



# Agent Based Model in SAS Environment for Rail Transit System Alignment Determination

I Made Indradjaja Brunner<sup>1</sup>

Teknik Lingkungan, Universitas Agung Podomoro  
Jakarta, Indonesia

Artikel Masuk : 9 Januari 2018

Artikel Diterima : 23 Januari 2018

Tersedia Online : 29 April 2018

**Abstract:** Transit system had been proposed for the urban area of Honolulu. One consideration to be determined is the alignment of the transit system. Decision to set the transit alignment will have influences on which areas will be served, who will be benefiting, as well as who will be impacted. Inputs for the decision usually conducted through public meetings, where community members are shown numbers of maps with pre-set routes. That approach could lead to a rather subjective decision by the community members. This paper attempts to discuss the utilization of grid map in determining the best alignment for rail transit system in Honolulu, Hawaii. It tries to use a more objective approach using various data derived from thematic maps. Overlaid maps are aggregated into a uniform 0.1-square mile vector based grid map system in GIS environment. The large dataset in the GIS environment is analyzed and manipulated using SAS software. The SAS procedure is applied to select the location of the alignment using a rational and deterministic approach. Grid cells that are superior compared to the others are selected based on several predefined criteria. Location of the dominant cells indicates possible transit alignment. The SAS procedure is designed to allow a transient vector called the GUIDE (Grid Unit with Intelligent Directional Expertise) agent to analyze several cells at its vicinity and to move towards a cell with the highest value. Each time the agent landed on a cell, it left a mark. The chain of those marks shows location for the transit alignment. This study shows that the combination of ArcGIS and SAS allows a robust analysis of spatial data and manipulation of its datasets, which can be used to run a simulation mimicking the Agent-Based Modelling. This study also opens up further study possibilities by increasing number of factors analyzed by the agent, as well as creating a composite value of multi-factors.

**Keywords:** GIS, Hawaii, SAS, Transit Alignment Analysis

## Introduction

The urban area of the City and County of Honolulu is the most densely populated region in the state of Hawaii. In 2016, there were estimated 987,000 population lived on Oahu Island, where the City and County of Honolulu are located (United States Census Bureau, 2018). They settled in an island of 1,545 square km in the middle of the Pacific Ocean. Motor vehicles were the main transportation means on the island where more than

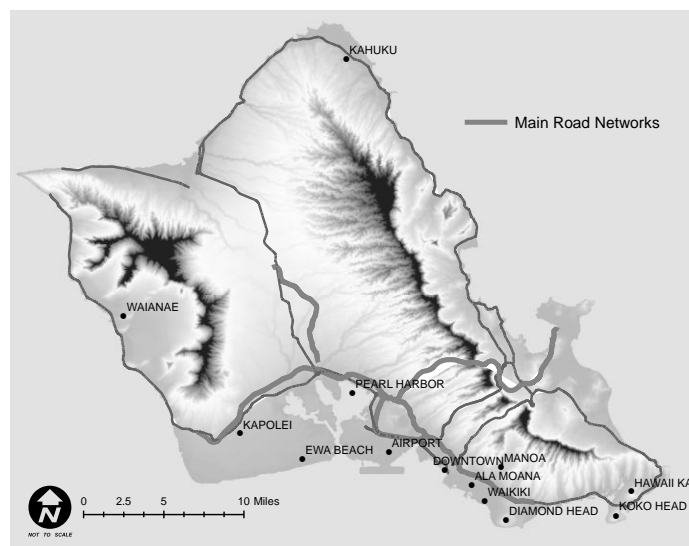
---

<sup>1</sup> Korespondensi Penulis: Teknik Lingkungan, Universitas Agung Podomoro, Jakarta, Indonesia  
Email: imade.brunner@gmail.com

770,000 registered vehicles were reported in 2015 (DBEDT, 2017). The essential freeway is the H1 which extends along the East to West corridor on the South shore of the island. The H1 is serving commuters from the second city of Kapolei in the West, and commuters from Hawaii Kai area in the East to the urban center of Honolulu in the middle. Traffic congestion is a daily routine during morning and afternoon peak hours along this freeway. A couple of years ago, to alleviate the congestion and shorten commute time, a fully automatic guided rail system was planned to accommodate around a half of a million Honolulu residences and tourist as well. The proposed rail transit is to serve commuters from Kapolei to urban Honolulu as well as the Honolulu International airport. Figure 1 depicts a map of Oahu which shows the location of the main roadway networks, including the West to East corridor of H1 freeway connecting Kapolei, the airport, Honolulu downtown, Waikiki, Hawaii Kai, among others.

