

Khanh Do

8 Douglass Park, Boston, MA, 02118 • 714-837-2430 • kh.do@northeastern.edu

EDUCATION

University of California, Riverside

Riverside, CA

Ph.D. Chemical and Environmental Engineering

2023

Dissertation title: “[Computational and Geospatial Approaches to Investigate Multi-scale Air Quality Trends in Southern California](#)”

Advisor: Professor Cesunica E. Ivey

M.S. Electrical Engineering

2023

Specialization: Robotics and Computer Vision

University of California, Berkeley

Berkeley, CA

B.S. Chemical Engineering

2015

EMPLOYMENT

Northeastern University, Boston

Boston, MA

Postdoctoral Research

2023 - Present

RESEARCH INTERESTS

My research focuses on predicting and investigating air quality using machine learning methods and deterministic models. I am interested in working with large datasets to recognize the patterns, explore relationships between variables, and apply machine learning and other technological advancements, such as GPU computing to enhance the computational efficiency of deterministic models.

Keywords: Air quality and modeling; machine learning; computer vision; personal exposure; data assimilation; high-performance computing; GPU computing; air quality; climate; model development; downscaling; bias correction;

PROFESSIONAL DEVELOPMENT

- Providing academic mentorship to six Ph.D. students, overseeing their progress, and offering guidance on research projects
- Assisting Principal Investigators in developing and preparing proposals for NSF, NOAA, and EPA funding opportunities
- Delivering a specialized training course, “Training Course on Seamless Prediction of Air Pollution in Africa” at the Egyptian Meteorological Authority (EMA), Egypt. Organized by the World Meteorological Organization, this course focused on instructing participants on running and evaluating the WRF-Chem model for the African domain

RESEARCH EXPERIENCE

Shared Socioeconomic Pathways high-resolution dynamically downscaling

- A study of 30 years of climate change, from 2060 to 2089, using Pseudo Global Warming method to obtain projections of future climate data based on Shared Socio-Economic Pathways (SSP) and WRF to dynamically downscaling. This project provides high resolution future meteorological data for the entire USA (36 km resolution), the Eastern US (6 km resolution), and North Carolina and Virginia (1 km resolution), which are further used for health assessment.

Chemical Transport Model Enhancement

- Utilize machine learning to accelerate chemical transport model by substituting its most time-consuming modules and enhance the accuracy of chemical transport models through machine learning bias correction (funded by iSUPER)
- Involve the development of the early version of UFS-SRW APP from NOAA (funded by NOAA)
- Implement GPU computing into the CMAQ gas solver and utilize a large number of CUDA cores to improve CMAQ computational efficiency (funded by NSF CDS&E)

Environmental Justice and Environmental Health Assessment

- Study the highly impacted areas and explore the correlation between personal exposure and household income in Boston neighborhoods and develop a new cumulative environmental justice score system (funded by iSUPER)
- Investigate personal exposure to PM_{2.5} in the Inland Empire to quantify the air pollutants and explain how personal habits, ethics, incomes, and communities of color drive their exposure to personal resolution (funded by the Sloan Foundation)

- Quantify the impact of personal mobility and housing characteristics on daily PM_{2.5} exposures and well-being for West San Bernardino community members and develop individualized resilience plans for community collaborators to support future PM_{2.5} exposure reduction (funded by CARB)

Machine Learning Assisted Air Quality Predictions

- Enhance low-cost sensor uncertainties by optimizing the installed locations and minimizing the total number of sensors required to cover the most land area, thereby maximizing machine learning prediction accuracy (funded by iSUPER)
- Investigate the impacts of meteorology on ground-level ozone and PM_{2.5} and determine a set of meteorological contributing factors to ozone formation by analyzing historical air quality and meteorology data and constructing an empirical model from historical data to project air quality trends (funded by SCAQMD)
- Utilize pseudo global warming method combined with WRF to downscale the data from the global climate model with various SSP scenarios
- Assemble and operate BAM 1020 and perform collocation calibration for low-cost air quality sensors

