PREDICTION OF CHANGES IN EPHEMEROPTERAN COMMUNITIES - A TRANSITION MATRIX APPROACH

Jan Leps[1], Tomás Soldán[2] and Vladimír Landa[2]

[1]*Laboratory of Biomathematical Methods*
[2]*Institute of Entomology, Czechoslavak Academy of Sciences, Branisovská 31, 370 05 Ceské Budejovice, Czechoslovakia*

ABSTRACT

Quantitative monthly samples, taken from 150 evenly distributed localities in the Labe basin, Czechoslovakia, during two research periods in the 1950's and 1970's were classified using a saprobiological classification (Czechoslovak National Standards Nos 830532 and 830602) and the TWINSPAN method (numerical hierarchical polythetic classifications). Transition matrices, i.e. matrices of transition probabilities between particular types, were constructed for each of the methods of mayfly community classification used. All the "matrices" reflect the tendency of the water quality to gradually deteriorate as indicated by changes in mayfly communities. The classification by saprobity classes seems to be too coarse for predictive purposes (most localities fall into betamesosaprobity class). Combined information from matrices based on saprobial index classes and TWINSPAN provided reasonable description of the process.

INTRODUCTION

Changes in animal communities represent one of the main areas of contemporary ecological research. Among other purposes the composition of some communities sensitively reflects the state of environment as well as the whole ecosystem. Since the nymphs of mayflies constitute one of the most important groups of freshwater invertebrates the changes in their natural communities can indicate substantial and long-term changes of water quality. Moreover, based on the mayfly community changes in the past or at present, changes in aquatic ecosystems in the future and trends influencing water quality can be predicted.

The transition matrix models are the simplest method for predicting further development of ecological systems. They have been frequently used for study of plant succession (Hulst 1979) but they are undoubtedly applicable to any type of dynamical change (Usher 1979, 1981) although applications in animal ecology are rare. The transition matrix models are mainly descriptive tools (low resolution models) that do not take into account the mechanisms of particular change. Their predictions should be considered as linear extrapolations of present trends only. They stem from the following simplifying assumptions:

1. The state of a system may be assigned to one or several non-overlapping classes (states);
2. The probability that the system will change during a time step from the state j to the state i (transition probability $P_{i,j}$) depends on the present state of the system only (Markovian assumption; the models are sometimes termed Markovian). In ecological terms $P_{i,j}$ means the probability that the community j will be replaced by a community type i during a time step; and
3. The transition probabilities are constant over time (time homogeneity assumption).

I.C. Campbell (ed.) Mayflies and Stoneflies, 281–287.

In this paper the above methods are applied to the evaluation and prediction of mayfly community changes and the results from classical saprobiological methods are compared with those based on numerical classification of mayfly communities.

METHODS

Sampling

During the two phases (1950-1965 and 1970-1985) of an intensive faunal investigation of aquatic insects in the Labe (Elbs) basin in Czechoslovakia more than 400,000 specimens of mayfly larvae were collected from 150 localities. Material is deposited in the Institute of Entomology in Ceské Budejovice, Czechoslovakia. These localities are evenly distributed within all the types of aquatic biotopes at various altitudes (for details and list of localities see Landa 1984, Landa and Soldán 1988). Samples were taken quantitatively ('kick' samples, Surber samples and occasionally drift nets) for one year during each phase of research in all seasons (monthly sampling except for December and January).

Classification of samples

To evaluate both the quantitative and qualitative changes in mayfly communities at these 150 localities within the past 20–30 years and to prepare data for the construction of transition matrices, the following methods of classification of mayfly communities were adopted:

1. Classification to saprobiological classes with a determination of "resulting" saprobity for every locality in the respective period: saprobiological classes were defined according to Czecholslovak National Standard No. 830532. Localities investigated were placed into 4 classes, as follows: x (xenosaprobity, water of the best quality): o (oligosaprobity); bmes (beta-mesosaprobity); and ames (alpha-mesosaprobity, water of a very low quality). Localities where no mayflies were found during the second phase of research (4 cases) were tentatively classified as 'alpha-mesosoprobial' although the water quality may have been even worse (polysaprobity category). For details on the determination of resulting saprobity of individual localities on the basis of mayfly nymphs see Zelinka and Marvan (1961), Braasch and Jacob (1976) and Russev (1979).

