IHGC98
First International Health Geographics Conference
Baltimore, Maryland, USA
(Maritime Institute)
16-18 October, 1998

Event Sponsor
The Johns Hopkins School of Public Health
Environmental Systems Research Institute, Inc.

SPATIAL FEATURING OF EPIDEMIOLOGICAL DATA OF A GENERAL POPULATION SAMPLE LIVING IN CENTRAL ITALY

R. della Maggiore (R.dellaMaggiore@cnuce.cnr.it), U. Mammini (U.Mammini@cnuce.cnr.it)
Geographical Information System Group - National Research Council (CNR), CNUCE Institute
Via S. Maria 36 I56126 Pisa - Italy

S. Baldacci, F. Di Pede, L. Carrozzi, A. Angino, G. Viegi (viegig@nsifc.ifc.pi.cnr.it)
Pulmonary Environmental Epidemiology Group - CNR, Institute of Clinical Physiology
Pisa - Italy

S. Petruzzelli (Stepetru@nsifc.ifc.pi.cnr.it), M. Luperini
Cardiopulmonary Department - University of Pisa

R. Barale (R.Barale@geog.unipi.it), L. Possenti
Environmental Science Department - University of Pisa

An epidemiological survey related to atmospheric pollution was carried out during 1991-1992 on a sample of about 2,800 people. Data were collected by means of standardized questionnaires, gathering information on the individual and his family, with special regard to respiratory aspects. Blood samples were taken at the interview and allergometric tests performed, as well as lung function tests. The data collected during the survey were widely inspected during these years within various disciplinary fields, such as epidemiology, lung disease, immunology, allergy, biochemistry, and cytogenetics. A specific characteristic of the study is the detection of antibodies anti-adducts to DNA of benzopyrene in the sera.
A geographical approach is currently being applied. A major goal is to develop a geographical model for air quality, derived from the effects of air pollution on human beings. To begin with we are investigating on relationships between health and the environment, analyzing the distribution of a respiratory symptom like wheeze, and the association between estimates of air pollution and the presence of circulating antibodies against specific DNA adducts due to airborne pollutants.
The method we are adopting for the geographical approach is as follows. In order to analyze the relationship between epidemiological data and the territory we decided to map the people who took part in the survey. These people are grouped per family, with perhaps more than one family residing in the same building, thus the same address may be common to many people, the number of presences ranging from 1 to more than 10. This enables us to build indexes (derived from epidemiological data) that express values relating to geographical points,

provided that each address is coded as a single geographical point. We thus obtain a dense scatter of points with associated values, on which geostatistical inspections can be performed to assess the fitness of the sample for a geographical approach, and possibly to adjust the sample. The final step of such mapping is to build thematic surfaces obtained by kriging techniques.

The study area comprises two contiguous geographical sites (parts of the municipalities of Pisa and Cascina, Tuscany), with subareas. The original sample was taken on a census section basis, with the assumption that the main pollutant source was an almost linear one, namely the main road with heavy traffic that crosses east-west and connects Pisa to Florence. To ensure the geographical robustness of the sample a propedeutic check is needed. In fact, since the epidemiological survey is a follow-up of a preceding survey performed in 1985-1988, some people have changed address in the period between the two surveys. Thus the people who have moved need to be identified and if necessary excluded from the sample, as the information they carry may no longer pertain to the site where they now reside (note that the original aim of the survey was not a specifically geographical one).

After classical statistical analyses have been performed, we are now applying geostatistic calculations. We also consider the noise due to the fact that human beings may be carrying misleading geographical information, because of their mobility, and we intend reduce it by using the detailed information collected with the diary of daily activity pattern for each subject of the survey.

As far as the cartographic support is concerned, the local administrations of Tuscany are being equipped with detailed numeric cartography (scale 1:2000), and GIS (Geographical Information System) data sets are growing fast (though they need to be refined). However as regards our two sites, numeric cartography is unfortunately only available for Pisa, while for Cascina we can only rely on data in CAD (Computer Aided Design) format. This meant that we had to manually input the address points of the Cascina site, while we were able to use automatic geocoding techniques for the Pisa one - ArcWiew/ArcInfo (ESRI) was the system adopted for GIS operations.

In addition to epidemiological data we can rely on other data from different sources. These are: estimates of the air quality obtained by the use of lichens as bioindicators (data from a survey performed by the Environmental Science Department of the University of Pisa); statistical data from the Census performed in 1991 by the Italian National Institute of Statistics (ISTAT); climatic factors, namely the predominating trend of winds during the daily periods of heavy traffic, to be compared with the direction of the main linear pollution source; and measures of air pollutants gathered by monitoring stations installed by local administrations.

At the end of the preparation phase we expect to have a suitable basis for spatial analysis containing both data describing the physical features of the territory such as environmental parameters and air pollution due to human activities, and measures of the state of the health of people and bioindicators.

