

APPLICATION OF THE GIS-BASED 3D MODELING OF MULTIFLAT BUILDINGS TO ASSESS THE PREVALENCE OF TUBERCULOSIS ON A CITY SCALE

I. Kuznetsov^{1,2}, E. Panidi^{1,*}, V. Korovka², A. Yakovlenko³

¹ Department of Cartography and Geoinformatics, Institute of Earth Sciences, Saint Petersburg State University, St. Petersburg, Russia - panidi@ya.ru, e.panidi@spbu.ru

² Saint Petersburg State Research Institute of Phthisiopulmonology, St. Petersburg, Russia

³ Interdistrict Petrogradsko-Primorskiy Tuberculosis Dispensary No. 3, St. Petersburg, Russia

Commission 4, WG 7

KEY WORDS: Medical Cartography, Infectious Disease, GIS-based Data Management, Medical Data Management System, 3D City Modeling.

ABSTRACT

This article is aimed at expanding and deepening knowledge in GIS analysis for medical professionals. Key task of described research is to elaborate a methodology of 3D mapping and visualization of the multiflat buildings in order to study most socially valuable diseases on the apartment scale in the St. Petersburg city. The use of this methodology allows to avoid the aggregation of geographical information within one building, and, on the other hand, allows to move from a general assessment of the prevalence of the disease to specific cases. In this case, the methodology is considered as primary health care support. The paper describes elaborated approach to detailed 3D mapping of multiple disease hotbeds in multiflat buildings. Main benefit of the proposed set of data processing and mapping techniques is the capability of apartment-scale connectivity evaluation of the hotbeds inside multiflat buildings.

1. INTRODUCTION

Application of map-based (Gordon, Womersley, 1997; Lesnykh, Mel'nikova, 2019) and GIS-based (Mayer, 1983; Gatrell, Bailey, 1996; Huang, Wang, 2012) data analysis, as well as any other appropriate analytical tools can be used to study and forecast infectious diseases spread and to save resources when fighting the disease. The coronavirus pandemic of 2019-2021 shows it clearly, while more than 221,5 million people have been infected worldwide and more than 4,5 million of them have died as of beginning of the September 2020, accordingly to the Johns Hopkins University (<https://coronavirus.jhu.edu/map.html>).

The topic of GIS technology use in medicine has been studied by us since the 2019. We have developed our own geocoder upon Nominatim OpenStreetMap geocoding engine (Kuznetsov et al., 2020a), and have used Web GIS to detect the problems and successes in the fight against tuberculosis in St. Petersburg (Kuznetsov et al., 2020b). Initially, we worked on detection of the tuberculosis infection and other diseases hotbeds, based on residence data of patients. This gave us the opportunity to determine the hotspots location of the tuberculosis infection with building location accuracy.

However, in practice, it turned out that such accuracy is not enough to consider the development of the disease. Repeatedly we met multiflat buildings, where the number of apartments was close to 500 and there were several hotbeds of infection. It was not clear whether these hotbeds were common, i.e., composed by patients who were familiar with each other, or whether the infection was introduced in different ways. In addition, when working with a planar map image (a thematic map without 3D elements), it is impossible to state unequivocally how dangerous a particular territory is. Let's assume that there are as many tuberculosis patients living in one building as there were not in

the entire district. In this case, a false idea is created about the real state of affairs in the building.

With the help of 3D visualization technology, it is possible to data distribute correctly throughout the building, to form most accurate idea of the current process of the disease spread. We decided to expand our database by introducing the apartment numbers of patients. Using this information, we are able to expand the scope of our research, although it will add the need for data encryption. Basing on this database, we became capable to visualise the hotbeds not just on a 2D map as it is implemented in other different research projects (Chen et al., 2021; Dutta et al., 2021; Wen et al., 2021), but also to apply 3D modelling of individual urban areas where the incidence rate is high; while 3D mapping and visualization in medical cartography appears not elaborated technique.