The proposed transit alignments for the rail system are exposed to the community through several public meetings. During the meetings, the program initiator shows several possible alignments printed in large maps which can be viewed by the community. Later, public will be asked to select which alignments suit them best. This is a common practice in gathering public opinion to select an alternative from many choices. Nonetheless, this practice tends to see the option based on romanticism, attachments, and visual recognition of the sites rather than understanding the underlying conditions or actualities.

The study was conducted based on the transit rail system in Honolulu, Hawaii. This paper attempts to find out other possible ways to determine transit alignments by avoiding subjectivity judgments. It describes the use of GIS and SAS program in determining the best possible route for rail transit alignment. The Geographic Information System (GIS) is applied due to its ability in manipulating digital maps from various types of vector maps (points, lines, or polygons) into a vector grid map, as well as intersecting several overlaid thematic maps into a new thematic map. The use of GIS in analyzing multiple factors has been widely used in many fields, including location hazard location (Raghuvanshi, Negassa, & Kala, 2015), forest fire (Setiawan, Mahmud, Mansor, Mohamed Shariff, & Nuruddin, 2004), among others.



*Source: Author Analysis, 2018*

**Figure 1. Map of the Island of Oahu**

A uniform grid map of Oahu Island was developed using a square pattern. The vector map system is chosen rather than a raster-based one since vector based grid can have multiple data in each cell. Each grid had a size of a 0.1-square mile and populated with various thematic maps, such as population distribution, job locations, land use characteristics, road networks, and other environmental considerations (Brunner, Kim, & Yamashita, 2011).

Nonetheless, GIS to some extent have difficulties in manipulating database with too many variables and records. In this sense, the use of SAS software that can handle almost limitless numbers of data and variables can be handy. SAS itself is command-based software where manipulation of variables or records can be done using specific commands and macro instructions. The database from GIS environment, normally in .dbf format can be imported seamlessly into SAS environment. Once they are in SAS system, the database can be manipulated, and the results can be exported back to .dbf format and to be read by the GIS system.

The checkerboard pattern of vector grid map is mimicking the Agent-Based Model (ABM) environment (Colella, Klopfer, & Resnick, 2001). The ABM model has been widely using in simulating urban development (Benenson & Omer, 2001; Kohler, Van West, Carr, & Langton, 1996; Magliocca, McConnell, Walls, & Safirova, 2012). Moreover, ABM is also widely discussed in traffic mobility simulation (Auld & Mohammadian, 2012; Bakillah, Liang, & Zipf, 2012; Bazzan & Klügl, 2014; Chen & Cheng, 2010). In this paper, the SAS system is applied to create an ABM type of simulation, where an "Agent" should be able to move among "Patches" based on several predetermined "Environment." Agent in this model is a transient vector called the GUIDE (Grid Unit with Intelligent Directional Expertise), acting as patches are the grid map itself. While, the environment is the characteristics of grid cell based on its land use, population density, and other variables.

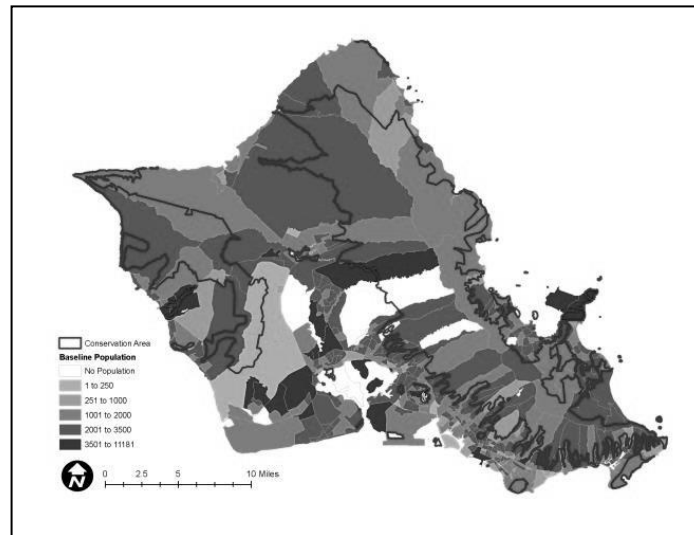
A key problem that needs to be solved with the ABM simulation is to determine the location of grid cells that are superiors compared to their surroundings. Those dominant cells can be linked to each other to create a pathway, which is believed to depict the best alignment for the transit system. This study does not attempt to determine the location of rail stations. Samanta & Jha (2008), however, developed an integrated mathematical and GIS model for optimizing station locations. Li, Li, Fan, & Deng (2017) have been discussing the use of a 1 km x 1 km grid in ArcGIS to support their study in locating the effecting car-sharing location.