IHGC98

Spatial Featuring of Epidemiological Data of a General population Sample Living in Central Italy

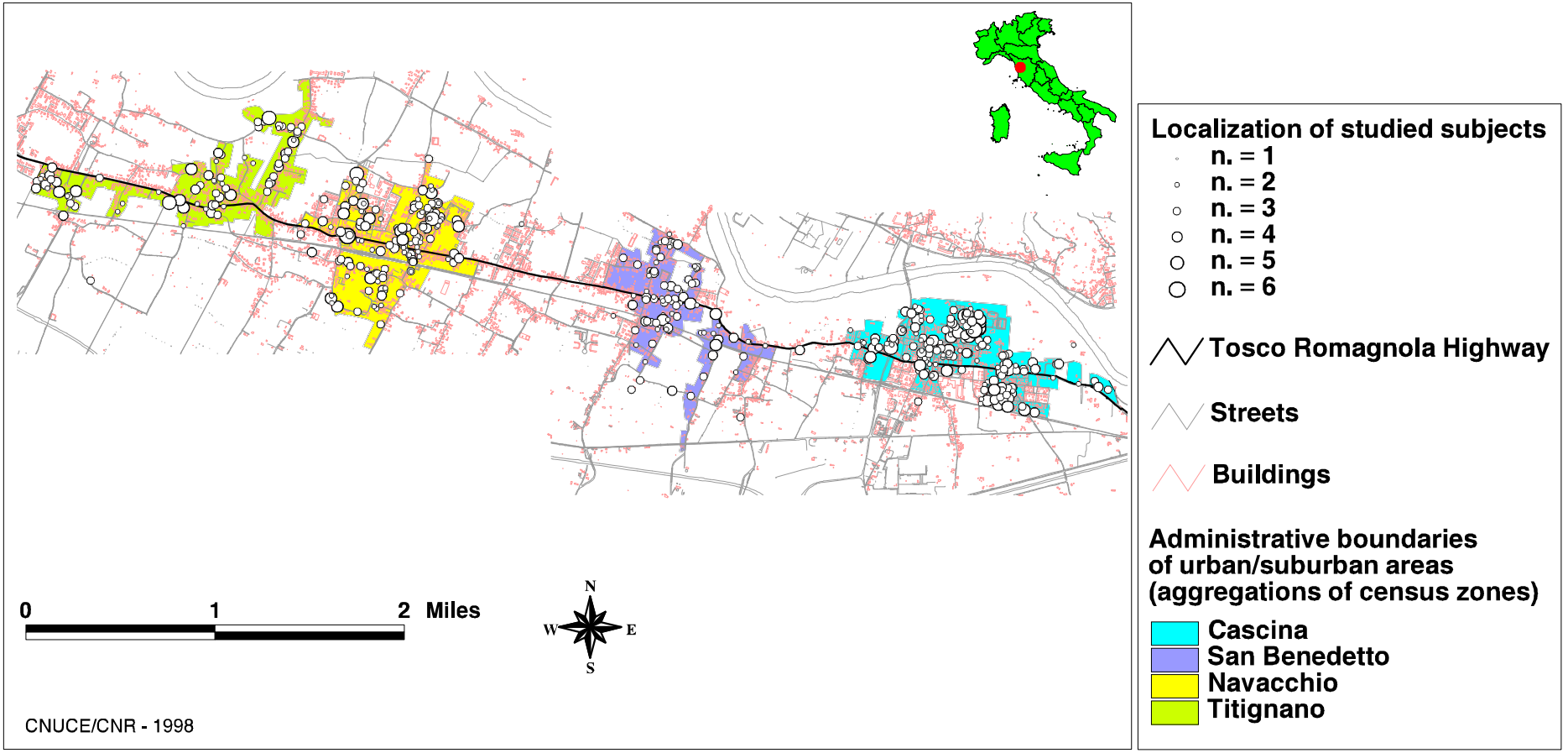

Fig. 1 - Geographical distribution of the studied subjects

Figure 1 shows the localisation of people who underwent Ab-BP-DNA test. The geographical constraints of the initial sampling were the nearness to Tosco Romagnola highway of the residences of studied people (within 800 m) and their belonging to selected census zones (supposed to be homogeneous for social status). Four urban/suburban areas constitute the aggregations of the selected census zones. Tosco Romagnola highway is heavy trafficked and it is considered the main pollutant source in the neighbourhoods.
The map shows a clustered scatter of the points of measure (residences of the studied people). The points represent a number of subjects ranging from 1 to 6. A total of 1018 people correspond to 482 different points of measure. People sharing the same residence may belong to different families.

