A Spatial Analysis of Health and Longevity of Taiwan People

Abstract

Initiated in 1995, Taiwan's National Health Insurance (NHI) program now covers over 99.6% of its residents, ensuring widespread medical access. Despite this, regional disparities in medical resource allocation persist. This study investigates the potential urban-rural divide in life expectancy and healthcare utilization. Drawing data from the Ministry of the Interior (population registration records), NHI Research Database (medical utilization), and Ministry of Health and Welfare (leading causes of death), we employ spatial analysis, visualization tools, and the standardized mortality ratio for assessing regional disparities. Our findings reveal distinct regional mortality differences in Taiwan, with lower rates in northern counties and higher in mountainous regions. However, healthcare utilization shows no significant regional variations. Notably, patterns of overall mortality rates and primary death causes demonstrate spatial clustering.

Keywords: National Health Insurance, Regional Disparity, Standardized Mortality Ratio, Medical Utilization, Major Death Causes

1. Introduction

Since the end of World War II, Taiwan has undergone a rapid demographic transition. Economic growth and medical advancements have been pivotal in driving these demographic changes, which have led to steep reductions in fertility and mortality rates. Specifically, Taiwan's total fertility rate plummeted in the 1960s, dropping below the replacement level of 2.0 in 1985 and stabilizing between 1 and 1.2 after 2000. These declining fertility rates have transformed family structures, with the core family (a couple with or without children) becoming predominant. Traditionally, the elderly are cared for by family members; however, alternative arrangements are now necessary. Concurrently, reduced mortality rates point to an extended life expectancy. Intriguingly, Taiwan's trend in life extension differs from European and American countries, which have been decelerating recently. Since 1960, the life expectancy of Taiwan residents has grown by approximately 18 years, increasing steadily at a rate of 0.2 and 0.3 years annually since the early 1980s (Figure 1, with lines derived via locally weighted scatterplot smoothing, or LOWESS).

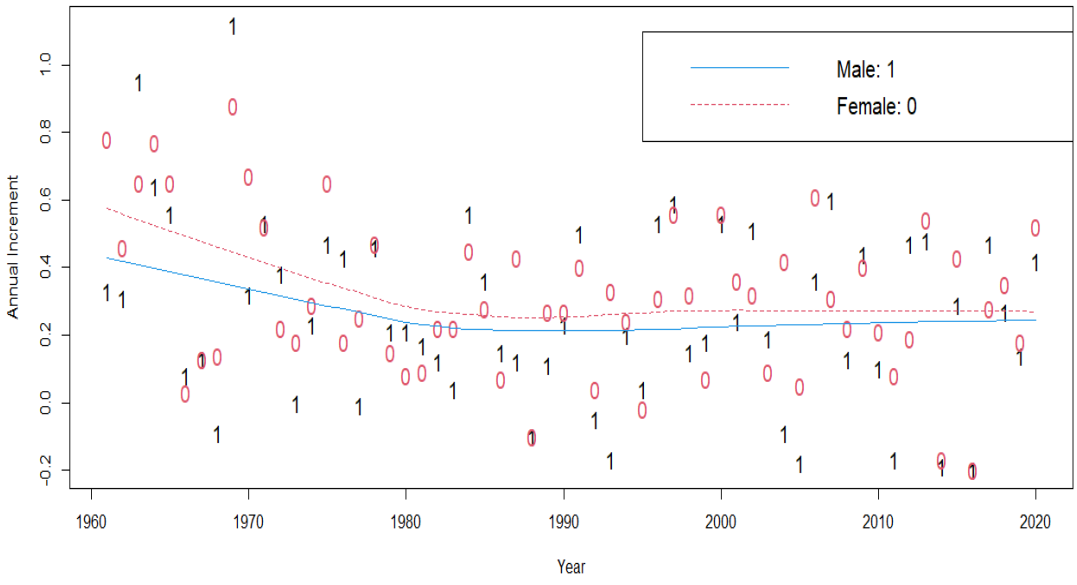

Figure 1. Annual Increment of Life Expectancy in Taiwan

Given the low fertility and mortality rates, the aging of Taiwan's population is inevitable. The percentage of individuals aged 65 and over is soaring: from 6.2% in 1990, it is projected to hit 20% by 2025 — a threefold increase in just 35 years (Source: National Development Council). Addressing the challenges posed by an aging demographic, Taiwan has rolled out and is further

developing various social insurance programs. The National Health Insurance (NHI) stands out as the most renowned, ensuring healthcare access for all, irrespective of socio-economic status. There's a consensus that the NHI has bolstered the health and lifespan of Taiwan's populace (Chang, 2012). Past research indicates that the NHI has diminished health disparities, with life expectancy rising more significantly for those with initially higher mortality rates (Wen et al. 2008). Keng and Sheu (2013) drew parallels, observing the most pronounced benefits among the least healthy elderly in areas such as mortality and self-assessed health. They also determined that the NHI enhances longevity and bridges the mortality gap between education levels, though it does not impact functional limitations uniformly.

Interestingly, regional differences in mortality in Taiwan seem to overshadow educational disparities. For instance, the male life expectancy in Taipei City (Taiwan's capital) was 81.43 in 2020, nearly nine years longer than in Taitung County (72.46) — a remote eastern region with a smaller populace. The disparity in female life expectancy between these regions is narrower, at about 7.6 years. This variation can largely be attributed to resource abundance in metropolitan areas, which offers superior medical care and employment prospects. In line with prior studies, examining the influence of Taiwan's social insurance on mortality becomes pertinent, especially in the wake of over 25 years of NHI implementation. Our goal is to ascertain if regional lifespan disparities have grown or diminished over time and assess if these findings can guide public policy and resource allocation decisions.

In this study, we apply spatial analysis tools to evaluate the health of Taiwanese individuals using data from the NHI Research Database (NHIRD), regional population figures from the Ministry of the Interior, and cause-of-death statistics from the Ministry of Health and Welfare. We aim to identify potential urban-rural disparities in life expectancy and medical utilization at both county and township levels, and to track how regional mortality disparity evolve. Given that no official estimates exist for life expectancy at the township level, we use the standardized mortality ratio (SMR) as the mortality measure. Additionally, we introduce two types of Geographically Weighted Regression (GWR) to study the relationship between mortality rates and healthcare utilization.

This paper is structured as follows: In Section 2, we detail our data and methodology, emphasizing the use of spatial analysis tools to determine whether regional mortality rates conform to spatial homogeneity conditions. Section 3 delves into the spatial characteristics of regional mortality rates and medical utilization. In Section 4, we shift our focus to an in-depth analysis of mortality rate trends, also integrating the application of GWR to township-level data to probe the

relationship between mortality rates and medical utilization. Finally, Section 5 offers our conclusions, discussing the spatial attributes of Taiwanese health.

2. Data and Methodology

Taiwan launched universal national health insurance in 1995, and currently, more than 99.6% of Taiwan's residents (population over 23 million) are enrolled in this program. The density of medical institutions in Taiwan is quite high, with over 20,000 medical institutions, including approximately 500 hospitals (with beds) and 20,000+ clinics (without beds). Every township has at least one medical institution and Taiwanese residents have good healthcare accessibility. The NHI has become a part of everyday life in Taiwan.

The Taiwanese government has collected health insurance-related data (e.g., inpatient and outpatient claims) from the NHIRD (Hsieh et al. 2019; Lee et al. 2021). Many researchers have used the NHIRD to study the health and medical utilization of the Taiwanese people (Hsing and Ioannidis 2015; Yue et al. 2018). We focus on outpatient records, including the number of visits, medical costs per visit, and total medical expenses per year. However, owing to the limitations of data acquisition, we can only assess NHI data from 2005 to 2013.