## **Methods**

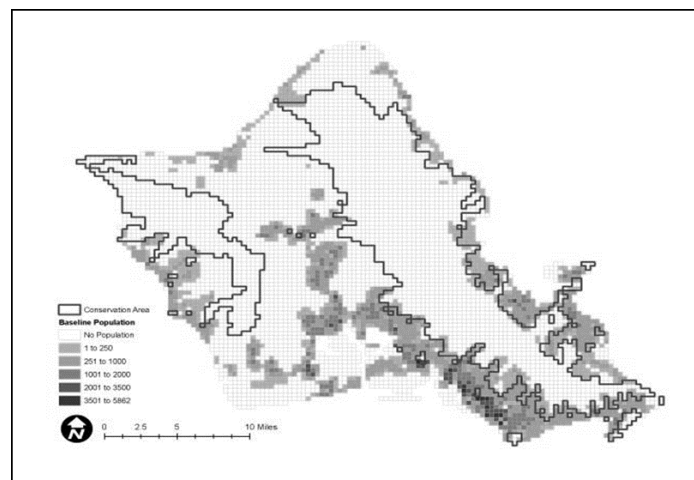
### **Development of Grid Map**

The grid map is a vector map, which means each cell is a polygon and it can have multiple data. Using the ArcGIS software from the Environmental Systems Research Institute (ESRI), a uniform 0.1-square mile (0.316 mile x 0.316 mile) grid map is utilized as the base to convert different thematic maps. Each cell of the grid map is populated with information derived from various thematic maps derived from numerous data sources. The use of 0.1-square mile grid cell will allow any point within the grid is reachable by anyone located within it. This is because each grid cell is within the 0.25 mile radius walking distance. Moreover, a uniform grid map allows population data of a grid cell to be reflected directly as population density and could be compared with other cells to show population density distribution. Figure 2 shows the comparison of population distribution on Oahu by (a) census block group and (b) grid-based map. Some of the census block groups cover conservation areas where limited or no population resides. Nonetheless, the map shows that a block group population is well distributed within its area disregard the type of land use. Meanwhile, the grid map shows that population is located where most of them reside;

conservation area shows an insufficient number of population. The same mapping system has been generated and utilized in several publications (Brunner et al., 2011; Kim, Brunner, & Yamashita, 2006; Kim, Uyeno, Yamashita, & Brunner, 2011). As a comparison, a larger 0.25-square mile grid map system has been used in analyzing bus transit system in Logan, Utah (Ramirez & Seneviratne, 1996), and south Placer County, California (Milam & Luo, 2008).



(a)



(b)

Source: Brunner, 2010

**Figure 2. (a) Baseline Population Distribution on Oahu by Blockgroups and (b) Grid-Based Structure**

The use of relatively small size cell grid is very important for the analysis of transit involving human movements. Studies suggest that distance plays a crucial role in mobility. The distance between the centroid of a grid cell and the centroid of its adjacent cells is between 0.333 to 0.470 miles. This distance is within a comfortable walking distance of  $\frac{1}{4}$  to a  $\frac{1}{2}$  mile as suggested by some literature (Dittmar & Ohland, 2004; NJ Transit, 1994; RPA, 1997). The distance between the centroid of a grid cell and the centroid of its second neighbor grids is between 0.667 to 0.940 miles. This distance is further than the

comfortable walking distance but still within the “walking impact zone” of  $\frac{3}{4}$  miles as revealed in the study of ridership among housing and commercial developments in Canada (TCRP, 1995).

