A Predictive Uncertainty Model for Field-Based Survey Maps Using Generalized Linear Models

Stefan Leyk and Niklaus E. Zimmermann

Swiss Federal Research Institute WSL, Zuercherstrasse 111,
CH-8903 Birmensdorf, Switzerland
{stefan.leyk,zimmerma}@wsl.ch

Abstract. In this paper we present an approach for predictive uncertainty modeling in field-based survey maps using Generalized Linear Models (GLM). Frequently, *inherent* uncertainty, especially in historical maps, makes the interpretation of objects very difficult. Such maps are of great value, but usually only few reference data are available. Consequently, the process of map interpretation could be greatly improved by the knowledge of uncertainty and its variation in space. To predict *inherent* uncertainty in forest cover information of the Swiss topographic map series from the 19th century we formulate rules from several predictors. These are topography-dependent variables and distance measures from old road networks. It is hypothesized that these rules best describe the errors of historical field work and hence the mapping quality. The uncertainty measure, the dependent variable, was derived from local map comparisons within moving windows of different sizes using a local community map as a reference map. The derivation of local *Kappa coefficient* and *percent correctly classified* from these enlarged sample plots takes the local distortion of the map into account. This allows an objective and spatially oriented comparison. Models fitted with uncertainty measures from 100m windows best described the relationship to the explanatory variables. A significant prediction potential for local uncertainty could thus be observed. The explained deviance by the Kappa-based model reached more than 40 percent. The correlation between predictions by the model and independent observations was ρ=0.76. Consequently, an improvement of the model to 47 percent, indicated by the G-value, was calculated. The model allows the spatial-oriented prediction of *inherent* uncertainty within different regions of comparable conditions. The integration of more study areas will result in more general rules for objective evaluation of the entire topographic map. The method can be applied for the evaluation of any field-based map which is used for subsequent applications such as land cover change assessments.

1 Introduction

Historical spatial information frequently represents essential input for landscape change analysis in geographical information systems (GISs). In this context, historical maps are unique witnesses of past landscape configurations. They provide histori-

M.J. Egenhofer, C. Freksa, and H.J. Miller (Eds.): GIScience 2004, LNCS 3234, pp. 191–205, 2004.

cal information on a landscape before aerial photography. The potential of historical documents has recently been recognized by various authors from different fields such as landscape management [5], landscape reconstruction, ecology or forestry [21] and GIS [3, 24].

The objects, in particular natural objects, and their delineations represented in such historical documents usually contain uncertainty to a high degree. This can be referred to as *inherent uncertainty* [24] or *production-oriented uncertainty* [19]. This type of uncertainty is complex and contains different concepts, such as *vagueness*, *ambiguity* and *error*. Different perspectives have contributed to these concepts such as philosophy [28], information theory [18], remote sensing [12], and GIS [10, 31].

Inherent uncertainty manifests itself in a spatial deviation from reality for whatever reason. Often in historical maps only this pattern of deviation can be assessed without any additional knowledge. This is due to the fact that detailed historical background information for thematic interpretation is hardly accessible. One way to assess the accuracy of a map with regard to a reference map is to carry out a quantitative map comparison. Thereby traditional techniques of accuracy assessment using error or confusion matrices produce statistical accuracy measures. There is an abundance of literature describing these techniques in detail [8, 11, 26], for a review in land cover classification). Recently these approaches were extended to assess area estimates from generalized confusion matrices [20] or to derive local summary statistics for geo-statistical analysis [4]. Furthermore combinations of confusion matrix approaches and fuzzy set operations [16, 17, 30] were shown to be flexible methods for map comparison if spatial entities have vague definitions or multi-memberships.

The above mentioned techniques allow the evaluation of map accuracies in general. But the spatial distribution of accuracies can only be considered within an area for which a reference map exists. Such combined methods provide no detailed knowledge of the uncertainty distribution over larger areas, which are not entirely covered by the reference map. In many cases the amount of *inherent* uncertainty varies greatly throughout the considered area depending on local conditions [25]. For the incorporation of historical maps into land cover change assessments, such knowledge would greatly improve the evaluation process. It would support the informed interpretation of the historically mapped objects. To enable this, the knowledge of the *inherent uncertainty* and its variation in space needs to be extrapolated to the whole historical map. This knowledge can be derived from a limited number of local reference maps. Conceptually, this can be done if rules or models can be developed, which link the local uncertainty to a set of generally available predictors. These linkages can be extrapolated into space or to areas not covered by the reference maps.

We present the development of a predictive model for mapping the *inherent* uncertainty of the forest cover of a historical national topographic map of Switzerland originating from the 19th century. These maps were surveyed in extensive field campaigns. Thus we hypothesize that the errors of such mapping efforts can be related to a set of topography-dependent predictors. These are slope, aspect, elevation, as well as distance and visibility from old road networks. The chosen explanatory variables represent factors that are assumed to have been limiting factors for field based surveys, such as visibility or accessibility. By doing so, we aim at finding more general rules, which best describe the quality and the associated errors of topographic field mapping during the 19th century.We use Generalized Linear Models (GLMs) and moving windows to explain the uncertainty found at moving window positions from a set of terrain-derived variables.

