COENOSES 5 (1): 37-43, 1990

SIMULATING THE DISTRIBUTION OF PLANT COMMUNITIES IN AN ALPINE LANDSCAPE

H.S. Fischer, Swiss Federal Institute of Technology, Department of Geography, Winterthurerstr. 190, CH-8057 Zurich, Switzerland

Keywords: Alpine vegetation, Distribution, Modelling, Simulation, Switzerland, Vegetation types

Abstract. A probabilistic vegetation model for the simulation of the distribution of vegetation types is described. The performance and limitation of the model are discussed. An application is shown in an alpine landscape in Switzerland.

Introduction

In recent years vegetation mapping gained importance in the documentation of changes in the landscape, in land use management and nature conservation. For the purpose of mapping, vegetation is usually classified into vegetation types. More than half a century ago Braun-Blanquet (1928) proposed that the distribution of vegetation types is widely determined by the distribution of the habitat factors. Only a few attempts have been made hitherto to verify this hypothesis. Box (1981) developed a global model for the distribution of formations based on a parallelepiped classifier (PPD classifier: Mather, 1987). A similar approach was used by Binz and Wildi (1988) to simulate the spatial distribution of plant communities on a regional scale. In Central European phytosociology the habitat of vegetation is described as ranges of the different habitat variables (for an example see Tab. 1 in Goodall and Feoli 1989). The PPD classifier in fact accords with this kind of vegetation description. The performance of the model described in Binz and Wildi (1988) is reasonable only at the level of plant formations in the sense of Ellenberg and Mueller-Dombois (1974). At the finer level of associations, the model widely failed, because PPD classification is a qualitative method, which is appropriate for coarse classifications.

Agriculture, forestry, tourism and nature protection are competing interests of man in the landscape. Growing tourism causes changes within the originally well balanced system. This may result in an increase of natural threats. Therefore, careful planning the development of this area is a major issue to the authorities. An ecological landscape model can help to solve this problem. Our aim is to build up a vegetation model, based on a quantitative description of the vegetation types and their habitat, which can be used for the prediction of changes in the vegetation based on the expected or planned changes in habitat conditions. The model simulates the distribution of associations in the sense of the Braun-Blanquet school at a regional scale as does the model of Binz and Wildi (1988), but uses a quantitative approach instead of the qualitative PPD classification.

Study area

The test site for our model is the same as in Binz and Wildi (1988). It is a 100 km^2 area within the district of Davos in the central Alps of Switzerland. The altitude ranges from 1500 m a.s.l. to 3000 m a.s.l., thus including the subalpine, alpine and nival vegetation belt. A high diversity in habitat conditions is typical to this region, as there are all major rock types of the alps represented (*e.g.*, gneiss, dolomite and serpentine), different types of land use, such as forestry, intensive and extensive meadows and pastures. One part of the area is heavily used by summer and winter tourism while another part is under avalanche danger in winter and is therefore not developed for tourism. A detailed description of the land use history of Davos is given by Günter (1985).

The data

In order to investigate the functional relationship among natural elements and human activities, a project was launched as part of the Man and Biosphere project (Wildi and Ewald 1986, Wildi in this issue). Many different thematical maps are available from the Swiss MaB-project, representing a source of information for a geographical information system (GIS).

A basic variable for the simulation of vegetation types is a digital terrain model (DTM), which has been derived from the topographical map on a 100 m grid with an altitudinal precision of $\pm$10 m. By means of bicubic interpolation, a grid width of 50 m was obtained. The altitude a.s.l. is an index for all altitude dependent climatic variables, *e.g.*, temperature. From the DTM some more variables can be derived such as slope (a component of geomorphology) and radiation (an index for moisture and local climate). The other variables used for the model are: 14 rock types, 28 soil types (Krause, 1986), 6 land use classes, and snow cover recession derived from LANDSAT MSS satellite images (Keller,

1 Paper presented at the 2nd C.E.T.A. International Workshop on Mathematical Community Ecology, November 19-25, 1988, Gorizia, Italy.

