

Scottish Horticultural Research Institute,
Invergowrie, Dundee, Scotland

DATA BASE MANAGEMENT FOR NEMATOLOGICAL SURVEYS

by

D. J. F. BROWN, P. B. TOPHAM and C. E. TAYLOR

A collaborative research project between Professor F. Lamberti, Bari, Italy and Dr. C. E. Taylor, Invergowrie, Dundee, Scotland, financed by the North Atlantic Treaty Organisation (NATO grant No. 424) began in 1970. As part of this project a survey was completed in both countries to determine the geographical distribution of species of the family Longidoridae. The abiotic and biotic factors possibly influencing these distribution patterns were investigated for each species. The results of the survey in Britain have been published (Taylor and Brown, 1976) and data from the survey in Italy are being analysed.

It was realised that to store, retrieve and finally to analyse the data obtained in a survey of the Longidoridae in Britain would require the utilisation of computer systems and techniques. The computer systems used and developed during the British survey (Topham *et al.*, 1975) have been modified to allow for their more efficient and economical use at the micro or macro level. At macro level the systems is presently being used to investigate the geographical distribution of the Longidoridae in Europe. The following is an outline of the procedures used in several British nematode surveys and which can be adapted for use in other countries, or used on a wider scale as for the European Plant Parasitic Nematode Survey (EPPNS).

Data acquisition

Data for the Longidoridae survey in Britain were obtained as records from existing or published sources and from soil samples

Scottish Horticultural Research Institute, Invergowrie, Dundee		
EUROPEAN PLANT PARASITIC NEMATODE SURVEY		
Collector:		Nema Sp. (per litre soil):
Collection No:		
Date:		
Grid Ref:		
Geo Code:		
UTM Grid Zone:		
Source: Sample Record		
Locality:		
Country:		

Vegetation	Comments
Virus/Disease/ Damage/Assoc.	
Altitude:	
Soil Type:	
Percentages sand:	
silt:	
clay:	
moisture:	
Organic matter:	
pH	

Fig. 1 - The European Plant Parasitic Nematode Survey (EPPNS) standard record card.

collected specifically for the survey. After the Longidoridae survey had commenced, a survey of nematodes associated with forest and woodland trees (Boag, 1974) was initiated and data obtained for each survey were shared providing both survey teams with additional data.

Records and soil samples can be obtained from numerous sources if instructions for the collection and recording of data are available for contributors. The minimum information required to constitute a single record or unit of data acquisition is a map grid reference, geocode, Universal Transverse Mercator (UTM) grid zone and a list of those nematode species found at that particular locality. This information allows a presence and absence dot distribution map to be prepared for the nematode species being investigated. Any record or sample containing this information set out in a methodical format can be used. However, a form of standardised input is preferable. A standardised record card (Fig. 1) has been prepared for use with the EPPNS (Brown *et al.*, 1978).

Upon receipt the accumulated data is coded for computer handling. For the EPPNS each species of *Longidorus*, *Paralongidorus* and *Xiphinema* has been assigned a numeric code. The biotic and abiotic data are also assigned numeric or alphanumeric codes based upon the scheme used for the British Longidoridae survey (Taylor and Brown, 1976) (Table I). These codes must be prepared with caution for, although combining categories by computer operations is straightforward, separating categories into two or more subcate-

Table I - *Computer codes used for data handling with the British Longidoridae Survey.*

Code	Primary vegetation type	Code	Soil type
GP	Permanent pasture	S---	Sand
DH	Deciduous woodland	L---	Loam
DH	Hedgerow	SI--	Silt
CW	Coniferous woodland	C---	Clay
MW	Mixed woodland	SL--	Sandy loam
SC	Scrubland	SICL	Silty clay loam
SM	Moorland		etc
A-	Arable		
AB	brassica	1.0 to 9.0	Soil pH
AC	cereals	0 to 100	Percentage soil moisture
A1	leys	0 to 100	Percentage sand
AS	soft fruit	0 to 100	Percentage silt
	etc	0 to 100	Percentage clay
	Virus association	0 to 9,999 m	Altitude
A---	Arabidopsis mosaic		
-S---	Strawberry latent ringspot		
--T-	Tomato black ring		
---R	Raspberry ringspot		

gories requires new data to be input to the computer for every pre-existing unit of information. It is expedient to assign codes to any category which may be distinguished at the time of input, accepting that species concepts may subsequently change so that species may lapse into synonymy or records may include a species *sensu lato* or *sensu stricto*.