GLMs are mathematical extensions of ordinary least-square regression models that do not force data into unnatural scales. Thereby they allow for non-linearity and non-constant variances [22]. This is an important prerequisite, since our dependent variables rate the amount of error in a moving window at a scale ranging from 0 (=perfect fit) to 1 (no agreement at all). Thus the dependent variable is bounded, and cannot easily be treated with ordinary least-square regression. GLMs specifically address this problem and represent thus a suitable approach for our study. These statistical methods are successfully applied to different fields such as predictive habitat modeling in ecology [2, 14, 15], or forest inventories based on remote sensing [23].

2 Materials and Methods

2.1 Study Area and Historical Maps

The historical Swiss National topographic map evaluated here is the first edition of the so-called Siegfried map series. It was published between 1870 and 1920 at scales of 1:25,000 for the Swiss Plateau and 1:50,000 for the mountain regions. As a new feature compared to earlier maps the forest cover was delineated during the field survey. Thus this map is the only source representing forest cover throughout the entire area of Switzerland at this time.

Our study area is the community of Pontresina (Canton Grisons), which is located in an interior valley of the eastern Swiss Alps (Figure 1). As reference data were scarce, the choice of the area was limited. We found only a few areas for which reference data exist. The available reference map for the community of Pontresina belongs to a map series that preceded the official cadastral mapping in Switzerland at that time. They were now mapped as community maps, produced at a scale of 1:5,000 or 1:10,000 with a high degree of detail. These plans are the most reliable spatial representations of landscape objects for the time considered and the instructions for data survey were much more detailed (including forest). Thus we used the forest cover of these maps as a reference to test against the forest cover information of the Siegfried map. Yet there may be an unknown disagreement in the thematic meaning of *forest* leading to a certain spatial effect of uncertainty. In order to exclude additional sources of error, we chose community maps where the date of publication or production did not differ more than 5 years from those of the Siegfried maps.

Both maps were scanned and geo-referenced on the basis of the present-day Swiss national topographic map. Local distortions made the correct registration of the Siegfried map rather difficult. The resulting RMSE using the affine transformation procedure for geo-rectifying was 9.2m. The forest cover was extracted from both datasets to be treated in vector or raster format. The spatial resolution of the raster form of the maps was 1.25m.

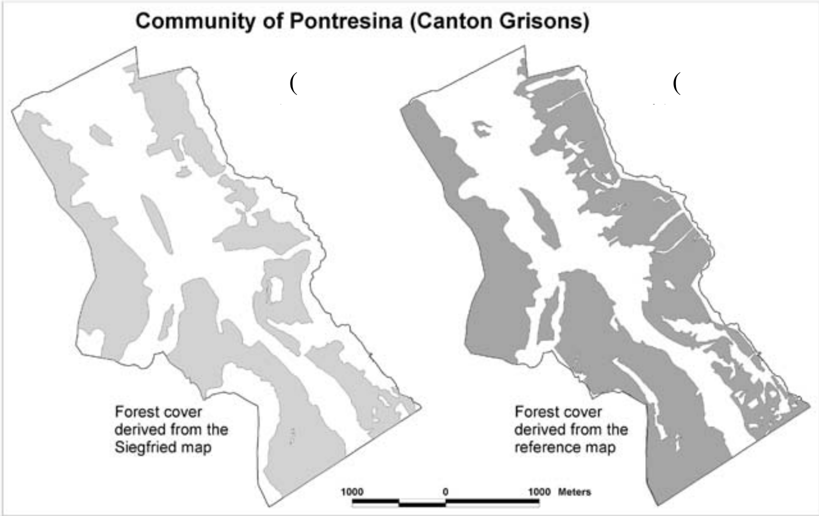


Fig. 1. The forest cover information from (b) the local community map of the study area (Pontresina) was used as a reference for map comparison with (a) the forest cover from the Siegfried map

2.2 Uncertainty Measures: The Dependent Variables

We first derived global measures, such as *percent correctly classified* (PCC) and the *Kappa coefficient* (κ) (based on [11]), from the confusion matrix by pixel-wise comparisons. To examine whether there are trends of uncertainty in relation to topographic gradients, such as slope or elevation, we derived the same measures within different strata (i.e., classes of slope). We found that uncertainty increased along these gradients, which indicated an existing relationship. The measures had to be weighted by the proportion of forest boundary per forest area unit within each stratum since uncertainty became apparent along these lines. This was necessary, because in strata where uncertainty was increasing, larger forest patches (with smaller proportions of forest boundary) occurred.

We had two options to generate a spatially distributed measure of uncertainty: (1) a pixel by pixel evaluation resulting in binary codes of 0 (both maps agree), or 1 (the two maps disagree in representing forest/non-forest), and (2) an aggregated evaluation indicating the degree of uncertainty within a moving window as illustrated in figure 2 (with continuous values between 0 to 1). We decided to use the second approach. Thus we calculated PCC and κ within rectangular moving windows of four different dimensions: 30x30m, 60x60m, 100x100m, and 180x180m. Uncertainty was then defined as 1-PCC and 1-κ. By this, the locally generated statistics provide gradual measures of local disagreement between the two maps (Figure 2).