We use the *qgis2threejs* module for *QGIS* (<https://qgis2threejs.readthedocs.io>). With this technology, we hope to solve the problem of understanding of the origin of hotbeds and discover the relationships between individual hotbeds. In the future, this technology will become an integral part of the study methodology applied to the detection and investigation of socially valuable disease hotbeds. The need of development of this methodology is due to a number of reasons:

1. The methodology complements previous experience in of GIS use in medical research, allows us to refine the information obtained and to reach a new research level
2. The 3D visualization allows us to process and map more detailed information about patients, complementing the hotbeds of tuberculosis with hotbeds of HIV, hepatitis and tuberculosis children positive samples

* Corresponding author

Since the 3D visualization of multiflat buildings assumes significant costs of 3D objects creating, we decided to limit ourselves by experimental areas. These areas, should be distinguished either by geographical isolation first of all, or by extremely high rates of morbidity and prevalence of infection. We did this intentionally for two reasons:

1. Geographical isolation excludes the possibility of unlimited movement of patients
2. The study of territories where the indicators are overestimated several times requires an immediate response and analysis from the medical community

As an experimental area, we identified the Yuntolovo microdistrict (Fig. 1) in the northwestern suburb of St. Petersburg (second biggest city in Russia). This territory was chosen according to a number of important reasons:

1. The territory of Yuntolovo is removed from the relatively dense development of the northern part of St. Petersburg. This should reduce the transit flows of people who could bring the infection to the territory
2. This area was built up in the period 2015-2018, and the first residents appeared here only in 2019. This allows us to trace the appearance of the disease, excluding the study of earlier hotbeds of the disease
3. A small number of residential buildings. At the time of 2021, there are 38 residential buildings, 2 kindergartens and 1 school in operation. This allows us not to spend large resources on collecting and processing information



Figure 1. Location of the Yuntolovo microdistrict (red hatched area). Islands to the southeast are the central area of St. Petersburg city.

This area is not the only that has been subjected to experimental treatment. In addition to the territory in Yuntolovo, we explored another area. It is located in the east of St. Petersburg, in the Nevsky district, two buildings on Chudnovsky Street No. 8/1 and 8/2 (Fig. 2, 3). These buildings are notable because over past 15 years, more than 70 HIV infection cases have been detected here. Residents who live in these buildings are in close proximity to large hotbeds of HIV infection, which poses a great threat in the event of accidental sexual acts. However, only few tuberculosis cases were detected there. Tuberculosis is extremely dangerous for HIV patients, because it aggravates the course of the disease.



Figure 2. Overview image of buildings No. 8/1 and 8/2 on Chudnovsky Street.



Figure 3. 3D models of buildings No. 8/1 and 8/2 on Chudnovsky Street.

The buildings belong to the so-called standard series of buildings. Such buildings were built in the Soviet Union in 1960-1990. The layouts, sizes and number of rooms in all old multiflat buildings are always the same, and the plans can be easily found on the Internet, they are publicly available.

The areas that does not fall under any of the previously agreed rules was also considered, however we decided not to incorporate these results into the study. One building, for example, is located in the Admiralteysky district of St. Petersburg. Its peculiarity is that this building is divided into residential part and a district orphanage part. In this building, increased rates of tuberculosis and HIV infection were also noted. Since this building has the only common address, it is impossible to determine the source of the hotbed of the disease unambiguously. Information on this area was sent to the district tuberculosis dispensary for clarification and additional collection.

2. COLLECTION OF MEDICAL STATISTICS DATA

The issue of data collection has become very important for us. There is no practice in the Russian medical community to observe information about a patient that is not directly related to the disease. Although the legal right to collect such information is enshrined in federal Law No. 152 (Federal law, 2006), medical professionals rightly ignore such an order, preferring to fill in much more important parameters about the patient's health. Information about the addresses of patients is also incomplete.