To analyze regional longevity, we employ two datasets provided by the Taiwanese government. One represents the household registration segmented by age and gender, while the other details all-cause and leading-cause mortality rates. These datasets originate from the Ministry of the Interior and the Ministry of Health and Welfare, respectively. We compare the all-cause and leading-cause mortality rates in different regions (with respect to counties and townships). Although the Taiwanese government regularly publishes official statistics on life expectancy, this data is available only at the national and county levels. Therefore, to differentiate longevity across counties and townships, we utilize age-specific mortality rates. Given the extensive range of age groups, our study adopted mortality indices for regional comparisons. To mitigate the influence of age distribution, we opted for standardized indices, which we will define subsequently.

In addition to spatial analysis tools, we apply Exploratory Data Analysis (EDA) to explore the health of Taiwanese residents at the county and township levels. EDA leverages statistical graphics and data visualization to dissect primary data characteristics without delving into formal modeling (Tukey 1977). Representing an interactive endeavor, EDA allows researchers to deploy a suite of analytical techniques, including data cleaning and aggregation. It's an essential preliminary step in data analysis. Despite its significance, EDA is complex and demands robust experience, domain knowledge, and sharp analytical skills. With the advent of Big Data, the prominence of EDA has surged (Milo and Somech 2020). Scholars have employed EDA to unearth

data errors and decipher correlations between pertinent variables. As a case in point, Indrakumari et al. (2020) utilized EDA on data concerning heart stroke and vascular disease. Their findings identified four classifications of chest pain, a vital variable in recognizing heart diseases.

There are two types of mortality standardization: direct and indirect. In this study, we chose the Standardized Death Rate (SDR) and Standardized Mortality Ratio (SMR), which are direct and indirect methods, respectively. Both standardization methods require a standard population. The SDR adjusts the age structure of the target population in the same manner as that of the standard population:

$$\mathrm{SDR}_j = \frac{\sum_x P_x^s \times m_x^j}{P^s}, \tag{1}$$

In this equation, P^s represents the size of the standard population, P_x^s denotes the size of age x in the standard population, and m_x^j indicates the central mortality rate of age x in population j. The SDR is a weighted average of the age-specific death rates of the standard population and is frequently expressed as the number of deaths per 100,000 individuals.

On the other hand, SMR is the most popular indirect method, often used in epidemiology, and it is defined as follows:

$$\mathrm{SMR}_j = \frac{D^j}{\sum_x P_x^j \times m_x^s}, \tag{2}$$

In this equation, D^j represents the observed number of deaths for population j; P_x^j denotes the population size of age x for population j; and m_x^s indicates the central mortality rate of age x for the standard population. $\mathrm{SMR_j}$ is used to compare the mortality rates of population j with those of the standard population; if $\mathrm{SMR_j} = 1$, the overall mortality rates of population j are approximately the same as those of the standard population. Similarly, if $\mathrm{SMR_j}$ is smaller or larger than one, this implies that the overall mortality rates of population j are lower or higher than those of the standard population, respectively. The possible range of $\mathrm{SMR_j}$ is $(0, \infty)$.

We consider two types of spatial analysis: one for judging spatial homogeneity and the other for describing the spatial correlation between the objective variable and explanatory variables. Moran's I (Moran, 1950) is the most popular index for measuring spatial autocorrelation, and is defined as follows:

$$I = \frac{n}{\sum_i \sum_j w_{ij}} \times \frac{\sum_i \sum_j w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{\sum_i (x_i - \bar{x})^2} \tag{3}$$

In this equation, x_i denotes the observed value of region i, $\bar{x}$ represents the average, w_{ij} is the weight between regions i and j, and n indicates the number of spatial units indexed by i and j.

Moran's I is usually used to measure spatial autocorrelation and it ranges between −1 and 1. A positive Moran's I value signifies positive spatial autocorrelation. The closer this value is to one, the stronger the relationship between neighboring regions. If Moran's I significantly deviates from 0, it suggests the presence of hotspots, which are areas with elevated mortality rates, or potential spatial correlations. Moran's I offers reliable testing power when evaluating spatial homogeneity (Bivand et al. 2009). Yet, the efficacy of this testing power can be swayed by factors like population size and the way "neighborhood" is defined (Oden 1955). In spatial metrics such as Moran's I, contiguity-based spatial weights are frequently utilized. Contiguity here refers to two spatial units that share a boundary. The "Queen" and "Rook" are prevalent contiguity measures, named after their analogous movements on a chessboard. The number of neighboring units identified by the "queen" method will always be the same as or exceed those identified by the "rook" method. Hence, the "queen" approach encapsulates more spatial information, resulting in greater analytical power. For our analysis, we've selected the "queen" method for determining contiguous neighbors.

GWR can be treated as a spatial version of the linear regression model, with a dependent variable *y* expressed via a linear function of a set of *p* independent variables, $x_1, x_2, \ldots, x_p$, or as follows:

$$\mathrm{Y}_i = \beta_{i0} + \sum_{k=1}^{p} \beta_{ik} x_{ik} + \varepsilon_i \tag{4}$$

In this equation, β_{ik} and x_{ik} represent the parameters and observed values of the independent variables $k\,(k = 1, \ldots, p)$, whereas, ε_i denotes the error terms at location *i* (*i* = 1, 2, … *n*). In other words, each location had its own regression model in the GWR model. In addition, ε_i is generally assumed to follow a normal distribution with zero mean and constant variance σ^2 (i.e., $\varepsilon_i \sim N(0, \sigma^2)$), similar to that in linear regression. As there are more parameters than observations in GWR, the method of parameter estimation is slightly different. We posit that nearby data of each location usually have similar attributes; thus, the Weighted Least Squares (WLS) method is suitable for parameter estimation, with different weightings for each location.

The coefficient estimates of all GWR variables are obtained from a moving data window, and all parameter estimates are derived from a fixed range of observations. The optimal width (i.e., bandwidth) of the moving windows in a GWR is usually determined by cross-validation; however, the GWR estimation can still be unsatisfactory and produce a distorted relationship. Most modifications to GWR are based on the selection of bandwidth, and choosing different bandwidths

for each independent variable is a possible alternative approach. Multiscale geographically weighted regression models (MGWR) and conditional geographically weighted regressions (CGWR) are two methods for adjusting the bandwidth (Fotheringham et al., 2017; Leong and Yue, 2017). We apply the original GWR and CGWR to Taiwanese data and compare their estimation results.

3. Exploratory Data Analysis of County Data

In this section, we apply EDA tools (especially visualization tools) to explore the spatial attributes of mortality and healthcare utilization among Taiwanese residents. Taiwan has 20 counties, 368 townships, and slightly more than 23 million people. In general, the population density of west coast townships is higher, and not many residents live in middle and eastern Taiwan or mountainous areas (Figure 3-1). For example, by the end of 2021, the most crowded township was Yonghe District (in New Taipei City), with over 36,000 people per square kilometer, and the least populated township was Taoyuan District (in Kaohsiung City), with only five people per square kilometer. The regional difference in population density is surprisingly large, approximately 7,000 times. Interestingly, Taiwan is not a big island (with an area around 36,000 square kilometers) but the regional differences in mortality are quite significant, as detailed by the results generated using the SMR as an EDA tool.