This allowed us to test the power of variables expressing the topographic configuration in explaining the degrees of uncertainty and to test how the explanatory power varies with spatial aggregation.

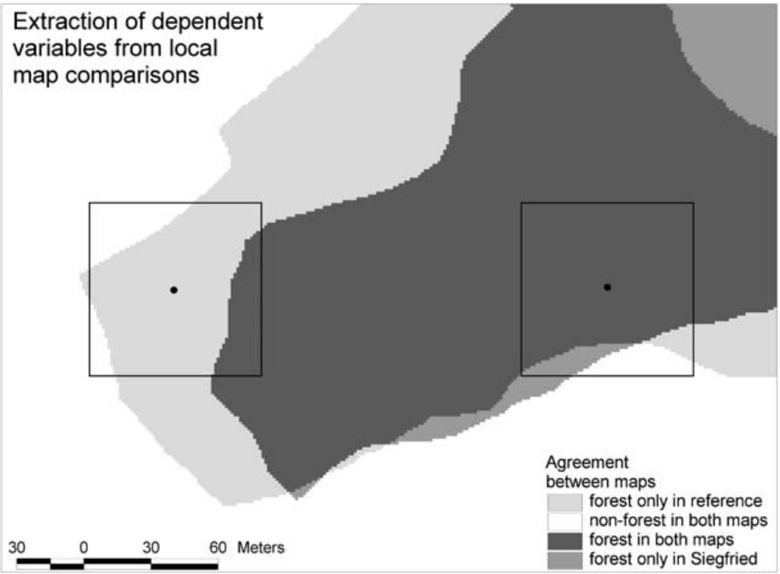


Fig. 2. The dependent variables were locally extracted from pixel-wise map comparisons within moving windows

The equations we used to derive PCC and κ from a confusion matrix of q classes occurring in both the Siegfried and the reference maps have the forms (Eqn. 1 and 2):

$$PCC = P_C = \sum_{i=1}^{q} p_{ii} \qquad (1)$$

and

$$\kappa = \frac{P_c - P_e}{1 - P_e} = \frac{P_c - \sum_{i=1}^{q} p_{i+} p_{+i}}{1 - \sum_{i=1}^{q} p_{i+} p_{+i}}, \qquad (2)$$

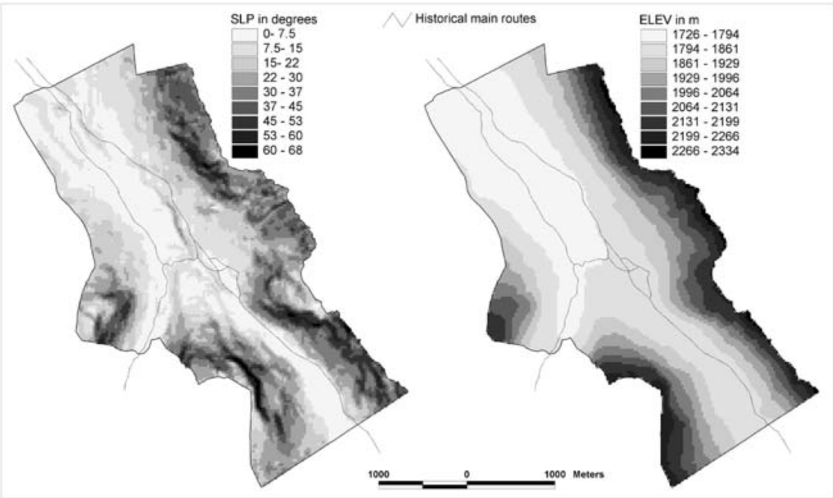
with
$$p_{i+} = \sum_{k=1}^{q} p_{ik} \qquad p_{+i} = \sum_{k=1}^{q} p_{ki},$$

where q is the number of classes considered for map comparison, p_{ii} is the proportion of class i that is correctly classified, P_e is the expected proportion of agreement due to chance, p_{i+} is the marginal proportion of row i and p_{+i} is the marginal proportion of column i of the confusion matrix.

In order to prevent pseudo-replication and to reduce spatial autocorrelation, we sampled the grids of the dependent variables at regular distances of at least the length of the moving window. The mesh distances of the sample locations and the resulting sample sizes, which were considerably reduced with increasing window size, were: 30m (9946 points), 60m (2481 points), 100m (1109 points) and 180m (277 points) (Table 2).

2.3 Independent Variables from Topographic Conditions

In order to find explanatory variables for the prediction of the occurrence of uncertainty in field-mapped forest cover the situation of topographic field survey more than 100 years ago has to be considered. Those conditions that had a strong influence on the quality of mapping will represent the variables with the highest explanatory potential. We included topographic conditions such as elevation (ELEV) (Figure 3b) and elevation difference from the lowest point within the study area (ELEVD) from the 25m DEM of Switzerland and its derivatives slope (SLP) (Figure 3a) and aspect (ASP).



