International Council for
the Exploration of the Sea

CM 2000/S:12
Theme Session S: Temporal and Spatial Trends in the Distribution of Contaminants and their Biologic Effects in the ICES Area

ANALYSIS OF ICES LONG-TERM DATA ON DISEASES OF NORTH SEA DAB (*LIMANDA LIMANDA*) IN RELATION TO CONTAMINANTS AND OTHER ENVIRONMENTAL FACTORS

W. Wosniok, T. Lang, V. Dethlefsen, S.W. Feist, A.H. McVicar, S. Mellergaard, A.D. Vethaak

Abstract

ICES data on the prevalence of grossly visible diseases (lymphocystis, epidermal hyperplasia/papilloma, acute/healing skin ulcerations) of dab (*Limanda limanda*) submitted by Member Countries were statistically analysed with respect to potential relations with contaminants in water, sediments and biota as well as nutrients, water temperature, salinity, oxygen content and catch per unit effort (CPUE). Data were extracted from the ICES Environmental Data Centre, the ICES Oceanography Data Centre and the ICES Fishery Databanks. The analysis was carried out for three regions located in the south-eastern, central and north-western North Sea which were selected on the basis of the availability of disease data. The time span considered partly covered almost two decades. Non-parametric interpolation techniques were used to obtain the necessary uniform time pattern for all time series. Parameter estimates and significances within a logistic model for the disease prevalences were calculated by means of a bootstrap procedure which accounted for the need to interpolate within observed time series. A variety of factors, including contaminants, were identified as being significantly related to the disease prevalence. However, depending on area, time range and data availability, different sets of factors were identified. This reflects the multifactorial aetiology of the diseases covered, but can also be attributed to some high correlations among the explaining quantities.

W. Wosniok: University of Bremen, Institute of Statistics, P.O. Box 330440, D-28334 Bremen, Germany [tel: +49 421 2183471, fax: +49 421 2184020, e-mail: wwosniok@math.uni-bremen.de]. T. Lang and V. Dethlefsen: Federal Research Centre for Fisheries, Institute for Fisheries Ecology, Deichstraße 12, D-27472 Cuxhaven, Germany [tel: +49 4721 38034, fax: +49 4721 53583, e-mail: first initial. surname@t-online.de]. S.W. Feist: CEFAS, The Weymouth Laboratory, The Nothe, Weymouth, Dorset, DT48UB, U.K. [tel: +44 1305 206662, fax: +44 1305 206601, e-mail: S.W. Feist@cefas.co.uk]. A.H. McVicar: The Marine Laboratory P.O. Box 101, Victoria Road, Aberdeen AB119DB, Scotland [tel: +44 1224 295605, fax: +44 1224 295511, e-mail: mcvicara@marlab.ac.uk]. S. Mellergaard: Danish Institute for Fisheries Research, Fish Disease Laboratory, Stigbojlen 4, DK-1879 Frederiksberg C, Denmark [tel: +45 35 352767, fax: +45 35 282711, e-mail: sme@dfu.min.dk]. A.D. Vethaak: RIKZ, Ecotoxicology Section, P.O. Box 8039, 4330 EA Middelburg, The Netherlands [tel: +31 118 672311, fax: +31 118 672311, e-mail: A.D.Vethaak@rikz.rws.minvenw.nl]

Introduction

Studies on the prevalence and spatial distribution of diseases of wild marine fish are for many years a component of national monitoring programmes in the ICES area aiming at an assessment of the quality of the marine environment. More recently, studies on externally visible fish diseases have also been incorporated in the suite of techniques recommended for biological

effects monitoring under the OSPAR Joint Assessment and Monitoring Programme (JAMP). Regular fish disease surveys are carried out according to standardised and intercalibrated ICES methodologies established through the work of the ICES Working Group on Pathology and Diseases of Marine Organisms (WGPDMO) (Dethlefsen *et al.* 1986, ICES 1989, Bucke *et al.* 1996, Lang and Dethlefsen 1996, Lang and Mellergaard 1999, Wosniok *et al.* 1999, Lang and Wosniok 2000).

Disease prevalence data generated within these programmes are submitted by ICES Member Countries to the ICES Environmental Data Centre and are regularly assessed by the WGPDMO. Standard procedures for data reporting, submission, validation and subsequent statistical analysis have been developed and successfully applied over the past years (Wosniok *et al.* 1999, Lang and Wosniok 2000). So far, the ICES fish disease data consist of information on externally visible diseases of the common dab (*Limanda limanda*) and the European flounder (*Platichthys flesus*) from the North Sea and adjacent areas, including the Irish Sea, English Channel, Baltic Sea, and some distant reference areas (e.g., Icelandic waters). Disease data are available from approximately 425.00 specimens, covering a time span of almost two decades for certain areas (e.g. German Bight, Dogger Bank). The majority of data stems from studies on diseases of dab (approx. 400.000 specimens), considerably less data are available for flounder (approx. 26.000 specimens) (Wosniok *et al.* 1999). It is envisaged, however, that the databank will be extended in the future to cover larger geographical areas and other fish species examined for diseases.

In a recent statistical analysis (Wosniok *et al.* 1999), temporal trends in the prevalence of three externally visible diseases, namely lymphocystis (dab and flounder), epidermal hyperplasia/papilloma (dab) and acute/healing skin ulcers (dab and flounder) have been established for areas in the North Sea and the western Baltic Sea. Marked spatial differences were identified with respect to both the absolute levels and the temporal (including seasonal) changes in the disease prevalence, helping to identify areas of concern which differ from other areas and which were, therefore, considered to deserve particular attention in the future. The authors emphasised, however, that the results of the analysis did not provide any information on possible natural and/or anthropogenic causes of the observed trends and that, therefore, a more integrated holistic data analysis should be carried out aiming at an assessment of possible cause-effect relationships between the disease prevalence and a range of environmental parameters known or suspected to be involved in the disease aetiology.

As a consequent next step, the WGPDMO elaborated an overview on data available in the different ICES databanks (ICES Environmental Data Centre, ICES Oceanography Data Centre and ICES Fishery Databanks) which were considered to be useful for such a holistic data analysis. A pilot study was undertaken subsequently, using a subset of data extracted from the ICES databanks for a multivariate statistical analysis on the relationship between disease prevalences and potentially explanatory abiotic and biotic factors in an area in the south-eastern North Sea including the German Bight. Although some shortcomings were identified in the analysis (e.g., a striking lack of ICES data for certain parameters including contaminants, methodological problems with the interpolation of data required for temporal trend analysis, correlations between certain parameters), the results of the pilot study were considered promising since, for a number of parameters included in the analysis, a close relationship with temporal variation in the disease prevalence could be identified worth to be investigated in greater detail (Lang and Wosniok 2000).

