Acta Bot. Fennica 162: 5–9 ISBN 951-9469-56-7 ISSN 0001-5369
Helsinki 15 January 1999

From dot maps to bitmaps: *Atlas Florae Europaeae* goes digital

Tapani Lahti & Raino Lampinen

Lahti, T. & Lampinen, R., Botanical Museum, Finnish Museum of Natural History, P.O. Box 7, FIN-00014 Helsinki University, Finland

Received 14 November 1997, accepted 15 June 1998

We describe the construction of a digital database from the distribution data available on the printed maps in *Atlas Florae Europaeae* (*AFE*) volumes, eleven of them published so far. This database is designed to be at the core of the *AFE* data management system in the future, both for editing the printed volumes and delivery of plant distribution data in digital format.

Key words: databases, floristic research, phytogeography

INTRODUCTION

Atlas Florae Europaeae (*AFE*), a project for mapping the distribution of vascular plants in Europe, was launched in 1965 as a collaborative effort of European botanists (*see* Jalas & Suominen 1967). One of its primary goals has been to complement the primarily taxonomic information of *Flora Europaea* (*FE*; Tutin *et al.* 1964, 1968,..., 1980, 1993) by providing a more detailed view of plant distribution patterns across Europe. The *AFE* Secretariat, responsible for the collection of distribution data from the national contributors and compilation of distribution maps, has been permanently located in Helsinki, Finland. By the end of 1997, eleven volumes of *AFE* with 2 927 distribution maps have been published (Jalas & Suominen 1972,..., 1994, Jalas *et al.* 1996).

During the first thirty years of *AFE*, the basic technology for the collection of distribution data and publication of distribution maps has not changed very much. Recent advances in data processing technology, as well as requests to make *AFE* data available in a format suitable for computerised statistical analyses, have made it necessary to reassess the data management scheme of *AFE*. There are also a number of inconsistencies in the mapping grid system used by *AFE* that should be resolved in order to make the distribution data more useful in a wider perspective.

As a response to these needs, the construction of the digital *AFE* database was started at the Botanical Museum of the University of Helsinki in 1992, in close collaboration with the *AFE* Secretariat. Here we shall describe the techniques used in the conversion of the printed distribution maps into a digital database, and shortly summarize the future prospects of *AFE*.

MATERIAL AND METHODS

The structure of the distribution data

The starting point for the construction of the digital database were the printed distribution maps in the published

AFE volumes. At the time of the initial database construction, there were nine of them, with a total of 2 109 distribution maps. Later volumes 10 and 11 have been processed after publication according to the same principles.

Apart from some minor changes in the dot locations, the *AFE* base map used for the visualisation of distribution patterns has remained the same in all volumes. However, only that portion of the map needed to show the distribution of the taxon in question has been printed. Thus, there are, in addition to the full map, eight different captions of the base map used (top, bottom, left and right halves; NE, SE, SW and NW corners).

On the *AFE* base map, a total of 4 419 possible dot locations have been defined as a modification of the standard Universal Transverse Mercator (UTM) grid. The basic size of the grid cell is a rectangle of 50×50 km^2, with wider and narrower cells located along the buffer zones of the six degree longitudinal bands of the UTM grid scheme. Depending on the status of the taxon in each grid cell, either an empty cell or one of the 16 symbols indicating the different status categories (Table 1) is shown on the map. Some of the symbols have been reserved for showing the distribution of several taxa on the same base map, and even joint occurrences in the same grid cell.

Interpretation of the distribution maps

After comparing the tradeoffs between the price and performance of various techniques for extracting the distribution data from the printed maps into the database, we decided to adopt a fairly simple (and, above all, inexpensive) solution. Only a part of the map originals were available to us in the Museum archives, and their large size made them less than optimal for the interpretation process. Instead, we chose to use the printed *AFE* volumes with distribution maps of a more suitable size.

First, one copy of each *AFE* volume was cut into separate sheets. Each map was then scanned into a black-and-white bitmap file with a resolution of 300 dots per inch. With this resolution, the diameter of the map symbols on the bitmap image was approximately 16 pixels, which we considered to be large enough for the automation of the interpretation process.

Table 1. Number and proportion of symbols in the different taxon status categories used in *Atlas Florae Europaeae* (volumes 1 to 11).