(a) (b)

Fig. 3. The independent variables were extracted from topographic gradients such as (a) SLP or (b) ELEV and distances to routes or forest boundaries

In order to link the pixel values to the moving windows analyzed, we calculated the standard deviation of these variables within each window position using focal functions of the ArcGrid module. Thus we obtained measures of variability for elevation (ELEVCH), slope (SLPCH) and aspect (ASPCH). A next variable derived was the "pathdistance" from any location to the closest main route (STRD), which was accessible with the required equipment. We hypothesize that the distance of locations along a historical road to the considered location contains high predictive potential. We also included a measure of visibility of a location (VIS) from the positions of the main routes. Further, we derived the distance from any location in the study area to the closest forest boundary of the reference map (FORD(ref)). This variable is assumed to hold predictive power since it indicates the distance to the main source of uncertainty or, in other words, to the location where spatial patterns of uncertainty can be assessed. Since this variable is only available in the test area, the resulting model is rather an *analytical* model. We derived this measure also from the Siegfried map (FORD) in order to enable a *predictive* model for other areas. We expect that the latter will result in slightly lower model accuracies. All topographic information was derived at a resolution of 25m. A summary of all independent variables used is given in table 1. The four sets of sample points of the dependent variable were intersected with each explanatory variable in a GIS.

Table 1. Description of the independent variables used in the analytical and predictive uncertainty model

Variable	Description
ELEV	Elevation from DEM 25
ELEVCH	Standard deviation of elevation within moving window
SLP	Slope (degree) as derivative of the DEM
SLPCH	Standard deviation of slope (degree) within moving window
ASPCH	Standard deviation of aspect (degree) in moving window
STRD	Pathdistance from moving window center to nearest road
VIS	Visibility of locations from street positions
ELEVD	Elevation difference between moving window center and lowest location
FORD	Distance from moving window center to nearest forest boundary in Siegfried
FORD(ref)	Distance from moving window center to nearest forest boundary in the reference map

With the explanatory variables listed in table 1, we intended to reflect factors that may have influenced the generation of uncertainty during the historical field survey. The main routes within accessible zones such as valleys with flat terrain are assumed to be the locations the topographer passed during the survey. With increasing horizontal and vertical difference, with increasing slope and with changes of aspect from these main routes the uncertainty in mapping the landscape objects (including forest cover) is expected to grow considerably. Reasons are the limited accessibility, the reduced visibility and consequently the increased difficulty in identifying landscape objects (e.g., forest borders). These were due to the conditions of field work at the time: limited transportation capabilities and heavy equipment. Examination of the distance from the closest forest boundary (FORD) showed that locations close to the boundary carry a much higher probability of *inherent* uncertainty than those further away. This measure is expected to have significant predictive power. This is due to the fact that visibility and other terrain features mentioned are not relevant if there is no forest at all or if the location is within a very large forest patch, far away from any boundary.

We do not know from which position the cartographers worked, and many cartographers possibly used different mapping approaches. Therefore, we assume that the resulting predictive model retains a significant unexplained variation of the mapped uncertainties. The remaining variation may include several additional aspects. Apart from others these are weather conditions, seasonal differences, subjective preferences and interests of the topographer and his expertise. Furthermore, the quality of the manual map transcription and reproduction may be responsible for unexplained variation.

2.4 The Statistical Model

In the following we present the formulation of the generalized linear model. The predictor variables X_j *(j=1, ..., p)* are combined to produce a linear predictor (*LP*), which is related to the expected mean value ($\mu = E(Y)$) of the response variable Y through a link function *g()*:

$$g(E(Y)) = LP = \alpha + X^T\beta \ , \qquad (3)$$

where α is a constant called the regression intercept, X represents a vector of p predictor variables ($X_1,\ldots,X_p$) of any possible power T, and β denotes a vector of p regression coefficients ($\beta_1,\ldots, \beta_p$) which are determined for each predictor.

Prior to selecting an appropriate model, the empirical probability distribution of the response variable has to be tested and compared to the theoretical distribution. Our dependent variable follows a binomial distribution since the original measures were given in presence/absence terms of errors from a pixel-by-pixel evaluation. With increasing window size these measures are increasingly weighted due to their local environment and will be included as weighted response variables in the model.

The link function we thus used was the logistic link often termed logit regression model [22]. This link can be described as:

$$g(E(Y)) = \log(\mu/(1-\mu)) \ , \qquad (4)$$

which is the logarithm of odds, a model widely used for binomial data [7]. In GLMs the linear combination of predictive or explanatory variables is related to the mean of the response variable through this link function to transform them to linearity and to maintain the prediction values within the range of coherent values for the response [15]. Thus the general logistic regression model we used has the form:

$$\text{logit} = \log(\mu/(1-\mu)) = \alpha + X^T\beta \ . \qquad (5)$$

We included linear terms as well as quadratic powers and interactions of the explanatory variables. A maximum-likelihood (ML) estimator is used for fitting the model. We plotted the standardized residuals against the fitted values to identify unexpected patterns in the deviance. All calculations were performed using Splus.