1987). A vegetation map distinguishing 63 vegetation types (Zumbühl 1986) represents the reference state to calibrate our model. The classification of the vegetation is based on compositional criteria, but it also considers agricultural yield. All these variables are stored in the GIS in the form of grid data with a grid size of 50 by 50 m.

The method

The vegetation model is part of the landscape simulation model of Davos in the Swiss MaB-project. It is a static and deterministic model. In this model, some variables are independent and invariant, such as relief, geology, macro-climate, snow cover recession, and land use. A cascade of state variables is simulated by connecting these variables (see Wildi for more details on the model in this issue). One of these state variables is vegetation. The vegetation map is simulated by superimposing the maps of habitat conditions (Fig. 1) and analyzing information from them. Using scientific expertise, additional variables such as agricultural yield, or a parameter measuring the value for nature conservation, can be derived from the vegetation map.

Our method for the simulation of vegetation is probabilistic. The vegetation is simulated by calculating the multivariate state conditional probability of occurrence for all 63 vegetation types. The result is a vector of the probabilities for the vegetation types to occur at given habitat conditions. The vegetation type with the highest probability is taken as the simulation result. The calculation of probabilities is based on univariate contingency tables (Tab. 1), sampled from the recorded habitat and vegetation maps. The contingency tables are adjusted column-wise to express estimates of the univariate state conditional probabilies $p(V_i|x)$. These are

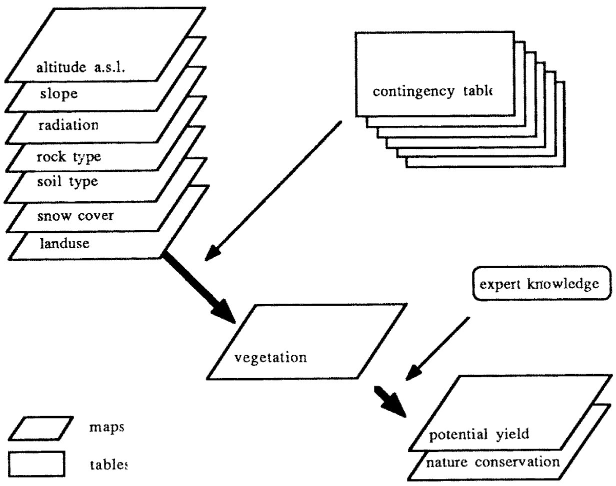

Fig. 1. The vegetation model.

the estimates of the probabilities of vegetation type V_i as site x (Tab. 2). The row-wise adjusted contingency tables contain the estimates of the probability $p(x|V_i)$ of finding site condition x at points with vegetation type V_i. The relativized row totals of the state conditional probabilities ($p^*(V_i)$ in Tab. 2) are an estimate for the prior probabilities of vegetation types, assuming that the values of the habitat variables where equaly distributed.

The multivariate state conditional probability could be estimated with a multivariate contingency table. Using 6 habitat variables with 10 categories each, this would result in a contingency table of exhaustive size including 10^6 cells. A set of approximately 10^7 observations would be needed to get reliable results. Since this is not available, the multivariate state conditional probability $p(V_i|\vec{x})$ is calculated according to the Bayes formula (Duda and Hard 1973):

Table 1. Frequency distribution of the 15 forest communities within 7 altitude classes.

Altitude [m]	1600	1700	1800	1900	2000	2100	2200	2300	$f_{i.}$	$p(V_i)$ [°/oo]
1 Melico-Piceetum	57	158	131	89	26	.	.	.	461	72
2 Veronico latifoliae-Piceetum	1	90	23	.	.	.	.	.	114	18
3 Sphagno-Piceetum calamagr.	143	562	635	501	314	81	.	.	2236	350
4 Sphagno-Piceetum adeno.	40	107	71	76	69	16	.	.	379	59
5 Calamgrostidi variae-Piceetum	.	1	8	20	14	1	.	.	44	7
6 Larici-Piceetum	5	199	530	475	287	154	6	.	1656	259
7 Larici-Pinetum cembrae	.	.	.	11	44	45	2	.	102	16
8 Erico-Pinetum montanae (Dolom)	.	1	13	10	10	.	.	.	34	5
9 Erico-Pinetum montanae (Serp.)	.	73	74	67	22	.	.	.	236	37
10 Sphagno-Pinetum montanae	.	.	.	.	1	.	.	.	1	<1
11 Legfoehren (Silikat)	.	8	83	99	120	82	36	.	428	67
12 Legfoehren (Dolomit)	.	.	5	22	32	26	.	.	85	13
13 Legfoehren (Serpentin)	.	25	63	63	79	87	52	1	370	58
14 Alnetum vir./A.-Cicerbitetum	.	9	44	114	49	12	.	.	228	36
15 Pioneerforest	.	5	3	5	.	.	.	.	13	2
$f_{.k}$	246	1238	1683	1552	1067	504	96	1	6387	
$p(x_k)$ [°/oo]	39	194	264	243	167	79	15	<1		1000