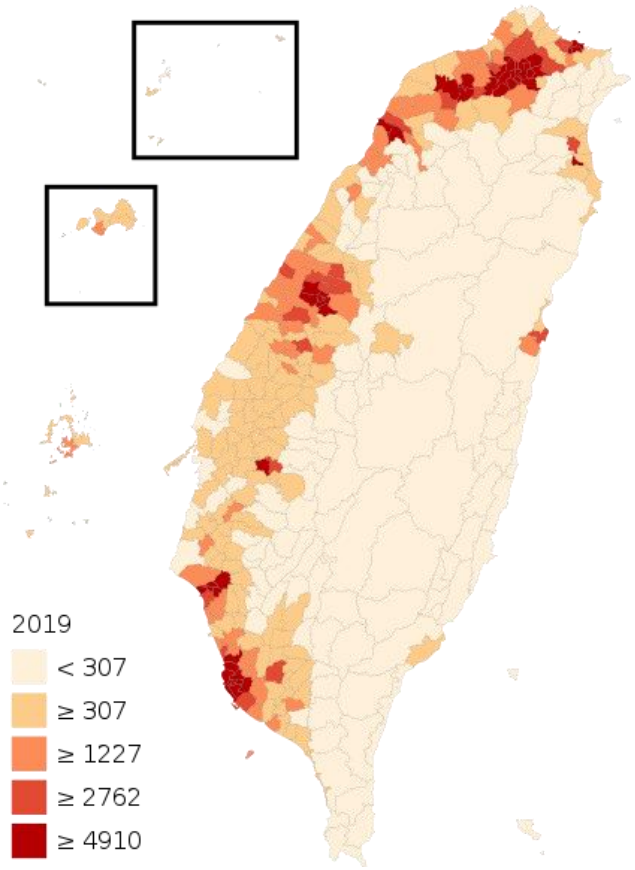

Figure 3-1. Population Density of Taiwan Township (Unit: people per km^2)

Source: Wikipedia

The boxplot of county-level SMR from 1974 to 2018 is depicted in Figure 3-2. Each year's SMR uses that year's national population as the standard. Counties with smaller populations, like Penhu and Taitung (100,000 to 220,000 residents), show more variations. Taipei City has the lowest SMR at about 0.8, while Taitung County peaks at around 1.3. This suggests that Taipei's mortality rate is roughly 60% of Taitung's. Significant SMR disparities also exist among the six major cities with populations over 2 million, as illustrated by female mortality rates in Figure 3-3. For instance, Taipei residents have a mortality rate about 70% of those in Kaohsiung.

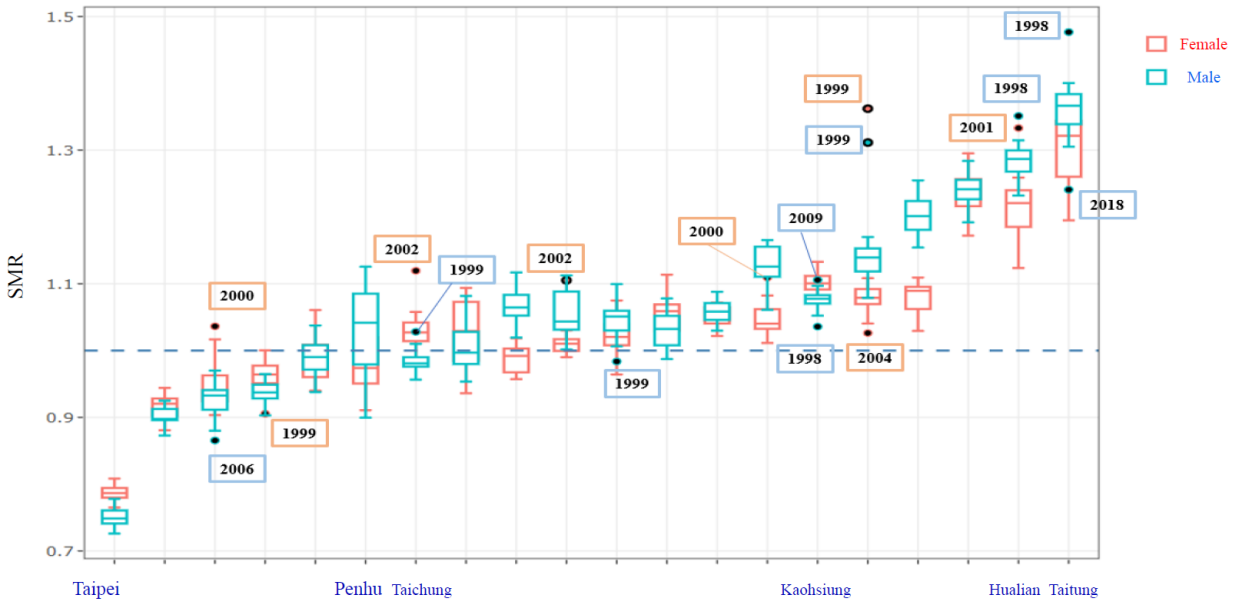

Figure 3-2. SMR of Taiwan's Counties (1974-2018)

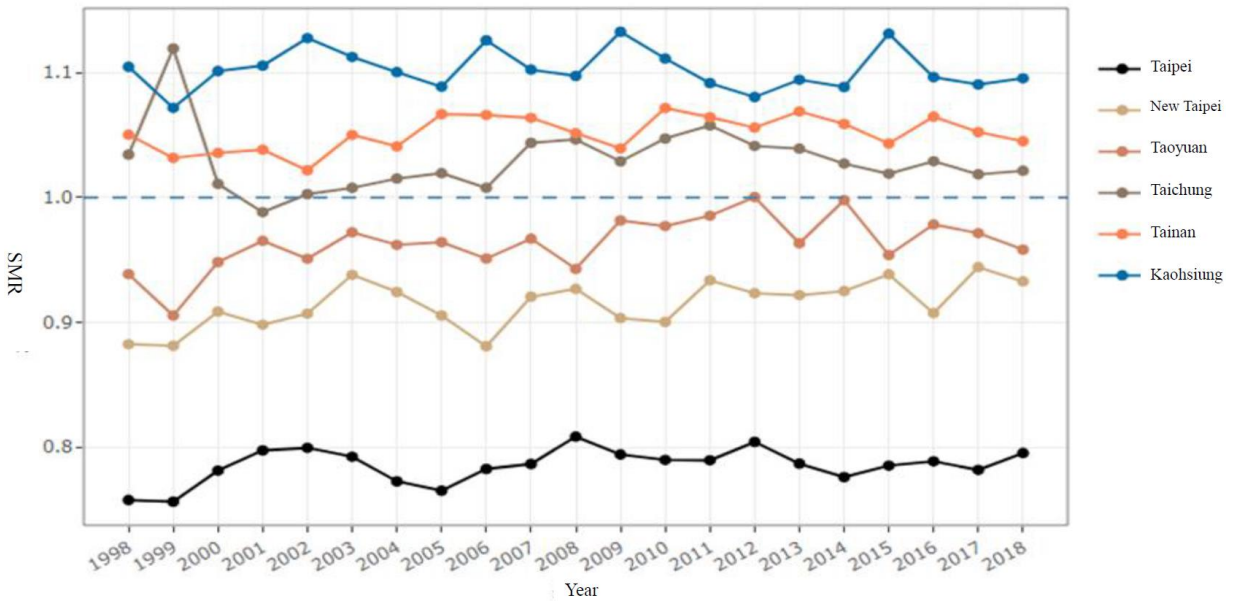

Figure 3-3. SMR of Taiwan's Six Major Counties (1998-2018)

EDA aids in hypothesis formulation and offers data collection guidance, vital for data analysis. The SMR highlights significant mortality disparities between counties, prompting exploration into associated factors. Initially, we sort the SMR by latitude, with Figure 3-4 presenting the 2020 county SMR. Taiwan, positioned between 22.5°–25°N latitude and 119.5°–121°W longitude, experiences a subtropical climate in the north and tropical in the south. Lifespan often shows a negative correlation with temperature. This holds true in Taiwan, as mortality rates (SMR) in counties decline with increased latitude. Typically, northern counties (in green) register the lowest mortality rates, in contrast to the higher rates in southern and eastern (warmer) counties. Figure 3-

4, a bubble plot, introduces a third dimension, population, to the scatterplot — a format the United Nations and OECD frequently utilize. Evaluating the 2020 SMR by population (Figure 3-5), we discern a negative correlation with population size. However, the association with population is slightly less pronounced than with latitude. Notably, the SMRs of major southern cities, Tainan and Kaohsiung, exceed those in northern non-major regions like Hsinchu City and Hsinchu County.

