Citation: Yao Z, Wei H, Perugu H (2017) A Heuristic Approach for Quantifying Household Travel GHG Emissions Using GPS Survey and Spatial Correlations-The Cincinnati Case Study. Environ Pollut Climate Change 1: 111.

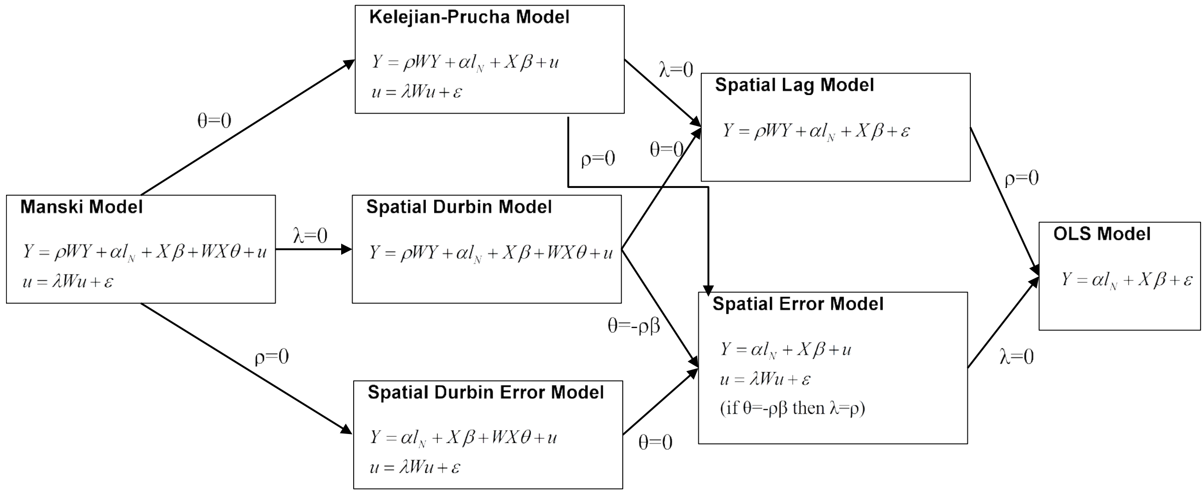

Figure 3: The relationships between spatial dependence models for cross-section data.

Variables	Estimate	Std. Error	t	p-value	Pr(>|t|)
(Intercept)	-1.61E-01	1.08E-01	-1.482	0.138675	
ACRES	7.61E-05	4.48E-05	1.697	0.09016	.
AT	1.87E-01	4.71E-02	3.965	8.10E-05	***
POP	4.57E-04	1.01E-04	4.535	6.80E-06	***
TOTAL_HH	4.20E-04	2.17E-04	1.932	0.053826	.
TOTAL_EMPL	-3.78E-05	2.48E-05	-1.52	0.128934	
POP_DENSIT	-2.13E-02	4.87E-03	-4.369	1.44E-05	***
EMP_DENSIT	4.21E-04	2.84E-04	1.485	0.13811	
TOTAL_AUTO	4.96E-04	1.04E-04	4.772	2.24E-06	***
EMP_M	2.40E-01	7.07E-02	3.39	0.00074	***
AVGWK	1.58E-01	8.50E-02	1.862	0.062971	.
AVGAUTO	-1.94E-01	7.48E-02	-2.594	0.009688	**
Avg_CarbEM	-7.54E+01	2.11E+01	-3.582	0.000365	***
Avg_TRIPSP	1.99E-02	3.56E-03	5.589	3.32E-08	***

(Residual standard error: 0.5001 on 679 degrees of freedom Multiple R-squared: 0.8002, Adjusted R-squared: 0.7964 F-statistic: 209.2 on 13 and 679 DF, p-value: <2.2e-16)

Table 2: OLS regression model and coefficients.

k number for model cross validation is 10. However, since there are 693 TAZs in our dataset, k=9 is used to ensure each "fold" is equal.

Since the data are randomly assigned to a number of 'folds'. Each fold is removed, in turn, while the remaining data is used to refit the regression model and the deleted observations are predicted. Table 3 shows the residual sum of squares and mean square. Figure 5 is the validation plot showing the removed (folded) vs. fitted data. The validation plot shows a good validation since each removed vs. fitted data flows similar 45 degree line. Overall, the OLS model is validated and it is a good fit.