All data are qualitative (presence or absence of nematode species, vegetation type, soil type etc.) or quantitative (nematode population size, altitude, soil pH etc.) and flexibility in handling can allow quantitative data to be transformed to qualitative data by prescribing class limits to the quantitative data. During the British Longidoridae survey a number of computer routines were written to create these transformations. Using such a routine allowed soils to be classified according to the American Soil Science Society soil texture classification scheme (Anon, 1965) from their percentage composition of sand, silt and clay.

Data input

Having acquired and coded the survey data, this coded information is prepared as computer input. Input media can be 80 column punch cards or punched paper tape, or data can be entered directly into the computer using a Key-to-Disc service system. This work is shared between scientific, data processing and computer service centre staff as convenience dictates when each successive batch of data arrives at the survey headquarters.

Data manipulation

At the Scottish Horticultural Research Institute, as several scientists were involved with nematological surveys, the information from the British Longidoridae survey was kept as a series of files each with special reference to one research project, together with a master file of physical information. This facility may be utilised during the EPPNS. Separate files may be created for each participating country to allow rapid access to each country's particular data set, also data required for mapping on a European scale can be kept as a master file. For any one investigation within a research

project such records as were required were copied from these various sources and merged into a new file used as input into mapping programmes or survey packages. The Edinburgh Multiple Access System (EMAS) allows this to be undertaken interactively from a teletype terminal. Short FORTRAN programs were typed into the computer directly with longer programs being submitted as card input. Compilation and running was carried out in foreground or batch mode using short commands. The system provided automatic allocation of storage space for files, automatic archiving to magnetic tape of chosen files and also backup to protect files from system crashes. Line printer copies of files could be obtained by typing LIST (file name, LP). The EDITOR facility permitted a new updated version of a file to be created alongside the existing file and operates by moving a cursor in any direction through the text of a file and inserting or deleting text with respect to its position.

These facilities have allowed data which have been stored continuously since being first input to the computer to be corrected for typing errors, validated in various ways, reformatted, merged with files containing additional information from the same samples and worked with continuously without any manual intervention for seven years. The amount of data in archive storage is effectively unlimited and almost 6 M bytes (six megabytes) of space in virtual memory is available for each user's files, although some of this space has to be kept as work space. The multi-access system can also be used for assembling job control, calling in packages, program directives and data for submission to the batch queue of an IBM 370 computer at Newcastle-upon-Tyne, England.

Computer mapping programs

For the British Longidoridae survey a simple FORTRAN program was written to convert the British Ordnance Survey National Grid map references to co-ordinates of a data array and either accumulate counts of the numbers of positive records or record that element of the array as being positive. The data array was then printed out on the line printer with a superimposed grid to correspond to the 100 km squares of the Ordnance Survey National Grid, or converted to characters and punched out on 80 column cards to create the input to the mapping system used by the Biological Records Centre (Perring, 1971) at the Institute of Terrestrial Ecology,

Monks Wood Experimental Station, Abbots Ripton, England. The Biological Records Centre mapping system utilises a card reader linked to an electric typewriter which has equal vertical and horizontal spacing. Each punched card contains information for one horizontal line on the typewriter page and also carries a series of codes enabling the pre-determined mapping symbols to be printed in the appropriate positions along that line. Pre-printed maps are normally used in the typewriter and as the map of Britain is the equivalent of 100 lines deep and 70 lines wide a pack of 100 punched cards can contain all of the information. Ireland is mapped on a different grid and therefore a separate base map and pack of 50 punched cards were required. After the Irish map and data have passed through the system the map can be superimposed on the larger British map (Fig. 2).