Status	Count	%
Native	581 815	91.7
Unknown	5 663	0.9
Established alien	36 852	5.8
Possibly extinct	4 365	0.7
Extinct	3 184	0.5
Record uncertain	2 758	0.4
Total	634 637	100.0

The next step in the interpretation process was performed by a customized application program that analyzed the scanned bitmap image of each distribution map. It utilized a file containing the locations of the grid cells on the base map to calculate the corresponding pixel location on the bitmap image.

Inside a circle surrounding the calculated pixel location of each grid cell, the proportion of black pixels was calculated. If the proportion was higher than a given threshold value, the program recorded a dot location into the output file. There was an intermediate optimum for the threshold value, due to the fact that some patterns on the base maps (e.g., the coastline of the Netherlands) were complex enough to generate artefactual dot locations even without any symbol on the original map.

Only native occurrences (black dots) were recorded into the output file. Automatic interpretation of the other symbols shown in Table 1 was considered too complicated given the time and resources available, therefore they were added manually. As can be seen from Table 1, this approach was accceptable due to the fact that black dots were the dominant symbol type on the distribution maps.

The final step in the interpretation process was done manually by comparing the scanned bitmap image and its interpretation on a computer screen. Missing or incorrect symbols (including all the other 15 symbols except black dots) were added with simple point-and-click mouse operations. The average time for the visual inspection of one distribution map was about three minutes, which we considered satisfactory.

Structure of the database

After the interpretation process was completed, the distribution data were loaded into a relational database for further processing. At the basic level, the structure of the *AFE* database is very simple. The main data table (for database terminology *see* e.g., Pankhurst 1991) comprises only three primary fields: taxon name, *AFE* grid cell name, and taxon status in the cell. By using the standard technology of relational databases, these fields are then linked to supporting datasets providing the necessary information for the taxonomic hierarchy of the plant taxa and geographical location of the grid cells.

Problems

The manual map production process has made it possible to deviate from the predefined mapping scheme and to make adjustments to distribution maps whenever considered necessary by the *AFE* Secretariat. During the construction of the digital database, these deviations became apparent.

The correspondence between a symbol on a map and a record in the database is based on the name of the *AFE* grid cell. Some of the printed distribution maps contain symbols at locations which are not among the 4 419 possible symbol locations on the *AFE* base map. Often these sym-