Table 2. State conditional probabilities $p(V_i|x_k)$ [‰].

Altitude [m]	1600	1700	1800	1900	2000	2100	2200	2300	Σ	$p^*(V_i)$
1 Melico-Piceetum	232	128	78	57	24	.	.	.	519	65
2 Veronico latifoliae-Piceetum	4	73	14	.	.	.	.	.	91	11
3 Sphagno-Piceetum calamagr.	581	454	377	323	294	161	.	.	2190	274
4 Sphagno-Piceetum adeno.	163	86	42	49	65	32	.	.	438	55
5 Calamgrostidi variae-Piceetum	.	1	5	13	13	2	.	.	34	4
6 Larici-Piceetum	20	161	315	306	269	306	63	.	1440	180
7 Larici-Pinetum cembrae	.	.	.	7	41	89	21	.	158	20
8 Erico-Pinetum montanae (Dolom)	.	1	8	6	9	.	.	.	24	3
9 Erico-Pinetum montanae (Serp.)	.	59	44	43	21	.	.	.	167	21
10 Sphagno-Pinetum montanae	.	.	.	.	1	.	.	.	1	<1
11 Legfoehren (Silikat)	.	6	49	64	112	163	375	.	769	96
12 Legfoehren (Dolomit)	.	.	3	14	30	52	.	.	99	12
13 Legfoehren (Serpentin)	.	20	37	41	74	173	542	1000	1887	236
14 Alnetum vir./A.-Cicerbitetum	.	7	26	73	46	24	.	.	176	22
15 Pioneerforest	.	4	2	3	.	.	.	.	9	1
Sum	1000	1000	1000	1000	1000	1000	1000	1000	8000	
‰	125	125	125	125	125	125	125	125		1000

$$p(V_i|\vec{x}) = \frac{p(\vec{x}|V_i)\,\hat{p}(V_i)}{p(\vec{x})}$$

where:

$\vec{x}$ is the vector of the habitat variables.

$\hat{p}(V_i) = \frac{f_{i\cdot}}{f_{\cdot\cdot}}$ (estimate of the prior probability of vegetation type V_i; *i.e.*, the relative frequency of vegetation type V_i)

$p(\vec{x})$ = the prior probability for site type $\vec{x}$ to occur (constant at any given point)

Under the assumption that the conditional univariate probabilities $\hat{p}(x^u|V_i)$ are independent, the multivariate state conditional probability can be estimated as:

$$\hat{p}(\vec{x}|V_i) = \prod_{u=1}^{q} \hat{p}(x^u|V_i)$$

$\hat{p}(u^u|V_i) = f^u_{ik}/f^u_{i.}$ (estimate of the conditional probability of site factor x^u at given vegetation type V_i.)

The model can be used to simulate the present vegetation map. In order to use the model for predicting the state of vegetation under different land use strategies or for different areas, the prior probabilities of the different vegetation types have to be estimated. In the univariate case this can be done by adjusting the frequency table to the observed frequency of the habitat variables:

$$g^u_{if} = \frac{f^u_{ik}}{f^u_{.k}\, f^u_{.k}{}^*}$$

where

g^u_{ik} is the adjusted frequency.

$f^u_{.k}{}^*$ is the observed frequency of the kth category of the uth habitat variable in the area to be simulated.