Spatial regression analysis results

The spatial regression models are estimated using the maximum likelihood method. Table 4 shows the variable coefficients using the OLS, SAR, SEM, SDM, SDEM, KPM, and MAM. The coefficients that are not spatially dependent (i.e., Avg_CarbEM, Avg_TRIPSP) are quite similar. And the spatially dependent variables have more variations in the coefficient. This is expected because each of the models has different assumptions and is of different forms as shown in Figure 3.

Goodness of fit measures for candidate models

The goodness of fit measures in spatial regression models is slightly more complex due to the lack of standard measures such as R^2. However, commonly used goodness of fit measures is the information-based measures. The information-based goodness of fit measures utilizes several model performance measures and rank based on the values. The model with the lowest rank is considered a better fit than others. Table 5 shows the information based measures and their ranks for OLS, SAR, SEM, SDM, SDEM, KPM and MAM models. This ranking utilized AIC, Log Likelihood and Moran's I on Residuals as measures. For all three criteria, smaller values are better. Therefore, the SDEM model has the lowest summation of ranks and it fits the data better.

Discussion

A spatial regression-based modeling framework was developed based on finding the minimal model residuals and multiple information-based measures of fit. The goodness of fit measures in spatial regression models is slightly more complex due to the lack of standard measures such as the R^2. However, a common goodness of fit measures is the