IHGC98

Spatial Featuring of Epidemiological Data of a General population Sample Living in Central Italy

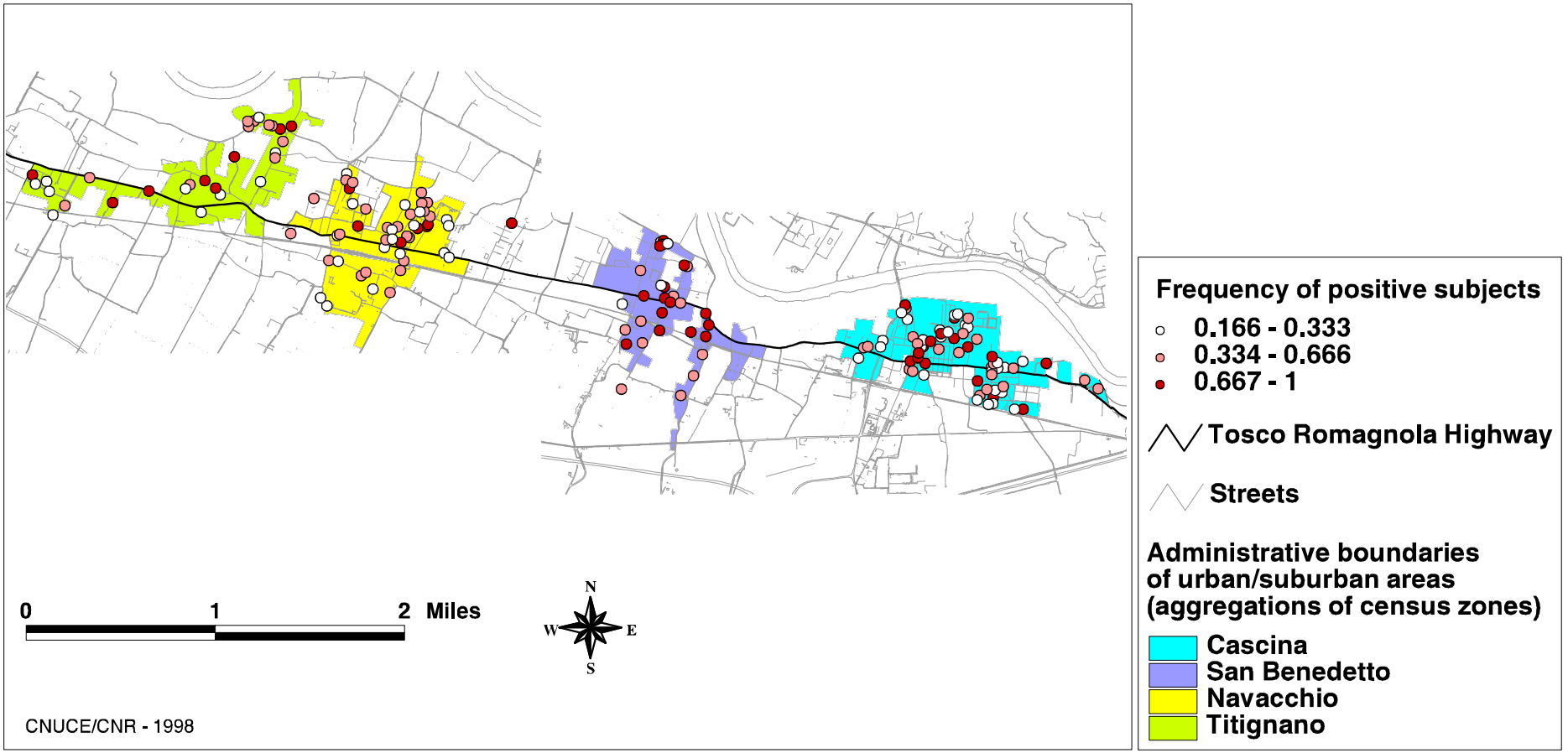

Fig. 2: frequency of positive subjects to Ab-BP-DNA test

In Figure 2 the frequency of positive subjects to Ab-BP-DNA is assumed as an index which is represented by tertiles. A total of 169 points of measure is mapped (the points corresponding to residences in which only negative subjects live are not shown).