The row totals of the adjusted matrix are proportional to the adjusted prior probabilities. The adjusted prior probability $\hat{P}^u(V_i)$ with respect to habitat variable u is:

$$\hat{P}^u(V_i) = \frac{g^u_{i.}}{g^u_{..}}$$

The geometric mean of these univariate adjusted prior probabilities $\hat{P}^u(V_i)$ can be used as an estimate of the multivariate adjusted prior probability:

$$\hat{P}(V_i) = \sqrt[q]{\prod_{u=1}^{q} \hat{P}^u(V_i)}$$

The assumption for this approach is that the contingency tables are based on independent observations. This does not hold for our data because of the presence of spatial autocorrelation. To solve this problem, the spatial autocorrelograms of the habitat maps are calculated, according to the BW-statistics of Cliff and Ord (1981). The distance is sought at which the autocorrelation is zero. For a distance of about 150 m the autocorrelation for most variables was found to be close to zero, so the contingency tables were calculated with a grid width of 150 m. With this approach, the influence of spatial autocorrelation on the contingency tables is minimized.

A special case in the variable 'land use'. At a certain level of thematic resolution, land use is widely determining vegetation. For example, if a site is mowed, there must be a meadow community on it. If the land use is timber production, there must be a forest community. Based on this land use classification, five submodels were defined, one for each of the land use types. For each of these submodels, the best combination of habitat variables was selected.

Results

The results obtained with the model show 70% of the pixels of the simulated vegetation map agree with the ascertained vegetation map. 90% of the pixels agreed at least with a similar vegetation type. In this context two vegetation types were regarded as similar, if transitions between them can be found in the field. The list of transitions given by Zumbühl (1986) was used to define the similarity. With the second criterion allowance is made for a certain latitude of judgment when mapping vegetation. The comparison matrix in Fig. 2 is a row wise adjusted frequency table comparing the ascertained vegetation map with the simulated one. Dark gray indicates high frequency, white zero values. From bottom left to top right the matrix contains 10 forest types, 16 meadow types, 6 intensively used pastures, 11 extensively used pastures, 13 types of alpine grassland and 6 types of rock vegetation. For almost all vegetation types the highest frequency can be found on the diagonal of the matrix. This illustrates good performance of the model. A few vegetation types are, however, not simulated in a reasonable way. These are

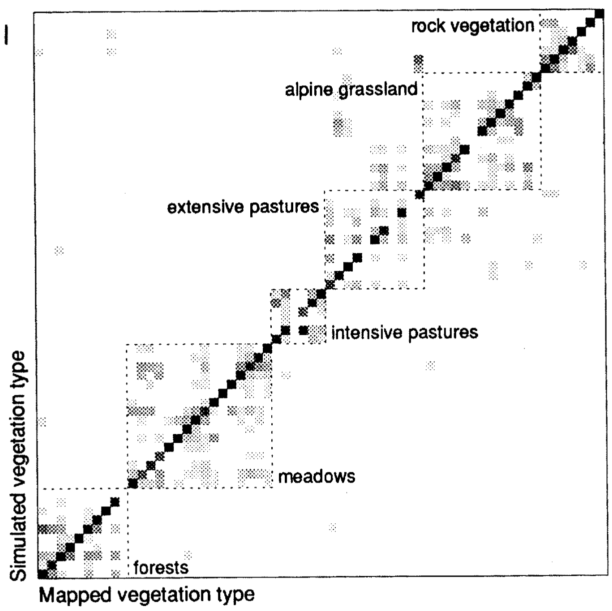

Fig. 2. Comparison matrix. See the details in the text.

types which either occur only in small patches, compared to the spatial resolution of the habitat data, or are extremely rare in the area of investigation.

While checking the performance of the model by comparing the results with the observations in the field during the summer of 1988, I found the following facts limiting the performance of the model:

1) Spatial precision, *i.e.*, the precision with which borders between habitat and vegetation types are recorded on the map, is limited by the capability of the mapper to determine his own position in the field. Unprecise drawing of boundaries causes misclassification.