### **Locating Dominant Cells**

Data for determining the dominant cells are derived from the grid map database that has been discussed earlier. A grid cell is considered as a dominant if it meets all of the following criteria, such as has a population size greater than 1,000 populations per grid cell; represents a low-income population group greater than 90 populations per grid cell; has an employment number greater than 500 job counts per grid cell; has potential to be built a Transit Oriented Development (TOD); connected to existing roads; has a higher potential to be a congested area; not an environmentally sensitive area; and shows an existence of key attractor place. Those criteria are a modification of the New Starts Program for transportation planning in the US (Banai, 2006; FTA, 2004). The cells are selected to address health, safety, and environmental concerns as well as convenience related to time and distance between home, work, school, shopping, and other human activity places (Filho, 1996). Effective land-use planning should consider the three basic systems that influence urban spatial structure: human activity, land development, and environmental concern (Berke, Godschalk, Kaiser, & Rodriguez, 2006).

### **The Guide**

After the dominant cells were identified, the locations of possible alignments were determined using a rational and deterministic approach to find the best direction based on predetermined rules. The main components of this model are the grid-based map, and a transient vector called the GUIDE (Grid Unit with Intelligent Directional Expertise) agent. The GUIDE could be seen as an agent in the ABM environment.

The based-grid map consists of squared-shape cells, which are arranged in a checkerboard platform with a unique x-axis and y-ordinate positions. Each grid cell contains several parameters with their unique value.

### **Findings and Discussion**

The GUIDE agent is designed to move in North-East-South direction only and to make a selection based on values of its vicinity grids. Its path indicates a trail of grids with higher values (e.g. population number or job counts) that also suggests the location of transit alignment. The GUIDE has to keep on moving to the next highest value grid until it cannot move any further. The GUIDE is designed to assess not only the values of its adjacent grids but up to two grids away from its origin. In level one analysis, the GUIDE is capable of comparing only the adjacent cells from its current location. In level two, the value of an adjacent cell is influenced by grids next to it. Figure 3 shows an example of a grid map with cell's unique value and GUIDE's origin position at grid F4, as well as illustrates the value of each adjacent grid in different assessment levels.

For level one analysis, results in Figure 3 suggest that the GUIDE would move from its original location at grid F4 to G4. In level two, it would move to grid G3. The assessments suggest that GUIDE would move to two different new positions. Those results are used to generate more alternative for the locations of transit alignments.

	E	F	G	H	I
2		3	1	4	
3		2	3	2	
4			4	1	
5		1	3	1	
6		1	3	2	

Value of Adjacent Grids	Adjusted Value of Adjacent Grid based on Level Two Analysis
[F3] = 2	$[F3_{adj}] = ([F3] + [F2] + [G2])/3 = (2+3+1)/3 = 2$
[G3] = 3	$[G3_{adj}] = ([G3] + [G2] + [H2] + [H3] + [H4])/5 = (3+1+4+2+1)/5 = 2.2$
[G4] = 4	$[G4_{adj}] = ([G4] + [H3] + [H4] + [H5])/4 = (4+2+1+1)/4 = 2$
[G5] = 3	$[G5_{adj}] = ([G5] + [G6] + [H4] + [H5] + [H6])/5 = (3+3+1+1+2)/5 = 2$
[F5] = 1	$[F5_{adj}] = ([F5] + [F6] + [G6])/3 = (1+1+3)/3 = 1.7$

