| Stuttgarter Beitr. Naturk. | Ser. A | Nr. 456 | 139–145 | Stuttgart, 31. 12. 1990 |
|---|---|---|---|---|

# Numerical Analysis of Data on Lichen Distribution in Baden-Württemberg: a Preliminary Outline

By Martin Pietschmann, Stuttgart

With 4 figures

## Summary

"Numerical techniques" used to discern putative causalities in lichen distributional patterns in Baden-Württemberg (Southwestern Germany) are outlined in a manner understandable by non-specialists. They present current interactions occurring between organisms and environmental factors in graphic displays. Examples are given for two phenomena: species redundancy and dimensionality of the structure of lichen vegetation. Both examples are major branches within the strategy of applying the "zero-hypothesis", the assumption that there is no distributional pattern at all, in the analysis of large scale distribution data. Future results are expected to yield new aspects for sampling design, which may then better fit the microclimate dependent ecology of cryptogams.

## Zusammenfassung

„Numerische Techniken" in der Analyse der Flechtenverbreitung in Baden-Württemberg (Südwestdeutschland) werden als ein Hilfsmittel vorgestellt, mit dem die Ursachen der Flechtenareale beschrieben und in verständlicher Form auch dem Nicht-Spezialisten vermittelt werden können. Mit ihrer Hilfe können existierende Wechselwirkungen zwischen den vorkommenden Organismen und Faktorennetzen graphisch dargestellt werden. Für die beiden Phänomene der Artenredundanz und der strukturellen Dimensionalität der Flechtenvegetation werden Beispiele aufgeführt. Beide Beispiele stellen Hauptwege einer Strategie dar, welche die folgende Nullhypothese anwendet: Das großräumige Vorkommen von Flechten weist keine wahrnehmbaren Verbreitungsmuster auf. Die zu erwartenden Ergebnisse können unter Umständen neue Wege zu einer verbesserten Aufnahmetechnik weisen, die der Ökologie von Kryptogamen besser gerecht werden kann.

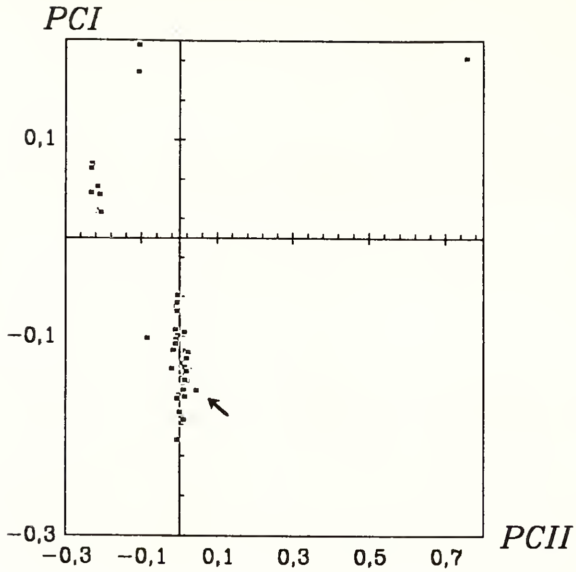
## 1. Introduction

In modern times all sciences are charaterized by a vast increase in data and knowledge. When I started school about 25 years ago, no lichen floras of Central Europe existed, whereas nowadays I am lucky to find these as well as explicit data on lichen distribution, e. g. in Baden-Württemberg (Wirth 1987). But I would like to ask what the aims – beyond distributional aspects – are of drawing maps which look far too rough to be able to compete with detailed low-scale recording of distribution.

I think, everybody knows from his own personal field experience that lichens in general are strikingly sensitive organisms which can indicate what is occurring in the surrounding environment.

Yet, how can one mediate between the lichen specialist and the laymen, who by looking at maps and pictures may hardly be able to imagine what an immense degree of information was lost, if lichens would disappear?