The present paper provides an updated overview of the data available in the ICES databanks which could be used for an integrated holistic analysis of the ICES fish disease data. It further presents information on results derived from an extended statistical analysis involving three

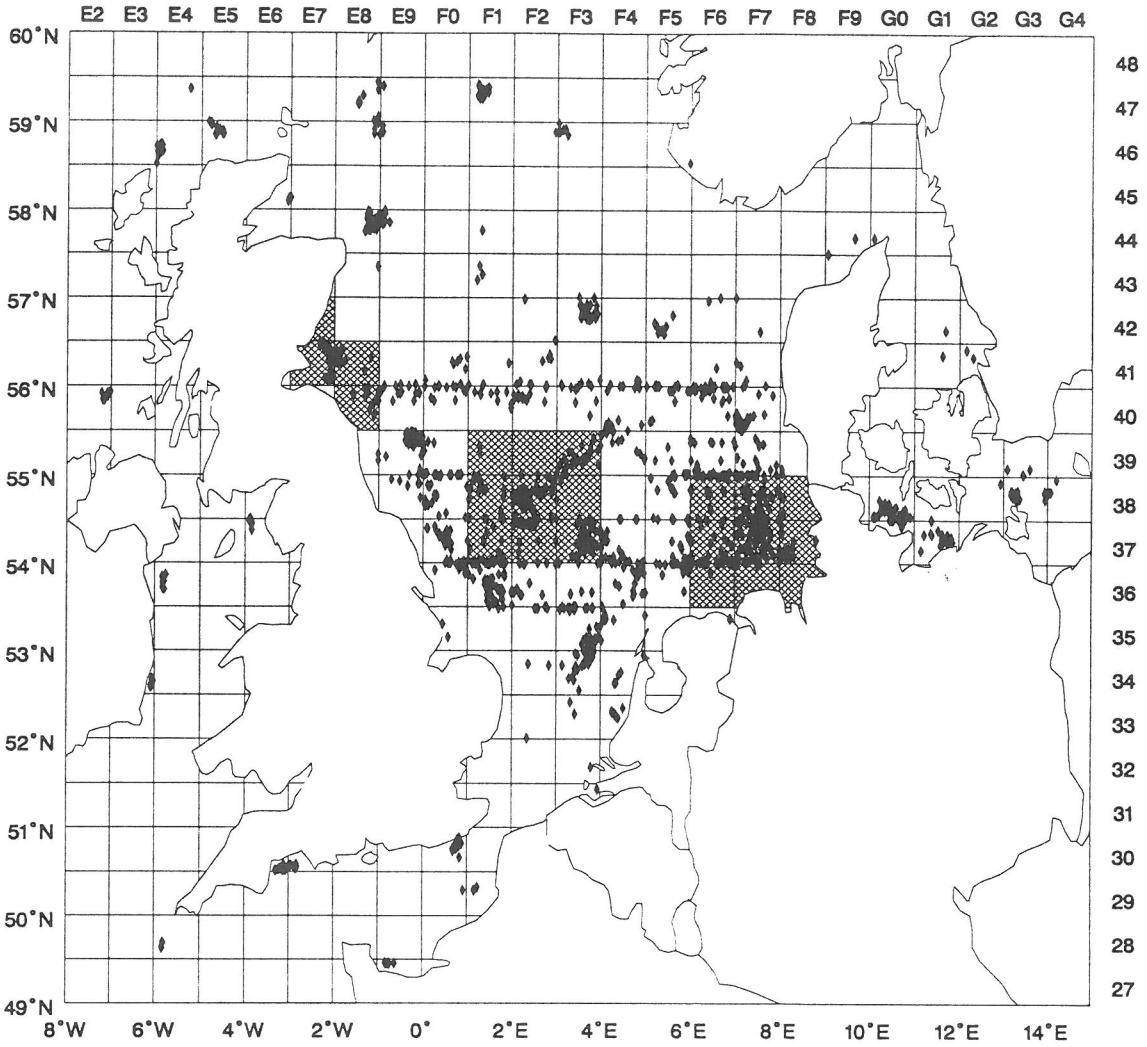

Figure 1: Location of the areas (grey shading) used for the statistical analysis of the ICES data on diseases of North Sea dab (*Limanda limanda*) in relation to contaminants and other environmental factors (Area 1: south-eastern North Sea, Area 2: central North Sea, Area 3: North-western North Sea). Each diamond indicates one sample of dab examined for diseases in the period 1981-1999.

geographical areas in the south-eastern, central and north-western North Sea, which were selected on the basis of the availability of dab disease data. New statistical procedures were applied in order to investigate the relationship between the disease prevalence and parameters for which data were extracted from the ICES Databanks and to overcome methodological problems identified in earlier data analyses. Particular emphasis was given to the assessment of the role of environmental contaminants as potential explanatory factors for the disease prevalence. However, effects of other factors such as water temperature, salinity, oxygen content, nutrients in water, and fish density as derived from data on catch per unit effort (CPUE) for dab were also analysed.

Material and Methods

Figure 1 shows the geographical areas used for the statistical analysis. Areas were selected for which a considerable amount of disease data is available and which differ both in the absolute disease prevalences and the temporal trends recorded over the past years (Lang and Wosniok 2000). The availability of disease data is marked in the figure by diamonds, each of which represents one sample of fish examined for diseases. In order to obtain a sufficient amount of data, large areas were selected, consisting of 4-9 ICES Statistical Rectangles. Area 1 was located in the south-eastern North Sea and included the German Bight, Area 2 in the central North Sea included the Dogger Bank, and Area 3 in the north-western North Sea included the Firth of Forth region

For comparative purposes and due to their abundance in the areas covered, only disease data for female dab, 20-24 cm total length, were used. The analysis focused on prevalence data for externally visible diseases of dab (lymphocystis, epidermal hyperplasia, acute/healing skin ulcerations), which were quantified and reported to the ICES Environmental Data Centre by ICES Member Countries (Denmark, Germany, The Netherlands, U.K.) according to ICES guidelines for fish disease surveys (ICES 1989, Bucke *et al.* 1996).

For the statistical analysis, a logistic model (McCullagh and Nelder 1989) was used to describe the relationship between the prevalence of fish diseases (lymphocystis, epidermal hyperplasia/papilloma, acute/healing skin ulcerations) and potential explaining factors. If the observation of an explaining factor was missing for the data of a fish disease observation, an estimate for the required value together with its standard error was obtained by interpolating within the time series of factor values. No interpolation was done in gaps of more than three years of length, also no extrapolation outside the observed time range of a factor was done. Interpolation was performed by a Gaussian kernel smoother, using generalised crossvalidation (Hastie and Tibshirani, 1990) to determine the smoothing parameter. A bootstrap procedure (Efron and Tibshirani, 1993, Hall, 1992) was used to account for the effect of using interpolated values. Here, modified replicates of the original data set were generated by adding a normally distributed error term to each interpolated value. The error term had a mean of zero and a standard deviation equal to the local interpolation standard error. For each replicate data set a logistic analysis was calculated. Two thousand replicates were found to produce stable estimates, however, for safety reasons, 4000 replicates were used. Significance levels for the estimated parameters (the coefficients in the logistic model) were obtained from the empirical distributions of the estimates.

As many of the potential explaining factors had not been observed over the whole time range, a joint or joint stepwise analysis of all factors was not possible. Factors with very short observation ranges (less than 3 years), very few (less than 4) observations, or with gaps of more than 3 years length within the observation series were generally excluded from the multivariate analysis. The remaining factors were grouped according to the length of observation periods and categories. This led to groups containing hydrographic and nutrient quantities, CPUE, contaminants/heavy metals in sea water, sediment, blue mussel (*Mytilus edulis*) tissue, dab liver and muscle. Within each category of factors a stepwise backwards procedure was used to eliminate non-significant factors. The first group to be considered was the group of hydrographic/nutrient quantities, which had the longest observation period. The factors identified in the first group as having significant influence on the disease prevalence were incorporated in all subsequent estimation steps, throughout using the initially obtained model coefficients as offset terms in the logistic model. This procedure was repeated in an analogue way when considering the next groups. In this way

use was made of all available data, and information derived from long data series could be exploited also when assessing the impact of a factor which had been recorded only during a relatively short period.

