

Article

Quantifying Contributions of Factors and Their Interactions to Aerosol Acidity with a Multiple Linear Regression-Based Framework: A Case Study in the Pearl River Delta, China

Hong Ling ¹, Mingqi Deng ¹, Qi Zhang ², Lei Xu ³, Shuzhen Su ⁴, Xihua Li ⁴, Liming Yang ⁵, Jingying Mao ⁶ and Shiguo Jia ^{1,7,*}

¹ School of Atmospheric Sciences, Sun Yat-sen University and Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai), Zhuhai 519082, China; lingh23@mail.sysu.edu.cn (H.L.); dengmq5@mail2.sysu.edu.cn (M.D.); jiazhg3@mail.sysu.edu.cn (S.J.)

² Tianjin Academy of Eco-Environmental Sciences, Tianjin 300191, China; zhangq358@mail2.sysu.edu.cn

³ Environmental Engineering Assessment Center of Ministry of Ecology and Environment, Beijing 100041, China; xulei@acee.org.cn

⁴ Guangdong Dongguan Ecological Environment Monitoring Station, Dongguan 523009, China; ymh0724@163.com (S.S.); bblala@me.com (X.L.)

⁵ Department of Chemical and Biomolecular Engineering, National University of Singapore, Singapore 117576, Singapore; cheylm@nus.edu.sg

⁶ Scientific Research Academy of Guangxi Environmental Protection, Nanning 530022, China; 13978185061@163.com

⁷ Guangdong Provincial Field Observation and Research Station for Climate Environment and Air Quality Change in the Pearl River Estuary, Guangzhou 510275, China

* Correspondence: jiazhg3@mail.sysu.edu.cn; Tel.: +86-756-3668166

S1. Description of MLR analysis

We established a systematic approach to quantify the contributions of individual variables and their interactions towards the explained variance (R^2) of the model. Our objective was to dissect the individual and combined effects of the variables on the dependent variable, pH, and to gain a deeper insight into the underlying relationships within the data.

Individual Variable Contributions: Initially, we assessed the impact of each predictor variable independently. For this, we calculated the change in R^2 (ΔR^2) resulting from the exclusion of each variable from the full model. This step was crucial to identify the variables that had the most significant individual impact on explaining the variance in pH.

Coefficient Stability in Interaction Analysis: Our initial phase of the MLR analysis solely included the predictor variables, allowing us to obtain their initial coefficients. As we progressed to introducing two-variable interactions, we made a deliberate choice to keep the coefficients of these original variables constant. This decision was made to maintain the integrity and stability of the model. Allowing the coefficients of the original variables to vary with the introduction of two-variable interactions could have introduced undue complexity, leading to unstable and potentially non-representative values for these coefficients. By keeping them constant, we ensured the robustness of the model and that the coefficients accurately reflected the true influence of each variable. This methodology was consistently applied not only in the context of two-variable interactions but also when analyzing more complex three-way and four-way interactions, further preserving the model's stability and reliability.

Two-Way Interaction Contributions: To explore the combined effects of pairs of variables, we introduced two-way interaction terms into the model. Each bivariate interaction term was analyzed by excluding it from a model that included all original variables and other interaction terms, and the change in R^2 was noted. The sum of ΔR^2 values across

Citation: Ling, H.; Deng, M.;

Zhang, Q.; Xu, L.; Su, S.; Li, X.;

Yang, L.; Mao, J.; Jia, S.

Quantifying Contributions of Factors and Their Interactions to Aerosol Acidity with a Multiple-Linear-Regression-Based Framework: A Case Study in the Pearl River Delta, China. *Atmosphere* 2024, 15, 172. <https://doi.org/10.3390/atmos15020172>

Academic Editor: Ian Colbeck

Received: 20 December 2023

Revised: 18 January 2024

Accepted: 25 January 2024

Published: 29 January 2024



Copyright: © 2023 by the authors.

Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

all two-way interactions provided a measure of the total contribution of bivariate interactions to the model.

Three-Way and Four-Way Interaction Contributions: Following a similar methodology, we extended our analysis to three-way (trivariate) and four-way (quadrivariate) interactions. We included interaction terms representing all possible combinations of three and four variables, respectively. The contribution of each set of interactions was assessed by observing the change in R^2 upon their exclusion from the model. This allowed us to evaluate the complex synergistic effects of multiple variables on pH.

Relative and Specific Contributions: For each variable and interaction term, we calculated their relative contributions by comparing their ΔR^2 to the total ΔR^2 of all terms. This relative measure was then used to determine the specific contribution of each term to the overall R^2 of the model.