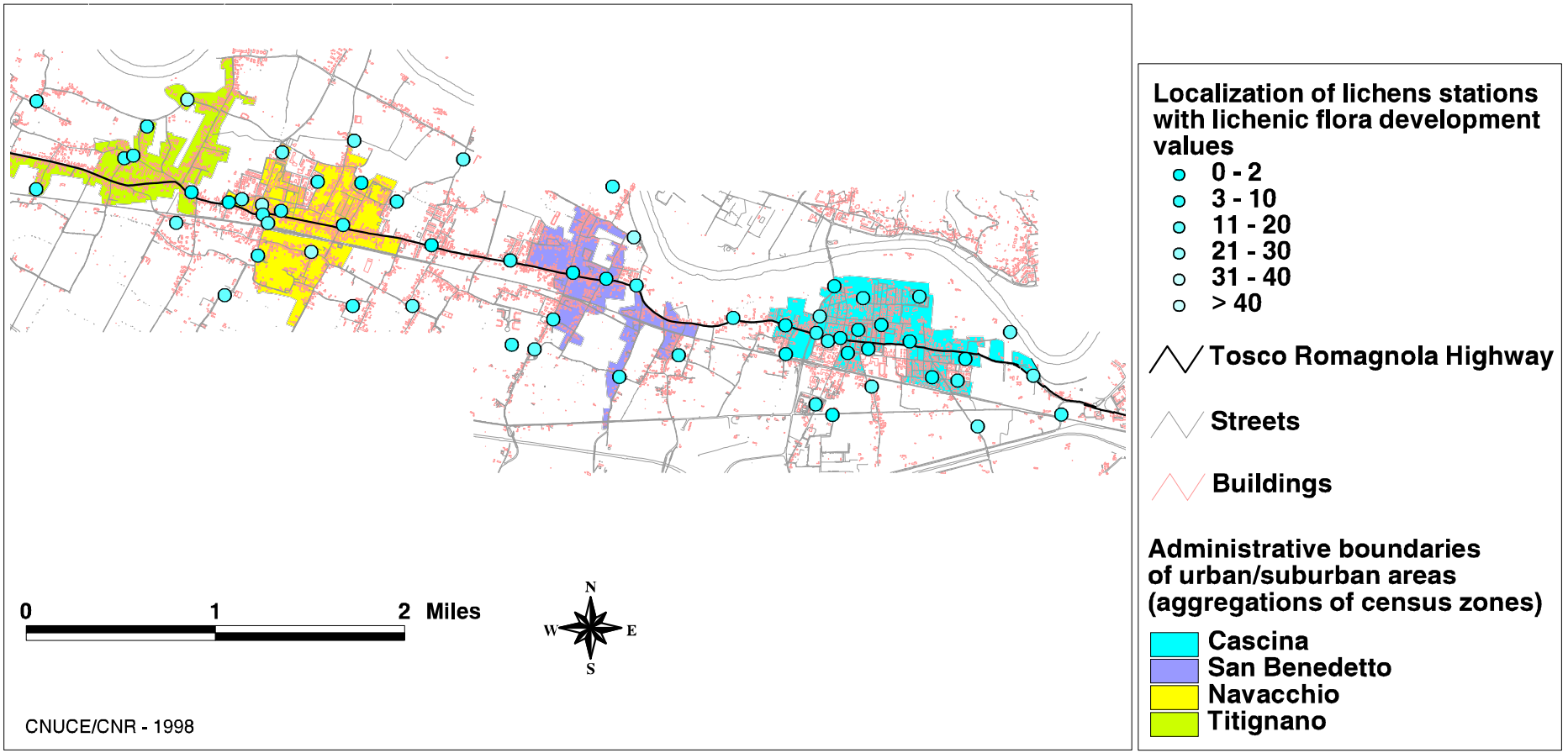

Fig. 3: lichens distribution and values

The survey on lichens was conducted approximately on the same area of the epidemiological survey, with less points of measure but a more regular distribution of the sample. A total of 69 points were mapped, corresponding to the trees examined for lichenic flora development. The geographical constraint was again the distance from Tosco Romagnola highway (within 1 Km) in the zone covered by the four urban/suburban areas. Figure 3 shows the distribution of the points of measure with the values assigned according to lichenic index, as shown in the legend.