The method use here to estimate model coefficients from observation periods of different lengths requires a specific interpretation of the coefficients, as can most easily be illustrated by an example. Suppose that two slightly correlated factors, denoted by X and Y, both influence the disease prevalence, and that factor X was observed over the whole study period, while factor Y was observed only during the second half of the period. Then, due to the correlation and missing observed data for Y, a part of the Y effect will be attributed to X when estimating the effect of X over the whole period. When estimating the Y effect from the second half of the observation period, it will be most likely be wrongly estimated, as part of it is already incorporated in the offset term for X. Hence the estimated Y coefficient will not have the correct size, where the deviation from the correct size depends on the type of correlation (positive, negative), on the forms which the time series of X and Y take, and, of course, on the true size of the effects. Even the sign of the estimated Y coefficient could be reversed, so that generally for estimated logistic coefficients for factors with small observation periods the main information is given by the p value of the coefficient, not by the value itself or its sign. Fortunately, if, different from the scenario discussed so far, the second factor, Y, has no effect on the target quantity or is not correlated with X, there is no danger that a significant p value for Y arises erroneously.

Results

Figures 2 a-c provide an overview of data available in the ICES Environmental Data Centre, the ICES Oceanographic Data Centre and Fishery Databanks which were due to their nature considered to be appropriate for an analysis of the relationship between the fish diseases prevalences and potentially explaining environmental factors. The overview is given for the period 1981-1999, the time span for which disease data are available (exception Area 3, in which the data series started in 1984). A list explaining the abbreviations used in Figures 2 a-c, 3 a-c and in Table 1 is provided in Table 2.

From the diagrams it can bee seen that there is a considerable lack of ICES data for certain parameters. In particular, there is only little data on contaminants in water, sediment and biota which cover only relatively short periods of time. Considerably more information is available on parameters like oceanographic, nutrient, CPUE and disease data. Most data are available for Area 1 (south-eastern North Sea), followed by Area 2 (central North Sea) and Area 3 (north-western North Sea).

Table 1 summarises the results of the multivariate analysis described above by geographical area (Areas 1-3) and disease (lymphocystis, epidermal hyperplasia/papilloma, acute/healing skin ulcerations). Only those explaining factors are included in the table, for which a significant relationship with the temporal change in the prevalence of at least one of the three diseases was detected. The degree of significance is indicated by the p value. For the reasons explained above under Material and Methods, the direction of the effects, i.e. whether there was a positive or a negative relationship between the prevalence and the observed or interpolated factor values, is only given for those factors which covered the complete time range for which data was available. The time range and number of data points incorporated in the analysis is also shown in the table for each factor.

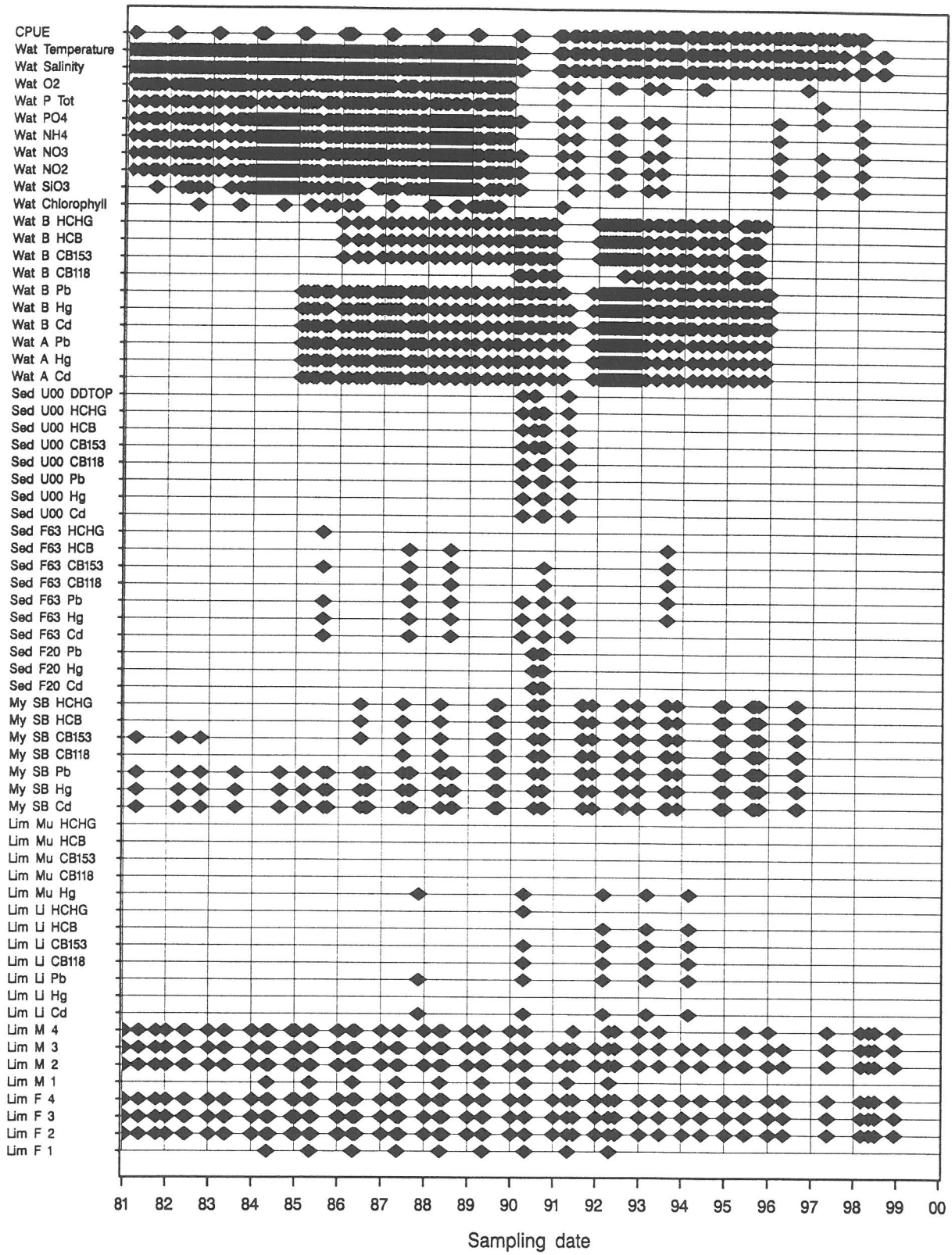

Figure 2a: Available data in Area 1 (cf. Figure 1). Each diamond indicates one data sample. The meaning of the abbreviations on the vertical axis is given in Table 2.

