Dasymetric mapping for an improved modeling of diseases

G. Boo, S.I. Fabrikant, S. Leyk

Research on diseases is contingent upon strict privacy regulations, which require anonymizing the identity of surveyed individuals. Information on the location of disease outcomes is therefore available only in the form of administrative summaries, in order to avoid disclosing the residential address of respondents (Beale et al., 2008). However, thematic maps based on aggregated disease data erroneously suggest that disease outcomes are homogeneously distributed within the administrative boundaries. As a matter of fact, disease outcomes typically occur where people live. In the Swiss context, for instance, populations and hence diseases are distributed within the residential area of municipalities, usually in the lower parts of the river valleys.

Die Erforschung von Krankheiten unterliegt strengen Datenschutzbestimmungen. Diese sehen die Anonymisierung der Identität der Befragten vor. Informationen über den Speicherort der Krankheitsergebnisse sind deshalb nur in Form von Zusammenfassungen verfügbar. Dadurch soll verhindert werden, dass die Wohnadresse der Befragten offengelegt wird. Die thematischen Karten, die sich auf aggregierte Krankheitsdaten stützen, deuten daher fälschlicherweise darauf hin, dass die Krankheitsergebnisse innerhalb der administrativen Grenzen homogen verteilt sind. Krankheitsergebnisse finden sich jedoch typischerweise dort, wo Menschen leben. In der Schweiz z.B. sind die Bevölkerung und damit die Erkrankungen in den Wohngebieten der Gemeinden verteilt.

La recherche de maladies est soumise à de très strictes dispositions légales concernant la protection des données. Celles-ci stipulent l'anonymisation de l'identité des personnes interrogées. Les informations concernant le lieu de stockage des résultats des maladies ne sont donc disponibles que sous forme de résumés. De cette façon on évitera de divulguer l'adresse des personnes interrogées. Les cartes thématiques qui se basent sur des données de maladies agrégées suggèrent faussement que les résultats des maladies sont réparties homogènement à l'intérieur des limites administratives. Mais les résultats des maladies se trouvent typiquement aux endroits où vivent des gens. En Suisse par exemple la population et par conséquent les maladies sont réparties dans les lieux d'habitation des communes.

La ricerca sulle malattie sottostà a normative molto rigorose a livello di privacy che prevedono l'anonimizzazione dell'identità dei partecipanti al sondaggio. Di conseguenza, le informazioni sulla localizzazione dei risultati delle malattie sono unicamente disponibili sotto forma di compendi amministrativi per evitare di risalire all'indirizzo dove abitano i partecipanti al sondaggio (Beale et al., 2008). Tuttavia, le carte tematiche basate sui dati aggregati delle malattie suggeriscono erroneamente che i risultati sulle malattie sono distribuiti in modo omogeneo entro le frontiere amministrative. In realtà, tali risultati sulle malattie sono riportabili a dove la gente è solita abitare. Per esempio, nel contesto svizzero, la popolazione e quindi le malattie si riscontrano nelle zone residenziali dei comuni, solitamente nella parte bassa delle vallate con fiumi.

Dasymetric mapping is a cartographic technique, developed in Russia in the 19th century, which aims at capturing the heterogeneous distribution of population within administrative units (Eicher and Brewer, 2001). In mapping diseases, this technique re-allocates aggregated disease data within administrative boundaries, by using related ancillary areal data on land use. A simple binary dasymetric refinement of aggregated disease data results, for example, in dividing municipalities into areas of residential versus non-residential land use. Disease counts are then re-allocated to the residential area of municipalities. This simple areal interpolation produces a more realistic cartographic representation of the spatial distribution of diseases.

This binary dasymetric technique can also be used for refining the location of disease explanatory factors and thus improving the statistical modeling of diseases (Boo et al., 2015). Refining disease explanatory factors consists in the recalculation of density and distance indicators as the spatial unit of reference changes from the municipality unit to its part of residential area. In fact, density indicators (e.g., population density) are calculated based on the extent of the reference unit as denominator. On the other hand, distance indicators (e.g., distance from health care facilities) are generally calculated based on the location of the centroid of the reference unit.

The map in Figure 1 summarizes the changes influencing the recalculation of density and distance indicators. The binary dasymetric refinement of density indicators is influenced by the extent of the residential area of the municipality. Changes are therefore more significant in the Alps and Jura Mountains, because municipalities are characterized by non-residential land uses dictated by the rugged relief. Similar results occur when refining distance indicators, which are associated with the shift from the centroid of the municipality boundary to the centroid of the residential area. Again,

