

## INTRODUCTION

### PARTICLE NUMBER CONCENTRATION (PNC) MEASUREMENTS<sup>1</sup>

- PNC: Measure of ultrafine particles (UFP; particles smaller than 100 nm)
- Obtained using data from mobile monitoring and fixed sites across the United States and a land-use regression (LUR) modeling framework
- Predict PNC at ~6 million residential census blocks in the contiguous United States
- Intraurban variation captured through highly spatially resolved mobile measurements in three US cities
- Interurban trends captured through longer-term fixed-site data in multiple urban locations across the United States
- Urban-rural gradients captured through longer-term data from rural locations

### US CENTERS FOR DISEASE CONTROL AND PREVENTION TYPE 2 DIABETES MELLITUS (T2DM) INCIDENCE DATA<sup>2</sup>

- United States T2DM incidence data obtained through county-level measurements from the Centers for Disease Control and Prevention
- T2DM incidence data correlated with ultrafine particle (UFP) concentration measurements from the prior year

## METHODS

- UFP estimates (available through county-level, yearly gridded PNC datafiles, calculated at a 1 km x 1 km resolution) (Figure 1a) and T2DM incidence (available through county-level, yearly, incidence number and percentages datafiles) (Figure 1b) were matched and correlated through a multiple linear regression analysis
- Controls included in analysis: age (18-44, 45-64, 65-74, 75+), gender (male, female), ethnicity and race (Hispanic, non-Hispanic White, non-Hispanic Black, non-Hispanic Asian), and education (less than high school, high school, and more than high school)

## RESULTS/DISCUSSION

### MULTIPLE LINEAR REGRESSION MODEL

- Pearson correlation coefficient analysis (using  $\frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}}$ ) between variables indicated presence of multi-collinearity (Figure 2)
- T2DM and UFP dataset features were scaled and translated using Python Scikit-Learn MinMax scaling method
- Variables with significant p-values and variance inflation factors were dropped
- Multiple linear regression model summary illustrated in Table 1

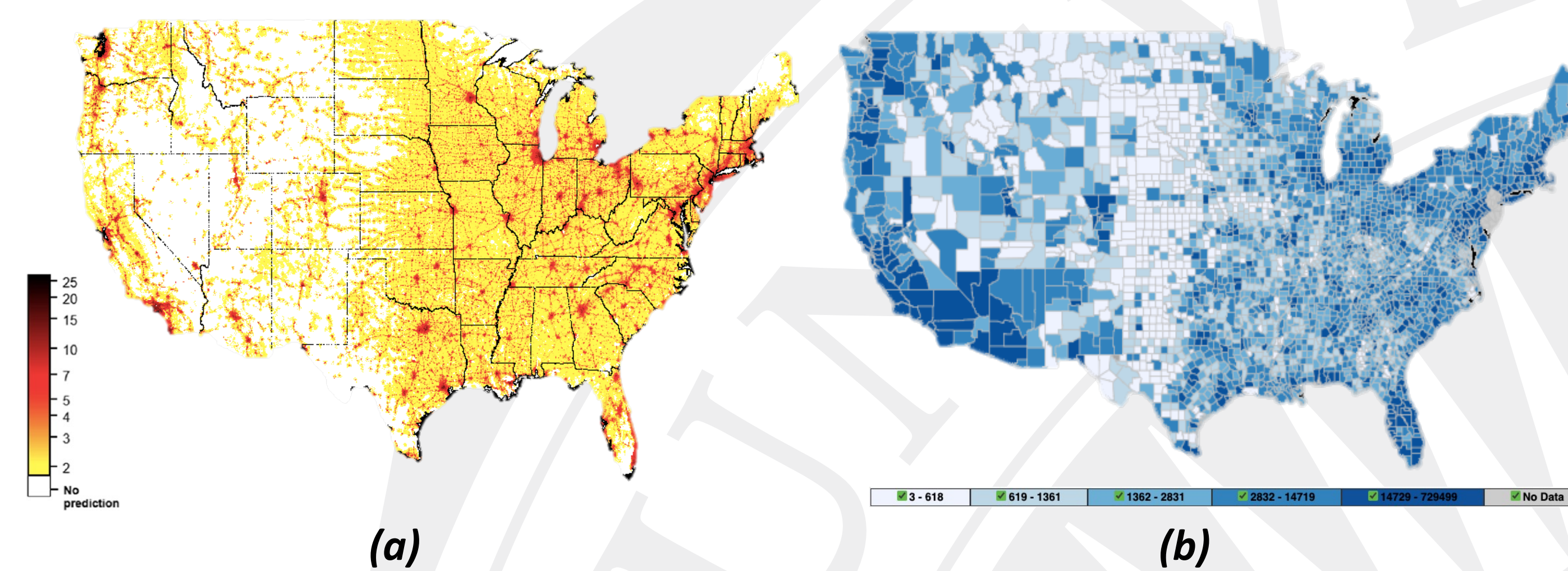


Figure 1. (a) LUR PNC (cm<sup>-3</sup>) United States coverage<sup>1</sup> (b) CDC United States T2DM Incidence<sup>2</sup>