2. Classification of water quality according to saprobial index by Pantle-Buck which is correlated with oxygen content classes (see Czechoslovak National Standards Nos 830532 and 830602). Localities with saprobial index < 1.00 are considered as those with a very high (best) water quality, 1.01–1.50 indicates slightly polluted water, 1.51–2.00 moderately polluted water and localities with saprobial index higher than 2.01 are exposed to considerable pollution at least for several months a year.

3. Twinspan method (Hill 1979). Twinspan is an hierarchial divisive polythetic classification method, developed originally for classification of plant communities. The set of releves is successively split in two parts, each part is split in two parts again until either some hierarchial level is reached or the size of the group drops below a stated threshold. Each dichotomy is characterized by indicator species. The dichotomies are often asymmetric. Consequently, we adopted the size of a group as a stopping rule. The localities without mayflies were assigned to a separate group. The classifications were compared mutually using the coefficient of Goodman and Kruskal (1954) which ranges in value from 0 to 1 with the latter corresponding to identical classifications.

Transition matrices

A simple transition matrix model was used for prediction of further composition of the set of localities:

$$N_{t+1} = P. N_t,$$

where N_t is the state vector (column) describing the composition of the set of localities in time t ($N_{i,t}$ is the number of localities of i-th type in time t) and P is the matrix of transition probabilities (transition matrix), $P_{i,j}$ is the probability of transition from state j to state i. The time step is twenty years.

The matrices of transition probabilities were estimated simply by:

$$P_{i,j} = \frac{\text{number of transitions from j to i}}{\text{number of localities in state j in fifties}}$$

The localities without mayflies were found in the second research period only. Therefore, we have no empirical data to estimate the transition on probabilities from these "dead" waters by any other types. Consequently, we stated arbitrarily the probability of remaining in the "dead" state to be 1 (and all the probabilities of transition from the "dead" state to be zero). This approach is well-based, because the "dead" localities are usually not able to recover without an application of some protective measures.

The predictions were done for two time steps only (i.e. for 1990's and 2010's). The general trend may be illustrated by results of eigenvalue and eigenvector analysis of the matrix P. The dominant eigenvalue always equals one and the corresponding eigenvector (determining here the composition after many time steps, theoretically for $t \rightarrow \infty$) is composed of zeros in all positions but that corresponding to "dead" waters (consequence of stating $P_{i,i} = 1$ for i corresponding to "dead" water). In terms of Markovian models, the "dead" state is the only absorbing state. The rate of disappearance of "live" localities is characterized by "eigenvalue quotient" $1/\lambda_2$, λ_2 is the second largest eigenvalue (see e.g. Usher 1979). The relative representation of types of localities with mayflies is proportional to the elements of the eigenvector corresponding to second largest eigenvalue. These proportions are listed in the tables as illustrations of the trend (not as a prediction).

RESULTS

Classification of samples

Three classifications were obtained, two traditional ones and the numerical one (TWINSPAN). Results of the TWINSPAN classification are presented in Fig. 1. The particular groups are ecologically well-defined and may be described as follows:

Group 1 –	Localities with the best water quality without any pollution: mostly montane streams of high altitudes (900 to 1000 m a.s.l.).
Group 2	Localities with a very high water quality, montane and submontane lakes and pools sometimes exposed to acidification, particularly at altitudes above 750 m a.s.l. in the mountains.
Group 3	Localities with good water quality, mostly submontane streams, brooks and smaller rivers at altitudes of 500–750 m a.s.l. showing the highest species diversity and a wide spectrum of aquatic habitats with only local sources of pollution.
Group 4	Localities of a very low water quality permanently exposed to pollution. Foothill and lowland brooks and streams, canals and dredges, some of them even eutrophied. Low species diversity localities, inhabited mostly by species with a very wide ecological range, resistant to pollution.
Group 5	Localities with moderately polluted water (mostly "cumulative" pollution). Large lowland river localities up to the altitudes of 250–500 m a.s.l. showing a relatively high species diversity of mostly stenotopic riverine species with disjunctive areas.