A program, CAMAP, developed in several versions for computer areal mapping, by the Geography Department, Edinburgh University, was also utilised. The survey grid references again were converted into coordinates, and an array of symbols was passed to the program defining a background symbol for negative map squares and other symbols for positive map squares. This information together with information defining the land areas of the British Isles on the same map grid basis, constituted the input to CAMAP. The CAMAP output is automatically routed to a special printer which employs an extended character set at 10 lines to the inch, thus giving « square » maps (Fig. 3).

Another version of CAMAP processed data by regions such as the agricultural extension services areas or administrative counties. Input consisted of data defining the regions on a base map, numerical values for some attribute of the regions and information defining the symbols to be used.

A survey, on a 1 km basis of the 100 km map square which includes the SHRI at Dundee is being prepared and the techniques being used are identical to those used for earlier mapping.

Computer programs for data tabulation and analysis

Data files were prepared to be processed by a standard statistical package (SPSS) developed at Stanford University, United States of America, especially for survey work. This « Statistical package for the Social Sciences » (Nie *et al.*, 1975) program, which was available

at the Edinburgh Regional Computing Centre, was used for two purposes. Data could be reclassified e.g. nematode population counts were reclassified as presence or absence data and vegetation types were placed into broader categories such as arable or deciduous woodland. With this reclassified data two or three way tabulations of data were prepared and statistical analyses for independence of the cross classifications were undertaken. This permitted possible relationships among nematode species and also between nematode species and various biotic and abiotic factors to be examined.