ADDITIONAL RELEVANT EXPERIENCE

1. GPU assisted image processing for high-resolution traffic footage
2. Modeling and evaluating the expected waiting time of M/M/1 and M/G/1 data center
3. Vehicular detection, counting, and classification
4. Vehicular emission detection using image processing technique
5. Modeling atmospheric dispersion to quantify the ozone concentrations in Riverside, CA
6. Experience with high performance computing platforms: Chyenne, Derecho, NERSC, Anvil, Discovery

AWARDS AND HONORS

- GAANN Fellowship (2019 - 2022)
- Esther F. Hays Graduate Fellowship Award (2020 - 2021)

- Salim Khan Graduate Award (2019 - 2020)
- Outstanding Teaching Assistant Award (2023), recognized for excellence in teaching in Chemical and Environmental Engineering Department

PUBLICATIONS

1. **Do K**, Yang Zhang, Siqi Ma, and Daniel Tong. “Assessing Air Pollution Exposure Disparities in Disadvantaged Communities of Greater Boston: A New Cumulative Environmental Justice Score System.” *Environmental Research Letters*. <https://doi.org/10.1088/1748-9326/adb16c>
2. Duncan Quevedo, **Do K**, Jose Rodriguez Borbon, Bryan M. Wong, and Ivey CE (2025). “GPU Implementation of a Gas-Phase Chemistry Solver in the CMAQ Chemical Transport Model.”, *Environmental Science & Technology Air*. <https://pubs.acs.org/doi/10.1021/acsestair.4c00181>
3. **Do K** and Yang Zhang. “Assessing Health Risks and Disparities in Air Pollution and Socioeconomic Status across the United States.” *Environmental Pollution under revisions*
4. **Do K**, George Delic, Jose Rodriguez Borbon, Duncan Quevedo, Bryan M. Wong, and Ivey CE (2024). “Graphics Processing Unit Assisted Computation for a Gas-Phase Chemical Solver in a Regulatory Chemical Transport Model.” <https://eartharxiv.org/repository/view/6323/>
5. Torres I, **Do K**, Delgado A, Mourad C, and Ivey CE (2024). “Indoor and Ambient Influences on PM_{2.5} Exposure and Well-being for a Rail Impacted Community and Implications for Personal Protections.” *Environmental Research Letters*. <https://doi.org/10.1088/1748-9326/ad90f5>
6. **Do K**, Daniel Schuch, and Yang Zhang (2024). “Practice Instruction for Online-Coupled WRF-Chem Version 4.6.0. Training Manual for the Hands-on Training Session on WRF-Chem.” <https://etrp.wmo.int/mod/resource/view.php?id=27085>
7. Zihan Zhu, **Do K**, Cesunica E. Ivey, Don R. Collins (2024). “Assessing CMAQ Model Discrepancies in a Heavily Polluted Air Basin Using UAV Vertical Profiles and Sensitivity Analyses.” *Environmental Science: Atmospheres*. <https://doi.org/10.1039/d4ea00004h>
8. **Do K**, Yeganeh AK, Gao Z, and Ivey CE (2024). “Performance of Machine Learning for Ozone Modeling in Southern California during the COVID-19 Shutdown.” *Environmental Science: Atmospheres*. <https://doi.org/10.1039/D3EA00159H>