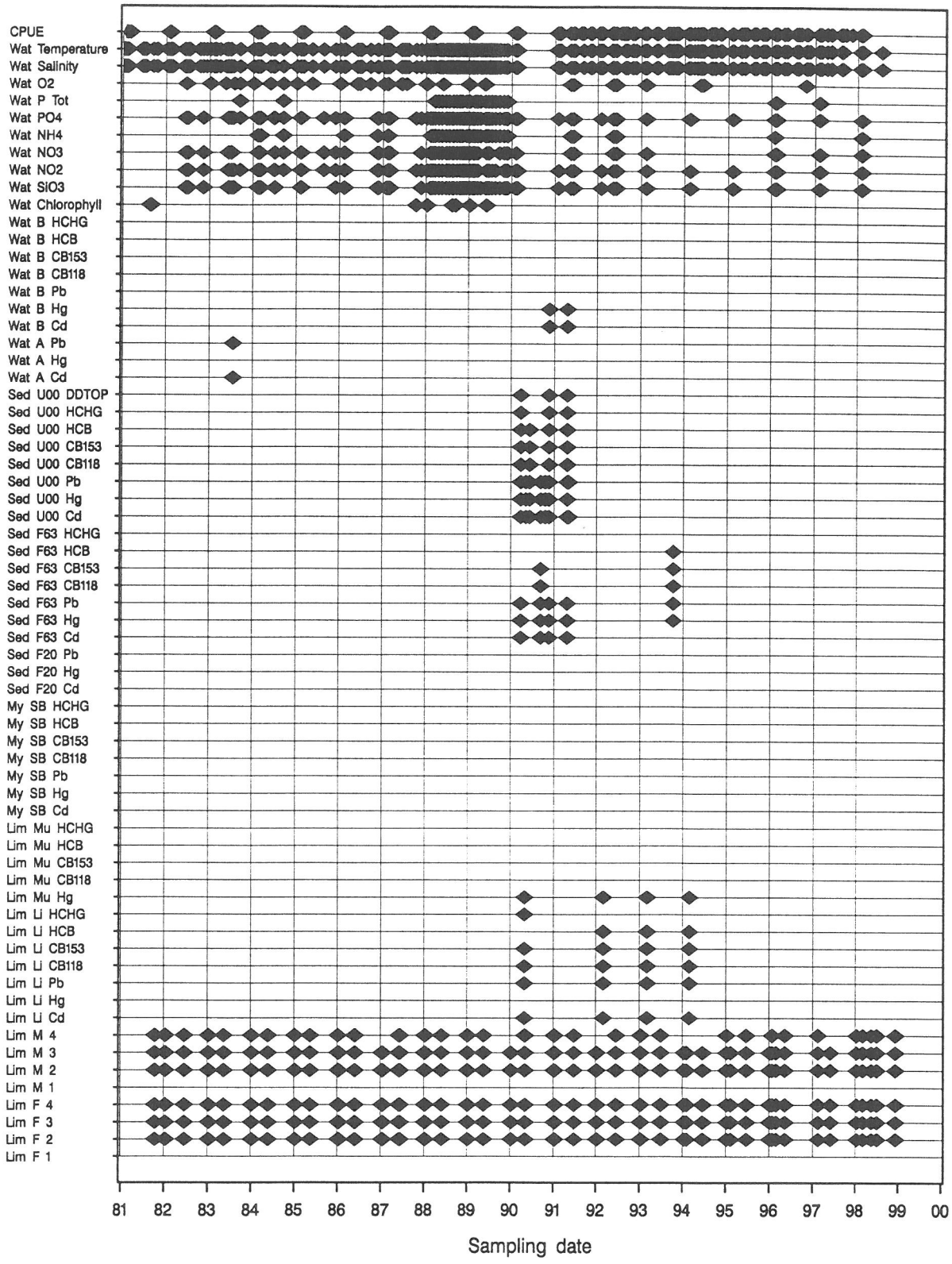

Figure 2b: Available data in Area 2 (cf. Figure 1). Each diamond indicates one data sample. The meaning of the abbreviations on the vertical axis is given in Table 2.

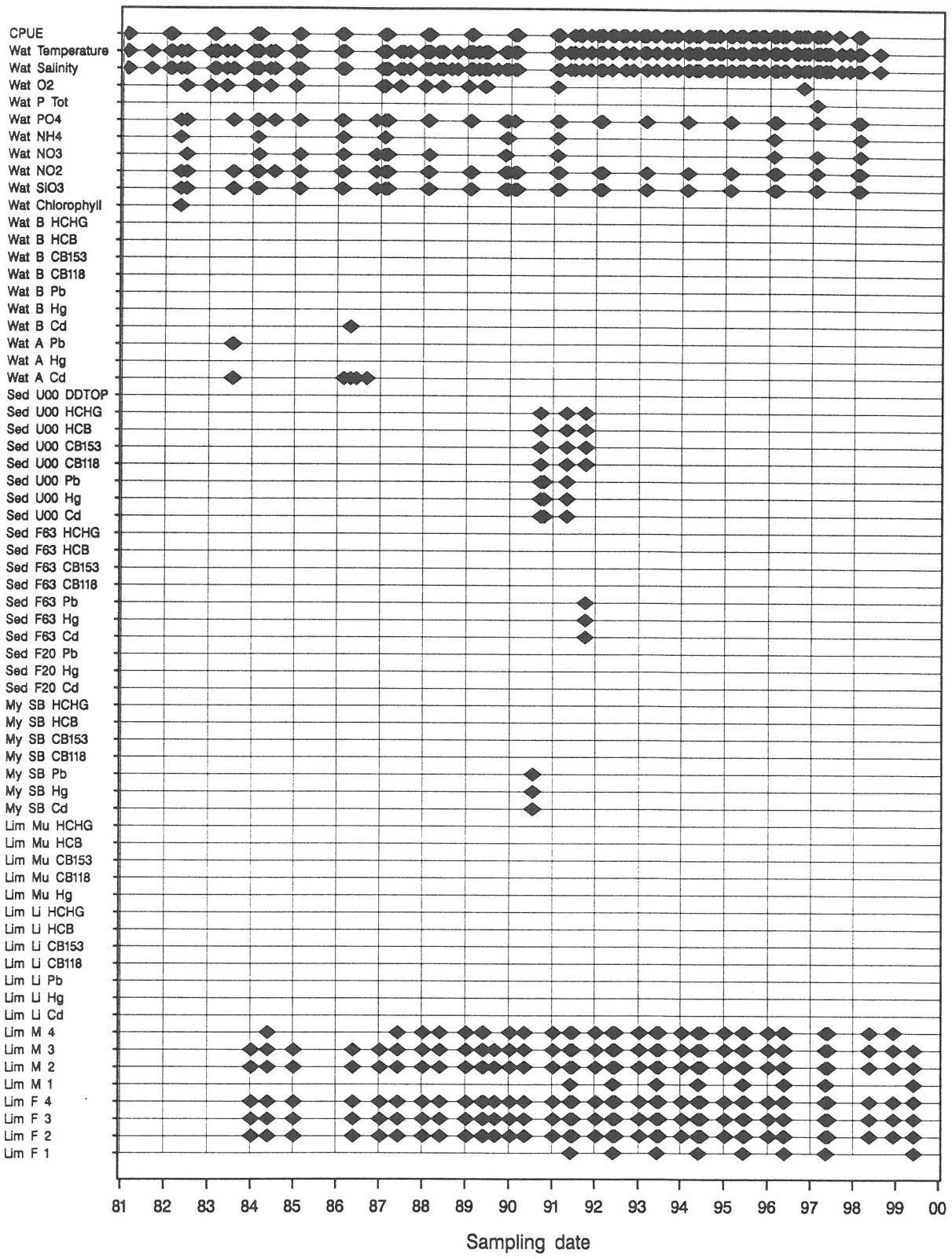

Figure 2c: Available data in Area 3 (cf. Figure 1). Each diamond indicates one data sample. The meaning of the abbreviations on the vertical axis is given in Table 2.