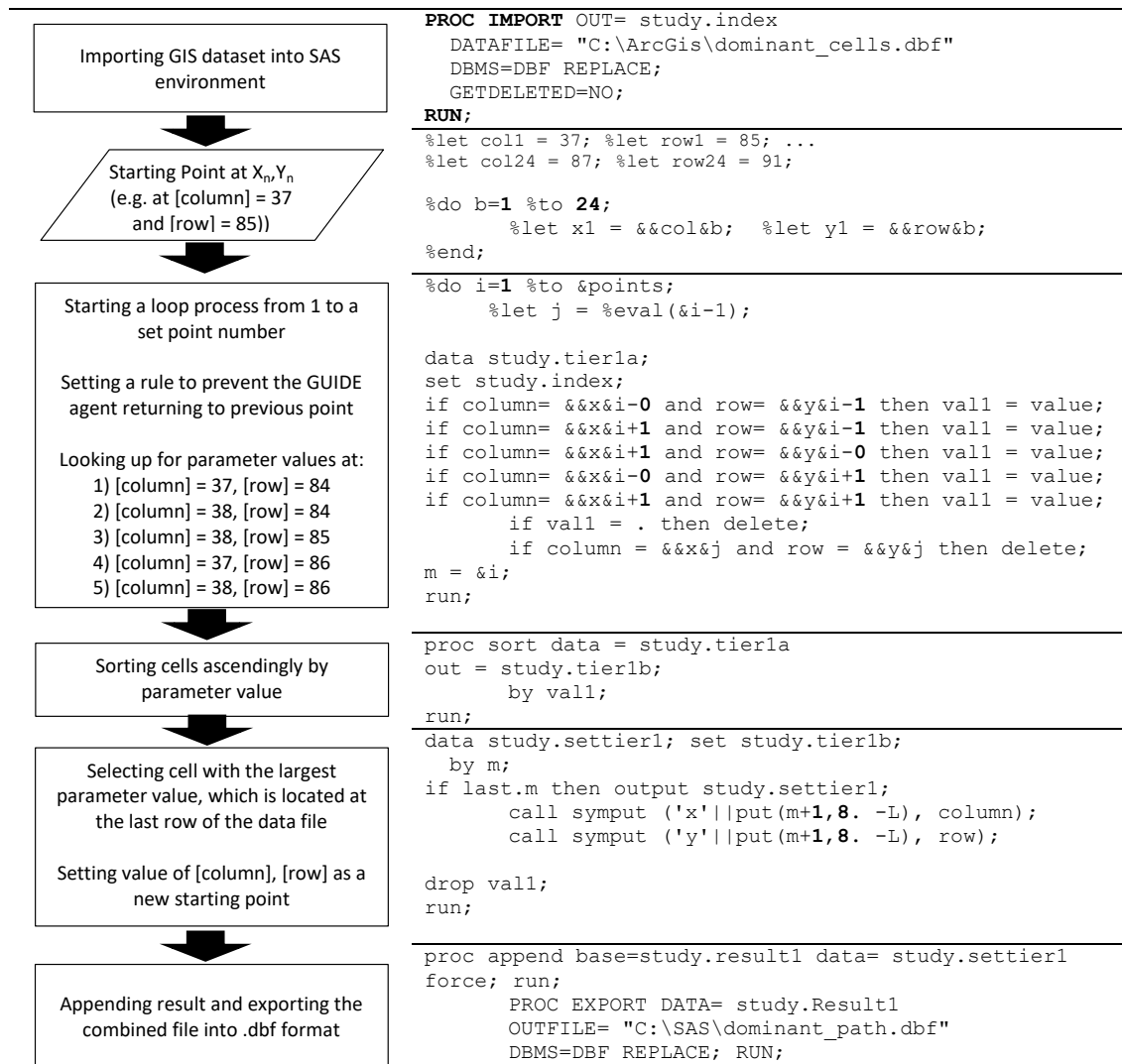
Source: Brunner, 2010

**Figure 3. Grid Values Based on Different Level of Analysis**

In this study, a SAS macro is designed to perform several tasks, which include: a) importing data from ArcGIS into SAS environment where the map database has several information, including [Column] showing X-coordinate of cell in the grid map, [Row] showing Y-coordinate of cell in the grid map, [Value] showing value of parameter (e.g. population, job counts); b) releasing the GUIDE at several starting points; c) analyzing value of set parameter of the vicinity cells; d) selecting the highest value of the selected cells; e) combining cells with highest values, and f) exporting the result into .dbf format that can be linked with ArcGIS system. Figure 4 explains processes taken in the SAS environment, as well as its procedures. The procedures show that the GUIDE has an ability to search for the highest value at one column and or one row away. Allowing the GUIDE to search just the cells at its vicinity could lead to a branch with a dead end. Assuming that there are two possible routes from a node, say one is heading north and other is heading to the east. The North route has a better value at the beginning but the value is shrinking fast and shortly becomes zero. The East route as a less favorable value at the beginning compared to the North branch. However, the East branch has a more steady value and longer possible route. The GUIDE once it released will pick the North route as the selected choice, and soon it will stop since the surrounding cell are empty. This could be the problem with GUIDE that has very short eyesight. To overcome this issue, the capability of GUIDE can be extended to detect the values of cells at two, three, or even more steps beyond its current position. The GUIDE can be designed to move to the next cell with a less favorable value but knowing it will lead to a longer route.

Figure 5 shows the results of releasing GUIDE at a point (letter “A” on the map) in West side of the island. After released and made some movements, the GUIDE was able to select route X for its first choice. The route X is relatively short and lead to nowhere at one point. The GUIDE then had to be released again at a junction between branch X and Y. Here, the GUIDE was able to select route Y which is much longer than X but it also lead to a full stop at a point in the South shore. The GUIDE is once more released at a junction between branch Y and Z. Finally, it can select route Z as a much longer option until it reaches the final destination at the east side of the island. This is a good example of an issue with a capability of the GUIDE to see cells only at its vicinity.