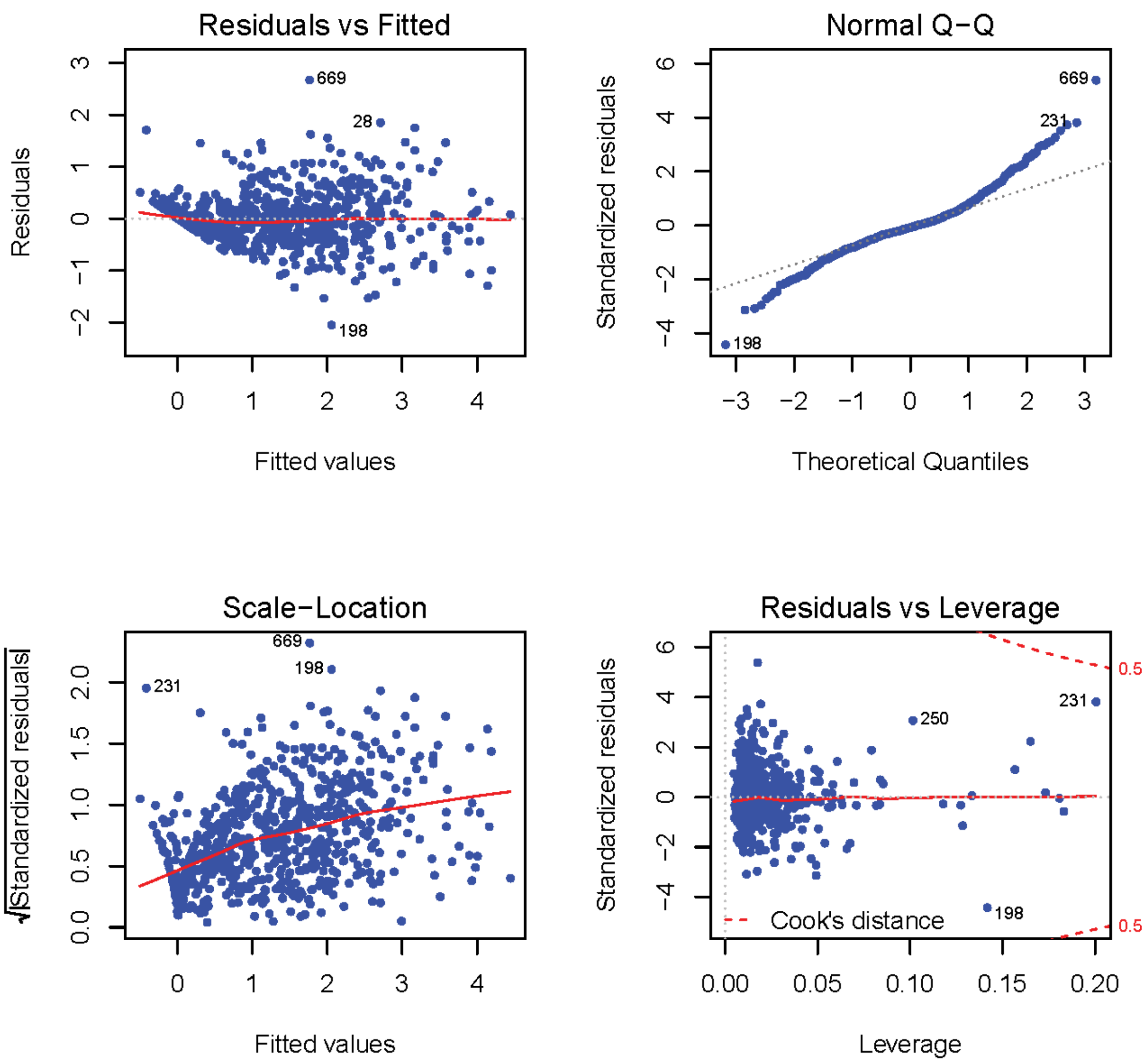

Figure 4: Diagnose plot for OLS regression model.

information-based measures. The information-based goodness of fit measures utilizes several model performance measures and ranks based on the values. The model with the lowest rank is considered a better fit than others. The information-based measures and their ranks for OLS, SAR, SEM, SDM, SDEM, KPM and MAM models are summarized and presented.

OLS model has an R^2 (coefficient of determination) of 0.8, which is a good fit. However, when examining the residuals on diagnosis plots, it was found that the residuals are still spatially correlated. This suggests that spatial models can fit the data better and reduce the residual spatial correlation. After performing spatial regressions, the information-based measure of fit based on AIC, log likelihood and Moran's I on residuals are compared and the best model fitting the given dataset is the Spatial Durbin Error Model. The SDEM has the lowest AIC and Moran's I on residuals compared to other candidate models.

This study has provided a proof of concept for the proposed methodology and solid foundation for the modeling land use changes, and GHG emission analysis. It has been proven that the proposed method has the capability to reveal the dynamic linkage between land use, transportation, and emissions. The findings from this research provide insights on how land-uses planning alternatives built on adopted policies and enforced development regulations correlate with travel patterns and their sequential GHG emissions. The level of specificity, such as the land use change and GHG emission analysis presented in this study enables more data and indicators to be developed. Such data and indicators can be incorporated into decision makers' plans, policies and ultimately regulations and its possible integration with project level review processes.

Conclusion

While the results from this study offer specific recommendations as to which types of land use planning policy practices are most highly

Fold	Residual Sum of Squares	Residual Mean Square
1	21.20	0.28
2	17.30	0.22
3	21.50	0.28
4	11.50	0.15
5	11.00	0.14
6	24.80	0.32
7	26.60	0.34
8	22.70	0.30
9	23.20	0.30
Average	**19.98**	**0.26**

Table 3: The 9-fold cross validation residuals.

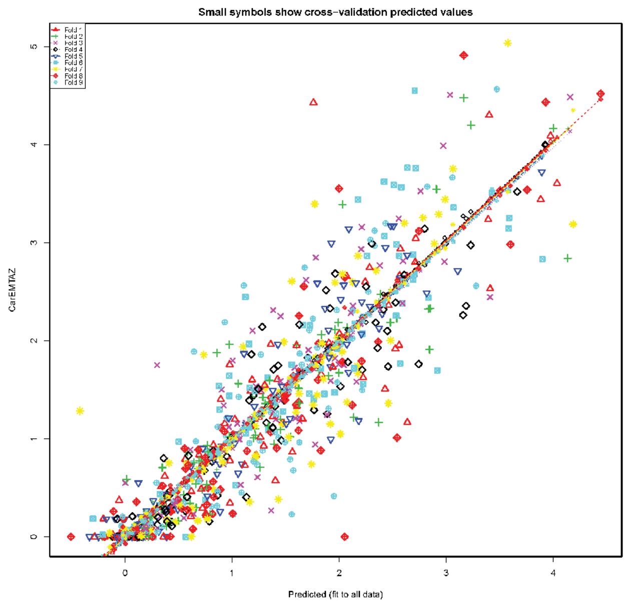

Figure 5: The 9-fold cross-validation results.