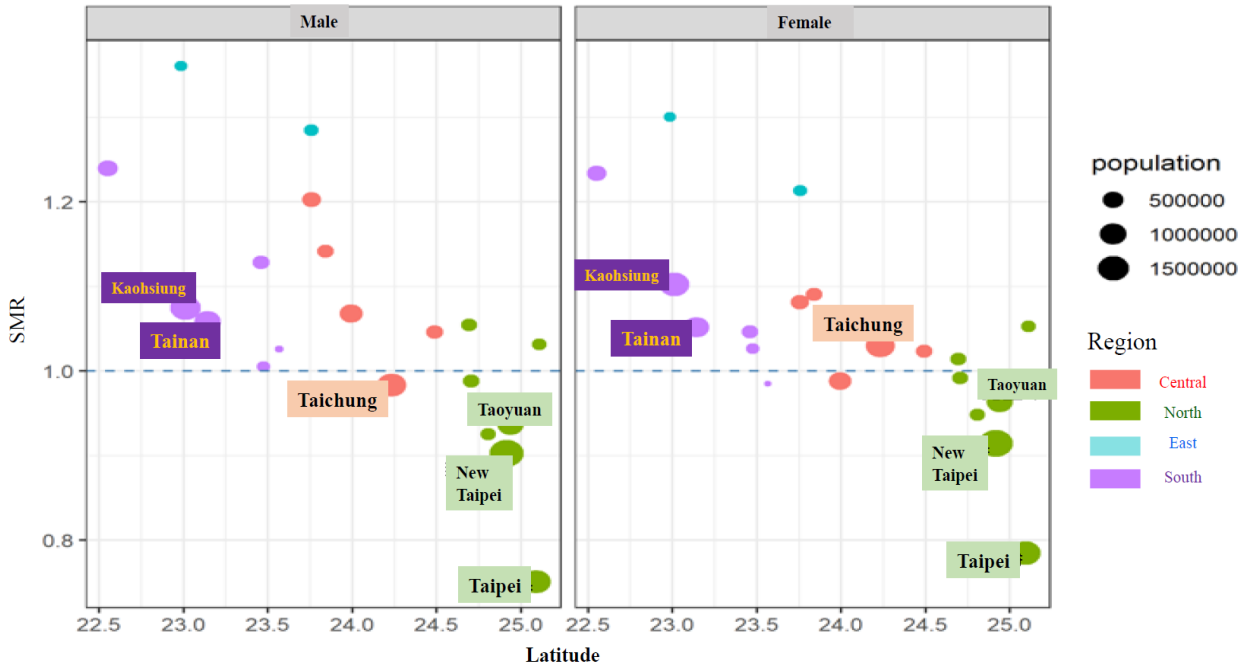

Figure 3-4. County SMR and Latitude (2020)

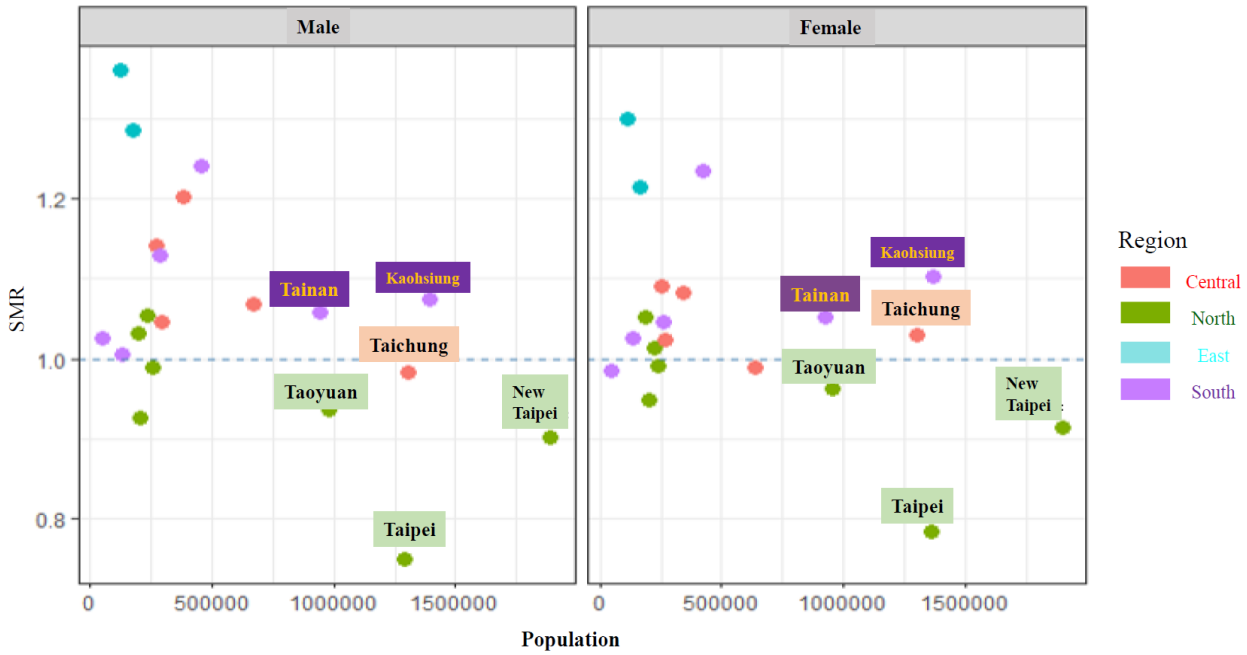

Figure 3-5. County SMR and Population (2020)