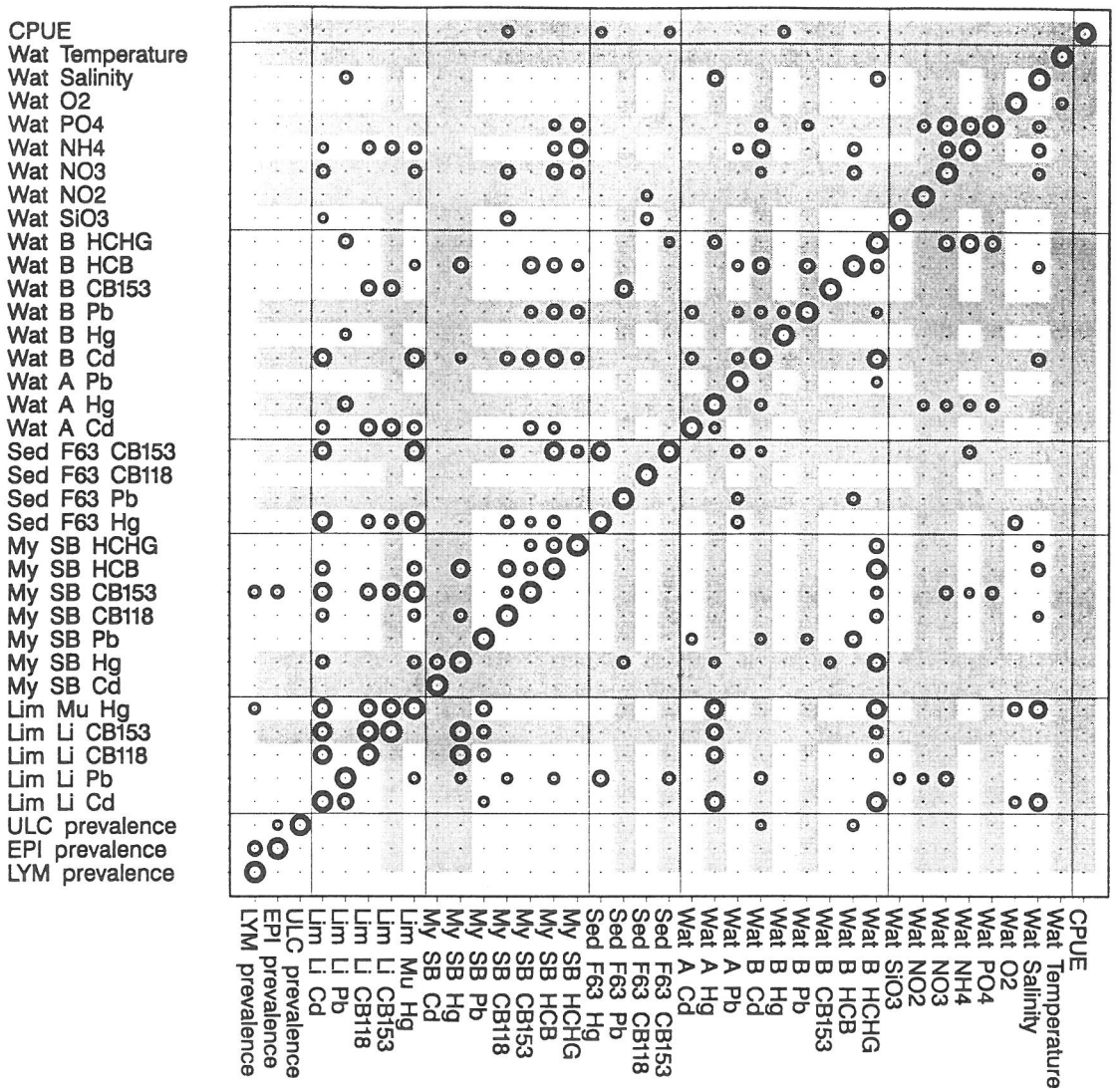

Figure 3a: Pearson correlations between factors in Area 1 (see Figure 1). Factor abbreviations are explained in Table 2. The outer diameter of the rings is proportional to the absolute correlation size, rings on the diagonal correspond to a correlation of one. Positive correlations are shown above the diagonal, negative ones below. Only significant correlations (α=0.05) with size > 0.50 are displayed. Shaded areas indicate factors which have been found to affect one or more disease prevalences significantly (see Table 1).

For Area 1, a variety of significant relationships was detected, partly reflecting the fact that here data for many, though by no means for all, parameters is available (see Figure 2 a). There was a consistent seasonal effect in that the prevalence of all three diseases was significantly lower in Season 2 (October-March) as compared to Season 1 (April-September). Another consistent feature affecting all diseases was the negative effect of NO_3. Water temperature, PO_4, NO_2 and CPUE were significantly related to two of the diseases, all except CPUE showed either a positive or negative effect. Contaminants in water, sediment, blue mussel tissue and dab liver were only related to one of the diseases each, in most cases to epidermal hyperplasia/papilloma.

In Area 2, only CPUE was related to all diseases, however, not in a consistent way. For lymphocystis and epidermal hyperplasia/papilloma a positive relationship was found, whilst the effect of CPUE was negative for acute/healing skin ulcerations. Silicate levels were positively related to the prevalence of lymphocystis and epidermal hyperplasia/papilloma, salinity was negatively related to lymphocystis, but positively to acute/healing skin ulcerations. The other