To understand the problem of the patient's address determining, it is worth paying attention to the extremely important terms for understanding the "address of residence" and "residence address of Russian citizens", since these terms differ significantly from European and American ones. Formally, Russian citizens can

legally reside in any city of the country without having to register with state authorities. Such an address will be called the address of the real residence (residential address). The address of a citizen's residence in Russia is not regulated by law. However, in this case, they cannot officially receive public services. To do this, it is necessary to issue a residence address – the place of residence that is indicated in the citizen's passport and is officially registered by the state. In practice, this system works only when you are receiving an important public service, for example, buying an apartment or a house, registering a child in school or kindergarten. In the medical industry, such a nuance is often omitted and medical services are provided under a state health insurance policy. In this case, the residential address is not important. Formally, Russian citizens have the right of free medical care, including treatment for tuberculosis and HIV, it must (officially) be carried out at the patient's residence address. The place of actual residence and the place of registration of the patient may differ. A conflict arises because of this: a patient without a residence address in St. Petersburg who lives in St. Petersburg may come to the city polyclinic, where he is treated accordingly to the address of his real residence in St. Petersburg. There is a real problem of multiple addresses. For example, some patients live in different regions of Russia and come to St. Petersburg exclusively for treatment. In this case, another region address gets into the database. When compiling reports for regulatory authorities, this problem is not noticed. However, when trying to process it (for example, using GIS technologies), difficulties arise.

Unfortunately, there is no way to unequivocally answer an obvious and simple question: how many patients actually form the hotbed of the disease in a particular area. With the current system of addresses collecting, it is necessary to focus primarily on the address of the real residence, which can be established only in personal communication with the patient. At the same time, the patient may intentionally or unintentionally hide the address of the real residence. Therefore, special documents were drawn up that allow us to get from the patient exactly the necessary information that is needed in the study, including all the addresses where the patient happens to be in one way or another (place of work, place of study, place of residence of friends, and relatives). Our colleagues from the Interdistrict tuberculosis dispensary No. 3 in St. Petersburg to collect information about our patients. Information about Chudnovsky Street area was reported by the City Tuberculosis Dispensary.

When collecting data on buildings, we relied exclusively on open sources of information about the housing stock, and used publicly available Web maps. Our geospatial database incorporates information from Rosreestr (Russian land and real estate cadastre service). With its help, we received information about the cadastral numbers of buildings and apartments. The numbers were served as the primary key in the database. We mined addresses and apartment numbers from free data sources, such as *OSM* maps (<https://www.openstreetmap.org>) and *2GIS* (<https://2gis.ru/spb>) and *Yandex* maps (<https://yandex.ru/maps/2/saint-petersburg/>). In addition, data was collected for each entrance of building. The numbers of the first and last apartment in the entrance (building block), the number of floors, and the floor plans of apartments were taken into account. The number of floors is especially important, since the number of floors can be different in different building blocks, especially in new houses built after 2010 according to non-standard projects.

We are going to share this information with colleagues from the Faculty of Sociology, as well as add additional information about

the private lives of patients. In theory, this will allow us to study the problem of socially valuable diseases more deeply not only in St. Petersburg, but also in Russia in whole.

3. VECTORIZATION AND GEOCODING OF SPATIAL DATA

Vectorization of buildings for each area was performed in the manual mode. This was done for a number of reasons:

1. The absence of vector building features divided into blocks
2. It is necessary to get a vector layer containing correct information about each block. This information can only be obtained if the objects are not processed automatically

In addition, the territories of the study areas are small, and thus the vectorisation process did not take too much time. The addition of point address features for every patient with tuberculosis infection was also performed manually, with the exception of cases in the area of Chudnovsky Street, where there were significant volume of patients, and the area was limited to 2 residential buildings. To implement this task, we used our own previously development geocoding module for *QGIS* (Kuznetsov, 2020a).