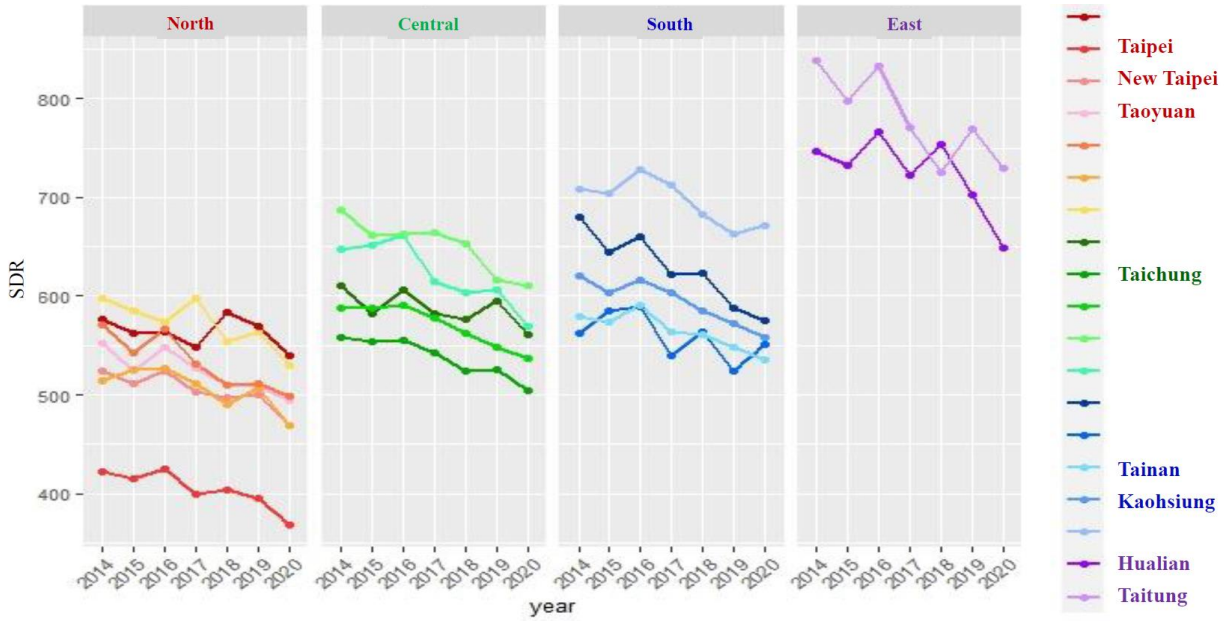

Figure 3-6. SDR of Different Regions (2014-2020)

In the analysis mentioned, we could employ the SDR, though the relative disparities between counties appear less pronounced. However, the SDR facilitates more straightforward detection of temporal shifts in mortality. We base the standard population for SDR calculation on the World Health Organization's WHO2000 (Ahmad et al. 2001). Figure 3-6 displays the SDR for four Taiwanese regions from 2014 to 2020. All counties present declining mortality rates, with similar rates of decrease. While regional mortality differences stand out, the rate of mortality changes across counties remains consistent.

To probe the rate of change in mortality rates further, we might consider the correlation of county SDR across different years (Figure 3-7). The intent behind Figure 3.7 is to highlight the substantial correlation between the SDRs of any two given years, irrespective of gender. Although the correlation coefficients for females are slightly lower than for males, all values are close to or surpass 0.7, signaling a robust linear relationship. This suggests that county mortality disparities exhibit little variation over time, consistent with the observations from Figure 3-6.

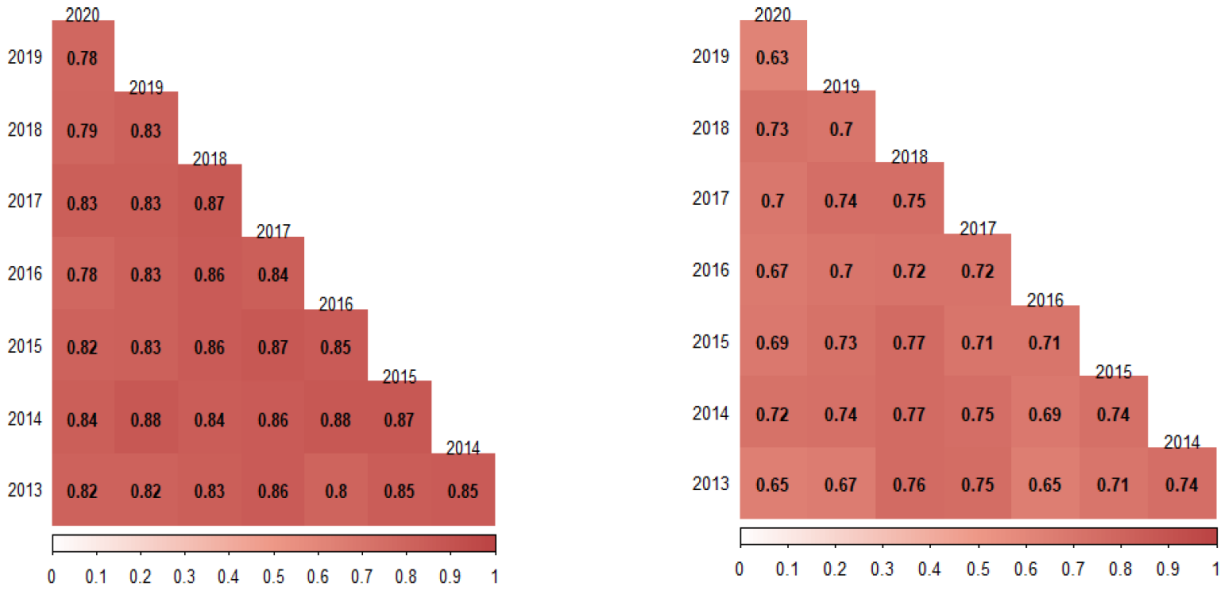

Figure 3-7. Correlation of County SDR (2013-2020; Left: Male, Right: Female)

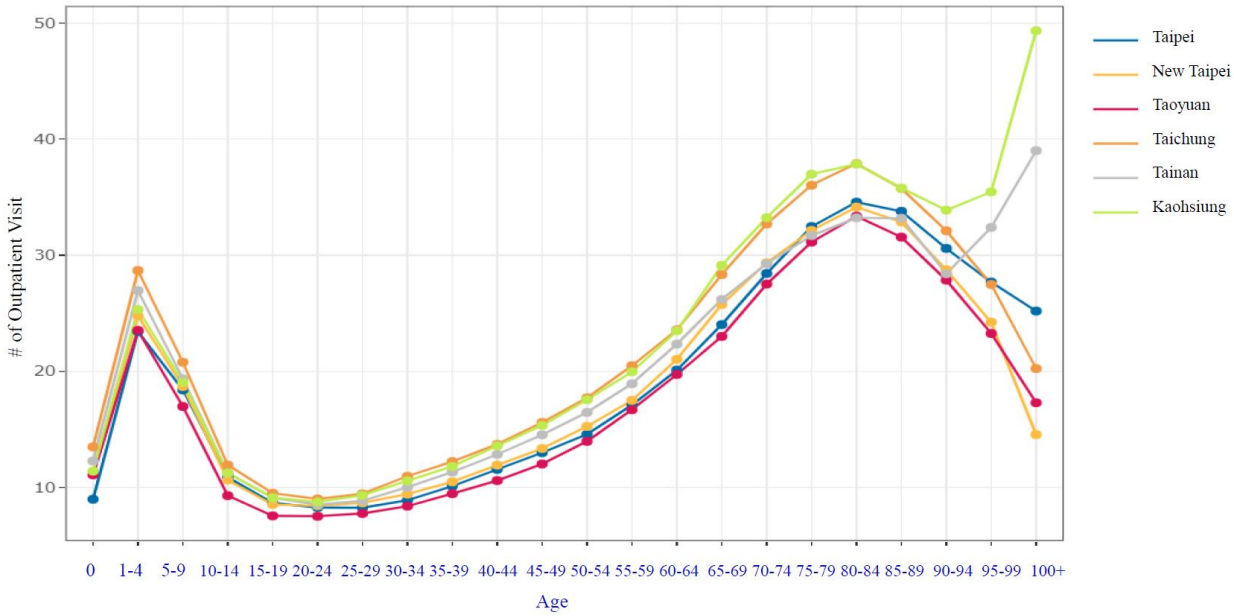

Figure 3-8. Numbers of Outpatient Visit in Six Major Cities (2005-2013)

Unlike mortality rates, county-level medical utilization does not show clear systematic trends with respect to counties. We use the number of outpatient visits in six major cities as an example (Figure 3-8). The number of outpatient visits is highly correlated with age and increases monotonically from ages to 25-29 to 80-84. In addition, age group 1-4 had the most outpatient visits among the younger age groups. However, the number of outpatient visits is not related to county. This pattern diverges from Figures 3-3 and 3-4, where major southern cities report elevated mortality rates. It's important to mention that the results for outpatient visits span only from 2005 to 2013 owing to constraints in accessing NHI data.