9. Ziqi Gao, **Do K**, Zongrun Li, Xiangyu Jiang, Kamal J. Maji, Cesunica E. Ivey, Armistead G. Russell (2024). "Predicting PM_{2.5} Levels and Exceedance Days Using Machine Learning Methods." *Atmospheric Environment*.
<https://doi.org/10.1016/j.atmosenv.2024.120396>
10. Duncan Quevedo, Ziqi Gao, **Do K**, Roya Bahreini, Don Collins, Cesunica Ivey (2024). "Multidecadal Analysis of Meteorological and Emissions Regimes for PM_{2.5} Across California." *Environmental Science & Technology Air*.
<https://pubs.acs.org/doi/10.1021/acsestair.3c00019>
11. Ziqi Gao, Cesunica E. Ivey, Charles L. Blanchard, **Do K**, Sang-Mi Lee, Armistead G. Russell (2023). "Emissions and meteorological impacts on PM_{2.5} species concentrations in Southern California using Generalized Additive Modeling." *Science of the Total Environment*. <http://dx.doi.org/10.2139/ssrn.4377297>
12. Ziqi Gao, Cesunica E. Ivey, Charles L. Blanchard, **Do K**, Sang-Mi Lee, Armistead G. Russell (2023). "Emissions, meteorological and climate impacts on PM_{2.5} levels in Southern California using a generalized additive model: Historic trends and future estimates." *Chemosphere*. <https://doi.org/10.1016/j.chemosphere.2023.138385>
13. **Do K**, Mahish M, Kashfi Yeganeh A, Gao Z, Blanchard CL, and Ivey CE (2023). "A Machine Learning Approach to Quantify the Impact of Meteorology on Tropospheric Ozone in the Inland Empire, CA." *Environmental Science: Atmospheres*.
<https://doi.org/10.1039/D2EA00077F>
14. Md Hasibul Hasan, Haofei Yu, Cesunica Ivey, Ajay Pillarisetti, Ziyang Yuan, **Do K**, and Yi Li (2023). Unexpected Performance Improvements of Nitrogen Dioxide and Ozone Sensors by Including Carbon Monoxide Sensor Signal.
<https://doi.org/10.1021/acsomega.2c07734>
15. Gao Z, Ivey CE, Blanchard CL, **Do K**, Lee S-M, and Russell AG (2022). "Separating emissions and meteorological impacts on peak ozone concentrations in Southern California using generalized additive modeling." *Environmental Pollution*, 119503.
<https://doi.org/10.1016/j.envpol.2022.119503>
16. Ziqi Gao, Yifeng Wang, Petros Vasilakos, Cesunica E. Ivey, **Do K**, and Armistead G. Russell (2022) "Predicting peak daily maximum 8h ozone and linkages to emissions and meteorology in Southern California using machine learning methods." *Geoscientific Model Development*, [10.5194/gmd-15-9015-2022](https://doi.org/10.5194/gmd-15-9015-2022)
17. **Do K**, Yu H, Velasquez J, Grell-Brisk M, Smith H, and Ivey CE (2021). A data-driven approach for characterizing community-scale air pollution exposure disparities in inland

Southern California. *Journal of Aerosol Science*, Article 105704.
<https://doi.org/10.1016/j.jaerosci.2020.105704>

18. Ivey CE, Gao Z, **Do K**, Kashfi Yeganeh A, Russell AG, Blanchard CL, and Lee S-M (2020). Impacts of the 2020 COVID-19 Shutdown Measures on Ozone Production in the Los Angeles Basin. *ChemRxiv. Preprint*. <https://doi.org/10.26434/chemrxiv.12805367.v1>

TECHNICAL REPORTS

1. **Do K** and Yang Zhang, “NOAA Semi-annual Report for UFS-SRW APP”, Grant No. NA22OAR4590515
2. Ivey CE, Nguyen A, Xu RM, **Do K**, Hao P, and Barth M. “Hyperlocal Monitoring of Traffic-Related Air Pollution to Assess Near-Term Impacts of Sustainable Transportation Interventions, final report.” National Center for Sustainable Transportation, Contract No. DOT 69A3551747114, Feb 2023.
3. Ivey CE, Blanchard CL, Russell AG, Gao Z, and **Do K**. “Ozone Meteorology Study, final report.” South Coast Air Quality Management District, Contract No. 20058, January 2022.

CONFERENCE ACTIVITIES

Presenter:

1. “Assessing Health Risks and Disparities in Air Pollution and Socioeconomic Status across the United States” Poster presentation at 23rd Annual CMAS Conference, Chapel Hill, NC, October 2024.
2. “Quantify the Impacts of Air Pollution on Disadvantaged Communities in Greater Boston with a New Cumulative Environmental Justice Score System” Poster presentation at BARI 2024, Boston, MA, April 2024.
3. “[GPU-Assisted Computation for a Gas-Phase Chemical Solver in CMAQ](#)” Oral presentation at 22nd Annual CMAS Conference, Chapel Hill, NC, October 2023.
4. “[Machine Learning with Spatial Interpolation Techniques for Constructing 2-Dimensional Ozone Concentrations in Southern California during the COVID-19 Shutdown](#)” Oral presentation at 22nd Annual CMAS Conference, Chapel Hill, NC, October 2023.