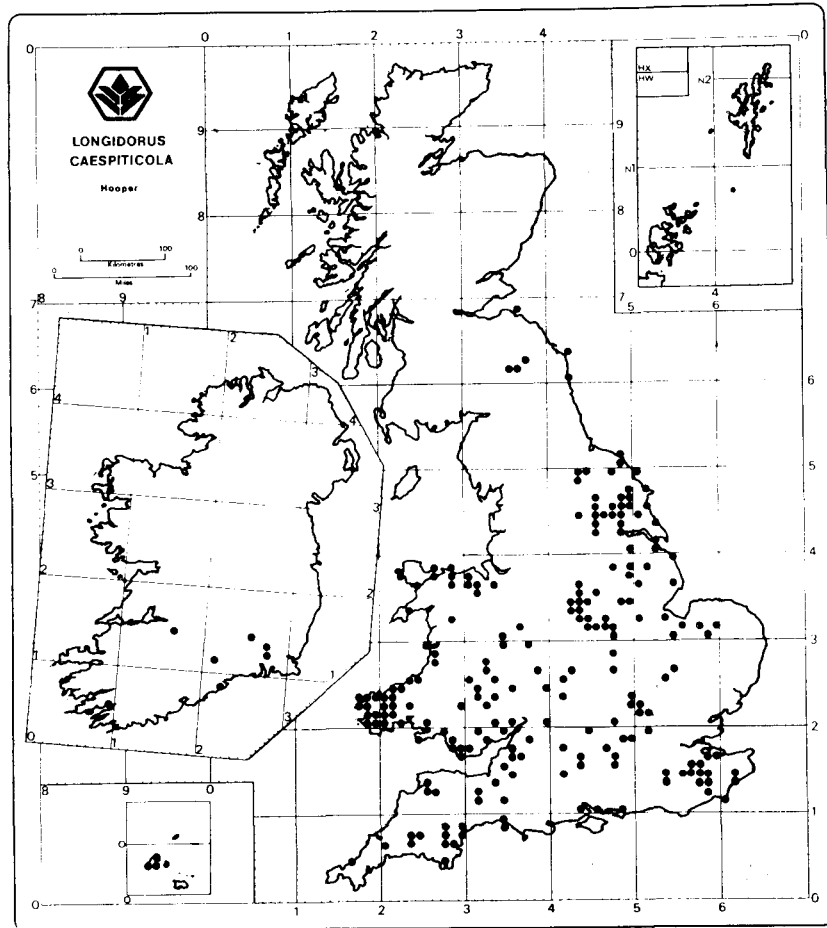


Fig. 2 - The geographical distribution of *Longidorus caespiticola* in Great Britain displayed on a base-map supplied by the Biological Records Centre, Monks Wood Experimental Station, Abbots Ripton, England.



Fig. 3 - The geographical distribution of *Xiphinema diversicaudatum* in the United Kingdom displayed on a map produced by CAMAP, Edinburgh Regional Computing Centre, Edinburgh, Scotland.

Conclusion and discussion

Nematological surveys carried out in many countries of the world have attempted to identify nematode species associated with specific crops or crop diseases whilst lists have been compiled of nematode species occurring in defined geographical areas (Dao *et al.*, 1970). Less frequently nematological surveys have been undertaken to identify and to plot systematically the distribution of nematode species occurring in a defined geographical area and to examine various biotic and abiotic factors possibly subscribing to the distribution patterns (Alphey and Boag, 1976; Dalmasso, 1970; Taylor and Brown, 1976). Biological recording, especially for conservation purposes, has been undertaken by a number of organisations in Europe particularly by botanical, entomological and ornithological groups (Heath and Perring, 1975) using various mapping schemes in order to accurately plot the geographical distributions of the organisms. Though the geographical areas investigated have usually been for countries, the various recording groups in the individual countries have been encouraged by the European Invertebrate Survey (Heath, 1971) to combine their information in an attempt to produce distribution dot maps at the continental level.

By developing a computerised data handling system and by adopting and modifying techniques and mapping systems used by these various conservation groups it has been possible for the first time to map accurately and analyse distribution data of agricultural pest organisms during the British Longidoridae survey. These same techniques and systems, which have proved to be both effective and economical, are being used to map and analyse data at various levels at the 1 km map square level within a 100 km map square; by the extension services regions; at the country level within Europe on a 10 km map square basis, and at the continental level (Fig. 4) using 50 km map squares. The mapping techniques used on all of these projects are the selection of appropriate records from a master file, the preparation of a base map and the production of a short FORTRAN program to produce input for the mapping program. The EPPNS data handling system which is based on the techniques and systems described here and used by the British Longidoridae survey can be easily exploited by other workers, for example the individual States of the USA could be mapped on a 10 km map

square basis and the USA as a whole could be mapped on a 50 km map square basis.

The maps, which are of interest in their own right, may be produced as an atlas (Heath *et al.*, 1977) and are of value when examining some of the biological and ecological factors affecting the distributions of the nematode species investigated. They may be consulted when formulating Phytosanitary regulations, and by adopting an appropriate size of map square such as 10 km on a country basis it may be possible to introduce Phytosanitary rules regulating plant and soil movement, not only between countries, but also within

