0.01
17	0.01	0.00	0.00	0.03	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.02	0.04	0.01	0.85

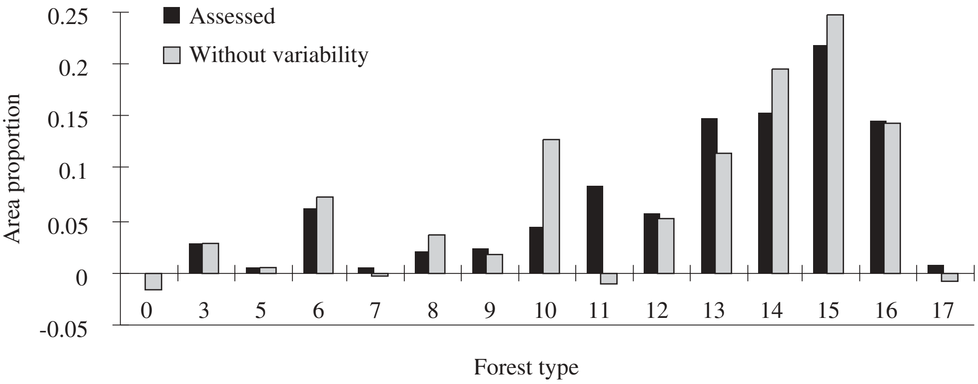


Figure 12. Proportion of the area representing the types of forest determined with (a) assessed data and (b) corrected with the inverse uncertainty distribution. The corrected distribution corresponds to the distribution which would occur if the data did not have any variability , i.e. would have been certain. Definition of the codes, see Table 6.

Table 6. Code for the forest types

0	Missing	10	Plenter high forest
3	Permanently unstocked	11	Irregular high forest
4	Temporary unstocked	12	Young growth/thicket
5	Ride and slopes	13	Pole wood
6	Permanently open	14	Young timber
7	Chestnut and other plantations	15	Medium timber
8	Coppice forest	16	Old timber
9	Coppice forest with standards	17	Incomplete description

Table 7. Swiss mean result values of models. They were determined with the area proportions based on (a) assessed data and (b) corrected with the inverse uncertainty distribution (see Figures 5,7,9,1, and 12).

	Category proportion of fulfilled stability criteria	Biotope rating	Ecotone value	Natural characteristics	Type of forest
Assessed	1.93	2.64	2.07	2.75	12.74
Without variability	1.76	2.74	2.24	2.84	12.84

4.4.3 Discussion

This study investigated how uncertainties in the input variables of several different models employed in the NFI affect the outcome of these models. The study is based on the contingency tables of the check assessment. These do not provide a distribution of errors around the true value, since the qualitative assessments do not present an objective true reference value. The distributions reflect only the dispersion of the subjective assessments. However, this is not very important for the quality of the models, as long as the same mean subjective assessment enters the model. This is true for the investigated models, since the developers of the models participated in the survey, the training, and the check assessment, and consequently were very familiar with the survey.

The contingency tables can only present the variability of the same attributes within the survey. They combine the variability of the first team and the control team. The distribution contains, as a consequence, a much higher dispersion than the actual uncertainty distribution of the survey teams. It is to be expected that the uncertainty distributions of the results from the models also have a lower dispersion; therefore, the shift of area proportions would be less. For example, negative area proportions should not occur any further if the actual uncertainty distribution can be used.

On the other hand, not all uncertainty distributions of all input variables were known. In addition, the variables presently assumed to be certain also contain a particular variability, which in turn should have led to an increase in the result variability.

4.4.3.1 Consequences

The examined models show a significant, but not aggravated sensitivity towards the data uncertainties at the magnitude they occurred. The uncertainty distribution of the result values suggests, at first glance, that the results from the models are not very reliable. This is certainly true for individual sample plots. The goal of a sampling inventory, however, is not to provide information about individual plots, but rather to present information about larger units. A relatively small deviation between the area proportion as well as the mean values (Table 7) indicate that the deviations compensate each other.