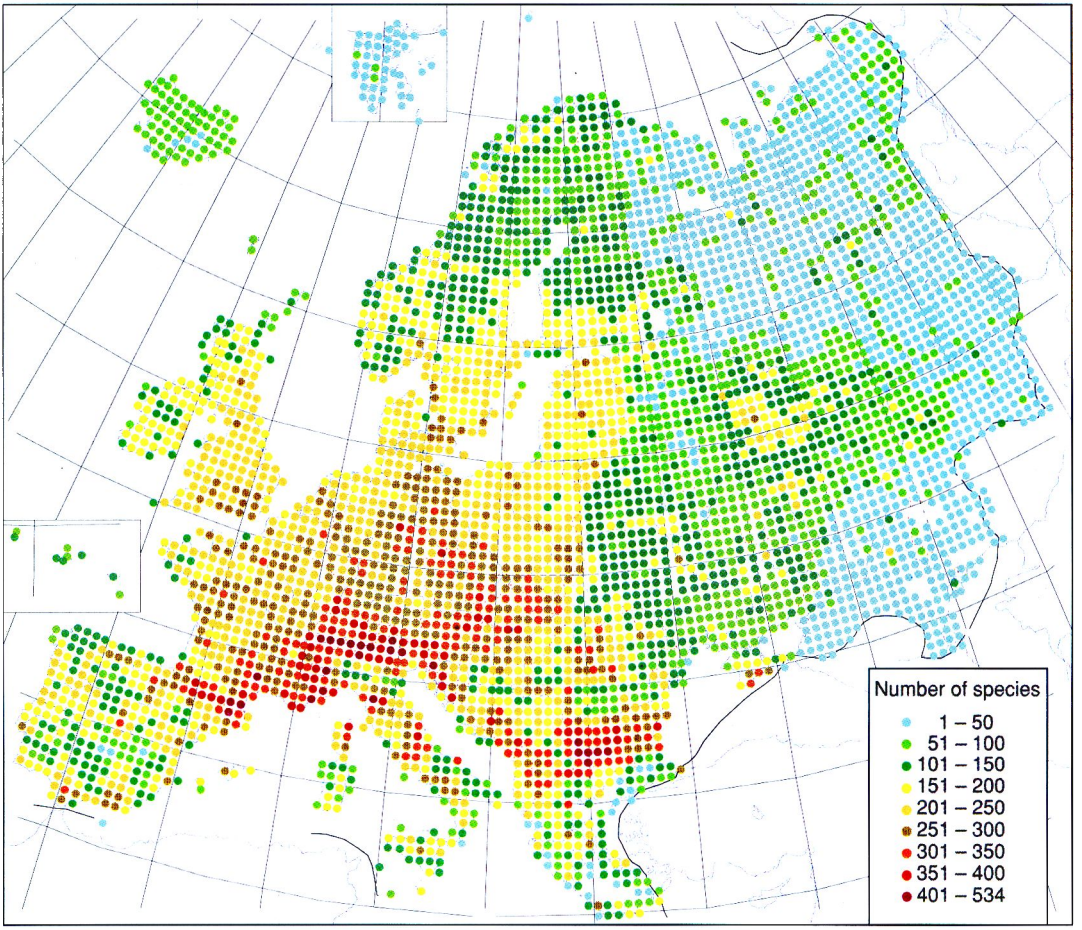

Fig. 1. The number of species in the AFE grid cells, according to the *Atlas Florae Europaeae* volumes 1–11.

bols indicate interesting occurrences of the taxon along the coastline or on the islands of the Baltic Sea or the Mediterranean region. Probably the most conspicuous of these maps is that of *Fumaria munbyi* (map 2085 in Vol. 9), where the only dot is in a location not defined on the *AFE* base map.

The reason for these locations being problematic results from the fact that the grid as defined on the *AFE* base map covers only continental regions and the largest islands. However, we do not see any technical reasons why the aquatic areas could not be used as well. Therefore, we are currently working on a new *AFE* base map that covers the whole of Europe, both land and water. For the time being, the maps containing nonstandard dot locations have been recorded, but these dots do not have a corresponding record in a database yet.

RESULTS

As a result of the interpretation of *AFE* volumes 1 to 11 published so far, we obtained a database containing a total of 634 637 status records of taxa in six categories (Table 1). This database covers only less than 20% of the European vascular flora (*AFE* Vol. 12 will complete the taxa of Flora Europaea Vol. 1), and is therefore far from complete as such. This fact should be kept in mind when interpreting the results shown below.

The digital format of the *AFE* distribution data makes it possible to analyse and visualize plant distribution patterns easily with modern computer tools. Some of these options are shown in the digital version of *Atlas Florae Europaeae*, known as *AFE* Update, which is a commercial software product running on Windows 95/NT computers.

At the simplest level, the database can be used for visualizing the distribution of one single taxon on the base map. Although this seems like a reconstruction of the printed *AFE* volumes, the database goes somewhat further. Unlike the static

maps in the printed *AFE* volumes, it is easy to show the distribution of *any* taxon, including higher level taxa like genera and families. Sometimes even a species-level distribution map may be in demand, like in Hill's (1997) attempt to visualize the distribution of *Cakile maritima* from the printed maps of four separate subspecies. With the digital *AFE* database available, we hope to prevent the "taxonomic tigers" from eating Hill and his fellow ecologists alive.

Statistical overview maps are also easily calculated from the database. Fig. 1 shows the number of species in the UTM grid cells in different parts of Europe. The highest species numbers in the mountain regions of Central Europe (sometimes enhanced by the coastal effect) are obviously quite accurate, as well as low figures in Northern Europe, Iceland and Svalbard. On the other hand, the steep gradient between Eastern and Western Europe largely reflects the variation in the amount of field data available. This map, together with a number of other examples of statistical patterns in the flora, is also available at the *AFE* World-Wide Web site (http://www.helsinki.fi/kmus/afe.html).

With the availability of the *AFE* database, statistical analyses of the European flora should become easier to perform. In addition to analyses of the flora as such (e.g., Birks 1976, Pedersen 1990, Myklestad & Birks 1993, Preston & Hill 1997), distribution data in digital format can also be related to environmental variables in order to obtain further insight into phytogeographical patterns (e.g., Huntley *et al.* 1995, Sykes *et al.* 1996).

DISCUSSION

The first stage of the construction of the *AFE* database can now be considered completed. All the distribution maps published so far in the printed volumes are also available in digital format. We are now ready to move to the next stage.