2) For the purpose of thematic mapping one has to decide what the minimum area unit of an individual vegetation patch should be, in order to be displayed on the map. This spatial resolution is independent of the pixel size in the GIS. It only need to be coarser than the pixel size to enable the storage of the data in the computer. Different spatial resolution of the habitat variables (*e.g.*, vegetation: 1 ha, soil type: 4 ha) causes a limitation for simulating the distribution of relatively small patches of vegetation.

3) Classification of continuous variables (*e.g.*, vegetation and soil) is done independently. Even if there is a perfect deterministic relationship between two variables, there may be a high degree of uncertainty in predicting the class of one variable from the class of the other if the thresholds for the classification do not coincide. Better results could be expected if the classification is based on a canonical, predictive procedure.

4) And finally simple mapping errors contribute to the misclassifications of the model. In these cases, the simulated vegetation may represent the real vegetation type, whereas the recorded reference map is wrong. Or the habitat variables are recorded incorrectly so the model can't simulate the real type.

The model described above is based on categorical variables. Continuous variables are transformed to interval classes and treated in the same way as for example soil types. If a vegetation type does not occur on a certain category, this does not necessarily mean that it is impossible to occur. An example is the richest soil types which are nowhere used for forestry inside the perimeter of the test site today. Nevertheless, forest can grow on these soil types. The model will not be able to simulate the vegetation in this case. From old literature we know that the forest type, which could grow on these rich soils (*Alnus incana* forests), no longer exists in this region at that altitude. Therefore, no reliable prediction can be made in these cases, because no 'training areas' for a model are available.

A great advantage of the categorical approach is that no assumption about a special species or community response model is necessary. The simulation model is based on the observed frequency distribution which can be unimodal, multimodal, symmetric, skewed or of any