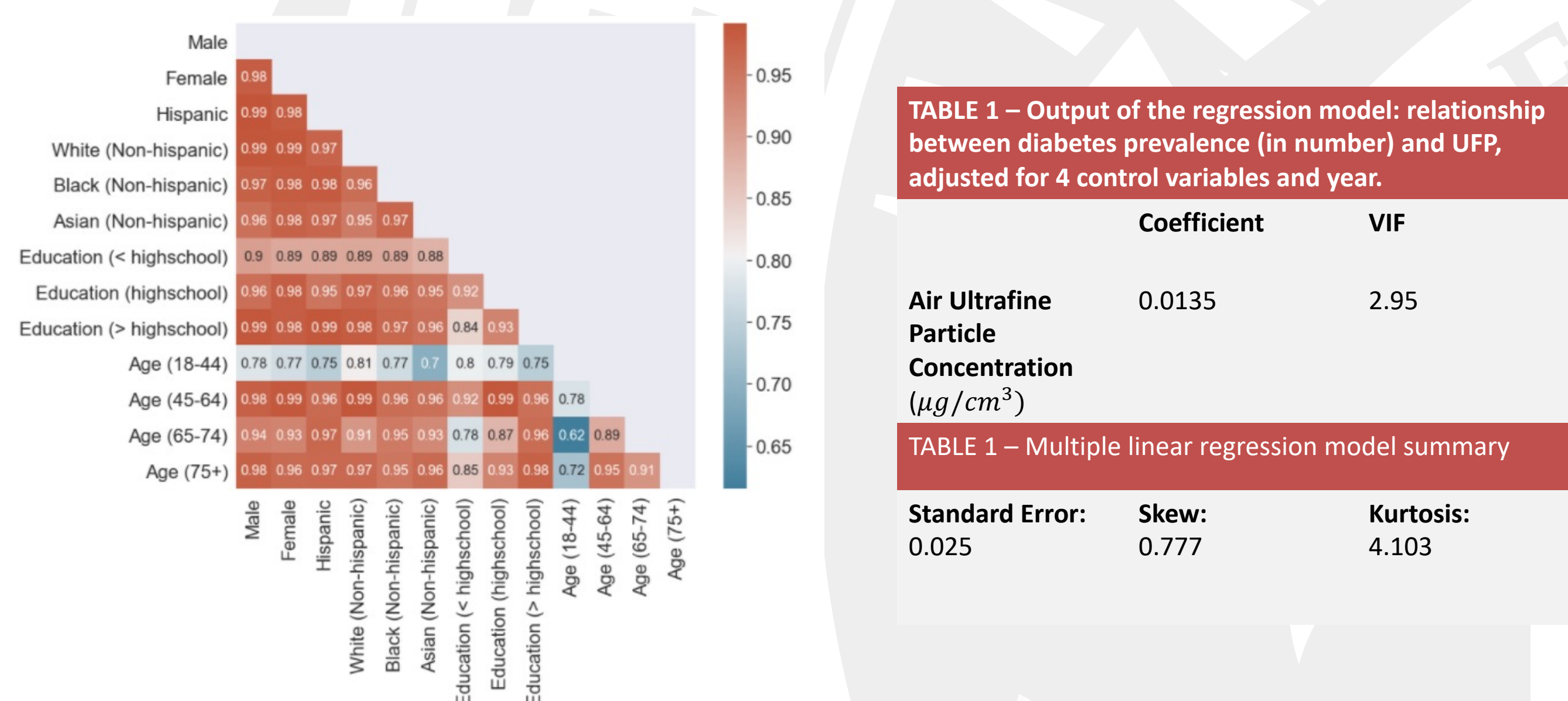


Figure 2. Pearson correlation coefficient control variable analysis results Table 1. Adjusted regression model results summary

TABLE 1 – Output of the regression model: relationship between diabetes prevalence (in number) and UFP, adjusted for 4 control variables and year.

|  | Coefficient | VIF  |
|--|-------------|------|
| Air Ultrafine Particle Concentration (µg/cm <sup>3</sup> ) | 0.0135      | 2.95 |
| Standard Error:  | 0.025       |      |
| Skew:  | 0.777       |      |
| Kurtosis:  | 4.103       |      |

## CONCLUSIONS

- The UFP T2DM linear regression model results are consistent with previous studies, where positive linear relationships between PM<sub>2.5</sub> and T2DM were revealed<sup>3</sup>
- Although UFP estimates were spatially widespread, data has limited temporal availability, thus individual-level diabetes risk factors could not be assessed
- Future work includes implementation of additional control variables (e.g. physical activity, income)
- Future analysis will compare T2DM incidence with previous year to better understand UFP exposure related incidence

## REFERENCES

- [1] High-Spatial-Resolution Estimates of Ultrafine Particle Concentrations across the Continental United States  
 Provat K. Saha, Steve Hankey, Julian D. Marshall, Allen L. Robinson, and Albert A. Presto  
*Environmental Science & Technology* 2021 55 (15), 10320-10331  
 DOI: 10.1021/acs.est.1c03237
- [2] US Centers for Disease Control and Prevention: Diabetes Data and Statistics  
 Accessed September 12, 2021.  
<https://www.cdc.gov/diabetes/data/index.html>
- [3] The relationship between air pollution and diabetes: A study on the municipalities of the Metropolitan City of Milan  
 Greta Meroni, Alessandra Valerio, Marika Vezzoli, Edoardo Croci, and Michele O. Carruba. *Diabetes Res Clin Pract.* 2021 Apr;174:108748.  
 DOI: 10.1016/j.diabres.2021.108748.

## ACKNOWLEDGEMENTS

This work was supported by the National Library of Medicine (T15LM5920634), the Undergraduate Research Opportunities Program, the National COVID Cohort Collaborative (N3C) Supplement (3UL1TR002538-03S3), National Institute of Diabetes and Digestive and Kidney Diseases Ruth L. Kirschstein National Research Service Award of the National Institutes of Health under award number [5T32DK110966-04] and the National Center for Advancing Translational Sciences (UL1TR002538). The Center of Excellence for Exposure Health Informatics was established from efforts of the Utah PRISMS Informatics Center (NIH/NIBIB U54EB021973). Computational resources were provided by the University of Utah Center for High Performance Computing, partially funded by the NIH Shared Instrumentation Grant (1S10OD021644-01A1).