One promising possibility might be to inform laymen that „patterns and processes" evident in lichen vegetation reveal intrinsic networks between the natural landscape and the living organisms. This approach implies the use of an explicit and



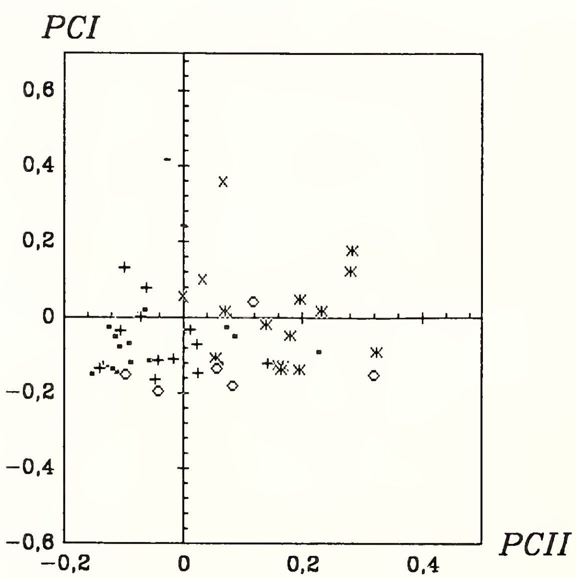




Fig. 1. PCA-ordination scatter diagrams performed on large scale and low scale data. Results of any analysis obviously depend on scale. The diagram above indicates strongly discontinous data of fairly simple structure. The one below shows results of an ordination of the point cluster indicated by an arrow in the upper diagram. Overlay symbols indicate compositional structure of high dimensionality not visible in two-dimensional space.

consistent medium or language, which must be capable of filling the hiatus between the specialist and the non-specialist. What I have in my mind for this task of course is the abstract but condensed and powerful language of "numerical techniques".

I will try to describe one major problem, that may perhaps be of special importance for the evaluation of lichen distribution data: Can we expect highly structured lichen area-types to exist, when it is known that lichens are cryptogams sensitive to microclimate?

First let me state simple hypotheses on what is happening, based on data from 318 map grids and about 1000 lichen species in Baden-Württemberg: Is there a random or a well structured distribution? Are there major ecological factors which are of a high explanatory value or not?

Fig. 1 shows two fairly different scatter diagrams. The upper ordination scatter diagram shows two well separated point clusters and four obvious outlayers. The one below is a principal component analysis of the point cluster indicated above by an arrow.

I do not want to give detailed descriptions of the main strategy, the way to manage such a great mass of data and the possibilities of interpreting subsequent numerically derived results. What I am going to discuss are two major branches within my strategy of analysing these data.