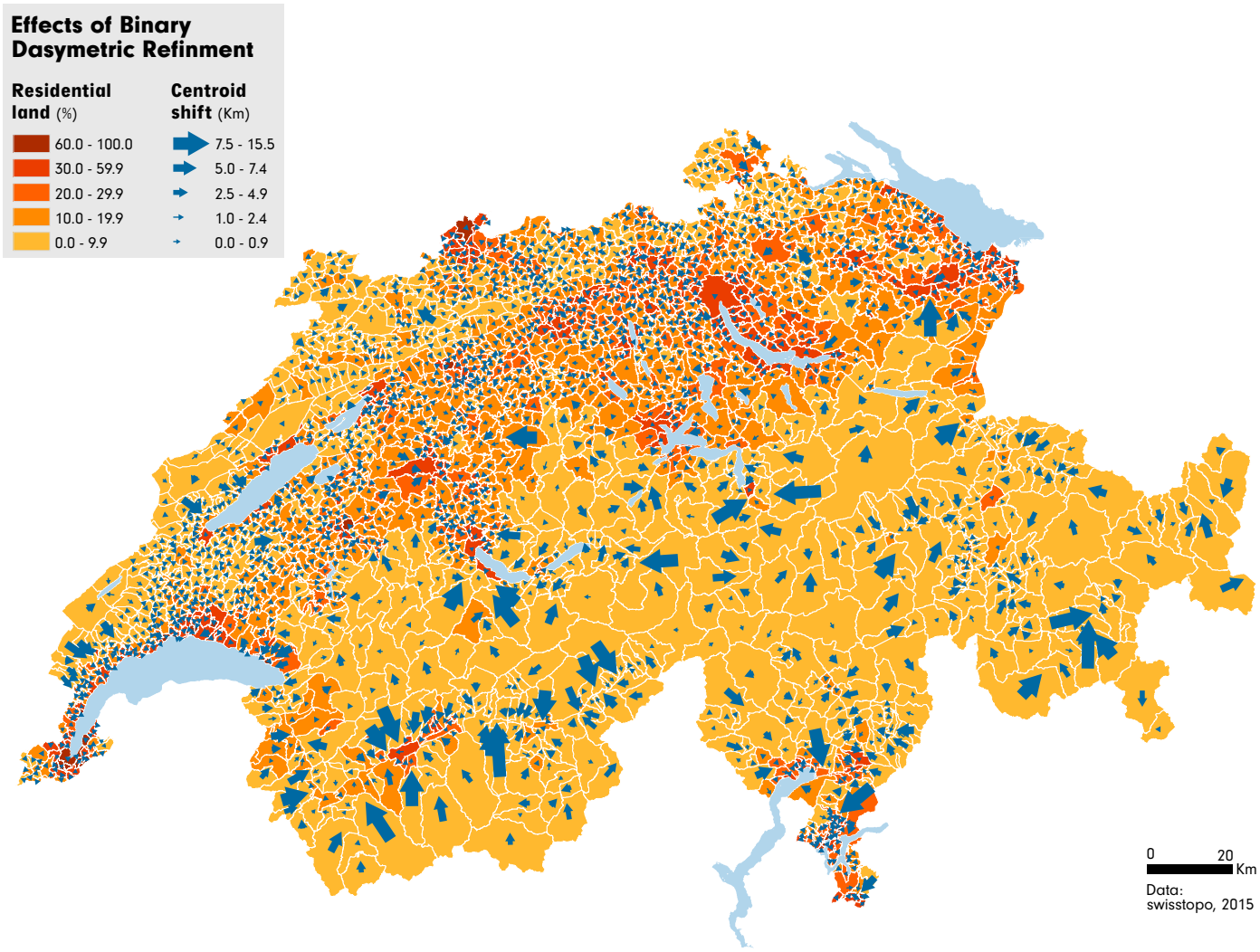

Fig. 1: The binary dasymetric refinement of disease explanatory factors requires recalculating density and distance indicators because the spatial units of reference are changed (after Boo et al., 2016).

the greatest shifts occur in municipalities where the residential development is restricted by the relief structure, generally in the Alps.

The recalculation of density and distance indicators has different effects across Switzerland because of specific land-use patterns associated with the relief structure. However, preliminary tests on the binary dasymetric refinement of selected density and distance indicators confirm significant improvements in the statistical modeling of diseases (Boo et al., 2015; Boo et al., 2016). The results of this cartographically-inspired research advocate the use of more realistic spatial units of reference for representing the distribution of diseases and associated explanatory factors. Further extensive research is needed to identify the full potential of the application of dasymetric mapping to the epidemiological modeling of diseases in different geographical contexts.

References:

Beale, L., Abellan, J.J., Hodgson, S., Jarup, L., 2008. Methodologic issues and approaches to spatial epidemiology. Environ. Health Perspect. 116, 1105–1110.

Boo, G., Fabrikant, S.I., Leyk, S., 2015. A novel approach to veterinary spatial epidemiology: dasymetric refinment of the Swiss Dog Tumor Registry data. ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences II-3/W5, 263–269.

Boo, G., Fabrikant, S.I., Leyk, S., Pospischil, A., Graf, R., 2016. A Dasymetric Framework for Improved Bias Assessment in the Swiss Dog Cancer Registry Data. Spat. Spatio-Temporal Epidemiol. 16 (under review).

Eicher, C.L., Brewer, C.A., 2001. Dasymetric mapping and areal interpolation: implementation and evaluation. Cartogr. Geogr. Inf. Sci. 28, 125–138.

Gianluca Boo
Sara I. Fabrikant
Department of Geography
University of Zurich, Switzerland
gianluca.boo@geo.uzh.ch
sara.fabrikant@geo.uzh.ch

Stefan Leyk
Geography Department
University of Colorado at Boulder (CO),
United States
stefan.leyk@colorado.edu