Spatio-Temporal Analysis of Premature Mortality Trends in the United States

Yelena Ogneva-Himmelberger
Department of International Development, Community and Environment
Clark University
Worcester, MA, USA
e-mail: yogneva@clarku.edu

***Abstract*— This paper applies geospatial data analytics to explore trends in premature mortality in the United States. Premature mortality, or Years of Potential Life Lost (YPLL), is one of the public health measures that focuses on deaths that happened at younger ages (before the age of 75 years) and thus could have been prevented. We used publicly available YPLL data for 2005-2016 for 3080 United States counties with spatio-temporal data mining tools in Geographic Information Systems (GIS) software to create a space-time cube and to find temporal trends in this measure. Our preliminary results indicate that 22% of counties experienced a statistically significant upward trend in YPLL, 24% experienced a statistically significant downward trend, and the remaining 54% did not experience any monotonic trend in YPLL measure. These findings can help county-level department of public health with developing targeted interventions to reverse upward trends in YPLL measure.**

Keywords-GIS; Spatial and spatio-temporal statistics; trend analysis; health; premature mortality.

I. INTRODUCTION

Premature mortality, or Years of Potential Life Lost (YPLL) is one of the public health measures that focuses on deaths that happened at younger ages (in the United States, before the age of 75 years) and thus could have been prevented. The Center for Disease Control collects annual county-level mortality data and provides state-level analysis. However, more detailed, local-level analysis is needed in order to conduct surveillance of temporal trends in premature mortality, and to evaluate the effectiveness of program interventions [1]. The goal of this paper is to identify temporal trends in premature mortality in the United States, using county-level data and space-time techniques in Geographic Information Systems (GIS). The rest of the paper is organized as follows: Section II provides a description of our data sources and methods. Section III reports our preliminary results, and Section IV discusses future research steps.

II. DATA AND METHODS

The YPLL measure was obtained from the County Health Rankings website [2] for 2005-2016. Since premature deaths are relatively rare events, YPLL is based on a three-year period, rather than on a single year. In this dataset, YPLL is a rate per 100,000 people and is age- adjusted to the 2000 US population. The rate is calculated as the number of total years of potential life lost for deaths that occurred amongst people who reside in a county under age 75, divided by the aggregate population under age 75 for the three years. The number of years lost is calculated for each death individually, and is based on the age at the time of death. The younger the person, the higher the number of years lost. To map the YPLL data, the GIS layer of county boundaries was downloaded from the U.S. Census Bureau [3] and YPLL tables were joined to the GIS layer using county Federal Information Processing Standards codes. There are 3142 counties in the United States, but 62 of them were missing YPLL data.

To identify spatio-temporal trends in YPLL, first, a space-time cube was created in ArcGIS Pro [4]. A space-time cube is a collection of spatial units (in this case, counties) layered vertically according to time. The bottom layer of the cube corresponds to 2005, the earliest year in the dataset, and the top layer of the cube corresponds to 2016, the latest year. Thus, a particular county at a given year is referred to as a bin within the space-time cube.

We applied the Mann-Kendall trend test to identify statistically significant temporal trends in YPLL for each spatial bin. This test compares values within each spatial bin over time and calculates changes between each consecutive time steps [5]. It identifies consistently increasing or decreasing trends over time at a location. Previous research applied this technique to analyze temporal trends in ground water level and precipitation [6]-[8], aridity [9], air temperature [10], traffic accidents [11] and crime [12]. We apply this widely used technique in a new context related to human health.

The output from the Mann-Kendall test is a spatial layer showing each spatial bin belonging to one of the seven categories: Up trend – 99% confidence; Up trend – 95% confidence; Up trend –90% confidence; Down trend – 99% confidence; Down trend – 95% confidence; Down trend – 90% confidence; No significant trend. The confidence level is determined based on the z score and the p-value of the trend.

To visualize the trends, we mapped seven trend categories and calculated the number of counties within each category.

 ISBN: 978-1-61208-762-7

III. RESULTS

To create the space-time cube, we first created a separate YPLL map for each year. An example of such map is provided in Figure 1.

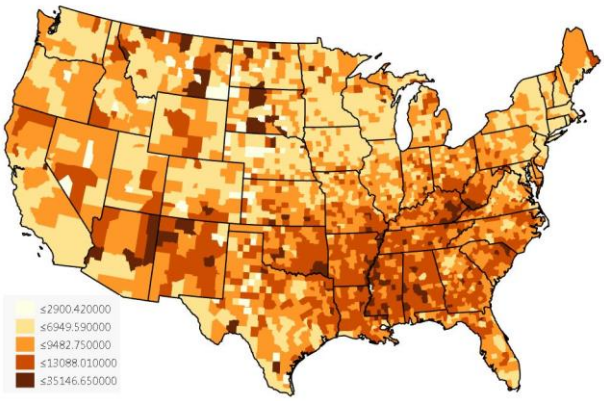

Figure 1. Premature mortality rate in continental United States in 2016.