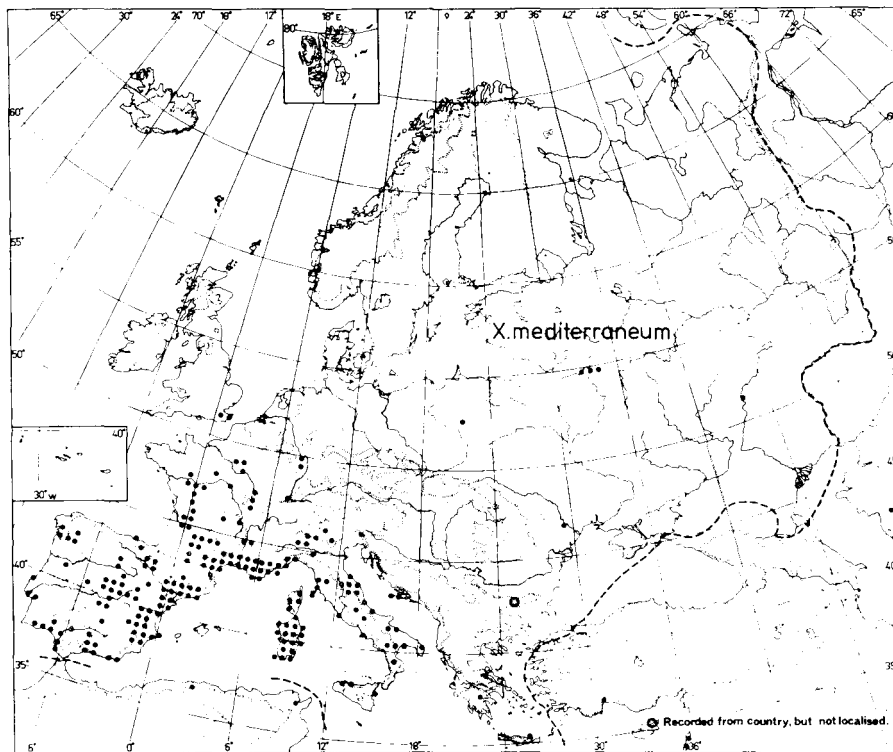


Fig. 4 - The distribution in Europe of *Xiphinema pachtaicum* (this figure was prepared before Siddiqi and Lamberti (1977) made *X. mediterraneum* a junior synonym of *X. pachtaicum*.) displayed on a base-map supplied by the European Invertebrate Survey, Biological Records Centre, Monks Wood Experimental Station, Abbots Ripton, England. (Information used to prepare this map was obtained from selected references. Many research papers indicate the occurrence of a species in a country but do not contain sufficient information to plot the distribution accurately.)

countries. The maps may also be used to help create more efficient control measures or to assess whether eradication schemes for selected species are possible. Economic yield losses due to nematode damage may be more accurately forecast; more accurate estimation of the scale and extent of nematological problems by examination of the nematode species distributions may enable more efficient use to be made of research resources.

S U M M A R Y

Several surveys of plant parasitic nematodes have been undertaken at the SHRI, Dundee, Scotland culminating in a proposal for a European Plant Parasitic Nematode Survey (EPPNS). The data base management system used for all surveys was based upon standard data processing facilities for file creation and management. Data, collected for the surveys as soil samples or from published work were interfaced to the computing packages by a minimal use of specially written programs. The choice of facilities which were freely available at the Edinburgh Regional Computing Centre permitted the research teams to systematize data recording and processing and to subordinate this aspect to map production and the assessment of ecological relationships.

The system described can be adapted to, and is being used for, records from a single field, farm or crop, records from agricultural advisory regions within the British Isles, nationally based records as well as the EPPNS. The European Survey is initially concerned with providing maps of geographical distribution of *Longidorus*, *Paralongidorus* and *Xiphinema* species but later will be extended to other plant parasitic groups. This involves international cooperation and will utilise the efficient and economical computerised data processing system established for the British surveys.

R I A S S U N T O

Elaborazione dei dati per le indagini sulla distribuzione dei nematodi.