We first fitted a full model, using all explanatory variables (linear and quadratic terms and interactions), then we applied stepwise regression–allowing for both backward and forward selection–in order to optimize the final model. For the analysis of the significance of eliminating or adding terms, the Akaike information criterion (AIC) was used. Thereby, given a fitted model object, individual terms of the model are removed and the respective effect is assessed in comparison to the previous model where the aim is to minimize the AIC [14].

We used χ^2 approximations [22] and AIC measures to test the deviance reduction associated with each variable for significance at a given confidence level (0.05). In addition we tested whether the model coefficients differ significantly from zero using the standard error associated with the estimated model coefficients for a Student t-test. We used the D^2 value (percent deviance explained) to evaluate the model fit, calculated as (*Null.Deviance* – *Residual.Deviance*) / *Null.Deviance*, where the *Null.Deviance* is the deviance of the model with the intercept only, and the *Residual.Deviance* is the deviance that remains unexplained after all final variables have been included. This is equivalent to the R^2 of a linear least-squares regression model. In addition we used the *adjusted* D^2 (derived from *adjusted* R^2 [29]) to take into account the number of observations (n) used for fitting the model and the number of explanatory variables (p):

$$D^2_{adj.} = 1-((n-1)/(n-p))\times(1-D^2)\,. \tag{6}$$

To produce an uncertainty distribution map the fitted model has to be cartographically represented. For implementation of the model in GIS the inverse of the link function has to be applied to transform the values back to the scale of the original response variable. In our case the inverse logistic transformation is required, which has the form:

$$p(y) = \exp(LP)/(1+\exp(LP))\,, \tag{7}$$

where *LP* is the linear predictor fitted by logistic regression (Eq. 4). In applying this retransformation we obtained values between 0 and 1, which is the same as the original response values. This inverse relationship has bias, which can be corrected by using a Taylor series approximation. However, this correction is rarely done in environmental applications. In cases where the mean of the estimated variable is within the range which can be considered fairly linear after transformation (between 0.2 and 0.8) the bias is expected to be very low. Nevertheless, the bias correction has to be done if this condition is not fulfilled.

In order to test the predictive power of the models we split the data set into two parts following the split-sample approach [27]. One part (50%) was used for calibration (training data), the other one (50%) to evaluate the model predictions and to measure the adequacy of the model (test data). Goodness-of-fit measures can then be used to evaluate the fit between the predictions and observations of the evaluation data set. We calculated the non-parametric rank correlation coefficient of Spearman (ρ Rho) and tested it for significance. Furthermore the G-value [1] was used to test the relative improvement of the model over a null model (i.e., the mean of the dependent variable within the calibration data subset).

First, we developed *analytical* models for PCC- and κ-based uncertainty in which the variable FORD(ref) was included as an independent variable in order to examine its explanatory power. Next, we calibrated *predictive* models applicable to areas other than the test site. Thus we replaced FORD(ref) by FORD for both PCC- and κ –based *predictive* models. All models were developed from split-sample data sets.

3 Results

In our study area, the *analytical* models from the 100m windows best describe the relationship between *inherent* uncertainty and the explanatory variables (Figure 4) when compared with the test data. At smaller window sizes, we did not find equally good explanations from the predictors. The model based on the 160m window data set is also inferior (see Table 2 for a comparison of the results). We therefore only present detailed results for the *predictive* models for this 100m window data set, comparing these models with the *analytical* models, respectively.

Table 2. Comparison of PCC-based predictive models calibrated from different window sizes

Window size for PCC extraction	pixel-wise	30m	60m	100m	180m
mean PCC	0.78	0.78	0.78	0.78	0.78
Sample size	9946	9946	2481	1109	277
Goodness of fit: D^2 (if model is significant)	0.16	0.21	0.25	0.31	0.29

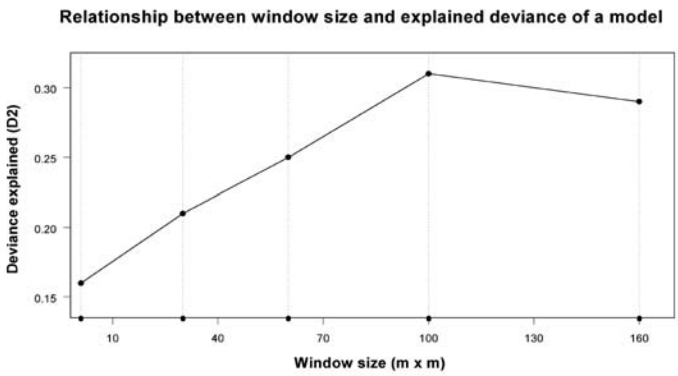


Fig. 4. Relationship between moving window size and deviance explained (D^2) by the fitted model (example for PCC-based models); models were optimized for 100m windows

The *analytical* models for the 100m window (Table 3) indicated significant explanatory power for the uncertainty measures. The step-wise regression resulted in a total deviance reduction of 41 percent for PCC and 46 percent for κ (Table 3, D^2 and *adjusted* D^2). The following variables contributed to the explanation of the fractions of uncertainty in the presented examples: FORD(ref), SLPCH, STRD for PCC-based uncertainty and FORD(ref), SLPCH, ELEVCH, ELEVD and STRD for κ-based uncertainty. When compared with the test data, we received a high agreement for predicted uncertainty (PCC: ρ=0.76, κ: ρ=0.77) within the test site, and the G-values indicated significant improvements over the null model (PCC: G=37%, κ: G=47%, Table 3). Due to the use of distance to the reference forest boundary the results are not applicable to areas outside of the study area.