4. Spatial Analysis of Township Health and Mortality

In the previous section, we explore the EDA of county-level data. Here, we apply spatial analysis to township-level data. The decision to forgo spatial analysis for county data stems from an insufficient number of data points (20 counties), while the township data, with 368 points, is ample. Using Moran's I, we assess the spatial homogeneity of township mortality in 2020 (Figure 4-1). Townships in central mountainous areas display elevated mortality rates, potentially indicating a hotspot or cluster. Both male and female cases have significant p-values for Moran's I, revealing that the distribution of township SMR is spatially heterogeneous, suggesting spatial clustering. The disparity between townships with the highest and lowest SMR values, both for men and women, is pronounced, far exceeding county differences seen in Figure 3-2.

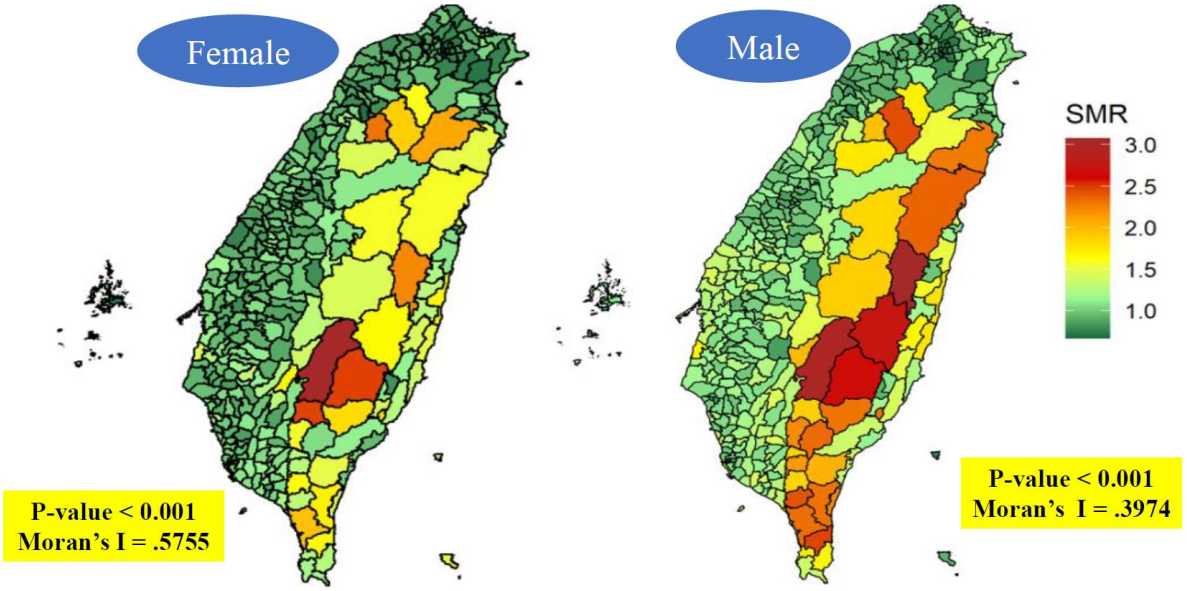

Figure 4-1. Township SMR and Moran's I (2020)

Further, we apply Moran's I to the SMR for five leading causes of death in Taiwan (Table 4-1). All tests return significant Moran's I values, negating the spatial homogeneity hypothesis. While spatial clustering of cancer mortality rates aligns with prior research (Hu and Lay, 2006), the findings for other primary death causes are unexpected. We theorize that certain areas may have environmentally related risk factors driving higher mortality rates, but this does not seem to apply universally. With traffic and falls accounting for approximately 45% and 20% of accidental deaths respectively, certain townships appear more susceptible to fatal car accidents. This necessitates further research into potential environmental risk factors for leading causes of death.

Table 4-1. Township SMR of Major Death Cause and Moran's I (2020)

SMR	Moran's I	p-value
Cancer	0.379	< 0.05
Heart Disease	0.340	< 0.05
Accident	0.275	< 0.05
Cerebrovascular Disease	0.156	< 0.05
Pneumonia	0.117	< 0.05

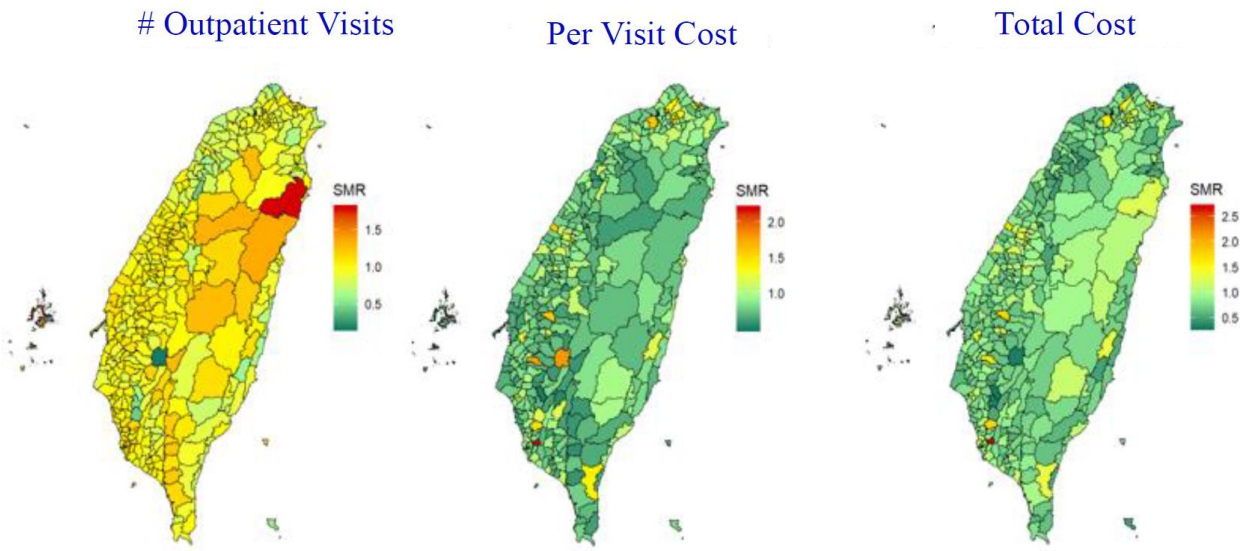

Figure 4-2. Medical Utilization (SMR) and Moran's I (2020)

Adapting the SMR definition in Equation (2), we can also define standardized indices for medical utilization. For example, to acquire a standardized number of outpatient visits, we can use the number of total outpatient visits for population j to replace D^j, and the number of outpatient visits for people aged x in the standard population to replace m_x^s. To simply the discussion, we still use the term "SMR" for the standardized medical utilization. Moran's I statistics for the number of outpatient visits, cost per outpatient visit, and total cost of outpatient visits indicates spatial homogeneity (Figure 4-2). Like the analysis at the county level, no obvious regional differences, in healthcare utilization at the township level, exist.