Coefficients	OLS	SAR	SEM	SDM	SDEM	KPM	MAM
(Intercept)	-1.61E-01	-1.38E-01	-1.61E-01	-6.37E-02	-1.61E-01	-1.42E-01	-3.17E-02
ACRES	7.61E-05	5.66E-05	7.68E-05	2.17E-04	7.68E-05	7.37E-05	2.41E-04
AT	1.87E-01	1.93E-01	1.87E-01	3.00E-01	1.87E-01	1.88E-01	3.11E-01
POP	4.57E-04	4.59E-04	4.57E-04	5.50E-04	4.57E-04	4.75E-04	5.77E-04
TOTAL_HH	4.20E-04	4.68E-04	4.20E-04	6.05E-04	4.20E-04	4.65E-04	6.33E-04
TOTAL_EMPL	-3.78E-05	-4.22E-05	-3.78E-05	-5.08E-05	-3.78E-05	-4.19E-05	-4.96E-05
POP_DENSIT	-2.13E-02	-2.18E-02	-2.13E-02	-2.55E-02	-2.13E-02	-2.21E-02	-2.68E-02
EMP_DENSIT	4.21E-04	3.78E-04	4.21E-04	3.75E-04	4.21E-04	3.72E-04	3.22E-04
TOTAL_AUTO	4.96E-04	4.78E-04	4.95E-04	2.56E-04	4.95E-04	4.56E-04	2.08E-04
EMP_M	2.40E-01	2.36E-01	2.40E-01	1.87E-01	2.40E-01	2.35E-01	1.83E-01
AVGWK	1.58E-01	1.70E-01	1.59E-01	1.91E-01	1.59E-01	1.74E-01	1.84E-01
AVGAUTO	-1.94E-01	-1.62E-01	-1.94E-01	-8.03E-02	-1.94E-01	-1.65E-01	-6.22E-02
Avg_CarbEM	-7.54E+01	-7.54E+01	-7.54E+01	-7.58E+01	-7.54E+01	-7.47E+01	-7.46E+01
Avg_TRIPSP	1.99E-02	2.02E-02	1.99E-02	1.98E-02	1.99E-02	1.99E-02	1.96E-0

(OLS: Ordinary Least Square; SAR: Spatial Autoregressive Model; SEM: Spatial Error Model; SDM: Spatial Durbin Model; SDEM: Spatial Durbin Error Model; KPM: Kelejian Prucha Model; MAM: Manski Model)

Table 4: Model coefficients comparison for OLS, SAR, SEM and SDM models.

Model Type	AIC	Rank	Log Likelihood	Rank	Moran's I on Residuals	Rank	Total Rank
Ordinary Least Square **(OLS)**	1092.5	4	-533.2	5	+0.026397503	5	14
Spatial Autoregressive Model **(SAR)**	1088.4	3	-530.2	3	+0.056595454	6	12
Spatial Error Model **(SEM)**	1093.1	5	-532.6	4	-0.000634720	1	10
Spatial Durbin Model **(SDM)**	1065.3	2	-507.7	2	-0.008763889	4	8
Spatial Durbin Error Model **(SDEM)**	1064.1	**1**	-532.6	**4**	-0.000634720	**1**	**6**
Kelejian-Prucha Model **(KPM)**	N/A	3.5	N/A	3.5	-0.000340069	2	9
Manski Model **(MAM)**	1093.8	5	-506.4	1	-0.0052871289	3	9

Table 5: Information-based measure of fit for spatial models.

associated with a higher amount of VMT, GHG emissions, there are also some potential to reveal policy impacts that can be applied to integrated land use and transportation sustainability practices. The results of this research are expected to add to the existing body of knowledge to enable faster and easier methods of examining the impact of adaptive planning strategies on alleviating the effects of household travel GHG emissions. The spatial cross-sectional regression model is developed through the integration of actual and scenario based land use visioning and planning, demographical changes, transportation emission analysis, and computer forecasting and evaluation of future scenarios. This research makes it possible to assess the household travel GHG footprint and provides models, data for possible GHG emission mitigation through land use policies and changes. Although the results may be pertaining to the specific dataset but it helps transportation decision makers to better connect the land use development and its related household socioeconomics with their GHG emission characteristics. Particular, the household travel GHG emission quantification results made its contribution to the current body of knowledge on the following: (1) provides accurate GHG emission results by using the best available traffic activity data inputs (VSP distributions) for emission modeling; (2) provides connections between household socioeconomics and their travel GHG footprint. The research suggests important potential to provide solid grounds for analyzing, modeling of sustainable community strategies, adaptive planning policies, and many other policy-making applications.