AFE Vol. 12, about to be published in 1998, will still be transferred into a digital format by scanning and interpreting the printed distribution maps as described above. It will complete the taxa of *FE* Vol. 1. *FE* Vol. 2 begins with the large and taxonomically demanding family Rosaceae. For the mapping of Rosaceae and the taxa that follow it, a new editorial system for *AFE* will be implemented in 1998. It is fully computerized and based on the *AFE* database, thus simplifying the joint production of printed and digital versions of *AFE* in the future. The new editorial system also enables the collection of distribution data from national contributors directly in digital format. These topics will be discussed in more detail elsewhere.

The very long timespan of the *AFE* project (at least a century with the present mapping scheme) creates a number of problems that the database approach may alleviate. Older distribution maps become obsolete long before the last ones will be published, and obviously none of us living today will see the distribution map of *Hammarbya paludosa* (the last species in *FE* Vol. 5) printed on an *AFE* volume.

The dynamic nature of the *AFE* database makes it possible to update distribution data for the published taxa, as well as to make distribution data, collected from national databases, available as preliminary maps independently of the editorial process. If this approach gains wider acceptance, it implies that the *AFE* database may gradually transform into a European floristic surveillance network that will expand the original mission of *Atlas Florae Europaeae*.

Acknowledgements: We are grateful to Drs. Juha Suominen and Pertti Uotila for their valuable comments on the manuscript. We also thank Dr. Mark Hill (Institute of Terrestrial Ecology, Monks Wood, England)) for kindly providing us a file containing the locations of the grid cells on the *AFE* base map. The map showing the number of species in the *AFE* grid cells was created by the DMAP software package, developed by Dr. Alan Morton.

REFERENCES

Birks, H. J. B. 1976: The distribution of European pteridophytes: a numerical analysis. — New Phytologist 77: 257–287.

Hill, M. O. 1997: Book review: Atlas Florae Europaeae, Vol. 11, Cruciferae (Ricotia to Raphanus). — Watsonia 21: 403.

Huntley, B., Berry, P. M., Cramer, W. & McDonald, A. P. 1995: Modelling present and potential future ranges of some European higher plants using climate response surfaces. — Journal of Biogeography 22: 967–1001.

Jalas, J. & Suominen, J. 1967: Mapping the distribution of European vascular plants. — Memoranda Soc. pro Fauna Flora Fennica 43: 60–72.

Jalas, J. & Suominen, J. (eds.) 1972, 1973, 1976, 1979, 1980, 1983, 1986, 1989, 1991, 1994: Atlas Florae Europaeae. Vol. 1–10. — The Commitee for Mapping the Flora of Europe and Societas Biologica Fennica Vanamo, Helsinki.

Jalas, J., Suominen, J. & Lampinen, R. (eds.) 1996: Atlas Florae Europaeae. Vol. 11. — The Committee for Mapping the Flora of Europe and Societas Biologica Fennica Vanamo, Helsinki. 310 pp.

Myklestad, Å. & Birks, H. J. B. 1993: A numerical analysis of the distribution patterns of Salix L. species in Europe. — Journal of Biogeography 20: 1–32.

Pankhurst, R. J. 1991: Practical taxonomic computing. — Cambridge University Press, Cambridge. 202 pp.

Pedersen, B. 1990: Distributional patterns of vascular plants in Fennoscandia: a numerical approach. — Nordic Journal of Botany 10: 163–189.

Preston, C. D. & Hill, M. O. 1997: The geographical relationships of British and Irish vascular plants. — Botanical Journal of the Linnean Society 124: 1–120.

Sykes, M. T., Prentice, I. C. & Cramer, W. 1996: A bioclimatic model for the potential distributions of north European tree species under present and future climates. — Journal of Biogeography 23: 203–233.

Tutin, T. G., Heywood, V. H., Burges, N. A., Valentine, D. H., Walters, S. M. & Webb, D. A. (eds.) 1964: Flora Europaea. Vol. 1. — Cambridge University Press, Cambridge.

Tutin, T. G., Heywood, V. H., Burges, N. A., Moore, D. M., Valentine, D. H., Walters, S. M. & Webb, D. A. (eds.) 1968, 1972, 1976, 1980: Flora Europaea. Vol. 2–5. — Cambridge University Press, Cambridge.

Tutin, T. G., Burges, N. A., Chater, A. O., Edmondson, J. R., Heywood, V. H., Moore, D. M., Valentine, D. H., Walters, S. M. & Webb, D. A. (eds.) 1993: Flora Europaea. Vol. 1, ed. 2. — Cambridge University Press, Cambridge.