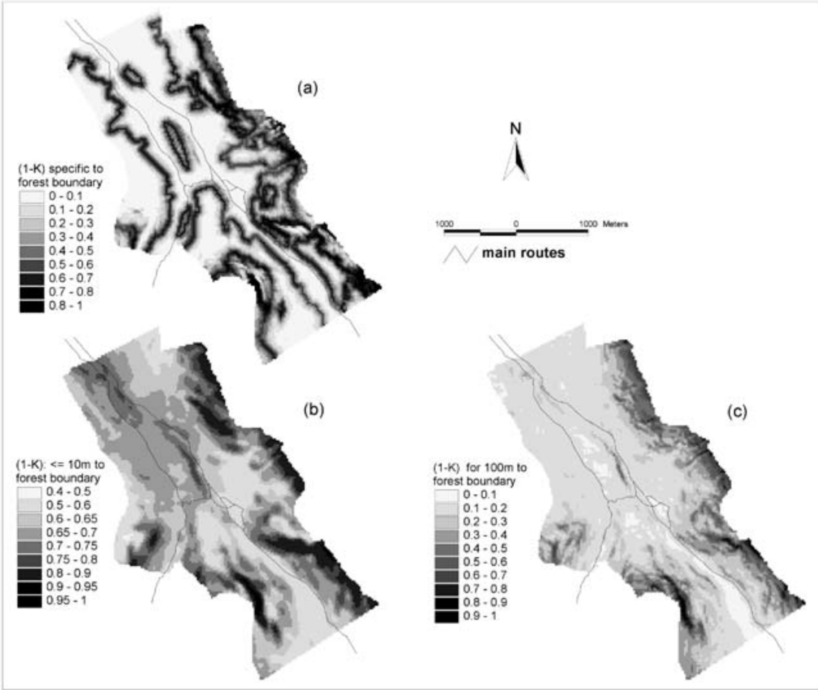


Fig. 5. Predictive uncertainty model for 1-κ as response: (a) site-specific uncertainty, with FORD as predictor, (b) uncertainty predictions with assumed constant values of FORD of 10m, and (c) the same as (b) but for FORD=100m

The calibration of a *predictive* PCC-based model resulted in a slight decrease in model qualities (D^2=0.35, *adjusted* D^2=0.34) and model accuracies (*ρ, G)* when compared to the left-out 554 points. Even though the FORD variable was assumed to be of significantly lower predictive power than the FORD(ref) variable, the *predictive* model still shows similar trends and accuracies (ρ=0.71, and G=0.27). The results of the PCC- *predictive* uncertainty model are presented in Table 3. The following variables contributed significantly to this model: FORD, ELEVD, ELEVCH, SLPCH and STRD.

The κ–based *predictive* uncertainty model revealed a much better model quality D^2=0.41, *adjusted* D^2=0.40), and the model accuracy is almost comparable to the κ-based *analytical* model. Testing the model using the test data resulted in comparably high ρ and G values (ρ=0.75, G=0.47, see Table 3 for an overview). The variables with significant contribution in this case were: FORD, ELEVD, SLP and SLPCH.

Table 3. Results of uncertainty modeling; the *analytical* models (M(pcc(ref)) and M(kap(ref))) and the *predictive* models (M(pcc) and M(kap)) were calibrated on 555 points using uncertainty from 100x100m moving windows, and evaluated on the second subset of 554 points

	Analytical model M(pcc(ref))	Predictive model M(pcc)	Analytical model M(kap(ref))	Predictive model M(kap)
Model quality				
D^2	0.41	0.35	0.46	0.41
Adjusted D^2	0.41	0.34	0.46	0.40
AIC	1359.1	1136.2	1595.9	1442.7
Model test				
Spearman's rank correlation ρ	0.76	0.71	0.77	0.76
G-value	0.37	0.27	0.47	0.47
Model parameters				
Constant	-4.884E-01	-6.210E-01	2.590E-01	8.749E+01
FORD(ref)	-3.368E-02	-	-3.881E-02	-
$(\text{FORD(ref)})^2$	-	-	1.977E-05	-
FORD	-	-1.996E-02	-	-3.157E-02
$(\text{FORD})^2$	-	2.476E-05	-	2.948E-05
ELEVD	-	-3.890E-03	-7.253E-03	-9.375E-03
$(\text{ELEVD})^2$	-	1.150E-05	-	1.521E-05
ELEVCH	-	9.262E-03	-	-
SLP	-	-	-	2.985E-02
SLPCH	1.694E-02	4.195E-03	1.776E-01	8.452E-02
STRD	3.446E-04	6.244E-04	3.715E-03	-
FORD(ref) : SLPCH	1.539E-03	-	-	-
FORD(ref) : ELEVD	-	-	5.974E-05	-
FORD(ref) : STRD	2.373E-05	-	-	-
FORD : SLPCH	-	8.292E-04	-	-
FORD : ELEVD	-	-	-	3.361E-05
ELEVCH : STRD	-	-	-1.295E-04	-
ELEVCH : ELEVD	-	-	2.925E-04	-
SLPCH : STRD	-	-	-1.974E-04	-