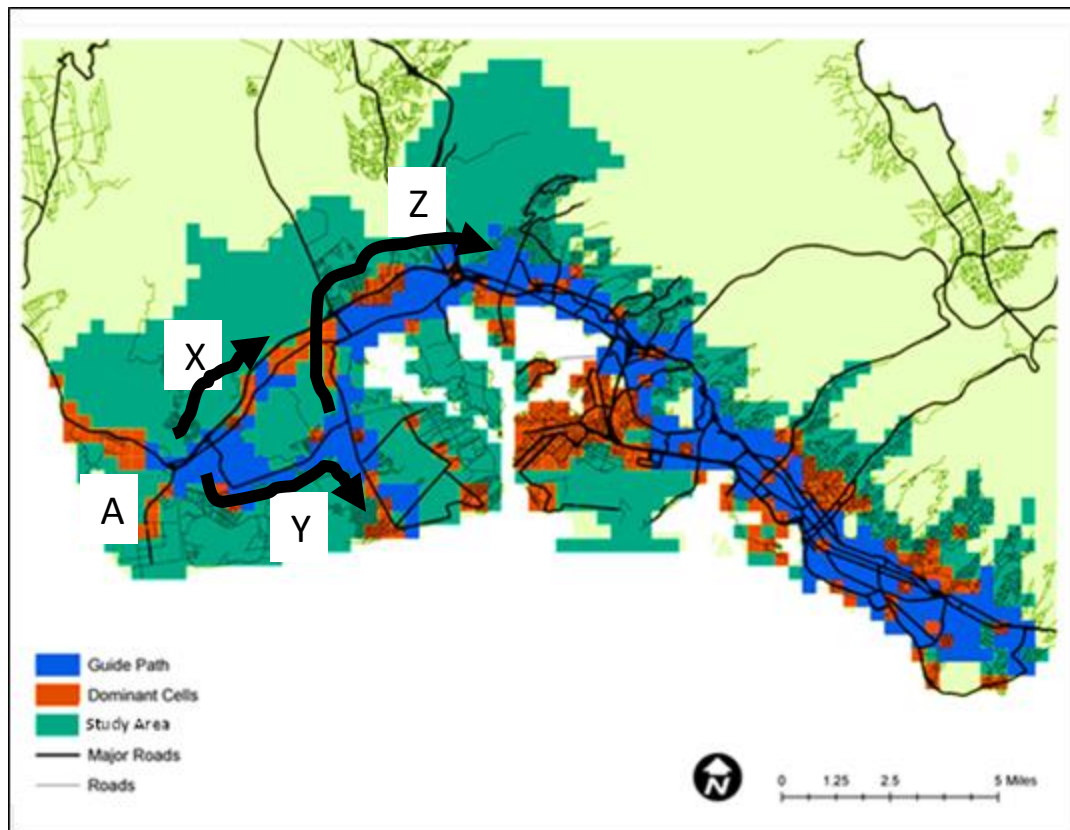




Source: Brunner, 2010

**Figure 4. Flow Chart and SAS Procedures**

This study also shows that it is possible to use SAS in calculating dataset to support spatial analysis. The “symput” and “put” procedures could be used to set the next cell location of the GUIDE. This finding could be useful for those who are familiar with SAS and want to add new procedures to her/his library. The results of SAS procedures are also proven to be able to connect with GIS environment. It increases the capability of analyzing massive dataset and yet the results can be shown on a map for a better understanding. Planners are always challenged to give a better judgment. One way to find out possible alternatives is by using massive datasets, analyzing them, and come up with a simple yet robust result. The result should be understandable by lay people and could lead to a solution. In this case the use of maps, tables, or other figures could be the choice.



*Source: Brunner, 2010*

**Figure 5. Grid Map with Dominant Cells and Guide Pathways**

The use the GUIDE can also be utilized in selecting not only alignment for transit, but also in determining route of water main transmission line, toll road, electricity line, among others that required site selection using spatial analysis.

## Conclusion

It is proven that SAS can be used as an ABM simulator software, although the main purpose of SAS is for statistical and database software. Hence, this paper shows that the GUIDE agent can allocate the next best cell independently as a representative of the optimal transit alignment. The GUIDE is equipped with a capability of determining and moving to the next cell with the highest value, which in this case is some population. The procedures give the possible transit alignments which are not determined by personal judgments.