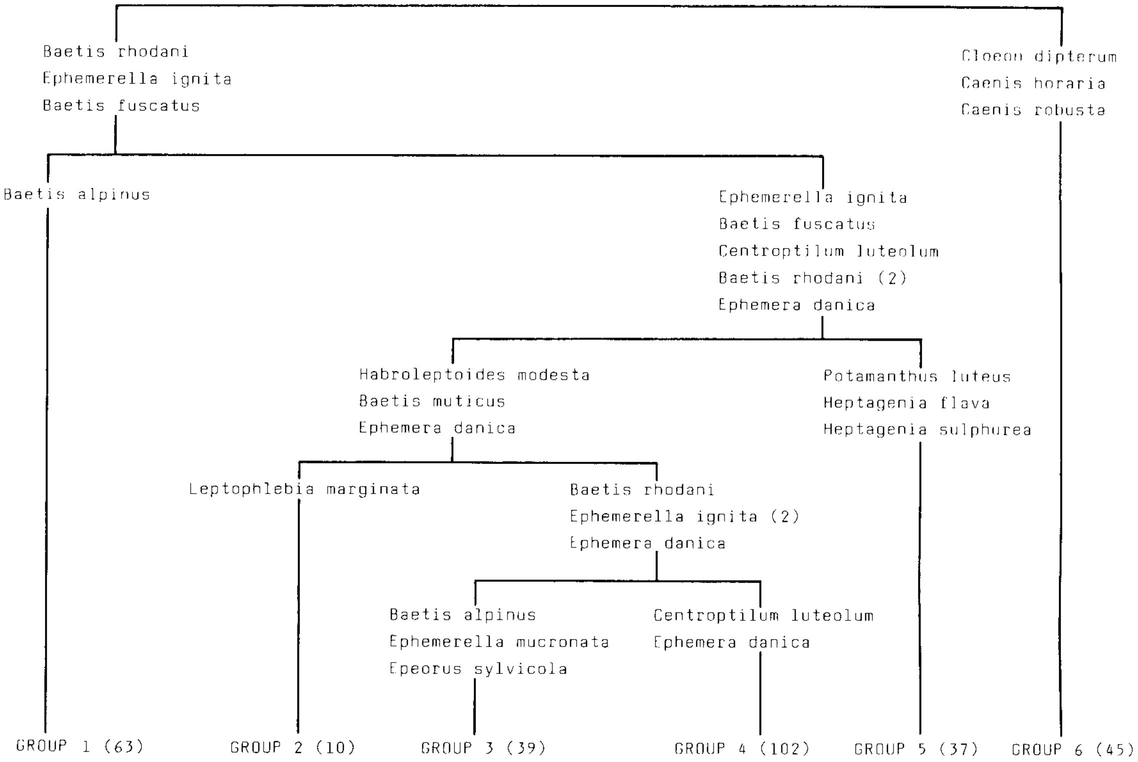

Fig. 1. A dendrogram representing the results of the TWINSPAN classification of sites based on mayfly assemblages.

Group 6 Considerably polluted and mostly eutrophied localities of lowland and foothill lakes and fish-ponds, inhabited by several species with adaptation to survive partial or long-term anoxia.

Mutual similarities (coefficient of Goodman & Kruskal 1954) of these three classifications are as follows: classification 1 and 2: similarity 0.41; 1 and 3: 0.24; 2 and 3: 0.38. The most similar are classifications 1 and 2 – both are based on the same indication values. A relatively high similarity was found between classifications 2 and 3, based on completely different principles. Differing from the previous two, the TWINSPAN classification reflects not only the trophic status of the locality but other ecolcgical factors (particularly those related to the elevation).