As expected, FORD was associated with the most significant deviance reduction in all cases. ELEVD and STRD were highly correlated, which occasionally resulted in a rather random mutual exclusion of one of the two. The same selection behavior was observed for ELEVCH and SLP in all models, since they naturally reflect similar topographic features. ASPFOC and VIS were excluded from all calibration processes since they did not contribute significantly to deviance reduction.

Figure 5 shows the spatial predictions of κ-based uncertainty models. In Figure 5a, we included FORD as predictor. Thus the map represents the spatial uncertainty inherent to this particular map explained by a range of topographic predictors. It is apparent that certain forest boundary zones show less uncertainty than others. In Figures 5b and 5c, predictions are displayed with assumed constant values of FORD (10m and 100m, respectively). By this, the predictions reveal the nature of the modifying effect of the terrain on the forest boundary as mapped in the 19th century and its associated *inherent* uncertainty.

4 Discussion

Overall, the four models calibrated from the 100m sample raster all resulted in models capable of explaining spatial pattern in uncertainty *inherent* in historical and other field based maps. We noticed that the κ-based uncertainty model proved superior both in quality and accuracy when compared to the PCC-based model. This is in agreement with recent research based on confusion matrix evaluations. PCC overestimates the agreement between maps in not accounting for chance agreement [6]. The Kappa coefficient measures the improvement of the *proportion correctly classified* over mere chance agreement [6]. However, κ fails to do so if one class far exceeds the others [9]. PCC showed higher overall agreement across the maps (mean of PCC=0.77) than did κ (mean of κ =0.62). This reflects the over-optimistic map accuracies, thus resulting in generally lower uncertainties and less well spread values across the scale from 0 to 1, compared to κ. This combination of drawbacks for PCC may have caused the difference between the two models.

The inclusion of VIS and ASPCH as explanatory variables did not provide any valuable contribution. The assumptions behind and the extraction processes need to be further tested. All other variables contributed significantly to the models. ELEV and STRD showed comparably high correlations, which is due to the fact that streets climb along the valley. The inclusion of FORD as predictor constrained the uncertainty to the source of error (i.e., the presence of forest boundary, irrespective of terrain features). Figure 5 demonstrates that SLP and SLPCH had high explanatory power since the predicted uncertainty distribution shows a trend along the slope gradient (compare with Figure 3).

We observed single prediction values that clearly deviated from the observed uncertainty in the evaluation subset. This is mainly caused by the many unknown historical factors which resulted in mapping uncertainty. A surveyor may occasionally have added unexpected detail or neglected easy to observe features. Thus, our model only captures the general trends of uncertainty as influenced by topographic features. Despite the rather general approach, the results provide useful insights into the historical mapping process and represent valuable evaluation support for modern use of historical maps. The choice of using window-based summary statistics proved to reflect the disagreement between both maps better than the pixel-wise global comparison. This was the result of changing the window size to optimize the predictive power of the model. One important reason for this may be that windows of 100m are less affected by (sub-pixel) geo-referencing errors. Additionally, the effects of the terrain upon accuracy seem to only be visible at a minimal area considered.

5 Conclusions

When developing the models, we aimed for simplicity and generality. We are aware that, by using additional variables, the model could have been fitted with higher accuracies within the study area. However, such models would be less easily applicable to other areas covered by the Siegfried map. From testing the model against independent data we conclude that the model is applicable to areas of similar topographic nature. Preliminary map comparisons within non-mountainous regions, such as the Swiss Plateau, indicated much higher accuracies, which is what we expect from the fitted models in mountainous terrain. Still, it remains to be tested, how well the existing models perform in non-mountainous terrain.

Historical maps contain valuable landscape information. However, there is an unknown uncertainty inherent in them. To date, only few attempts were made to analyze this uncertainty [21, 24]. In order to make informed use of such documents for land use or land cover change assessment or modeling, it is mandatory, though, to know the quality and inherent uncertainty of these historical data sources. In this paper we have developed a method to predictively map such inherent uncertainty in space.

This method can be applied to any map developed from field surveys under challenging conditions. The topography-related explanatory model variables showed satisfying predictions when the window size was optimized. The spatial predictions of *inherent* uncertainty can thus be used for the evaluation of historical data within different regions by defining fuzzy membership functions and expressing the "possibility of uncertainty" at any given location. Thus the method we presented is very well suited for incorporation into subsequent applications in a larger context to increase the objectivity of the research. We conclude that GLMs represent a very flexible tool for a range of applications. It remains to be tested, whether additional explanatory variables have the potential to improve such models, or if different uncertainty measures such as the NMI [13] are easier to be modeled spatially.