Nonetheless, this study shows that the capability of the agent in reading the cell data needs to be improved. So far, it can only read one type of data and use it as a factor in determining the next dominant cell. Further, a composite value of several underlying data needs to be established. The composite value of each cell should be used as the factor for an agent to analyze. This would open up another study on how to determine the weight of each type of data before they can be combined into a single value.

This study also demonstrates that SAS and GIS software are complementing each other. SAS is best for handling datasets, while ArcGIS is best for spatial analysis. Although ArcGIS has some capability in manipulating datasets, it has some limitation and somewhat



slow in handling large datasets. There is a SAS Bridge program that allows users to link SAS and ArcGIS. Nonetheless, some compatibility issues deter users from applying for this useful program. On the other hand, Proc Import and Proc Export in SAS can always be applied in the case that SAS Bridge is not available.

The SAS procedure utilized in this study is similar to the principle of Agent-Based Modeling (ABM). The GUIDE is an agent which can move base on several rules; the grid-based map acts as the environment. It has some characteristics for the agent to move and explore. The procedure was design as a deterministic model. Nonetheless, this opens an opportunity for developing a model based on a probabilistic approach. The development of both approaches, however, allows users to model and to study human behavior, such as in a case of emergency evacuation.

## References

- Auld, J., & Mohammadian, A. K. (2012). Activity planning processes in the agent-based dynamic activity planning and travel scheduling (ADAPTS) model. *Transportation Research Part A: Policy and Practice*, 46(8), 1386–1403. doi:10.1016/j.tra.2012.05.017.
- Bakillah, M., Liang, S. H. L., & Zipf, A. (2012). *Toward coupling sensor data and volunteered geographic information (VGI) with agent-based transport simulation in the context of smart cities. Proceedings of the First ACM SIGSPATIAL Workshop on Sensor Web Enablement - SWE '12*. doi:10.1145/2451716.2451719.
- Banai, R. (2006). Public transportation decision-making: A case analysis of the Memphis light rail corridor and route selection with analytic hierarchy process. *Journal of Public Transportation*, 9(2), 1–24. Retrieved from <http://www.nctr.usf.edu/jpt/pdf/JPT9-2Banai.pdf>.
- Bazzan, A. L. C., & Klügl, F. (2014). A review on agent-based technology for traffic and transportation. *The Knowledge Engineering Review*, 29(3), 375–403. doi:10.1017/S0269888913000118.
- Benenson, I., & Omer, I. (2001). *Agent-based modeling of residential distribution. Department of Geography and Human Environment*. Retrieved from [http://www.demogr.mpg.de/papers/workshops/010221\\_paper01.pdf](http://www.demogr.mpg.de/papers/workshops/010221_paper01.pdf).
- Berke, P. R., Godschalk, D. R., Kaiser, E. J., & Rodriguez, D. A. (2006). *Urban land use planning* (Fifth Edit). Baltimore: University of Illinois Press.
- Brunner, I. M. (2010). *The integration of multiple criteria decision making (MCDM) and geographic information system (GIS) for transit planning in Honolulu*. University of Hawaii, Honolulu.
- Brunner, I. M., Kim, K., & Yamashita, E. (2011). *Analytic hierarchy process and geographic information systems to identify optimal transit alignments. Transportation Research Record* (Vol. 2215). doi:10.3141/2215-06.
- Chen, B., & Cheng, H. H. (2010). A review of the applications of agent technology in traffic and transportation systems. *IEEE Transactions on Intelligent Transportation Systems*, 11(2), 485–497. doi:10.1109/TITS.2010.2048313.
- Colella, V. S., Klopfer, E., & Resnick, M. (2001). *Adventures in modeling: Exploring complex, dynamic systems with StarLogo*. New York: Teachers College Press.
- Department of Business Economic Development & Tourism (DBEDT). (2017). *2016 State of Hawaii data book*. Department of Business, Economic Development & Tourism. Retrieved from <http://dbedt.hawaii.gov/economic/databook/db2016/>.
- Dittmar, H., & Ohland, G. (2004). *The new transit town: Best practices in transit-oriented development*. (H. Dittmar & G. Ohland, Eds.). Washington, DC: Island Press.
- Federal Transit Administration (FTA). (2004). *Guidelines and standards for assessing transit-supportive land use*. Washington, DC.
- Filho, A. (1996). *A road location method using noise contours for land use planning around airports*. Carleton University, Ottawa Ontario.
- Kim, K., Brunner, I. M., & Yamashita, E. (2006). Influence of land use, population, employment, and economic activity on accidents. *Transportation Research Record*, 1953(1953), 56–64. doi:10.3141/1953-07.
- Kim, K., Uyeno, R., Yamashita, E., & Brunner, I. M. (2011). Methodological issues in the study of tourism and