Transition matrices

Transition matrices are presented in Tables 1, 2 and 3. In Tables 1 and 2, the categories are order-

Table 1. Estimated transition matrix for mayfly communities according to saprobial classes ('resulting' saprobity)

To/From	x	o	bmes	lwm
x	0.76	0.11	0.02	0
o	0.10	0.58	0.04	0
bmes	0.14	0.26	0.91	0
lwm	0	0.05	0.03	1.0

x = xenosaprobity, o = oligosaprobity, bmes = betamesosaprobity, lwm = localities without mayflies and is equivalent to alphamesosaprobity

Table 2. Estimated transition matrix for mayfly communities according to saprobial index classes. 'lwm' = localities without mayflies

To/From	< 1.0	1.01–1.5	1.51–2.0 > 2.0 lwm		
< 1.0	0.63	0.20	0	0	0
1.01–1.5	0.25	0.57	0.09	0	0
1.51–2.0	0.09	0.23	0.67	0.06	0
> 2.0	0	0	0.22	0.88	0
lwm	0.03	0	0.02	0.06	1.0

Table 3. Estimated transition matrix for mayfly communities according to TWINSPAN classification (classes 1–6 corresponding to Fig. 1). 'lwm' = localities without mayflies

To/From	1	2	3	4	5	6	lwm
1	0.77	0	0.06	0	0	0	0
2	0	0.60	0	0	0.05	0.05	0
3	0.14	0	0.82	0.04	0.05	0	0
4	0.06	0.20	0.06	0.90	0.05	0	0
5	0	0	0.06	0.02	0.75	0	0
6	0	0.20	0	0	0.05	0.95	0
lwm	0.03	0	0	0.04	0.05	0	1.0

ed according to their position on the "saprobiological" gradient, i.e. individual locality shifts between categories of saprobial index or resulting saprobity values. In both cases, the below-diagonal elements exceed the above-diagonal ones (i.e. $P_{j,i} > P_{i,j}$ for $i > j$). In Table 2, this j,i i,j "rule" holds in all cases, in Table 1, the only exception is $P_{2,1} > P_{1,2}$ (0.11 and 0.10 respectively). This means that continuous shift toward higher saprobity appears. The pattern of transitions based on TWINSPAN defined types (Table 3) is more complicated. However, the general tendency of shifting toward more eutrophied waters remain unchanged.

In all cases, the highest $P_{i,i}$ (i.e. the strongest tendency to remain in the same state) is performed by low-quality waters – betamesosaprobial waters, waters with saprobial index higher than 2 and groups 4 and 6 of the TWINSPAN classification.

The frequencies of particular types in the two research periods and predictions are listed in Tables 4, 5 and 6. All of them predict increase of localities without mayflies and decrease in number of best water quality localities (particularly localities with saprobial index smaller than 1 and TWINSPAN group 1, less pronounced decrease is exhibited by xenosaprobial localities). In TWINSPAN based predictions the number of 3rd group localities (i.e. localities with relatively good water quality) increases temporarily (particularly because of high $P_{3,1}$, i.e. high transition probability from the best water quality localities) and very slight increase of very low quality waters (group 6) is expected. Similarly, continuing increase of number of localities with saprobial index higher than 2 is expected. The tendency toward increase of eutrophication is stressed by the trend predicted by the eigenvector analysis (last columns in Tables 4, 5 and 6). Note that only proportions of localities with mayflies are in question.