Then, we combined all eleven maps into a space-time cube. It consisted of 3080 locations (counties) and 11 time slices, resulting in 33880 space-time bins. The minimum, mean, and maximum values of YPLL in the cube were 2817, 7963 and 35147 years/100,000 people, respectively. The Mann-Kendall test results are shown in Table 1 and Figure 2.

TABLE 1. NUMBER OF COUNTIES IN EACH TREND CATEGORY

Trend	Number of counties
decreasing - 99% confidence	346
decreasing - 95% confidence	258
decreasing - 90% confidence	149
no trend	1643
increasing - 90% confidence	157
increasing - 95% confidence	249
increasing - 99% confidence	278

These results indicate that 684 (22%) counties experienced a statistically significant upward trend in YPLL, 753 (24%) counties experienced a statistically significant downward trend, and the remaining 1643 (54%) counties did not experience any monotonic trend in YPLL measure. These findings are important, as they highlight areas with the alarming trend (statistically significant increase in YPLL rates over time) and can help local departments of public health develop targeted interventions to reverse these trends.

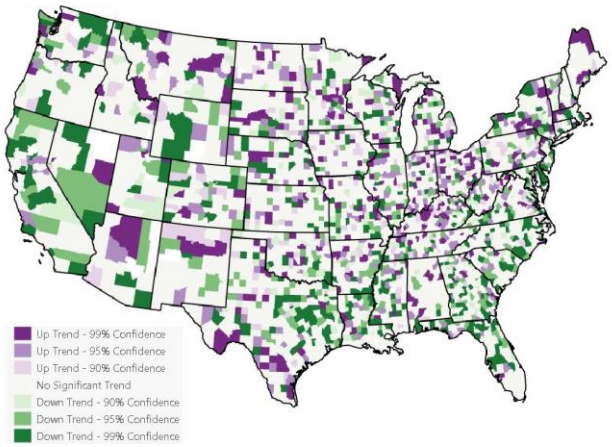

Figure 2. Mann-Kendall trends in premature mortality rates in continental United States, 2005-2016.

IV. CONCLUSION

In this paper, we identified temporal trends in premature mortality in the United States, using county-level data and space-time techniques in Geographic Information Systems. Our preliminary results indicate that in each state, there are counties with both upward and downward YPLL trends, sometimes next to each other. As the next step, it would be important to select several states for an in-depth analysis of the relationship between socio-economic and demographic factors and health outcomes to gain a better understanding of factors related to local variations in YPLL rate. This work could provide additional insights for more effective public health interventions.

REFERENCES

[1] Centers for Disease Control and Prevention. "Premature mortality in the United States: Public health issues in the use of years of potential life lost. MMWR Morb Mortal Wkly Rep. 1986;35(suppl 2), pp.1S-11S.

[2] County Health Rankings. Available from: https://www.countyhealthrankings.org/ [retreived: February, 2020]

[3] U.S. Census Bureau. Available from: https://www.census.gov/geo/maps-data/data/cbf/cbf_counties.html [retreived: February, 2020]

[4] ESRI, 2019. ArcGIS Pro Help. Available from: https://pro.arcgis.com/en/pro-app/help/main/welcome-to-the-arcgis-pro-app-help.htm [retreived: February, 2020]

[5] M. G. Kendall, Rank correlation methods, 2 ed. Oxford, England: Hafner Publishing Co. 1955.

[6] R. Agarwal and P. K. Garg, "Statistical assessment of groundwater resources and long term trend using geospatial techniques," 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 2016, pp. 1808-1811.

[7] A. Chandrakar, D. Khare, and R. Krishan, "Assessment of spatial and temporal trends of long term precipitation over

 ISBN: 978-1-61208-762-7

Kharun watershed, Chhattisgarh, India," Environmental Processes, vol. 4, no. 4, pp. 959-974, December 2017.

[8] S. Kumar, D. Machiwal, and D. Dayal, "Spatial modelling of rainfall trends using satellite datasets and geographic information system," Hydrological Sciences Journal, vol. 62, no. 10, pp. 1636-1653, July 2017.

[9] F. J. Moral, L. L. Paniagua, F. J. Rebollo, and A. García-Martín, "Spatial analysis of the annual and seasonal aridity trends in Extremadura, southwestern Spain," Theoretical and Applied Climatology, vol. 130, no. 3, pp. 917-932, November 2017.

[10] F. Viola, L. Liuzzo, L. V. Noto, F. Lo Conti, and G. La Loggia, "Spatial distribution of temperature trends in Sicily," International Journal of Climatology, vol. 34, no. 1, pp. 1-17, 2014.

[11] M. K. Rahman, T. Crawford, and T. W. Schmidlin, "Spatio-temporal analysis of road traffic accident fatality in Bangladesh integrating newspaper accounts and gridded population data," GeoJournal, July 2017.

[12] L. Ross, Y. Ogneva-Himmelberger, and C. Starr, "The use of geographic information systems for real-time monitoring of comprehensive community initiatives," Justice Research and Policy, April 2019; DOI: 10.1177/1525107119843259

 ISBN: 978-1-61208-762-7