This comprehensive analysis provided us with a nuanced understanding of the role played by each variable and their interactions in our regression model. It highlighted not only the individual importance of each variable but also how combinations of variables interacted to influence the model's explanatory power. The results from this analysis are presented in detail in the accompanying tables and figures in this supplementary material.

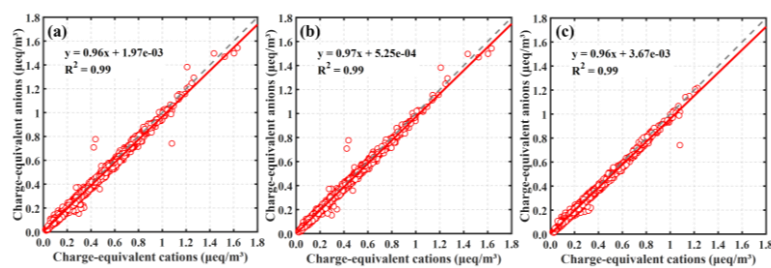


Figure S1. Comparison of charge-equivalent cations and anions of water-soluble inorganic species: (a) All Samples, (b) Daytime, and (c) Nighttime.

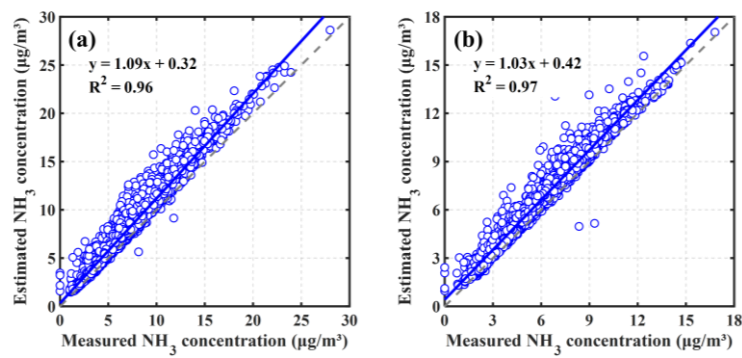


Figure S2. Comparison of measured NH_3 concentrations and estimates derived from ISORROPIA-II during daytime and nighttime.

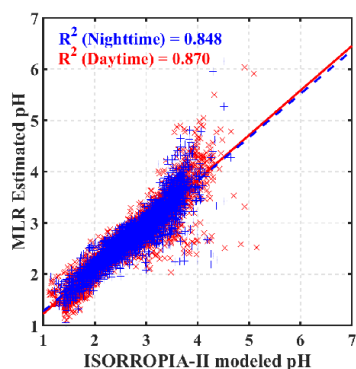


Figure S3. Comparative of aerosol pH as calculated by ISORROPIA-II and estimated through MLR for daytime and nighttime samples separately. The resultant MLR equations are: (1) $\text{pH}_{\text{day}} = 17.07 + 0.30 \times \text{Na}^+ - 0.08 \times \text{SO}_4^{2-} + 0.05 \times \text{TNH}_x + 0.02 \times \text{TNO}_3 + 0.05 \times \text{TCl} + 0.18 \times \text{Ca}^{2+} + 0.05 \times \text{K}^+ + 0.53 \times \text{Mg}^{2+} + 1.44 \times \text{RH} - 0.05 \times \text{T}$ and (2) $\text{pH}_{\text{night}} = 17.24 + 0.27 \times \text{Na}^+ - 0.08 \times \text{SO}_4^{2-} + 0.07 \times \text{TNH}_x + 0.01 \times \text{TNO}_3 + 0.05 \times \text{TCl} + 0.32 \times \text{Ca}^{2+} + 0.08 \times \text{K}^+ + 0.46 \times \text{Mg}^{2+} + 1.53 \times \text{RH} - 0.05 \times \text{T}$.

Table S1. Comprehensive overview of parameter ranges and units employed in MLR analysis.

Parameter	Minimum	Maximum	Unit ¹
Na ⁺	0	7.178	µg/m ³
SO ₄ ^{2−}	0.01	27.234	µg/m ³
TNH _x	1.4037	38.263	µg/m ³
TNO ₃	0	58.031	µg/m ³
TCl	0	26.796	µg/m ³
Ca ²⁺	0	2.3990001	µg/m ³
K ⁺	0	29.468001	µg/m ³
Mg ²⁺	0	2.7620001	µg/m ³
RH	0.16594	0.8975	Dimensionless
T	277.41	311.07	Kelvin

Note: Concentrations below the detection limit have been recorded as 0 for the purposes of this analysis.