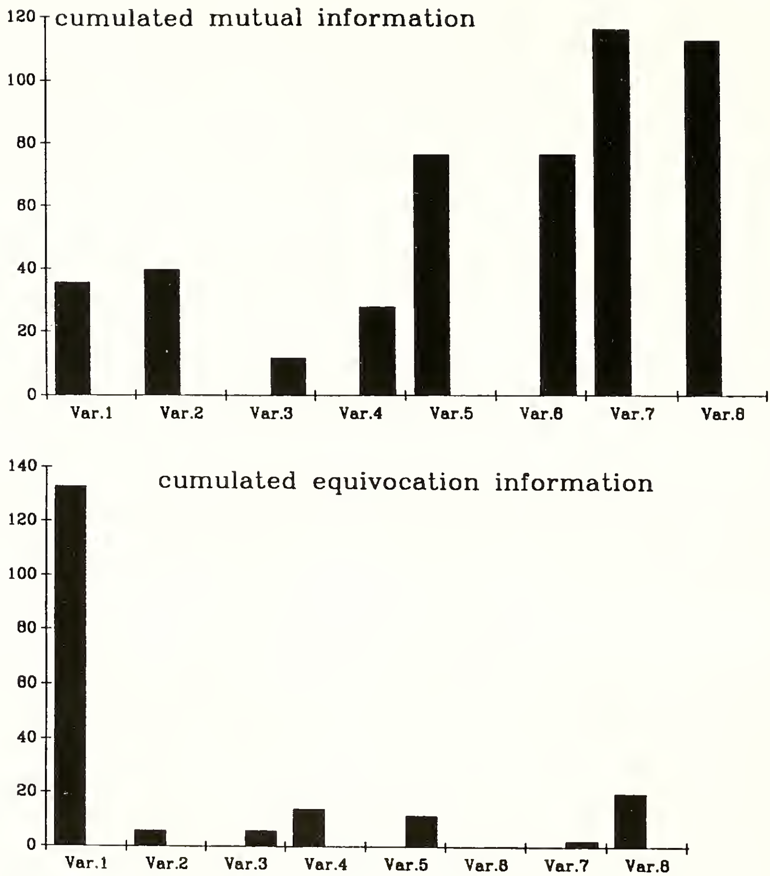
## 2. Species redundancy

It is well known that different species may have a similar distribution because of affinities in their ecological demands. Thus, it is possible to reduce a long species list by a criterion of redundancy (FEOLI et al. 1984). Figure 2 shows results achieved by applying two different information redundancy criteria to the same data set, the mutual information I (Au; Az) and the equivocation information E (Au; Az). Both diagrams describe the same data set, but the results look rather different. The first describes the mutual enclosement of two frequency distribution sets of variables, the so called "intersection of information". The latter describes the unique part of two frequency distributions, i. e. that part of information which is not influenced by other variables. Ranking by equivocation information indicates a strong asymmetry of information present in the variables considered.

Figure 3 shows the well known VENN diagram which symbolizes the relation between the two redundancy criteria. The degree of overlapping information (mutual information) depends on the individual types of information I (Au) and I (Az). The total uncertainty [joint information I (Au; Az)] also depends on the degree of intersection of the frequency distribution. How can the relation of these two redundancy criteria to lichen distribution data be described?

Assuming a general influence of man's activities, e. g. $SO_2$-emission, there should exist a large group of species, let us say around *Hypogymnia physodes*, which is favoured by the general decline of natural woodland species because they are able to colonize stands where lichens grow poorly. This is the part of intersection of information which corresponds to the absence of species indicating high air quality, e. g. the group around *Lobaria pulmonaria*.

On the other hand, there may exist rare assembleges of species which do not correlate with air quality factors but represent suppressed patterns depending on the intactness of natural environments which are clearly underrepresented in a man influenced world. Due to its uniqueness such an assemblage provides very specific



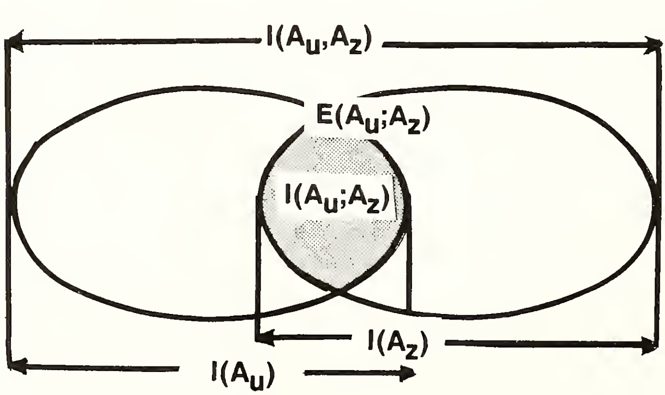




Fig. 2. (above) Bar diagrams of two different redundancy criteria applied to the same data set.

Fig. 3. (below) VENNs diagram describing relations between different functions of information.