Numerose indagini sulla distribuzione di nematodi fitopatogeni, condotte presso l'Istituto Scozzese di Ricerche Orticole (S.H.R.I.) a Dundee, hanno portato ad una proposta per una indagine su scala europea (EPPNS). Il sistema di elaborazione primaria dei dati, utilizzato in tutte le indagini, si basava su normali attrezzature per l'elaborazione dei dati per la schedatura e per la loro utilizzazione. I dati ottenuti da campioni di terreno o da pubblicazioni sono stati trasferiti sulle schede dell'elaboratore mediante minimo impiego di programmi scritti *ad hoc*. La scelta delle attrezzature disponibili al Centro Regionale di Calcolo di Edimburgo hanno permesso ai gruppi di ricerca di sistematizzare la registrazione ed elaborazione dei dati e di subordinare questo aspetto alla produzione di carte per la valutazione delle relazioni ecologiche. Il sistema descritto può essere adattato, e di fatto lo è, ai dati provenienti da singoli campi, aziende o colture, ai dati derivanti dai servizi tecnici agrari regionali nell'ambito delle Isole Britanniche, nonché a dati a livello nazionale o di EPPNS. L'indagine distributiva europea è, inizialmente, rivolta alla produzione di carte della distribuzione geografica delle specie di *Longidorus*, *Paralongidorus* e *Xiphinema*, ma in seguito sarà estesa ad altri gruppi di nematodi fitopatogeni.

Ciò evidentemente richiede cooperazione internazionale e potrà giovare dell'efficiente ed economico sistema computerizzato per l'elaborazione dei dati già messo a punto per le indagini britanniche.

L I T E R A T U R E C I T E D

- ALPHEY T.J.W. and BOAG B., 1976 - Distribution of trichodorid nematodes in Great Britain. *Ann. appl. Biol.*, 84: 371-381.
- ANON., 1965 - Committee report: glossary of soil science terms. *Proc. Soil. Sci. Soc. Am.*, 29: 330-351.
- BOAG B., 1974 - Nematodes associated with forest and woodland trees in Scotland. *Ann. appl. Biol.*, 77: 41-50.
- BROWN D.J.F., BOAG B. and HEATH J., 1978 - The European Plant Parasitic Nematode Survey - Instructions for Participants. *Occasional Publication No. 3*. Scottish Horticultural Research Institute, Invergowrie, Dundee, Scotland, 16 pp.
- DALMASSO A., 1970 - Influence directe de quelques facteurs écologiques sur l'activité biologique et la distribution des espèces Françaises de la famille des Longidoridae (Nematoda - Dorylaimida). *Ann. Zool. Ecol. Anim.*, 2: 163-200.
- DAO F., OOSTENBRINK M. and VIETS H.A., 1970 - A list of nematode surveys made for agricultural purposes. *Versl. Meded. plziektenk. Dienst Wageningen Ser No. 415*: 84 pp.
- HEATH J., 1971 - The European Invertebrate Survey. *Acta ent. fenn.*, 28: 27-29.
- HEATH J. and PERRING F., 1975 - Biological recording in Europe. *I.C.I. Endeavour*, 34: 103-108.
- HEATH J., BROWN D.J.F. and BOAG B., 1977 - Provisional Atlas of the Nematodes of the British Isles. Parts 1-3. *European Invertebrate Survey, Biological Records Centre, Inst. Terrestrial Ecol., Monks Wood Exp. Sta., England*. E.W. Classey Ltd., Faringdon, 76 pp.
- NIE N., HULL C.H., JENKINS J.G., STEINBRENNER K. and BENT D.H., 1975 - S.P.S.S. Statistical Package for the Social Sciences. McGraw-Hill, New York.
- PERRING F., 1971 - The Biological Records Centre - a data centre. *Biol. J. Lin. Soc. Lond.*, 3: 237-243.
- SIDDIQI M.R. and LAMBERTI F., 1977 - *Xiphinema mediterraneum* Martelli et Lamberti a junior synonym of *X. pachtaicum* (Tulaganov) Kirjanova. *Nematol. medit.*, 5: 133-135.
- TAYLOR C.E. and BROWN D.J.F., 1976 - The geographical distribution of *Xiphinema* and *Longidorus* nematodes in the British Isles and Ireland. *Ann. appl. Biol.*, 84: 383-402.
- TOPHAM P.B., BOAG B. and BROWN D.J.F., 1975 - Computerised data processing applied to nematode surveys. *J. of Biogeography*, 2: 45-48.

Accepted for publication on 30 April, 1978.