### 39 Agent Based Model in SAS Environment for Rail Transit System Alignment Determination

- natural, a geospatial analysis of Oahu Hawaii. In *Introduction of Qualitative Analysis for Tourism Industry* (pp. 251–273). Okinawa, Japan: Nansei Media. (Original in Japanese).
- Kohler, T. ., Van West, C. ., Carr, E. ., & Langton, C. . (1996). *Agent-based modeling of prehistoric settlement systems in the Northern American Southwest*. Retrieved from [https://www.researchgate.net/publication/238176557\\_Agent-based\\_modelling\\_of\\_prehistoric\\_settlement\\_systems\\_in\\_the\\_Northern\\_American\\_Southwest%27%27T\\_hird\\_In](https://www.researchgate.net/publication/238176557_Agent-based_modelling_of_prehistoric_settlement_systems_in_the_Northern_American_Southwest%27%27T_hird_In).
- Li, W., Li, Y., Fan, J., & Deng, H. (2017). Siting of carsharing stations based on spatial multi-criteria evaluation: A case study of Shanghai EVCARD. *Sustainability*, 9(1). doi:10.3390/su9010152.
- Magliocca, N., McConnell, V., Walls, M., & Safirova, E. (2012). *Explaining sprawl with an agent-based model of exurban land and housing markets. Resources for the Future Discussion Paper*. Washington, DC. doi:10.2139/ssrn.2086735.
- Milam, R. T., & Luo, T. (Anna). (2008). Land use-based transit planning. *Transportation Research Record: Journal of the Transportation Research Board*, 143–148. doi:10.3141/2063-17.
- NJ Transit. (1994). *Planning for transit-friendly land use: A handbook for New Jersey communities*. Berkeley: University of California.
- Raghuvanshi, T. K., Negassa, L., & Kala, P. M. (2015). GIS based grid overlay method versus modeling approach—a comparative study for landslide hazard zonation (LHZ) in Meta Robi district of West Sh owa Zone in Ethiopia. *Egyptian Journal of Remote Sensing and Space Science*, 18(2), 235–250. doi:10.1016/j.ejrs.2015.08.001.
- Ramirez, A. I., & Seneviratne, P. N. (1996). Transit route design applications using geographic information systems. *Transportation Research Record: Journal of the Transportation Research Board*, 1557, 10–14. doi:10.3141/1557-02.
- Regional Plan Association (RPA). (1997). *Building transit-friendly communities a design and development strategy for the Tri-State Metropolitan region*. Retrieved from <http://www.rpa.org/pdf/tfc01.pdf>.
- Samanta, S., & Jha, M. K. (2008). Identifying feasible locations for rail transit stations: Two-stage analytical model. *Transportation Research Record: Journal of the Transportation Research Board*, 2063(1), 81–88. doi:10.3141/2063-10.
- Setiawan, I., Mahmud, A. R., Mansor, S., Mohamed Shariff, A. R., & Nuruddin, A. A. (2004). GIS-grid-based and multi-criteria analysis for identifying and mapping peat swamp forest fire hazard in Pahang, Malaysia. *Disaster Prevention and Management: An International Journal*, 13(5), 379–386. doi:10.1108/09653560410568507.
- Transit Cooperative Research Program (TCRP). (1995). *Research results digest: An evaluation of the relationships between transit and urban form*. Retrieved from [http://onlinepubs.trb.org/Onlinepubs/tcrp/tcrp\\_rrd\\_05.pdf](http://onlinepubs.trb.org/Onlinepubs/tcrp/tcrp_rrd_05.pdf).
- United States Census Bureau. (2018). American fact finder Honolulu county, Hawaii. Retrieved from [https://factfinder.census.gov/faces/nav/jsf/pages/community\\_facts.xhtml?src=bkmk#](https://factfinder.census.gov/faces/nav/jsf/pages/community_facts.xhtml?src=bkmk#).