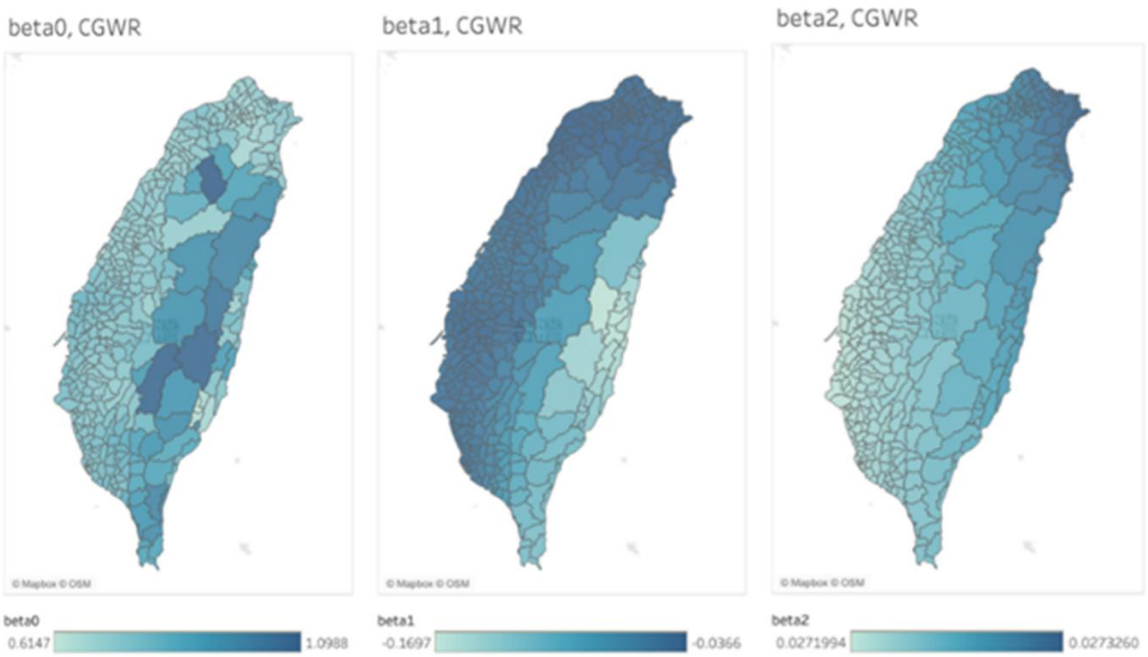

Figure 4-3. Parameter Estimates of CGWR Model

We use township SMR as the target variable Y in the GWR analysis. We choose population, population density, number of hospitals (not clinics), number of beds, and medical utilization as independent variables (three variables in Figure 4-2). After a series of variable selections, we find that the population (X_1) and number of outpatient visits (X_2) are the most relevant variables; thus, we keep them in the GWR models. As stated previously, the GWR model assigns distinct coefficients to each region, allowing independent variable impacts to vary regionally. For example, the coefficient of population X_1 (β_1 in Figure 4-3) is negative and less on the east coast than the west coast. This signifies a more pronounced decrease in SMR value concerning population size on the east coast compared to the west coast.

The R^2 (or R-squared) values of GWR and CGWR models are 0.84 and 0.94, respectively. As the fitted result of the CGWR is slightly better, we use it to explain the parameter estimation results. The estimated intercept is very close to the SMR (Figure 4-1) and townships in mountainous areas have higher mortality rates. The coefficient of population (β_1) is negative, indicating that a larger population leads to lower mortality, aligning with Figure 3-5. In addition, β_1 increases from east to west, indicating that the marginal effect of the population is stronger in the east. However, the coefficient of the number of outpatient visits (β_2) is positive, indicating that more outpatient visits are associated with higher mortality and that its marginal effect is stronger in northeast townships.

Residual diagnostics are performed to evaluate the GWR and CGWR models. To simply the discussion, we solely focus on the scatterplot of fitted values and standardized residuals, or $\hat{y}_i$ vs.

$\hat{\varepsilon}_i$ (Figure 4-4). If there are too many outliers or extreme outliers (i.e., values that are too large or too small), the model may not be suitable. In particular, we often focus on standardized residuals larger or smaller than three, and the probability of observing these types of outliers is approximately 0.3% under the normality assumption. The GWR model has six outliers with values outside (–3, 3), and one outlier is close to six, suggesting that the GWR model requires some adjustments. By contrast, the CGWR model has three outliers outside (–3, 3) and no extreme outliers.

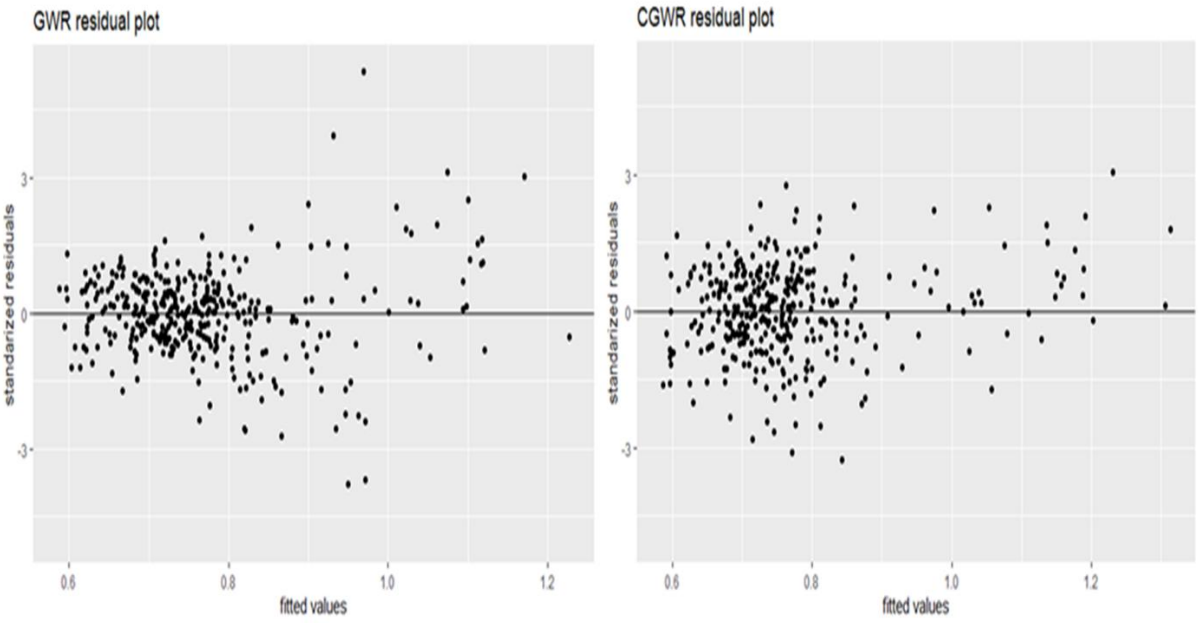

Figure 4-4. Residual Plots of GWR and CGWR Models

5. Discussions and Conclusion

Regions across countries, even in smaller ones like Taiwan, exhibit variations in death rates. For men and women, the life expectancy gap at the county level approaches 9 and 8 years, respectively. As Taiwan's government dedicates resources to bridging regional disparities, we anticipate a gradual reduction in these mortality differences. Recently, Taiwan implemented the National Health Insurance (NHI), a policy contributing to the extended life expectancy of its residents. Yet, the trajectory of regional disparity remains uncertain. To gain insight, we employ tools in EDA and spatial analysis, using data from the NHI Research Database and two other government sources.