Due to the uncertainties of the input data, the models for the qualitative evaluation of the forest functions, namely the models for the protection forest, biotopes, ecotones, and natural characteristics, estimate the output variables slightly too negatively. The proportion of fulfilled stability criteria, as well as biotope and ecotone values seem, on average, to be too small. This is reflected in the average values determined from the area proportions (Table 7). Thus, the statements given in the result volume publication for the second NFI should be considered as conservative.

The misclassification of the forest types (Table 5 and Figure 12), in particular plenter forests and irregular forests, could in some cases be of importance, since some analyses in the NFI use this attribute as a stratifying attribute.

4.4.3.2 Outlook

The software developed in this study for the uncertainty analysis using Monte-Carlo simulations provides an instrument which allows efficient uncertainty analyses of other models that use the data from current and future inventories.

In order to estimate model uncertainties comprehensively, check assessments with other important attributes should be conducted during the next inventory. Especially desirable is to establish the error distribution. This could be achieved if the check assessment would be conducted by a particularly well-trained expert group, which would ideally be involved in the analysis later on. Another possibility consists in multiple surveys (more than ten surveys) of a subset of all samples. The most frequently selected outcome of an attribute would be considered the true outcome.

The sensitivity of the model towards data uncertainties, as they were shown by this study, leads one to consider that the models also react sensitively to uncertainties within the structure and the models' parameters (e.g., the interval limits within the decision trees). It is advisable to assess the influence of uncertainties with respects to model structures and parameter values, and for the classification of the result values using a sensitivity analysis (see for example BUGMANN 1994; LISCHKE 1992), if these models are used further or new models of this type are developed. The method introduced here provides the necessary instruments for this as well.

4.4.4 Literature

Brzeziecki, B.; Kienast, F.; Wildi, O. 1993: A simulated map of the potential natural forest vegetation of Switzerland. J. Veg. Sci. 4 (4):499–508.

Brzeziecki, B.; Kienast, F.; Wildi, O. 1995: Modelling potential impacts of climate change on the spatial distribution of zonal forest communities in Switzerland. J. Veg. Sci. 6 (2):257–268.

Bugmann, H. 1994. On the ecology of mountainous forests in a changing climate: A simulation study. Ph.D. thesis, Dissertation, ETH.

EAFV. 1988. Schweizerisches Landesforstinventar. Ergebnisse der Erstaufnahme 1982–1986. Vol. 305, Ber. Eidgenöss. Forsch.anst. Wald Schnee Landsch. Birmensdorf: Eidgenöss. Anst. forst. Vers.wes. in Zusammenarbeit mit dem Bundesamt für Forstwesen und Landschaftsschutz. 375 p.

Johnson, M.E. 1987. Multivariate statistical simulation, Wiley Series in probability and mathematical statistics. New York a.o.: John Wiley & Sons. 230 p.

Kienast, F.; Brzeziecki, B.; Wildi, O. 1994: Computergestützte Simulation der räumlichen Verbreitung naturnaher Waldgesellschaften in der Schweiz. Schweiz. Z. Forstwes. 145 (4): 293–309.

Kienast, F.; Brzeziecki, B.; Wildi, O. 1996: Long-term adaptation potential of Central European mountain forests to climate change: a GIS-assisted sensitivity assessment. For. Ecol. Manage. 80:133–153.

Lischke, H. 1992: A model to simulate the population dynamics of the codling moth (Cydia pomonella L.(Lepidoptera, Torticidae)): Parameter Estimation and sensitivity analysis. Acta Hortic. 313: 331–338.

Pahl-Wostl, C.; Jaeger, C. C.; Rayner, S.; Schaer, C.; Van Asselt, M.; Imboden, D. M.; Vckovski, A. 1997: Regional integrated assessment of climate change and the problem of indeteminacy. In: Cebon, P., Dahinden, U., Davies, H., Imboden, D., Jaeger, C. (Editor): A view from the Alps: Regional perspectives on climate change. Boston: MIT Press. 435–492.