References

1. Agterberg F.P.: Trend surface analysis. In: G.L. Gaile and C.J. Willmott (eds.): Spatial Statistics and Model. Reide, Dordrecht, The Netherlands (1984) 147–171
2. Austin M.P.: Spatial prediction of species distribution: An interface between ecological theory and statistical modeling. Ecological Modelling 157 (2002) 189-207
3. Brown D.G.: Classification and boundary vagueness in mapping pre-settlement forest types. International Journal of Geographical Information Science 12(2) (1998) 105-129
4. Brunsdon C., Fotheringham S. and Charlton M.: Geographically weighted local statistics applied to binary data. In: M.J. Egenhofer and D.M. Mark (eds.): Geographic Information Science, GIScience 2002, Boulder, LNCS 2478. Springer, New York (2002) 38-51
5. Cissel J.H., Swanson F.J., Weisberg P.J.: Landscape management using historical fire regimes: Blue River, Oregon. Ecological Applications 9 (1999) 1217-1231
6. Congalton R.G.: A review of assessing the accuracy of classifications of remotely sensed data. Remote Sensing of Environment 37 (1991) 35-46
7. Dobson A.J.: An Introduction to Generalized Linear Models. Second Edition, Chapman and Hall/CRC, New York, (2002)
8. Fielding A.H.: How should accuracy be measured? In: A. Fielding (ed.): Mashine Learning Methods for Ecological Applications. Kluwer Academic Publishers (1999) 209-223
9. Fielding A.H. and Bell J.F.: A review of methods for the assessment of prediction errors in conservation presence/absence models. Environmental Conservation 24(1) (1997) 38-49

10. Fisher P.: Models of uncertainty in spatial data. In: Longley P., Goodchild M.F., Maguire D. and Rhind D. (eds.): Geographical Information Systems: Principles, Techniques, Management and Applications (1). Wiley & Sons, New York (1999) 191-205
11. Foody G.M.: Status of land cover classification accuracy assessment. Remote Sensing of Environment 80 (2002) 185-201
12. Foody G.M. and Atkinson P.M. (eds.): Uncertainty in Remote Sensing and GIS. Wiley (2002)
13. Forbes A.D.: Classification algorithm evaluation: five performance measures based on confusion matrices. Journal of Clinical Monitoring 11 (1995) 189-206
14. Guisan A., Edwards T.C. and Hastie T.: Generalized linear and generalized additive models in studies of species distribution: setting the scene. Ecological Modelling 157 (2002) 89-100
15. Guisan A. and Zimmermann N.E.: Predictive habitat distribution models in ecology. Ecological Modelling 135 (2000) 147-186
16. Hagen A.: Fuzzy set approach to assessing similarity of categorical maps. International Journal of Geographical Information Science 17(3) (2003) 235-249
17. Jäger G. and Benz U.: Measures of classification accuracy based on fuzzy similarity. IEEE Transactions on GeoScience and Remote Sensing 38 (2000) 1462-1467
18. Klir G.J. and Wierman M.J.: Uncertainty-Based Information. Springer (1998)
19. Leyk S., Boesch R. and Weibel R.: A conceptual framework for uncertainty investigation in map-based land cover change modelling. Transactions in GIS (2004) forthcoming
20. Lewis H.G. and Brown M.: A generalised confusion matrix for assessing area estimates from remotely sensed data. International Journal of Remote Sensing 22 (2001) 3223-3235
21. Manies K.L., Mladenoff D.J. and Nordheim E.V.: Assessing large-scale surveyor variability in the historic forest data of the original U.S. Public Land Surveys. Canadian Journal of Forest Research 31 (2001) 1719-1730
22. McCullagh P. and Nelder J.A.: Generalized Linear Models. Second Edition, Chapman and Hall, London (1989)
23. Moisen G.G., Edwards T.C.: Use of generalized linear models and digital data in a forest inventory of northern Utah. Journal of Agricultural, Biological and Environmental Statistics 4 (1999) 372-390
24. Plewe B.: The nature of uncertainty in historical geographic information. Transactions in GIS 6(4) (2002) 431-456
25. Steele B.M., Winne J.C. and Redmond R.L.: Estimation and mapping of misclassification probabilities for thematic land cover maps. Remote Sensing of Environment 66 (1998) 192-202
26. Stehman S.V. and Czaplewski R.L.: Design and analysis for thematic map accuracy assessment: fundamental principles. Remote Sensing of Environment 64 (1998) 331-344
27. Van Houwelingen J.C. and Le Cessie S.: Predictive value of statistical models. Statistics in Medicine 9 (1990) 1303–1325
28. Varzi A.C.: Vagueness in geography. Philosophy and Geography 4(1) (2001) 49-65
29. Weisberg S.: Applied Linear Regression. Wiley, New York (1980)
30. Woodcock C.E. and Gopal S.: Fuzzy set theory and thematic maps: accuracy assessment and area estimation. International Journal of Geographical Information Science 14(2) (2000) 153-172
31. Zhang J. and Goodchild M.F.: Uncertainty in Geographical Information. Taylor & Francis, London (2002)