Table 4. Percentual representation of particular types in the set of releves (saprobial classes); the last column concerns only localities with mayflies. 'lwm' = localities without mayflies

Saprobial classes	Observed		Predicted		
	1950's	1970's	1990's	2010's	t → ∞
x	19	17	16	15	13
o	13	12	12	11	11
bmes	68	68	67	66	76
lwm	0	3	5	8	

Table 5. Percentual representation of particular types in the set of releves (saprobial index); the last column concerns only localities with mayflies. 'lwm' = localities without mayflies

Saprobial index	Observed		Predicted		
	1950's	1970's	1990's	2010's	t → ∞
< 1.0	21	17	15	13	5
1.01–1.5	20	20	19	17	9
1.51–2.0	37	33	30	28	22
> 2.0	22	27	31	34	64
lwm	0	3	5	8	

Table 6. Percentual representation of particular types in the set of releves (TWINSPAN); the last column concerns only localities with mayflies. 'lwm' = localities without mayflies

TWINSPAN goup	Observed		Predicted		
	1950's	1970's	1990's	2010's	t → ∞
1	23	19	15	13	3
2	3	3	3	3	7
3	12	15	17	17	11
4	34	34	34	34	29
5	13	11	10	10	5
6	15	15	16	16	45
lwm	0	3	5	7	

DISCUSSION

There are two basic problems in application of transition matrix models in community ecology (see e.g. Usher 1979, 1981): (1) The definition of types and (2) The possible bias of predictions caused by violation of the Markovian assumptions and/or of the time homogeneity assumption.

1. We have used three classifications for the definition of types. The first and second one are based on indicator values of particular species with respect to saprobity (the indication values of particular species had been stated according to the long-term field experience and were subsequently modified in Czechoslavak National Standards). This approach is based on the same principles as the weighted average method of direct gradient analysis by Whittaker (1978) or as the indicator values of Ellenberg (1974). The subjectivity of these methods has been pointed out many times (e.g. Gauch 1982). The above mentioned methods are used for ordination – in our case further difficulty lies in the necessity to split the gradient into discrete categories. On the other hand, the indicator values enable the relationship between particular environmental factors (saprobity here) and community composition, to be stated.

2. We used the transition matrices mainly as a descriptive tool. The mechanisms and causes of community change are not included in the model. Because of the large-scale nature of the model the causes of particular changes differ considerably including consequences of global changes in precipitation pH, local changes in water pollution, and reflect even changes in economic activities, like changes in agricultural practices. The trend might be considerably changed by taking some measures against eutrophication. The predictions are based on the time homogeneity assumption. Hence they have to be taken conditionally – i.e. what would happen if the present trends continue. Similarly Finn (1985) found the transition matrices useful for the analysis of large-scale land use changes. Our matrices represent the (absorbing) Markov chains, because the probability of remaining in the state "without mayflies" is 1. However, this value was not estimated from the data, but was deduced from the fact that the "dead waters" are not able to recover if protective measures are not taken.

Although the selection of localities investigated seems to be very representative for description and/or definition of present trends of water quality deterioration, some limitations must be taken into account.

1. Water quality indication is based solely on mayflies and does not consider other aquatic organisms. In spite of their relatively very high indicator value and high abundance, mayfly nymphs may be used reliably to determine changes of relatively clean, running waters predominantly of lotic-erosive biotopes. Consequently, taking into consideration the restricted usefulness of mayflies in lotic-depositional and especially lentic biotopes, indication of water quality at these biotopes is less detailed although basic information on change is available.

2. Some changes of water quality cannot be determined at all. For instance, increasing air-borne acidity of water is not reflected by any changes in mayfly communities in slightly acid waters. Also pronounced changes (e.g. elimination of *Baetis* and increasing abundance of *Siphlonurus* and *Leptophlebia* currently occurring in Europe – See Harmanen 1980) are not detectable by present methods which are only suitable for determining gross organic pollution.

3. Heavily polluted localities where no mayflies were found during the first phase of research were not considered at all. The ratio of such localities to localities with mayflies is estimated as 1: 10–15 in localities chosen by chance.

The results obtained by comparison of the above three methods show the classical saprobial method to be very rough with only limited informative value for predictive purposes. "Resulting saprobity" classes are very wide (more than 2/3 of localities fall into betemesosaprobity class) and do not reflect minor changes of mayfly communities and transitional situations. Only very pronounced changes caused by heavy pollution lead to shifts among individual classes. That is why the predictive changes (Table 4) are very inconspicuous.