5. “Superior Performance of Convolutional Neural Network for Predicting PM_{2.5} Concentrations and Exceedances under Sparse Data Availability” Poster Presentation at AEESP 2023, Boston, MA, June 2023.
6. “[A Machine Learning Approach to Quantify the Impact of Meteorology on Tropospheric Ozone in the Inland Empire, CA](#)” Oral presentation at 21st Annual CMAS Conference, Chapel Hill, NC, October 2022.
7. “[GPU-Assisted CMAQ Simulations](#)” Oral presentation at 21st Annual CMAS Conference, Chapel Hill, NC, October 2022.
8. “Applying Machine Learning and Chemical Transport Model to investigate the influences of meteorology on PM_{2.5} and Ozone” Poster Presentation at 38th Meeting of the American Association for Aerosol Research, October 2020.
9. “High-Resolution Personal Exposure Monitoring of PM_{2.5} in Inland Southern California” Poster Presentation at 37th Meeting of the American Association for Aerosol Research, Portland, Oregon, October 2019.

Co-author:

10. “Multi-Scale Air Quality Modeling Using WRF-Chem-GHG over Northeastern Africa” Oral presentation at 23rd Annual CMAS Conference, Chapel Hill, NC, October 2024.
11. “GPU Implementation of a Gas-Phase Chemistry Solver in the CMAQ Chemical Transport Model” Oral presentation at 23rd Annual CMAS Conference, Chapel Hill, NC, October 2024.
12. “Graphics Processing Unit Assisted Computation for a Gas-Phase Chemical Solver in a Regulatory Chemical Transport Model.” Poster Presentation at UC Berkeley, BASC Symposium 2024, Berkeley, CA, March 2024.
13. “Compounding Risks in Disparately Impacted Communities: California Case Studies and Future Recommendations.” Oral Presentation at AGU Fall Meeting 2023, San Francisco, CA, December 2023.
14. “UAV Measurements in a Heavily Burdened Air Basin to Understand Meteorological and Emissions Uncertainties in CMAQ” Poster Presentation at MAC-MAQ 2023, Davis, CA, September 2023.
15. “Air Pollution Exposure Mitigation for the Protection of Impacted Communities” Oral Presentation at AEESP 2023, Boston, MA, June 2023.
16. “Vertical Ozone Profiles Measurement in Riverside, CA” Poster Presentation at AGU Fall Meeting 2021, December 2021.

17. “Application of Machine Learning for Future Air Quality Predictions in Southern California.” Poster Presentation at 38th Meeting of the American Association for Aerosol Research, October 2020.
18. “High-Temporal Resolution Personal Exposure Pilot Study in Inland Southern California.” Poster Presentation at the Air Sensors International Conference, Pasadena, CA, May 2020.
19. “CMAQ-Enhanced Estimates of Personal Exposure to Diesel Particulates.” Poster Presentation at the Community Modeling and Analysis System Meeting, Chapel Hill, NC, October 2019.
20. “Field Evaluation and calibration of a Six-parameter Low-cost Sensor System in Northwestern and Southeastern US” Poster Presentation at the 37th Meeting of the American Association for Aerosol Research, October 2019.

EXTERNAL COMPUTATIONAL RESOURCES

1. National Energy Research Scientific Computing Center (NERSC) with 1,500 CPU Node hours
2. NCAR-Wyoming Supercomputing Center repository UNTE0007, “SSP Downscaling and RRFs-CMAQ Development.” 500,000 CPU hours and 1,500 GPU hours

SERVICE ACTIVITIES

- CMAS 2023 and 2024 poster evaluating committee
- Journal Review: *Environmental Science & Technology*, *Nature Sustainability*, *GeoHealth*, *Atmosphere*, *Applied sciences*, *Environmental Research Letter*
- Invited talk at Northeastern University CEE Graduate Expo 2024
- Teacher Assistant (CHE 117 Separation Processes, enrollment of 49), University of California, Riverside (2022)
- Guest Lecture at Carnegie Mellon University, 2023
- Technical advisor to Center for Community Action and Environmental Justice (Jurupa Valley, CA) and The Air I Breathe (Colton, CA)
- Community outreach in San Bernardino, CA
- Student assistance at 2019 AAAR conference