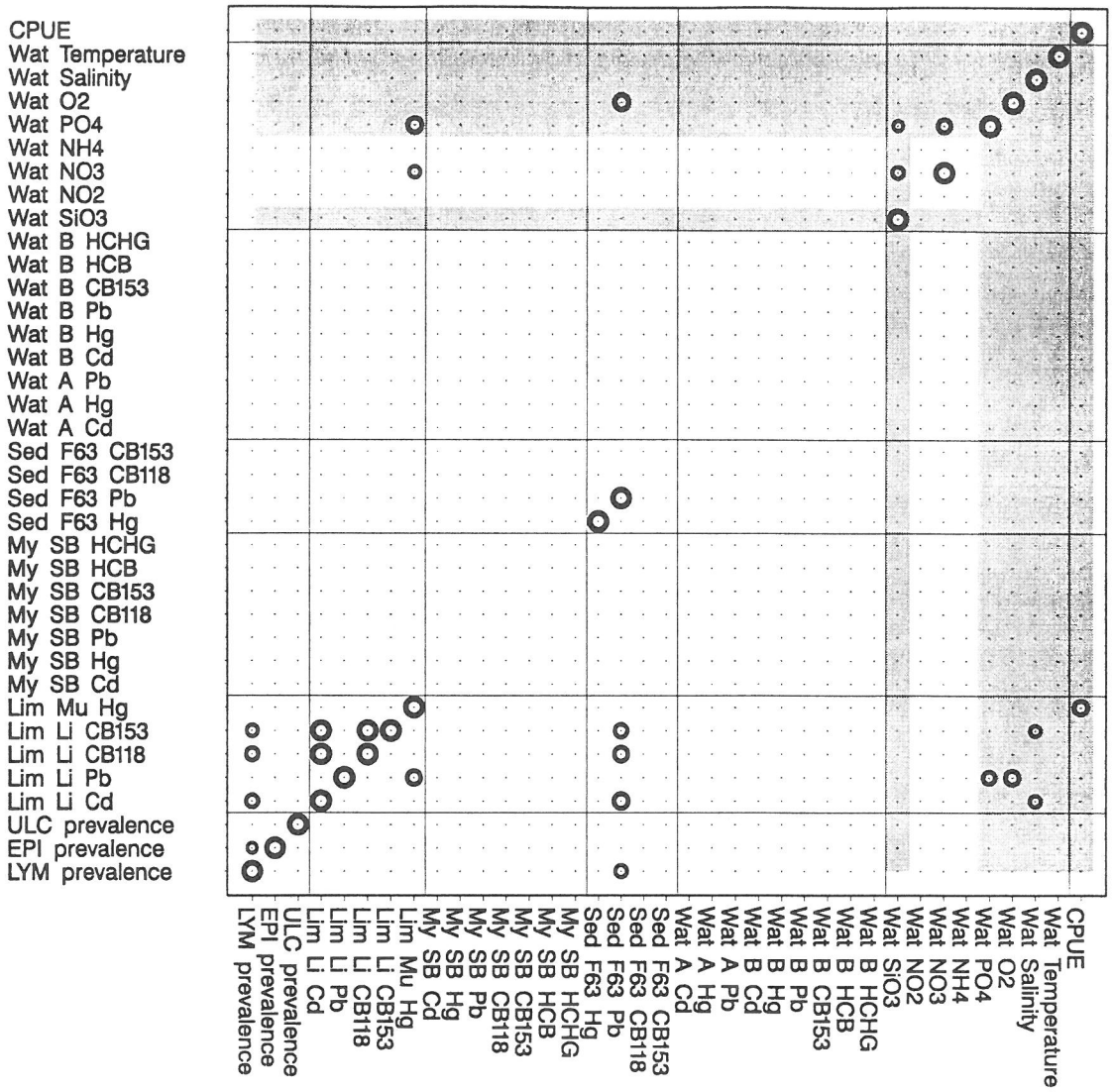

Figure 3b: Pearson correlations between factors in Area 2 (see Figure 1). Factor abbreviations are explained in Table 2. The outer diameter of the rings is proportional to the absolute correlation size, rings on the diagonal correspond to a correlation of one. Missing entries on the diagonal denote lack of data for the factor in that row/column. Positive correlations are shown above the diagonal, negative ones below. Only significant correlations (α=0.05) with size > 0.50 are displayed. Shaded areas indicate factors which have been found to affect one or more disease prevalences significantly (see Table 1).

significant explaining factors were only related to one of the diseases. The only contaminant with a significant relationship to the disease prevalence was CB 153 in dab liver tissue. However, this had to do with the lack of appropriate contaminant data in the ICES Environmental Data Centre.

Only three explaining factors were significantly related to the disease prevalence in Area 3. As in Area 1, a strong and consistent seasonal effect was evident. However, in contrast to Area 1, the prevalences of all three diseases were higher in Season 2 (October-March) than in Season 1 (April-September). Water temperature was positively related to the prevalence of lymphocystis and epidermal hyperplasia/papilloma, for lymphocystis again contrasting the findings for Area 1. Largely due to the lack of appropriate data, no relationships with contaminants were identified.

In the course of the analysis it became obvious that many of the potentially explaining factors under consideration were highly correlated. However, correlations among the explaining factors can lead to ambiguous or erroneous conclusions, since existing relationships between single