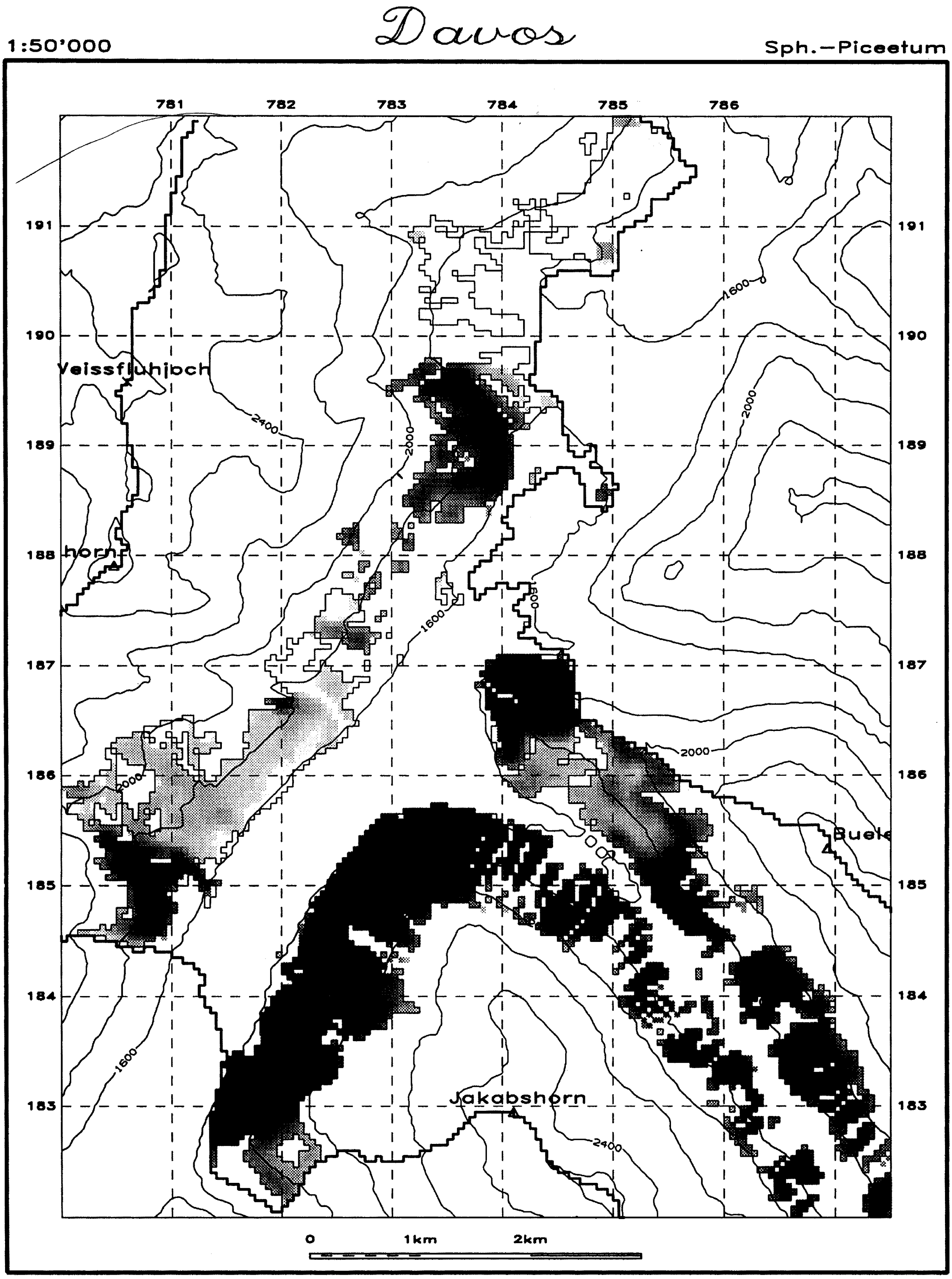

Fig. 3. The probability map of Sphagneto-Piceetum in the Davos site. The probability increases from white space (improbable occurrence) to black space (occurrence with certainty).

other form. The assumption of independence, used in calculating the multivariate state conditional probabilites, seems to be in contradiction to Walter's ''Gesetz der relativen Standortskonstanz'' (law of relative site constancy, Walter *et al.* 1953), which presupposes a dependency of site factors. Two points have to be mentioned in this connection. Walter's law of relative site constancy is concerned with species. Depending on the classification system, vegetation types can have a much smaller range in habitat conditions than species. By means of an appropriate fine classification system, the relations can be approximated with an independent model. Dependence may result from using inadequate site factors. Using geographical latitude and aspect, for instance, will obviously result in a high degree of dependence. A species growing in the cold north at south slopes will probably be found in warmer regions at north slopes. The reason for this is that aspect and latitude do not influence vegetation directly, but are closely related to relevant site factors as moisture and temperature. This illustrates the importance of carefully selecting the variables for the model.

Conclusions

Science is building models. Our idea of the world can be regarded as a mental model, the exact formulation of the mental model as an algorithm. The realization of such an algorithm on a computer is what is called a 'simulation'. So the comparison of the performance of such a model with the observed reality can help one to decide whether the mental model is appropriate. Such a model can, for example, help to answer the question whether the system can be described in static, deterministic terms or if feedback mechanisms and dynamics play a major role.

By omitting one variable after the other in the simulation one can estimate the importance of variables in the model and by extension in the simulated ecosystem. The increase in simulation errors is an index of the importance. The model can be used as a tool for vegetation mapping. Based on vegetation maps from subareas, or point information, a vegetation map can be generated for the whole area, if the information about habitat conditions is available. In the 1988 November issue of the National Geographic there is a computer-generated topographical map of the Mount Everest area (Washburn, 1988). Within the next five years most of the topographical maps of Switzerland will be available in digitalized form. It is reasonable to assume that the same will be true for other thematic maps in the near future. If the external variable 'land use' in the model described above is substituted by multispectral satellite data, the model could be used as a tool for remote sensing vegetation mapping and as an ecological monitoring system. In the alpine region this would require a high resolution DTM for the radiometric correction of the influece of the relief on the video values in the scanner image. The grid size needs to be smaller than the resolution of the satellite scanner in order to calculate aspect and slope on the spatial scale of the satellite image. Unfortunately, this is not available at the moment four our area, but it is expected within the next few years. Consequently we cannot test the performance of the model for this application at present.