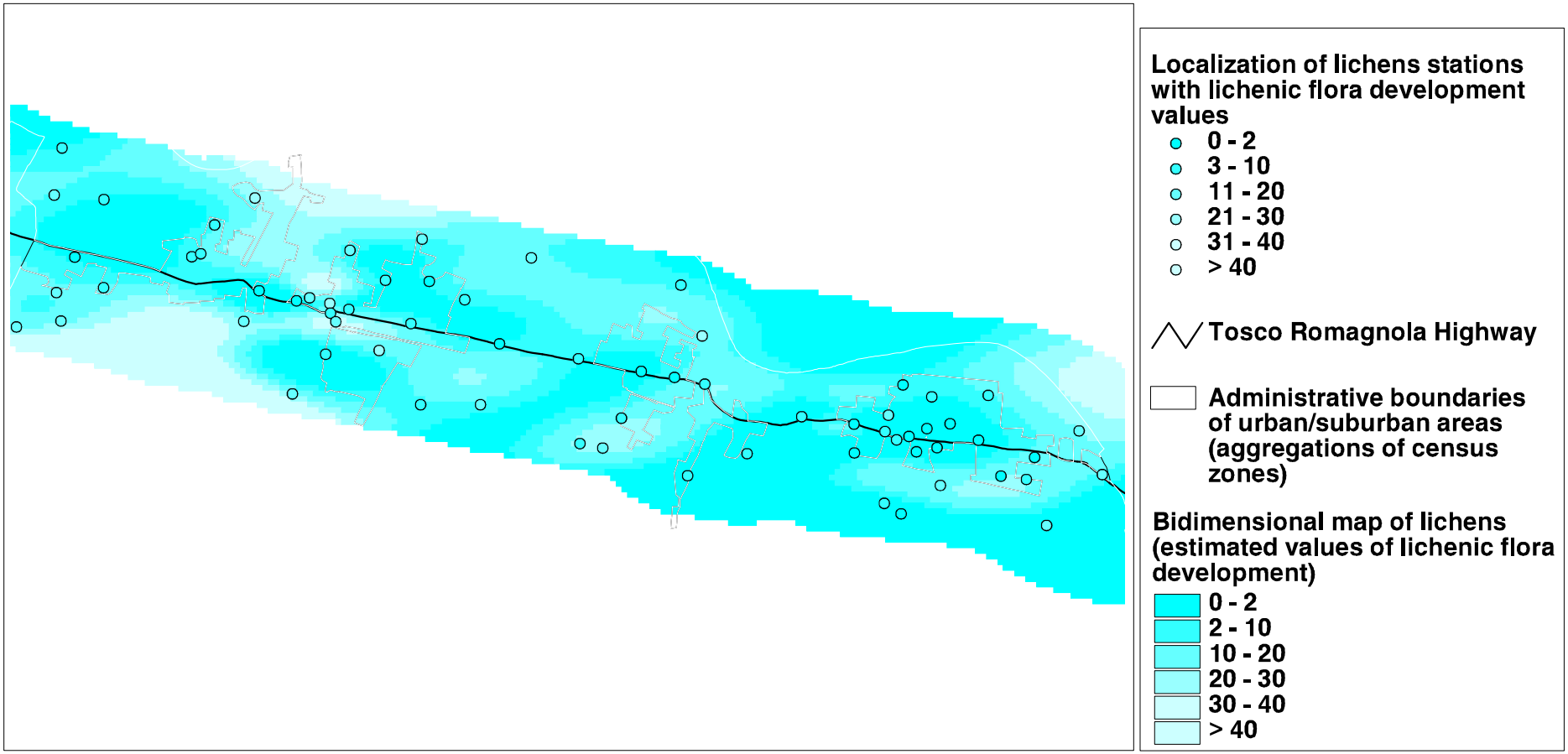

Fig. 4: air quality according to lichens

The values of the lichenic flora development were interpolated (spline method) to obtain the equipotential surfaces shown in figure 4; generating points are also shown, with values expressed according to the same colour legend as the surface. It is worth to note that the surface values should be considered more reliable in those zones where the points of measure are denser. Thus, at a visual spatial analysis, the nearness to Tosco Romagnola highway can be assumed as a factor of higher air pollution.
The dimension of the grid cell is 50 m. The mapped area is a buffer of 1 km around Tosco Romagnola highway. The administrative boundaries of the areas on which epidemiological data are available are overlaid for reference.

Fig. 5: map of distribution of positive subjects (nonsmokers only)

The frequency values of positive subjects to Ab-BP-DNA test were used to obtain equipotential surfaces by interpolation. Nonsmokers only were considered for a total of 741 individuals, corresponding to 408 points of measure. Smokers were excluded after a statistical study indicating the smoke of tobacco as a major cause of positivity. The map shown is the composition of the maps obtained for each of the four areas. The dimension of the grid cell is 10 m.
By analysing this map, there is an influence of Tosco Romagnola highway on the positivity of subjects to Ab-BP-DNA test, although it is not so evident as in the case of lichens (Fig. 4).