In total, the specialists managed to collect complete information on 13 patients in Yuntolovo (Fig. 4) and more than 70 cases of HIV in the Nevsky district. They formed our test database, when performing information processing. As a result, we got 2 vector datasets: building polygons (more precisely, polygons of individual blocks of each building) and points of the disease hotbeds.



Figure 4. Distribution of hotbeds in Yuntolovo.

The need to divide buildings into separate blocks is caused not only by the desire to display the correct number of floors in the building. A separate entrance to the building (separate block) is already a large focus of a socially valuable disease, since this space is used by a limited number of people who directly live in the block. At the same time, it is not a primary task to consider individual entrances. The key goal is to move to the level of individual floors and even apartments. Thus, it is possible to obtain a complete composition of spatial data about patients – from the city area to a separate apartment. Thus, we can trace the path of infection development and its impact onto the area in a whole.

4. VISUALIZATION OF INFECTION HOTBEDS

As noted earlier, the key goal of using another methodology for disease hotbeds visualization is to switch to specific cases and their location in a multiflat apartment building. To begin with, we have connected the *qgis2threejs* module (*qgis2threejs* QGIS plugin). This allows us to visualize information in a 3D easily and quickly. The heights of the buildings were determined by the number of floors in each building block. In some cases, different blocks have different numbers of floors and apartments. It was quite easy to calculate the heights of the buildings in our case, because many buildings in the USSR and in Russia have a standard height between the ceilings (with the exception of built before 1955, which did not participate in the study). For clarity, this parameter was set to 2 times of the actual size – 5.6 meters. Multiplying the height by the number of floors, we got the height of the necessary building. It was much more difficult to determine the position of the apartment on the floor. Since we did not have floor plans for new buildings in Yuntolovo, the height of the point (the hotbed of infection) was set basing on the apartment number. Knowing the number of apartments in one block, as well as the number of floors of the block, we can determine the position of the patient's apartment in the building according to the following principle: the ordinal number of apartments on the floor is always calculated from left to right, from the smallest to the largest value. This approach is used only if it is not possible to obtain floor plans to use in the study. As a result, we can generate an imagery that approximately demonstrates the positions of the disease hotbeds in the buildings (Fig. 5).



Figure 5. Approximate location of hotbeds in buildings in Yuntolovo.

Now it is worth considering the buildings in the eastern part of St. Petersburg, which are located on Chudnovsky Street. Both buildings were built in 1989, have 9 floors in each block and an equal number of apartments on each floor (4 apartments per floor). Each apartment has a number in the planning structure. All the plans were obtained from the Internet and are available publicly (https://www.kvmetr.ru/information/homes_series/). The figure below (Fig. 6) shows the distribution of tuberculosis and HIV infection hotbeds in two buildings. In fact, this example illustrates the largest source of HIV infection in the Nevsky district of St. Petersburg. It is worth noting that the data were obtained for different periods. Data on tuberculosis reflect information for the last 5 years, and on HIV infection—for the last 20 years. Since a person with HIV can live long enough from the moment of infection (if the disease is treated in a timely manner), this information is quite relevant today. We mark HIV cases with red dots, and tuberculosis patients and deceased tuberculosis patients with blue and black dots, respectively. This picture makes it very clear that the mortality rate for tuberculosis can be

significantly increased in the area where people with HIV infection live, so these areas should be under special control.

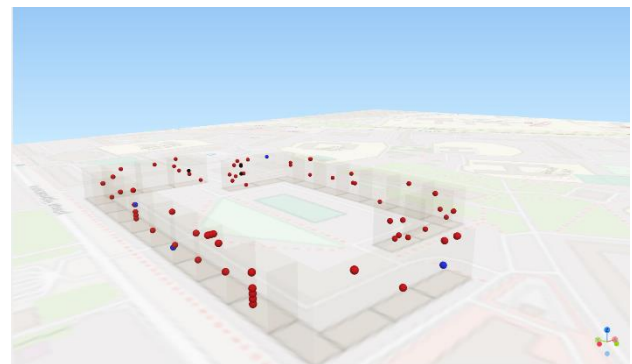


Figure 6. The location of the hotbeds of infection in the buildings along Chudnovsky Street №8/1 and №8/2.