On the other hand, the classification based on saprobial index can be easily compared with the TWINSPAN classification. Classes of saprobial index values (defined according to water quality categories – see material and methods) evidently reflect moderate changes although they need not definitively mean change of individual classes. Predictions based on the saprobial index (Table 5) seems to be much more detailed than those by saprobial classes (Table 4).

The TWINSPAN classification provided ecologically interpretable results. However, the classification reflects not only the saprobiological status of the locality. This method provided the most detailed insight into transitional structure of the system.

In general, it seems that there is no single best method; in our case the combined information from matrices based on saprobial index and TWINSPAN provided a reasonable description of the process.

ACKNOWLEDGEMENTS

Our sincere thanks are due to Dr. I.C. Campbell of Water Studies Centre, Chisholm Institute of Technology, Caulfield East, Australia, who corrected our English text.

REFERENCES

Braasch, D. and Jacob, U. (1976). Die Verwendung von Ephemeropteren (Insecta) der DDR als Indikatoren für die Wasserg te. *Ent. Nachr.* 20: 101–109.

Czechoslovak National Standard No. 830532 (1966). Evaluation of surface water quality and its classification. Praha. [in Czech].

Czechoslovak National Standard No. 830532, (1980). *Biological Analysis of Surface Water.* Part VI: Determination of saprobial index according to Pantle-Buck. Praha 1980. [In Czech.]

Ellenberg, H. (1974). Zeigerwerte der Gefässpflanzen Mitteleuropas. *Scripta Geobotanica* 9: 1–97.

Finn, J.T. (1985). Analysis of land use change statistics through the use of Markov chains. *In* 1985 Summer Computer Simulation Conference. Chicago 22-26 July 1985. pp. sep. 1–10.

Gauch, H.G. (1982). *Multivariate Analysis in Community Ecology.* Cambridge U. P., Cambridge.

Goodman, L.A. and Kruskal, W.H. (1954). Measures of association of cross classifications. *J. Amer. Stat. Ass.* 49: 732–764.

Harmanen, M. (1980). Der Einfluss sauer Gewässer auf den Bestand der Ephemeropteren – und Plecopterenfauna. *Gewäss. Abwäss.* 66/67: 130–136.

Hill, M.O. (1979). *TWINSPAN, a FORTRAN Program for Arranging Multivariate Data in an Ordered Two-way Table by Classification of the Individuals and Attributes.* Ecology and Systematics. Cornell Univ. Ithaca.

Hulst, R. van (1979). On the dynamics of vegetation: Markov chains as models of succession. *Vegetatio* 40: 3–14.

Landa, V. (1984). Studies on aquatic insects of Czechoslovakia with regards to changes in the quality of water in the past 20-30 years. *Proc. 4th Int. Conf. Ephemeroptera, Bechyne*, 1983.

Landa, V. and Soldán, T. (1988). Distribution of mayflies (Ephemeroptera) in Czechoslovakia and its changes in connection with water quality changes in the Elbe basin. *Trans. Czechoslov. Acad. Sci.* in print.

Russev, B. (1979). Die Anpassungsfähigkeit der Ephemeropteren an die Verunreinigung der Gewässer und die Möglichkeit ihrer Ausnützung als Limnosaprobe Bioindikatoren. *Proc. 2nd Intern. Conf. Ephemeroptera, Krakow*, 1975.

Usher, M.B. (1979). Markovian approaches to ecological succession. *J. Anim. Ecol.* 48: 413–426.

Usher, M.B. (1981). Modelling ecological succession, with particular reference to Markovian models. *Vegetatio* 46: 11–18.

Whittaker, R.H. (1973). Direct gradient analysis: techniques. *In* R.H. Whittaker (ed.) *Handbook of Vegetation Science.* 5: 7–31. Junk, The Hague.

Zelinka M. and Marvan P. (1961). Zur Präzisierung der biologischen Klassifikation der Reinheit der Fliessender Gewässer. *Arch. Hydrobiol.* 57: 389–407.