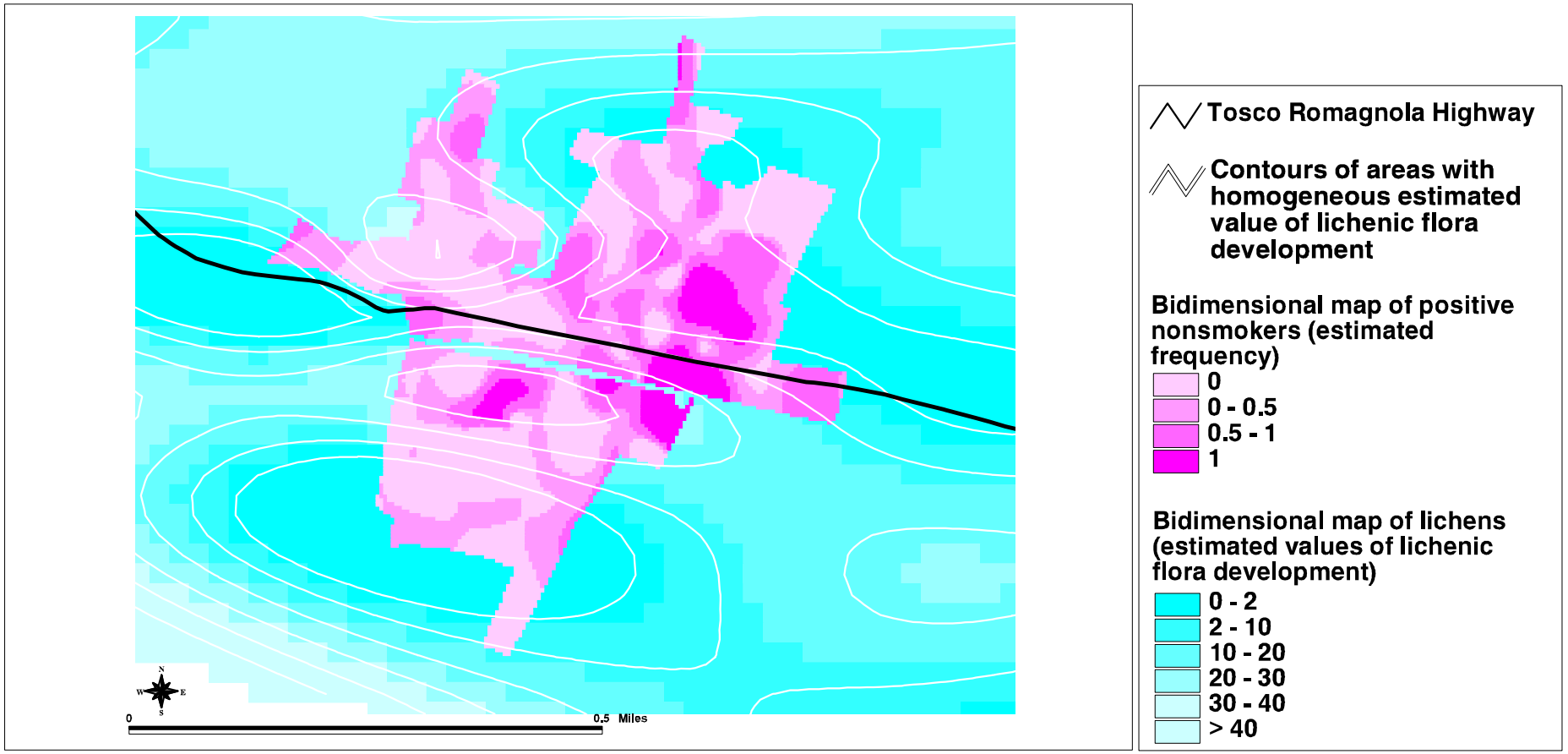

Fig. 6: comparison of lichenic flora development and Ab-BP-DNA test results

Figure 6 shows, for Navacchio area as an example, the superimposition of the map derived from Ab-BP-DNA test to interpolated lichens distribution map. Contour lines of lichens surfaces are then overlaid to allow understanding the envelope of the hidden surfaces. In this way it is possible to try a visual comparison of the two phenomena. Some degree of correlation is apparent.

Conclusions and developments

These preliminary results indicate that visual spatial data analysis seems to be a valid tool for inspecting geographical features relating to epidemiological data. The use of interactive tools is recommended as they allow to easily investigate the characteristics of data much more than the few proposed maps can suggest.

So far we have applied only to one result of the epidemiological survey, nevertheless we have developed a methodology to map individuals on the territory and to get indexes with geographical meaning from individual data. The results obtained by using these indexes encourage to further proceed with this approach. The quality of the samples can be refined by filtering the noise on the subjects, e.g. as performed for smokers it is possible to discriminate people on the basis of sedentarity; the large number of studied subjects allow to get still meaningful subset after filtering.

Data sources

Base map:	Carta 2000 – Tuscany Region (Italy)
Administrative boundaries:	Municipality of Cascina
Epidemiological data:	Survey performed in 1991-1993 by CNR (Italian National Research Council) and ENEL (Italian National Electricity Society)
Lichens:	Biomonitoring survey performed in November 1994 by Environmental Science Department - University of Pisa

Bibliography

Petruzzelli S. et al.
Serum antibodies to benzo(a)pyrene diol epoxide-DNA adducts in the general population: effects of air pollution, tobacco smoking, and family history of lung diseases. Cancer research Vol. 58:4122-4126 1998

Viegi G., Paoletti P., Carrozzi L., Vellutini M., Diviggiano E., Di Pede C., Pistelli G., Giuntini C., Lebowitz MD.
Prevalence rates of respiratory symptoms in Italian general population samples exposed to different levels of air pollution. Environ Health Perspect 94: 95-96, 1991.

Viegi G., Carrozzi L., Paoletti P., Diviggiano E., Baldacci S., Modena P., Pedreschi M., Mammini U., Di Pede C., Giuntini C.
Effects of the home environment on respiratory symptoms of a general population sample in Middle Italy. Arch Environ Health 47: 64-70, 1992.