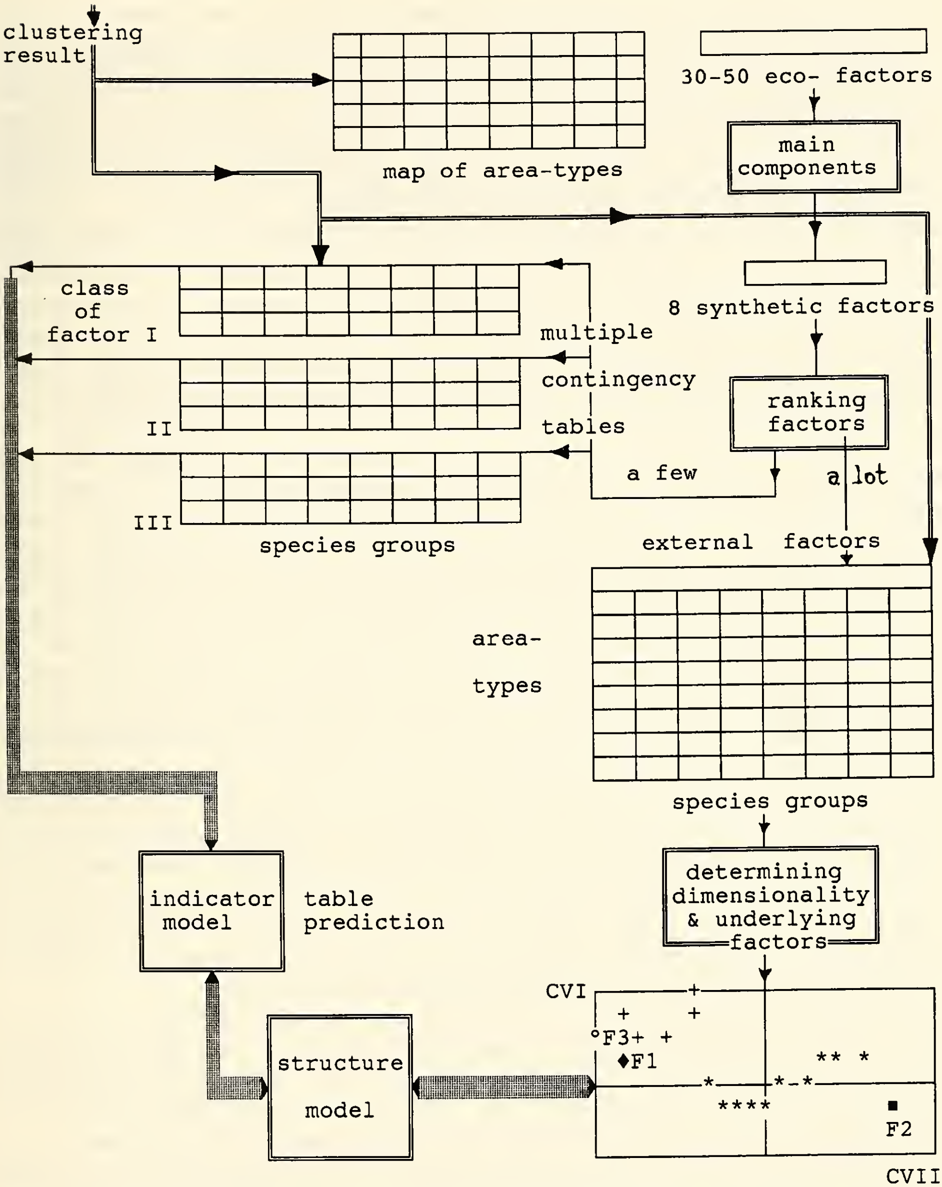



Fig. 4. Flow chart of the ultimate step of numerical analysis. At the top information enters resulting from preliminary clustering of species versus map grids. The major dichotomy of simple to complex variation pattern is shown in the center.

information, which is only partially dependent on major environmental trends. The environmental biologist is clearly asked in this case to give a sound answer to which of the two redundancy criteria is preferred in order to reduce a long species list without loss of considerable information.

## 3. The underlying structure of ecological factors

It seems impossible to derive a priori decisions about continuity and discontinuity of ecological factors determining structure of distribution, the so called "dimensionality of structure" (Feoli & Orloci 1979), from the considered data set (318 map grids and about 1000 lichen species). It is more instructive to analyze factors and species simultaneously, which alone is the ecologically proper approach, than to factors and species jointly as performed by Haeupler (1974).

If area-types and species groups are identified by clustering criteria a simple question arises, which yet may be difficult to answer. Are assignments like these satisfactory?: „A species group is the representative of a given area-type." On the contrary, I think we should demand stability of classificatory results (area types and species groups) and ask in the sequel, what makes lichen vegetation show the revealed distribution patterns. In doing this, we will start to analyse factor/species groups contingencies.