The greatest attention of medical specialists was attracted by the block with numerous cases of HIV infection, in which cases of deaths of tuberculosis patients were registered (Fig. 7). It is noteworthy that a person with tuberculosis lives in a neighbouring block, where there are no cases of HIV infection, but he continues to be successfully treated and is alive to this day. This image demonstrates clearly the failure of work on the hotbed at a separate address. If the parameters are averaged during the general analysis, then this image shows the most burdened places where the disease is actively developing. Thus, it is possible to assess the level of work with the population and move to a personal level of responsibility for poor-quality performance of work by medical personnel.

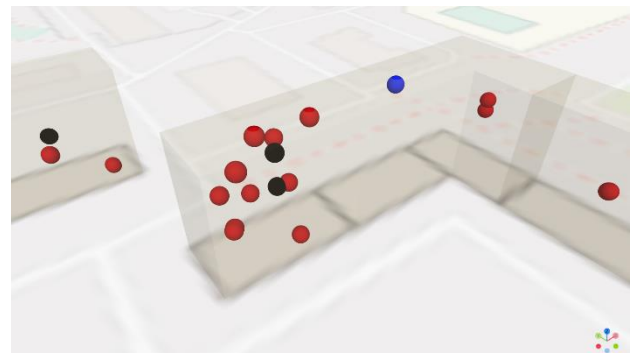


Figure 7. The most problematic part of the building No. 8/2.

Since the standard layout plans of the buildings are available publicly, we tried to systematize the data on HIV in order to determine the specific hotbed of the disease. To do this, the floor plan was tied to the block. After that, the apartments were removed from the plan. The figure below (Fig. 8) shows the number of cases of HIV infection detected in a separate apartment. For example, two or more people infected with HIV live on the third and fifth floors. The intensity of the colour indicates the occupancy of the apartment. Gained data demonstrated the need to use floor plans of in the subsequent stages of work.

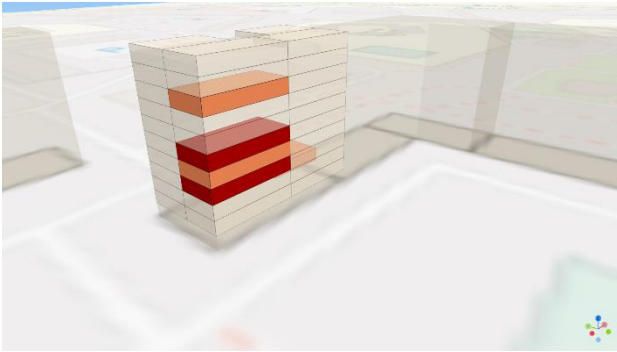


Figure 8. Linking hotbeds to apartments according to floor plans.

5. CONCLUSIONS

Without any doubt, the 3D visualization methodology allows to understand the essence of the processes of diseases development occurring in the studied areas. As a first approximation, the methodology should be tested on all the burdened areas of St. Petersburg. There are quite a few such territories both in the historical part of the city, and in new buildings on the outskirts and in the nearest agglomeration.

It is also necessary to consider the methodology application in more mass cases, when the number of buildings is measured in hundreds. This will require the introduction of new methods for creating vector layers and obtaining new data from district services. It is also required to calculate the feasibility of such measurements and the method of their effective implementation in the work of medical clinics. Now, we receive extremely positive feedback from medical specialists. According to many primary health care specialists, it is precisely such an understandable visualization of data that is sometimes lacking when analysing information in the discovered area.