Openshaw S., Clarke G.
Developing spatial analysis functions relevant to GIS environments
Spatial Analytical Perspectives on GIS – Edited by Fischer M., Scholten H.J. and Unwin D. – ISBN 0-7484-0340-X

Cislaghi C., Nimis P.L.
Lichens, air pollution and lung cancer
NATURE - Vol.387 - 29 May 1997

Smans M. and Estève J.
Practical approaches to disease mapping
Geographical & Environmental Epidemiology – Edited by Elliot P., Cuzick J, English D. and Stern R. – ISBN 0-19-262235-8

Poster map

An epidemiological survey related to atmospheric polluti
during 1991-1992 on a sample of about 2,800 people. D
by means of standardized questionnaires, gathering in
individual and his family, with special regard to respirat
samples were taken at the interview and allergometric t
well as lung function tests. The data collected during
widely inspected during these years within various discip
as epidemiology, lung disease, immunology, allergy,
cytogenetics. A specific characteristic of the study is
antibodies anti-adducts to DNA of benzopyrene in the ser
A geographical approach is currently being applied. A
develop a geographical model for air quality, derived f
air pollution on human beings. To begin with we are
relationships between health and the environmen
distribution of a respiratory symptom like wheeze, an
between estimates of air pollution and the presen
antibodies against specific DNA adducts due to airborne
The method we are adopting for the geographical approa
order to analyze the relationship between epidemiolog
territory we decided to map the people who took part in
people are grouped per family, with perhaps more than o
in the same building, thus the same address may be
people, the number of presences ranging from 1 to m
enables us to build indexes (derived from epidemio
express values relating to geographical points, provided t
coded as a single geographical point. We thus obtain
points with associated values, on which geostatistical i
performed to assess the fitness of the sample for a geog
and possibly to adjust the sample. The final step of su
build thematic surfaces obtained by kriging techniques.
The study area comprises two contiguous geographical
municipalities of Pisa and Cascina, Tuscany), with suba
sample was taken on a census section basis, with the as
main pollutant source was an almost linear one, namely t
heavy traffic that crosses east-west and connects Pisa
ensure the geographical robustness of the sample a pro
needed. In fact, since the epidemiological survey is
preceding survey performed in 1985-1988, some peo
address in the period between the two surveys. Thus the
moved need to be identified and if necessary excluded fr
the information they carry may no longer pertain to the si
reside (note that the original aim of the survey was
geographical one).
After classical statistical analyses have been perform
applying geostatistic calculations. We also consider the
fact that human beings may be carrying mislead
information, because of their mobility, and we intend red
detailed information collected with the diary of daily a
each subject of the survey.
As far as the cartographic support is concerned, the local
Tuscany are being equipped with detailed numeric c
1:2000), and GIS (Geographical Information System) dat
fast (though they need to be refined). However as rega
numeric cartography is unfortunately only available fo
Cascina we can only rely on data in CAD (Comput
format. This meant that we had to manually input the ad
Cascina site, while we were able to use automatic geocod
the Pisa one - ArcWiew/ArcInfo (ESRI) was the system
operations.
In addition to epidemiological data we can rely on other d
sources. These are: estimates of the air quality obtain
lichens as bioindicators (data from a survey pe
Environmental Science Department of the University o
data from the Census performed in 1991 by the Italian N
Statistics (ISTAT); climatic factors, namely the predo
winds during the daily periods of heavy traffic, to be c
direction of the main linear pollution source; and measur
gathered by monitoring stations installed by local adminis
At the end of the preparation phase we expect to have a
spatial analysis containing both data describing the phys
territory such as environmental parameters and air pollu
activities, and measures of the state of the healt
bioindicators.

Spatial featuring of epidemiological data of a general population sample living in Central Italy

R. della Maggiore, U. Mammini - Geographical Information System Group - National Research Council (CNR), CNUCE Institute
S. Baldacci, F. Di Pede, L. Carrozzi, A. Angino, G. Viegi - Pulmonary Environmental Epidemiology Group - National Research Council (CNR), Institute of Clinical Physiology
S. Petruzzelli, M. Luperini - Cardiopulmonary Department - University of Pisa
R. Barale, L. Possenti - Environmental Science Department - University of Pisa

IHGC98
First International Health Geographics Conference
Baltimore, Maryland, USA
(Maritime Institute)
16-18 October, 1998

IHGC98

Spatial Featuring of Epidemiological Data of a General population Sample Living in Central Italy

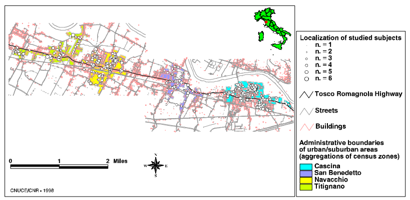

Fig. 1 - Geographical distribution of the studied subjects

IHGC98

Spatial Featuring of Epidemiological Data of a General population Sample Living in Central Italy

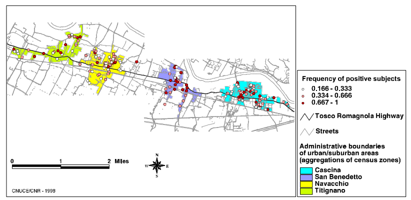

Fig. 2: frequency of positive subjects to Ab-BP-DNA test

IHGC98

Spatial Featuring of Epidemiological Data of a General population Sample Living in Central Italy