Our findings, through EDA, reveal persistent regional differences in mortality within Taiwan but no notable variances in medical utilization. Typically, mortality rates drop in northern regions

and escalate in mountainous territories. Areas with larger populations experience reduced mortality rates. For township-level data spatial analysis, consistent spatial clustering appears across overall and major causes of death. The consistency between EDA and GWR results boosts our confidence in the analysis. Our study's empirical findings suggest potential hotspots with higher mortality risk in Taiwan. Furthermore, GWR models emphasize the influence of population size and outpatient visits on mortality rates, along with regional variances.

Spatial analysis tools, primarily developed for regional exploration, can also chart mortality rate time trends using standardization indices such as SMR and SDR. A commonality in the mortality improvement trend across counties suggests that regional mortality differences persist. This finding prompts questions about resource allocation strategies. Should resources shift to high mortality areas or do existing policies suffice? The mortality rates are related not only to the urban-suburban divide but also to geographical differences between the north and south. This suggests that policies should adopt a localized strategy to effectively address these disparities. For example, there are 23 medical centers with relatively abundant medicals resources, only two of which are not in the north or major cities. Also, persisting regional differences in mortality could lead to significant life expectancy gaps, presenting potential insurance risks for life insurers.

The consistent trend in mortality improvement across counties can refine mortality models for Taiwan's smaller areas. The Lee-Carter model, for instance, may produce skewed estimates for limited populations. Enhancing its accuracy might involve incorporating data from nearby counties, given the uniformity in mortality trends. Exploring this method's feasibility remains a focus for future research. Moreover, adaptations, like a spatiotemporal modification of the Lee-Carter model, warrant consideration.

Following the Personal Data Protection Act, Taiwan's government revises the access guidelines to NHI data in 2018. Researchers now assess NHI data on-site at the Health and Welfare Data Science Center, Ministry of Health and Welfare. The NHI data for this study comes from before 2018, covering the period 2005 to 2013. Since the data does not receive updates, the medical utilization analysis might not reflect current trends. Efforts to incorporate NHI data post-2014 face challenges due to the new regulations that demand additional time, funds, and personnel. Besides NHI data, exploration continues for other reliable data sources to understand the health and medical utilization of Taiwanese residents at the county and township levels. Additionally, this study employs an isotropic kernel function in the GWR. However, Taiwan's central mountain range may introduce direction-dependent (anisotropic) spatial properties in mortality patterns. It remains advisable to compare analysis outcomes using both isotropic and anisotropic models.

References

Ahmad, O.B., Boschi-Pinto, C., Lopez, A.D., Murray, C.J., Lozano, R., and Inoue, M. (2001). Age Standardization of Rates: A New WHO Standard. Geneva: *World Health Organization*, 9(10), 1–14.

Bivand, R., Müller, W.G., and Reder, M. (2009) Power Calculations for Global and Local Moran's I. *Computational Statistics and Data Analysis*, 53: 2859-2872.

Chang, S. (2012) The Effect of Taiwan's National Health Insurance on Mortality of the Elderly: Revisited. *Health Economics*, 21: 1257–1270.

Fotheringham, A.S., Yang, W., and Kang, W. (2017) Multiscale Geographically Weighted Regression (MGWR). *Annals of the American Association of Geographers*, 107(6): 1247–1265.

Fu, W.J., Jiang, P.K., Zhou, G.M., & Zhao, K.L. (2014) Using Moran's I and GIS to Study the Spatial Pattern of Forest Litter Carbon Density in a Subtropical Region of Southeastern China. *Biogeosciences*, 11(8): 2401–2409.

Hsieh, C., Su, C., Shao, S., Sung, S., Lin, S., Kao Yang, Y., and Lai, E.C. (2019) Taiwan's National Health Insurance Research Database: past and future. *Clinical Epidemiology*, 11:349–358.

Hsing, A.W. and Ioannidis, J.P.A. (2015) Nationwide Population Science Lessons from the Taiwan National Health Insurance Research Database. *JAMA Internal Medicine*, 175(9): 1527–1529.

Hu, L. and Lay, J. (2006) Spatial Analysis of Female Cancers in Taiwan. *Journal of Taiwan Geographic Information Science*, 4: 39–55.

Indrakumari, R., Poongodi, T., and Jena, S.R. (2020) Heart Disease Prediction using Exploratory Data Analysis. *Procedia Computer Science*, 173: 130–139.

Keng, S. and Sheu, S. (2013) The Effect of National Health Insurance on Mortality and the SES–Health Gradient: Evidence from the Elderly in Taiwan. *Health Economics*, 22: 52–72.

Lee, R.D. and Carter, L.R. (1992) Modeling and Forecasting US Mortality. *Journal of the American Statistical Association*, 87(419): 659–671.

Lee, P., Kao, F., Liang, F., Lee, Y., Li, S., and Lu, T. (2021) Existing Data Sources in Clinical Epidemiology: The Taiwan National Health Insurance Laboratory Databases. *Clinical Epidemiology*, 13:175–181.

Leong, Y.-Y. and Yue, J.C. (2017) A Modification to Geographically Weighted Regression. *International Journal of Health Geographics*, 16(1).

Li, N. and Lee, R.D. (2005) Coherent Mortality Forecasts for a Group of Populations: An Extension of the Lee-Carter Method. *Demography*; 42(3): 575–594.

Milo, T. and Somech, A. (2020) Automating Exploratory Data Analysis via Machine Learning: An Overview. *Proceedings of the 2020 ACM SIGMOD International Conference on Management of Data*, 2617–2622.

Moran, P. (1950) A Test for the Serial Independence of Residuals. *Biometrika*, 37(1/2): 178–181.

Oden, N. (1955) Adjusting Moran's I for population density. *Statistics in Medicine*, 14(1): 17–26.

Shi, H., Zhang, L., and Liu, J. (2006) A New Spatial-attribute Weighting Function for Geographically Weighted Regression. *Canadian Journal of Forest Research,* 36(4): 996–1005.

Shi, Y. (2021) Forecasting Mortality Rates with the Adaptive Spatial Temporal Autoregressive Model. *Journal of Forecasting*, 40(3): 528–546.

Tango, T., and Takahashi, K. (2005) A Flexibly Shaped Spatial Scan Statistic for Detecting Clusters. *International Journal of Health Geographics, 4*(1), 11.

Tukey, J.W. (1977) Exploratory Data Analysis, Addison-Wesley Publishing Company.

Wang, H.C. and Yue, J.C., and Chong, C. (2018) Mortality Models and Longevity Risk for Small Populations. *Insurance: Mathematics and Economics*, 78: 351–359.

Wen, C.P., Tsai, S.P., and Chung, W.I. (2008). A 10-Year Experience with Universal Health Insurance in Taiwan: Measuring Changes in Health and Health Disparity. *Annals of Internal Medicine,* 148(4): 258–267.

Yu, H., Fotheringham, A.S., Li, Z., Oshan, T., Kang, W., and Wolf, L.J. (2019) Inference in Multiscale Geographically Weighted Regression. *Geographical Analysis,* 52: 87–106.

Yue, J.C., Wang, H., Liang, Y., and Su, W. (2018) Using Taiwan National Health Insurance Database to Model Cancer Incidence and Mortality Rates. *Insurance: Mathematics and Economics*, 78: 318–324.

Yue, J.C., Wang, H., and Wang, T. (2021) Using Graduation to Modify the Estimation of Lee-Carter Model for Small Populations, *North American Actuarial Journal*, 25: sup1, S410–S420.