For reforestation it is important to know what natural forest type would grow at a certain size. With the help of these models a forest map can be generated even for sites where there is currently no forestry. For the purpose of nature conservation, it may be interesting to know the probability of occurrence of specific vegetation types. With the model presented it is possible to produce such a probability map. Fig. 3 is an example for this application.

Acknowledgements. I would like to thank O. Wildi, F.A. Bemmerlein and C. Renner for valuable discussions and text reviewing.

Glossary of notation

i	subscript referring to vegetation type i [1, 2, ... m]
k	subscript referring to category k of a habitat variable [1, 2, ... n]
u	superscript referring to habitat variable u [1, 2, ... q]
V_i	vegetation type i
x^u	habitat variable u
f^u_{ik}	element of contingency table u expressing the frequency of vegetation type i with the k-th category of habitat variable u.
$f^u_{.k}$, $f^u_{i.}$, $f^u_{..}$	row, column and grand total of the contingency table u
g^u_{ik}	element of the adjusted contingency table u
$p\ (x^u)$	univariate prior probability of the occurrence of habitat condition u
$p(\vec{x})$	multivariate prior probability of the occurrence of habitat condition x.
$p(V_i\|x^u)$	univariate conditional probability of the occurrence of vegetation type i
$p(V_i\|\vec{x})$	multivariate conditional probability of the occurrence of vegetation type i
$\hat{p}\ (...)$	estimate of a probability

REFERENCES

Binz, H.R. and O. Wildi. 1988. Das Simulationsmodell MaB-Davos. Schlussber. Schweiz. MaB-Programm Nr. 33. Bundesamt für Umweltschutz. CH-3003 Bern.

Box, E.O. 1981. *Macroclimate and Plant Forms: An Introduction to Predictive Modelling in Phytogeography.* Junk, The Hague.

Braun-Blanquet, J. 1928. *Pflanzensoziologie. Grundzüge der Vegetationskunde.* Berlin, Wien, New York.

Cliff, A.D. and J.K. Ord. 1981. *Spatial Processes. Models and Applications.* Pion, London.

Duda, R.O. and P.E. Hard. 1973. *Pattern Classification and*

Scene Analysis. New York. Wiley.

ELLENBERG, H. and D. MUELLER-DOMBOIS. 1974. *Aims and Methods of Vegetation Ecology*. Wiley, New York.

GOODALL, D.W. and E. FEOLI. 1988. Application of probabilistic methods in the analysis of phytosociological data. Coenoses 3: 1-10.

KELLER, M. 1987. *Ausaperungskartierung mit Landsat-MSS Daten zur Erfassung ökologischer Einflussgrössen im Gebirge*. Diss. Univ. Zürich.

KRAUSE, M. 1986. *Die Böden von Davos: Ertragspotential, Belastbarkeit und Gefährdung durch Nutzungsänderungen*. Schlussber. Schweiz. MaB-Programm Nr. 18. Bundesamt für Umweltschutz. CH-3003 Bern.

MATHER, P.M. 1987. *Computer Processing of Remotely-sensed Images*. New York. Wiley.

WALTER, H. and E. WALTER. 1953. Einige allgemeine Ergebnisse unserer Forschungsreise nach Südwestafrika 1952/53: Das Gesetz der relativen Standortskonstanz; das Wesen der Pflanzengemeinschaften. Ber. Deutsch. Bot. Ges. 66: 228-236.

WASHBURN, B. 1988. Mount Everest. Surveying the third pole. Nat. Geographic 174: 653-659.

WILDI, O. und K. EWALD. 1986. *Der Naturraum und dessen Nutzung im alpinen Tourismusgebiet von Davos. Ergebnisse des MaB-Projektes Davos*. Eidg. Anst. Forstl. Versuchswes., Ber. Nr. 289. CH-8903 Birmensdorf.

ZUMBÜHL, G. 1986. Vegetationskartierung des MaB6-Testgebietes Davos. In: Landolt, E. *et al.* (ed.), *Vegetationskartierung und Untersuchungen zum landwirtschaftlichen Ertrag im MaB6-Gebiet Davos*, pp. 13-113. Veröff. Geobot. Inst. ETH, Stiftung Rübel, Zürich 88a.

Manuscript received: June 1989