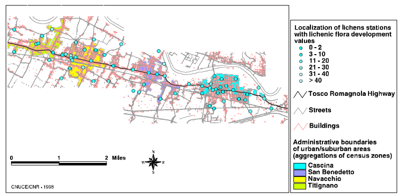
Localization of lichens stations with lichenic flora development values
0 - 2
3 - 10
11 - 20
21 - 30
31 - 40
> 40
Tosco Romagnola Highway
Streets
Buildings
Administrative boundaries of urban/suburban areas (aggregations of census zones)
Cascina
San Benedetto
Navacchio
Titignano
Miles

Fig. 3: lichens distribution and values

IHGC98

Spatial Featuring of Epidemiological Data of a General population Sample Living in Central Italy

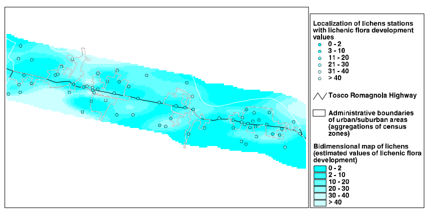
Localization of lichens stations with lichenic flora development values
0 - 2
3 - 10
11 - 20
21 - 30
31 - 40
> 40
Tosco Romagnola Highway
Administrative boundaries of urban/suburban areas (aggregations of census zones)
Bidimensional map of lichens (estimated values of lichenic flora development)
0 - 2
2 - 10
10 - 20
20 - 30
30 - 40
> 40

Fig. 4: air quality according to lichens

IHGC98

Spatial Featuring of Epidemiological Data of a General population Sample Living in Central Italy

Fig. 5: map of distribution of positive subjects (nonsmokers only)

IHGC98

Spatial Featuring of Epidemiological Data of a General population Sample Living in Central Italy

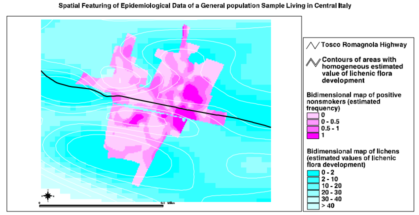
Tosco Romagnola Highway
Contours of areas with homogeneous estimated value of lichenic flora development
Bidimensional map of positive nonsmokers (estimated frequency)
0
0 - 0.5
0.5 - 1
1
Bidimensional map of lichens (estimated values of lichenic flora development)
0 - 2
2 - 10
10 - 20
20 - 30
30 - 40
> 40

Fig. 6: comparison of lichenic flora development and Ab-BP-DNA test results

Conclusions and developments

These preliminary results indicate that visual spatial data analysis seems to be a valid tool for inspecting geographical features relating to epidemiological data. The use of interactive tools is recommended as they allow to easily investigate the characteristics of data much more than the few proposed maps can suggest.
So far we have applied only to one result of the epidemiological survey, nevertheless we have developed a methodology to map individuals on the territory and to get indexes with geographical meaning from individual data. The results obtained by using these indexes encourage to further proceed with this approach. The quality of the samples can be refined by filtering the noise on the subjects, e.g. as performed for smokers it is possible to discriminate people on the basis of sedentarity; the large number of studied subjects allow to get still meaningful subset after filtering.

Data sources

Base map: Carta2000 - Tuscany Region (Italy)

Administrative boundaries: [illegible]

Epidemiological data: Survey performed in 1991-1993 by CNR (Italian National Research Council) and ENEL (Italian National Electricity Society)

Lichens: Biomonitoring survey performed in November 1994 by Environmental Science Department - University of Pisa

Bibliography

Petruzzelli S. et al.
Serum antibodies to benzo(a)pyrene diol epoxide-DNA adducts in the general population: effects of air pollution, tobacco smoking, and family history of lung diseases. Cancer research Vol. 58:4122-4126 1998

Viegi G., Paoletti P., Carrozzi L., Vellutini M., Diviggiano E., Di Pede C., Pistelli G., Giuntini C., Lebowitz MD.
Prevalence rates of respiratory symptoms in italian general population samples exposed to different levels of air pollution. Environ Health Perspect 94:95-96, 1991.

Viegi G., Carrozzi L., Paoletti P., Diviggiano E., Baldacci S., Modena P., Pedreschi M., Mammini U., Di Pede C., Giuntini C.
Effects of the home environment on respiratory symptoms of a general population sample in Middle Italy. Arch Environ Health 47:64-70, 1992.

Openshaw S., Clarke G.
Developing spatial analysis functions relevant to GIS environments. Spatial Analytical Perspectives on GIS - Edited by Fischer M., Scholten H.J. and Unwin D. - ISBN 0-7484-0340-X

Cislaghi C., Nimis P.L.
Lichens, air pollution and lung cancer. NATURE Vol.387 29 May 1997

Smans M. and Estève J.
Practical approaches to disease mapping. Geographical & Environmental Epidemiology - Edited by Elliot P., Cuzick J, English D. and Stern R. - ISBN 0-19-262235-8