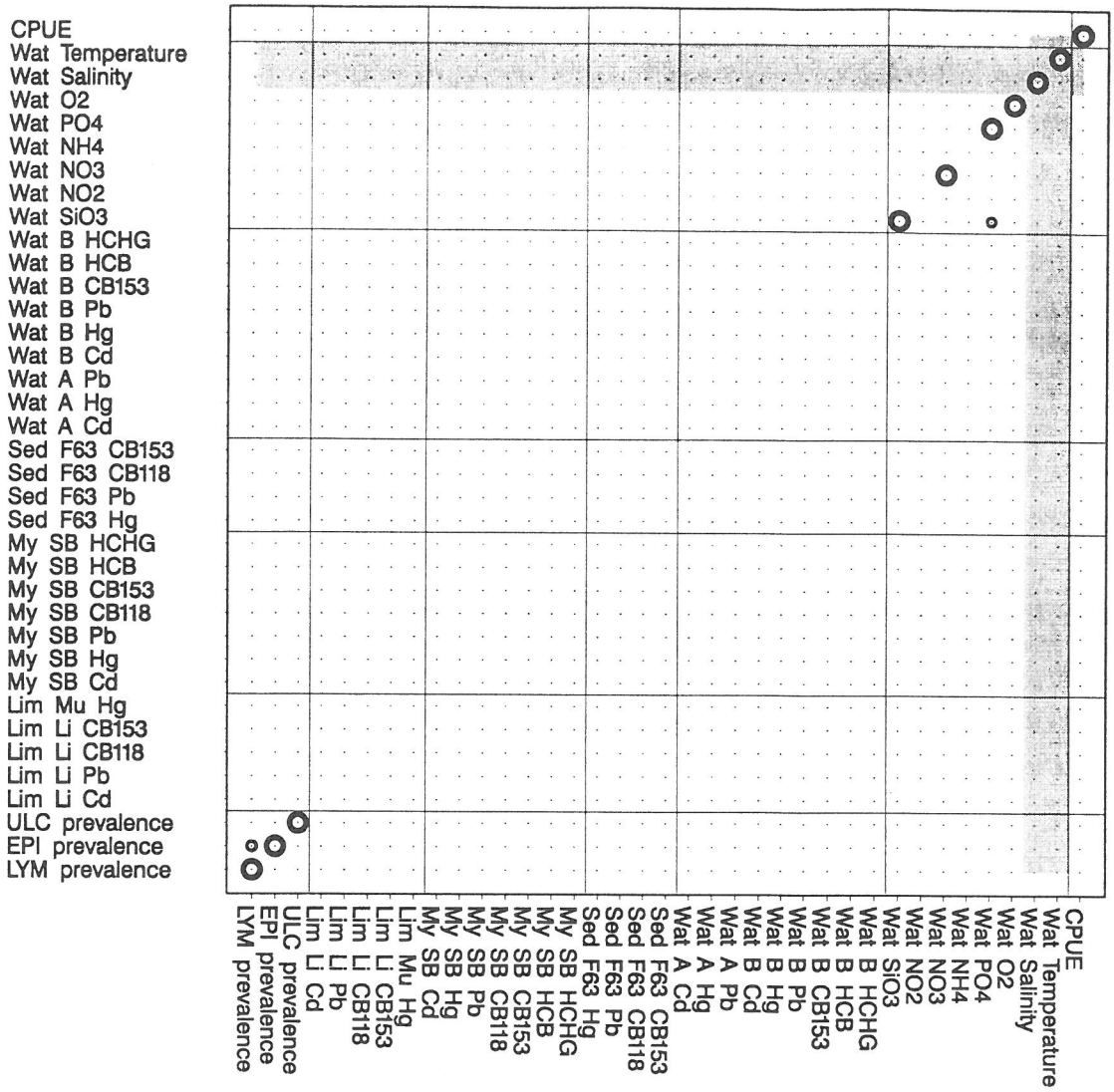

Figure 3c: Pearson correlations between factors in Area 3 (see Figure 1). Factor abbreviations are explained in Table 2. The outer diameter of the rings is proportional to the absolute correlation size, rings on the diagonal correspond to a correlation of one. Missing entries on the diagonal denote lack of data for the factor in that row/column. Positive correlations are shown above the diagonal, negative ones below. Only significant correlations (α=0.05) with size > 0.50 are displayed. Shaded areas indicate factors which have been found to affect one or more disease prevalences significantly (see Table 1).

factors and the disease prevalence might be obscured if instead of the effective factor another one is included in the model which is only statistically but not causally related to the effective factor. In such a case, only one, not both factors will enter the model, because the additional inclusion of the second one would not improve the model fit sufficiently strong to justify the inclusion. Significant correlations between potentially explaining factors and between the disease prevalences themselves are shown as an interpretation aid for the three areas (Figures 3 a-c). When assessing the results of the analysis given in Table 1, it has to be considered that factors significantly correlated with those factors appearing in the full models could also be candidates for the explanation of the disease prevalences. It should be noted that the usual way do deal with correlated explaining factors, e.g. the transformation into principal components or similar quantities with uncorrelated structures is not feasible here, as these methods need full observations for all data records, without missing values.

Discussion

The results of the multivariate analysis revealed a number of significant relationships between the prevalence of the three diseases of North Sea dab studied and potentially explaining environmental factors. Within areas, a few factors were identified with a more or less consistent effect on the prevalence of at least two of the diseases. However, none of these factors exerted the same effect in all areas, not even for one of the diseases. That means, that, from the present analysis, there is no clear indication for the existence of a single or a few underlying factors that drive the disease prevalence in all areas in the same way. Among other explanations, this can be attributed to the following:

- The disease prevalence can be affected by various factors inducing the same change in prevalence (concept of a multifactorial disease aetiology/pathogenesis).

- The effect of factors might differ between areas (e.g., contaminants may be present, but levels in some areas may be below a toxicologically relevant threshold concentration, fish may have different tolerance levels due to accommodation, genetic predisposition etc.)

- The availability of data differs between areas (e.g., if contaminant data is not available, no effects can be attributed to contaminants).

Another aspect that has to be considered when interpreting the findings of the present analysis is the correlation between the potentially explaining factors (see Figures 3 a-c). Examples are the correlations identified in Areas 1 and 2 between nutrients in seawater and concentrations of anorganic and organic contaminants in water, sediments and biota. If strong correlations between factors suspected to cause changes in disease prevalence are present, it is impossible to identify the 'really' responsible factor. This can be considered as a general problem in studies on relationships in ecosystems where factors under study cannot be adjusted according to a balanced sampling design.

Figures 2 a-c illustrate the considerable lack of ICES data. It is self-evident that a lack of data hinders a statistical analysis. Even if data interpolation is considered feasible and if interpolated values are assigned an additional variance by means of a bootstrap procedure, they cannot replace real observed values and, therefore, may introduce a considerable bias and lead to misleading results and interpretation. A lack of data was particularly evident for ICES data on contaminants in sediments and biota, making it almost impossible to investigate relationships between contaminants and fish diseases. Furthermore, apart form the long-term data on fish diseases, there is currently only little data on other biological effects considered to be associated with exposure to contaminants available in the ICES Environmental Data Centre.

If the ICES Environmental Data Centre is to be used as data source for internationally co-ordinated assessments on biological effects of contaminants, much more historic and current data on contaminants and their effects need to be incorporated. There is no doubt that such data are available in national databanks maintained by ICES Member Countries. However, the countries have to be convinced to submit their data, provided that they have been generated according to internationally agreed quality assurance procedures.

The availability of a more comprehensive data set would improve the spatial and temporal data coverage and would very likely facilitate an analysis based on a smaller geographical scale than that used in the present analysis. The use of such large areas creates problems since conditions are normally not the same over the entire area. For instance, Area 1 in the south-eastern North

Sea includes estuarine, coastal and offshore regions the conditions of which are quite different. Additional data would, furthermore, minimise the need for temporal interpolation and would improve the power of any analysis. Also, the chance of randomly occurring high correlations among potentially explaining factors is high when the data series are short, which in turn means that by increasing the amount of data there is a better perspective to identify the 'real' relationships.

Literature cited

BUCKE, D., VETHAAK, A.D., LANG, T. and MELLERGAARD, S. (1996). Common diseases and parasites of fish in the North Atlantic: Training guide for identification. ICES Techniques in Marine Environmental Sciences 19, 27 pp.

DETHLEFSEN, V., EGIDIUS, E. and McVICAR, A.H. (1986). Methodology of fish disease surveys. Report of an ICES Sea-going Workshop held on RV ‚Anton Dohrn' 3-12 January 1984. ICES Cooperative Research Report 140, 33 pp.

EFRON, B. and TIBSHIRANI, R.J. (1993). An introduction to the bootstrap. Chapman & Hall, New York.

HALL, P. (1992). The bootstrap and Edgeworth expansion. Springer-Verlag, New York.

HASTIE, T.J. and TIBSHIRANI, R.J. (1990). Generalized additive models. Chapman & Hall, London.

ICES 1989. Methodology of fish disease surveys. Report of an ICES Sea-going Workshop held on RV U/F 'Argos', 16-23 April 1988. ICES Cooperative Research Report, 166, 33 pp.

LANG, T. and DETHLEFSEN, V. (1996). Fish disease monitoring - a valuable tool for pollution assessment ? ICES C.M. 1996/E:17, 18 pp.

LANG, T. and MELLERGAARD, S. (1999). The BMB/ICES Sea-Going Workshop ‚Fish Diseases and Parasites in the Baltic Sea' – introduction and conclusions. ICES Journal of Marine Science 56: 129-133.

LANG, T. and WOSNIOK, W. (2000). ICES data for an holistic analysis of fish disease prevalence. *In*: ICES 2000, Report of the ICES Advisory Committee on the Marine Environment, 1999. ICES Coop. Res. Rep. 239: 193-209.

McCULLAGH, P. and NELDER, J.A. (1989). Generalized linear models. 2nd edition, chapman & Hall, London.

WOSNIOK, W., LANG, T., VETHAAK, A.D., des CLERS, S., MELLERGAARD, S. (1999). Statistical analysis of fish disease prevalence data from the ICES Environmental Data Centre. *In*: ICES 1999, Report of the ICES Advisory Committee on the Marine Environment 1998, ICES Coop. Res. Rep. 233: 297-327.