Acknowledgement

The authors are appreciative of the support by U.S. EPA and Ohio-Kentucky-Indiana Regional Council of Governments. Any opinions expressed in this paper are the author(s)'s point of view and do not necessarily reflect the views of the Agency.

References

1. USEPA. Fast Facts: U.S. Transportation sector greenhouse gas emissions 1900-2010. Facts EPA 420-F12-063, United States Environmental Protection Agency, Washington, DC.
2. USEPA (2012) Using MOVES to prepare emission inventories in state implementation plans and transportation conformity. Washington DC.
3. USEPA (2010) Transportation conformity guidance for quantitative hot-spot analyses in PM2.5 and PM10 nonattainment and maintenance areas. Technical guidance EPA-420-B-10-040, USEPA, Washington, DC.
4. FHWA (2010) Transportation conformity: A basic guide for state and local officials. Washington DC.
5. Guohua S, Lei Y, Ziqianli W (2008) A practical modeling approach for evaluation of fuel efficiency for road traffic. Transportation Research Board Annual Meeting.
6. Frey H, Nagui R, Haibo Z (2008) Link-based emission factors for heavy-duty diesel trucks based on real-world data. TRR: Journal of the Transportation Research Board 2058: 23-32.
7. Fulper`C, Hart C, Warila J, Koupal J, Kishan S, et al. (2011) Development of real-world data for MOVES development: The Houston Drayage activity characterization study. Transportation Research Annual Meeting.
8. Yao Z, Wei H, Perugu H, Liu H, Li Z (2014) Sensitivity analysis of project level MOVES running emission rates for light and heavy duty vehicles. JTTE 1: 81-96.
9. Harikishan P, Heng W, Zhuo Y (2016) Integrated data-driven modeling to estimate PM 2.5 pollution from heavy-duty truck transportation activity over metropolitan area. Transportation Research Part D: Transport and Environment 46: 114-127.
10. Zhuo Y, Heng W, Hao L, Zhixia L (2013) Statistical vehicle specific power profiling for urban freeways. Elsevier.
11. Robinsson R, Edgar V, Paula R, Eduardo B (2016). Influence of driving patterns on vehicle emissions: A case study for Latin American cities. Transportation Research Part D: Transport and Environment 43: 192-206.
12. William B, Wayne S, Shauna H, Randall G (2000) Modeling regional mobile source emissions in a geographic information system framework. Transportation Research Part C: Emerging Technologies 8: 205-229.
13. Bhat C, Conoor S, Nerella S (2003) GIS-integrated traffic input models for mobile-source emissions analysis. Texas Department of Transportation, Austin.
14. Mittal N, Golias M, Boile M, Spasovic L, Ozbay K (2004) Estimating truck volumes on state highways – A statistical approach. Transportation Research Board Annual Meeting.
15. Wang H, Jang K, Chan CY (2010) Analysis and Multi-level modeling of truck freight Demand. Strategic Highway Research Program 2 (SHRP2) Freight Modeling and Data Symposium.
16. Elhorst JP (2010) Spatial panel model. Handbook of Applied Spatial Analysis. Springer Berlin Heidelberg, Berlin.
17. Frazier C, Kockelman KM (2006) Spatial Econometric models for panel data: Incorporating spatial and temporal data. TRR: Journal of the Transportation Research Board 1902: 80-90.
18. Wang Y, Kockelman K, Wang X (2012) Anticipation of land use change through use of geographically weighted regression models for discrete response. TRR: Journal of the Transportation Research Board 111-123.
19. Wang Y, Kockelman K, Damien P (2013) A spatial autoregressive multinomial probit model for anticipating land-use change in Austin, Texas. Ann Reg Sci.
20. Zhou B, Kockelman K (2008) Neighborhood impacts on land use change: a multinomial logit model of spatial relationships. Ann Reg Sci 42: 321-340.
21. Zhou B, Kockelman K (2009) Predicting the distribution of households and employment: a seemingly unrelated regression model with two spatial processes. J Trans Geogr 17: 369–376.
22. Parent O, LeSage J (2010) A spatial dynamic panel model with random effects applied to commuting times. Transportation Research Part B: Methodological 44: 633-645.
23. Perugu H, Wei H, Rohne A (2012) Modeling roadway link PM2.5 emissions with