Main benefit for the medical administrations and institutions can be gained as a result of the methodology implementation, is the capability of apartment-scale detection of the diseases hotbeds; and evaluation of hotbeds connectivity in the cases of multihotbed buildings.

In addition, relying on this methodology, we can quickly respond to mistakes and failures in the medical monitoring organization, which will reduce the time for making important decisions. It allows to aggregate a large amount of related information about various diseases. However, while we have tested only a principal possibility of 3D representation of the medical statistics data we are operating, the graphical/mapping visualization of multiple attributes can be recognized as a future work challenge.

In the future, it is planned to consider the possibility of integrating this visualization technique into specialised cartographic software product, the need for which is currently already ripe in the Russian medical community. Work is already underway in this direction. Commercial and government software developers do not have the necessary skills and specialists to create such solutions. Currently, there is an extremely shortage of medical information systems on the market with the ability to use maps and 3D data visualization.

In fact, developed methodology will serve not only as a source of displaying real data, but also will allow processing and calculating relative indicators of the prevalence of the disease for their further display on maps. 3D visualization should be

considered as a necessary addition in areas of a burdened epidemiological situation.

REFERENCES

- Chen, Y., Li, Q., Karimian, H., Chen, X., Li, X., 2021: Spatio-temporal distribution characteristics and influencing factors of COVID-19 in China. *Scientific Reports*, 11(1), 3717. doi:10.1038/s41598-021-83166-4
- Dutta, I., Basu, T., Das, A., 2021: Spatial analysis of COVID-19 incidence and its determinants using spatial modeling: A study on India. *Environmental Challenges*, 4, 100096. doi:10.1016/j.envc.2021.100096
- Federal law on personal data #152 of 27.07.2006 (in Russian)
- Gatrell, A.C., Bailey, T.C., 1996: Interactive spatial data analysis in medical geography. *Social Science and Medicine*, 42(6), 843-855. doi:10.1016/0277-9536(95)00183-2
- Gordon, A., Womersley, J., 1997: The use of mapping in public health and planning health services. *Journal of Public Health*, 19(2), 139-147. doi:10.1093/oxfordjournals.pubmed.a024601
- Huang, X., Wang, J., 2012: Analysis on the learning hot spots about AIDS evidence-based medical evidence abroad based on SPSS knowledge map. *Geomatics and Information Science of Wuhan University*, 37(SUPPL. 2), 70-73.
- Kuznetsov, I., Panidi, E., Kolesnikov, A., Kikin, P., Korovka, V., Galkin, V., 2020a. GIS-based infectious disease data management on a city scale, case study of St. Petersburg, Russia. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLIII-B3-2020, 1463-1467. doi:10.5194/isprs-archives-XLIII-B3-2020-1463-2020
- Kuznetsov, I., Panidi, E., Korovka, V., Galkin, V., Voronov, D., 2020b. Web-based representation and management of infectious disease data on a city scale, case study of St. Petersburg, Russia. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLIV-3/W1-2020, 87-91. doi:10.5194/isprs-archives-XLIV-3-W1-2020-87-2020
- Lesnykh, S.I., Mel'nikova, O.V., 2019: Generation of databases and visualization of current epidemiological information for purposes of a medical-ecological monitoring of a region. *Geography and Natural Resources*, 40(2), 115-121. doi:10.1134/S1875372819020033
- Mayer, J.D., 1983: The role of spatial analysis and geographic data in the detection of disease causation. *Social Science and Medicine*, 17(16), 1213-1221. doi:10.1016/0277-9536(83)90014-X.
- Wen, C., Huang, X., Feng, L., Chen, L., Hu, W., Lai, Y., Hao, Y., 2021: High-resolution age-specific mapping of the two-week illness prevalence rate based on the National Health Services Survey and geostatistical analysis: a case study in Guangdong province, China. *International Journal of Health Geographics*, 20(1), 20. doi:10.1186/s12942-021-00273-1