	Area	Parameter	Disease						Number of data points	Time range
			Lymphocystis		Epidermal hyperplasia / papilloma		Acute/healing ulcerations			
			p value	Direction of effect	p value	Direction of effect	p value	Direction of effect		
	Area 1: south-eastern North Sea									
1		Season 2	< 0.0003	-	< 0.0003	-	0.0010	-	44	01/81-05/97
2		Temperature	< 0.0003	-	< 0.0003	-			44	01/81-05/97
3		PO4	0.0095	-	0.0015	-			44	01/81-05/97
4		NO2	< 0.0003	+	< 0.0003	+			44	01/81-05/97
5		NO3	0.0005	-	< 0.0003	-	0.0010	-	44	01/81-05/97
6		Wat A Hg	0.0040						26	05/85-06/96
7		Wat B Cd					0.0110		26	05/85-06/96
8		Wat B Pb			< 0.0003				26	05/85-06/96
9		Sed F63 CB153			0.0115				22	12/85-06/93
10		Sed F63 Pb			0.0100				22	12/85-06/93
11		My Cd			< 0.0003				42	05/81-05/97
12		My Hg			< 0.0003				42	05/81-05/97
13		My HCB	0.0055						42	05/81-05/97
14		Lim Li CB153			0.0010				9	05/90-06/93
15		CPUE	0.0095	+			0.0150	-	43	05/81-05/97
	Area 2: central North Sea									
1		Season 2					< 0.0003	-	43	10/81-01/98
2		Temperature					< 0.0003	+	43	10/81-01/98
3		PO4	< 0.0003	-			0.0035	+	43	10/81-01/98
4		SiO4	< 0.0003	+	< 0.0003	+			43	10/81-01/98
5		O2			0.0010	+			43	10/81-01/98
6		Salinity	0.0500	-			0.0035	+	43	10/81-01/98
7		Lim Li CB118					<0.0003		10	08/90-01/94
8		Lim Li CB153	< 0.0003						10	08/90-01/94
9		CPUE	< 0.0003	+	< 0.0003	+	0.0305	-	43	10/81-01/98
	Area 3: north-western North Sea									
1		Season 2	< 0.0003	+	0.0095	+	0.0028	+	35	01/84-05/98
2		Temperature	< 0.0003	+			< 0.0003	+	35	01/84-05/98
3		Salinity			0.0290	-			35	01/84-05/98

Table 1: Results of the multivariate statistical analysis of the ICES data on diseases of North Sea dab (*Limanda limanda*) in relation to contaminants and other potentially explaining environmental factors.

Season 2	Winter season, October - March
CPUE	Catch per unit effort
Wat Temperature	Water temperature in °C
Wat Salinity	Salinity
Wat O2	Dissolved oxygen in sea water
Wat P Tot	Total phosphorus in sea water
Wat PO4	Phosphate in sea water
Wat NH4	Ammonium in sea water
Wat NO3	Nitrate in sea water
Wat NO2	Nitrite in sea water
Wat SiO4	Silicate in sea water
Wat Chlorophyll	Chlorophyll-α in sea water
Wat B HCHG	γ-HCH (lindane) in sea water, before filtration
Wat B HCB	Hexachlorobenzene in sea water, before filtration
Wat B CB153	Chlorbiphenyl 153 in sea water, before filtration
Wat B CB118	Chlorbiphenyl 118 in sea water, before filtration
Wat B Pb	Lead in sea water, before filtration
Wat B Hg	Mercury in sea water, before filtration
Wat B Cd	Cadmium in sea water, before filtration
Wat A Pb	Lead in sea water, after filtration
Wat A Hg	Mercury in sea water, after filtration
Wat A Cd	Cadmium in sea water, after filtration
Sed U00 DDTOP	DDT (o,p') in sediment, undefined fraction
Sed U00 HCHG	γ-HCH (lindane) in sediment, undefined fraction
Sed U00 HCB	Hexachlorobenzene in sediment, undefined fraction
Sed U00 CB153	Chlorbiphenyl 153 in sediment, undefined fraction
Sed U00 CB118	Chlorbiphenyl 118 in sediment, undefined fraction
Sed U00 Pb	Lead in sediment, undefined fraction
Sed U00 Hg	Mercury in sediment, undefined fraction
Sed U00 Cd	Cadmium in sediment, undefined fraction
Sed F63 HCHG	γ-HCH (lindane) in sediment, fraction < 63µ
Sed F63 HCB	Hexachlorobenzene in sediment, fraction < 63µ
Sed F63 CB153	Chlorbiphenyl 153 in sediment, fraction < 63µ
Sed F63 CB118	Chlorbiphenyl 118 in sediment, fraction < 63µ
Sed F63 Pb	Lead in sediment, fraction < 63µ
Sed F63 Hg	Mercury in sediment, fraction < 63µ
Sed F63 Cd	Cadmium in sediment, fraction < 63µ
Sed F20 Pb	Lead in sediment, fraction < 20µ
Sed F20 Hg	Mercury in sediment, fraction < 20µ
Sed F20 Cd	Cadmium in sediment, fraction < 20µ
My SB HCHG	γ-HCH (lindane) in soft body of *Mytilus edulis*
My SB HCB	Hexachlorobenzene in soft body of *Mytilus edulis*
My SB CB153	Chlorbiphenyl 153 in soft body of *Mytilus edulis*
My SB CB118	Chlorbiphenyl 118 in soft body of *Mytilus edulis*
My SB Pb	Lead in soft body of *Mytilus edulis*
My SB Hg	Mercury in soft body of *Mytilus edulis*
My SB Cd	Cadmium in soft body of *Mytilus edulis*
Lim Mu HCHG	γ-HCH (lindane) in *Limanda limanda* muscle
Lim Mu HCB	Hexachlorobenzene in *Limanda limanda* muscle
Lim Mu CB153	Chlorbiphenyl 153 in *Limanda limanda* muscle
Lim Mu CB118	Chlorbiphenyl 118 in *Limanda limanda* muscle
Lim Mu Hg	Mercury in *Limanda limanda* muscle
Lim Li HCHG	γ-HCH (lindane) in *Limanda limanda* liver
Lim Li HCB	Hexachlorobenzene in *Limanda limanda* liver
Lim Li CB153	Chlorbiphenyl 153 in *Limanda limanda* liver
Lim Li CB118	Chlorbiphenyl 118 in *Limanda limanda* liver
Lim Li Pb	Lead in *Limanda limanda* liver
Lim Li Hg	Mercury in *Limanda limanda* liver
Lim Li Cd	Cadmium in *Limanda limanda* liver
Lim M 4	disease prevalence for male *Limanda limanda* with length ≥ 25 cm
Lim M 3	disease prevalence for male *Limanda limanda* with length 20-24 cm
Lim M 2	disease prevalence for male *Limanda limanda* with length 15-19 cm
Lim M 1	disease prevalence for male *Limanda limanda* with length ≤ 19 cm
Lim F 4	disease prevalence for female *Limanda limanda* with length ≥ 25 cm
Lim F 3	disease prevalence for female *Limanda limanda* with length 20-24 cm
Lim F 2	disease prevalence for female *Limanda limanda* with length 15-19 cm
Lim F 1	disease prevalence for female *Limanda limanda* with length ≤ 19 cm

Table 2: Abbreviations used in Figures and Tables. Abbreviations appear in the same order as in Figures 2 and 3.