Different strategies are proposed depending on whether vegetation structure shows clear trends or a seemingly chaotic, inobvious variation in composition. Figure 4 shows a flow chart of the whole process. Here again, analysis of c. 30 different ecological factors which are expected to show interesting interactions with lichen distribution, is another difficult task. For instance, before starting the analysis, we have to sort out logical and partially logical factor to factor dependencies (Green 1979). For example there may exist a large scale-dependency between elevation and annual precipitation in Baden-Württemberg. This is why any statistical analysis of factor interactions and factor/species interactions is going to fail, simply because of the assumption of an independency of variables. I will therefore analyse ecological factors in the sequel by extracting independent synthetic principal component (PC) axes which are capable of summarizing the structure of factors in a fairly powerful way. Within this step success of extracting PC-axes can be estimated and it can be decided, if it is possible to disregard a large number of extracted variables.

If the number of important axes is low, unidimensional factor/species group-contingency tables are sufficient to deduce an indicator-model of maximum simplicity (Daget et al. 1972). This classification by single factors can of course be displayed graphically by ordination scatter diagrams or grid map overlays.

If the number of important axes is high, the dimensionality of species concentrations versus area types has to be analysed in order to build up a structural model of lichen vegetation in Baden-Württemberg. Applied to both cases, the zero-hypothesis is stated as: "There is no structure at all." (Orloci & Kenkel 1985). Inference of mans influence can be stated indirectly by the dimensionality of vegetation.

In both cases structured tables, similar to those used in phytosociology, will be used. The deviation from random variation can be calculated from these tables.

## 4. Conclusions

Both discussed approaches are intended to yield comprehensive tabular and graphical representation of results. Thus, the results themselves can serve as an efficient and convincing medium by using a language which everybody is able to read, i. e. graphics.

I anticipate results which can serve as a baseline for estimating the efficiency of our sampling design. For instance, deleting species abundancy can be a matter of argument. On the other hand, is it worthwhile to map each of the c. 1000 lichen species to gain conclusions about large scale variation patterns, e. g. of environmental stress? Or, is it sufficient to confine the sampling procedure to a set of indicator species? Our tentative predictive model can help in solving important questions like these.

Planning of mapping projects in the future can be infered indirectly from results obtained in Baden-Württemberg. Thus this pilot project can be a step forward towards an integration of lichenological knowledge into plans of nature protection and other mapping projects carried out by higher plants specialists or zoologists. New bridges could be built if we are able to state results explicitly from a sound data base.

## 5. Literature

DAGET, P., GODRON, M. & GUILLERM, J. L. (1972): Profils écologiques et information mutuelle entre espèces et facteurs écologiques. – *In*: R. TÜXEN (ed.): Grundfragen und Methoden in der Pflanzensoziologie. – Ber. Intern. Symposion **1970**: 113–119; The Hague (Junk).

FEOLI, E., LAGONEGRO, M. & ORLOCI, L. (1984): Information analysis of vegetation data. – 143 p.; The Hague, Boston & Lancaster (Junk).

FEOLI, E. & ORLOCI, L. (1979): Analysis of concentration and detection of underlaying factors in structured tables. – Vegetatio **40**: 49–54; Dordrecht.

GREEN, R. H. (1979): Sampling design and statistical methods for environmental biologists. – 257 p.; New York, Chichester, Brisbane & Toronto (John Wiley & Sons).

HAEUPLER, H. (1974): Analyse von Punktrasterkarten. – Scripta Geobot. **8**: 1–141; Göttingen.

ORLOCI, L. & KENKEL, N. (1985): Introduction to data analysis. – 340 p.; Burtonsville (International Publishing House).

WIRTH, V. (1987): Die Flechten Baden-Württembergs. Verbreitungsatlas. – 528 p.; Stuttgart (Ulmer).

Author's addresss:

MARTIN PIETSCHMANN, Staatl. Museum für Naturkunde (Museum am Löwentor), Rosenstein 1, D-7000 Stuttgart 1, Germany.